



THE UNIVERSITY
of ADELAIDE

**Clonal proliferation of hepatocytes during
chronic hepatitis B virus infection**

Mr. Thomas Tu

B.Sc. (Hons) (Biomedical Science)

A thesis submitted to The University of Adelaide for
The Degree of Doctor of Philosophy

School of Molecular and Biomedical Sciences

The University of Adelaide, Australia

October 2012

Table of contents

Title page	i
Table of contents	ii
Abstract	x
Declaration	xii
Acknowledgements	xiii
Publications and presentations arising from this thesis	xv
Awards arising from this thesis	xviii
Abbreviations	xix
1. Introduction	1
1.1 Epidemiology of Hepatitis B Virus (HBV) infection	2
1.2 HBV structure and replication	2
1.3 Natural history of HBV infection	5
1.3.1 Acute HBV infection	5
1.3.2 Chronic HBV infection - Immune tolerance phase	6
1.3.3 Chronic HBV infection - Immune clearance phase	7
1.3.4 Chronic HBV infection - Immune control phase and reactivation	9
1.3.5 Disease outcomes of chronic HBV infection	10
1.4 Liver anatomy and microstructure	11
1.5 Hepatocyte repopulation in the liver	11
1.6 Tracing hepatocyte lineages	13
1.7 Histology of HCC and HCC-associated lesions	17
1.8 Molecular changes during HCC progression	18
1.9 HBV and HCC progression	21
1.9.1 The role of HBV proteins in HCC progression	22
1.9.2 The role of HBV DNA integration in HCC progression	23
1.9.3 The role of HBV-associated inflammation in HCC progression	25
1.10 HBV-associated disease progression as a Darwinian process	27
1.11 Summary and significance	31
1.12 Aims and hypotheses	31
2. Materials and Methods	32
2.1 Collection of liver tissues	33
2.2 Tissue fixation and paraffin wax-embedding	33
2.3 Liver tissue histological staining and analysis	34
2.3.1 Haematoxylin and Eosin staining	34
2.3.2 Periodic acid-Schiff Diastase staining	35
2.3.3 Gordon and Sweet's reticulin staining	35
2.3.4 Haematoxylin van Gieson staining	36
2.3.5 Perls' Prussian Blue Iron staining	36
2.3.6 Fouchet's Bile pigment staining	36
2.4 Detection of HBcAg and HBsAg in liver sections by immunohistochemistry	37
2.5 Dual detection of HBsAg and cytokeratin 19 in liver sections by dual immunofluorescence	38

2.6	Detection of HBV DNA in liver sections using <i>in situ</i> hybridisation	39
2.6.1	Probe production	39
2.6.2	Pre-hybridisation and hybridisation solutions	39
2.6.3	<i>In situ</i> hybridisation	40
2.7	Extraction of total DNA from snap-frozen liver tissue	41
2.8	Extraction of total DNA from paraffin wax-embedded liver tissue sections	42
2.9	Extraction of nuclear DNA from liver samples	42
2.10	Extraction of total DNA from serum samples	43
2.11	Optical Density Spectroscopy	43
2.12	Amplification and sequencing of HBV DNA	43
2.12.1	PCR primer design	43
2.12.2	PCR conditions	44
2.12.3	DNA sequencing	44
2.13	Enumeration of HBV DNA or virus-cell junction copy number by qPCR	45
2.13.1	Primer design	45
2.13.2	qPCR conditions	45
2.13.3	Generation of plasmid DNA standards for Beta Globin	45
2.14	invPCR	46
2.15	invPCR of high molecular weight DNA	48
2.16	Gel extraction in 96-well format	49
2.17	Semi-DOP PCR	49
2.18	qPCR analysis of virus-cell junctions	50
2.19	Computer simulation	51
2.20	Laser microdissection of foci of hepatocytes	51
2.21	Total DNA extraction and invPCR from tissue isolated by laser-microdissection	51
2.22	Calculation of clone size from invPCR of DNA extracts from liver sections	52
2.23	Detection of PreS-mutants in hepatocytes isolated by laser-microdissection	52
2.24	Quantification of methylated HBV DNA CpG regions in hepatocytes isolated by laser-microdissection	53
2.25	Estimation of global methylation of DNA in hepatocytes isolated by laser-microdissection	53
2.26	Imaging mass spectrometry	54
2.27	High pressure Liquid Chromatography Electrospray Ionisation Linear Ion Trap Orbitrap mass spectrometry (HPLC-nESI-LTQ Orbitrap MS)	55
2.28	High Pressure Liquid Chromatography Matrix-assisted Laser Desorption Ionisation Time of Flight tandem Mass Spectrometry (HPLC-MALDI-TOF MS/MS)	56
2.28.1	HPLC and Fraction Collection	56
2.28.2	MALDI-TOF MS of HPLC Fractions	57
2.28.3	Tandem MS of HPLC Fractions	58
2.29	Data processing and visualisation of the distribution of known proteins in tissue sections	58
3.	Methods Optimisation	60
3.1	Introduction	61
3.2	Results	62

3.2.1	Optimisation of genomic-based techniques	62
3.2.1.1	Approximation of nucleus concentration	62
3.2.1.2	Optimisation of DNA extraction from liver tissue	63
3.2.1.2.1	Total DNA extraction of 5 mg liver tissue fragments	63
3.2.1.2.2	Total DNA extraction of EtOH-fixed paraffin wax-embedded liver tissue sections	64
3.2.1.2.3	Nuclear DNA extraction	65
3.2.1.2.4	Total DNA extraction of hepatocyte foci isolated by laser-microdissection	65
3.2.1.3	Optimisation of semi-degenerate oligonucleotide primer PCR	66
3.2.1.4	Optimisation of invPCR	67
3.2.1.5	Gel purification of high-molecular weight DNA using ultrasonication or partial restriction enzyme digestion	68
3.2.1.6	Human invPCR design and efficiency of inversion reaction	69
3.2.1.7	Optimisation of detection of long interspersed element (LINE) promoter methylation	70
3.2.1.8	HBV DNA CpG island methylation	71
3.2.2	Optimisation of proteomic-based techniques	72
3.2.2.1	Optimisation of imaging mass spectrometry (IMS)	72
3.2.2.2	Optimisation of HPLC-nESI-LTQ Orbitrap MS	74
3.2.2.3	Optimisation of HPLC-MALDI-TOF MS/MS	75
3.2.2.4	Optimisation of visualising the distribution of known proteins in liver tissue sections	76
3.2.3	Optimisation of RNA extraction from frozen liver tissues	77
3.3	Discussion	78
3.3.1	DNA extraction	78
3.3.2	Detection of virus-cell DNA junctions	78
3.3.3	Epigenetic genotyping	79
3.3.4	Proteomic phenotyping	80
3.3.5	Summary	82
4.	Characterisation of liver tissues	84
4.1	Introduction	85
4.2	Experimental Outline	86
4.3	Results	88
4.3.1	Patient details and liver histology reports	88
4.3.2	HBV DNA analysis	90
4.3.3	Detection and localisation of HBV DNA by <i>in situ</i> hybridisation	91
4.3.4	Detection of HBV antigens by immunohistochemistry	92
4.3.5	HBsAg and Cytokeratin 19 (CK19) localisation by dual immunofluorescence	94
4.4	Discussion	95
4.4.1	Classification of patient tissues	95
4.4.2	Virological aspects of HBV patients	96
4.4.3	HBV antigen expression in cells other than hepatocytes	97
4.4.4	Summary	98

5.	Detection of hepatocyte clones in HBV-infected liver tissues	99
5.1	Introduction	100
5.2	Experimental Outline	102
5.3	Results	104
5.3.1	Primer design and restriction enzymes selection for invPCR	104
5.3.2	Calculation of theoretical maximum fraction of observable virus-cell junctions	105
5.3.3	Large hepatocyte clones were observed in DNA extracts from 5 mg of liver fragments by invPCR	108
5.3.4	Characterisation of large clones found in patients XA and CY	109
5.3.5	InvPCR analysis of DNA extracted from sections of ethanol-fixed liver tissue	111
5.3.6	Predominantly unique subsets of virus-cell junctions were observed using the <i>DpnII</i> and the <i>NcoI</i> invPCR designs	113
5.3.7	Observed integration sites of HBV DNA are clustered around the expected right-hand end of the dsDNA form	113
5.3.8	Observed integration sites are random within the cellular DNA with respect to chromosomes and genes and do not affect clone size	114
5.3.9	Observed integration sites (but not clone sizes) are weakly associated with structural/matrix attachment regions (S/MAR)	115
5.4	Discussion	116
5.4.1	PCR-amplified DNA artefacts	118
5.4.2	Changes in phenotype directly caused by HBV DNA integration	119
5.4.3	Clonal proliferation due to stem cell activation and proliferation	121
5.4.4	Increased liver turnover over the course of infection increases stochastic clonal proliferation or clonal proliferation of hepatocytes with survival advantage	122
6.	Computer simulation of hepatocyte proliferation during chronic HBV infection	124
6.1	Introduction	125
6.1.1	Assumptions	126
6.1.1.1	The liver grows 10-fold during development from neonate to full-grown adult	126
6.1.1.2	Hepatocyte number in the adult liver remains constant	127
6.1.1.3	All hepatocytes are infected at the start of infection	127
6.1.1.4	Survival advantages are heritable	127
6.1.2	Program description	128
6.1.3	Parameters	129
6.1.3.1	LiverArray size	129
6.1.3.2	Connectivity	130
6.1.3.3	Focal cell death	131
6.1.3.3	Hepatocytes with a survival advantage	132
6.1.3.4	Stem-cell populations	133
6.2	Results	133

6.2.1	The current model replicates the results of a previous stochastic simulation	133
6.2.2	The effect of LiverArray size on complexity, maximum clone size and clone frequency	134
6.2.3	Connectivity weakly affects clonal proliferation	134
6.2.4	Stochastic clonal proliferation is highly predictable during liver turnover	135
6.2.5	Clonal proliferation due to liver growth during development is similar to that of 10 liver turnovers	137
6.2.6	Focal liver cell death inhibits clonal proliferation	137
6.2.7	Clonal proliferation is exquisitely sensitive to the introduction of survival advantage	138
6.2.8	The effect of oval cells is undetectable	139
6.2.9	Estimation of maximum liver turnovers in chronic HBV patients	139
6.3	Discussion	140
7.	Characterisation of hepatocyte foci	144
7.1	Introduction	145
7.1.1	Detectable virus-cell DNA junctions in cells other than hepatocytes	146
7.1.2	Histological changes in hepatocytes that have undergone clonal proliferation	146
7.1.3	Association between expression of HBV DNA and clonal proliferation of hepatocytes	147
7.1.4	Mechanisms of heterogeneous expression of HBV antigens	147
7.1.5	Altered cellular protein expression in hepatocytes	148
7.1.6	Aims and hypotheses	148
7.2	Experimental Outline	149
7.3	Results	150
7.3.1	Isolation of liver cell subpopulations by laser-microdissection	150
7.3.2	InvPCR and qPCR analysis of non-hepatocyte cell populations for HBV DNA	151
7.3.3	Virus-cell junctions were not frequently detected in hepatocytes with large cell change (LCC)	152
7.3.4	Virus-cell DNA junctions were 2 times more likely to be detected in foci of HBsAg-negative hepatocytes compared to foci of HBsAg-positive hepatocytes	152
7.3.5	Detected HBV DNA integrations display no preference for particular genes or proximity to S/MAR	153
7.3.6	HBsAg-positive hepatocytes contain two times more HBV DNA genomes per cell compared to HBsAg-negative hepatocytes	154
7.3.7	HBV PreS1/S2 mutants were predominantly detected in HBsAg-positive hepatocytes	156
7.3.8	No significant difference in methylation of HBV DNA was observed between HBsAg-positive and -negative foci of hepatocytes	156
7.3.9	Cellular protein expression was altered in foci of HBsAg-positive hepatocytes compared to foci of HBsAg-negative hepatocytes	157

7.3.10	Altered cellular protein expression due to zonal heterogeneity was not observed in liver tissue sections by IMS	159
7.3.11	Cellular protein distribution is heterogeneous in liver tissue sections from patients with chronic HBV infection compared to HBV non-infected controls	161
7.4	Discussion	162
7.4.1	Identification of the cell populations that had undergone the clonal proliferation detected by invPCR	162
7.4.2	HBsAg expression and clonal proliferation	163
7.4.3	Heterogeneous expression of HBsAg	165
7.4.4	Heterogeneous expression of cellular proteins	166
7.4.5	Summary	169
8.	Concluding remarks	170
8.1	Summary	171
8.2	Future work	173
8.3	Impact on other fields of study	176
8.4	Conclusion	177
9.	Appendices	178
9.1	Sequence of pHBV1.3BB4.5	179
9.2	Melting temperature of Southern hybridisation probe	181
9.3	Efficiency of self-ligation of excised virus-cell DNA junctions in the invPCR protocol	182
9.3.1	Assumptions	182
9.3.2	Calculation of factor j	182
9.3.3	Calculation of factor i	183
9.3.4	Calculation of j/i	185
9.4	Copies of Ethics Approval forms	186
9.4.1	Royal Adelaide Hospital HREC approval for Protocol #070301	186
9.4.2	NSW Health HREC approvals for Protocols #X10-0072 and #HREC/10/RPAH/130	187
9.4.3	Fox Chase Cancer Center Institutional Review Board approvals for Protocols #31-281 and #08-801	189
9.4.4	Southern Clinical HREC approval for Protocol #171.11 and amendments	192
9.4.5	Royal Adelaide Hospital HREC approval for Protocol #050803	193
9.5	Proteins detected by HPLC-nESI-LTQ-Orbitrap XL analysis of trypsin-digested, lysates of hepatocytes isolated by laser-microdissection of NHL tissue	194
9.5.1	VivaSpin-purified lysates of hepatocytes	194
9.5.2	Pepclean-purified lysates of hepatocytes	199
9.6	Proteins detected by HPLC MALDI-TOF MS/MS results from trypsin-digested, PepSpin C18 Spin column-purified of hepatocytes isolated by laser-microdissection of NHL tissue	203
9.7	The HBV DNA sequence detected in each patient with chronic HBV infection	206
9.7.1	Patient C	206

9.7.2	Patient CN	206
9.7.3	Patient CY	207
9.7.4	Patient CYRY	207
9.7.5	Patient DG	208
9.7.6	Patient FMC1	208
9.7.7	Patient FMC2	208
9.7.8	Patient FMC3	208
9.7.9	Patient FMC4	208
9.7.10	Patient FMC5	209
9.7.11	Patient FMC6	209
9.7.12	Patient FMC7	209
9.7.13	Patient GS1	209
9.7.14	Patient GS2	210
9.7.15	Patient GS3	210
9.7.16	Patient GS4	210
9.7.17	Patient GS5	210
9.7.18	Patient HN	211
9.7.19	Patient HS	211
9.7.20	Patient L	211
9.7.21	Patient MH	212
9.7.22	Patient NT	212
9.7.23	Patient SAAO	212
9.7.24	Patient WN	212
9.7.25	Patient XA	213
9.7.26	Patient Y2	213
9.7.27	Patient Y3	213
9.7.28	Patient Y4	214
9.7.29	Patient Y5	214
9.7.30	Patient Y6	214
9.8	Sequences of virus-cell DNA junctions	215
9.8.1	Patient C	215
9.8.2	Patient CN	216
9.8.3	Patient CY	216
9.8.4	Patient CYRY	217
9.8.5	Patient DG	217
9.8.6	Patient FMC1	218
9.8.7	Patient FMC2	218
9.8.8	Patient FMC3	218
9.8.9	Patient FMC4	219
9.8.10	Patient FMC5	219
9.8.11	Patient FMC6	219
9.8.12	Patient FMC7	219
9.8.13	Patient GS1	220
9.8.14	Patient GS2	220
9.8.15	Patient GS3	220
9.8.16	Patient GS4	220

9.8.17	Patient GS5	220
9.8.18	Patient HN	220
9.8.19	Patient HS	220
9.8.20	Patient L	220
9.8.21	Patient MH	221
9.8.22	Patient NT	221
9.8.23	Patient SAAO	221
9.8.24	Patient WN	222
9.8.25	Patient XA	223
9.8.26	Patient Y2	224
9.8.27	Patient Y3	225
9.8.28	Patient Y4	225
9.8.29	Patient Y5	227
9.8.30	Patient Y6	228
9.9	Computer programming	230
9.9.1	Random number generation	230
9.9.2	Cell selection	230
9.9.3	Source code for the simulation of stochastic clonal proliferation	230
9.9.4	Source code for the 2D visualisation of stochastic clonal proliferation	237
9.10	Ion maps of mass peaks corresponding to proteins identified by IMS and HPLC-nESI-LTQ Orbitrap MS in multiple liver tissue sections of HBV-negative and -positive patients	242
10.	References	249

Abstract

Chronic hepatitis B virus (HBV) infection causes liver disease that can progress to cirrhosis and hepatocellular carcinoma (HCC). Changes in the hepatocyte population that occur from the early immune-tolerant stage of infection to late-stage disease outcomes remain unclear. We hypothesised that some hepatocytes lose HBV antigen expression and escape the HBV-specific immune response, allowing them to undergo clonal proliferation. Clonal proliferation of altered hepatocytes may be a marker of disease progression and may have a direct role in the development of HCC.

Liver tissues from 30 patients were analysed, including patients with early-stage HBV infection, late-stage infection with cirrhosis, or with HCC. Unique virus-cell DNA junctions formed by the integration of HBV DNA into the host cell genome were detected using inverse nested PCR (invPCR). The copy number of unique virus-cell DNA junctions was used as a measure of clonal proliferation of hepatocytes. A computer simulation of a liver undergoing stochastic liver turnover was used to determine if the hepatocyte clones observed by invPCR could have been formed by random chance. Immunohistochemistry for HBV surface antigen (HBsAg) expression and Imaging Mass Spectrometry (IMS) for cellular protein expression were carried out to detect cellular changes that may be associated with clonal proliferation.

Significantly ($p < 0.01$) larger clones were observed by invPCR in liver DNA extracts of patients with late-stage HBV-associated disease (≤ 280000 hepatocytes) compared to patients in early-stage HBV infection (8-1124 hepatocytes). Computer simulations indicated that stochastic turnover could not produce clones of > 10000 hepatocytes, suggesting that the hepatocytes that had formed large clones had a survival advantage. No significant difference in the extent of clonal proliferation was observed in foci of HBsAg-positive and -negative hepatocytes isolated by laser-microdissection. Heterogeneous expression of cellular proteins was detected using IMS in hepatocytes with apparently normal histology.

These results indicate that clonal proliferation of hepatocytes with survival advantage does occur in the hepatocyte population during chronic HBV infection and can be detected before histological changes are evident in the hepatocytes of patients with both early- and late-stage disease. Consistent with our hypothesis, larger hepatocyte clones were associated with disease progression. The cause of the clonal proliferation remains unknown. Contrary to our

hypothesis, loss of HBsAg expression was not associated with increased clonal proliferation, suggesting that escape from HBsAg-specific immune attack is not a survival advantage. While not investigated in this thesis, the loss of expression of other HBV antigens may provide a survival advantage. Heterogeneous cellular protein expression suggests that hepatocyte phenotype has been altered in some hepatocytes. However, we could not show using invPCR approaches that the foci of hepatocytes with altered cellular protein expression were clonal.

In conclusion, this research has provided groundwork in determining the relationship between the clonal proliferation of hepatocytes, altered hepatocyte phenotype and HBV-associated disease progression. Further studies into the molecular causes of clonal proliferation of hepatocytes with survival advantages could elucidate pathways of HBV-associated disease progression and novel ways to curb the evolution of the hepatocyte population to a less pathogenic state.

Declaration

NAME: Thomas Tu

PROGRAM: Doctor of Philosophy (Ph.D.)

I declare that no material from this work has been submitted for the award of any other degree or diploma in the University of Adelaide or any other tertiary institution. To the best of my knowledge and belief, this work contains no material previously published or written by another person, except where due reference has been made in the text.

I consent to this copy of my thesis being made available for circulation and photocopying for purposes of study and research in accordance with the rules established by the University of Adelaide and subject to the provisions of the Copyright Act 1968.

I consent for the digital version of my thesis to be made available on the Internet via the University of Adelaide digital research repository, the Barr Smith Library catalogue, the Australian Digital Thesis Program and through web search engines, unless permission has been granted by the University of Adelaide to restrict access for a period of time.

Thomas Tu

October, 2012

Acknowledgements

First and foremost, I would like to acknowledge A/Prof. **Allison Jilbert**, for accepting me into her laboratory to do my Ph.D. project and nurturing my growth as a person and a scientist. Her guidance and encouragement has honed me and my research into something that I am proud of.

Also, I acknowledge my co-supervisors: Emeritus Prof. **William Mason** (FCCC), who not only supervised me during the early parts of my candidature, but also has provided me with bench experience, feedback and helpful conversations through my time as a Ph.D. student; and Dr. **Uwe Stroehler**, who brought new knowledge to my project and has supported me with his mentoring.

I have made almost-familial bonds with past and present lab mates, to whom I express deep thanks for support and conversations, scientific and not. Special thanks to Ms. **Catherine Scougall**, who has eagerly helped out in technical lab work, thesis editing and proofreading; Ms. **Huey Chi Low**, who has laid down many foundation techniques required for this project; and Dr. **Georget Reaiche-Miller**, whose experience in the lab and scientific life has continually helped me throughout my candidature.

I am also indebted to the significant contribution by collaborators to this project: A/Prof. **Hugh Harley** (RAH), Prof. **Geoff McCaughan** (Centenary Institute), and Drs. **Nick Shackel** (Centenary Institute), **John Chen** (FMC), and **Andrew Ruskiewicz** (IMVS) for providing access to liver tissues; A/Prof. **Andrew Clouston** (UQ) for histological interpretation; **eResearch SA** for computing time; Mr. **Ashleigh Brook** and Mr. **Luke Maurits** for invaluable help on mathematics and programming; Mr. **Johan Gustaffson**, Ms. **Yin Ying Ho** and A/Prof. **Peter Hoffmann** (APC) for their collaboration, conversations about proteomics theory and conducting of the mass-spectrometry experiments; and to the **patients who have volunteered their liver tissues** that this research could be possible.

I acknowledge my financial support during my Ph.D. through a Faculty of Science Divisional Scholarship. I also would like to thank the staff and students at the School of Molecular and Biomedical Sciences, and the University of Adelaide for their support throughout my student life.

On a personal note, I would like to thank all my friends and family. In particular, I would like to acknowledge my mother, Mrs. **Tee Wu**, whose upbringing, sacrifices and hot meals have made me who I am today. Finally, thanks to Ms. **Miriam Kleinig** for her tender warmth, support and love.

“A poet once said, "The whole universe is in a glass of wine." We will probably never know in what sense he meant that, for poets do not write to be understood. But it is true that if we look at a glass of wine closely enough we see the entire universe. There are the things of physics: the twisting liquid which evaporates depending on the wind and weather, the reflections in the glass, and our imagination adds the atoms. The glass is a distillation of the Earth's rocks, and in its composition we see the secrets of the universe's age, and the evolution of stars. What strange arrays of chemicals are in the wine? How did they come to be? There are the ferments, the enzymes, the substrates, and the products. There in wine is found the great generalization: all life is fermentation. Nobody can discover the chemistry of wine without discovering, as did Louis Pasteur, the cause of much disease. How vivid is the claret, pressing its existence into the consciousness that watches it! If our small minds, for some convenience, divide this glass of wine, this universe, into parts — physics, biology, geology, astronomy, psychology, and so on — remember that Nature does not know it! So let us put it all back together, not forgetting ultimately what it is for. Let it give us one more final pleasure: drink it and forget it all!”

- *Richard P Feynman, Volume I, 3-10, The relation of Physics to other sciences*

Publications and presentations resulting from this thesis

Manuscripts in preparation

Tu, T., Mason, W.S., Low, H.C., Clouston, A., Stroehrer, U., and Jilbert A.R.

Clonal proliferation of hepatocytes is associated with disease progression in chronic hepatitis B virus infection.

Tu, T., Gustafsson, J., Ho, Y.Y., Mason, W.S., Clouston, A., Stroehrer, U., and Jilbert A.R.

Heterogeneous viral and cellular protein expression in hepatocytes during chronic hepatitis B virus infection – analysis by imaging mass-spectrometry.

Tu, T., Ruszkiewicz, A., Clouston, A., and Jilbert A.R.

Liver cell coexpression of hepatitis B surface antigen and cytokeratin 19 during chronic HBV infection.

Oral presentations at international conferences

Tu, T., Mason, W.S., Low, H., Clouston, A., Stroehrer, U., and Jilbert A.R. (2012)

Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.

14th International Symposium on Viral Hepatitis and Liver Disease, Shanghai International Convention Centre, 24th June 2012

Tu, T., Mason, W.S., Low, H., Clouston, A., Stroehrer, U., and Jilbert A.R. (2010)

Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.

The International Meeting for the Molecular Biology of Hepatitis B Viruses, Academia Sinica, Taipei, Taiwan, 12th October 2010

Poster presentations at international conferences

Tu, T., Mason, W.S., Low, H., Clouston, A., Grosse, A., Stroehrer, U., and Jilbert A.R. (2009)

Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.

The International Meeting for the Molecular Biology of Hepatitis B Viruses, Vinci International Convention Centre, Tours, Loire Valley, France, 30th August 2009.

Oral presentations at national conferences

Tu, T., Mason, W.S., Low, H., Clouston, A., Stroehrer, U., and Jilbert A.R. (2011)

Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.

Australian Centre for HIV and Hepatitis Virology Workshop 2011, Twin Waters Resort, Sunshine Coast, QLD, 1st June 2011.

Tu, T., Mason, W.S., Low, H., Clouston, A., Grosse, A., Stroehler, U., and Jilbert A.R. (2009)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
Australian Centre for HIV and Hepatitis Virology Workshop 2009, Crowne Plaza, Terrigal,
NSW, 3rd June 2009.

Tu, T., Low, H., Grosse, A., Miao, Y., Mason, W.S., and Jilbert A.R. (2008)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
Australian Centre for HIV and Hepatitis Virology Workshop 2008, Novotel Barossa Valley
Resort, SA, 4th June 2008.

Tu, T., Low, H., Miao, Y., Mason, W.S., and Jilbert A.R. (2007)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
Australian Centre for Hepatitis Virology Workshop 2007, Burnet Institute, Melbourne, 15th May
2007.

Oral presentations at awards presentations

Tu, T., Low, H., Grosse, A., Miao, Y., Mason, W.S., and Jilbert A.R. (2008)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
The Australian Society for Microbiology (SA Branch) 2008 Student Awards Night, Flinders
Medical Centre, Adelaide, 26th November 2008.

Poster presentations at national conferences

Tu, T., Mason, W.S., Low, H., Clouston, A., Grosse, A., Stroehler, U., and Jilbert A.R. (2009)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
Australian Society of Microbiology meeting 2009, Perth Convention Exhibition Centre, Perth,
6th July 2009.

Tu, T., Low, H., Miao, Y., Mason, W.S., and Jilbert A.R. (2007)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
4th Australian Virology Group Meeting 2007, Kingfisher Bay Resort, Frazer Island, 9th
December 2007.

Oral presentations in a seminar series

Tu, T., Mason, W.S., Low, H., Clouston, A., Stroehler, U., and Jilbert A.R. (2009)
Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.
Medical Infectious Diseases Research Seminar Series, Institute of Medical and Veterinary
Science (Now SA Pathology), Adelaide, 27th August 2010.

Tu, T., Mason, W.S., Low, H., Clouston, A., Grosse, A., Stroehler, U., and Jilbert A.R. (2009)

Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.

Infectious Diseases Laboratories Research Seminar Series, Institute of Medical and Veterinary Science (Now SA Pathology), Adelaide, 24th April 2009.

Tu, T., Low, H., Grosse, A., Miao, Y., Mason, W.S., and Jilbert A.R. (2008)

Clonal proliferation of hepatocytes in chronic hepatitis B virus infection.

Infectious Diseases Laboratories Research Seminar Series, Institute of Medical and Veterinary Science (Now SA Pathology), Adelaide, 13th June 2008.

Bold represents presenting author

Awards resulting from this thesis

Invited Guest Speaker Honorarium from 14th International Symposium on Viral Hepatitis and Liver Disease, valued at USD \$1500 (Awarded in 2012)

Integrated Science Young Achievement award at Australian Centre for HIV and Hepatitis Virology Workshop 2011, valued at AUD \$500 (Awarded in 2011)

Student Travel Grant Award, International Meeting for Molecular Biology of HBV, valued at USD \$800 and USD \$1100 (Awarded in 2009 & 2010)

ASM SA Branch Student Award, valued at AUD \$1500 (Awarded in 2008)

Abbreviations

Abbreviation	Full Meaning
%GC	Guanosine and Cytosine content (expressed as a percentage)
~	Approximately
<	Less than
=	Equal to
>	Greater than
≥	Greater than or equal to
°C	degree(s) Celsius
µg	Microgram(s)
µL	Microlitre(s)
µM	Micromolar
µm	Micrometre(s)
5'-CG-3'	Cytosine-guanadine dinucleotide sequences
Ab	Antibodies
ABC	Avidin-Biotin Complex
ABI	Applied Biosystems
ACN	Acetonitrile
AFP	Alpha Fetoprotein
ALT	Amino liver transferase
APC	Adenomatous Polyposis Coli
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BRDU	5-bromo-2'-deoxyuridine
BSA	Bovine serum albumin
cccDNA	Covalently closed circular DNA
CCl ₄	Carbon tetrachloride
C _{die}	Cell nominated for cell death
CHCA	α-cyano-4-hydroxycinnamic acid
CK19	Cytokeratin 19
C _{live}	Cell nominated to replace dying cell
cm	Centimetre(s)
CNTF	Ciliary Neurotrophic Factor
COBRA	Combined Bisuphite and Restriction Analysis

COX	Cytochrome c oxidase
CPM	Counts per minute
CV	Central Vein
Da	Dalton(s)
DAB	3,3'-Diaminobenzidine tetrahydrochloride
dCTP	Deoxycytidine
DHBV	Duck hepatitis B virus
DMNT	DNA methyltransferase
DNA	Deoxyribonucleic acid
DOP	Degenerate Oligonucleotide Primer
DR	Direct repeat sequence
dsDNA	Double-stranded linear DNA
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EAA	Ethanol acetic acid
EDTA	Ethylene-diamine-tetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic Reticulum
EST	Database of Expressed Sequence Tags
EtOH	Ethanol
F	Frequency
FA	Formic Acid
FAH	Foci of altered hepatocytes
FCCC	Fox Chase Cancer Center
FH	Fulminant Hepatitis
g	Gram(s)
GGH	Ground Glass Hepatocytes
H&E	Haematoxylin & eosin
H ₂ O	Water
HBcAg	Hepatitis B core antigen
HBeAg	Hepatitis B e antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HBx	Hepatitis B x protein
HCC	Hepatocellular carcinoma

HCl	Hydrochloric acid
HCV	Hepatitis C Virus
HPLC	High-pressure liquid chromatography
HPLC-nESI-LTQ	High Pressure Liquid Chromotography Electrospray Ionisation
Orbitrap MS	Linear Ion Trap Orbitrap Mass Spectrometry
hr	Hour(s)
HREC	Human Research Ethics Committee
hTERT	Human Telomere Reverse Transcriptase
ID	Unique identifier
IMS	Imaging Mass-Spectrometry
IMVS	Institute of Medical and Veterinary Science
invPCR	Inverse nested PCR
kbp	Kilobase pair
kDa	Kilodalton(s)
L	Litre(s)
<i>lacZ</i>	Beta-galactosidease gene
LB	Luria broth
LCC	Large cell change
LHBs	Large Hepatitis B surface antigen
LINE	Long interspersed elements
LTQ	Linear Ion Trap
M	Molar
m/z	Mass:Charge ratio
mA	Milliampere(s)
MALDI	Matrix-assisted laser-desorption/ionisation
mg	Milligram(s)
MgCl ₂	Magnesium chloride
MHBs	Medium Hepatitis B surface antigen
MHC	Major histocompatibility complex
min	Minute(s)
mL	Millilitre(s)
mM	Millimolar
mm	Millimetre(s)
MPN	Most Probable Number
mRNA	Messenger Ribonucleic Acid

MS	Mass spectrometry
MW	Molecular weight
NaCl	Sodium chloride
NB	Needle Biopsy
NBT	Nitro blue tetrazolium chloride
NCBI	National Center for Biotechnology Information
ND	Not Detected
NEB	New England Biolabs
nESI	Nano-electrospray ionisation
NHEJ	Non-homologous end joining
NHL	Normal human liver
nm	nanometre(s)
NP40	Nonidet-P40
nt	Nucleotide(s)
O/N	Overnight
ORF	Open reading frame
PAS-D	Periodic acid schiff - diastase
PBS	Phosphate-buffered saline
PCA1	Procollagenase alpha 1
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PD-1	Programmed Death 1
PD-1L	Programmed Death-1 Ligand
PEN	Polyethylene Napthalate
pgRNA	Pregenomic RNA
PK	Proteinase K
pM	Picomolar
pol	Hepatitis B polymerase protein
PT	Portal tract
qPCR	Quantitative Polymerase Chain Reaction
R	Resistant
RB1	Retinoblastoma-1
rcDNA	Relaxed circular DNA
RE	Restriction enzyme(s)
RI	Replicative intermediates

RNA	Ribonucleic acid
<i>rnd</i>	Random number
RP-HPLC	Reverse Phase High Pressure Liquid Chromatography
rpm	Revolutions per minute
RT	Room temperature
s	Second(s)
S	Sensitive
S/MAR	Structural and Matrix Attachment Regions
S/N	Supernatant
SA	Survival advantage
SA _{die}	Survival advantage of the cell nominated for cell death
SA _{live}	Survival advantage of the cell nominated to replace dying cell
SCC	Small cell change
SD	Standard Deviation
SDS	Sodium dodecyl sulphate
SHBs	Small Hepatitis B surface antigen
SINE	Short interspersed elements
SSC	Saline sodium citrate
TAE	Tris-acetate EDTA
TE	Tris-EDTA
TGF	Tumour Growth Factor
T _m	Melting temperature
TO	Turnover(s)
TOF	Time of Flight
TOF/TOF	Tandem Time of Flight
TRF1/2	Telomere Repeat Binding Factor1/2
Tris	2-amino-2-hydroxymethyl-1,3-propanediol
U	unit(s)
ULN	Upper limit normal
UV	Ultraviolet
V	Volt(s)
WHO	World health organisation
WHV	Woodchuck hepatitis virus
X-Gal	Bromo-chloro-indolyl-galactopyranoside
α -HBcAb	Antibodies specific against hepatitis B core antigen

α -HBeAb	Antibodies specific against hepatitis B e antigen
α -HBsAb	Antibodies specific against hepatitis B surface antigen
ϵ	Epsilon region

1 - Introduction

Hepatitis B is a prevalent human infection, of which the causative agent is hepatitis B virus (HBV). One third of the world's population has been exposed to HBV and ~400 million people are chronically infected (WHO 2000). Chronically infected patients have a 25% cumulative lifetime risk of death from either hepatocellular carcinoma (HCC) or cirrhosis. This translates to ~600000 million deaths annually attributable to HBV infection (WHO 2000; Lai 2002). HCC and cirrhosis generally manifest only after several years to decades of chronic HBV infection. The changes that occur in the hepatocyte population of patients over the course of chronic HBV infection are not yet fully understood. Therefore, one of the major aims of this research was to investigate the kinetics of hepatocyte repopulation in patients with chronic HBV infection using integrated HBV DNA as a marker for hepatocyte lineages.

1.1 Epidemiology of HBV infection

The worldwide distribution of HBV infection is heterogeneous, as shown in Figure 1.1. In developed regions, such as Western Europe, North America and Australia, chronic HBV infection occurs in <0.5% of the population; whereas in developing regions, such as in China, Southeast Asia and sub-Saharan Africa, chronic HBV infection occurs to 8-20% of the population (WHO 2000).

HBV is spread by blood and body fluid exchange, and is highly transmissible due to high virus titres, sometimes exceeding 10^9 HBV DNA copies per mL of serum (Jardi, Rodriguez et al. 2001; Stelzl, Muller et al. 2004; Allice, Cerutti et al. 2007). Common routes of infection include needle sharing, blood-to-blood contact and unprotected sex (Chen, Wang et al. 2000; WHO 2000; Farrell and Dienstag 2002; Lavanchy 2004).

1.2 HBV structure and replication

HBV, the prototypic member of the *Hepadnaviridae* family, is an enveloped, double-stranded DNA virus (Dane, Cameron et al. 1970; Tiollais, Charnay et al. 1981; Gust 1986). As shown in Figure 1.2, HBV virions are composed of a 32-36 nm icosahedral nucleocapsid surrounded by a 42nm host-derived lipid envelope, in which various forms of HBV surface antigen (HBsAg) are embedded. The nucleocapsid of HBV virions contains a ~3.2 kb relaxed circular DNA (rcDNA) genome, the structure of which is shown in Figure 1.3. During viral replication, subviral HBsAg particles are produced in excess of infectious particles and are secreted as 20 nm spheres or filaments of variable length (Dane, Cameron et al. 1970).

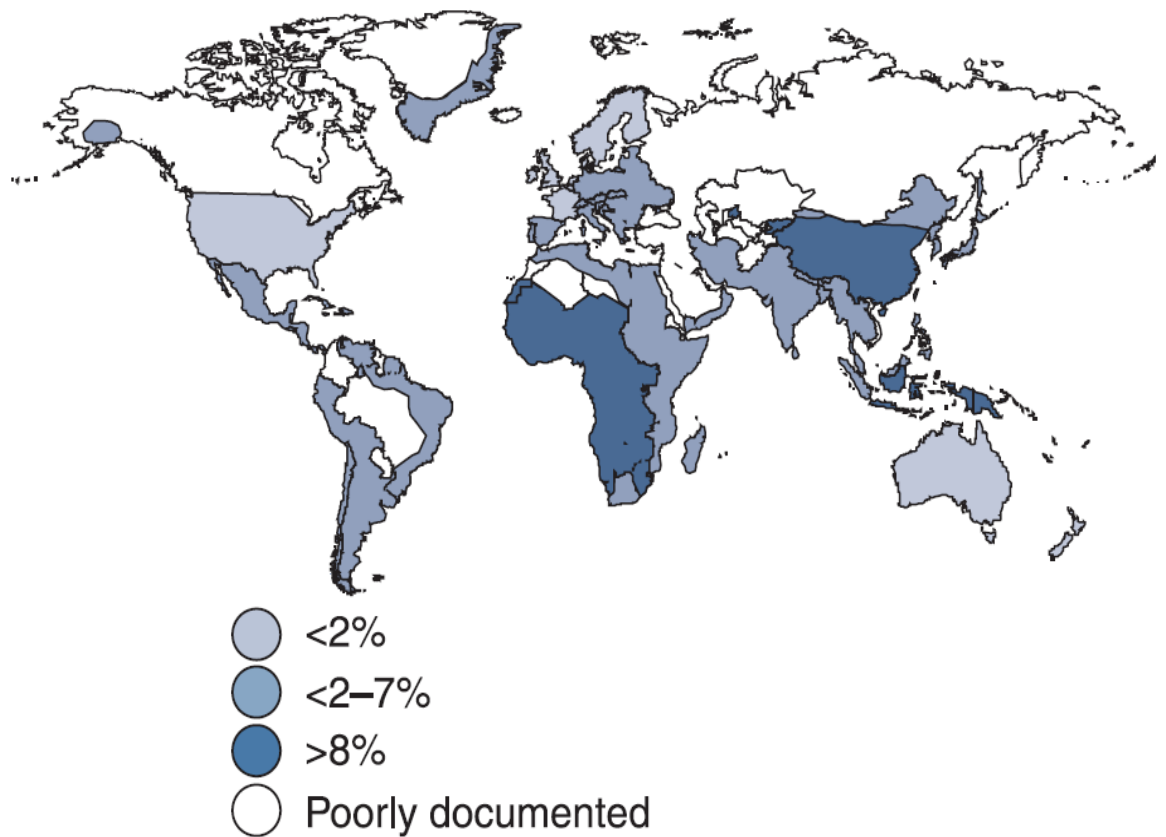


Figure 1.1. The global distribution of HBV infection. The map shows the worldwide prevalence of patients in the population positive for serum HBV surface antigen, an indicator of present HBV infection, is heterogeneously distributed throughout the world. Figure reproduced with permission from Lavanchy (2004). Data derived from the World Health Organisation (WHO, 2000).

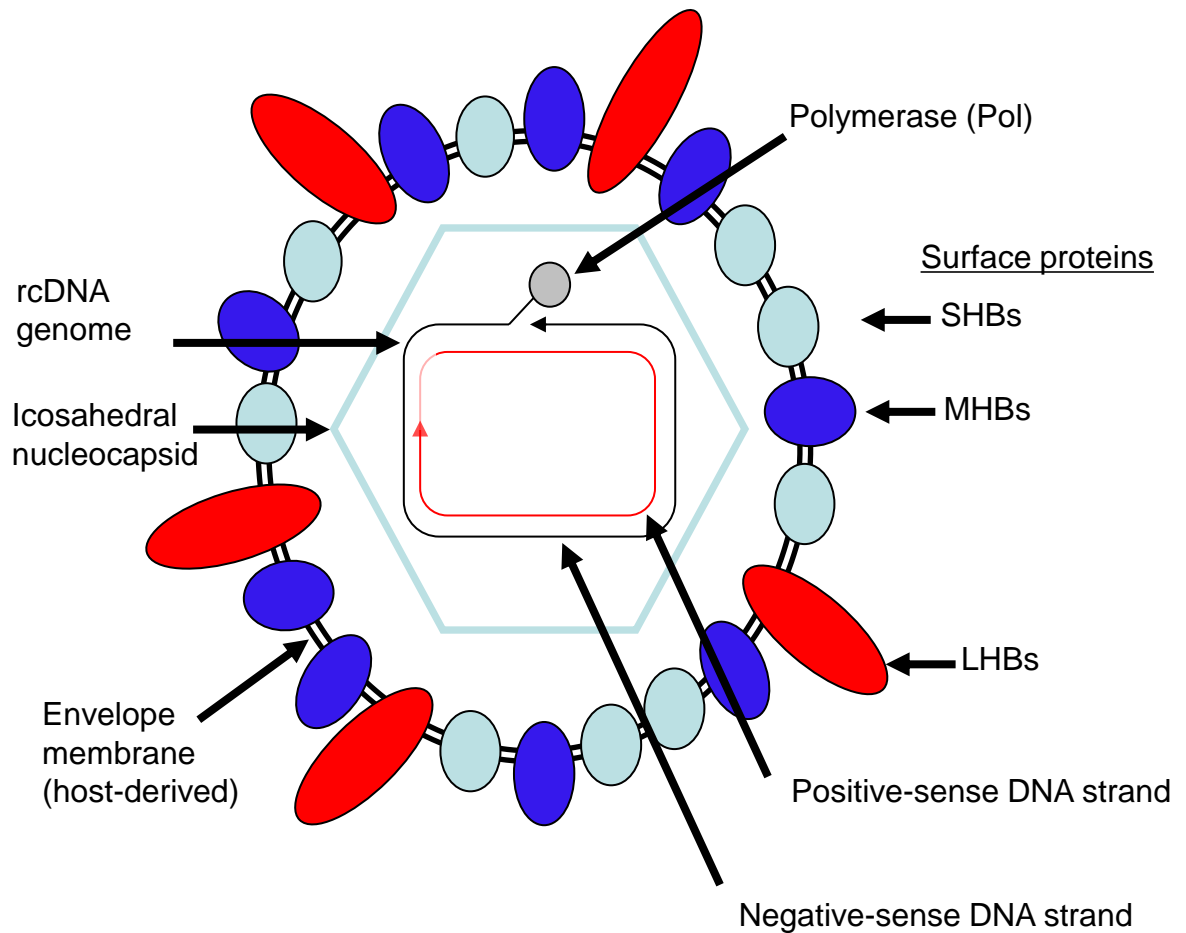


Figure 1.2. Schematic representation of the HBV virion. Large, medium and small HBV surface proteins (LHBs, MHBs, and SHBs respectively) are present in the virus envelope surrounding the icosahedral nucleocapsid (light blue hexagon) that contains the rcDNA genome. The HBV polymerase (Pol) (grey circle) is covalently attached to the 5' end of the negative-sense strand of the HBV DNA genome. Figure based on previous figure from Kann (2002).

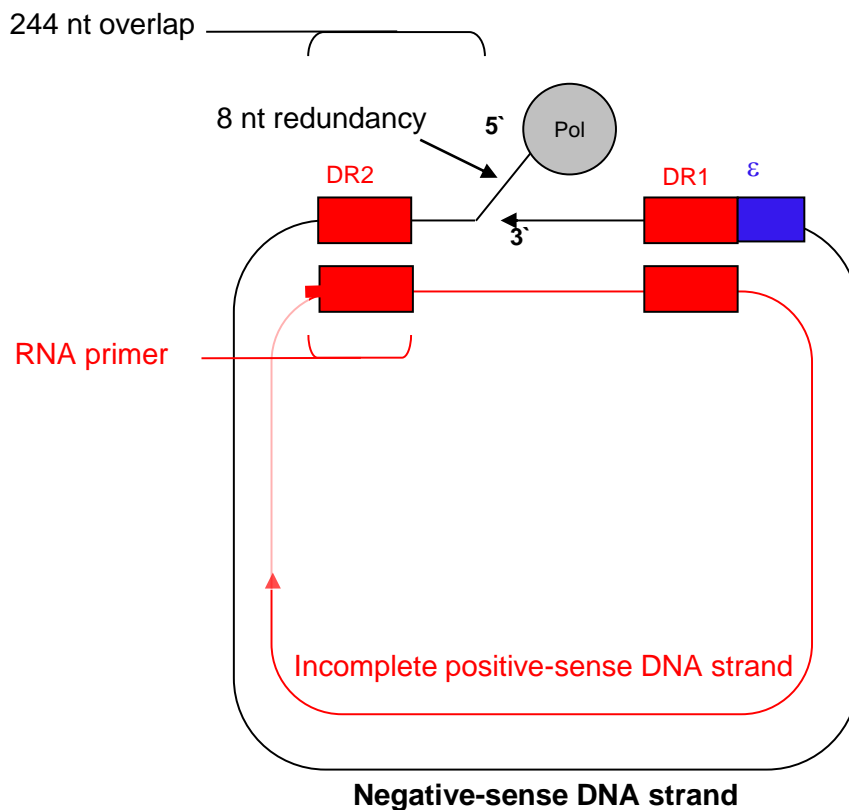


Figure 1.3. Schematic representation of the rcDNA HBV genome. The HBV genome is kept in an rcDNA form by a 244 bp overlap of the negative-sense (-) and positive-sense (-) DNA strands. The 5' end of the negative-sense DNA strand has an 8 nt redundant region. The HBV Pol is covalently attached to the 5' end of the negative-sense DNA strand by its Tyr₆₃ residue. The positive-sense DNA strand contains a 5' capped 18 nt oligoribonucleotide primer (labelled RNA primer) derived from pregenomic RNA (pgRNA). The 3' end of the positive strand is incomplete, 600-1200 nt short of the full length of the genome (Kann, 2002). Essential structures for HBV replication are shown: the epsilon region (ϵ , blue box), which forms a stem-loop secondary structure when transcribed into pgRNA; and two 12 nt direct repeat sequences (red boxes), designated DR1 and DR2. Figure based on previous figure from Kann (2002).

Subviral HBsAg particles do not contain a viral nucleocapsid and are therefore non-infectious.

Upon entering the bloodstream of a newly infected individual HBV virions preferentially bind to and enter hepatocytes through a series of unknown binding and attachment processes resulting in receptor-mediated endocytosis (De Meyer, Gong et al. 1997; Treichel, Meyer zum Buschenfelde et al. 1997; Paran, Geiger et al. 2001; Schulze, Gripon et al. 2007). The HBV rcDNA genome is transported into the nucleus where it is converted into covalently closed circular DNA (cccDNA) (Tuttleman, Pourcel et al. 1986; Newbold, Xin et al. 1995). cccDNA is a highly stable episomal molecule that acts as a template for the transcription of HBV mRNAs and pregenomic RNA (pgRNA) (Mason, Aldrich et al. 1982).

The positive-sense strand of cccDNA contains 4 overlapping open reading frames (ORF): surface, pre-core/core, polymerase and x ORF, as shown in Figure 1.4. The HBV mRNA transcripts have a common poly-adenylation site and are translated into 5 proteins: HBsAg, of which there are large (L-HBs), medium (M-HBs) and small (S-HBs) forms; HBV core antigen (HBcAg), which forms the HBV nucleocapsid; HBV e antigen (HBeAg), which is translated from pre-core/core transcript and is a secreted form of HBcAg; HBV x protein (HBx), which can transactivate certain cellular genes; and HBV polymerase (Pol), a viral polymerase that is essential for virus replication (Will, Reiser et al. 1987; Seeger and Mason 2000; Jilbert, Burrell et al. 2002).

pgRNA is 150 nt longer than HBV genomic length and, in addition to being an mRNA for HBcAg and Pol, acts as a template for reverse transcription to produce new HBV genomes (Buscher, Reiser et al. 1985). pgRNA transcripts have a poly A tail and contain an RNA stem-loop structure formed by an inverted repeat region termed the epsilon region (ϵ) (Knaus and Nassal 1993). As pgRNA is 150 nt longer than the HBV genome, it contains 2 copies of ϵ , near the 5' and 3' ends of the pgRNA transcript respectively. Furthermore, two 11 nt direct repeat (DR1 and DR2) regions exist on the 5' and 3' ends of the pgRNA and are crucial in HBV DNA replication (Loeb, Tian et al. 1996; Rieger and Nassal 1996). Again due to the longer-than-genome length of pgRNA, there are 2 copies of the DR1 region per pgRNA transcript. The relative positions of these features are shown in Figure 1.5A.

Following transcription of pgRNA, the binding of HBV Pol to the 5' ϵ induces the formation of immature nucleocapsids around the Pol-pgRNA complex (Bartenschlager, Junker-

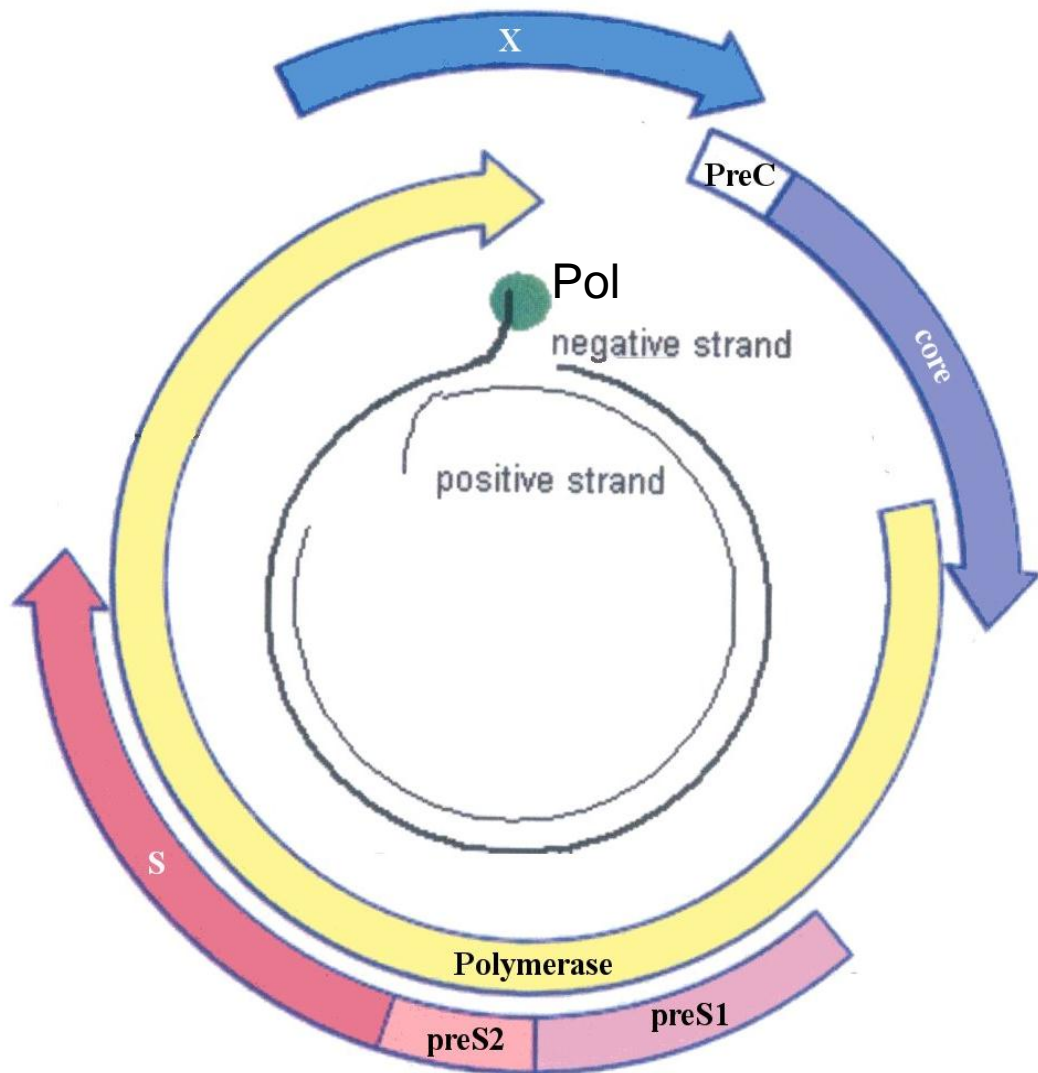


Figure 1.4. Overlapping open reading frames (ORF) and associated gene products present in the HBV genome. Coloured arrows point in the direction of translation. LHBs, MHBs and SHBs are translated from the preS1 + preS2 + S, preS2 + S and S alone ORF, respectively. Pol is encoded by the polymerase ORF. HBcAg is translated from the core ORF alone. HBeAg is formed by the cleavage of the translated product from PreC + core ORF. HBxAg is translated from the X ORF. Figure reproduced with permission from "Hepatitis B Virus, Human Virus Guide", CL Lai & S Locarnini (Eds.), 2002, © International Medical Press.

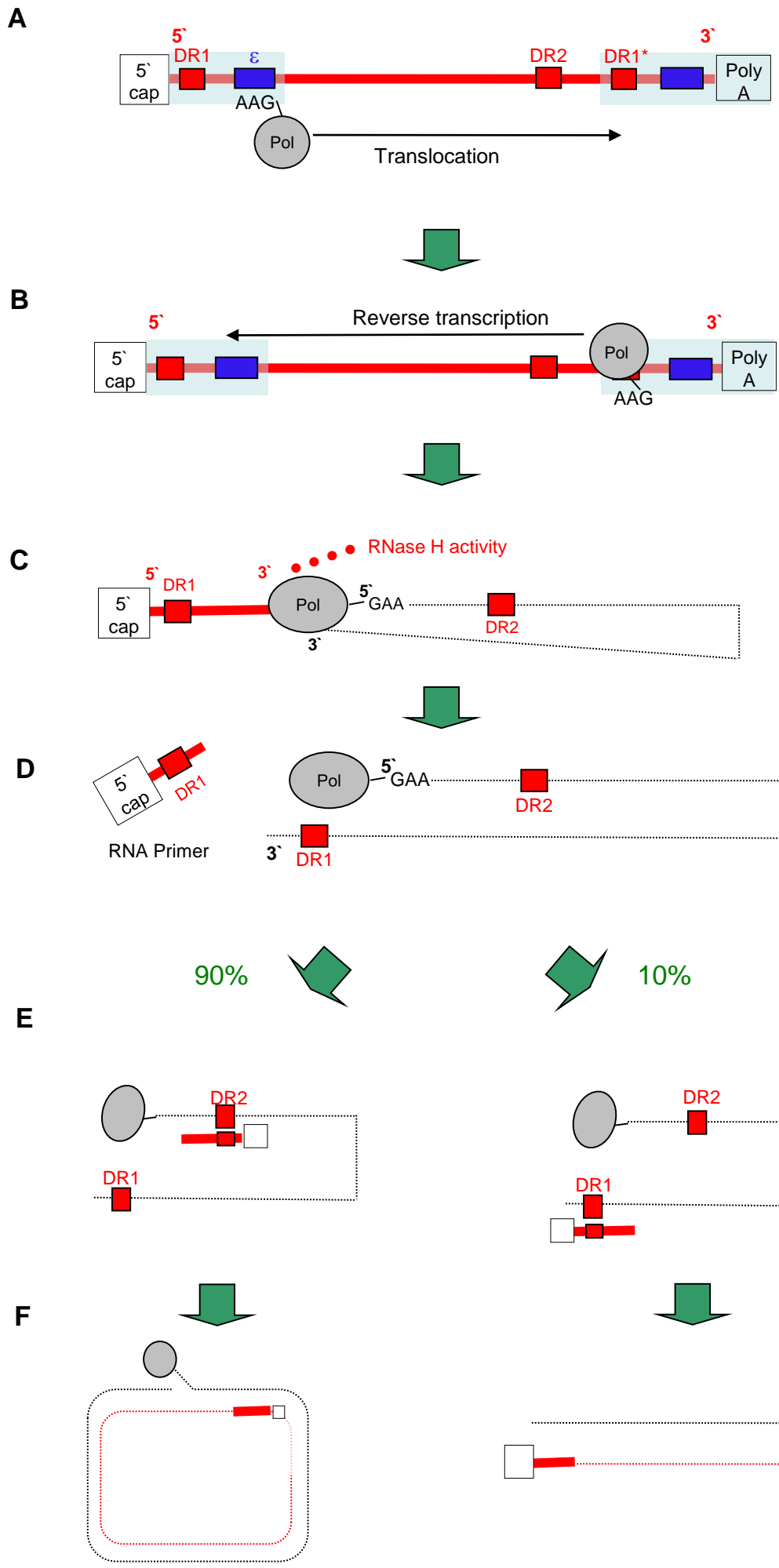


Figure 1.5. Schematic diagram of the reverse transcription of HBV. Reverse transcription of the 5'-capped pgRNA transcript (-) occurs within immature nucleocapsids. pgRNA is greater than genome length and thus contains redundant regions (light blue shading). During reverse transcription, Pol has three distinct functions: primer synthesis, RNA/DNA dependent DNA polymerisation, and RNaseH activity. After binding to the ϵ region of the pgRNA (1.5A), Pol synthesises a 3 nt oligonucleotide, 5'-GAA-3', using the ϵ region as a template (Kann, 2002). The trinucleotide primer is covalently attached to the Pol at its Tyr₆₃ residue. This nucleotide-pol complex then translocates to the DR1* region (1.5B). Pol reverse transcribes a negative-sense DNA strand (■■■) using the 3 nt oligonucleotide as a primer and the pgRNA as a template. The Pol hydrolyses the RNA template with its RNaseH activity lagging 18 nt behind the site of reverse transcription (1.5C). Since the Pol is covalently attached to the 5' end of the synthesised negative-sense DNA strand, a loop structure forms. The entire pgRNA transcript is hydrolysed up to 18 nt from the 5' end. The remaining RNA acts as a primer for the synthesis of the positive-sense DNA strand (■■■) (1.5D). ~90% of the time, the 18 nt RNA primer translocates to the DR2 sequence (1.5E), leading to the synthesis of rcDNA (1.5F). However, there is a 10% chance that the RNA primer does not move from the DR1 region (1.5G) in a phenomenon called “*in situ* priming”, and dsIDNA is formed instead (1.5H) (Yang *et al.* 1996). Figure based on previous figure by Jilbert *et al.* (2002).

Niepmann et al. 1990; Junker-Niepmann, Bartenschlager et al. 1990). Once encapsidation of the Pol-pgRNA complex occurs, a complex series of events (shown in Figure 1.5) leads to reverse transcription of the pgRNA by Pol into either an rcDNA or a double stranded linear DNA (dslDNA) genome (Seeger, Summers et al. 1991; Wang and Seeger 1992; Ganem, Pollack et al. 1994; Tavis, Perri et al. 1994; Loeb and Tian 1995).

Briefly, Pol has three functions: priming of reverse transcription; either DNA- or RNA-dependent DNA polymerisation; and RNaseH activity (Hirsch, Lavine et al. 1990; Radziwill, Tucker et al. 1990; Wang and Seeger 1992). The Pol primes DNA polymerisation by covalently attaching three nucleotides to its own particular tyrosine residue by using the 5' copy of ϵ as a template (Nassal and Rieger 1996). The Pol translocates to the 3' copy of the DR1 on the pgRNA transcript and begins RNA-dependent DNA polymerisation, using the pgRNA as a template (Rieger and Nassal 1996). Simultaneously, the RNaseH activity of the Pol degrades the pgRNA about 18 nt behind the site of polymerisation (Summers and Mason 1982). Thus, when the Pol reaches the end of the pgRNA, an 18 nt RNA strand remains, including the 5' copy of DR1. This RNA strand then normally translocates to DR2 to prime DNA-dependent DNA polymerisation to form rcDNA (Staprans, Loeb et al. 1991). About 10% of the time, translocation does not occur and dslDNA is formed. The HBV nucleocapsid then undergoes conformational changes and becomes a mature nucleocapsid.

The mature HBV nucleocapsid then either migrates to the nucleus (forming a pool of up to 30 copies of cccDNA per nucleus), or is enveloped in either the endoplasmic reticulum (ER) or intermediate-ER structures, and is exported from the cell via the Golgi apparatus (Patzer, Nakamura et al. 1986; Tuttleman, Pourcel et al. 1986; Huovila, Eder et al. 1992). Replication of HBV is not cytolytic *per se*, but immune responses against infected hepatocytes may induce cell death, as described in further detail in Section 1.3.

If dslDNA is transported to the nucleus of an infected hepatocyte, the HBV genome may randomly integrate into the host cell chromosome via non-homologous end joining (NHEJ) (Bill and Summers 2004). HBV DNA integration results in a replication-incompetent form of HBV, as longer than genome length pgRNA cannot be produced from integrated dslDNA. Integrated forms of duck hepatitis B virus (DHBV) DNA have been detected in 0.1-0.01% of hepatocytes in duck liver by 6 days post-infection (Yang and Summers 1999).

1.3 Natural history of HBV infection

1.3.1 Acute HBV infection

Acute HBV infections, by definition, are cleared by 6 months after initial infection. The serological pattern commonly seen in acute HBV infections is shown in Figure 1.6. As shown by chimpanzee models, during acute HBV infection virus may spread to infect almost the entire hepatocyte population (Berquist, Peterson et al. 1975; Murray, Wieland et al. 2005).

There is a lag period of 6-8 weeks after initial HBV infection before cell-mediated and humoral immune responses can be detected. Antibodies specific to HBcAg (α -HBcAb) are non-neutralising and are produced early in the course of HBV infection (Bowden 2002). Presence of serum α -HBcAb is simply a measure of past or current HBV infection. Antibodies specific to anti-HBeAg (α -HBeAb) are also produced and inflammatory cells infiltrate into the liver (Murray, Wieland et al. 2005; Asabe, Wieland et al. 2009). Jaundice and flu-like symptoms can occur in some patients during the immune response phase, but in many patients acute HBV infection is asymptomatic (Liaw, Yang et al. 1985).

Infiltrating CD8⁺ cytotoxic T-cells are the effector cells that lead to immune-mediated death of infected hepatocytes, although CD4⁺ helper T-cells are also required for efficient immune response and virus clearance (Bowden 2002; Asabe, Wieland et al. 2009). An overactive immune response against HBV-infected hepatocytes during this phase can cause extensive, multifocal necrosis and liver failure, termed fulminant hepatitis (Horney and Galambos 1977). Fulminant hepatitis is a rare, but deadly condition, occurring in ~1% of HBV patients, of which 53-83% die as a result or require liver transplantation (Bernuau, Goudeau et al. 1986; Tandon, Acharya et al. 1996).

Neutralising anti-HBsAg antibodies (α -HBsAb) are eventually raised by the humoral immune response and confer protective immunity, and so act as a clinical marker of successful virus clearance. After the HBV infection is cleared, there are no known long-term pathologies associated with acute HBV infections (Bowden, 2002). The mechanisms that lead to the successful clearance of HBV infection are still unclear. It appears that neutralising α -HBsAb is associated with clearance.

However, in 5-10% of unvaccinated adults and 90% of unvaccinated children less than 1 year of age, clearance of HBV DNA from the liver does not occur and a chronic HBV infection develops (Hyams 1995). Chronic HBV infection is distinct from acute HBV infection in that:

Acute infection phases

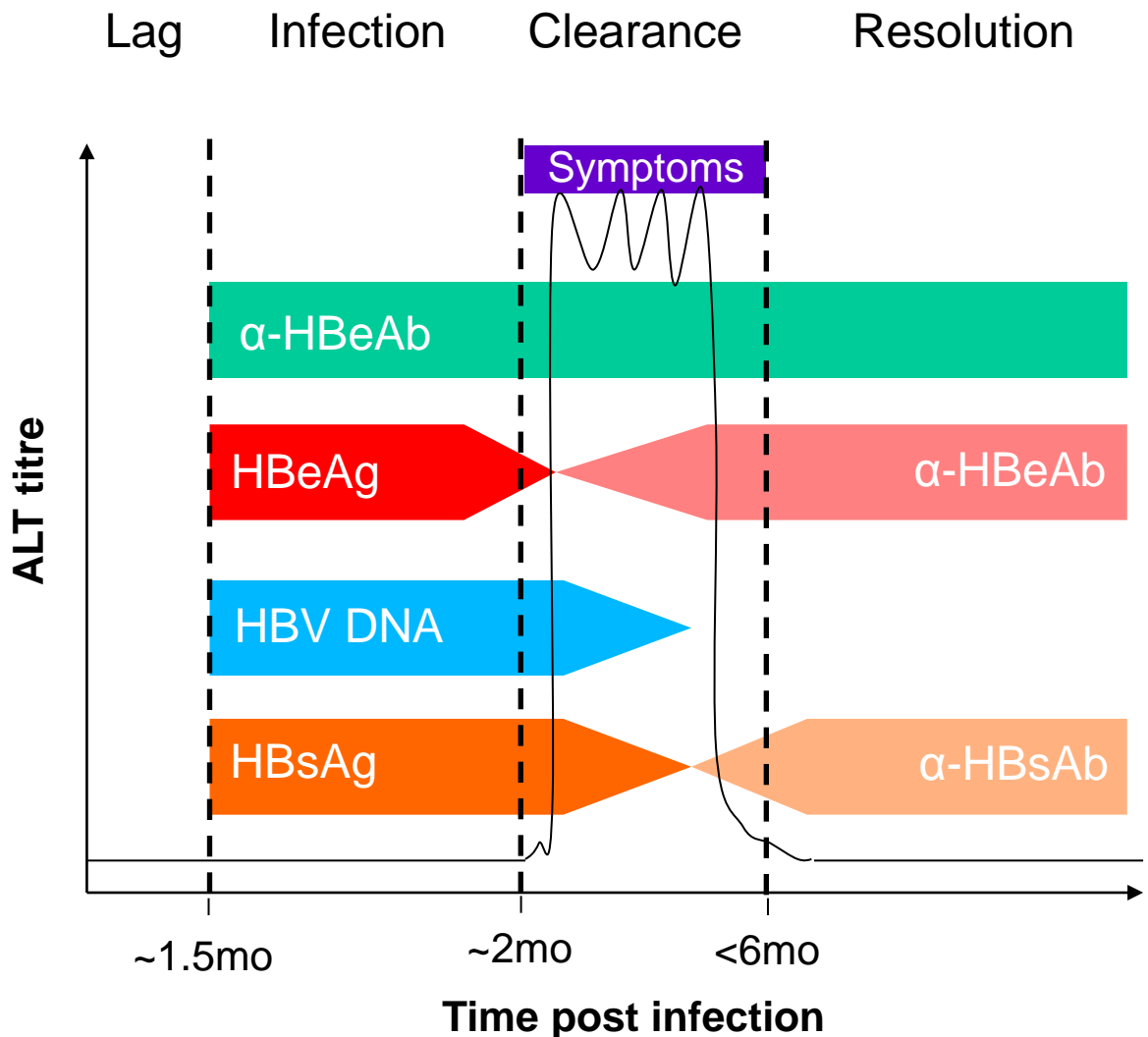


Figure 1.6. Serological and virological serum profile during acute HBV infection. In the lag phase of HBV infection, no serological markers or HBV DNA is detectable in the serum. After ~45 days serological and virological markers are detectable in the serum. Non-neutralising α -HBcAb are raised early in infection. HBeAg is observed early in infection, but seroconversion to α -HBeAb occurs during the clearance phase. Loss of HBeAg is associated with an increase in circulating serum ALT, symptomatic disease and a reduction in HBV DNA levels in the blood. With this drop in circulating virus, circulating HBsAg are neutralised by α -HBsAb, a marker of the resolution of HBV infection. Figure adapted from Bowden (2002).

circulating HBV virions and HBsAg particles can be detected for >6 months; recognition of infected hepatocytes by the immune system takes decades, compared to weeks in acute HBV infection; and the increased risk of cirrhosis and HCC compared to uninfected individuals (Yim and Lok 2006).

Chronic HBV infection has been categorised into four phases (although not all patients go through all four): immune tolerance, immune clearance, immune control, and reactivation. The serological profile of these phases is described in Figure 1.7.

1.3.2 Chronic HBV infection - Immune tolerance phase

The immune tolerance phase is characterised by high serum HBsAg levels and HBV DNA levels (up to 10^{10} copies of HBV copies/mL of serum). There is limited immune response against virus antigens, leading to low levels of hepatocyte death and normal serum amino liver transferase (ALT) levels (a marker for liver cell damage) (Yim and Lok 2006; Yuen and Lai 2008). In experimental chimpanzee models of HBV infection, both HBsAg and HBcAg can be detected in >95% of hepatocytes in early immune tolerance phase (Wieland, Spangenberg et al. 2004; Murray, Wieland et al. 2005; Asabe, Wieland et al. 2009). HBsAg localises to cellular membranes and cytoplasm, while HBcAg localises to the nucleus and/or the cytoplasm (Gowans and Burrell 1985), whereas HBeAg can be detected in the serum by ELISA (Yuen and Lai 2008).

The mechanism(s) of immune tolerance still remain unclear. HBeAg has been suggested as an inducer of immune tolerance (Milich and Liang 2003). Indeed, significant immune attack against hepatocytes during HBV infection is associated with the production of α -HBeAb. Furthermore, HBeAg-negative HBV mutants have been associated with several outbreaks of fulminant hepatitis (Kosaka, Takase et al. 1991; Liang, Hasegawa et al. 1991; Omata, Ehata et al. 1991; Sato, Suzuki et al. 1995).

Recently, HBcAg has also been suspected of acting as a tolerogen for HBV infection. A significantly greater expansion of CD4⁺ FoxP3⁺ regulatory T cells occurred after HBcAg stimulation *in vitro* of peripheral blood mononuclear cells (PBMC) of from chronically HBV-infected patients compared to patients who were immune to HBV (Barboza, Salmen et al. 2007). However, whether the expanded cells were functionally active was not studied.

Chronic infection phases

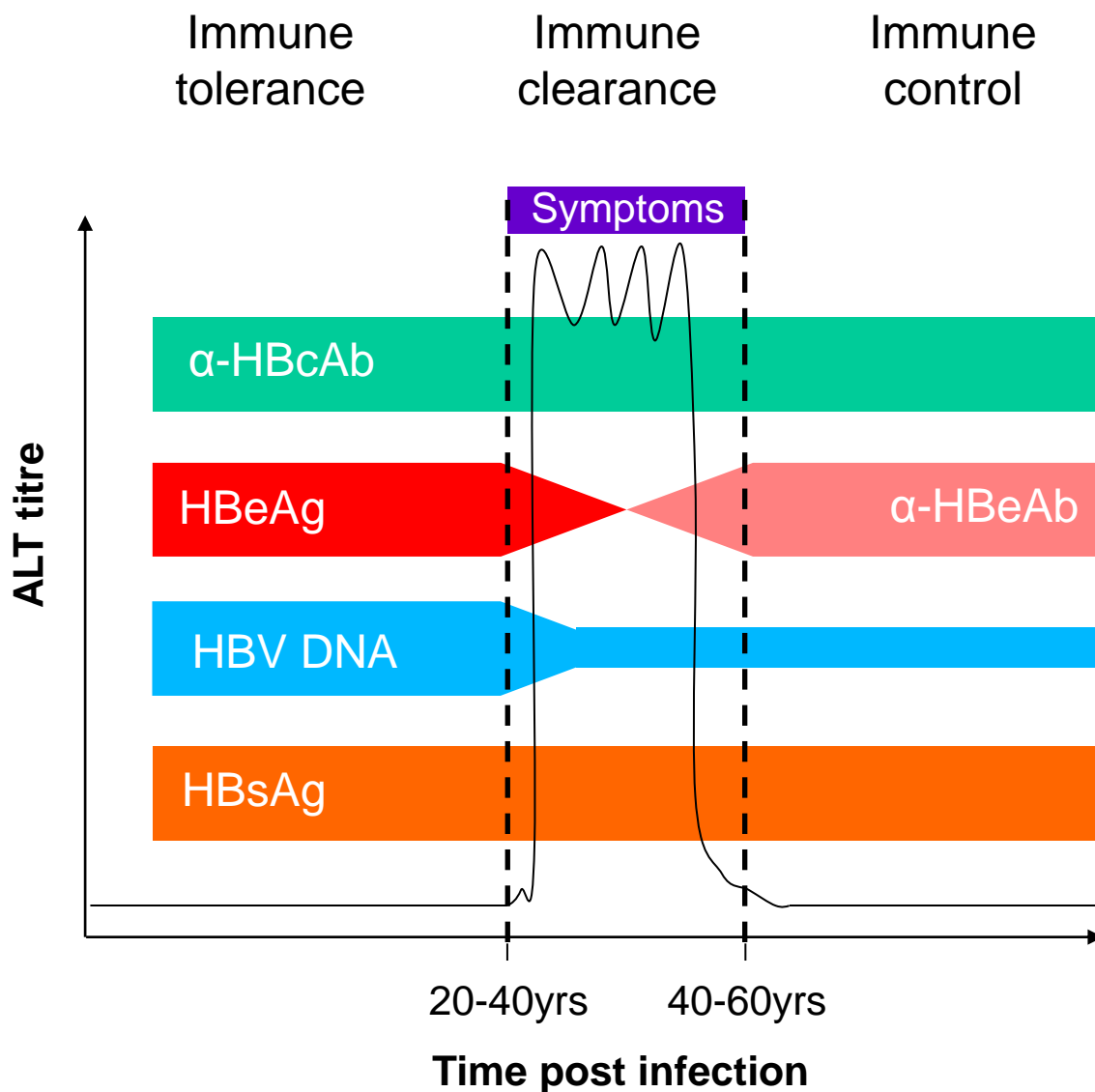


Figure 1.7. Serological and virological serum profile during chronic HBV infection. The immune tolerance phase of chronic HBV infection is serologically identical to the infection phase of acute HBV infection (shown in Figure 1.6), except that it persists for a much longer period, up to 40 years after the initial infection. Once an immune response against the HBV-infected hepatocytes is elicited, seroconversion of HBeAg to α -HBeAb can occur along with increasing serum ALT levels and the reduction, but not complete resolution, of HBV DNA. HBsAg remains prevalent in the serum of chronic HBV patients throughout infection. Figure adapted from Bowden (2002).

Anergy of T-cells may play a role in immune tolerance. Activated T-cells can be induced to become anergic through the Programmed-Death-1 (PD-1)/PD-1 ligand (PD-1L) pathway, as reviewed in (Keir, Francisco et al. 2007). Various chronic virus infections, including hepatitis C virus (HCV), human immunodeficiency virus (HIV) and indeed HBV infections, have been attributed to upregulation of this pathway (Golden-Mason, Palmer et al. 2007; Ha, West et al. 2008; Peng, Li et al. 2008). In particular for chronic HBV infection, two independent groups have reported a positive correlation between serum HBV DNA and expression of PD-1 in HBV-specific CD8⁺ T-cells (Evans, Riva et al. 2008; Peng, Li et al. 2008). Evans *et al.* also demonstrated that the CD8⁺ T-cells of 6/18 patients who progressed from the immune tolerance phase of chronic HBV infection had significantly down-regulated PD-1 expression (Evans, Riva et al. 2008).

The immune tolerance phase is not usually associated with liver disease and is commonly asymptomatic (Lindberg and Lindholm 1988; de Franchis, Meucci et al. 1993; Boxall, Sira et al. 2004; Hui, Leung et al. 2007). Case reports of children in the immune tolerance phase who have HBV-associated HCC exist, but are rare (Shimoda, Uchida et al. 1980; Tanaka, Miyamoto et al. 1986). A longitudinal study of 240 Taiwanese adults with chronic HBV infection in the immune tolerance phase revealed that the incidence of HCC was 0% and cirrhosis was 0.5% per year over 10.5 years (Chu, Karayiannis et al. 1985). This is lower than the average annual incidence of HCC (1-3%) and cirrhosis (2-6%) in adults with chronic HBV infection who are not in the immune tolerance phase (Villeneuve 2005; Yim and Lok 2006). Additionally, a study examining liver biopsy tissue from 40 patients in the immune tolerance phase found that all 40 patients had a METAVIR score of F0-F1, representing little or no fibrosis (Andreani, Serfaty et al. 2007). Chronically HBV-infected patients generally remain in the immune tolerance phase for 10-40 years.

1.3.3 Chronic HBV infection - Immune clearance phase

Through unknown trigger(s), the patient immune system recognises HBV antigens and mounts an immune response against HBV-infected hepatocytes. This heralds entry into the immune clearance phase of chronic HBV infection. As explained earlier, immune-mediated cell death occurs through HBV antigen-specific CD8⁺ cytotoxic T-cells activated by CD4⁺ helper T-cells (Eddleston and Mondelli 1986; Bowden 2002; Bertoletti and Gehring 2006). Concurrently, B-cells producing antibodies specific against HBV antigens, particularly HBeAg, play a crucial role in the immune clearance phase.

During the immune clearance phase when there is seroconversion from HBeAg to α -HBeAb, CD8⁺ cytotoxic T-cells specific for HBcAg, HBsAg and Pol antigens are raised, suggesting that HBeAg induces immune tolerance. For example, PBMC specific against HBsAg and HBcAg have been observed in 20% and 82% respectively of patients who have recovered from acute HBV infection, when tested by an *ex vivo* peptide-stimulated proliferation assay (Penna, Artini et al. 1996). Also, patients who had either cleared HBV infection or had HBeAg-negative chronic HBV infection have significantly higher levels of HBcAg-specific PBMC compared to those from patients without HBV infection (Vassilopoulos, Rapti et al. 2008). Furthermore, Mizukoshi *et al.* found that the blood of HBeAg-negative patients with chronic HBV infection contained greater numbers of peripheral T-cells that responded to HBV Pol stimulation compared to a HBeAg-positive cohort (~40% versus ~10%) of chronically HBV-infected patients (Mizukoshi, Sidney et al. 2004). Thus the loss of HBeAg often coincides with increased levels of immune attack against HBV-infected hepatocytes.

This immune-mediated destruction of hepatocytes causes flares of serum ALT levels and symptomatic liver disease (Tsai, Chen et al. 1992). Serum ALT levels are generally negatively correlated with serum HBV DNA levels. While the severity of flares has been positively correlated to the risk of HCC and cirrhosis (Yim and Lok 2006; Villa and Fattovich 2010), a large study showed that patients with ALT levels that were <2 times upper limit of normal (ULN) ALT levels are most at risk of developing HCC or cirrhosis (Lai, Hyatt et al. 2007; Tsang, Trinh et al. 2008; Wang, Lim et al. 2008). In addition, patients with ALT levels >0.5-1 times the ULN had the same risk of complications as those with ALT levels >2-6 times the ULN (Yuen, Yuan et al. 2005). This may be because the continual inflammation and immune attack on HBV-infected hepatocytes sets up conditions for the development of HCC and cirrhosis. This hypothesis is supported by the observation that chronic HBV patients with consistently high HBV DNA levels are more likely to progress to HCC and cirrhosis (Chen, Yang et al. 2006; Park, Kim et al. 2010). Furthermore, in patients with HCC and/or cirrhosis, high serum HBV DNA levels are predictive of a worse clinical outcome (Ohkubo, Kato et al. 2002; Ma, Wei et al. 2008; Yeh, So et al. 2010).

Humoral responses are also raised against HBV antigens during the immune clearance phase, in particular, the production of α -HBeAb. As mentioned previously, serum α -HBeAb is often associated with increased immune attack against HBV infected hepatocytes. Sustained seroconversion from serum HBeAg to α -HBeAb represents a watershed moment in the

progression of chronic HBV infection from the immune tolerance to the immune control phase.

1.3.4 Chronic HBV infection - Immune control phase and reactivation

The immune control phase represents the state wherein the immune system has successfully destroyed many of the hepatocytes that support HBV replication. It is characterised by both serum α -HBeAb and low to undetectable levels of serum HBV DNA. Serum ALT levels usually return to normal, although flares of HBV replication followed by spiking serum ALT levels resulting from immune-mediated cell death may still occur.

HBeAg-negative patients in the immune control phase are at lower risk of developing HCC or cirrhosis compared to HBeAg-positive patients (Yim and Lok 2006; Liaw, Lau et al. 2010). Chu *et al.* found that seroconversion to α -HBeAb at younger ages significantly decreased the incidence of cirrhosis in a cohort of 240 patients with chronic HBV infection after a 17 year follow-up period (Chu, Hung et al. 2004). In a chronically HBV infected cohort of 233 interferon-alpha and 233 untreated matched controls, both treated and untreated patients who seroconverted to α -HBeAb had significantly lower incidence of both HCC and cirrhosis compared to untreated HBeAg-positive patients after 17 years follow up (Lim and Kowdley 2007).

However, serum α -HBeAb does not provide absolute protection against HCC and cirrhosis. In an α -HBeAb-positive cohort, patients with raised ALT status active hepatitis were ~3 times more likely to develop HCC and cirrhosis than patients with inactive hepatitis after a median follow up period of 8.6 years (Hsu, Chien et al. 2002).

Most patients remain in the immune control phase, but some individuals can undergo reactivation. This may occur if the patient is immunocompromised through HIV infection, immunosuppressive medication or natural waning of the immune system over time (Hoofnagle 2009). This leads to sudden flares of serum and liver HBV DNA levels. HBV reinfects the hepatocytes and, due to existing partial immunity, HBV strains mutations in the core/pre-core ORF that prevent HBeAg production can be selected (Frelin, Wahlstrom et al. 2009). The survival rates of patients with reactivated HBV are worse than those in other stages of HBV infection (Sharma, Saini et al. 2005).

A minority of chronic HBV patients (~1% per year) undergo HBsAg seroconversion and produce α -HBsAb. As explained above, α -HBsAb are neutralising and protective against infection (Bortolotti, Guido et al. 2006). Serum α -HBsAb usually coincides with undetectable levels of HBV antigen-positive hepatocytes and undetectable levels of serum HBV DNA.

Changes in the hepatocyte population may also play a role in HBV clearance. For example, HBV negative hepatocytes can occur in the presence of viraemia (Burrell, Gowans et al. 1984; Huang, Hung et al. 2010; Mason, Liu et al. 2010), suggesting that some hepatocytes become resistant to HBV infection.

While the risk of disease is greatly reduced in patients that have developed α -HBsAb and cleared HBV infection, progression to HCC is still possible, especially if cirrhosis has been established by the time seroconversion occurs (Villa and Fattovich 2010). Greater age at HBsAg seroconversion and high inflammation in the liver are both positive predictors of progression to HCC.

1.3.5 Disease outcomes of chronic HBV infection

The most notable disease outcomes associated with chronic HBV infection are cirrhosis and HCC. An in-depth meta-study of epidemiological studies carried out in the Asia-pacific and Africa has shown that patients with chronic HBV infections have an annual incidence of cirrhosis and HCC of 1-2% and 0.27-1%, respectively (Lin, Robinson et al. 2005). This translates to ~600,000 deaths annually attributable to HBV infection (WHO 2000; Lai 2002).

Chronic HBV infection is frequently associated with fibrosis and cirrhosis. Fibrosis is the deposition of collagen strands in portal tracts and alongside hepatocytes (Crawford 2007). Cirrhosis occurs when the collagen strands become so prevalent as to divide the liver parenchyma into distinct nodules, leading to loss of hepatocytes, structurally abnormal nodules, and altered architecture, vascular anatomy and microcirculation (Anthony, Ishak et al. 1977; Anthony, Ishak et al. 1978). Patients with cirrhosis have decreased liver function, which leads to decompensated liver failure in 3-4% of cirrhotic patients each year (Liaw, Lin et al. 1989; Xu, Hu et al. 2003). Furthermore, patients with cirrhosis also progress to HCC at an annual rate of 3-7% (Lin, Robinson et al. 2005).

Chronic HBV infection is associated with a ~100-fold increased risk of HCC compared to uninfected controls (Beasley 1988). While cirrhosis can cause progression to HCC, reported

estimates of up to 50% of HCC cases are not associated with concurrent cirrhosis (Reviewed in (Fattovich, Stroffolini et al. 2004)). HCC represents a serious danger to chronic HBV patients as the annual mortality rate of patients with HCC is estimated to be ~50% (Lin, Robinson et al. 2005). The direct mechanisms by which HBV drives the progression of cirrhosis and HCC are not yet fully understood.

1.4 Liver anatomy and microstructure

As mentioned previously, hepatocytes are the main targets for HBV infection and are therefore targets of immune-mediated cell death during HBV clearance. Hepatocytes comprise the majority (~70%) of the cells in the liver. Hepatocytes are arranged in single-cell-thick plates in a sponge-like pattern (Roskams, Desmet et al. 2007). The liver is supplied with blood by two routes: 1) nutrient-rich blood from the intestines via the portal vein; and 2) oxygen-rich blood from the heart via the hepatic artery.

Figure 1.8 shows a common model of the liver microstructure, involving the division of the liver into functional subunits called lobules (Matsumoto and Kawakami 1982; Fawcett 1994; Teutsch, Schuerfeld et al. 1999; Teutsch 2005). In these subunits, a single central vein is surrounded by a number of portal veins, all draining blood into the central vein. Hepatocytes radiate from the central vein and form 3 general zones: Zone 1, the periportal hepatocytes; Zone 2, those further from the portal tracts; and Zone 3, the pericentral hepatocytes. Hepatocytes from each zone experience different levels of oxygen and nutrients, producing enzymatically- and metabolically-heterogeneous hepatocytes (Khan and Angus 1980; Jungermann 1988; Jungermann and Katz 1989). It is important to note that, however, there are no physical divisions between lobules in the human liver (Elias and Sokol 1953; Elias and Sherrick 1969; Matsumoto and Kawakami 1982).

1.5 Hepatocyte repopulation in the liver

If damaged, the liver can regenerate to replace dying hepatocytes. Hepatocyte transplantation into mice with defective hepatocyte development has shown that hepatocytes can undergo at least 69 mitotic events and still produce viable progeny hepatocytes *in vivo* (Rhim, Sandgren et al. 1994).

Hepatocyte regeneration is a complex and heavily regulated process. Liver transplants have been shown to regrow or shrink to maintain an age-, gender- and weight-dependent size of

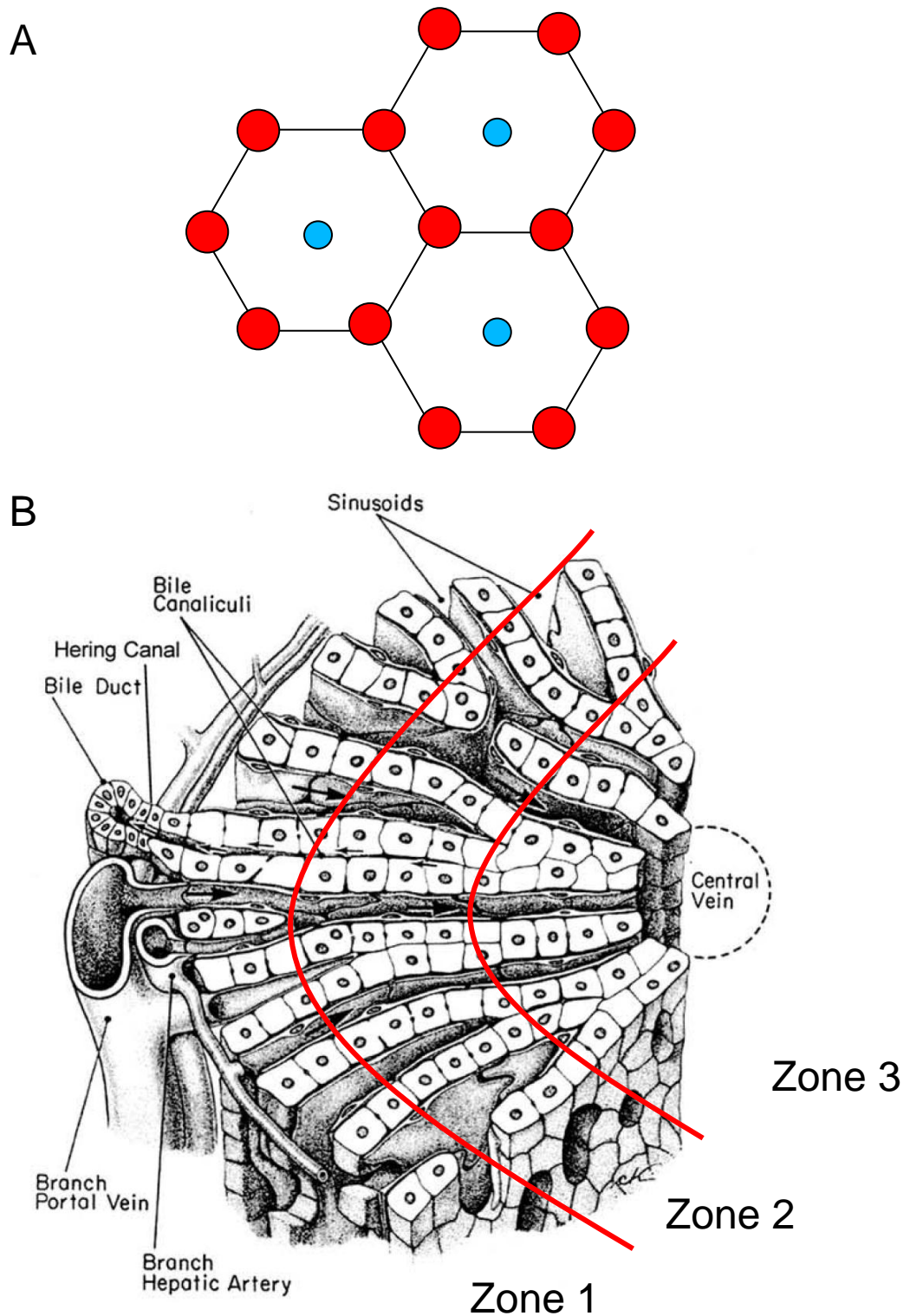


Figure 1.8. Lobule subunit microstructure. Lobules have been described as fundamental subunits of the liver. The relative arrangement of three lobules is shown in Figure 1.8A. A lobule is composed of 3-8 portal tracts (PT, red circles), draining into a common central vein (CV, blue circles). A more detailed diagram of the lobule is shown in Figure 1.8B. Blood flows (arrows) from the portal tracts, through sinusoids, across the hepatocytes and is eventually carried away into the central vein. Zones 1, 2 and 3 represent areas in which blood becomes progressively poorer in nutrients and oxygen. Figure adapted from Bloom and Fawcett, *A Textbook of Histology*, 10th edition with permission.

~2.5% of total body weight in liver transplant recipients (Van Thiel, Gavaler et al. 1987). Hepatocyte regeneration is also controlled by highly redundant mechanisms; no single gene knockout has been shown to completely eliminate hepatocyte regeneration in mice that have undergone two-thirds hepatectomy. However, the knockout of key genes has delayed hepatocyte regeneration (Koniaris, McKillop et al. 2003; Fausto, Campbell et al. 2006). Furthermore, in studies of partial hepatectomy performed in rats, blood-borne factors were not sufficient to induce hepatocyte proliferation in untreated rats (Rogers, Shaka et al. 1961).

The mechanism by which hepatocytes are repopulated during normal liver turnover is still a controversial topic. Several researchers propose the “streaming liver” model, wherein bipotent stem cells (oval cells), which are located close to bile-duct cells near the portal vein, supply the parenchyma with hepatocytes. This hypothesis was supported by observing the livers of rats injected with tritiated thymidine (Zajicek, Oren et al. 1985). As the rats grew older, the percentage of tritium-positive hepatocytes decreased near the portal vein (zone 1) and increased closer to the central vein (zone 3).

However, strong evidence against the “streaming liver” model was observed in a later study (Bralet, Branchereau et al. 1994). Rats underwent partial hepatectomy and were injected 24 hours later with a retrovirus containing an *E. coli* beta-galactosidase gene (*lacZ*). In some hepatocytes, *lacZ* was integrated into the host cell genome and expressed. Positive hepatocytes were seen in clusters equally distributed between zones 1-3, even >1 year later. The observations by Zajicek *et al.* (1985) were explained as being an artefact of the utilisation of labelled thymidine by the peri-central (zone 3) hepatocytes (Bralet, Branchereau et al. 1994; Grisham 1994). Peri-central hepatocytes undergo mitosis later than peri-portal hepatocytes, giving the impression of streaming hepatocytes.

Recent observations based on mutations in mitochondrial genes have revived the “streaming liver” model. Patches of hepatocytes with mutations in the cytochrome c oxidase (COX) gene have been detected by immunohistochemistry and PCR in normal human liver (NHL) (Fellous, Islam et al. 2009). The patches of hepatocytes were abutting the portal vein and contained the same mutation in the COX gene, suggesting that the hepatocytes were clonal. The authors hypothesise that to build up such numbers of mitochondria with the same mutation in all of their genomes would require a long-lived, stem cell population and this population is responsible for repopulation of the hepatocytes.

In a competing “random replacement” model, if minor liver damage occurs, neighbouring hepatocytes will respond by proliferating to replace nearby dead hepatocytes (Bralet, Branchereau et al. 1994). In cases where <10% of the liver is damaged, the mitosis of the hepatocytes beyond the local area of injury is not significantly altered (Macdonald, Rogers et al. 1963). However, if >10% of the liver is damaged, as seen in a model of partial hepatectomy in rats for example, widespread hepatocyte mitosis is induced (Bucher and Swaffield 1964). This widespread mitosis of hepatocytes was found to be inversely proportional to the distance from the hepatectomy site. In a porcine ¼ hepatectomy model, a greater number of hepatocytes adjacent to the surgical site were observed to have incorporated the thymidine-analogue bromodeoxyuridine than those distant from the surgical site (Sun, Chow et al. 1999).

Thus, the mechanisms of hepatocyte repopulation during chronic HBV infection remain controversial with these two competing models: the “streaming liver” model and the “random replacement” model. Hepatocyte repopulation is a focus of the research described in this thesis as it determines the liver microenvironment and thus may affect disease progression. Therefore, various methods have been used in this thesis and other studies described below in Section 1.6 to analyse hepatocyte lineages within normal and HBV-infected livers to help understand the evolution of the hepatocyte population.

1.6 Tracing hepatocyte lineages

In differentiated female cells, either maternal or paternal X-chromosomes are randomly inactivated and the inactivated copy is passed on to progeny cells. Studies have been conducted to detect the pattern of inactivation of X-linked genes in groups of hepatocytes. Using this method, studies have reported that approximately half of all regenerative nodules in HBV-infected cirrhotic liver tissue are clonal (Aihara, Noguchi et al. 1994). Also, this technique was used in conjunction with mass-spectrometry to show that expression of three unidentified protein peaks was associated with cirrhotic nodules that were shown to be clonal (Guedj, Dargere et al. 2006).

However, there are many limitations in using this approach to interpret the extent of clonal proliferation of hepatocytes. The assay differentiates hepatocyte clones using only two states and this limits accurate quantification of hepatocyte clone sizes. Also, only liver tissue from female patients heterozygous for X-linked genes can be assayed. This limits its application in HBV research, as the majority of cases of HBV-associated late-stage disease occur in males

(McMahon 2009; Schutte, Bornschein et al. 2009). Furthermore, X-chromosome inactivation in human somatic tissues occurs in the blastocyte and embryo stages of development (Park 1957; Park 1957), so these studies conflate the clonal proliferation associated with both development and HBV-associated disease.

As mentioned previously in Section 1.5, clonal proliferation of hepatocytes has been detected using polymorphisms in mitochondrial DNA as markers of progeny of hepatocytes in which mutations had occurred. Patches of hepatocytes negative for COX have been observed in the liver of HBV-uninfected patients (Fellous, Islam et al. 2009). Hepatocytes isolated from the same patches contained identical mutations in the COX gene, indicating that they were clonal. Similar studies conducted on hepatocytes in the livers of patients with cirrhosis showed that regenerative nodules are also clonal (Bioulac-Sage and Balabaud 2010; Lin, Lim et al. 2010).

Hypermethylation of the promoters of host genes has also been used as a marker for analysis of hepatocyte clonality. In eukaryotic genomes, gene promoter regions contain a greater concentration of cytosine-guanosine dinucleotide sequences (5'-CG-3') than the average over the entire genome (Reviewed in (Robertson and Jones 2000)). These concentrations of 5'-CG-3' are dubbed "CpG islands". Addition of a methyl group to the 5' carbon of cytosine residues mediates pre-transcriptional silencing of the downstream gene. 5'-CG-3' sequences not associated with CpG islands are methylated in normal physiological conditions. DNA methylation is mediated by host DNA methyltransferases (DNMT). DNMT2 and DNMT3 both methylate non-methylated 5'-CG-3' sequences *de novo*; whereas DNMT1 maintains methylation in post-mitotic genomes by converting CpG methylation on a single DNA strand into fully methylated 5'-CG-3' sequences. This methylation is inherited by progeny cells. Using this fact, tumours arising from the same progenitor cell population were confirmed by the hyper-methylation of a panel of 11 tumour suppressor genes in patients with multi-focal HCC (Nomoto, Kinoshita et al. 2007).

The previous three methods for detecting clonal proliferation have the common problem of determining the starting time of clonal proliferation. It is difficult to pinpoint exactly when X-chromosome gene inactivation, mutation in COX or hypermethylation of tumour-suppressor genes may have occurred during the lifetime of the patients. These assays could be influenced by the clonal proliferation of hepatocytes that occurs during liver growth and development. Thus, the influence of HBV infection on the clonal proliferation of hepatocytes is more difficult to measure by these methods.

Several groups have used integrated hepadnaviral DNA as a marker for hepatocyte lineages in both chronically and acutely HBV-infected animals and humans (Hino, Kitagawa et al. 1984; Tanaka, Esumi et al. 1988; Esumi, Tanaka et al. 1989; Yasui, Hino et al. 1992; Mason, Jilbert et al. 2005; Mason, Liu et al. 2010). Any clones detected by using integrated hepadnaviral DNA must have been formed after the initiation of HBV infection.

As discussed above, HBV dsDNA is occasionally produced during HBV replication. If the mature nucleocapsid containing dsDNA migrates back into the nucleus, HBV dsDNA can integrate into the cellular genomic DNA by NHEJ (Bill and Summers 2004). As shown in Figure 1.9, exposed ends caused by rare random breaks in the host cell chromosome and the exposed ends of the HBV dsDNA were found to act as the substrates to NHEJ. This observation has been strengthened by numerous independent studies showing that HBV integration sites are highly localised about the expected ends of the dsDNA form of HBV DNA, i.e. nt 1819 and nt 1826 in the HBV genome (Genbank Accession #AB241115) (Yaginuma, Kobayashi et al. 1985; Takada, Gotoh et al. 1990; Staprans, Loeb et al. 1991; Yang and Summers 1999; Kimbi, Kramvis et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). Therefore, HBV DNA integration occurs infrequently and in random sites of the host cell genome. Unique integrations can thus be used as a genetic marker for progeny hepatocytes derived from a unique hepatocyte in which the integration of HBV DNA occurred.

In the 1980s, several studies investigated liver tissues from patients with HCC by Southern blot hybridisation using a DNA probe specific for HBV DNA (Shafritz, Shouval et al. 1981; Hino, Kitagawa et al. 1984; Chen, Harrison et al. 1988; Esumi, Tanaka et al. 1989). All the hepatocytes of a given HCC tumour were observed to contain the same HBV DNA integration, while distinct from any other HBV DNA integration events detected in any other HCC. This was expected as tumours are presumed to be derived from a single hepatocyte in which cumulative oncogenic mutations had occurred. Furthermore, clones of hepatocytes in non-tumour tissue were detected by Southern blot hybridisation in chronically HBV-infected patients (with and without HCC) in 38 of 78 cases (Esumi, Tanaka et al. 1989).

In a small proportion (0-5.5%) of patients, who were in immune tolerance phase without HCC, liver DNA extracts were found by Southern blot hybridisation to contain integrated HBV DNA. Liver DNA extracts from these patients produced a diffuse HBV DNA signal of

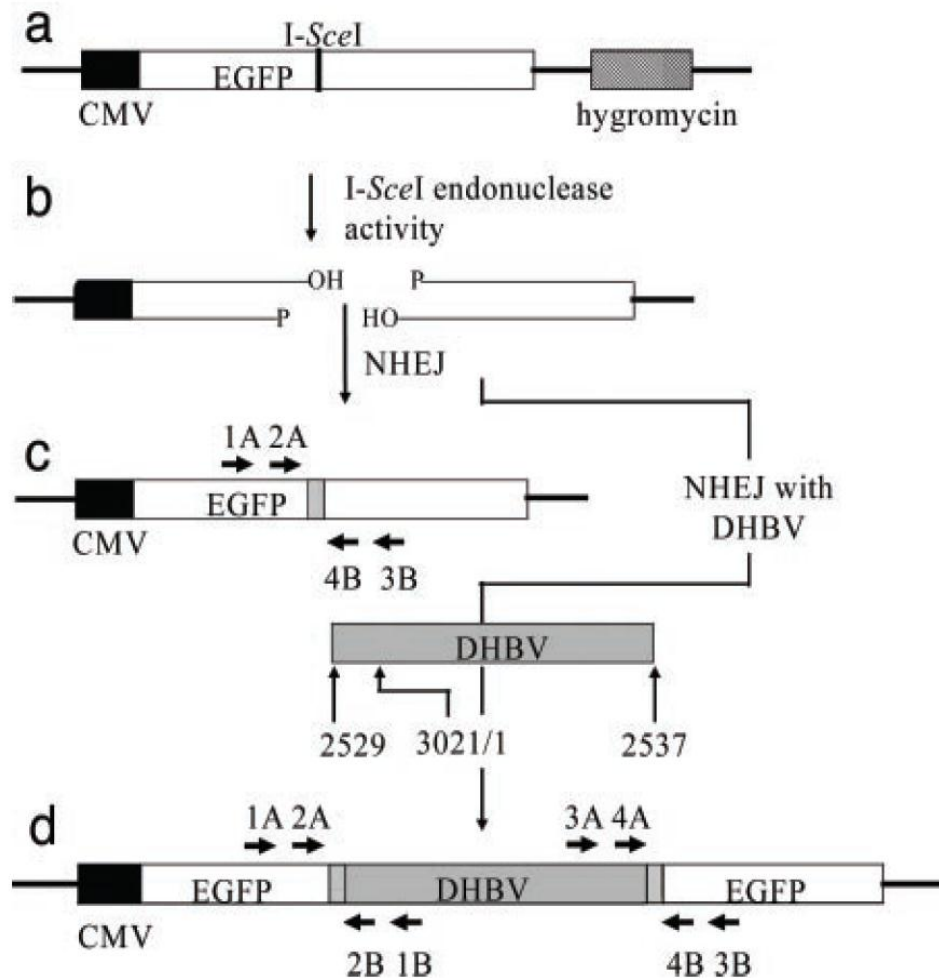


Figure 1.9. Integration of viral DNA sequences into cellular DNA by non-homologous end-joining (NHEJ). Bill and Summers (2004) showed that double-stranded breaks were targets of virus DNA integration inserting the 18 nt sequence of the restriction enzyme *I-SceI* into the genome of chicken hepatoma LMH cells (1.9A). Enhanced green fluorescent protein (EGFP)- and hygromycin-coding genes were also inserted as selection markers. When *I-SceI* was introduced into the nucleus, the DNA was restricted at the *I-SceI* site, as shown in Figure 1.9B. Cell repair mechanisms repaired this DNA break as shown in Figure 1.9C. In the process, small insertions (short grey box) were detected by PCR using the primers shown (short arrows), which is characteristic of NHEJ.

When the cells were transfected with a DHBV and *I-SceI*-coding DNA construct, insertions of DHBV dsDNA (long grey box) were detected by nested DHBV primer pairs. Nucleotide numbers above is based on the DHBV DNA sequence previously described in Mandart et al. (1984). As shown in 1.9D, virus-cell DNA junctions at the left hand end of DHBV DNA were detected using primer pairs 1A and 1B, and then 2A and 2B. Similarly, virus-cell DNA junctions at the right hand end of DHBV DNA were detected using primer pairs 3A and 3B, and then 4A and 4B. Again, small insertions indicating NHEJ were observed.

Figure is reproduced with permission from Bill and Summers (2004), © 2004 National Academy of Sciences, USA.

>3300bp, suggesting that while random HBV DNA had integrated, little clonal proliferation had occurred (Shafritz, Shouval et al. 1981; Hino, Kitagawa et al. 1984; Chen, Harrison et al. 1988). Fowler *et al.* detected a signal >3300 bp in liver DNA from 6/36 non-HCC patients in immune control stage, 4 of which showed distinct bands (Fowler, Greenfield et al. 1986).

The size of these clones were not quantified, but were necessarily large in order to be detected by the low sensitivity Southern blot hybridisation assay. The highest sensitivity of Southern blot hybridisation is 10-1000 fg (Tham, Chow et al. 1991; Leenman, Panzer-Grumayer et al. 2004), which is equivalent to 10^3 - 10^5 copies of the 3.2 kb HBV DNA genome. Thus, the data from the Southern blot hybridisation analyses of liver DNA extracts suggest that some clonal proliferation of hepatocytes occurs before tumour formation.

Greater sensitivity in detection of HBV integration has been achieved using PCR-based techniques. PCR using primers specific for HBV and Alu elements have been used to amplify virus-cell DNA junctions in both HCC tissue and hepatocytes of patients with HBV infection (Kimbi, Kramvis et al. 2005; Murakami, Saigo et al. 2005). Alu elements are common short interspersed nuclear elements (SINEs) present in primate genomes and number ~1000000 copies per human genome (Lander, Linton et al. 2001). These studies found HBV integration sites were observed close to the ORF of genes associated with HCC progression. However, Alu PCR may be inherently biased as Alu elements in the human genome are concentrated near ORFs (Grover, Majumder et al. 2003).

Yang and Summers developed an inverse nested PCR (invPCR) technique to detect integrated DHBV DNA fragments in DNA extracted from DHBV-infected duck liver (Yang and Summers 1999). Figure 1.10 shows a detailed explanation of the invPCR protocol. Restriction enzymes (RE) are used to excise the virus-cell DNA junction formed by virus DNA integration. Then, through a series of ligation and digestion reactions, the unknown cellular sequence is flanked with known viral sequences. The unknown cellular sequence included in the virus-cell DNA junction can then be determined by PCR using virus specific primers.

InvPCR has also been used to detect integrated WHV DNA in wild-infected woodchucks (Mason, Jilbert et al. 2005). Hepatocyte clones composed of >10000 cells containing the same virus-cell DNA junctions were observed. Clones of this size were predicted by computer models to be formed stochastically after ~124 liver turnovers. Proliferating cell nuclear antigen (PCNA) staining indices of liver biopsies taken from the WHV-infected woodchucks

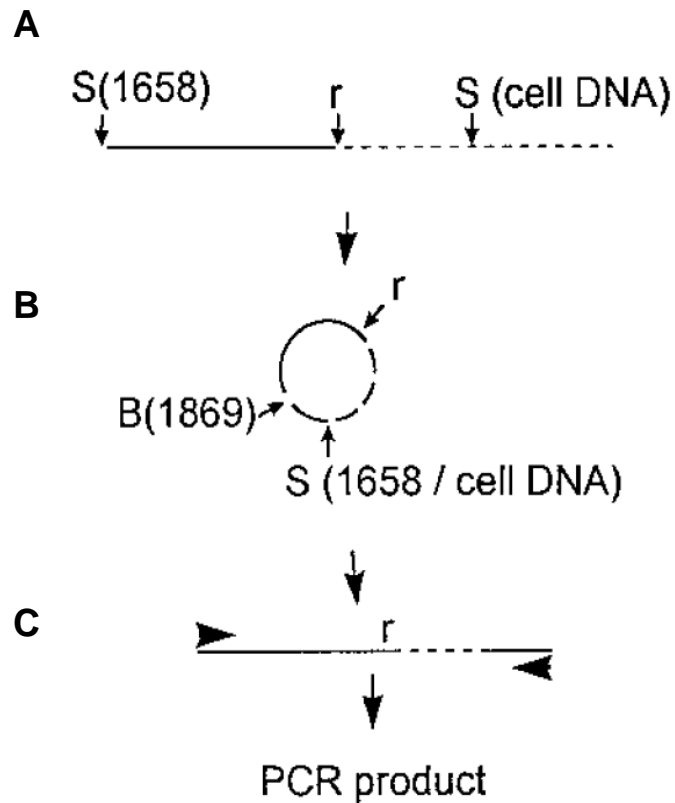


Figure 1.10. Inverse nested PCR. Nucleotide numbers above are based on the DHBV DNA sequence previously described in Mandart *et al.* (1984). As shown in Figure 1.10A, the virus-cell junction (*r*) is excised by digestion with *Sau3AI* at specific *Sau3AI* sites (*S*) in integrated DHBV sequence (solid line) at nt 1658 and at an unknown downstream site in the cellular DNA sequence (dashed line). The *Sau3AI* ends are ligated together and form a recombinant circle of DNA as shown in Figure 1.10B. This circle of DNA is linearised by digestion with *BspHI* at nt 1869. The resultant inverted linear DNA shown in Figure 1.10C has an unknown cellular sequence flanked by two known virus sequences. The inverted DNA are titrated and then amplified via nested PCR using primers designed in known virus regions. Figure reproduced with permission from Yang and Summers (1999).

suggested only 6-12 turnovers had occurred over their lifetime. Thus, these observations suggested that the large hepatocyte clones were composed of cells with a survival advantage in the context of the WHV-infected liver. Further work from this group showed that liver tissue from experimentally HBV-infected chimpanzees and from non-cirrhotic human patients with HCC also contain hepatocyte clones composed of >10000 cells (Mason, Low et al. 2009; Mason, Liu et al. 2010). This suggested that the clonal proliferation of hepatocytes with a survival advantage is present in all hepadnavirus infections. However, the phenotype of the cells that contain the virus-cell DNA junctions remains unknown.

The studies described in this thesis of changes in the hepatocyte population over the course of HBV infection aim to explain the long-term disease progression during chronic HBV infection, which include cirrhosis and HCC.

1.7 Histology of HCC and HCC-associated lesions

The most reliable method for HCC diagnosis is histological examination of a liver biopsy. Despite having highly variable histology (even within the same tumour), cells within HCC generally contain at least some of the following changes: high nuclear to cytoplasmic ratio; prominent nuclei and nucleoli; nuclear irregularity; moderate eosinophilia; and/or finely granular cytoplasm (Goodman and Terracciano 2007). Also, normal single cell-thick sinusoid structures are replaced with ≥ 2 cell-thick contorted cords, due to rapid, uncontrolled cellular proliferation.

Particular changes in the histological phenotype of hepatocytes have been observed before the development of HCC proper. What has been termed “large cell change (LCC)” has been associated with HCC development (Lee, Tsamandas et al. 1997; Goodman and Terracciano 2007). LCC is defined as hepatocyte enlargement with normal nuclear:cytoplasmic ratio, cellular hyperchromasia, pleomorphic or multiple nuclei, or any combination of the above. Despite the association with HCC, foci of hepatocytes with LCC express senescence-associated markers, such as p21, and have little to no proliferative activity (Lee, Tsamandas et al. 1997; Ikeda, Sasaki et al. 2009). Therefore, LCC probably do not represent a precursor for HCC development. Indeed, large cell changes are observed in disease states with low HCC risk, such as biliary liver disease and Wilson's disease, as well as HCC and cirrhosis (Le Bail, Bernard et al. 1997; Su, Benner et al. 1997).

Small cell changes (SCC), on the other hand, are more likely candidates as precursors for malignant growth. SCC are defined as cells of decreased size, higher than normal nuclear:cytoplasmic ratio; and nuclear pleomorphism. These hepatocytes generally have a greater proliferative index than normal, unlike the senescent hepatocytes with LCC (Watanabe, Okita et al. 1983; Le Bail, Bernard et al. 1997). Indeed, hepatocytes with small cell changes are associated with HCC in cirrhotic patients and increase in number upon development of HCC (Su, Benner et al. 1997; Su and Bannasch 2003). Other authors have shown that telomeres are shortest in HCC tissues, intermediate in hepatocytes with SCC, and only slightly shorter in hepatocytes with LCC compared to histologically normal hepatocytes (Kim, Oh et al. 2009).

Foci of altered hepatocytes (FAH) have been found to be greater predictors of low or high risk for HCC than SCC alone in patients with cirrhosis (Su, Benner et al. 1997). FAH is a general term for clusters of hepatocytes (hepatocyte foci) with altered histology. Several phenotypes of cell foci have been previously described (Bannasch 1986): clear cells containing high levels of glycogen; acidophilic cells containing high levels of both glycogen and smooth endoplasmic reticulum; tigroid cells containing low levels of glycogen and high levels of rough endoplasmic reticulum; basophilic cells containing no build-up of glycogen but high levels of rough endoplasmic reticulum; and intermediate cells containing intermediate levels of rough and smooth endoplasmic reticulum and glycogen. Mixed cell foci (i.e. foci of hepatocytes populated with >1 of the aforementioned cell types) were most predictive of HCC risk (Enzmann and Bannasch 1987; Su, Benner et al. 1997).

However, all of these dysplastic changes in hepatocytes are simply risk factors. In a study of 33 cirrhotic patients with viral hepatitis, 29 out of 33 dysplastic foci did not appear to progress to HCC by ultrasonic observation (Seki, Sakaguchi et al. 2000). Fifteen dysplastic foci disappeared altogether. Thus, molecular changes have been investigated to supplement histological data on HCC identification and progression.

1.8 Molecular changes during HCC progression

The molecular changes associated with the initiation and progression of HCC can come in many interconnected forms: epigenetic changes, such as gene promoter hypermethylation; altered expression of RNA transcripts; or changes in protein expression (Cha and Dematteo 2005; Feo, Frau et al. 2009).

As mentioned in Section 1.6, CpG methylation of genetic promoters can silence downstream transcription. A wide range of gene promoters have been found to have altered methylation status in HCC (Tischoff and Tannapfe 2008). Methylation of the adenomatous polyposis coli (APC) gene promoter has been shown to be significantly increased in low-grade dysplastic nodules compared to cirrhotic nodules in HBV-infected patients (Um, Kim et al. 2010). Promoters of the negative regulators of cell growth p16, HAI2 and RASSF1A were found to be methylated in 90% of HCC tissues and 56% of histologically-normal neighbouring tissues (Su, Zhao et al. 2008). These promoters are usually slightly methylated or unmethylated in uninfected liver tissues. Progressive hypermethylation in the promoter of RASSF1A from normal human tissue to dysplastic nodules to HCC has been shown by methylation CpG island amplification microarray (Gao, Kondo et al. 2008). Indeed, methylation-sensitive PCR of HCC tissue found that the greatest differences in hypermethylation compared to normal liver occurred in promoters for APC (81.7%), GSTP1 (76.7%), RASSF1A (66.7%), p16 (48.3%) (Lee, Lee et al. 2003). These studies suggest that CpG methylation of certain gene promoters is an early event in HCC formation.

While CpG islands in specific promoters are hypermethylated in HCC, inter-island 5'-CG-3' sequences are hypomethylated (Tischoff and Tannapfe 2008). This has been observed by numerous techniques. Total levels of 5'-methyl cytosine were significantly lower in DNA extracted from HCC compared to normal, non-cirrhotic and cirrhotic liver tissue (Lin, Hsieh et al. 2001). Also, DNA extracted from samples of HCC digested with the methylation-sensitive RE *HpaII* had significantly greater levels of fragment-end nucleotide incorporation compared to DNA from normal uninfected liver, indicating global DNA hypomethylation (Calvisi, Ladu et al. 2007). Finally, using combined bisulfite restriction analysis, long interspersed nuclear elements (LINE) promoter methylation was significantly increased in liver DNA extracted from rats with choline-induced HCC compared to untreated rats (Asada, Kotake et al. 2006).

Differential expression of RNA in HCC tissue compared with NHL tissue has also been investigated. Numerous studies using microarray analysis of mRNA transcripts have found hundreds of transcripts differentially expressed between HCC, HBV-infected tissue and NHL (Kim, Ye et al. 2004; Yuen, Yuan et al. 2005; Ieta, Ojima et al. 2007). These studies into differential mRNA expression have led to several candidate proteins for both the detection of HCC and the mechanisms of carcinogenesis.

A clinically relevant serum marker of HCC is alpha fetoprotein (AFP), which is currently used to monitor HCC formation in high-risk individuals (Yao, Dong et al. 2007; Marrero, Kudo et al. 2010; Stefaniuk, Cianciara et al. 2010). However, the positive and negative predictive values at ~70-60% and ~70-80% respectively is insufficient in many cases to accurately differentiate HCC and non-HCC populations (Soresi, Magliarisi et al. 2003; Snowberger, Chinnakotla et al. 2007; Murugavel, Mathews et al. 2008). Furthermore, the expression of AFP provides few hints into the mechanism behind HCC development.

Alterations in proliferation- and apoptosis-regulation pathways, such as the tumour growth factor (TGF)-, retinoblastoma (RB1)- and p53-dependent pathways, have been associated with HCC (Azechi, Nishida et al. 2001; Edamoto, Hara et al. 2003; Ito, Nishida et al. 2004; Rodrigues and Barry 2011). One of the canonical changes associated with up to 90% of all HCC of any etiology is altered Wnt pathway expression (Wong, Fan et al. 2001; Thompson and Monga 2007). This pathway regulates the expression of many proliferation-, metabolism- and restructuring-associated genes, such as *c-myc*, *c-jun*, cyclin D1 and matrix metalloproteases (Edamoto, Hara et al. 2003; Giles, van Es et al. 2003; Calvisi, Ladu et al. 2007; Lee, Park et al. 2007).

Several proteins are known to alter expression of the Wnt pathway. The dysregulation of the Wnt pathway has been shown in mice with liver-specific inactivation of APC, leading to HCC in 6/10 animals (Colnot, Decaens et al. 2004). Indeed, methylation of the APC promoter or APC inactivation has been shown to be common in HCC of various aetiologies including HBV-infection (Edamoto, Hara et al. 2003; Csepregi, Rocken et al. 2008). Deletion mutations in β -catenin also activate the Wnt pathway and are common, with up to 26% of human HCC tumour tissues of unknown aetiology having nuclear accumulation of mutated β -catenin (de La Coste, Romagnolo et al. 1998). Other groups showed that nuclear mutated β -catenin was associated with increase cellular proliferation, vascular invasion and tumour recurrence (Nhieu, Renard et al. 1999; Zulehner, Mikula et al. 2010).

However, alteration in the Wnt pathway is not sufficient for HCC, as the activation of the Wnt pathway via altered p53, APC, or β -catenin expression was observed in benign liver neoplasms (Chen, Jeng et al. 2002). Also, alteration in the Wnt pathway is significantly less common in HBV-associated HCC compared to other aetiologies, suggesting further changes in cell phenotype are necessary to produce HBV-associated HCC (Hsu, Jeng et al. 2000; Wong, Fan et al. 2001).

Expression of proteins, including heat shock protein 70, glypican-3 and glutamine synthetase, are altered early in carcinogenesis and have been described as candidate markers for early HCC (Lopez 2005; Sakamoto 2009). Genes, including upregulation of human telomere reverse transcriptase (hTERT) and downregulation of the Wnt pathway-associated protein E-cadherin, have been reported to be differentially expressed between early HCC and dysplastic foci in HCV-infected patients (Llovet, Chen et al. 2006).

During the development of HCC, some molecular changes in hepatocytes occur even before histological evidence of transformation (Trevisani *et al.* 2008; Hui 2009; Sakamoto 2009). For example, hTERT expression in dysplastic nodules was observed to be twice that of NHL (Llovet *et al.*, 2006). Another study found that the expression of telomere repeat binding factors 1 and 2 (TRF1/2) and TRF1-interacting nuclear protein were increasingly upregulated during the progression from normal liver tissue to dysplastic nodules to early HCC to late HCC (Oh, Kim et al. 2005). Additionally, Jun activation domain binding protein-1 and p27 were highly expressed in 50% of dysplastic nodules and early HCCs, but not expressed in histologically normal hepatocytes (Yachida, Imaida et al. 2010). However, in all of these studies, the gene expression differences were highly variable in the observed dysplastic nodules.

In summary, while there are common changes in particular pathways that drive HCC progression (e.g. RB1-, p53-, and Wnt-dependent pathways), changes in expression of these oncogenes are not present in all HCCs (Levy, Renard et al. 2002; Edamoto, Hara et al. 2003). Furthermore, changes in dysplastic nodules that are presumed precursors for HCC show even greater variability in expression. Also, the molecular expression profile of HCCs can vary depending on its aetiology (Delpuech, Trabut et al. 2002; Iizuka, Oka et al. 2004; Chen, Ho et al. 2010). This suggests that development of HCC may take multiple routes to arrive at the same outcome, as opposed to a linear path of mutations that must be navigated.

1.9 HBV and HCC progression

One of the strongest predictors for HCC is chronic infection with HBV and/or HCV (Goodman and Terracciano 2007; Tan, Yeh et al. 2008). In particular, chronic HBV infection is associated with a ~100-fold increased risk of HCC compared to uninfected controls (Beasley 1988). Within HBV-infected patients, factors that increase risk of HCC include the age of the patient, continual low levels of liver inflammation, HBeAg-positive serum and

persistently high HBV DNA serum levels ($>10^6$ copies/mL) (Chemin and Zoulim 2009). However, the HBV-associated mechanisms that drive HCC progression are multi-faceted, complex, and not yet fully elucidated. Proposed mechanisms include: direct and indirect oncogenesis caused by HBV products; alterations in DNA sequence and stability due to HBV integration; and induction of carcinogenesis as a bystander effect of the host immune response against HBV.

1.9.1 The role of HBV proteins in HCC progression

Firstly, the HBV components HBx and HBsAg have been reported to induce the progression of HCC. While HBx is a necessary factor for HBV replication (Tsuge, Hiraga et al. 2010), its exact role during replication is unknown. The major reported role for HBx is transactivation of host proteins, by which it may induce the initiation and progression of HCC (Feitelson, Reis et al. 2009; Brechot, Kremsdorf et al. 2010; Arzumanyan, Friedman et al. 2011). For example, expression of HBx in HepG2 cells increased degradation of β -catenin, a regulator of the Wnt pathway (Chen, Hu et al.). Furthermore, expression of HBx has been associated with upregulation of vasoinvasion-associated genes HIF1- α , metastasis-associated protein 1 and histone deacetylase 1 in both primary human tumour tissues and transgenic mouse models (Xie, Song et al. 2008; Yoo, Na et al. 2008).

However, while a large range of functions have been attributed to HBx, many of these studies involve overexpression of HBx outside of physiological conditions. For example, studies have suggested that HBx increases hTERT transcription (Lin, Chen et al. 2005; Qu, Zou et al. 2005; Zhang, Dong et al. 2005). However, the introduced HBx gene in all studies was under the control of a highly-expressing CMV promoter. The same promoter was used in many studies showing HBx inhibition of NF- κ B-dependent pathways (Bui-Nguyen, Pakala et al. 2010; Kim, Song et al. 2010; Lee, Kang et al. 2010; Zhou, Wang et al. 2010). Also, in a study showing HBx antagonising transactivation of tumour-suppressor p53, the HBx gene was over-expressed under a T7 promoter (Truant, Antunovic et al. 1995). As such, the exact role HBx plays in the initiation and progression of HCC is still poorly understood.

It has also been suggested that HBsAg plays a role in HBV-associated carcinogenesis. Overexpression of HBsAg in transgenic mice induces precancerous liver lesions and HCC in the majority of animals (Chisari, Filippi et al. 1987; Pasquinelli, Bhavani et al. 1992). Furthermore, ER stress-related mRNA transcripts were associated with expression of HBsAg with PreS-mutations in hepatocytes of patients with chronic HBV infections (Wang, Wu et al.

2003; Wang, Huang et al. 2006). ER stress has been shown to upregulate HCC-associated genes, such as Veg-F (Yang, Teng et al. 2009).

On the other hand, ER stress has also been associated with protective and anti-HCC effects (Dara, Ji et al. 2010; Malhi and Kaufman 2011). For example, the anti-HCC drug candidate oridonin induces the expression of ER stress-related proteins in HepG2 cells, suggesting a role of ER-stress in tumour apoptosis (Wang, Ye et al. 2010). Also, ER stress induced by the expression of HCV core protein triggers apoptosis of liver cells in both transfected liver cell lines and liver tissues in transgenic mice (Benali-Furet, Chami et al. 2005). Furthermore, excess linoleic acid has been reported to induce apoptosis in H4IIE hepatoma cells by inducing ER stress (Zhang, Xue et al. 2012).

Studies investigating the link between HBV-coded components and HCC support the role of HBV inducing progression of pre-existing HCC to a more malignant phenotype. However, HBV-associated HCC generally does not develop until decades after the initial infection, thus the role HBV antigens *per se* is not likely to be a strong driver in the initiation of HCC.

1.9.2 The role of HBV DNA integration in HCC progression

As previously mentioned, the dsDNA form of HBV can integrate into the host cell genome (Bill and Summers 2004). HCC progression may be induced by HBV DNA integration by direct insertional mutagenesis of oncogenes or by altering the host chromatin leading to chromosomal rearrangement or the loss or gain of large segments of chromosomes, termed chromosomal instability (CIN).

In woodchucks, almost all woodchucks with chronic WHV infection progress to HCC with 2-4 years (Tennant, Toshkov et al. 2004). Notably, WHV integration into the oncogene N-myc2 is commonly observed in the HCC tissues of these animals (Wei, Fourel et al. 1992; Wei, Ponzetto et al. 1992; Bruni, D'Ugo et al. 1999). In chronic HBV infection, the integration of HBV DNA into cellular oncogenes has been observed. Independent groups have shown that HBV-associated HCC tissue contained HBV DNA integration near key regulatory genes, including hTERT (Paterlini-Brechot, Saigo et al. 2003; Brechot 2004; Jiang, Jhunjunwala et al. 2012). Whole genome sequencing of liver tissue from patients with HBV-associated HCC showed that a significantly greater number of HBV DNA integration events had occurred in tumour tissue compared to non-tumour tissue where integration had occurred into oncogenes such as Mixed-Lineage Leukaemia 4 (Jiang, Jhunjunwala et al. 2012; Sung, Zheng et al.

2012). Changes in levels of mRNA expression of genes located adjacent to the HBV DNA integration site were also observed.

However, in the studies of human HCC, no single gene in particular was disrupted in all tumours by insertional mutagenesis caused by HBV DNA integration. Also, a recent study found no significant differences in DNA rearrangement or insertion into oncogenes caused by HBV DNA integration in HCC tissue compared to neighbouring non-tumour tissue (Jiang, Yang et al. 2012). This is consistent with data showing that HBV DNA integrates via sequence-independent NHEJ (Bill and Summers 2004).

Some studies suggested that continual integration of HBV DNA into the host genome induces CIN, thus initiation and progression of HCC. In many viruses, viral DNA integration into the host cell chromosome induces CIN (Reviewed in (Herath, Leggett et al. 2006; Nishida and Goel 2011)). CIN may be induced by the integration into and resultant dysfunction of structural and matrix attachment regions (S/MAR) of the cellular chromatin (Ohshima, Ohgami et al. 1998; Pett, Alazawi et al. 2004). S/MAR modulate the activity of gene enhancers *in cis* and interact with DNA-binding, transcription-associated proteins, such as DNA topoisomerase II (Boulikas 1993; Bode, Stengert-Iber et al. 1996). They are also associated with origins of replication during semi-conservative DNA replication due to the propensity for weak base-pairing. Motifs, such as low GC% content, topoisomerase II sites and unwinding elements, are conserved in S/MAR (Singh, Kramer et al. 1997; Frisch, Frech et al. 2002). HBV DNA integrations into S/MAR have been commonly observed in HCC associated with chronic HBV and WHV infections (D'Ugo, Bruni et al. 1998; Shera, Shera et al. 2001; Bruni, D'Ugo et al. 2003). The association between HBV DNA integration and CIN has led to the speculation that CIN induces HCC initiation and progression (Levy, Renard et al. 2002; Brechot 2004; Feitelson and Lee 2007).

However, the association between HBV DNA integration and CIN does not imply causation. For example, the link between these two phenomena may share a common cause. Oxidative stress, such as that experienced by chronically inflamed and tumour tissue, has been shown to increase double-strand DNA breakage in host cell (Shimoda, Nagashima et al. 1994; Barash, E et al. 2010; Reuter, Gupta et al. 2010). DNA breakage occurs especially in base unpairing regions in the DNA genome, including S/MAR (Legault, Tremblay et al. 1997; Pineau, Marchio et al. 1998; Liu, Ribocco-Lutkiewicz et al. 2003). Furthermore, DNA breaks have been shown to be the substrate for HBV DNA integration (Bill and Summers 2004). Indeed,

increased HBV integration has been observed in HBV-infected HepG2 cells under oxidative stress induced with hydrogen peroxide (Dandri, Burda et al. 2002). Thus, the association between HBV DNA integration and CIN may be caused by the pre-existing oxidative stress of tumour tissue, leading to both increased CIN and HBV DNA integration into S/MAR.

Indeed, there are multiple conflicting studies both rejecting (Yaginuma, Kobayashi et al. 1985; Takada, Gotoh et al. 1990; Kimbi, Kramvis et al. 2005; Jiang, Yang et al. 2012) and supporting (Rogler, Sherman et al. 1985; Hino, Tabata et al. 1991; Pineau, Marchio et al. 1998; Brechot, Gozuacik et al. 2000) chromosomal deletion and rearrangement during HBV DNA integration. Furthermore, hepatocytes with normal histology have been observed with integrated HBV DNA, suggesting that integration does not necessarily induce changes in phenotype in the host cell (Takada, Gotoh et al. 1990; Mason, Liu et al. 2010). Indeed, Jiang *et al.* showed that the number of tumour suppressor gene mutations was not correlated to the number of HBV DNA integration events in HCC tissue (Jiang, Yang et al. 2012). Also, mathematical models suggest that CIN in itself is not sufficient to cause HCC as the survival advantage a cell may attain via mutagenesis of oncogenes is outweighed by the probability of the loss of a required gene (Komarova 2004). Thus, while it may affect the progression of HCC, mutagenesis through HBV DNA integration remains a controversial explanation for the initiation of the majority of HCCs.

1.9.3 The role of HBV-associated inflammation in HCC progression

Chronic inflammation is known to play a role in cancer progression in various tissues, including the liver (Balkwill, Charles et al. 2005; Villa and Fattovich 2010; Hanahan and Weinberg 2011). Indeed, the constant low level inflammation associated with chronic immune attack against HBV-infected hepatocytes is a known risk factor for HCC development (Lai, Hyatt et al. 2007; Tsang, Trinh et al. 2008; Wang, Lim et al. 2008). There may be multiple mechanisms by which inflammation drives HCC progression: by increasing mutation rate; altering liver microenvironment to a pro-oncogenic state; and inducing the formation of cirrhosis.

As mentioned previously, oxidative stress in cells can induce double-stranded breaks (Shimoda, Nagashima et al. 1994; Barash, E et al. 2010; Reuter, Gupta et al. 2010). Chronically inflamed liver tissues contain high sources of oxidative stress, such as endothelial inducible nitric oxide synthase expression and reactive oxygen species produced by infiltrating immune cells (Jungst, Cheng et al. 2004; Kawanishi and Hiraku 2006; Farinati,

Piciocchi et al. 2010). Methionine adenosyltransferase-knockout mice have increased levels of hepatic oxidative stress and subsequently increased liver malignancies (Martinez-Chantar, Corrales et al. 2002). However, while these mice were more susceptible to the carcinogenic effects of carbon tetrachloride (CCl₄)-induced liver damage, the mice did not spontaneously develop HCC. Therefore, oxidative stress alone is not sufficient to explain HCC progression. Indeed, chronic oxidative stress is found in the livers of patients with Wilson's disease due to copper accumulation, however the risk of HCC is low in these patients (Portmann, Thompson et al. 2007).

The increased turnover of hepatocytes during chronic inflammation could also act to induce double-stranded breaks in cellular DNA and therefore increase rates of mutation. Mdr2-knockout mice lack a hepatic P-glycoprotein, which induces chronic periportal inflammation from 3 months, progressing to HCC by 12-15 months (Mauad, van Nieuwkerk et al. 1994). Increasing turnover rate through partial hepatectomy at 3 months induces HCC progression earlier by >1 month (Barash, E et al. 2010). However, some HCC tumours show intact oncogenes and tumour-suppressor genes, suggesting that DNA mutation is not necessary for transformation (Pasquinelli, Bhavani et al. 1992). As mentioned previously in Section 1.9.2, DNA mutation of oncogenes is not a strong driver of HCC initiation.

The chronic inflammatory microenvironment and associated cell death may induce proliferation of stem cells, which have been hypothesised to be progenitors for HCC (Alison, Islam et al. 2009; Mishra, Banker et al. 2009). As mentioned previously, during high levels of cell death, the liver stem cell population can be induced to proliferate and differentiate. Several researchers have drawn similarities between protein expression profiles of oval cells and cancer cells (Hsia, Evarts et al. 1992; Wu, Lau et al. 1999; Chiba, Kita et al. 2006; Mishra, Banker et al. 2009).

However, the presence of integrated HBV DNA in the majority of HBV-associated HCC tissues indicates that HCCs are derived from HBV-susceptible cells. The susceptibility of oval cells to HBV infection is still controversial. *In vivo* studies have shown that pluripotential murine embryonic stem and bi-potent HBC-3 progenitor liver cells had significantly lower expression of transfected HBV DNA compared to fully differentiated hepatocytes (Ott, Ma et al. 1999). Furthermore, HBV-specific mRNA transcripts have not been detected in non-hepatocyte liver cells, including bile duct cells, in the partial and greater than full-length HBV DNA transgenic mouse models (G7 and G26, respectively) (Aragona, Burk et al. 1996). This

suggests that non-permissive cells types impart negative transcriptional regulation on HBV expression.

Inflammation-associated signals, such as growth factors, chemokines, and cytokines, have been posited to alter the microenvironment of the chronically infected liver to a pro-oncogenic state (Budhu and Wang 2006; Gonda, Tu et al. 2009; Castello, Scala et al. 2010). For example, the proinflammatory cytokine macrophage migration inhibitory factor has been shown both to inhibit the transactivation properties of p53 and is upregulated with drug-induced liver injury (Bourdi, Reilly et al. 2002). Also, TNF- α is highly expressed during inflammation and can induce apoptosis in target cells. Chronic inflammation may select for cells that can escape this process. Indeed, dysregulation of the TNF response pathway is one of the hallmarks of HCC (Chen, He et al. 2003; Yamamoto, Nagano et al. 2004).

A critical factor for the formation of cirrhosis during chronic HBV infection is continual inflammation and hepatocyte death (Giannelli, Quaranta et al. 2003). As mentioned previously in Section 1.3.4, cirrhosis causes vascular rearrangement, leading to disturbed blood flow in liver nodules (Crawford 2007). The resultant stress of impaired nutrition and oxygen distribution to hepatocytes may contribute towards cirrhosis-associated HCC (Kobayashi, Ikeda et al. 2006). A strong association exists between cirrhosis and HCC with 80-90% of HCC patients from different aetiologies having concomitant cirrhosis (Fattovich, Stroffolini et al. 2004).

However, 40-50% of HCC patients with chronic HBV infection do not have concurrent cirrhosis (Reviewed in (Fattovich, Stroffolini et al. 2004)). Furthermore, the risk of HCC is associated with not only cirrhosis severity but also independently associated with patient variables, including age, platelet count, concurrent hepatitis virus infection and prothrombin activity (Velazquez, Rodriguez et al. 2003). Thus, cirrhosis does not necessarily lead to HCC development. This suggests that cirrhosis and fibrosis are not sufficient or necessary paths to HCC development.

1.10 HBV-associated disease progression as a Darwinian process

Not one of the proposed carcinogenic mechanisms described above appear to be sufficient in themselves to initiate and promote HCC. Instead, a complex combination of these mechanisms and their interactions with the host are likely to be driving progression towards HCC. During this progression, the hepatocyte population contains variation, undergoes

selective pressures and passes on changes through heredity, thus necessarily conforming to Darwinian evolution. As such, HCC development should be considered in an evolutionary framework.

Several groups have viewed the initiation, development and progression of cancer in the paradigm of evolutionary theory (Axelrod, Axelrod et al. 2006; Merlo, Pepper et al. 2006; Vincent and Gatenby 2008; Attolini and Michor 2009; Foo, Leder et al. 2011). These evolutionary models draw connections between ecologies containing various species and tissue microenvironments containing various cellular clones. In this paradigm, some cells proliferate due to the selective pressures of the surrounding microenvironment before neoplastic changes are apparent. These cells are altered in a way that increases the probability that cancer will develop.

Several studies support this model. Merlo *et al.* (2010) measured the clonal diversity of oesophageal epithelial cells using multiple independent methods in a cohort of 239 patients with Barrett's oesophagus. The oesophageal epithelial cells of patients with Barrett's oesophagus are chronically destroyed by stomach acid and undergo preneoplastic changes (Odze and Maley 2010). Reductions in clonal diversity as a result of increased clonal proliferation of rare cells significantly predicted the progression to oesophageal adenocarcinoma in these patients (Merlo, Shah et al. 2010).

Furthermore, large clones of histologically normal, but mutation-containing crypt cells were detected in intestinal biopsies of patients with ulcerative colitis, a risk-factor for colorectal cancer (Chen, Rabinovitch et al. 2005). Multiple, neighbouring crypts containing cells with mutations in the p53 gene were visualised by fluorescence *in situ* hybridisation, suggesting clonal proliferation of these cells.

Also, studies suggest that selection for cancer precursor cells occurs in a carcinogen-induced HCC rat model (Solt, Medline et al. 1977; Dragan, Laufer et al. 1993; Dragan, Campbell et al. 1994; Pitot, Dragan et al. 1996; Ou, Conolly et al. 2001). Rats were first treated with a single dose of the carcinogen diethyl-nitrosamine (DEN), then a short term of the general growth inhibitor 2-acetylaminofluorene (2-AAF) and finally a partial hepatectomy to induce proliferation of hepatocytes resistant to both DEN and 2-AAF (Solt, Medline et al. 1977). Rats treated with DEN, 2-AAF and partial hepatectomy all developed highly-proliferative basophilic hepatocyte foci with pre-neoplastic histological changes, occasionally resulting in

HCC. In the absence of any of those three factors, few foci developed. Also, in rats treated with sub-carcinogenic doses of DEN, foci of hepatocytes overexpressing gamma-S-transferase Pi (GSTP) and transforming growth factor-alpha have been observed, eventually accumulating preneoplastic lesions upon dosage with other carcinogens, such as phenobarbital (Dragan, Laufer et al. 1993; Dragan, Campbell et al. 1994; Dragan, Teeguarden et al. 1995; Pitot, Dragan et al. 1996; Ou, Conolly et al. 2001). Foci of hepatocytes overexpressing GSTP had an increased incorporation of BrDU, suggesting that they have increased proliferation and a survival advantage (Dragan, Hully et al. 1994). These studies suggest that preneoplastic hepatocytes with general resistance to chemical inhibitors are selected for prior to carcinogen-induced HCC.

If this model of cancer progression is correct, then the initiation and development of HBV-associated HCC occurs not late in HBV infection, but throughout the course of infection selected by changes in the liver microenvironment. Thus, evidence of clonal proliferation with some sort of survival advantage may be observed in the hepatocyte population.

In HBV infection specifically, London and Blumberg (1983) hypothesised that there are two types of hepatocytes in the liver population: fully differentiated sensitive (S) hepatocytes, which can support virus replication and are subject to immune-mediated cell death during infection; and less differentiated resistant (R) hepatocytes, which can resist infection and therefore avoid immune-mediated cell death (London and Blumberg 1983). Thus, rare R hepatocytes would clonally proliferate due to their survival advantage during immune clearance of infected hepatocytes. Less-differentiated R hepatocytes may be more susceptible to transformation and thus act as a mechanism for HBV-associated HCC. While it was not tested through experiment, this hypothesis is consistent with an interpretation of some mathematical models of HBV infection (Ciupe, Ribeiro et al. 2007).

The theory proposed by London and Blumberg was used to explain a wide range of observations, particularly the evolution of the hepatocyte population over the course of infection. For example, ~100% of hepatocytes are infected with HBV at the start of HBV infection (Berquist, Peterson et al. 1975; Murray, Wieland et al. 2005), but the percentage of hepatocytes supporting virus replication decreases over time to 50-15% (Gowans and Burrell 1985; Gowans, Burrell et al. 1985). HBV antigen- and DNA-negative hepatocytes are observed as foci; perhaps composed of the hypothetical R hepatocytes.

Mason *et al.* (2005) generalised the theory of clonal proliferation of hepatocytes by survival advantage. They proposed that hepatocytes that gain a hyper-responsiveness to growth signals, lose the ability to express virus antigens, fail to display virus antigens on MHC or otherwise resist immune-mediated cell death are likely to clonally proliferate in the environment of a HBV-infected liver (Mason, Jilbert *et al.* 2005). Indeed, clonal proliferation of hepatocytes has been observed with evidence of survival advantages in humans as well as chimpanzee and woodchucks models of HBV infection (Mason, Jilbert *et al.* 2005; Mason, Low *et al.* 2009; Mason, Liu *et al.* 2010).

This places a more active role on clonal proliferation as a route of disease progression as some survival advantages, such as hyper-responsiveness to growth signals or failure to express intrinsic antigens on MHC, may contribute to carcinogenesis. Supporting this theory is the observation of greater numbers of FAH in WHV antigen-negative hepatocytes compared to WHV antigen-positive hepatocytes (Xu, Yamamoto *et al.* 2007). Also, only a minority of HCC tumour tissue supports HBV antigen expression (Xuan, Xin *et al.* 2007). Thus, hepatocytes with loss of HBV virus expression may not only have a survival advantage, but also be more likely to contain preneoplastic changes.

The hypothesised clonal proliferation model is compatible with previously proposed mechanisms for carcinogenesis. Mutations and changes in gene expression associated with HCC progression may also provide survival advantages in the microenvironment of a liver with chronic HBV infection. Chronic inflammation may encourage clonal proliferation by multiple mechanisms: by increasing the amount of liver cell death to provide space for hepatocyte clones to expand; by increasing the mutation rate and thus the probability that a hepatocyte randomly develops a survival advantage; and by inducing oxidative stress, which alters the liver microenvironment and selection pressures on the hepatocytes.

Furthermore, this model may provide an explanation for the foci of HBV-negative hepatocytes observed in chronic HBV infection. Hepatocytes with cellular changes that inhibit HBV expression or replication may evade the immune response and therefore be more likely to replace hepatocytes that undergo cell-mediated cell death.

1.11 Summary and significance

In conclusion, HBV is a prevalent viral disease that is responsible for significant worldwide mortality through HCC and cirrhosis. The molecular and cellular evolution of the hepatocyte population during chronic HBV infection remains unknown.

We propose a model wherein HBV-infected hepatocytes are killed by immune cells and replaced by the mitosis of neighbouring hepatocytes. Some of these hepatocytes may have changes that confer a heritable survival advantage in the environment of the HBV infected liver. Hepatocytes with survival advantage may contain changes (possibly the very changes that have provided the survival advantage) that make it more likely that these hepatocytes become cancerous cells.

Thus, studies into the changing liver population that occurs within the chronically HBV-infected liver may provide insight into the long-term pathogenesis associated with HBV infection.

1.12 Aims and hypotheses

Therefore, the aims of the current project are:

1. To detect hepatocyte lineages using unique virus-cell DNA junctions produced by integrated HBV DNA,
2. To determine whether clonal proliferation of hepatocytes with survival advantage has occurred and,
3. To characterise the phenotypic changes that have occurred in the hepatocytes that have undergone clonal proliferation.

We hypothesise that clonal proliferation of hepatocytes with survival advantage occurs in the livers of chronic HBV patients and that one of these survival advantages is the loss of HBV antigen expression. Thus, hepatocytes in HBV-negative foci should have clonally proliferated to a greater extent than those in HBV-positive foci. Furthermore, we hypothesise that the changes in gene expression and DNA methylation in HBV-negative foci of hepatocytes include those associated with HCC initiation and progression.

2 – Materials and Methods

2.1 Collection of liver tissues

The collection and use of human tissues was conducted with ethics approval from the respective HRECs of Royal Adelaide Hospital (Protocol number 070301), and Southern Adelaide Clinical Research (Project ID #171.11). This study was classified as an exempt dealing by the Office of the Gene Technology Regulator and the Institutional Biosafety Committee (Identifier 20061205_995).

Liver samples were collected from patients clinically diagnosed with chronic HBV infection and were collected from a variety of sources. Liver tissue was collected from: nine patients from the Centenary Institute, Sydney, during resection prior to liver transplant; five patients from the University of Washington, Philadelphia, during resection of HCC tissue; three patients from The Royal Adelaide Hospital, Adelaide, during resection of HCC tissue or by needle biopsy. Also, five HBV-negative NHL samples were collected from patients during resection metastatic colon cancer at The Royal Adelaide Hospital, Adelaide. After collection, all tissue samples were immediately snap-frozen in liquid nitrogen and stored at -80 °C.

2.2 Tissue fixation and paraffin wax-embedding

Snap-frozen liver tissue was cut into pieces of ~1 mm in thickness, placed into plastic cassettes and (TechnoPlas) fixed in either 4% v/v formalin, 70% v/v ethanol (EtOH), or 3:1 v/v Ethanol (Merck):Acetic Acid (Merck) (EAA). Fixed liver tissues were then embedded in paraffin wax blocks.

Formalin fixation was carried out by placing cassettes in 4% v/v Analytical Reagent (AR) grade formalin (Merck) and storing at RT O/N. EtOH fixation was carried out by placing cassettes in ice-cold AR grade 70% EtOH (Merck) at 4°C O/N in 70% EtOH, then transferred to ice-cold absolute EtOH. EAA fixation was carried out by placing cassettes in ice-cold EAA for 30 min, and then transferred to ice-cold absolute EtOH.

Fixed tissues were then submitted to the Hanson Institute Centre for Neurological Diseases, SA Pathology (South Australia) for paraffin wax-embedding. Paraffin wax-

embedding was carried out in a Sakura FineTek VIP tissue processor (Sakura FineTek Inc.). The embedding protocol was as follows: 2 cycles with xylene at 40°C for 45min; 4 cycles with paraffin at 58 °C for 30 min.

2.3 Liver tissue histological staining and analysis

Formalin-fixed paraffin wax-embedded liver tissue was sectioned at 5 µm intervals and stained as outlined below by the Centre of Neuropathology, SA pathology. After sectioning, tissues were mounted onto glass slides. Tissues were then dewaxed using two 5 min washes in AR grade xylene, cleared with 2x1 min in 100% EtOH and rehydrated for 2 min in reverse osmosis water. The rehydrated tissue was stained using Haematoxylin and Eosin, Periodic acid Schiff-Diastase, Gordon and Sweet's reticulin stain and Haematoxylin Van Geison's stain as described in Section 2.3.1-2.3.6. After staining, tissues were dehydrated with two 1 min washes in absolute EtOH and cleared with two 5 min washes in xylene. Coverslips were mounted with Entellan[®] mounting medium (Merck). Subsequently, tissues were visualised by light microscopy.

Histological analysis of all liver tissues was performed by Associate Professor Andrew Clouston (University of Queensland & Envoi). METAVIR and Ishak scoring, both standard measures of histopathological progression during viral liver infection (Ishak, Baptista et al. 1995; Bedossa and Poynard 1996), were also measured.

2.3.1 Haematoxylin and Eosin staining

Haematoxylin and Eosin stain was used to elucidate general tissue morphology and to examine the extent of inflammation in tissues. The staining protocol is adapted from Gamble (Gamble 2008). Nuclei were stained for 4 min using Lillie Mayer alum haematoxylin (50 mg/mL (NH₄)Al(SO₄)₂, 5 mg/mL Haematoxylin, 1 mg/mL sodium iodate in 2% v/v acetic acid, 30% v/v glycerol, 1% v/v EtOH). Excess dye was washed off under running water for 5 min. Nuclei were differentiated by 30 sec treatment with 0.3% acid alcohol (0.3% v/v HCl in 70% v/v EtOH). Excess acid alcohol was washed off under running water for 5 min. Nuclei were blued with 2.5% v/v saturated lithium carbonate in water. Excess lithium carbonate solution was

washed off under running water for 5 min. Tissues were counter-stained for 2 min with alcohol acetified eosin phloxine solution (0.1% Eosin Y, 0.01% phloxine, 0.4% v/v acetic acid in 80% v/v EtOH).

2.3.2 Periodic acid Schiff-Diastase staining

Periodic acid Schiff-Diastase stain was used to localise activated macrophages in liver tissues. The staining protocol is adapted from Meyers *et al.* (Meyers, Fredenburgh *et al.* 2008). Tissues were treated with 2 mg/mL α -amylase type VI-B (Sigma) for 20 min. Excess amylase was washed off under running tap water for 5 min. Oxidisation of 1,2-glycol groups was performed by treating tissues with 10 mg/mL periodic acid for 5 min. Excess acid was washed off in distilled water (DW) for 5 min. Oxidised 1,2-glycol groups were stained magenta with Schiff's reagent (Australian Biostain). Excess Schiff's reagent was washed off under fast-running water for 5 min. Nuclei were counterstained by staining with Lillie Mayer alum haematoxylin for 1 min. Extra differentiation of nuclei was attained by 30 sec treatment of tissues with 0.3% acid alcohol. Acid alcohol was washed off under running water for 5 min. Nuclei were counter-stained with 2.5% v/v saturated lithium carbonate in water. Excess lithium carbonate solution was washed off under running water for 5 min.

2.3.3 Gordon and Sweet's Reticulin staining

Gordon and Sweet's Reticulin stain was used to detect the basement membrane of the liver sinusoids, allowing assessment of liver architecture. The staining protocol is adapted from Jones *et al.* (Jones, Bancroft *et al.* 2008). Reticulin fibres were oxidised by treating sections with acidified potassium permanganate (0.15% v/v sulphuric acid in 0.5% potassium permanganate) for 3 min. The potassium permanganate was washed off with DW and decolourised by 1 min wash with 2% w/v oxalic acid. Excess oxalic acid was washed off with DW. Further differentiation was achieved by a 10 min wash with 4% w/v iron alum. Excess iron alum was washed off with DW. Silver molecules were impregnated into tissues by a 15 sec wash with 10% aqueous silver nitrate solution. Excess silver nitrate acid was quickly washed off with DW. Impregnated silver was immediately reduced to metallic silver by a 2 min wash with 10% aqueous formalin. Excess formalin was washed off under running water for 2 min. Tissues

were fixed with a 2 min wash in 2% sodium thiosulphate. Excess sodium thiosulphate was washed off under running water for 2 min.

2.3.4 Haematoxylin Van Gieson staining

Van Gieson stain was used to stain collagen within the liver and determine the extent of fibrosis or cirrhosis. The staining protocol is adapted from Jones, *et al.* (Jones, Bancroft et al. 2008). Nuclei were counter-stained by staining with Lillie Mayer alum haematoxylin for 1 min then washed with 2.5% v/v lithium carbonate. Excess lithium carbonate solution was washed off under running water for 5 min. Tissues were treated with Van Gieson's solution (0.1% acid fuchsin in saturated picric acid) for 2 min. Excess Van Gieson's solution was washed off under fast-running water for 5 min.

2.3.5 Perls' Prussian Blue iron staining

Perls' Prussian Blue iron stain was used to detect haemosiderin deposits in the liver. The staining protocol is adapted from Churukian (Churukian 2008). To form insoluble ferric ion compounds, slides were washed with 2% potassium ferrocyanide in 2% HCl for 10 min. Excess potassium ferrocyanide was washed off by three 2 min washes with DW. Tissues were counter-stained with Neutral Red dye (10 mg/mL Neutral Red in 1% v/v acetic acid) for 2 min. Excess Neutral Red dye was washed off with three dip rinses in DW.

2.3.6 Fouchet's Bile pigment staining

Fouchet's bile pigment stain was used to detect bile deposits as a result of cholestasis or other liver injuries. The staining protocol is adapted from Churukian (Churukian 2008). Slides were washed for 10 min in an equal parts mixture of potassium ferrocyanide and HCl. Excess stain was washed with 3x5 min washes in DW. Tissues were counter-stained with Neutral Red dye for 2 min. Excess Neutral Red dye was washed off with three dip rinses in DW.

2.4 Detection of HBcAg and HBsAg in liver sections by immunohistochemistry

EAA- or EtOH-fixed paraffin wax-embedded liver tissue was sectioned at 5 μ m intervals. Sectioned tissue was dewaxed by 2x5 min washes in AR grade xylene, then rehydrated through 2x5 min washes in AR grade absolute EtOH, then 2x5 min washes in 1x PBS. Sections were ringed using a liquid-blocking SuperPAP pen (ProSciTech). To inactivate endogenous peroxidases present in the liver sections, ~150 μ L peroxidase suppressor solution (Thermo Scientific, Cat # PI-35000) was added to the sections. Slides were then incubated in a humid box for 30 min at RT. The solution was washed off with 2x3 min in 1x Tris buffered saline (TBS) with 0.5% Tween. Endogenous biotin and non-specific protein binding was blocked using ~150 μ L of 4 drops/mL Avidin (Vector Labs, Cat # SP-2001) diluted in Universal Blocking Solution (Thermo Scientific, Cat # 37535). Slides were incubated in a humid box at RT for 30 min. The blocking solution was washed off with 2x3 min washes in 1x TBS 0.5% Tween. To block non-specific antibody binding, tissues were incubated in a humid box for 15 min at RT with 5% normal swine serum (NSwS) or rabbit serum (NRS), both obtained from SA Pathology (formerly IMVS).

HBcAg or HBsAg were detected using a 1:500 dilution of polyclonal rabbit anti-HBcAg (Dako, Cat # B0586) or 1:50 monoclonal mouse anti-HBsAg (Dako, Cat # M3506) antibodies, respectively. Antibodies were diluted in 4 drops/mL Biotin (Vector Labs, Cat # SP-2001) in Universal Blocking Solution. Slides were incubated with antibody solution in a humid box for 1 hr at RT, then washed with 2x3 min washes in 1x TBS 0.5% Tween and 1x10 min wash in 1x PBS. Either 0.45% biotinylated swine anti-rabbit (Dako, Cat # E0431) and 0.45% NSwS diluted in 1x PBS (for HBcAg detection) or 0.45% biotinylated rabbit anti-mouse (Dako, Cat # E0413) and 0.45% NRS diluted in 1x PBS (for HBsAg detection) were added as secondary antibodies and subsequently incubated in a humid box at RT for 30 min. Excess antibodies were washed off with 1x PBS for 10 min. Secondary antibodies were detected using an avidin-biotin complex (ABC) detection kit (Thermo Scientific, Cat # 32020), as per the manufacturer's instructions. ABC solution was prepared, incubated for 30 min at RT, then added to tissue sections in a humid box for 30 min at RT. Excess antibodies were washed off with 1x PBS for 10 min. Vector NovaRed chromagenic reagent (Vector Labs, Cat # SK-4800) was prepared as per

manufacturer's instructions and 500 μ L of chromagenic solution was added to each slide. After incubation for 7 min at RT, slides were washed in deionised water for 5 min, counter-stained with Mayer's haematoxylin for 2 min, dip washed twice in 1x PBS and washed in deionised water for 5 min. Tissues were then dehydrated with 2x5 min washes in AR grade absolute EtOH and cleared with 2x5 min washes in AR grade xylene. Coverslips were mounted with Entellan[®] mounting medium. Subsequently, slides were examined by light microscopy.

2.5 Detection of HBsAg and cytokeratin 19 in liver sections by dual immunofluorescence

Formalin-fixed paraffin wax-embedded liver tissue was sectioned at 5 μ m intervals. The sectioned tissue was dewaxed by 2x5 min washes in AR grade xylene, cleared with 2x5 min washes in absolute EtOH, then rehydrated with 2x5 min washes in 1x PBS. Antigens were retrieved by incubation with 0.1% trypsin or 0.1% pepsin (Roche) in 0.1 M HCl at 37°C for 30 min. The retrieval solution was washed off with 2x3 min washes in 1xTBS 0.5% Tween. Non-specific protein binding was blocked using a 1:50 dilution of normal goat serum in ~150 μ L of Universal Blocking Solution (Thermo Scientific). Slides were incubated in a humid box at 37°C for 30 min. The blocking solution was washed off with 2x3 min washes in 1x TBS 0.5% Tween.

HBsAg was detected using a 1:50 dilution of monoclonal mouse anti-HBsAg (Dako, Cat # M3506) antibodies, diluted in Universal Blocking Solution. Slides were incubated with antibody solution in a humid box for 1 hr at 37°C, then washed with 2x3 min washes in 1xTBS 0.5% Tween and 1x10 min wash of 1xPBS. Cytokeratin-19 was detected in the same section using 1:25 affinity-purified polyclonal rabbit anti-cytokeratin-19 (Novus Biologicals, Cat # NB100-687), diluted in 1x PBS. Slides were incubated with antibody solution in a humid box for 1 hr at 37°C, then transferred to 4°C O/N. Excess antibodies were washed off with 2 washes of 1xPBS for 10 min. Secondary antibodies were then added as a combination of Alexa594-conjugated goat anti-mouse IgG1 (Invitrogen, Cat # A21125) and Alexa488-conjugated goat anti-rabbit (Invitrogen, Cat # A11034) at a dilution of 1:200 in 1xPBS, then incubated in a humid box at 37°C for 30 min. Excess antibodies were washed off with three washes of 1x PBS for 5 min. Slides were counter-stained with DAPI (Invitrogen, Cat #

D1306) for 20 min at RT in the dark. Excess DAPI was removed with 3x5 min washes in 1xPBS. Slides were mounted in 90% glycerol in 10 mM Tris in PBS, sealed with nail polish and visualised via fluorescence microscopy. A Nikon Eclipse Ti microscope was used in conjunction with NIS elements software (Nikon). The following excitation and emission filters were used to visualise each fluorophore:

Alexa594 - excitation = 532-587 nm, emission = 608-683 nm;

Alexa488 - excitation = 470-490 nm, emission = 520-560 nm;

DAPI - excitation = 330-380 nm, emission = 435-∞ nm.

2.6 Detection of HBV DNA in liver sections using *in situ* hybridisation

The protocol for the detection of HBV DNA in liver sections using *in situ* hybridisation is based on *in situ* hybridisation protocols outlined by Jilbert (Jilbert 2000) and those previously used by Mason *et al.* (Mason, Jilbert et al. 2005; Mason, Liu et al. 2010).

2.6.1 Probe production

Digoxigenin (DIG)-labelled DNA probes were produced using DIG-nick translation Mix (Roche, Cat # 11745816910). For each DIG-nick translation reaction, 1 µg of either pHBV1.3BB4.5 (Genbank accession #AF305422), a plasmid containing a 1.3 times genome-length HBV sequence in the plasmid vector pBlueBac4.5, (Figure 2.1), or the plasmid vector pBlueBac4.5 (Figure 2.2, Invitrogen, Cat # V1995-20) were used as templates. The DIG-labelled pHBV1.3BB4.5 and pBlueBac4.5 probes were produced as per the manufacturer's instructions. After EtOH precipitation, the probes were redissolved in DW at a concentration of 100 ng/µL.

2.6.2 Pre-hybridisation and hybridisation solutions

Pre-hybridisation solution was assembled in two parts: Mix A and Mix B. To make 200 µL of pre-hybridisation solution, Mix A was a solution of 100 µL deionised formamide, 5 µL 20 mg/mL salmon sperm DNA (Sigma), 10 µL 10 mg/mL tRNA (Roche) and 10 µL deionised water. Mix A was incubated for 5 min to 95°C, then cooled on ice immediately. Mix B was a solution of 40 µL 5x hybridisation buffer

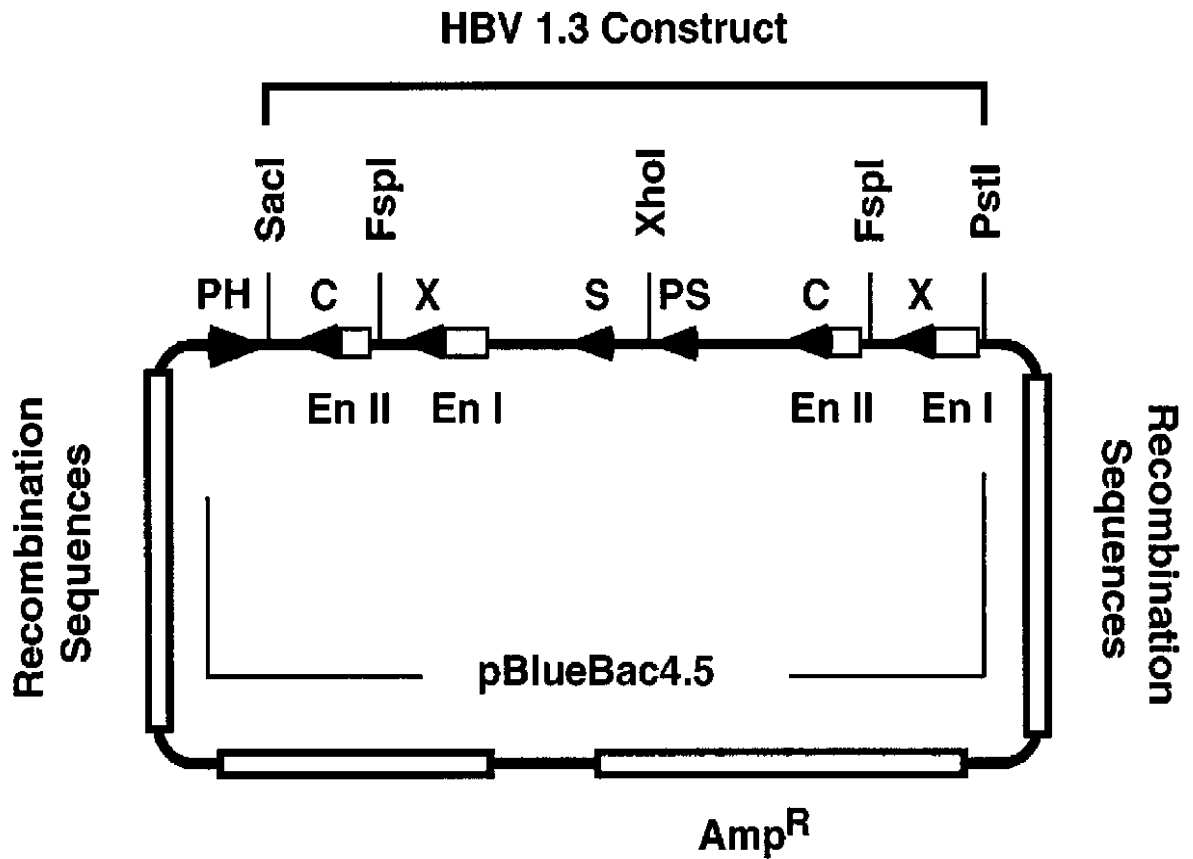
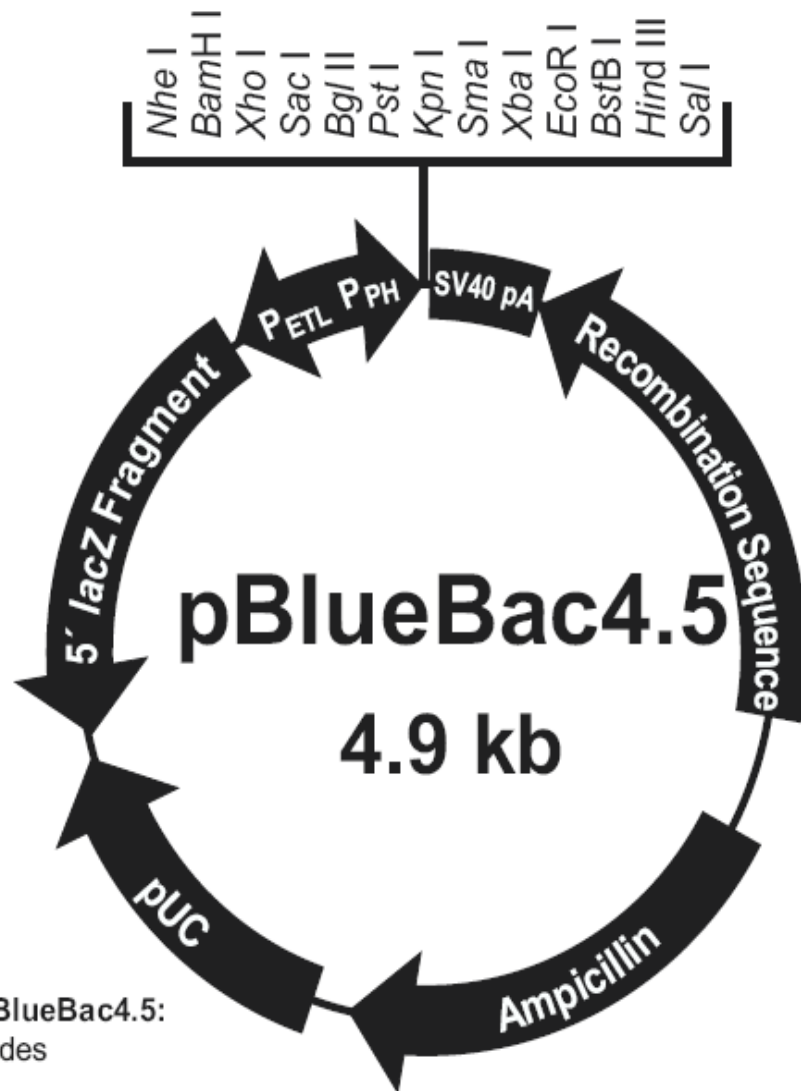


Figure 2.1. Plasmid map of pHBV1.3BB4.5. A 1.3 times genome-length sequence of HBV DNA (Genotype D, Genbank accession #AF305422) was subcloned into the baculovirus transfer vector pBlueBac4.5 (Invitrogen, Figure 2.2) by Delaney and Isom (1998). The 1.3-HBV insert is antisense with respect to the vector-associated baculovirus polyhedrin promoter (PH). The HBV-derived enhancers (En I and En II) and promoters (◄) for HBV core (C), X, pre-S (PS), and surface (S) open reading frames are shown. The direction of arrowheads show the direction of transcription. The Amp^R region encodes β -lactamase. Figure reproduced with permission from Delaney and Isom (1998).



Comments for pBlueBac4.5:
4940 nucleotides

Polyhedrin promoter (P_{PH}): bases 7-95
 Multiple cloning site: bases 132-198
 SV40 polyadenylation sequence: 199-326
 Recombination sequences (ORF1629): bases 1343-543
 Ampicillin resistance gene: bases 1730-2587
 pUC origin: bases 2735-3408
 5' *lacZ* fragment: bases 3526-4638 (C)
lacZ sequence homologous to *lacZ* sequence
 in Bac-N-Blue™ DNA: bases 3526-4414 (C)
 Early-to-late promoter (P_{ETL}): bases 4639-4940 (C)
 (C) = complementary strand

Figure 2.2. Plasmid map of pBlueBac4.5. The baculovirus transfer vector pBlueBac4.5 (Invitrogen, Cat # v1995-20) was used as a vector for a 1.3 times genome-length HBV sequence (Figure 2.1). The Ampicillin region encodes β -lactamase. Figure reproduced with permission from pBlueBac4.5 user manual (Invitrogen, version D 2003).

(0.1% polyvinylpyrrolidone (Sigma), 0.1% Ficoll (Amersham), 50 mM EDTA (Sigma), 50 mM sodium pyrophosphate (Sigma), and 500 mM Tris pH 8.0 (Sigma) in 10x saline sodium citrate buffer (SSC)), 4 μ L 1 M dithiothreitol (DTT) (Sigma), 20 μ L nuclease-free BSA (NEB, Cat # B9001S) and 11 μ L deionised water. Mix B was kept on ice, then added to Mix A once it had cooled on ice.

Hybridisation solution was made as above, but substituting deionised water in Mix A with 15 μ L of DIG-labelled probe (either pHBV1.3BB4.5 or pBlueBac4.5) to a final concentration of 7.5 ng/ μ L.

2.6.3 *In situ* hybridisation

Sections of EAA or EtOH-fixed liver tissue were dewaxed by 2x5 min washes in AR grade xylene, and subsequently rehydrated by a series of EtOH: 2x2 min in absolute EtOH, 1x2 min in 90% EtOH and 1x2 min in 70% EtOH. Tissues were fully rehydrated with a 1x5 min wash in 1xPBS. Tissues were fixed with 0.1% Electron Microscopy Grade glutaraldehyde (Unilab, Cat # 698500ML) diluted in 1xPBS for 30 min at 4°C, followed by 2x5 min washes in 1xPBS. To increase permeability of the cells, tissues were digested with 3 μ g/mL proteinase K (PK) (Roche, Cat # 03115879001) in 1x PK buffer for 1x15 min wash at 37°C, followed by 2x5 min washes in 1xPBS. PK was denatured by incubating the slides in 0.1xSSC for 5 min at 95°C. Tissues were fixed again with 0.1% glutaraldehyde in 1xPBS for 1x15 min at 4°C, followed by 2x5 min washes in 1xPBS at RT. To decrease background signal, amine groups were acetylated with 0.25% acetic anhydride (Sigma) in 0.1 M triethanolamine (Sigma), balanced to pH 8 for 10 min at RT, followed by 2x5 min washes in 2xSSC. Slides were dehydrated through a series of EtOH washes: 1x2 min in 70% EtOH, 1x2 min in 90% EtOH and 2x2 min in absolute EtOH. Slides were air-dried for >2 hr at 37°C.

Tissue sections were framed with Frame-Seal Incubation Chambers (Biorad, Cat # SLF-0201) and 75 μ L of pre-hybridisation mix was added to each section. Slides were incubated for 1 hr at 37°C in a humid chamber. Pre-hybridisation mix was removed by aspiration and replaced with 24 μ L of hybridisation mix. The incubation chambers

were then sealed, heated for 5 min at 75°C, and then incubated O/N at 34°C, ~30°C below the melting temperature (T_m) of the probes (Appendix 9.2).

Following hybridisation, the incubation chambers were removed with forceps and the slides were dip washed in 2xSSC. Non-specific binding of DIG-labelled probe was removed in a series of washes: 2x 30 min with 2xSSC at 58°C, then 2x30 min with 0.1xSSC at 58°C, ~14°C below the T_m of the probes (Appendix 9.2). Tissues were then blocked with 0.5% w/v Blocking Reagent (Roche) in DIG buffer 1 (100 mM Tris, 150 mM NaCl, pH 7.5) for 60 min at 37°C. DIG-labelled probes were detected by applying a 1:500 dilution of anti-DIG alkaline phosphatase-conjugated antibody in 0.5% Blocking Reagent in DIG buffer 1 and incubating for 60 min at 37°C in a humid chamber. Excess antibodies were washed off with 2x5 min washes in 1xPBS at RT. Bound antibodies were detected by incubating with freshly-made developing solution (337.5 µg/mL Nitro blue tetrazolium chloride (NBT) (Roche, Cat # 11383213001), 175 µg/mL 5-bromo-4-chloro-3-indolyl-phosphate, 4-toluidine (BCIP) (Roche, Cat # 10760994001) diluted in 100 mM Tris pH9.5, 100 mM NaCl, 50 mM MgCl₂) O/N at RT in the dark.

Excess developer was removed by dip washing slides in 1xPBS followed by 2x5 min washes in 1xPBS. Slides were counter-stained with Mayer's haematoxylin. Excess haematoxylin was removed with 3x5 min washes in 1xPBS. Slides were mounted in 90% glycerol in 10 mM Tris in PBS, sealed with nail polish and visualised by light microscopy.

2.7 Extraction of total DNA from snap-frozen liver tissues

For snap-frozen tissue, ~5 mg of liver tissue was cut on a sterile plastic Petri dish with a sterile scalpel blade and placed immediately in 400 µL digestion solution (100 mM NaCl, 0.5% SDS, 50 mM Tris pH 7.5, 10 mM EDTA, 2 mg/mL PK (Roche)).

Total DNA was then extracted by PK digestion of the liver tissue by incubating in a shaking thermomixer (Eppendorf) at 55°C for 2 hr. Total DNA was further extracted with 400 µL of a mixture of 25:24:1 UltraPure phenol:chloroform:isoamyl alcohol (Life Technologies, Cat # 15593-031). Finally, the DNA was precipitated by adding

35 μ L 3M sodium acetate pH 4.6, followed by 800 μ L AR Grade EtOH and incubating at either at -20°C O/N or -80°C for 4 hr. The precipitated DNA was pelleted by centrifugation at 14000 g, the pellet was launched, washed with 1 mL 70% EtOH, then vacuum-dried for 20 min and redissolved in 50 μ L of Elution Buffer (Qiagen).

2.8 *Extraction of total DNA from paraffin wax-embedded liver tissues*

For paraffin wax-embedded liver tissue, two 5 μ m sections were cut and placed into a 1.5 mL Eppendorf tube. To remove the paraffin wax, the sections of liver tissue were 2x5 min in 1 mL AR grade xylene, 2x5 min in 1 mL AR grade absolute EtOH, and then vacuum-dried. After addition of 400 μ L digestion solution, extraction proceeded as described in Section 2.7.

2.9 *Extraction of nuclear DNA from liver samples*

Nuclear DNA was isolated from snap-frozen liver tissue as previously described by Zhang *et al.* (2003). Approximately 5mg of liver tissue was homogenised in 100 μ L of homogenisation solution (10 mM Tris, 3 mM MgCl₂, 250 mM sucrose, and 0.05% Triton X-100) using an 1.5 mL Eppendorf tube pestle. This homogenisation step was designed to disrupt the plasma membrane of cells while leaving the cell nuclei intact. If nuclei were being visualised by fluorescence microscopy, 2 μ L of 100x DAPI (Invitrogen, Cat # D1306) was added. Nuclei were pelleted by centrifugation at 2000 rpm for 5 min. Pellets were washed again with 1 mL of homogenisation solution and then resuspended in 200 μ L of homogenisation solution. As a confirmation of the isolation of nuclei by fluorescence microscopy, a wet mount slide was prepared with 100 μ L of resuspended nuclei and DAPI-stained nuclei were visualised by confocal fluorescence microscopy. To extract DNA from the isolated nuclei, the resuspended pellet was digested with PK (Roche) by addition of 500 μ L 1 mg/mL PK in 10 mM Tris and 0.1% Triton X-100. The digestion solution was incubated at 50°C for 2 hr, then heat inactivated at 75°C for 15 min. The digestion solution was extracted with 660 μ L of a 25:24:1 mixture of UltraPure phenol:chloroform:isoamyl alcohol (Life Technologies). Precipitation and dissolving of the DNA was carried out as described in Section 2.7.

2.10 Extraction of total DNA from serum samples

Total DNA was extracted from serum using High Pure Viral Nucleic Acid Kit (Roche, Cat # 11858874001), as per the manufacturer's instructions. Extracted DNA was eluted into 40 μ L of elution buffer.

2.11 Spectrometry of DNA extracts

The purity and concentration of DNA extracts were analysed by spectrophotometry. Using a NanoDrop ND-1000 spectrophotometer (ThermoScientific), 1 μ L of DNA extract was placed on the glass stage and absorbance readings were recorded.

To quantify DNA from light absorbance readings, a modified Beer-Lambert equation is used by the Nanodrop software:

$$c = (A * e)/b$$

Where **c** is the nucleic acid concentration in ng/ μ L, **A** is the absorbance in AU, **e** is the wavelength-dependent extinction coefficient in ng-cm/ μ L, and **b** is the path length in cm. For double-stranded DNA, $e = 50$ ng-cm/ μ L.

2.12 Amplification and sequencing of HBV DNA

2.12.1 PCR Primer design

PCR primers capable of detecting all HBV genotypes were designed to amplify HBV sequences using conserved regions of aligned sequences from 18 selected HBV strains (Table 2.1) from previous studies (Bichko, Pushko et al. 1985; Takahashi, Akahane et al. 1998; Fujiwara, Tanaka et al. 2005; Sakamoto, Tanaka et al. 2006; Michitaka, Horiike et al. 2007; Suzuki, Akuta et al. 2007). Primers were designed to amplify two ~1 kb fragments, one upstream and one downstream from the expected integration junction at nt 1824 of the HBV DNA sequence with Genbank accession #AB241115. Primer sequences are summarised in Table 2.2.

Table 2.1. Selected HBV strains to used design primers in conserved DNA regions

Genbank Accession #	HBV genotype	Genome length (bp)	Reference
AB241115	A	3222	(Sakamoto <i>et al.</i> 2006)
AM282986	A	3222	(Direct submission by Panjaworayan <i>et al.</i> 2006)
AB241116	B	3179	(Sakamoto <i>et al.</i> 2006)
AB241117	B	3215	(Sakamoto <i>et al.</i> 2006)
AB014365	C	3215	(Takahashi <i>et al.</i> 1998)
AB014388	C	3161	(Takahashi <i>et al.</i> 1998)
AB014397	C	3185	(Takahashi <i>et al.</i> 1998)
AB014399	C	3215	(Takahashi <i>et al.</i> 1998)
AB014398	C	3167	(Takahashi <i>et al.</i> 1998)
AB288026	C	3215	(Direct submission by Yamazaki <i>et al.</i> 2006)
X02496	D	3182	(Bichko <i>et al.</i> 1985)
AB210822	D	3182	(Michitaka <i>et al.</i> 2007)
AB210820	D	3182	(Michitaka <i>et al.</i> 2007)
AB219529	E	3200	(Fujiwara <i>et al.</i> 2005)
AB219533	E	3218	(Fujiwara <i>et al.</i> 2005)
AB219534	E	3179	(Fujiwara <i>et al.</i> 2005)
AB275308	H	3215	(Direct submission by Chihara <i>et al.</i> 2006)
AB298362	H	3215	(Suzuki <i>et al.</i> 2007)

Table 2.2. Primers used to sequence regions of HBV DNA 1 kb upstream and downstream of the expected right-hand junction at nt 1832 (as mentioned in Section 1.6, Staprans *et al.* (1991) and Yang *et al.* (1999)

Primer name	Sequence (5' -> 3') ¹	Target on HBV sequence (nt) ²	T _m (°C) ³	Length (nt)
HBV Seq 1672F	cataagaggactccttgact	1672	51	20
HBV Seq 1796R	caatttatgcctacagcctcct	1796	55	22
HBV Seq 1813F	atgcaactttttcacctctgc	1813	55	21
HBV Seq 2026R	taaggcttctcgatacagagc	2026	54	21

¹ Primers sequence design based on consensus sequences of strains outline in Table 2.1.

² Based on HBV DNA sequence from Genbank Accession #AB241115.

³ Melting temperature (T_m) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.

Table 2.3. Primers used to quantify total HBV DNA in liver DNA extracts by qPCR.

Primer name	Sequence (5' -> 3')	T _m (°C) ³	Length (nt)
HBV F ¹	GTGTCTGCGGCGTTTTATCA	52	20
HBV R ¹	GACAAACGGGCAACATACCTT	52	21
Beta-globin F ²	GAAGAGCCAAGGACAGGTAC	54	20
Beta-globin R ²	CAACTTCATCCACGTTACCC	52	20

¹ Primers were previously designed by Mendy *et al.* (2006).

² Primers were previously designed by Vandegraaff *et al.* (2001).

³ Melting temperature (T_m) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.

2.12.2 PCR conditions

Approximately 100 ng of extracted total liver DNA or serum DNA was combined with 1x Amplitaq Gold PCR mix (Applied Biosystems, Cat # 4331816) that contained 2.5 U of hot-start TAQ polymerase, 2.5 mM MgCl₂ and 0.5 mM concentration of each PCR primer in a 50 µL reaction volume. PCR conditions were as follows: initial 10 min denaturation step at 95°C to activate hot-start polymerase; 35 cycles of 15 sec denaturation step at 95°C, 15 sec annealing step at 54°C and 3 min elongation step at 72°C; then a holding step at 15°C. PCR products were separated by gel electrophoresis, purified using QIAEX II Gel Extraction kit (Qiagen), as per the manufacturer's instructions, and then sequenced by Sanger sequencing, as described in Section 2.12.3.

2.12.3 DNA sequencing

DNA was sequenced using a Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Cat #4337457). Approximately 100 ng of purified DNA was added to a 10 µL sequencing reaction (1 µL BigDye Terminator mix, 1x sequencing buffer, 160 nM primer). Sequencing reaction conditions were as follows: initial 2 min denaturation step at 95°C; 25 cycles of 10 sec denaturation step at 95°C, 10 sec annealing step at 55°C and 4 min elongation step at 60°C; then a holding step at 15°C.

Unincorporated dye terminator nucleotides were removed using a CleanSeq magnetic bead kit (Agencourt, Cat # A29151). Briefly, 42 µL 85% EtOH and 10 µL of magnetic beads were added to each PCR tube used or each well of a 96-well plate and vortexed. The samples were then incubated on a 96-ring SPRIplate magnetic base (Agencourt, Cat # A29164) for 10 min. The magnetic beads were washed twice in 100 µL of 85% EtOH. The magnetic beads were then vacuum-dried for 30 min and sequencing products were redissolved in 80 µL of 0.1 mM EDTA. To read sequencing reaction products, 40 µL of redissolved product was submitted to the Australian Genomic Research Facility (AGRF) sequencing facility, Adelaide for capillary separation. Chromas Lite software (Technelysium) was used for analysis of generated chromatographs.

2.13 Enumeration of HBV DNA copy number by qPCR

2.13.1 Primer design

HBV DNA copy number in the total liver DNA extracts was quantified by qPCR. The sequences of the primers used are summarised in Table 2.3. HBV DNA specific primers that were used were previously designed by Mendy *et al.* (Mendy, Welzel *et al.* 2010). Primers specific for the single copy gene beta globin was used for standardisation as previously designed by Vandegraaff *et al.* (Vandegraaff, Kumar *et al.* 2001).

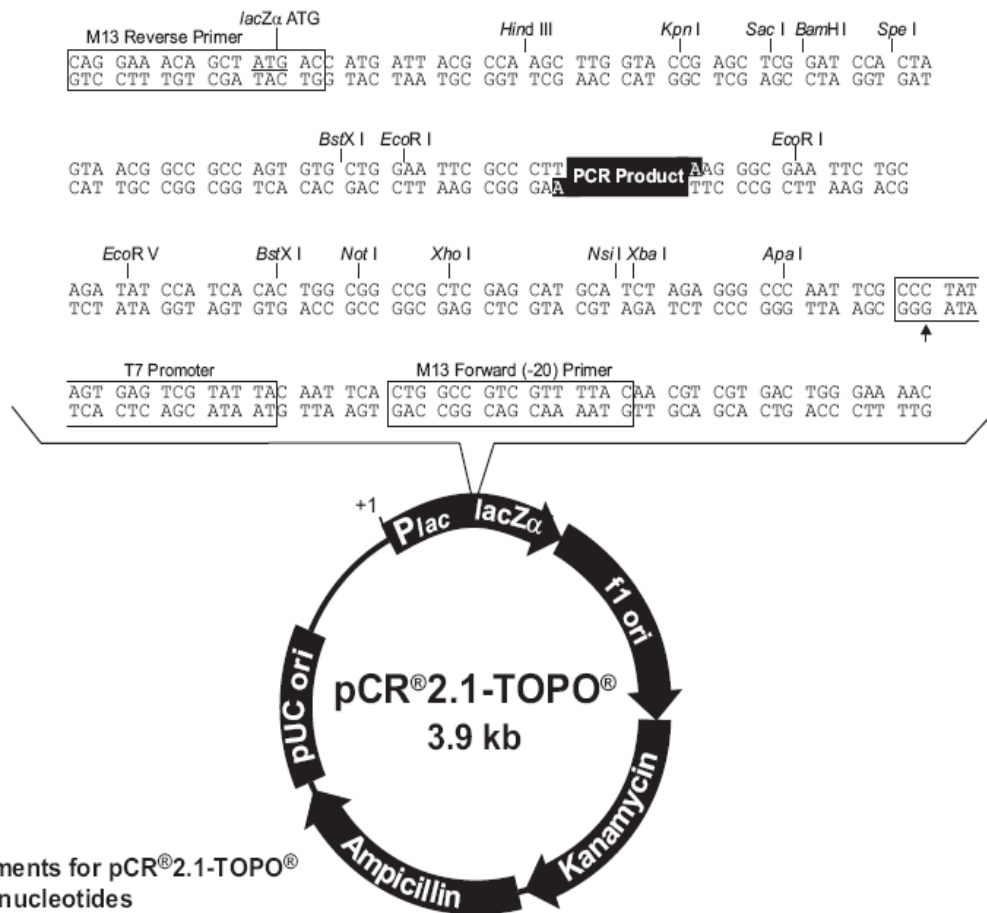
2.13.2 qPCR conditions

Five microlitres of DNA extract were added to 15 μ L of SybrGreen master mix (Applied Biosystems) containing 1x SybrGreen and 0.15 nM of each primer. qPCR was carried out using the following conditions: initial 10 min denaturation step at 95°C to activate the polymerase; 41 cycles of 15 sec denaturation step at 95°C, 1 min annealing and elongation step at 60°C with data collection at each cycle; then a melt curve step with a 0.3°C/sec ramping speed from 60°C to 95°C with data collection. Data was collected and analysed by StepOne Software v2.0 (Applied Biosystems).

2.13.3 Generation of plasmid DNA standards for Beta Globin

The beta globin gene was amplified by the primers outlined in Table 2.3, as described in Section 2.12.2. This amplified sequence was then purified using QIAEX II Gel Extraction kit (Qiagen, Cat # 20051), as per the manufacturer's instructions, and sequenced by Sanger sequencing, as described in Section 2.12.3 and then cloned into the pCR2.1-TOPO plasmid vector (Invitrogen, Cat # 46-0801) using TOPO TA cloning kit (Invitrogen, Cat # K4500-01), as per the manufacturer's instructions.

Briefly, 5 μ L of purified PCR product was added to 1 μ L TOPO vector, the ends of which contain a DNA ligase, facilitating cloning. Ligated DNA was transformed into OneShot TOP10 *E. coli* cells (Invitrogen) by heat-shock transformation, as per the manufacturer's instructions. As the pCR2.1-TOPO contains a ampicillin resistance gene (Figure 2.3), successful transformants were selected on LB agar containing 50



LacZα fragment: bases 1-547
M13 reverse priming site: bases 205-221
Multiple cloning site: bases 234-357
T7 promoter/priming site: bases 364-383
M13 Forward (-20) priming site: bases 391-406
f1 origin: bases 548-985
Kanamycin resistance ORF: bases 1319-2113
Ampicillin resistance ORF: bases 2131-2991
pUC origin: bases 3136-3809

Figure 2.3. Plasmid map of pCR®2.1-TOPO®. The TOPO TA cloning vector pCR®2.1-TOPO® (Invitrogen, Cat # 46-0801) was used as a vector PCR products amplified from the beta-globin gene. The Ampicillin region encodes β-lactamase. Figure reproduced with permission from TOPO TA cloning vector user manual (Invitrogen, version U 2006).

µg/mL ampicillin. The pCR2.1-TOPO cloning site is in the middle of a LacZ gene and thus colonies containing TOPO vectors with DNA inserts could be colour identified using 40 µL of 40 mg/mL X-gal (Roche, Cat # 11680293001) in dimethylformamide (Sigma, Cat # D4551) spread on each agar plate. Agar plates were incubated at 37°C O/N.

The pCR2.1-TOPO cloning site is flanked by M13 sequences. Thus, the insert products can be amplified and sequenced using M13 primers. White colonies were picked with a sterile toothpick and dipped firstly into a 25 µL Amplitaq Gold PCR mix containing M13 primers, and secondly into 10 mL Luria Broth containing 50 µg/mL ampicillin, which was incubated at 37°C O/N to culture the colony.

Products were amplified by PCR using conditions described in Section 2.12.2. PCR products were separated by gel electrophoresis, purified using QIAEX II Gel Extraction kit (Qiagen), as per the manufacturer's instructions, and sequenced as described in Section 2.12.3 to confirm the sequence of the insertion.

Plasmid DNA was extracted from 2 mL of the Luria Broth culture containing the bacteria with plasmid of interest using a QIAprep Spin Miniprep Kit (Qiagen, Cat # 27104). Plasmid concentration (and thus copy number) was determined using optical density spectrometry (as described in Section 2.11).

2.14 *invPCR*

Two inversion designs were used for the analysis of DNA extracts from HBV-infected liver tissue. One inversion design was used for the analysis of DNA extracts from WHV-infected woodchuck liver tissue. The designs are outlined in Table 2.4 and Figure 2.4. Control PCRs specific for inverted and un-inverted forms of a single copy cellular gene procollagen alpha I (Boyd, Weliky et al. 1986), were carried out to determine the efficiency and the yield of the inversion reaction (Figure 2.5). Primers for the procollagen alpha I reactions are summarised in Table 2.5.

DNA extracts analysed by *invPCR* included: 1) 1/10 of each DNA extract from 5 mg of liver tissues (Section 2.7); 2) 2.5/10 of each DNA extract from EtOH-fixed,

Table 2.4A. invPCR designs

Species	1 st cut RE	2 nd cut RE	3 rd cut RE	Forward Outer primer ³	Reverse Outer primer ³	Forward Inner primer ³	Reverse Inner primer ³	Patient ⁴						
Wood-chuck ¹	<i>SacI</i>	<i>HindIII</i>	<i>AflII</i>	S1 F	S1 R	S2 F	S2 R	WC366						
Human ²	<i>NcoI</i>	<i>BsiHKAI</i>	<i>SphI</i>	1380	1383	1381	1416	HN						
						1385	1382	C, CY, FMC1, FMC3, FMC5, FMC7, GS1, GS4, Y2, Y3, Y4, Y6, XA						
								1397	Y5					
						1414	1382	FMC2, FMC4, FMC6, GS2, L, MH						
						1395	1381	1416	GS3, GS5, HS, WN					
							1385	1382	CYRY, HS					
								1416	WN					
						1383/ 1395	1385	1382	CN, NT, SAAO, DG					
						Human ²	<i>DpnII</i>	<i>BsiHKAI</i>	<i>NcoI</i>	1380	DpnII A	1385	1383	CN, HN, SAAO, GS4
											DpnII B	1385	1383	XA
DpnII C	1385	1383	Y2, Y4, Y6, FMC3											

¹ Design was previously developed by Summers *et al.* (2003).

² Designs were previously developed by Mason *et al.* (2010).

³ Sequences of primers are summarised in Table 2.4B.

⁴ Various combinations of invPCR designs and primers were used for different patients due to the sequence variation within the infecting HBV DNA.

Table 2.4B. Sequences of PCR primers used for invPCR.

Primer name	Sequence (5' -> 3')	Target on HBV sequence (nt) ³	Tm (°C) ⁴	Length (nt)
S1 F ¹	AACTTAGTATGCTATGAATTA	2185	44	21
S1 R ¹	TCTAATAGCTGTATGGTGCA	2195	51	20
S2 F ¹	ACTAAATTGATAGCTTGGATGAGCTCTA	2165	56	28
S2 R ¹	GAGCATGTCCCTACCTGTAG	2217	54	20

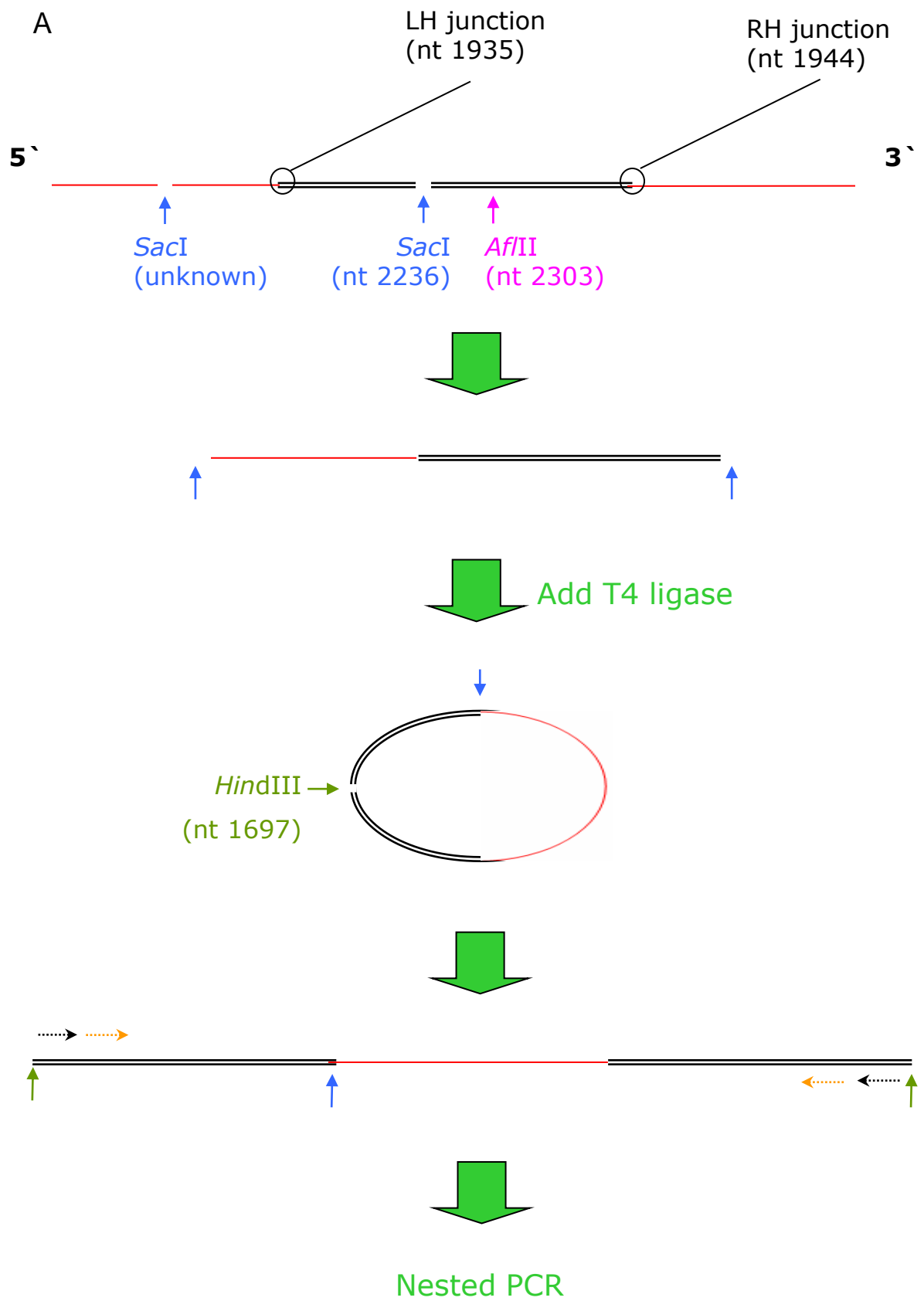
1380F ²	TTCGCTTCACCTCTGCACG	1603	58	19
1381F ²	TGGAGACCACCGTGAACG	1626	57	18
1382R ²	CACAGCCTAGCAGCCATGG	1390	58	19
1383R ²	AAAGGACGTCCCGCGCAG	1422	61	18
1385F ²	CGCATGGAGACCACCGTGA	1623	60	19
1395R ²	AAAGGACGTCCCGCGAAG	1422	58	18
1397R ²	CACACCCTAGCAGCCATGG	1390	58	19
1414F ²	CGCATGGAAACCACCGTGA	1623	58	19
1416R ²	AGTACAGCCTAGCAGCCAT	1388	55	19
DpnII AR	GGGACGAGAGAGTCCCAAG	1516	56	19
DpnII BR	GGGACGGTAGAGTCCCAAAC	1515	57	20
DpnII CR	GGAGAAGCGGGCGGTAGAGC	1516	62	20

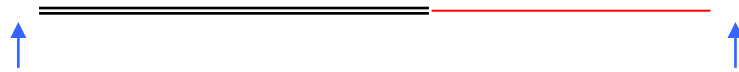
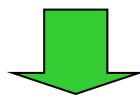
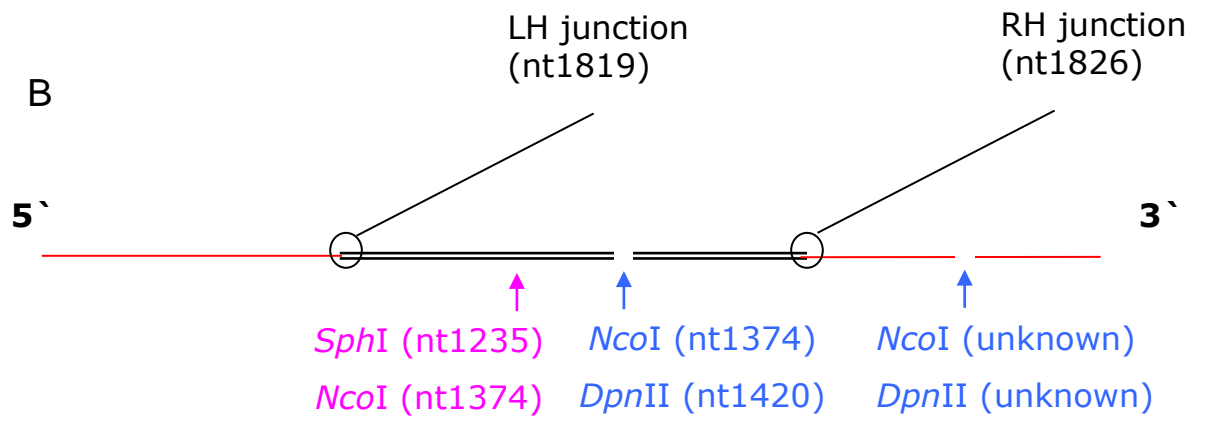
¹ Primer sequences previously designed by Summers *et al.* (2003).

² Primer sequences previously designed by Mason *et al.* (2010).

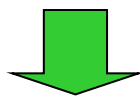
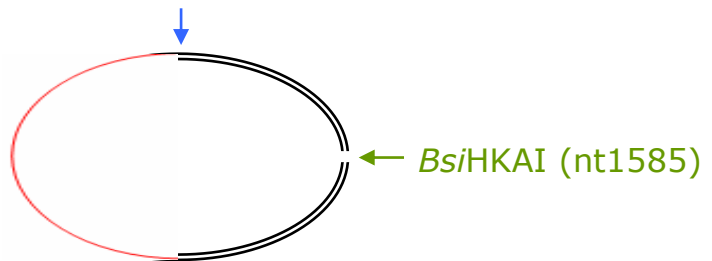
³ Based on sequences from Genbank Accession #J02442 for WHV DNA and Genbank Accession #AB241115 for HBV DNA.

⁴ Melting temperature (Tm) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.





Add T4 ligase



Nested PCR

Figure 2.4. Inversion of DNA extracted from WHV-infected woodchuck liver (2.4A) and HBV-infected human liver (2.4B). These designs have been used previously to detect virus-cell junctions in DNA extracts of liver tissue from woodchucks chronically infected with WHV (Mason *et al.* 2005) and from patients chronically infected with HBV (Mason *et al.* 2010).

To detect virus cell junctions in woodchuck liver tissue (2.4A), the left-hand virus-cell junction formed by the integration of WHV DNA (black double line) into cellular DNA (red) was excised with digestion with *SacI*. The excised product was diluted and circularised using T4 DNA ligase. The circularised product was cleaved with *HindIII* to form an inverted product, with known viral sequences surrounding unknown cellular sequences. The inverted products were subsequently cleaved with *AflIII* to reduce aberrant amplification of WHV DNA (explained in detail in Figure 2.6). The expected nucleotide numbers of dsDNA ends and restriction sites were predicted based on by the WHV sequence previously elucidated by Galibert *et al.* (1982) (Accession #J02442).

To detect virus cell junctions in human liver tissue (2.4B), the right-hand virus-cell junction formed by the integration of HBV DNA into cellular DNA was excised with digestion with either *NcoI* or *DpnII*. The excised product was diluted and circularised using T4 DNA ligase. The circularised product was subsequently cleaved with *BsiHKAI* to form an inverted product, with known viral sequences surrounding unknown cellular sequences. The inverted products were also cleaved with either *SphI* or *NcoI* to reduce aberrant amplification of HBV DNA (explained in detail in Figure 2.6). The expected nucleotide numbers of dsDNA ends and restriction sites were predicted based on Genbank Accession #AB241115.

Inverted products were then amplified by nested PCR primers (black and orange arrows), as described in greater detail in Figure 2.7.

Procollagen alpha I gene

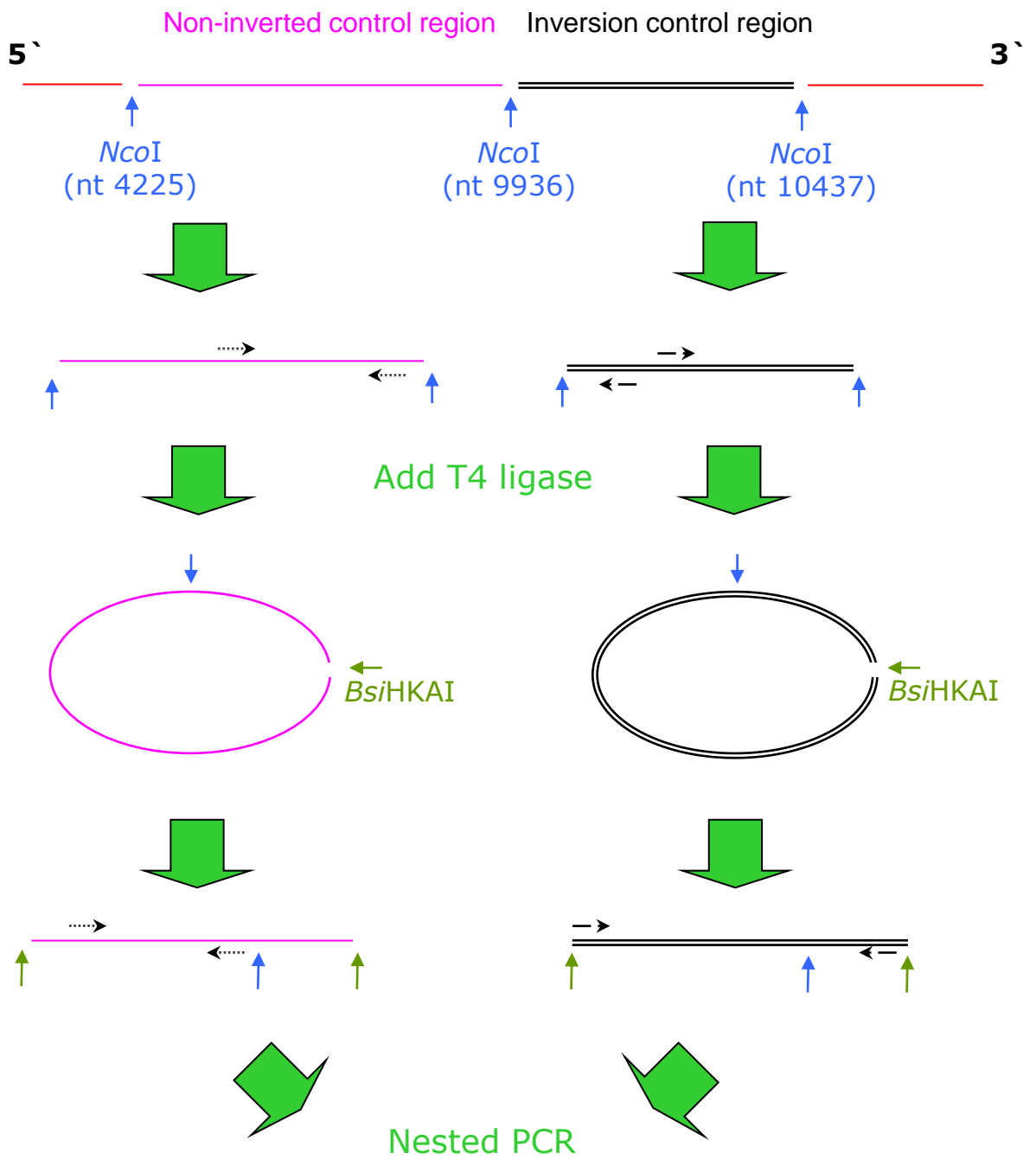


Figure 2.5. Control reactions for the inversion of DNA extracted from HBV-infected human liver. Two portions of the procollagen alpha 1 gene (red line) were excised using *NcoI*: the upstream section was assigned as the control uninverted region (pink); the downstream region was assigned as the inversion control region (black double line). The excised product was diluted and circularised using T4 DNA ligase. The circularised product was cleaved with *BsiHKAI* to form an inverted product. The inverted products were also cleaved with *SphI*.

The relative primer positions are shown for PCRs detecting the uninverted control region (short dashed arrow) and the inverted control region (long dashed arrow). Note that the product amplified from the uninverted region (403 bp) is formed regardless of whether inversion occurs, whereas the product amplified from the inverted region (434 bp) is only formed when inversion occurs. Inversion efficiency was defined as the ratio between number of inverted products to the number of uninverted products amplified from extracted DNA.

Table 2.5. Primers for control experiments on *NcoI* invPCR design

Primer name	Sequence (5' -> 3')¹	T_m (°C)²	Length (nt)
PC Inverted F outer	CCTGGCCTTCAGGGAATG	55	18
PC Inverted R outer	CACCAGCATCACCTATGTG	55	20
PC Inverted F inner	CTGGTGAACGTGGTGACG	56	18
PC Inverted R inner	AACCAAGAAGACTGGAGTGAGG	56	22
PC Uninverted F outer	CCTCCCCGTGACTGTAGTGT	58	20
PC Uninverted R outer	AAACAGCCCCTCTCTTCCTC	56	20
PC Uninverted F inner	CTGGGAGGTAGGGGTAGGAA	56	20
PC Uninverted R inner	CCCCAAACTATGGACCAAGA	54	20

¹ Single copy gene procollagen alpha I was used as a positive control for invPCR. Primer design based on the human procollagen alpha I sequence from Genbank accession #AF017178.

² Melting temperature (T_m) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.

paraffin wax-embedded sections of liver tissue (Section 2.8); and 3) 2.5/10 of each DNA extract from tissue isolated by laser-microdissection (Section 2.21). Each portion of DNA was digested in a 40 μ L reaction volume with 5 U of the 1st cut restriction enzyme (RE) for 1 hr and heat inactivated at 70°C for 20 min.

Reaction mixture was then made up to 450 μ L in 1x T4 DNA ligase buffer. 500U of T4 DNA ligase was added to circularise the DNA fragments produced by the previous digestion. The large 450 μ L ligation reaction volume favoured the self-ligation and circularisation of DNA fragments as opposed to intramolecular ligation (calculations are shown in full in Appendix 9.3). The ligation reaction was left to occur at RT for 2 hr. T4 ligase was heat inactivated at 70°C for 20 min. Denaturation of ligation was ensured by addition of 10 μ L of 10% SDS (Sigma, Cat # L57501kg). In order to increase efficiency of EtOH precipitation, 10 μ L of 5 M NaCl (Merck, Cat # 1.03404.0500) and 2 μ L of 20 mg/mL dextran (Sigma, Cat # D4133) was added. After pulse vortexing, 1 mL of 100% EtOH was added, tubes were inverted and stored at -20°C O/N. Precipitated DNA was pelleted by centrifugation at 14000 g for 10 min and washed with 1 mL of 100% EtOH. The DNA pellet was dried under vacuum at RT for 20 min, then redissolved in 39.5 μ L of a RE reaction mix containing 5 U 2nd cut RE. After incubation at 65°C for 1 hr, the reaction mixture was cooled to 37°C, 5U 3rd cut RE was added to decrease the amplification of cccDNA (Figure 2.6) and the sample was incubated at 37°C for 1 hr.

A quarter, i.e. 10 μ L, of the inverted DNA was serially diluted and distributed into a 96-well PCR plate (Biorad, Cat # MLP-9601) as described in Figure 2.7. In some instances when DNA concentrations were low, serial dilution was not done; instead the DNA extract was added directly into 1 mL of Amplitaq Gold PCR mix (Applied Biosystems), vortexed briefly and 10 μ L of mix was distributed to each well. Primers used in the PCR were dependent on the HBV DNA subtype associated with each patient, and are summarised in Table 2.5.

PCR conditions were identical to those described in Section 2.12.2. Products of this PCR were transferred to a second 96-well plate containing 1 mL of 1x GoTaq Flexi PCR mix (Promega, Cat # M8296) with 2.5 mM MgCl₂. Again, primers for this reaction varied from patient to patient and are summarised in Table 2.5. A 96-pin

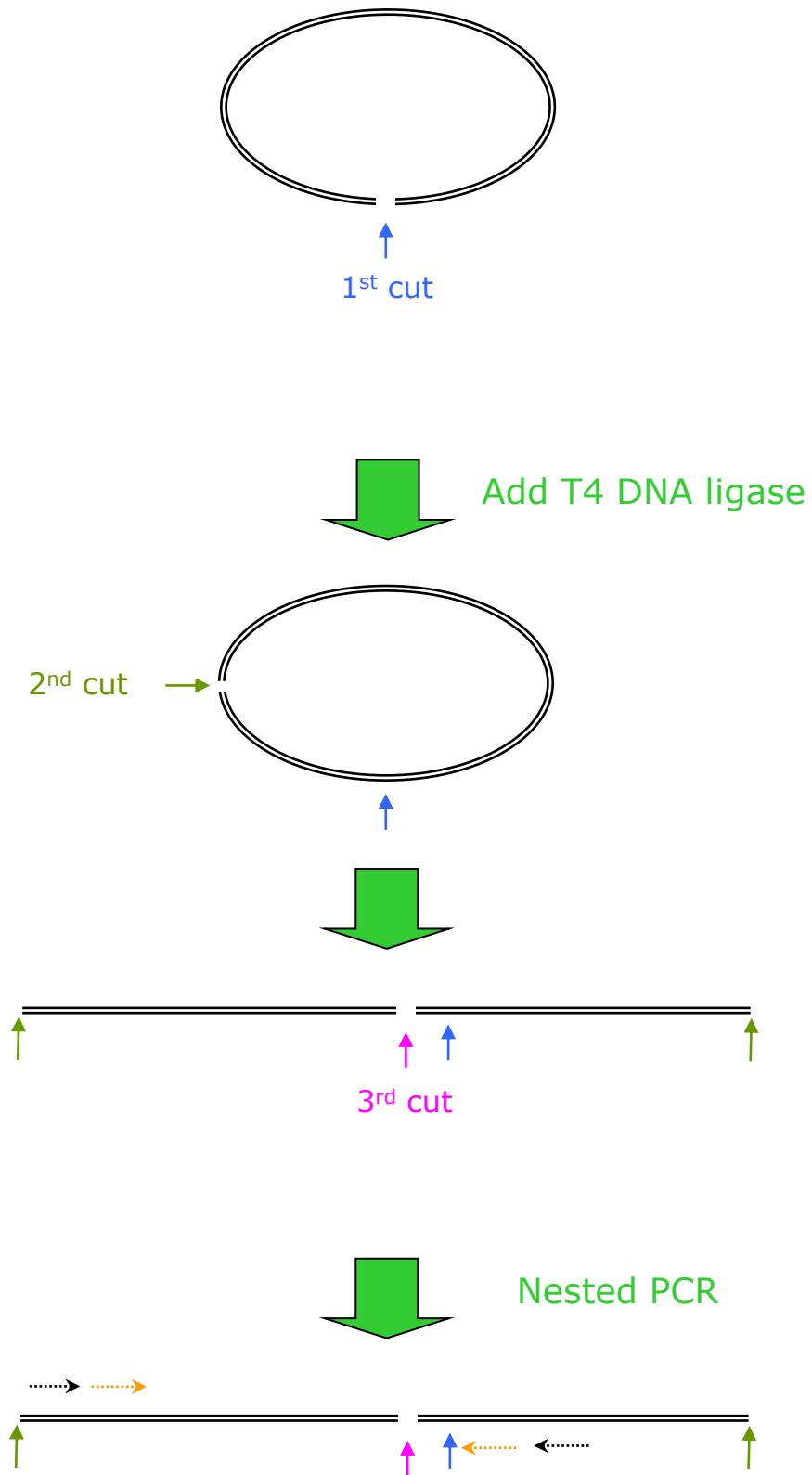
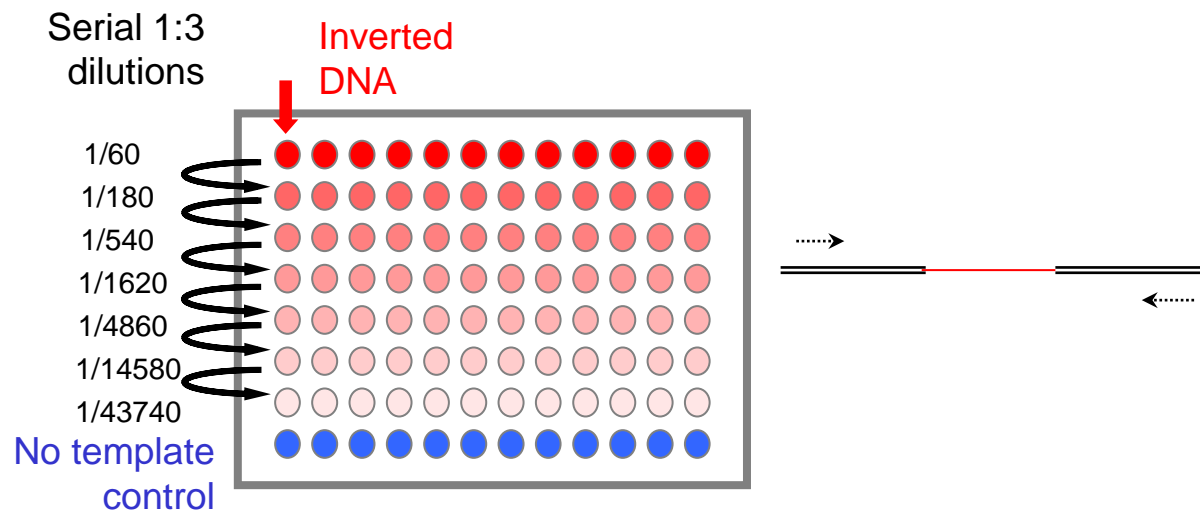
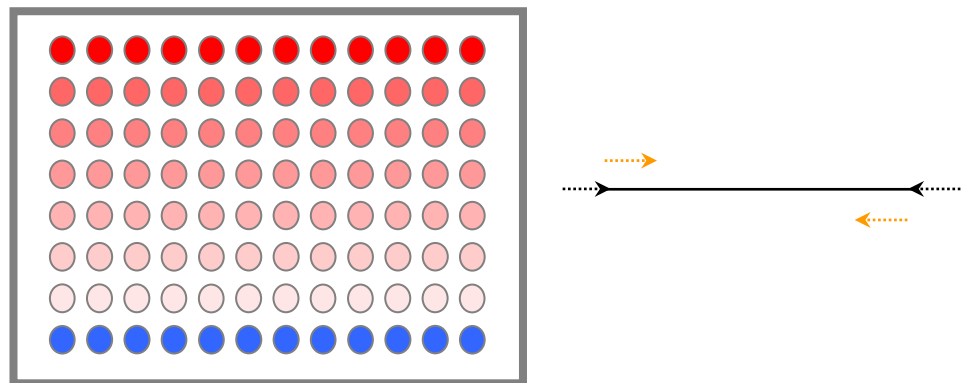


Figure 2.6. Inclusion of a third restriction enzyme during inversion limits the amplification of cccDNA. cccDNA (black double lines) can potentially be amplified during invPCR. If the first cut (blue arrows) is rejoined during the ligation reaction, then products could potentially be amplified by nested PCR. However, if a third restriction enzyme (pink arrows) is used to digest the inverted products, then nested primers (black and orange dashed arrows) are unable to amplify the cccDNA-derived inversion product.



Flamed pin transfer



Gel electrophoresis

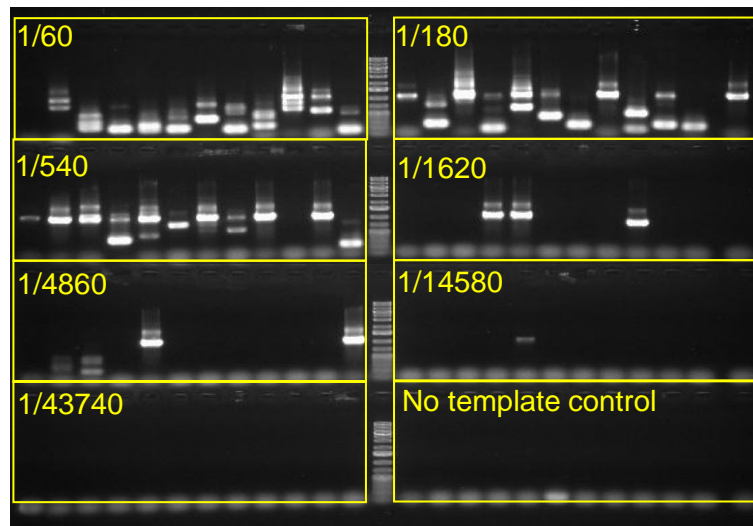


Figure 2.7. Nested PCR of inverted DNA. Inverted DNA was serially 1:3 diluted down the first column of a 96-well plate in a PCR mixture containing outer nested primers (Black dashed arrows). The numbers on the left represent the dilution with respect to the original DNA extract. PCR mixture was then distributed along the rows. Inverted products were amplified by PCR as shown on the right. A 96-pin replicator was sterilised by heating and used to transfer products of the PCR to another 96-well plate with each well containing 10 μ L of PCR mix with inner primers (orange dashed arrows). The products of the first PCR were then amplified by the second PCR. Products of the second PCR were then visualised by gel electrophoresis, isolated and sequenced. Repeated unique virus-cell junctions represented clones of hepatocytes, the size of which was determined by end-point dilution.

replicator (Nunc Nalgene, Cat # 250520) was heated until red hot with a bunsen burner and used to transfer PCR products to the second PCR. The nested PCR was carried out using the same conditions as Section 2.12.2, except that the initial denaturation step at 95°C was 2 min.

PCR products were analysed by gel electrophoresis, and then purified using QIAEX II Gel Extraction kit (Qiagen) with a modified protocol for a 96-well tray format (Section 2.16). Products were then Sanger sequenced and submitted to AGRF for capillary separation, as described in Sections 2.12.3.

Since each amplified product could represent greater than one template molecule, the copy number of each unique virus-cell DNA junction was calculated using the most probable number method via MPN calculator (Curiale 2004), freely available at <http://www.i2workout.com/mcuriale/mpn/index.html>. This calculator takes the number of positive products per dilution as input and returns the most probable number of virus-cell DNA junctions in the original extract with 95% confidence intervals.

2.15 *invPCR of high molecular weight DNA*

In some instances, low molecular weight HBV replicative intermediate DNA was suspected of producing artefacts during *invPCR*, so high molecular weight genomic DNA was isolated through agarose gel purification. Thus, 5 µL of total liver DNA extract was separated by electrophoresis through a 1% low melt agarose gel (Biorad, Cat # 161-3111) at 60 V for 2 hr. High molecular weight genomic DNA was excised with a disposable plastic drinking straw and placed in 1 mL 1st cut RE buffer with 0.1% Triton X-100. The agarose was equilibrated with 1st cut RE buffer at 4°C O/N. The supernatant was aspirated and 5U of the 1st cut RE was added. The high molecular weight DNA was digested with incubation at 37°C for 1hr. The 1st cut RE was then heat inactivated at 70°C for 20 min. The DNA was then purified using QiaQuick PCR purification kit (Qiagen, Cat # 28104), as per the manufacturer's protocol. *InvPCR* progressed from the ligation step of Section 2.14.

2.16 Gel extraction in 96-well format

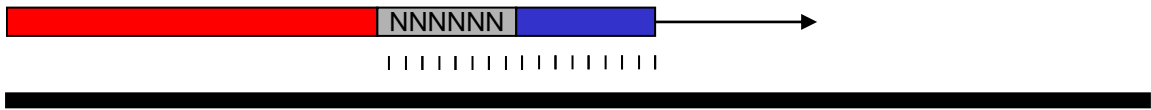
DNA products were excised from agarose gels using disposable drinking straws and placed into Thermowell® 96-well thick-skirted PCR plates (Costar, Cat # 6551). DNA was extracted from excised agarose gel using a modified protocol of described in for the QIAEX II Gel Extraction kit (Qiagen). To each well, 100 µL of QX1 Buffer (Qiagen), 5 µL of QiaexII glass beads (Qiagen), 1 µL of 10% Triton X-100, and 10 µL of 3 M sodium acetate were added. Agarose was dissolved at 55°C for 30 min with occasional vortexing. Beads were pelleted with centrifugation at 3000rpm for 3 min, then washed twice with 50 µL of QX1 Buffer containing 0.1% Triton X-100. Beads were then washed three times with 100 µL of PE buffer (Qiagen). Pellets were vacuum-dried for 20 min. To elute DNA from the dried glass beads, 30 µL of Elution Buffer (Qiagen) was added to each well. The tray was vortexed, and incubated for >5 min at RT. Beads were pelleted with centrifugation at 3000 rpm for 3 min, then the DNA-containing supernatant was collected for sequencing as described in Section 2.12.3.

2.17 Semi-DOP PCR

Semi-degenerate oligonucleotide primer (semi-DOP) PCR was used to detect unknown integrated HBV-cell DNA junctions as shown in Figures 2.8 and 2.9. The four degenerate oligonucleotide primers used were based on primers developed by Das *et al.* (Das, Noe et al. 2005) and all primers used are summarised in Table 2.6.

For control experiments, ~1 µg of total liver DNA extracted from normal human liver was spiked with ~15 ng (equivalent to $\sim 1.3 \times 10^9$ copies) of pHBV1.3BB4.5. Experimental samples were total DNA extracts from HBV-infected patients, gel-purified for high-molecular weight DNA as described in Section 2.15. A 1 mL of 1x Amplitaq Gold PCR mix (Applied Biosystems) with 2.5 mM MgCl₂ was made up containing one of four DOPs and one biotinylated HBV-specific outer primer. The DNA extract was diluted in a 96-well PCR plate as described in Figure 2.7. PCR conditions were as follows: initial 10 min denaturation step at 95°C to activate hot-start polymerase; six non-specific amplification cycles of 30 sec denaturation step at 95°C, annealing step at 30°C ramping up to 72°C at 0.2°C/sec, and 3 min elongation

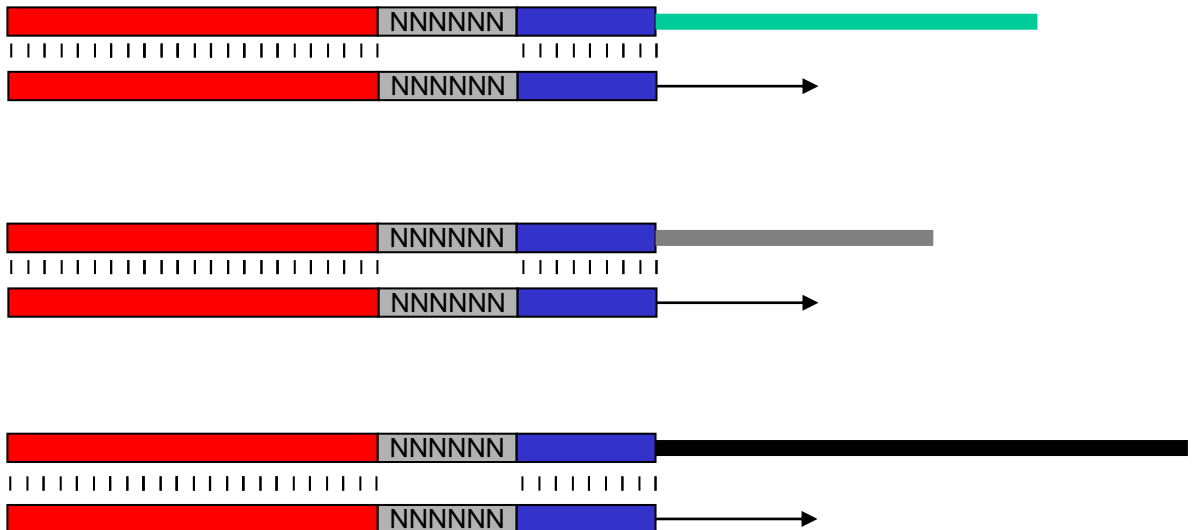
First phase - Random amplification



Low annealing temp =
annealing at random N and anchor regions



Second phase - Specific amplification



High annealing temp =
annealing at tag and anchor regions

Figure 2.8. Using semi-degenerate oligonucleotides to prime PCR amplification of unknown sequences. Degenerate oligonucleotide primer (DOP) PCR was first described by Telenius *et al.* (1992). DOP are composed of 3 regions: a tag region (red) that was not specific for any sequences within the template DNA; a random region (grey) made up of 6-10 random nucleotides; and an anchor region (blue) made up of a defined 4-6 base pairs, which acted as a firm 3' bond in second phase specific amplification. Random amplification occurs in the first phase of DOP-PCR; a low annealing temperature (~30°C) allows the primer to bind to the template DNA based on the short anchor region and the random region. The first phase is carried out for 5-10 cycles to increase the pool of randomly primed template. The second phase specifically amplifies products from the first phase; a high annealing temperature (~45°C) allows binding only at the tag and anchor regions.

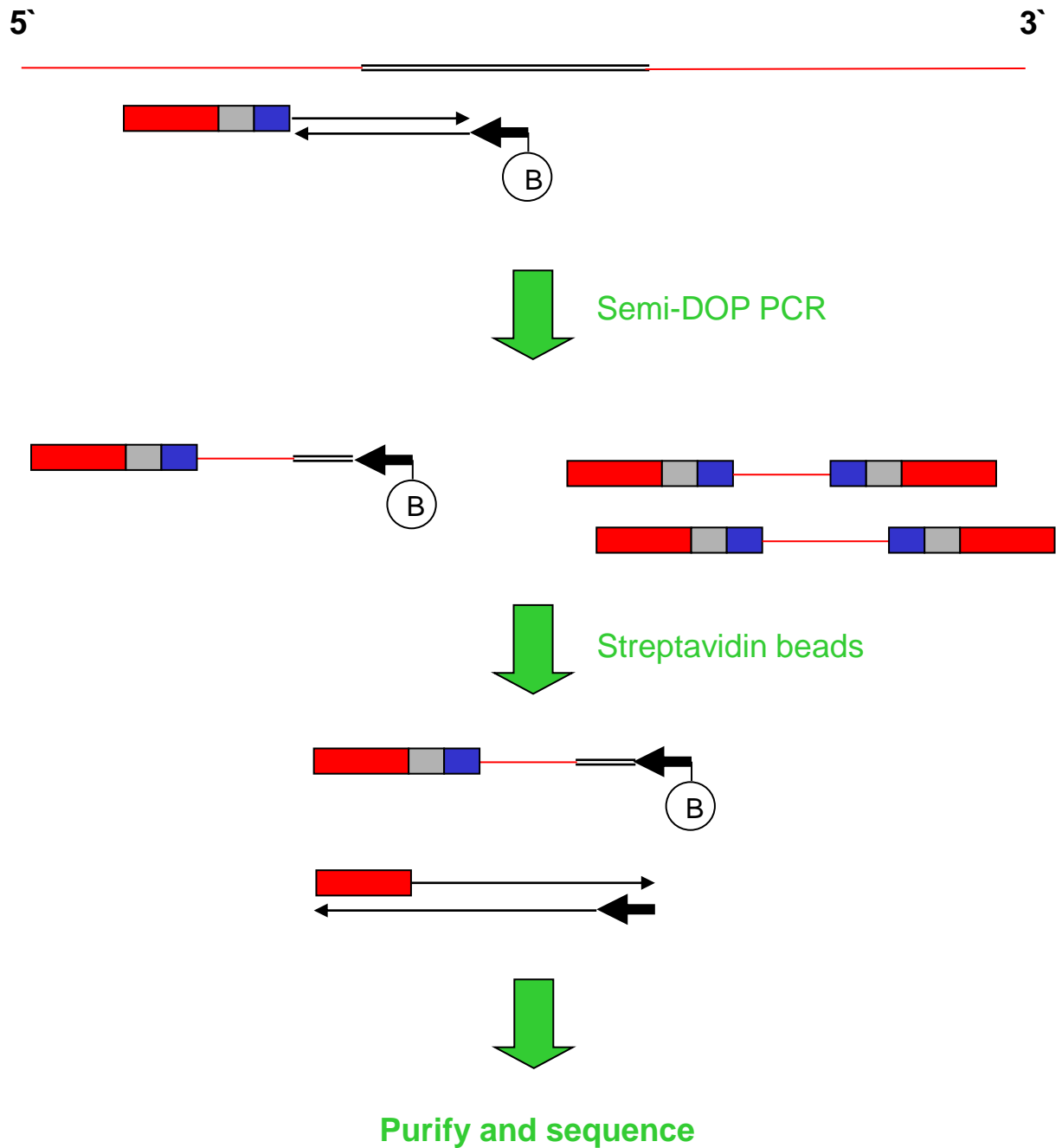


Figure 2.9. Use of semi-DOP PCR to detect virus-cell junctions. DOP were used to prime the cellular end of integrated HBV DNA. The other end of the product was primed with a HBV-specific, biotin 5'-capped primer. Products of the semi-DOP PCR are comprised of virus-cell junctions and DOP-DOP products. Streptavidin beads were used to specifically isolate the virus-cell junctions via the biotinylated HBV-specific primer. A nested reaction using primers specific for the DOP tag-region and HBV DNA was used to increase specificity and sensitivity. Products were then purified and sequenced using an inner primer specific for the HBV DNA.

Table 2.6. List of DOP (2.6A) and HBV-specific (2.6B) primers used for semi-DOP PCR.

Primer name	Sequence (5' -> 3')	T_m (°C)⁴	Length (nt)
Tag1-10-TCGA ¹	GGCCACGCGTCGACTAGTCANNNNNNNNNNTCGA	-	34
Tag1-10-TCGAC ¹	GGCCACGCGTCGACTAGTCANNNNNNNNNNTCGA C	-	35
Tag1-8-TCGA ¹	GGCCACGCGTCGACTAGTCANNNNNNNNTCGA	-	32
Tag1-8-TCGAC ¹	GGCCACGCGTCGACTAGTCANNNNNNNNTCGAC	-	33
Tag1 ¹	GGCCACGCGTCGACTAGTCA	62	20
Tag2-10-TCGAC ²	GGCCACGGCAGTGCTTCTCANNNNNNNNNNTCGA C	-	35
Tag2 ²	GGCCACGGCAGTGCTTCTCA	63	20
LHHBV Outer ³	(B) GAGCTGAGGCGGTGT	53	15
LHHBV Inner ³	AAAACGAGAGTAACTCCA	48	18
LHHBV Seq ³	AAAACGAGAGTAACTCCACAGTAGC	57	25

¹ Sequences were previously designed by Das *et al.* (2005).

² Tag2 was designed using freely available online software Primo Random Primer Design 3.4 (<http://www.changbioscience.com/primo/primor.html>, Chang Bioscience).

³ HBV-specific primers were designed based on the HBV sequence contained within pBB4.5HBV1.3 (Genbank accession #AF305422). The full sequence of pBB4.5HBV1.3 is shown in Appendix 9.1. (B) represents 5' biotinylation modification on the HBV-specific primer.

⁴ T_m was not listed for primers with degenerate sequences as melting temperature is sequence-dependent. Melting temperature calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations

step at 72°C; then 30 tag-specific amplification cycles of 30 sec denaturation step at 95°C, 30 sec annealing step at 45°C and 3 min elongation step at 72°C; then a holding step at 15°C.

Products amplified using the biotinylated HBV-specific primer were isolated using magnetised streptavidin-coated beads (Roche, Cat # 11641778001). Beads were washed three times with an equal volume TEN₁₀₀ (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl pH 7.5). Four volumes of TEN₁₀₀ and 10 µL of washed magnetic beads were added to each PCR reaction, then incubated at RT for 30 min with mixing to allow binding of amplified products to beads. Unbound PCR products were discarded from the bead mixture with two washes with 100 µL TEN₁₀₀₀ (10 mM Tris-HCl, 1 mM EDTA, 1 M NaCl pH 7.5). The beads were then resuspended in 20 µL PCR-grade gamma-irradiated DW.

For a nested reaction, a total of 1 µL of bead suspension was transferred to a second 96-well plate containing 9 µL of GoTaq PCR mix (Promega) per well. The primers used were specific to the tag region of the DOP and specific for a HBV sequence internal to the one used in the first PCR as shown in Figure 2.9. PCR was carried out using the following conditions: initial 2 min denaturation step at 95°C to activate hot start polymerase; 30 cycles of 30 sec denaturation step at 95°C, 30 sec annealing step at 50°C and 3 min elongation step at 72°C; then a holding step at 15°C. Products were analysed by gel electrophoresis, purified with QIAEXII beads, as per the manufacturer's instructions, and sequenced using an internal primer as described in sections 2.12.3.

2.18 *qPCR analysis of virus-cell DNA junctions*

The absolute numbers of specific virus-cell DNA junctions in total DNA extract were quantified using a plasmid standard curve of known copy numbers. Previously purified invPCR products (Section 2.16) were used to produce plasmid DNA standards, as described in Section 2.13.3.

In order to quantify virus-cell DNA junctions, qPCR primers were designed using Primer3 plus software (found at <http://bioinformatics.nl/cgi->

bin/primer3plus/primer3plus.cgi) to detect specific virus-cell DNA junctions that were previously observed by invPCR in DNA extracts from patients XA and CY. Forward primers were designed to straddle the virus-cell DNA junction, whereas the reverse primers were specific to the downstream cellular DNA sequence. The primers used are listed in Table 2.7. Conditions for qPCR were identical to those described in Section 2.13.2.

2.19 Computer simulation

A computer simulation was written in Java using freely available Jcreator LE (Xinox Software). In instances when laboratory personal computers did not have enough capacity to run programs with large arrays, simulations were run using the supercomputer facility Hydra (eResearchSA).

2.20 Laser dissection of foci of hepatocytes

Paraffin wax-embedded, EtOH-fixed liver tissue sections were confirmed by invPCR to contain detectable clones, as described in Section 2.14. Tissue sections that were positive by invPCR were then sectioned at 5 μm and mounted to polyethylene-naphthalate (PEN) membrane slides (Leica, Cat # 11505158). Adjacent sections were mounted onto glass slides and HBcAg or HBsAg were detected by immunohistochemistry as described in Section 2.5. Foci of hepatocytes stained positive and negative for HBV antigens were observed in adjacent sections, and then located in the unstained PEN membrane slide mounted sections. PEN-membrane slides were dewaxed by 2x5 min washes in AR grade xylene, then dried with 2x5 min washes in absolute EtOH. HBV antigen-positive and negative foci of hepatocytes were isolated in the caps of 200 μL PCR tubes (Eppendorf, Cat # 0030124537) using Leica AS Laser Microdissection microscope provided by Adelaide Microscopy (Figure 2.10).

2.21 Total DNA extraction and invPCR of tissue isolated by laser-microdissection

Tissues isolated by laser-microdissection were digested in 200 μL digestion solution by incubating in a shaking thermomixer (Eppendorf) at 55°C for >3 hr. To aid in

Table 2.7. Primers used to amplify specific virus-cell junctions in DNA extracts from patient CY and XA.

Primer name	Sequence (5' -> 3') ³	Tm (°C) ⁴	Length (nt)
XA clone F ¹	<u>AGCGCCATGCGGAAGATA</u>	57	18
XA clone R ²	GTGGAAGTGTCAAATTCTTGAGC	55	23
CY clone F ¹	<u>CACCATGCGCTTGTGCTT</u>	57	18
CY clone R ²	AAAAGGTGACTGAGAGCGTCA	56	21

¹ Forward primers were specific for each virus-cell junction.

² Reverse primers were specific for downstream cellular sequences.

³ HBV-specific regions in the primers are underlined.

⁴ Melting temperature calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations

Table 2.8. Primers used amplify HBV pre-S regions by hemi-nested PCR to detect Pre-S mutants (previously designed by Fan *et al.* (2001))

Primer name	Sequence (5' -> 3') ¹	Target on virus sequence (nt) ²	Tm (°C) ³	Length (nt)
PreS HBV F	GCGGGTCACCATATTCTTGG	2837	56	20
PreS HBV Ro	GAGTCTAGACTCTGCGGTAT	255	53	20
PreS HBV Ri	TAACACGAGCAGGGGTCCTA	199	57	20

¹ Primer sequences were based on the HBV genome sequence listed under Genbank accession number X51970.

² Based on HBV DNA sequence from Genbank Accession #AB241115.

³ Melting temperature (Tm) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.

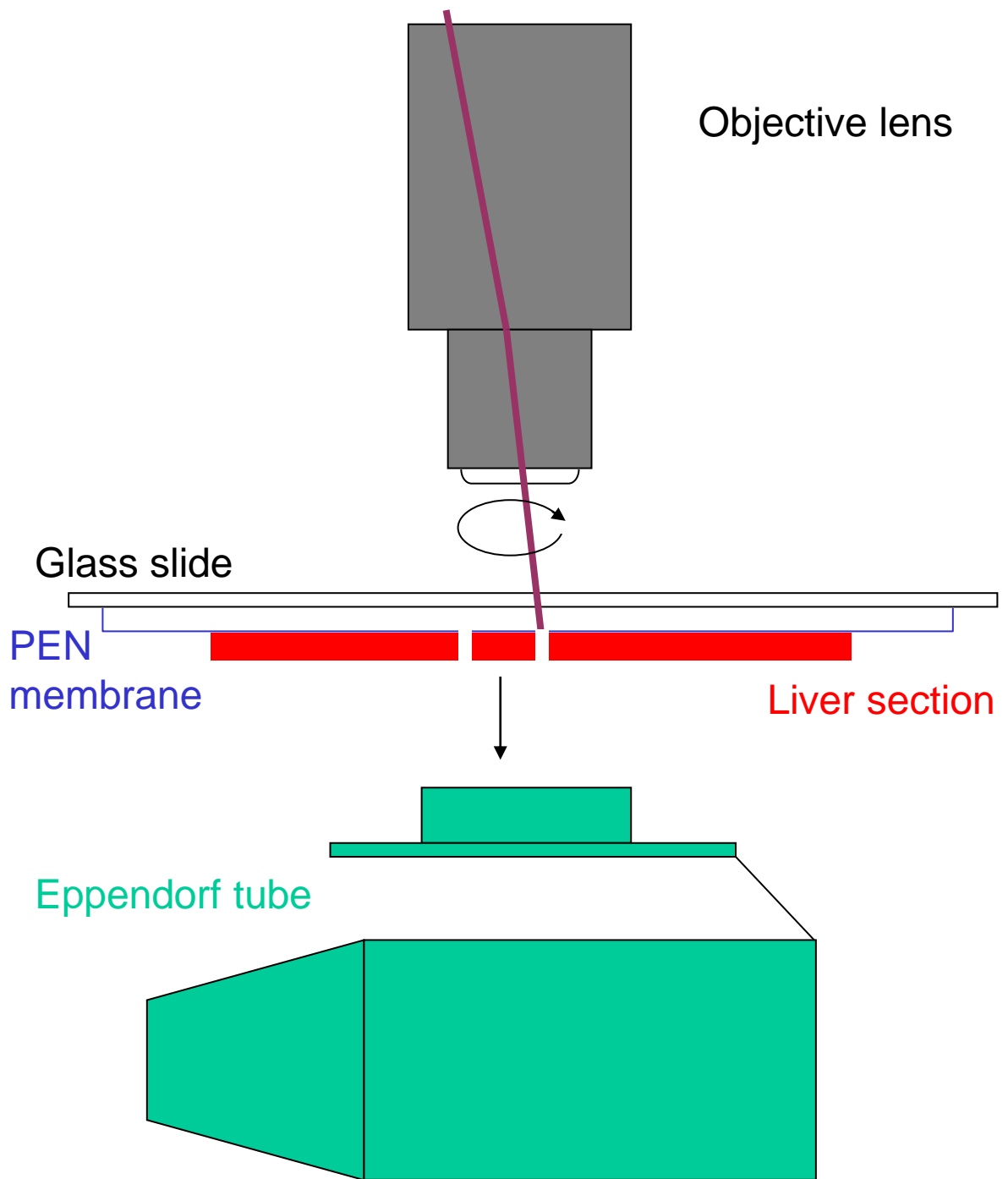


Figure 2.10. Set-up of the Leica LMD system. The ethanol-fixed, paraffin wax-embedded liver section (red) is mounted onto the polyethylene naphthalate (PEN) membrane of the glass slide and placed upside-down on the stage. The objective lens of the LMD microscope (grey) acts as a means of viewing the section and directing the UV laser beam (purple) onto the section. The liver tissue and PEN membrane are vaporised by the laser beam. Excised tissue sections drop by gravity into the lid of an 0.5 mL thin-walled Eppendorf tube.

DNA yield, 1 mg of carrier RNA and 500 µg of lambda phage DNA (Roche, Cat # 10745782001) or 500 µg pBlueBac4.5 plasmid DNA (Invitrogen) were added. 200 µL of gamma-irradiated PCR-grade DW was added to make up the volume to 400 µL. DNA extraction then progressed through (25:24:1) phenol:chloroform:isopropyl alcohol extraction, EtOH precipitation and redissolving steps, as described in section 2.7. A quarter of the DNA extract was analysed by invPCR as described in section 2.14.

2.22 Calculation of clone size from invPCR of DNA extracts from liver sections

Several assumptions were made in calculating the 3 dimensional size of hepatocyte clones detected in liver sections, via extracts of whole liver sections or tissue isolated by laser-microdissection: 1) invPCR had a 100% detection rate; 2) hepatocyte clones were spherical; and 3) hepatocyte density was the same as neighbouring tissue. Minimum hepatocyte clone size was calculated by assuming that the detected clonal hepatocytes had been taken from the largest diameter of a spherical clone (Figure 2.11). Approximate nucleus concentration (Conc_{Hep}) in the hepatocyte clone was calculated by dividing the nuclei count of the excised tissue by volume of the excised tissue.

2.23 Detection of preS-mutants in hepatocytes isolated by laser-microdissection

To detect mutations in Pre-S regions of HBV genomes extracted from hepatocytes isolated by laser-microdissection, primers previously designed by Fan *et al.* (Fan, Lu *et al.* 2001) were used as shown in Table 2.8. Ten µL of hepatocyte DNA extract were added to 110 µL of Amplitaq Gold PCR mix (Applied Biosystem) containing primers PreS HBV F and PreS HBV Ro (Table 2.8). The PCR mix was pipetted up and down and dispensed into 10 µL reactions and subjected to PCR conditions as described in section 2.12.2. PCR products were flame-pin transferred to a second 10 µL PCR of 1x GoTaq PCR mix (Promega) with 2.5 mM MgCl_2 containing primers PreS HBV F and PreS HBV Ri. PCR conditions used were identical to those as described in Section 2.12.2, except that the initial denaturation step was 2 min long. PCR Products were separated by agarose gel electrophoresis, extracted and sequenced using forward primer as described in Section 2.12.2.

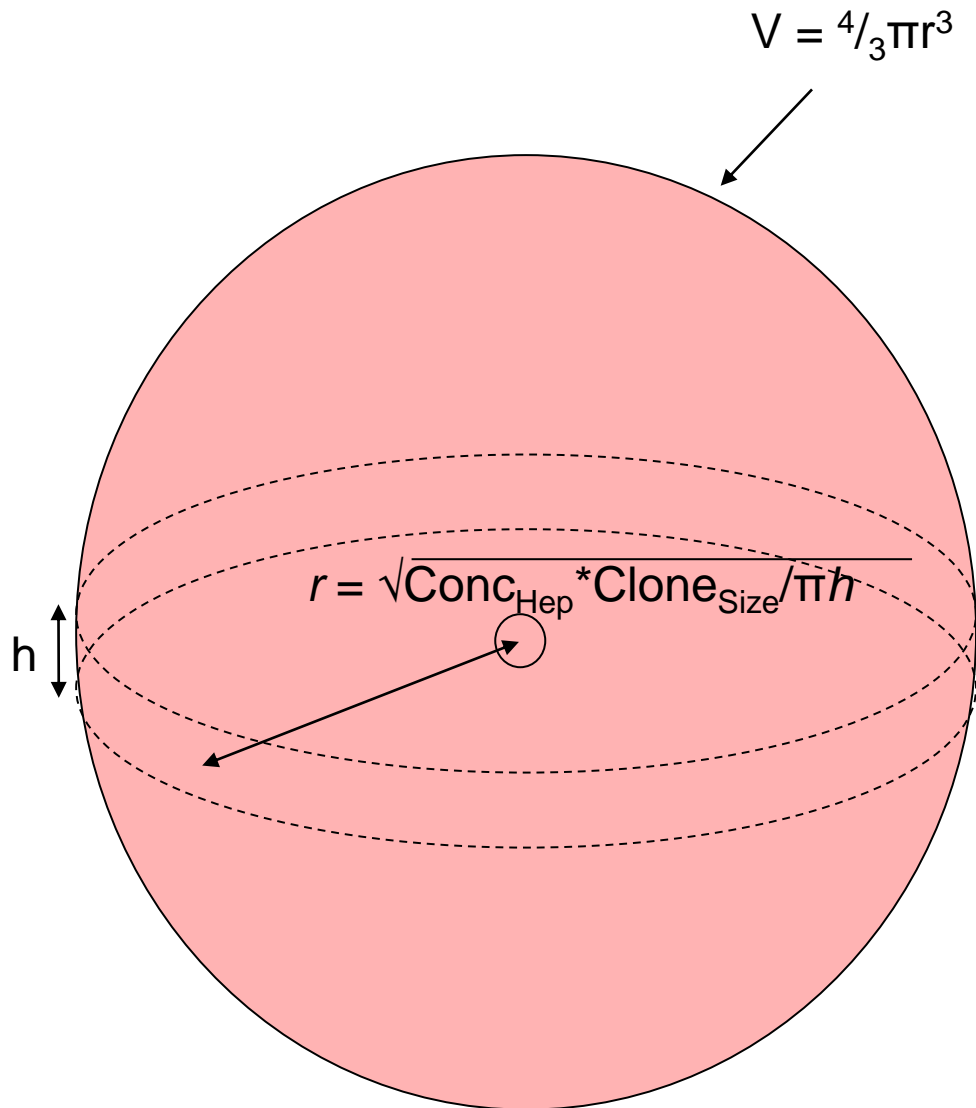


Figure 2.11. Calculation of 3-dimensional size of a hepatocyte clone from invPCR of a section of liver. During sectioning of the liver tissue, a cylindrical section of hepatocytes has effectively been taken from the spherical hepatocyte clone. The radius (r) of the cylinder, and thus the smallest spherical clone possible, was calculated by the equation: $r = \sqrt{\text{Conc}_{\text{Hep}} * \text{Clone}_{\text{Size}} / \pi h}$, where Conc_{Hep} is the experimentally derived concentration of hepatocytes (i.e. the number of nuclei per unit volume), $\text{Clone}_{\text{Size}}$ is the clone size detected by invPCR, and h is the thickness of the tissue section. The volume of spherical clone (V) was calculated by $V = \frac{4}{3}\pi r^3$. The number of cells in the spherical clone (N_{clone}) was then estimated by $N_{\text{clone}} = V / V_{\text{cell}}$, where V_{cell} is the volume of a hepatocyte.

Conc_{Hep} was calculated by dividing the nuclei count of a section of liver by the calculated volume of the section of liver (Section 3.2.1.6).

2.24 Quantification of methylated HBV DNA CpG regions in hepatocytes isolated by laser-microdissection

The methylation-sensitive RE *AciI* was used to determine the percentage of methylated HBV DNA in HBsAg-positive and -negative hepatocyte foci isolated by laser-microdissection (Figure 2.12). Four μL of extracted DNA was digested in 10 μL 1x RE mix, containing 10 U *EcoRI* (NEB) to linearise the HBV DNA molecules and 5 U *AciI* (NEB) to determine methylation. 5'-CCGC-3' sequences containing methylated cytosine bases would not be cleaved by *AciI*, whereas 5'-CCGC-3' sequences without methylated cytosine bases would be cleaved. Primers were designed to amplify sequences in CpG islands 1 and 2 in the HBV genome (Vivekanandan, Thomas et al. 2008) containing CCGC sequences (Table 2.9). Primers amplifying products not including CCGC sequences (Table 2.3) were used as a reference product. A comparative C_t qPCR assay was performed and the ratio between the C_t of *AciI*-cut and mock-cut assays was calculated.

For a negative control, pBB4.5HBV1.3 samples were treated with *SssI* methyltransferase (NEB, Cat # M0226), an enzyme that methylates all 5'-CG-3' dinucleotides, prior to *AciI* digestion. The ratio between the C_t of *AciI*-cut and mock-cut assays of *SssI*-treated plasmid DNA was used as a control to show that methylation does block digestions by *AciI*.

2.25 Estimation of global methylation of DNA from laser captured hepatocytes

This protocol was adapted from the method described by Yang *et al.* (Yang, Estecio et al. 2004). Global methylation was estimated by determining the methylation of promoters for long interspersed nucleotide elements (LINE). DNA extracts of HBsAg-positive and -negative hepatocytes were treated with bisulphite ions using Imprint DNA modification kit (Sigma, Cat # MOD50), as per the manufacturer's instructions.

Bisulphite treatment converts unmethylated cytosine bases into uracil. Bisulphite-treated DNA was amplified using LINE-1 promoter specific primers

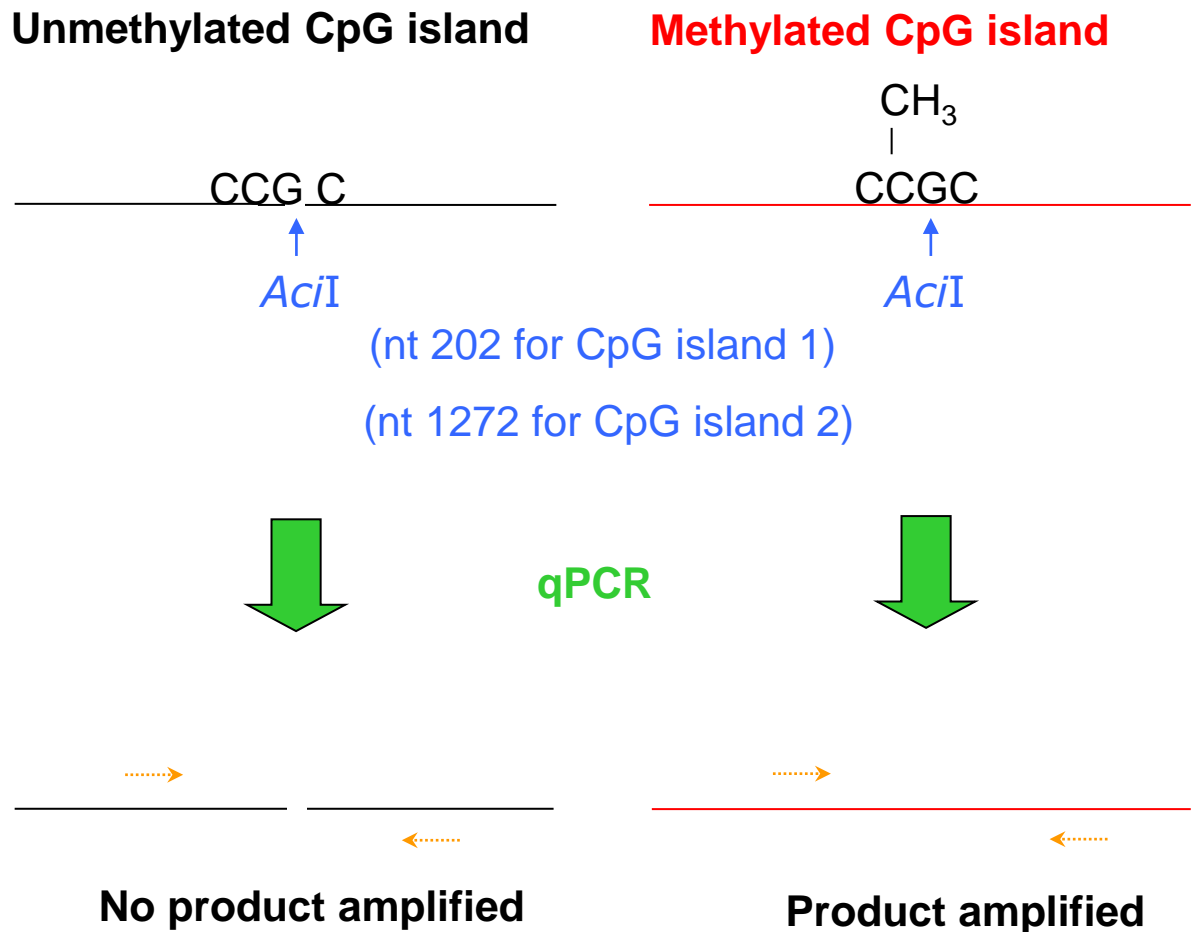


Figure 2.12. *AciI* methylation assay. The methylation sensitive restriction enzyme *AciI* was used to distinguish between non-methylated and methylated CpG islands in HBV DNA genomes extracted from laser-microdissected foci of hepatocytes. First, the DNA extract was digested with *AciI*. Unmethylated CpG motifs within *AciI* sites were cleaved, while methylated CpG motifs were protected. Primers spanning conserved *AciI* sites in CpG islands 1 and 2 of the HBV DNA genome were designed (Table 2.9). qPCR was carried out using these primers, thereby quantifying the level of cleavage at each site. The percentage of methylation was determined by the ratio of signal between *AciI*-digested samples to mock-digested samples (normalised by HBV DNA copy number).

Table 2.9. Sequences of primers used in *AciI* methylation assay.

Primer name	Sequence (5' -> 3')¹	Target on HBV sequence (nt)²	Tm (°C)³	Length (nt)
HBV CpG1 F	AGGACCCCTGCTCGTGTTAC	180	59	20
HBV CpG1 R	TTGAGAGAAGTCCACCACGA	273	55	20
HBV CpG2 F	CAAGTGTTTGCTGACGCAAC	1176	56	20
HBV CpG2 R	GGCTGCGAGCAAAACAAG	1306	55	18

¹ Primers sequence based on consensus sequences of strains outline in Table 2.1.

² Based on HBV DNA sequence from Genbank Accession #AB241115.

³ Melting temperature (Tm) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.

5'TTGAGTTGTGGTGGGTTTTATTTAG3' and 5'TCATCTCACTAAAAAATACC-AAACA3' (primer sequences designed by Yang *et al.* (Yang, Estecio et al. 2004)). PCR conditions were as described in section 2.12.2. Half of each reaction volume was digested with the RE *HinfI*, the restriction site of which is preserved if the promoter is methylated. *HinfI*-digested and undigested DNA separated by gel electrophoresis and analysed by Quantity One 1D gel analysis software (Biorad). The percentage of total DNA methylation was determined by the ratio of total DNA in the undigested lane to the DNA prone to *HinfI*-digestion.

2.26 *Imaging mass spectrometry (IMS)*

This protocol was adapted from Gustafsson *et al.* (Gustafsson, Oehler et al. 2011). Liver tissues were analysed by IMS in triplicate. EtOH-fixed paraffin wax-embedded liver tissue was sectioned at 5 μm intervals and mounted onto indium tin oxide conductive slides (Bruker Daltonics, Cat # 576352-25PAK). Sectioned tissue was dewaxed by 2x5 min washes in AR grade xylene, then rehydrated through a series of 2 min washes with an EtOH solutions of 100%, 90%, 80%, and then 70% AR grade EtOH, followed by 2 min wash in 10 mM NH_4HCO_3 . Either of two types of 20 ng/ μL trypsin were used to digest the sections: 1) 20 μg lyophilised modified porcine Trypsin (Promega, Cat # V5111) resuspended in 400 μL of 25 mM NH_4HCO_3 ; or 2) 0.5 $\mu\text{g}/\mu\text{L}$ Trypsin gold (Promega, Cat # V5280) in 5 mM NH_4HCO_3 , diluted 1:10 in 25 mM NH_4HCO_3 .

An ImagePrep station (Bruker Daltonics) was used to deposit the complete volume of trypsin by vibration vaporisation. Following digestion the ImagePrep station was used to deposit 7 mg/mL α -cyano-4-hydroxycinnamic acid (CHCA) in 50% v/v HPLC grade acetonitrile (ACN) (BDH, Cat # BDH6002-4) and 0.2% v/v trifluoroacetic acid (Merck, Cat # 108178). Control of deposition was via measurement of light scatter voltage, which is a surrogate for matrix crystal density on the slide surface. Mass spectrometry (MS) of tissues *in situ* was performed on an Ultraflex III MALDI-TOF/TOF instrument (Bruker Daltonics) operating in reflectron positive ion mode. Centre to centre acquisition distance was set at 100 μm . Acquisition on the MS instrument was controlled using flexImaging (V2.1, Bruker Daltonics) running an operator defined auto-execute sequence.

Spectra were processed using a chemical noise filter smoothing (width = 5 m/z) and TopHat background subtraction. If used downstream, spectra were internally calibrated on trypsin autolysis peaks. Ion intensity maps were generated in flexImaging and exported for presentation in MS PowerPoint.

2.27 High Pressure Liquid Chromatography Electrospray Ionisation Linear Ion Trap Orbitrap Mass Spectrometry (HPLC-nESI-LTQ Orbitrap MS)

EtOH-fixed slide-mounted liver tissue was isolated by laser microdissection and trypsin digested by one of two protocols.

- 1) The tissue was lysed in 100 μ L 6 M urea (Merck) at RT for 10 min. Lysate was applied to a VivaSpin 10 kDa centrifugal concentrator (Cole-Parmer, Cat # EW-36228-02) and centrifuged at 14 000 g for 30 min. Samples were then reduced and denatured by addition of 250 μ L denaturing buffer (6 M guanidine hydrochloride, 100 mM Tris, 5 mM EDTA, pH 8.0) containing 3 mg/mL DTT and incubation at 60°C for 1 hr. The denaturing buffer was washed out by centrifugation at 14 000 g for 30 min. Sample proteins were alkylated with addition of 250 μ L denaturing buffer containing 15 mg/mL iodoacetamide (GE Healthcare) and incubation at 37°C for 30 min. The alkylating buffer was washed out by centrifugation at 14 000 g for 30 min. Reduced and alkylated proteins were washed twice with 300 μ L 200 mM NH_4HCO_3 . Proteins were digested with 300 μ L of either 333 ng/mL or 33.3 ng/mL Trypsin gold (Promega) in 200 mM NH_4HCO_3 and incubation at 37°C O/N. Peptides were washed off the membrane twice by addition of 50 μ L 200 mM NH_4HCO_3 and centrifugation at 14000 g for 10 min.
- 2) The tissue was rehydrated in 10 μ L 25 mM NH_4HCO_3 at RT for 10 min. Proteins were digested with addition of 10 μ L of 10 μ g/mL Trypsin gold (Promega) at 37°C for 2hr. Digested peptides were then purified using PepClean C-18 Spin Columns (Thermo Scientific, Cat # PI-89873), as per manufacturer's instructions. Peptides were eluted into 40 μ L of 70% ACN.

Eluted peptides were freeze-dried to a volume <5 μL and dissolved to a volume of $\sim 10 \mu\text{L}$ in 3% ACN, 0.1% formic acid (FA, Sigma-Aldrich, Cat # F0507-1L). Peptides were then stored at -20°C until further analysis.

HPLC-nESI-LTQ-Orbitrap XL analysis was performed by Yin Ying Ho (Adelaide Proteomic Centre). An online Ultimate 3000 nHPLC system (Dionex) and LTQ-Orbitrap XL instrument (Thermo Scientific) was used to analyse the digested peptides. One to 5 μL of the peptide sample was loaded at 5 $\mu\text{L}/\text{min}$ flow rate using H_2O buffer onto an Acclaim Pepmap100 C-18 μ -precursor column cartridge (3 μm , 100 \AA , 300 μm I.D. x 5 mm). A micro-pump was brought in line with the precolumn and analytical separation was achieved using an Acclaim Pepmap100 C-18 column (2 μm , 100 \AA , 75 μm I.D. x 15 cm) running at 0.3 $\mu\text{L}/\text{min}$ over a 150 minute 0-55% B gradient of A) 0.1% v/v FA in 2% v/v ACN and B) 0.1% v/v FA in 80% v/v ACN. Eluted peptides were introduced by ESI into the LTQ-Orbitrap XL and the top 6 peptides in each scan were fragmented by CID. The following parameters were used: minimum signal = 1000; isolation width = 3.00; normalized collision energy = 35.0; charge state $\geq +2$.

RAW files produced by detection of masses were analysed by Qual Browser function of the Xcalibur program (v2.0.7 SP, ThermoScientific). Peptides were aligned using MASCOT analysis. The search parameters employed were as follows: precursor mass range = 350-5000 Da; protein database = SwissProt; maximum missed cleavages = 2; precursor mass tolerance = 0.5 Da; fragment mass tolerance = 0.8 Da; dynamic modifications = oxidation; static modifications = carbamidomethylation.

2.28 High Pressure Liquid Chromatography Matrix-assisted Laser Desorption Ionisation Time of Flight tandem Mass Spectrometry (HPLC-MALDI-TOF MS/MS)

2.28.1 HPLC and Fraction Collection

Alternatively, trypsin-digested peptides purified by Peplean columns were identified by HPLC in conjunction with MALDI-TOF mass-spectrometry. An UltiMate 3500 RS Nano/Cap System (Dionex) was operated using binary gradients of mobile phase

A (0.05 % TFA in 98 % H₂O, 2 % ACN) and B (0.04 % TFA in 80 % ACN, 20 % H₂O). The UltiMate 3000 Dionex System was controlled using the Hystar software platform (V3.2-SR2, Bruker Daltonics). A micro-WPS-3000 autosampler (Dionex) injected 5 μ L of purified trypsin-digested peptides onto an Acclaim Pepmap 100 trap column (75 μ m \times 2 cm, 3 μ m, 100 Å). Sample loading was performed with a loading pump with a flow rate of 3 μ L/min under aqueous conditions (0 % B). After desalting for 10 min, the trap column was switched inline to an Acclaim Pepmap100 C-18 analytical column (75 μ m \times 15 cm, 3 μ m, Dionex) at a flow rate of 300 nL/min. The peptides were eluted with an ACN gradient from 4 to 8% solvent B in 1 min, 8 to 42% solvent B in 44 min, 42 to 90% solvent B in 5 min followed by 90% B for 10 min.

Fifteen second fractions were collected onto a MTP 384 MALDI 800 μ mAnchorChip target (Bruker Daltonics) using a Proteiner Fraction Collector (Bruker Daltonics). The eluate from the analytical column was mixed online with CHCA matrix solution via a 500 mL Hamilton syringe (Upchurch) and syringe pump (Cole-Parmer, Illinois, USA) through a MicroTee (PEEK, 1/32 in, Upchurch Scientific). The matrix solution consisted of 748 μ L of 95% ACN in 0.1% TFA, 36 μ L of CHCA saturated in 90% ACN in 0.1% TFA, 8 μ L of 10% TFA and 8 μ L of 100 mM ammonium phosphate. One microliter of Bruker Daltonics Peptide Calibration Standard II (dissolved in 125 μ L of 0.1 % TFA) was mixed with 49 μ L of 30 % ACN in 0.1 % TFA and 150 μ L of matrix solution. Six hundred nL of this mixture was subsequently deposited onto calibration spots.

2.28.2 MALDI-TOF MS of HPLC Fractions

MS analysis was performed on an UltrafleXtreme MALDI-TOF/TOF system (Bruker Daltonics). Mass spectra were collected in reflectron positive ion mode, using WARP-LC (V1.2, Bruker Daltonics) interfaced with flexControl. Laser power was operator determined to provide optimal MS intensity and resolution. Three thousand laser shots were collected for both sample and calibration spots. Acquisition settings: *m/z* range of 700-4000, 5.0 \times reflector gain and 2.00 GS/s detector acquisition rate. Following MS collection, MS spectra were smoothed (Gaussian, 1 cycles, width - *m/z*

0.02), and baseline subtracted (TopHat). Peak masses and intensities were detected with flexAnalysis using the SNAP algorithm.

2.28.3 Tandem MS of HPLC Fractions

WARP-LC was used to calculate a compound list for automated MS/MS based on a preset LC-MALDI method. MS/MS instrument settings were defined by the LIFT method. Following automated MS/MS acquisition MS/MS spectra were exported to for MASCOT analysis (V2.3.02). The following search parameters were employed: taxonomy = *Homo sapiens*; protein database = Swiss-Prot; MS tolerance = 30 ppm; MS/MS tolerance = 0.8 Da; enzyme = Trypsin; missed cleavages = 2; dynamic modifications = oxidation; static modifications = carboamidomethylation. Matches were assigned with at least $p < 0.05$ significance threshold and an ion score > 25 . The molecular weight search and probability scores were used as the basis for ion annotation in the MS/MS spectra.

2.29 Data processing and visualisation of the distribution of known proteins in tissue sections

Methods used to process spectral data produced by IMS were as previously described by Gustafsson (Gustafsson 2012). Briefly, code was written in Visual Basic (flexAnalysis scripts) and Java (IonMapper) with the assistance of Mr. James Eddes (Adelaide Proteomics Centre). The acquired profile mass spectra from IMS outputs were processed by Gaussian smoothing (two cycles with a m/z width of 0.02) and baseline subtracted (TopHat) in flexAnalysis (V3.3, Bruker Daltonics, Bremen, Germany). Peaks were detected using the SNAP II algorithm and exported to ASCII text files using a custom flexAnalysis script. Each text file was saved with the name of its spectral coordinate to allow for spatially based data processing and image generation. A Java based script (NetBeans IDE6.9.1) was used to automate processing of the ASCII peak list files. A size ordered array of all the peaks in every peak list was generated first. The list was then grouped such that if the second peak in the combined peak list was within 0.2 Da from the first peak, then the two values were grouped together. If the difference was greater than the tolerance a new peak group was created. This process was completed for the entire combined peak list. The m/z

value for the peak groups was calculated as an Abundance Weighted Mean (AWM) using peak intensity. Ion intensity maps, based on signal to noise (S/N) values of peak group members, were then plotted if the mass peaks existed in >50 spectra.

The ion intensity maps were then aligned to the list of identified peptides by HPLC-nESI-LTQ Orbitrap MS. If the AWM of a peak group was within 0.2 Da of the mass peak identified by HPLC-nESI-LTQ Orbitrap MS, then they were paired together. In this way, both the distribution across a section and the peptide sequence, and therefore protein identity, were elucidated for each mass peak.

3 - Method Optimisation

3.1 Introduction

The studies described in this Chapter include optimisation of methods to analyse the genotype and phenotype of the hepatocyte population in the liver of patients with chronic HBV infection. These studies allowed us to elucidate the extent, cause and consequences of clonal proliferation of hepatocytes, and its potential impact on disease progression in the context of chronic HBV infection.

To achieve this aim, liver tissues from patients chronically infected with HBV were categorised by their disease state and characterised using a range of techniques. First, the extent of clonal proliferation of hepatocytes was determined using the invPCR technique. Next, changes in DNA methylation in HBV DNA promoter regions were analysed to determine whether CpG silencing of the genes expressing HBV antigens was associated with clonal proliferation. Furthermore, the protein expression of hepatocytes in these tissues was measured by mass spectrometry-based techniques to investigate if these could be linked to the observed clonal proliferation. The various methods developed in the course of this study can be broadly divided into two groups:

1. Genomic Methods

These include isolation and characterisation of extracted liver cell DNA from patients with chronic HBV infection.

Integrated HBV DNA was detected using invPCR and the unique virus-cell DNA junctions were used as a marker for individual hepatocyte lineages. The copy number of the unique virus-cell DNA junctions present in the individual extracts was then used to determine the extent to which hepatocytes had clonally proliferated during chronic HBV infection. Techniques were developed to measure changes in CpG methylation of HBV DNA promoters, which may downregulate HBV antigen expression and therefore give hepatocyte subpopulations within the liver a survival advantage against HBV-specific immune attack.

2. Proteomic Methods

These include experiments to determine if differences existed in patterns of cellular protein expression by comparing 1) cellular protein expression in sections of HBV-infected and control NHL and 2) different regions of hepatocytes in liver sections from the same HBV-infected patient. Using imaging mass-spectrometry (IMS), the distribution of a range of cellular proteins was visualised across a section of liver tissue and compared to the

distribution of HBV antigen expression as measured on an adjacent serial liver section. This allowed us to determine whether HBV antigen expression is associated with a change in expression of cellular proteins and whether these two phenomena were associated with clonal proliferation.

Some of the methods described in this Methods Optimisation Chapter were unsuccessful and were therefore not used in subsequent Chapters.

3.2 Results

3.2.1 Optimisation of Genomic Methods

3.2.1.1 Approximation of nucleus concentration

To approximate the yield of the DNA extracted from liver tissue sections, we sought to estimate the number of nuclei present in the sections of liver based on the area of the section. Furthermore, as outlined in Section 2.22, a formula was developed to calculate the 3-dimensional size of hepatocyte clones detected by invPCR analysis of DNA extracted from liver tissue sections. The assumptions in the calculation were that: 1) invPCR had a perfect detection rate; 2) the hepatocyte clone was spherical; and 3) the liver section in which the clone was detected was sampled through the centre of the clone. A missing quantity in the equation was the concentration of nuclei.

Therefore, the average concentration of nuclei was measured by light microscopy analysis of ethanol-fixed and H&E stained tissue sections, which were serial sections from the same liver in which hepatocyte clones were detected.

The number of nuclei in $49636 \mu\text{m}^2$ regions in a 20x field of magnification was enumerated in 5 separate fields of view for each of the 5 different liver sections. Fields of view containing at least 90% parenchymal tissue were selected at random for each section. Results from nuclei counts are summarised in Figure 3.1A. On average, a single nucleus was counted per $592 \mu\text{m}^2$. Since the section was $5 \mu\text{m}$ thick, this amounted to 0.344 counted nuclei per $1000 \mu\text{m}^3$.

Due to the thickness of the sectioned liver tissue and the expected size of nuclei, partial nuclei were also counted in this assay as shown in Figure 3.1B. The average diameter of each nucleus was measured using AnalySiS software to be $6.9 \mu\text{m}$, according to measurements of

A

Patient	Average nuclei count / 49636 μm^2	Average area/nucleus (μm^2)	Average nucleus diameter (μm)
Y2	85.6	581	6.24
Y3	77.8	633	6.39
Y4	107.6	459	7.22
Y5	81.4	610	7.18
Y6	73.6	676	7.56
Average	85.2	592	6.92

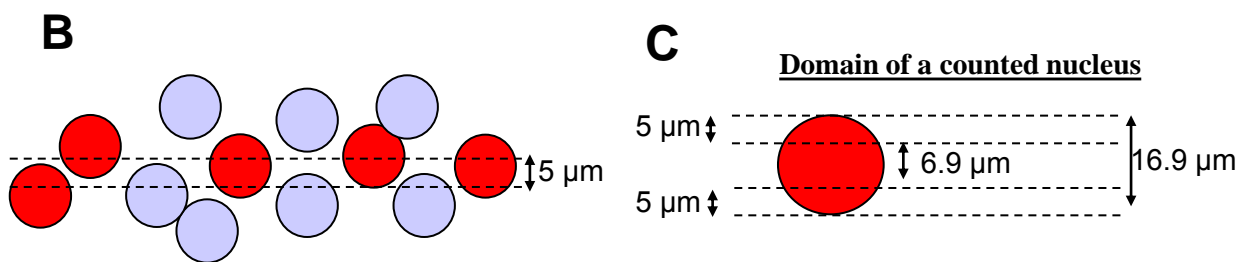


Figure 3.1. Calculation of average hepatocyte concentration by nuclei enumeration in ethanol-fixed tissues. Hepatocyte nuclei were counted in 5 tissues, averaged over 5 fields of view (Figure 3.1A). The diameter of the nuclei were also measured by AnalySiS software. Figure 3.1B shows that the nuclei that were counted within the 5 μm liver tissue section (red circles) were contained within the 5 μm liver section. Each nucleus could be detected by nuclei counts if it is sectioned within a 16.9 μm domain (3.1C).

125 nuclei (Figure 3.1A). As shown in Figure 3.1C, a nucleus is counted if the section is taken within 5 μm of the nucleus. A nucleus is therefore counted if a section cuts within 16.9 μm of its centre. On average, each 5 μm slice contains (5 μm /16.9 μm) or 29.6% of the nucleus. Thus, on average a whole nucleus was sampled every 592 μm^2 / 29.6%, i.e. a whole nucleus per 2000 μm^2 .

As a side-note, due to the spherical shape of a nucleus, a greater fraction of the nucleus would be sampled if the section taken is closer to the equator of the nucleus. However, during the extraction a large number of nuclei are sampled, so that the fraction of the nucleus sampled tends towards the average.

This calculation leads to an average nucleus concentration in the tissue of 0.102 whole nuclei/1000 μm^3 , or 1 whole nucleus per 9820 μm^3 . Assuming hepatocytes are spheres, the average intranuclear distance is therefore ~ 27 μm , which is consistent with previously reported hepatocyte diameter of 30-40 μm (Roskams, Desmet et al. 2007). Shrinkage of tissue during ethanol fixation may have led to a small decrease in the distance between hepatocyte nuclei.

3.2.1.2 Optimisation of DNA extraction from liver tissue

DNA extraction from liver tissues was central to the study. Isolation of the DNA enabled: determination of the HBV sequences associated with each patient and subsequent design of invPCR protocols; detection of hepatocyte clones based on virus-cell DNA junctions; and characterisation of phenotypic information, such as genetic mutations in the HBV DNA or changes in the methylation status of HBV DNA promoters. Extraction techniques were designed to isolate either nuclear or total DNA.

3.2.1.2.1 Total DNA extraction of 5 mg liver tissue fragments

Total DNA extraction was carried out on tissue from HBV-infected patients as described in Section 2.7. Briefly, 5 mg of snap-frozen liver tissue was placed in PK digestion solution. DNA-bound protein was digested by proteinase K (PK) with a 2hr incubation at 55°C, and the total nucleic acids were subsequently extracted twice with an equal volume of Tris-buffered phenol (pH 7.5) followed by a single phenol:chloroform:isopropyl alcohol (25:24:1) extraction. The nucleic acid in the aqueous phase was precipitated with two volumes of absolute ethanol and collected via centrifugation and then redissolved in distilled water. High

molecular weight DNA was visualised using agarose gel electrophoresis, as shown in Figure 3.2.

qPCR was performed on the extracted DNA samples as described in Section 2.13 to detect the copy numbers of the single copy cellular gene β - globin (BG). The primer sequences for this assay are listed in Table 2.1. The standard curves for this qPCR assay were generated by using standards of plasmids containing a single copy of BG, produced as described in Section 2.13.1. The standard curves, as shown in Figure 3.3A, indicate that BG qPCR assay was highly sensitive and could detect products to ≤ 10 copies per reaction. The melt curves for the BG qPCR assay shown in Figure 3.3B indicate that the BG qPCR assay specifically amplified only single PCR products.

The qPCR results from 10 HBV-infected patients are summarised in Table 3.1. Cellular DNA was present in most DNA extracts, except for patient NT. The copy numbers of BG observed in the DNA extracts were equivalent to a range of 0.3-1.5 mg of liver. qPCR readings were consistently lower than expected given the approximate weight of the tissues. This could potentially be due to contamination of the DNA extracts with an inhibitor of the qPCR polymerase. Alternatively, the copy number of plasmids used to produce the standard curve could have been greater than expected.

3.2.1.2.2 Total DNA extraction of EtOH-fixed paraffin wax-embedded liver tissue sections

Additionally, total DNA was extracted from 5 μm paraffin wax-embedded ethanol-fixed liver tissue sections, as described in Section 2.8. Briefly, two 5 μm -thick liver tissue sections were placed in a 1.5 mL tube and washed twice with xylene in order to remove the paraffin wax. The xylene was removed by two washes with absolute ethanol. After vacuum drying, total DNA was extracted from the liver tissue as summarised in Section 3.2.1.2.1, except with only a single extraction with phenol:chloroform:isopropyl alcohol (25:24:1).

Examples of the tissues section from which total DNA was extracted are shown in Figure 3.4. The area of the tissues was measured by AnalySiS image software. The approximate number of nuclei in each section was estimated by dividing the area of section by the 2000 μm^2 , the average area per whole nucleus as observed in Section 3.2.1.1. The calculated nuclei counts for each tissue section are summarised in Table 3.2.

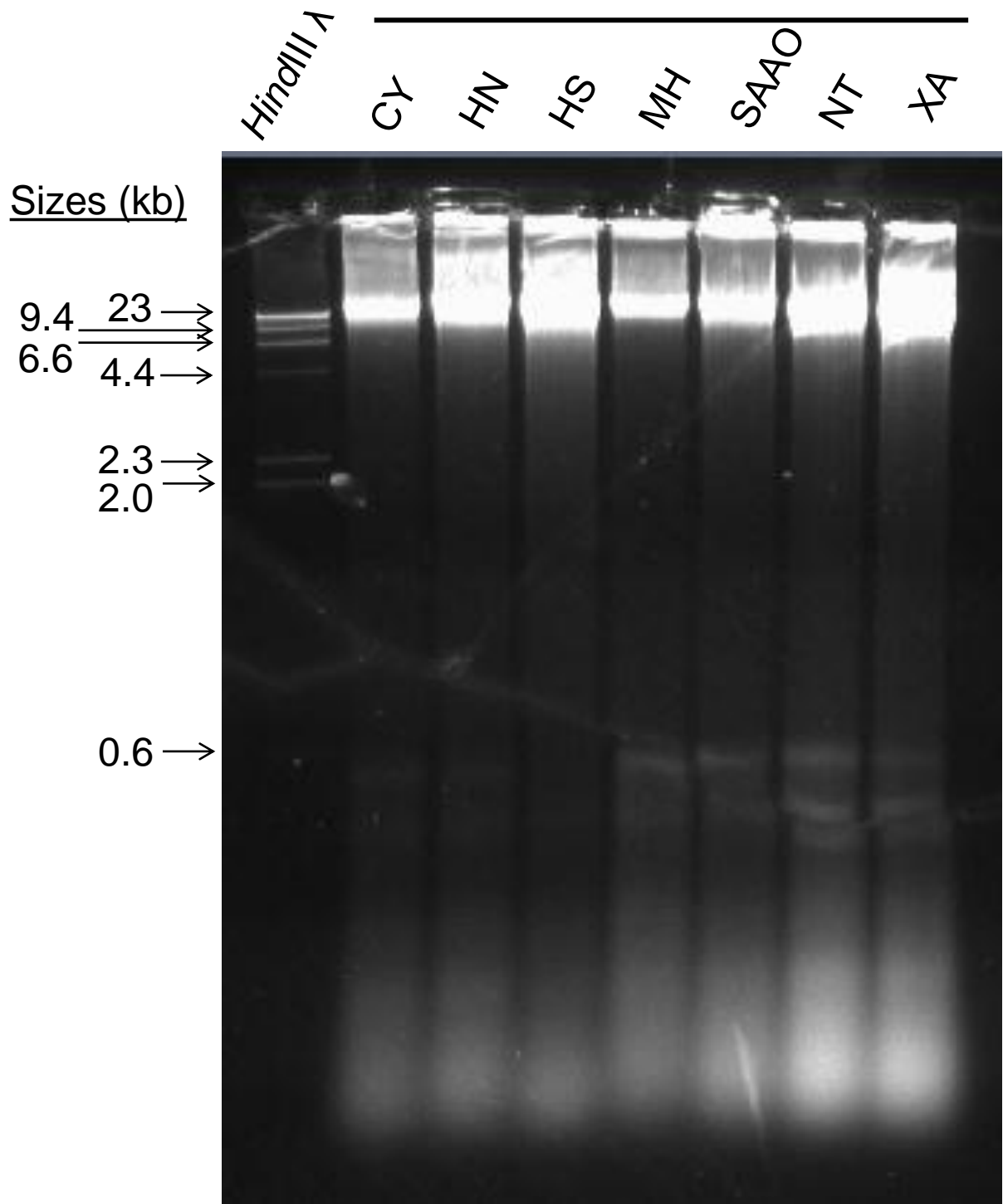
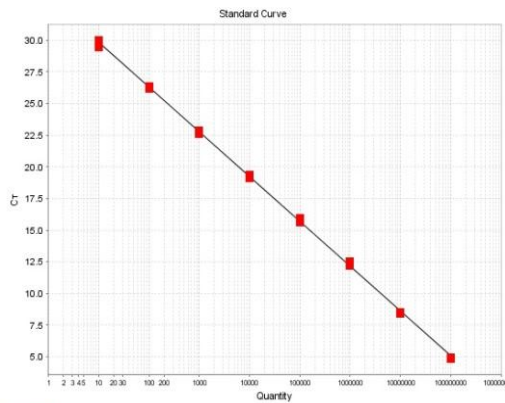
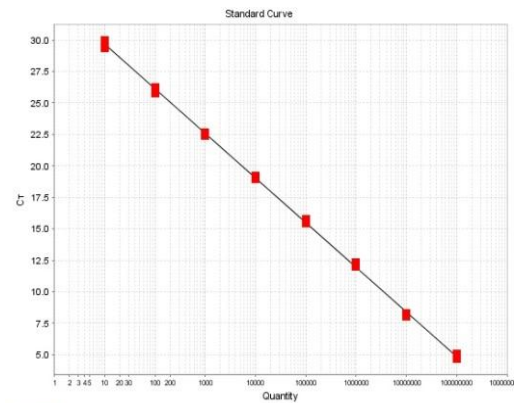


Figure 3.2. High molecular weight DNA was observed in gel electrophoresis of total DNA extraction of snap-frozen HBV-infected tissues. DNA was extracted from ~5 mg snap-frozen liver tissue as described in Section 2.7. Approximately 1 μ g of total DNA extract (quantified by optical densitometry) was run in each lane. One μ g of lambda phage DNA restricted with the restriction enzyme *HindIII* was used as a size marker.

A $R^2 > 0.999$ 

Efficiency >91%

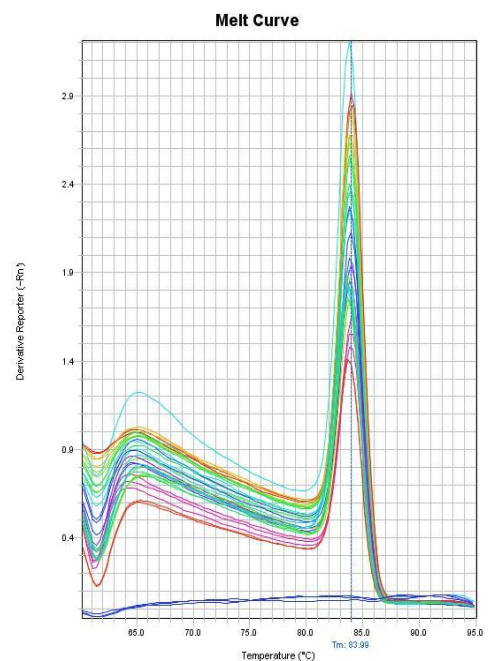
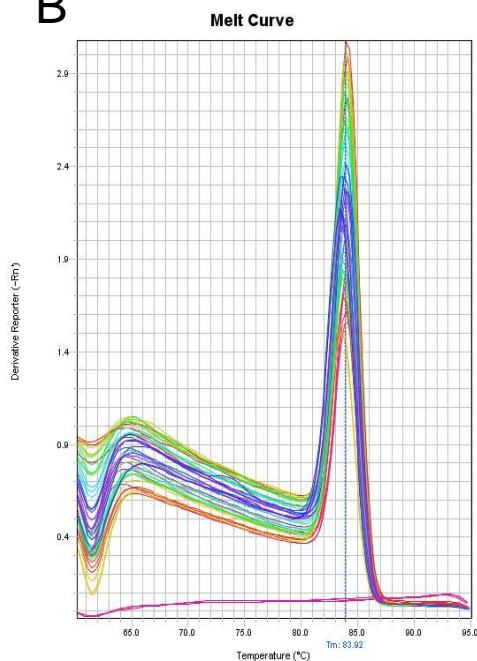
B $T_m = 84^\circ\text{C}$

Figure 3.3. Standard curves and melting curves for qPCR detecting the β -globin gene (BG) in total DNA extracts from liver tissue. The sequences of the primers used for amplifying BG are listed in Table 2.1. Due to the number of samples, two separate qPCR assays were run as described in Section 2.13. Each DNA extract was run in triplicate. The BG gene was subcloned into a pBlueBac vector as described in Section 2.13.1 to produce plasmid standard curves. Plasmid standards were diluted from 10^1 - 10^8 copies per reaction and quantified by qPCR as described in Section 2.13. As shown by the standard curve in Figure 3.3A, products were detectable to <10 copies per reaction. The melt curves in Figure 3.3B shows that only single products were amplified in all samples.

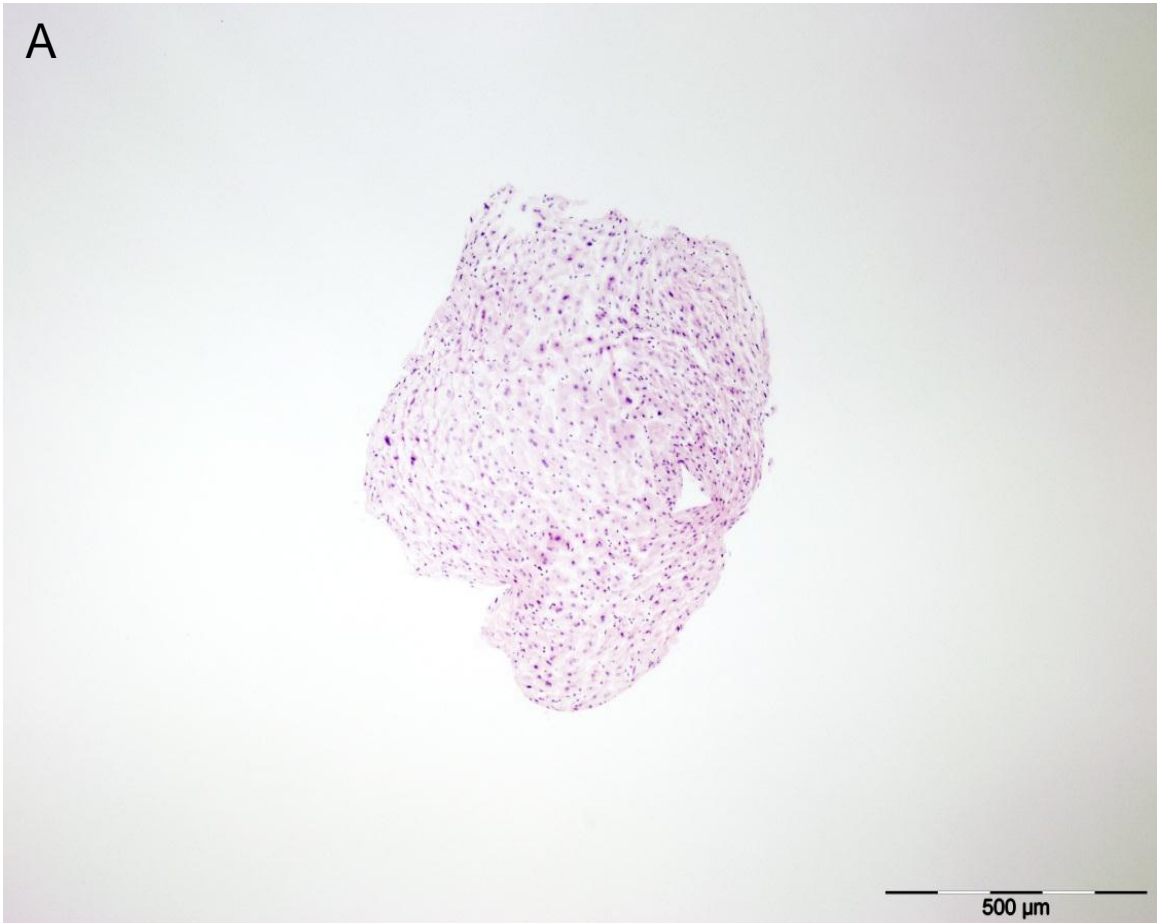
Table 3.1. qPCR analysis of the BG gene in DNA extracts from ~5 mg liver samples from patients with chronic HBV infection.

Patient	BG/extract of ~5 mg liver tissue (copies x 10⁵)¹	Equivalent in liver mass based on DNA content (mg)²
CN	2.2	0.8
CY	1.2	0.43
CYRY	2.1	0.7
HN	2.5	0.9
HS	4.2	1.5
MH	0.8	0.3
NT	Undetectable	
SAAO	2.6	0.9
WN	3.7	1.3
XA	2.9	1.1

¹ BG = BG. Copy number shown is the average calculated total copy number of BG in each extract.

² Assuming, i) $\sim 1.4 \times 10^5$ cells per mg of liver tissue (Sohlenius-Sternbeck 2006), ii) 6 pg of DNA per diploid human cell (NCBI, 2011), and iii) 2 copies of BG/cell. (Lawn, 1978)

A



B

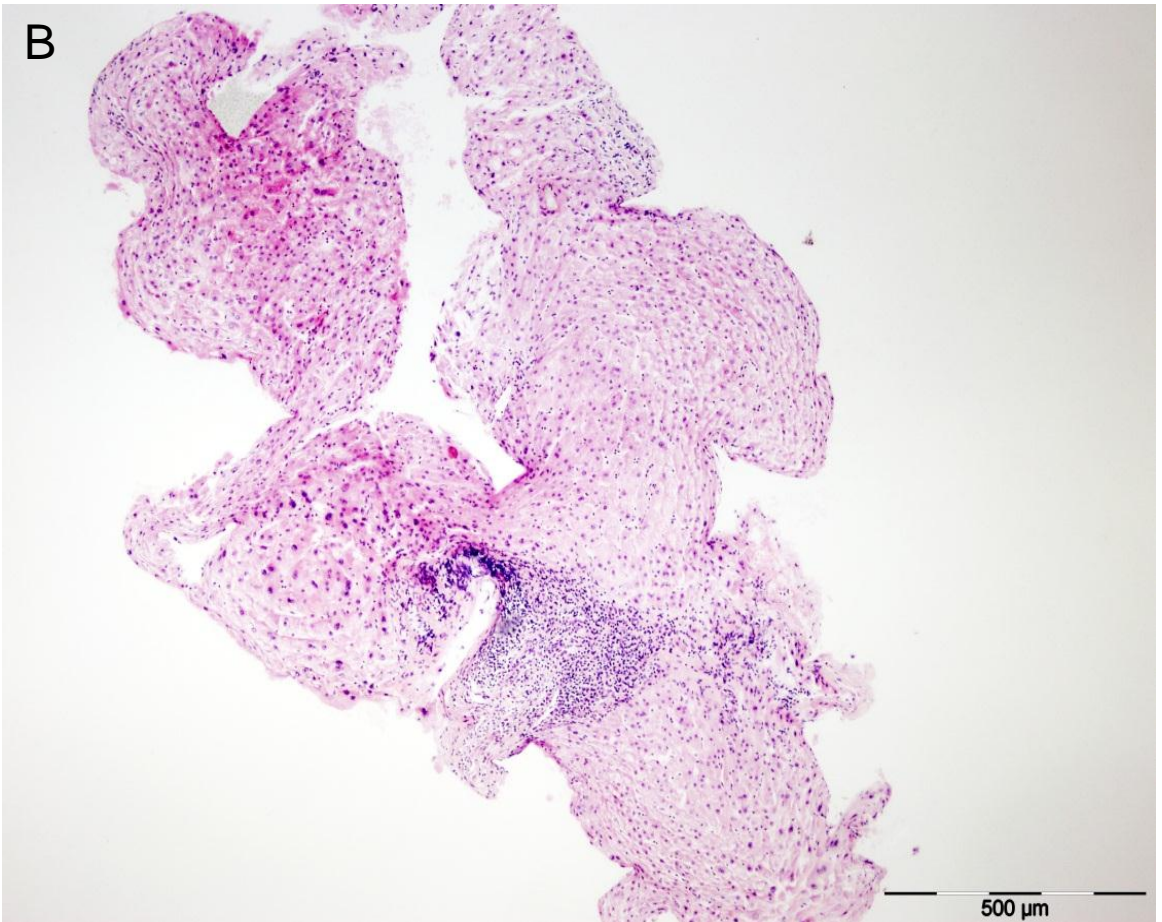


Figure 3.4. Total DNA was extracted from liver tissue sections of chronic HBV patients C (3.4A) and L (3.4B). Total DNA was extracted as described in Section 2.8 from unstained 5 μm -thick tissue sections that were serially sectioned from those seen above. These tissues sections were stained with H&E for visualisation. Magnification = 4x. Scale bar = 500 μm .

Table 3.2. BG qPCR showed that high levels of genomic DNA were isolated from liver tissue sections.

Patient¹	Approximate nuclei count²	Mean BG copies per extract³
C1	514	1650
C2	375	489
L1	1750	1750
L2	1990	3160
Total	4640	7049

¹ Two biological replicates for each tissue sample was analysed by qPCR.

² The approximate number of nuclei in DNA extract was estimated by dividing the area of 2 sections by the 2000 μm^2 , the average area per whole nucleus as observed in Section 3.2.1.1.

³ BG = BG. qPCR assays were run in duplicate for the liver DNA extract of each tissue section.

Additionally, the yield from the DNA extraction was analysed by qPCR. One microlitre of each 20 μ L DNA extract was analysed by qPCR to determine the copy number of BG as described above and in Section 2.13. Consistent with previous qPCR assays, the standard and melt curves shown in Figure 3.5 indicate a sensitivity of ≤ 10 copies per reaction and confirm that a single product was amplified. As shown in Table 3.2, the copy numbers of BG detected by qPCR in the DNA extracts were approximately double the calculated number of hepatocyte nuclei in each section. Since each nucleus should contain 2 copies of BG (Lawn, Fritsch et al. 1978), this suggested that the yield for DNA extraction of paraffin wax-embedded liver tissue was efficient.

3.2.1.2.3 Nuclear DNA extraction

Detection of virus-cell DNA junctions requires only chromosomal DNA. As there are significant amounts of viral RI DNA present in the cytoplasm of infected cells, invPCR assays may be inhibited or produce false positive virus-virus DNA artefacts. Therefore, we sought to isolate nuclear DNA but to exclude cytoplasmic DNA by extracting DNA from nuclei, as previously described in Section 2.9 (originally optimised by Zhang *et al.* (2003)).

Briefly, 5 mg of liver tissue was homogenised in 1 mL of homogenisation solution containing 250 mM sucrose and 0.1% Triton X-100 detergent. Nuclei were pelleted and washed in the homogenisation solution. The washed nuclei were lysed by PK digestion and extracted with phenol:chloroform:isopropyl alcohol (25:24:1). Extracted nucleic acids were ethanol-precipitated and redissolved in distilled water. The quantity and size range of the extracted DNA was analysed by gel electrophoresis.

Initially, pilot studies were carried out on NHL tissue. When washed nuclei were stained with DAPI and a wet slide mount was prepared, no intact nuclei were observed. A fifth of the extracted DNA was analysed by gel electrophoresis: however, no high molecular weight DNA could be observed. This suggested that this extraction technique had a low yield in our hands. Therefore, this technique was not used in subsequent studies.

3.2.1.2.4 Total DNA extraction of hepatocyte foci isolated by laser-microdissection

Finally, total DNA was extracted from hepatocyte foci isolated by laser-microdissection of 5 μ m paraffin-embedded ethanol-fixed liver tissue sections, as described in Section 2.21. Briefly, the 5 μ m-thick sections were mounted on polyethylene naphthalate membrane-coated glass slides (Leica). Sections were washed twice with xylene in order to remove the paraffin

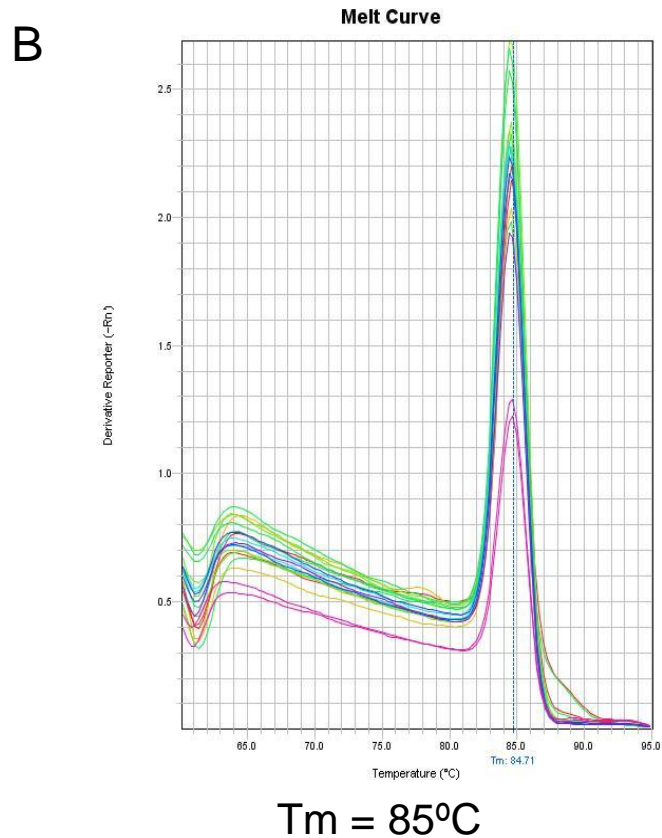
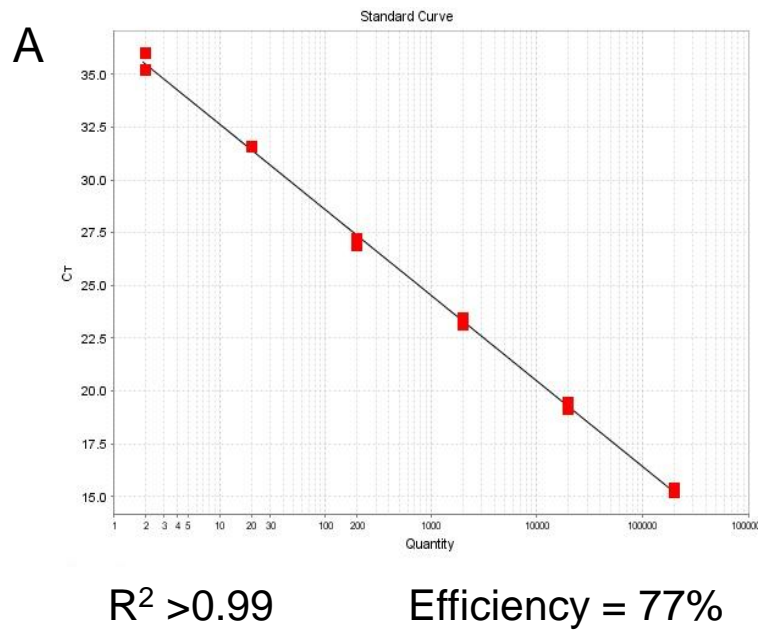


Figure 3.5. Standard and melting curves for qPCR detecting BG in total DNA extracts of 5 μm -thick liver tissue sections. The sequences of the primers used for amplifying BG are listed in Table 2.1. Each DNA extract was run in duplicate. The BG gene was subcloned into a pBlueBac vector as described in Section 2.13.1 to produce a plasmid standard curve. Plasmid standards were diluted from 2 to 2×10^5 copies per reaction and quantified by qPCR as described in Section 2.13. As shown by the standard curve in Figure 3.5A, products were detectable to < 2 copies per reaction. The melt curve in Figure 3.5B shows that only single products were amplified in all samples.

wax, and then twice with absolute ethanol to remove the xylene. After vacuum drying, hepatocyte foci were isolated by laser-microdissection and total DNA was extracted from the liver tissue by PK digestion, as summarised in Section 3.2.1.1.2 except with the addition of 1 mg of carrier RNA and 500 µg of lambda phage DNA as carrier.

The hepatocyte foci that were isolated by laser-microdissection were assigned unique identifiers and are shown in Figure 7.1. The area of each isolated focus was measured by ImageJ software. The approximate number of nuclei in each section was estimated by dividing the area of the focus by the 2000 µm², the average area per whole nucleus as observed in Section 3.2.1.1. The calculated nuclei counts per hepatocyte focus are summarised in Table 3.3.

qPCR for BG was carried out on 1 µL of each 20 µL DNA extract as described in Section 2.13. The standard and melt curves shown in Figure 3.6 indicate a sensitivity of 2 copies per reaction and specificity for the amplification of a single PCR product. The copy number of BG in each DNA extract are summarised in Table 3.3. For the majority of isolated hepatocyte foci, the copy number of BG detected by qPCR was almost double the calculated nuclei per hepatocyte focus. Since each nucleus should contain 2 copies of BG (Lawn, Fritsch et al. 1978), this suggested that the extraction from hepatocyte foci isolated by laser-microdissection is efficient.

3.2.1.3 Optimisation of semi-degenerate oligonucleotide primer (semi-DOP) PCR

Semi-DOP PCR has previously been used to sequence unknown virus sequences isolated from faecal samples (Arthur, Higgins et al. 2009). The random priming nature of DOP made it theoretically possible to use semi-DOP PCR to detect virus-cell DNA junctions. The semi-DOP PCR protocol is outlined in Section 2.17 and Figure 2.9. This protocol was optimised in collaboration with Drs. Jane Arthur and Rodney Ratcliff (Infectious Diseases Laboratories, SA Pathology).

Briefly, PCR amplification containing a biotinylated HBV-specific primer and a degenerate oligonucleotide primer (DOP) was carried out using total liver DNA extract as a template. The sequences of the primers used are shown in Table 2.6. The initial 10 cycles of the PCR were conducted with a low annealing temperature, allowing the DOPs to bind to the host cell genome based on a 3' 4-6 nt anchor region. This was followed by 30 cycles with a high annealing temperature to encourage specific amplification of random products. HBV-specific

Table 3.3. BG qPCR showed that high levels of genomic DNA were isolated from hepatocyte foci isolated by laser-microdissection.

Focus ID¹	Approximate nuclei count²	Mean BG copies per extract³
Y2 1n	147	254
Y2 10n	299	456
Y2 3p	132	75
Y2 4p	265	446
Y3 15n	477	934
Y3 17n	317	504
Y3 26n	1040	2070
Y3 13p	220	575
Y3 7p	522	63
Y3 8p	522	102
Y5 1n	195	775
Y5 5n	319	320
Y5 1p	610	2023
Y5 5p	355	1257
Y6 7n	185	239
Y6 7p	186	214
Average	5790	10304

¹ Unique ID were given to each of the hepatocyte foci that were excised. The hepatocyte foci that each foci ID was assigned is shown in Figure 7.1.

² The approximate number of nuclei in DNA extract was estimated by dividing the area of section by the 2000 μm^2 , the average area per whole nucleus as observed in Section 3.2.1.1.

³ qPCR assays were run in duplicate for the liver DNA extract of each tissue section.

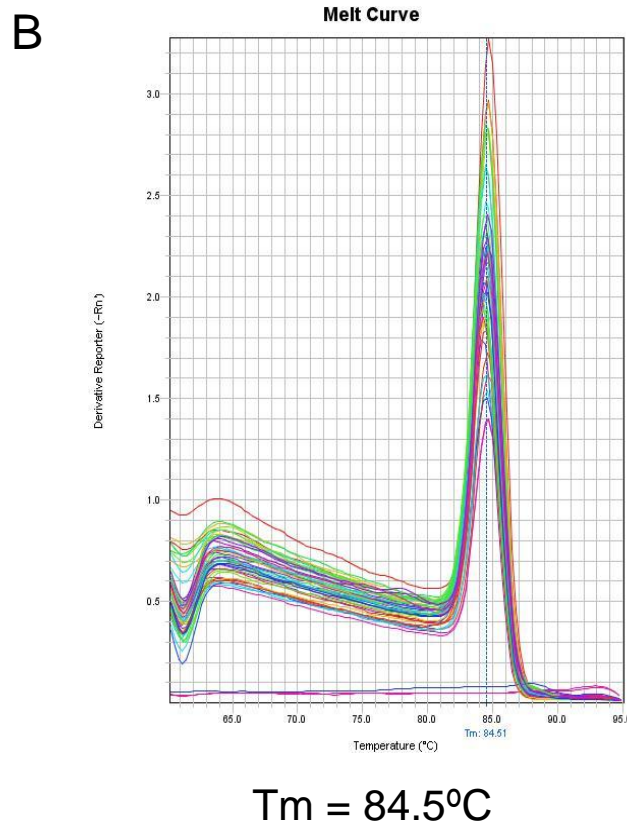
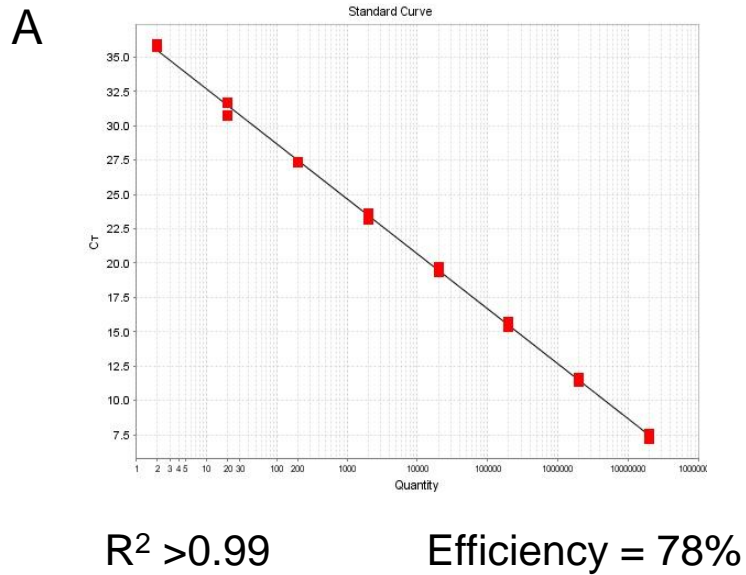


Figure 3.6. Standard and melting curves for qPCR detecting BG in total DNA extracts of hepatocytes isolated by laser-microdissection. The sequences of the primers used for amplifying BG are listed in Table 2.1. Each DNA extract was run in triplicate. The BG gene was subcloned into a pBlueBac vector as described in Section 2.13.1 to produce a plasmid standard curve. Plasmid standards were diluted from 10^1 - 10^8 copies per reaction and quantified by qPCR as described in Section 2.13. As shown by the standard curve in Figure 3.6A, products were detectable to <10 copies per reaction. The melt curve in Figure 3.6B shows that only single products were amplified in all samples.

products were isolated using streptavidin-coated magnetic beads. Products were amplified by a second round of PCR using nested primers to amplify the DNA attached to the beads as a template. Amplified products were isolated and DNA sequenced. For optimisation, total liver DNA from NHL was spiked with 1000 and 10,000 copies of a HBV-containing plasmid (pBB1.3HBV, plasmid map shown in Figure 2.1).

Upon gel electrophoresis as shown in Figure 3.7, multiple DNA bands could be visualised in both assays, including a 563 bp band that matched the expected PCR product of the HBV-plasmid junction in pBB1.3HBV. However, in negative controls where the total DNA extracted from NHL was not spiked with pBB1.3HBV DNA, the same 563 bp band was observed, suggesting that the product was an artefact of non-specific amplification. Furthermore, when the 563 bp band was isolated and used as a template for DNA sequencing with a third primer specific for the HBV DNA, no signal was observed. The other DOPs listed in Table 2.6 were evaluated with identical results. Therefore, semi-DOP PCR was not used in subsequent studies.

3.2.1.4 Optimisation of invPCR

InvPCR has previously been used to detect virus-cell DNA junctions in hepadnavirus-infected liver tissues from ducks, woodchucks, chimpanzees and humans (Yang and Summers 1999; Summers and Mason 2004; Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). Hepatocyte clones were detected by the identification of repeated virus-cell DNA junctions. Thus, we aimed to optimise this technique to determine the extent of clonal proliferation within HBV-infected human liver tissue.

A woodchuck DNA extract previously found to contain virus-cell DNA junctions was used to confirm that we could successfully apply the invPCR technique. The DNA extracted from a ~25 mg liver fragment of a chronically WHV-infected woodchuck was a generous gift from Dr. William Mason (FCCC). The woodchuck designated WC366 had been previously shown to contain detectable hepatocyte clones by invPCR (Mason, Jilbert et al. 2005).

The invPCR protocol was used as outlined by Mason *et al.* (2005). The following restriction enzymes (RE) were used in the inversion reaction of invPCR:

1st cut - *SacI*

2nd cut - *HindIII*

3rd cut - *AflIII*

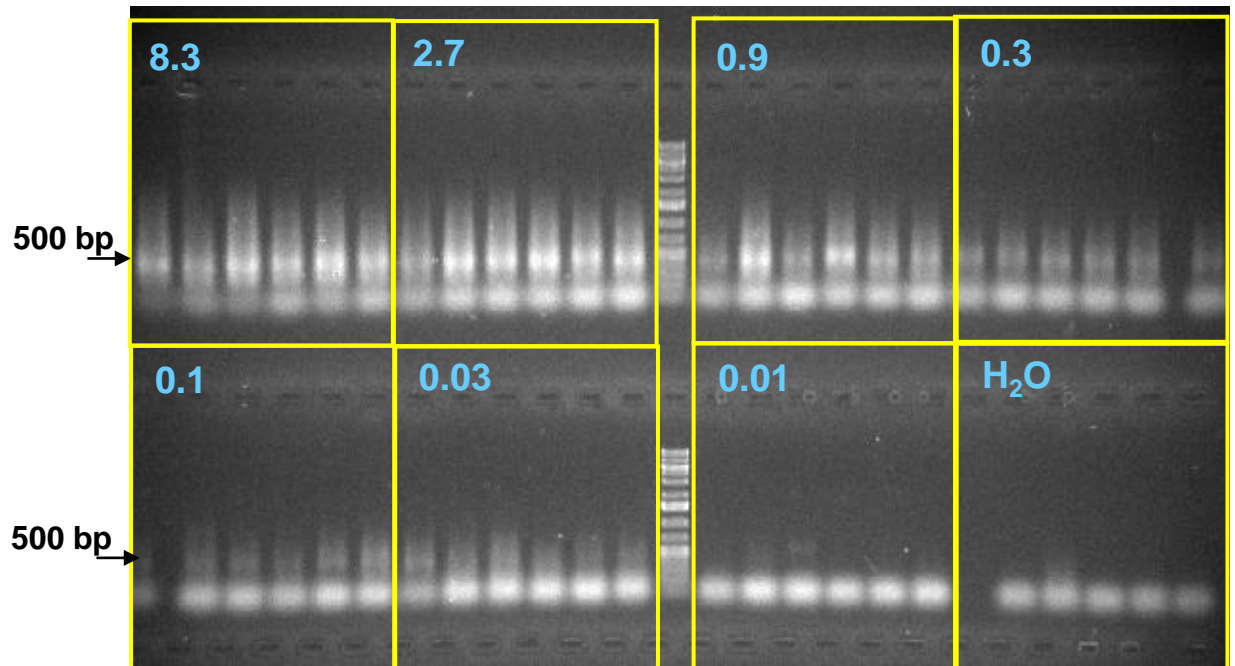
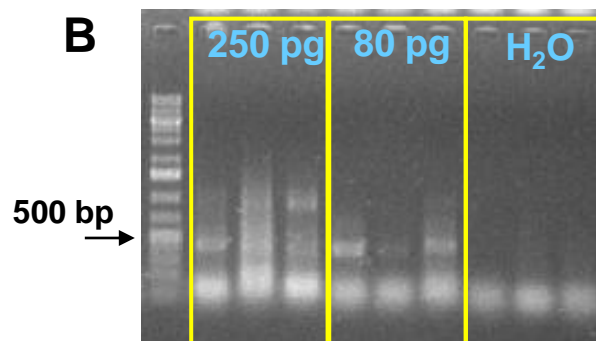
A**B**

Figure 3.7. Semi-DOP PCR detected a ~500 nt product in both plasmid-spiked (3.7A) and negative-control reactions (3.7B). One microgram of DNA extracted from uninfected human tissue was spiked with pBB4.5HBV1.3 and serially diluted before analysed by Semi-DOP PCR, as described in Section 2.17 (3.7A). The light blue numbers represent the approximate copy numbers of plasmid per reaction. The theoretical copy number of plasmids per reaction for each serial dilution is shown. Products of ~500 nt are seen in even very dilute samples. These products were also seen with low concentrations (250 pg and 80 pg) of genomic DNA without plasmid DNA (3.7B). The light blue numbers represent the mass of NHL DNA per reaction. This suggests that the product is a false positive result from amplification of the cellular DNA sequences.

The primers used are outlined in Table 2.4.

The resultant products, as shown in Figure 3.8, were isolated and DNA sequenced as outlined in Section 2.12. Resultant DNA sequences were aligned using analysis with BLAST (NCBI) to the previously sequenced WHV genome, Genbank accession #J02442 (Galibert, Chen et al. 1982). Twelve unique virus-cell DNA junctions were confirmed by sequence analysis (Results are summarised in Table 3.4).

The ends of the integrated WHV DNA were mapped with respect to the WHV sequence in Figure 3.9. The majority of integrations clustered around nt 1932, the expected left hand end of the major dsDNA form of WHV DNA. This data is consistent with previous studies of DNA extracts from this animal and in WHV-infected woodchucks in general (Mason, Jilbert et al. 2005).

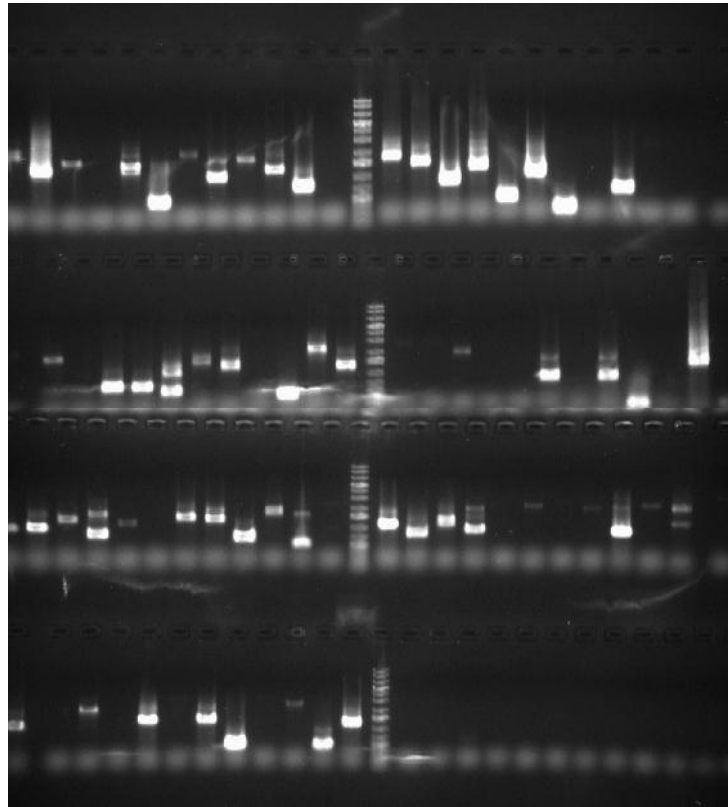
3.2.1.5 Gel purification of high molecular weight DNA using ultrasonification or partial RE digestion

Some DNA samples subjected to invPCR and subsequent DNA sequencing revealed high levels of false-positive virus-virus DNA junctions. It was proposed that these junctions arose due to cccDNA mutants or interactions between HBV RI DNA and cellular DNA during the ligation step. Thus, gel purification was used to potentially isolate high molecular weight chromosomal DNA from the lower molecular weight HBV DNA.

Agarose gel electrophoresis of 1 μ g of total liver DNA extract and isolation of the high molecular weight band for subsequent invPCR still amplified false positive virus-virus DNA junctions. This suggested that HBV replicative intermediates and cccDNA were being trapped within the high molecular weight band during gel electrophoresis. Therefore, the strands of host genomic DNA were fragmented by ultrasonification or partial RE digestion to allow more thorough separation from lower molecular weight HBV DNA.

Ultrasonification of DNA extracts was carried out in an ice bath with intervals of 10 sec “on”/ 10 sec “off” to prevent denaturation by heat. Samples that had been subjected to ultrasonification for <60 sec (i.e. 6 “on” intervals) produced no observable change in migration rate by gel electrophoresis, as shown in Figure 3.10A. However, ultrasonification for ≥ 70 sec led to a large range of fragments, resulting in a smear of DNA during gel electrophoresis.

A



B

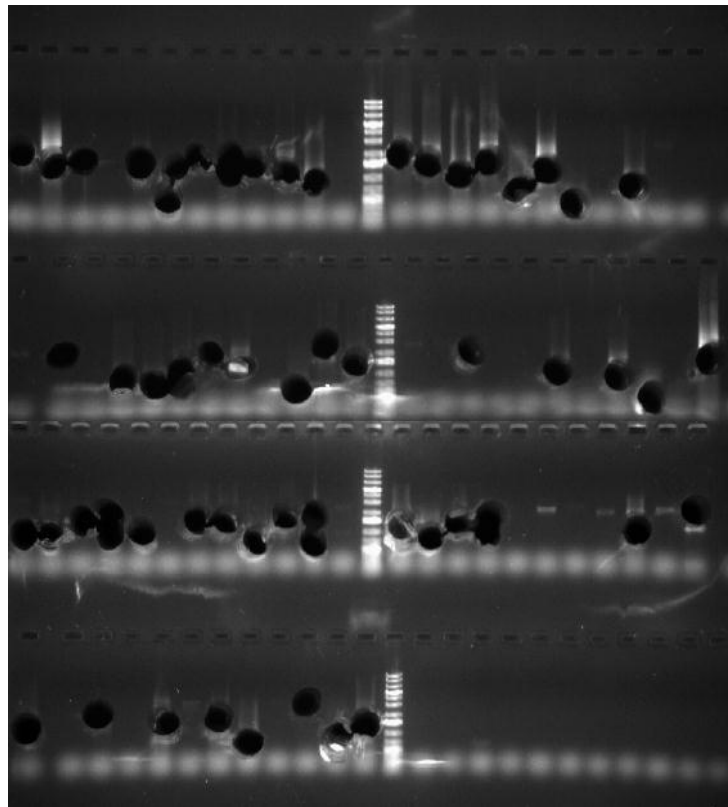


Figure 3.8. DNA products amplified by the *SacI* invPCR design in DNA extracted from a 25 mg woodchuck liver tissue fragment. DNA products were amplified by invPCR as described in Section 2.14 using liver DNA extracted from woodchuck WC366 (previously analysed in Mason *et al.* (2005)) as a template (3.8A) and isolated by cutting out from the agarose gel (3.8B).

Table 3.4. Virus-cell DNA junctions detected by the *Sac*I invPCR.

Clone ID	Clone size ¹	Junction wrt. WHV sequence (nt)
WC366 i	64	1935
WC366 ii	32	1921
WC366 iii	32	1932
WC366 iv	32	1937
WC366 ix	32	2008
WC366 v	32	1943
WC366 vi	32	1944
WC366 vii	32	1949
WC366 viii	32	1979
WC366 x	32	2041
WC366 xi	32	2078
WC366 xii	32	Unknown
WC366 xiii	32	Unknown
WC366 xiv	32	Unknown
WC366 xv	32	Unknown
WC366 xvi	64	2039

¹ Clone size as determined by invPCR and end-point dilution.

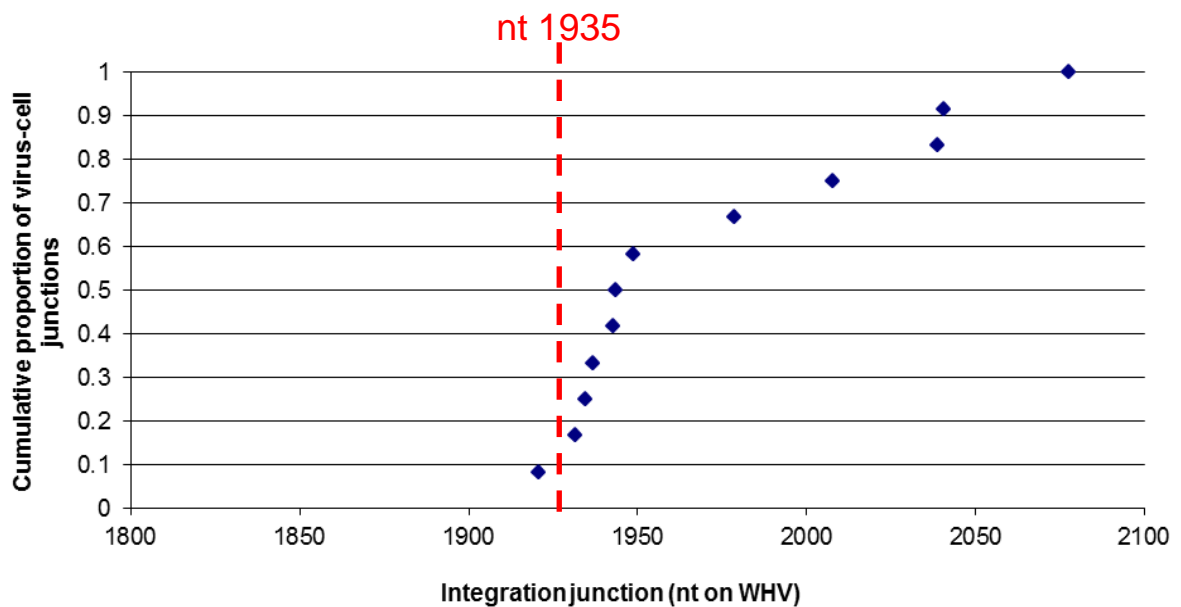


Figure 3.9. Integration sites of WHV DNA into the host cell chromosome matches the expected left hand end of the double-stranded linear DNA. The cumulative probability of integration was plotted against the site of integration with respect to the WHV DNA sequence as outlined by Galibert *et al.* (1982) (Genbank Accession #J02442). The probability of integration site rose and clustered before the expected left hand end of the double-stranded linear DNA form of WHV located at nt 1935 (Summers *et al.* 2003).

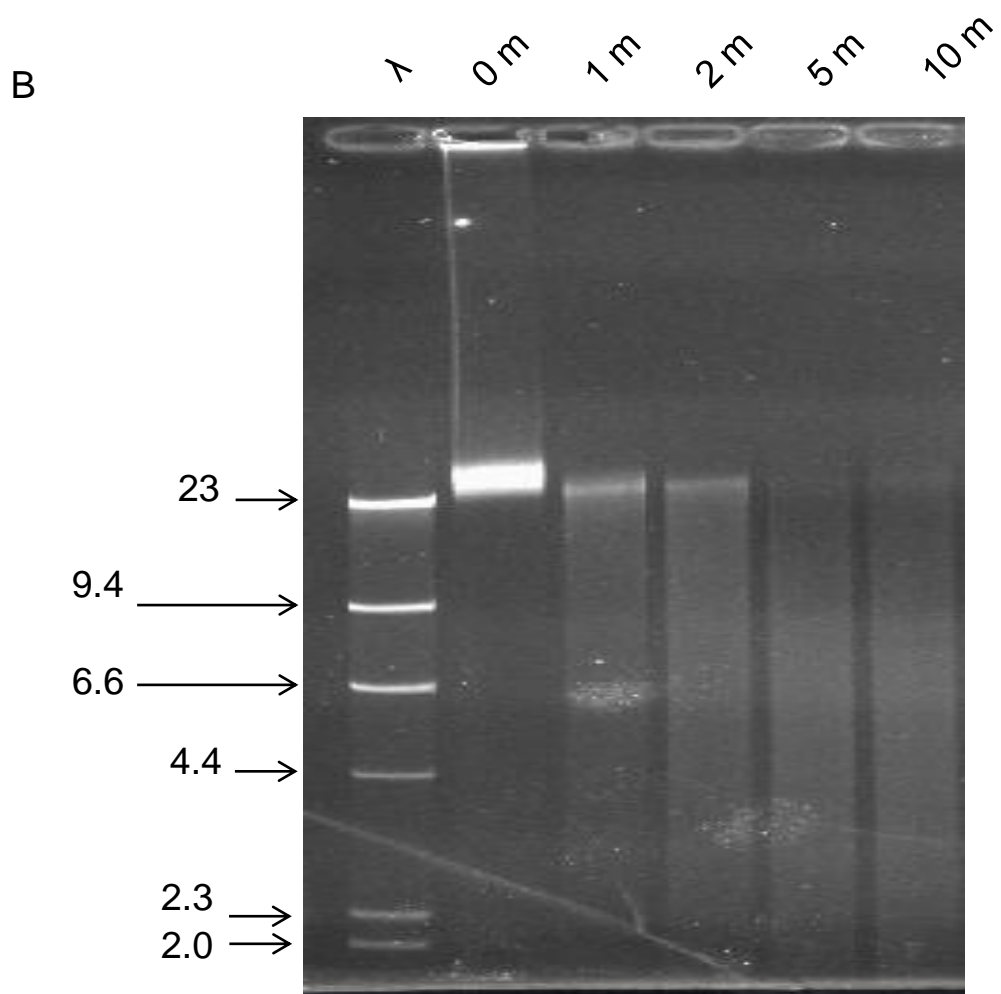
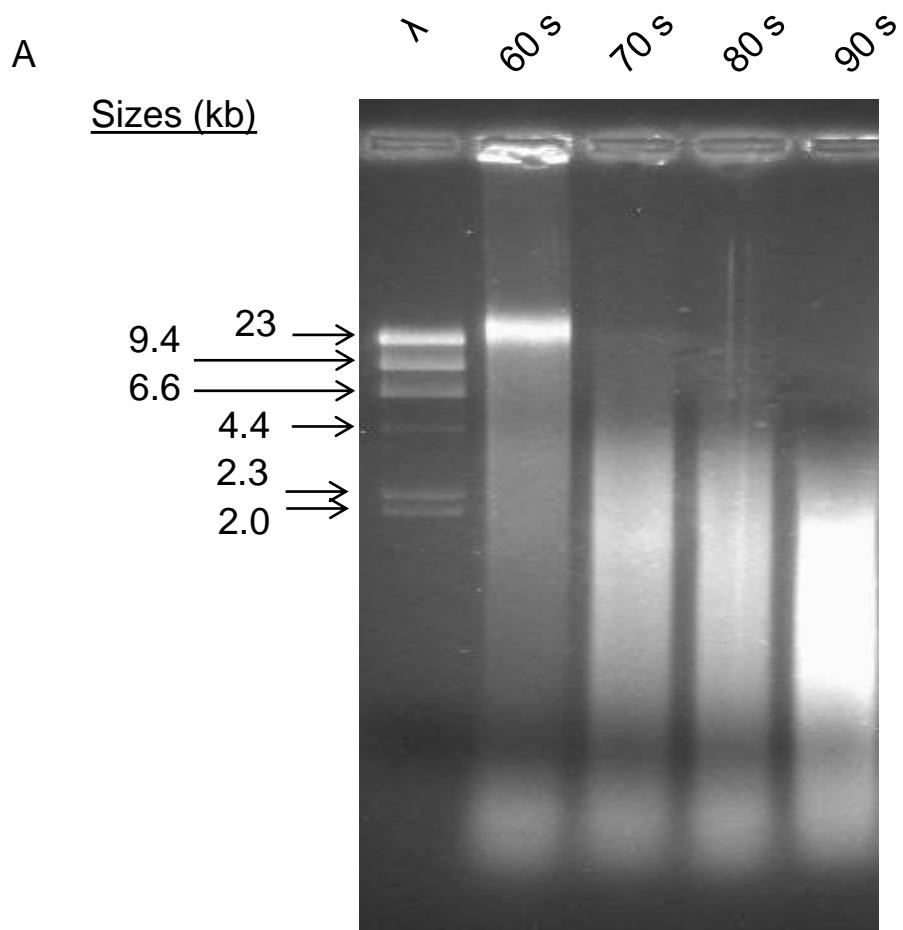


Figure 3.10. Attempts to increase separation of HBV replicative intermediates from high-molecular weight cellular DNA by ultrasonication (3.10A) or by partial enzyme digestion with *Bsa*WI (3.10B).

Total DNA extracts were exposed to an ultrasonication via a Bioruptor sonicator (Diagenode) on low settings for 60, 70, 80 and 90 sec, as described in Section 3.2.1.5. DNA extracts were then separated by gel electrophoresis (3.10A) showing that the amount of high molecular weight DNA very suddenly decreased between 60 and 70 sec. The range of sizes produced by 70 sec of exposure to ultrasonication was too large to separate genomic DNA from HBV DNA replicative intermediates.

Total DNA extracts were partially digested with 0.05 U of *Bsa*WI for 0, 1, 2 and 10 min, as described in Section 3.2.1.5. DNA extracts were then separated by gel electrophoresis (3.10B). Progressive decrease in size and intensity was seen in the high molecular weight DNA after greater digestion times, as expected. High molecular weight was thus isolated by gel electrophoresis after 1 min of digestion with *Bsa*WI and subjected to invPCR in the following studies.

Partial RE digestion was carried out with the addition of 0.5U of *Hind*III to DNA extracts and incubation at 37°C. Digestion was halted at intervals of 1, 2, 5 and 10 min with addition of 1 µL of 500 mM EDTA. As shown in Figure 3.10B, partial enzyme digestion with *Hind*III gave similar results with smears of DNA observed with ≥ 1 min of treatment.

For both treatments, when the smears of high molecular weight DNA were isolated from the agarose gels and then subjected to invPCR, false positive virus-virus DNA junctions were still amplified. Due to the detection of virus-cell DNA junctions in other liver samples, further efforts in the isolation of high molecular weight DNA were not pursued.

3.2.1.6 Human invPCR design and efficiency of the inversion reaction

A control experiment was designed to detect the inversion of the single copy cellular gene procollagenase α 1 (PCA1), as described in Section 2.14. This assay was used to quantify the efficiency of the circularisation reaction in the *Nco*I invPCR assay, as summarised in Table 2.4B. This information would improve the estimate of the clone sizes calculated from the number of virus-cell DNA junctions detected by invPCR.

The Genbank sequence AF017178 (Chu, de Wet et al. 1985) was used to design nested primers to amplify a ~400 bp product either from a PCA1 DNA sequence regardless of inversion or an inverted PCA1 DNA sequence, as shown in Figure 2.5. Copy numbers of each product were measured by diluting inverted DNA template into a 96-well nested PCR. The sequence of both inverted and uninverted products was confirmed. Copy numbers of inverted and uninverted DNA products were calculated using Most Probable Number (i.e. binomial distribution of zeroes) statistical analysis via an MPN calculator developed by Curiale (2004). The ratio of the copy numbers of inverted and uninverted products was defined as the inversion efficiency.

No inverted product was amplified in DNA extracts that had not been subjected to an inversion reaction. This strongly suggests that the amplified inverted products are not false-positive products from the invPCR assay.

The inversion efficiency of *Nco*I invPCR reactions was measured in duplicate DNA extracts from 10 different patients, as summarised in Table 3.5. The average calculated inversion efficiency for the 20 samples was 26%.

Table 3.5. Determining the efficiency of the invPCR inversion reaction by inverting a single-copy cell gene procollagenase 1 α (PCA1) as a template.

Extract	Efficiency (MPN%)¹	Upper 95% CI (%)²	Lower 95% CI (%)³
CN1	37	82	18
CN2	47	88	26
CY2	27	58	12
CY3	27	52	15
CYRY1	22	41	12
CYRY2	6	17	3
HN1	26	49	15
HN2	33	61	18
HS2	29	56	18
HS3	48	100	22
MH2	15	35	6
NT1	10	24	4
NT2	25	45	14
SAAO1	37	90	16
SAAO2	49	100	18
WN1	22	54	9
WN2	21	80	6
XA1	4	24	0.8
XA2	15	31	8
Average	26	57	13

¹ Inverted DNA was distributed on two 96-well plates and PCR amplified with nested primers specific for either an inverted region or an uninverted region of PCA1, as described in Section 2.14. Inversion efficiency was defined as the most probable number as calculated by MPN calculator (Curiale, 2000) of inverted products divided by the MPN of inverted products.

² The upper 95% confidence interval was calculated by dividing the upper 95% confidence interval of the uninverted MPN by the lower 95% confidence interval of the inverted MPN.

³ The lower 95% confidence interval was calculated by dividing the lower 95% confidence interval of the uninverted MPN by the upper 95% confidence interval of the inverted MPN.

Upper and lower 95% confidence intervals were calculated for the copy numbers of the inverted and uninverted products using the MPN calculator. To calculate the upper 95% confidence intervals for inversion efficiency, the upper 95% confidence interval of the uninverted MPN was divided by the lower 95% confidence interval of the inverted MPN. The converse was done to calculate the lower 95% confidence intervals for inversion efficiency. The average upper and lower 95% confidence intervals for inversion efficiency for all 20 samples were 57 and 13% respectively.

Therefore, the efficiency ($\pm 95\%$ CI) for *NcoI* inversion was 26 (13-57) %.

3.2.1.7 Detection of long interspersed element (LINE) promoter methylation

As outlined in the Section 1.8, alterations in the methylation of cellular DNA have been reported in the liver tissue of patients with HBV-associated HCC (Lin, Hsieh et al. 2001; Asada, Kotake et al. 2006; Calvisi, Ladu et al. 2007; Tischoff and Tannapfe 2008). This suggests that CpG methylation may be a driver of disease progression. To determine the overall CpG methylation status of the cellular genome, the methylation of LINE promoters was used as a surrogate marker for the amount of the overall CpG methylation in the genome, as previous studies have shown that CpG methylation of LINE promoters is directly proportional to global cellular DNA methylation (Yang, Estecio et al. 2004). Paradoxically, a decrease in global cellular DNA methylation is associated with an increase in cellular CpG-island methylation in HCC (Tischoff and Tannapfe 2008). Indeed, studies have shown that DNA extracted from HCC tissue has greater levels of methylation in CpG islands, but lower levels of methylation in inter-island regions compared to surrounding non-tumour tissue (Asada, Kotake et al. 2006; Lambert, Paliwal et al. 2010).

Therefore, attempts were made to determine whether cellular DNA methylation was associated with different patterns of HBsAg-expression in the liver tissue of patients with chronic HBV infection. The methylation of LINE promoters was measured as described previously (Yang, Estecio et al. 2004). Briefly, DNA extracts were bisulphite-treated using an Imprint DNA methylation kit (Sigma-Aldrich). Treated DNA was then used as a template for PCR amplification of a LINE-1 promoter using specific primers designed by Yang *et al.* with sequences of 5'TTGAGTTGTGGTGGGTTTTATTTAG3' and 5'TCATCTCACTAAAAA-TACCAAACA3' (Yang, Estecio et al. 2004). Half of the product was digested with the RE *HinfI*. The recognition site for *HinfI* in the DNA sequence is preserved if a LINE-1 promoter is methylated. However, the *HinfI* site is converted into a different DNA sequence by

bisulphite-treatment, if the CpG at that site is unmethylated. The undigested and digested PCR products were subjected to agarose gel electrophoresis and image analysis with ImageJ software was used to estimate the ratio of undigested product of each sample. The ratio between undigested and digested samples is an estimate of the percentage of total cellular DNA methylation.

DNA extracted from NHL was bisulphite-treated, PCR amplified, and then either mock digested or digested with *HinfI*, as previously described in Section 2.25. As shown in Figure 3.11, Products of the expected size were detected after digestion of the amplified DNA. ImageJ analysis showed that ~62% of the amplified product was resistant to *HinfI* digestion. This suggested that the overall methylation of cellular DNA was ~38%.

When DNA extracted from foci of 1000-10000 hepatocytes isolated by laser-microdissection were analysed by the same technique, no products could be amplified. This suggested that the small amounts of extracted DNA were damaged by bisulphite treatment. Therefore, the assay was not used in subsequent studies.

3.2.1.8 HBV DNA CpG island methylation

Methylation of CpG islands within HBV DNA may be a mechanism behind the observed heterogeneity in HBsAg expression within liver tissues of chronic HBV patients. In *in vitro* studies, methylation of cccDNA causes downregulation of HBV replication and antigen production (Vivikenandan et al., 2008). Furthermore, the proportion of cccDNA that is methylated is greater in HBeAg-negative patients compared to HBeAg-positive patients (Kaur et al., 2010).

As the copy number of HBV DNA per cell may be much lower than the copy number of LINE sequences per cell, a highly-sensitive RE-based assay was developed for the current study to analyse the methylation of HBV CpG islands. This protocol is described in Section 2.25.

The RE *AciI* recognises the site 5'-CCGC-3' and is blocked by CpG methylation. Sequences of selected HBV genomes of different genotypes were analysed by Methyl Primer Express v1.0 (Applied Biosystems). All genotypes had at least 2 CpG islands, which corresponded to previously described HBV DNA CpG islands 1 and 2 (Vivekanandan, Thomas et al. 2008). Alignment of HBV DNA sequences showed that HBV CpG islands 1

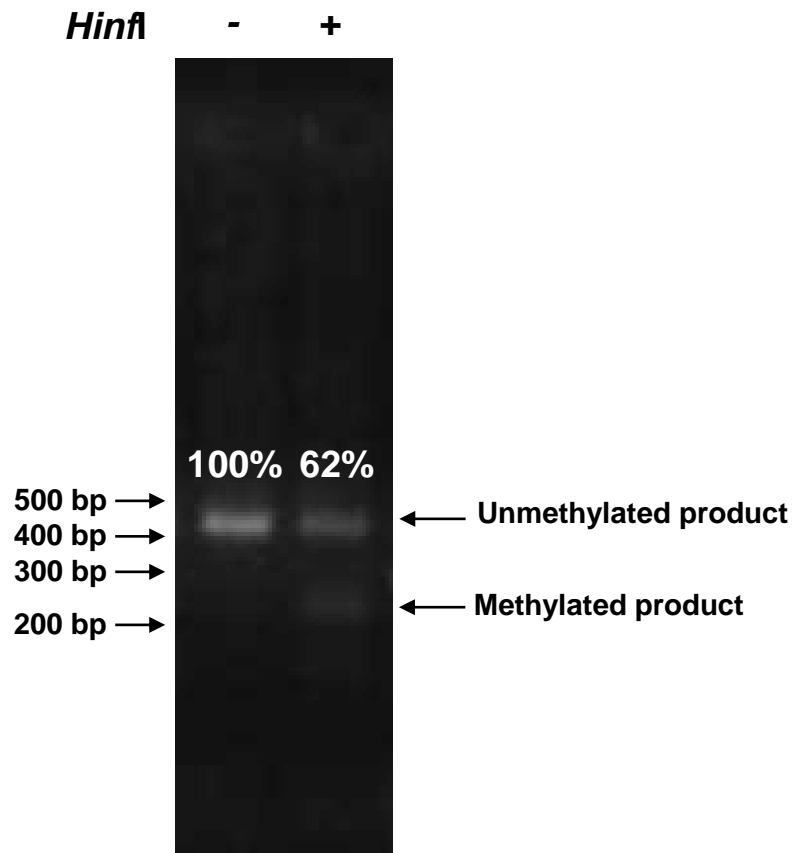


Figure 3.11. CpG Methylation of LINE promoter DNA can be detected by bisulphite treatment of DNA, PCR amplification followed by *Hin*I digestion. DNA extract from 5 mg of snap-frozen uninfected liver tissue was subjected to bisulphite treatment, followed by PCR to amplify LINE promoter DNA sequences and digestion with *Hin*I, as described in Section 2.25. Due to bisulphite treatment, only CpG methylated LINE 1 promoters contain preserved *Hin*I sites.

After digestion of PCR products with *Hin*I, ImageJ analysis detected a 38% drop in band intensity. This suggests that 62% of LINE promoter CpG islands are either unmethylated or mutated in the *Hin*I site.

and 2 both had conserved *AciI* restriction sites, surrounding which primers were designed, as described in Table 2.9.

To determine levels of CpG methylation in HBV DNA, DNA extracts were digested with *AciI*. Extracts were also digested with *EcoRI* to linearise cccDNA molecules and thus to improve the sensitivity of the downstream PCR by increasing denaturation of the DNA template. Digested DNA was used as a template for a qPCR assay containing either primers amplifying HBV CpG island 1, HBV CpG island 2 or a region without an *AciI* site. The level of CpG methylation of each site was calculated as the decrease between samples digested by *AciI* and *EcoRI* compared to *EcoRI* alone. HBV DNA copy number was determined by qPCR of a region of HBV DNA without an *AciI* site. Assays of each DNA extraction was normalised to HBV copy number.

To test the robustness of the assay, mixtures of methylated and unmethylated plasmids containing a 1.3mer of HBV DNA (pBB1.3HBV) were subjected to pilot experiments. Methylated plasmid was generated by incubation of pBB1.3HBV DNA with *M.SssI* CpG methyltransferase (NEB), which methylates cytosine residues in all 5'-CG-3' sequences. Methylated and unmethylated preparations of pBB1.3HBV were mixed in ratios of 0:4, 1:3, 2:2, 3:1 and 4:0. Mixtures were digested with RE and used as template for qPCR assays, as previously described.

The standard curves and melting curves for this assay are summarised in Figure 3.12A. The methylation levels calculated from the qPCR results are graphed in Figure 3.12B and do not significantly differ from the theoretical methylation value given the mixture of methylated and unmethylated pBB1.3HBV. These results indicate that the methylation assay is robust.

3.2.2 Optimisation of proteomic methods

3.2.2.1 Optimisation of imaging mass spectrometry (IMS)

Heterogeneous HBV antigen expression has been observed in liver tissue of patients with chronic HBV infection (Gowans and Burrell 1985; Gowans, Burrell et al. 1985; Mason, Liu et al. 2010). Heterogeneity occurs by some unknown mechanism despite the exposure of all hepatocytes to continual infection by circulating HBV virions. Thus, we hypothesise that the determinant for HBV antigen expression is cell-dependent. Therefore, cellular protein expression patterns were measured in hepatocytes with and without HBV-antigen expression

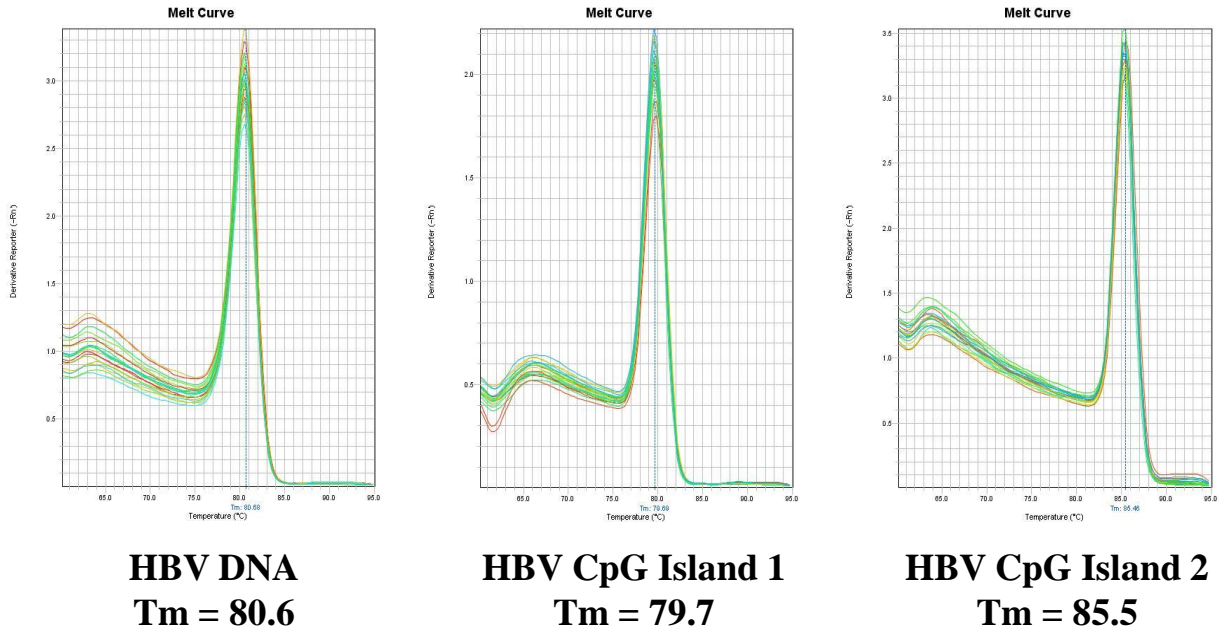
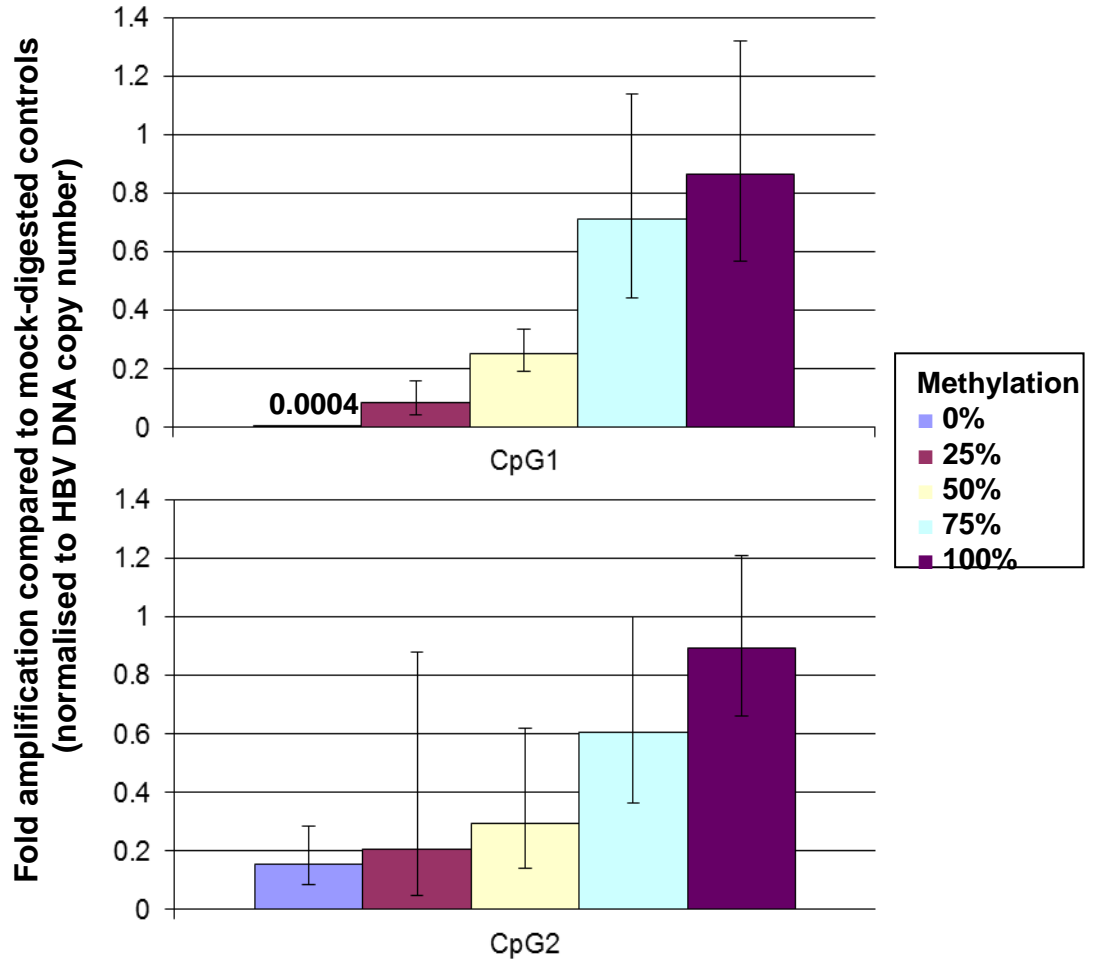
A**B**

Figure 3.12. CpG island methylation of HBV DNA was quantified using a *AciI* digestion and qPCR. CpG methylation of HBV DNA was quantified by a novel method as described in Section 2.24 using CpG methylation-sensitive restriction enzymes and qPCR. *SssI* CpG methyltransferase was used to add methyl groups to 5'-CG-3' dinucleotides in aliquots of pBB4.5HBV1.3. CpG-methylated plasmid was mixed with unmethylated plasmids in ratios such that mixtures contained 0, 25, 50, 75 and 100% of the CpG-methylated plasmid. Plasmid mixtures were either digested with methylation-sensitive digestion enzyme *AciI* or mock-digested. Digestion was measured by qPCR using primers amplifying products across *AciI* sites in CpG islands 1 and 2 of the HBV DNA genome. Primers amplifying HBV sequences not containing *AciI* sites were used for normalisation.

Melting curves for the three products are shown in 3.12A and indicate that single products were amplified. The fold changes in amplification signal from mock-digested to *AciI*-digested the plasmid mixtures are summarised for CpG island 1 and 2 assays in Figure 3.12B. Sample analysis was done in duplicate, error bars represent the maximum and minimum change in amplification based on the duplicate samples.

to identify differences in cell phenotype and to determine whether the changes in cellular protein expression affected disease progression.

To detect protein expression patterns of hepatocytes *in situ*, collaborations were set up with Mr. Johan Gustafsson (Adelaide Proteomics Centre) to optimise an IMS assay. Briefly, the IMS assay involves digestion of proteins with either modified porcine trypsin (Promega, Madison WI) or Trypsin Gold (Promega, Madison WI) on a tissue section and subsequent matrix-assisted laser-desorption and ionisation (MALDI) of digested peptides at 100 μm intervals. Mass spectra of the ionised peptides were produced by analysis with an Ultraflex III MALDI-tandem time of flight (TOF/TOF) instrument. A heat map of each detectable mass/charge value was generated over the surface of the liver tissue section.

For optimisation studies, sections of NHL were mounted onto slides and analysed, as previously described in Section 2.26. The sum spectra across the area of the same tissues digested with either porcine trypsin or Trypsin Gold were virtually identical.

The majority of the mass peaks detected were not due to trypsin autolysis. The sum spectra of the control tissues were compared to a list of trypsin autolysis mass peaks, previously reported by Harris *et al.* (2002). As shown in Figure 3.13, only 2 out of 7 expected autolysis mass peaks were observed above noise levels (arbitrarily set at A.U. >5). The distributions of the prominent autolysis mass peaks were concentrated on the edges of the tissues. Low-level autolysis mass peaks were homogeneously distributed throughout the section, so were likely to be expected introduced masses, as opposed to endogenous masses.

The sum mass spectra, i.e. the sum of the mass peak spectra taken at each point of the liver tissue, of NHL are shown in Figure 3.14. The spectra from the different sections within the same patient and different sections in different patients all showed similar peaks. This suggests that liver tissues in different patients who were not infected with HBV have similar protein expression profiles. The distributions of mass peaks detected in control NHL tissues were generally homogeneously distributed, as shown in Figure 3.15. This suggests few subpopulations of hepatocytes with altered protein expression were present in patients without chronic HBV infection.

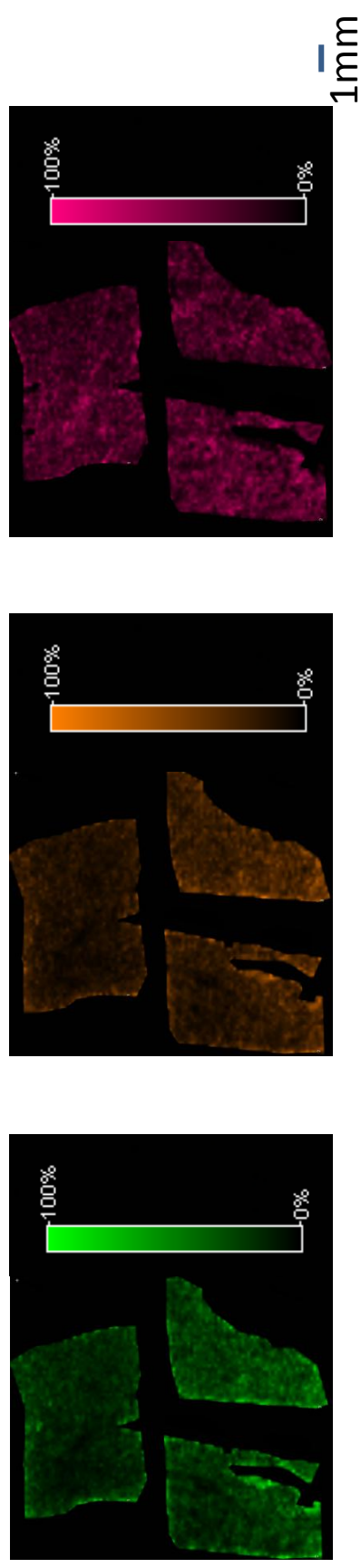
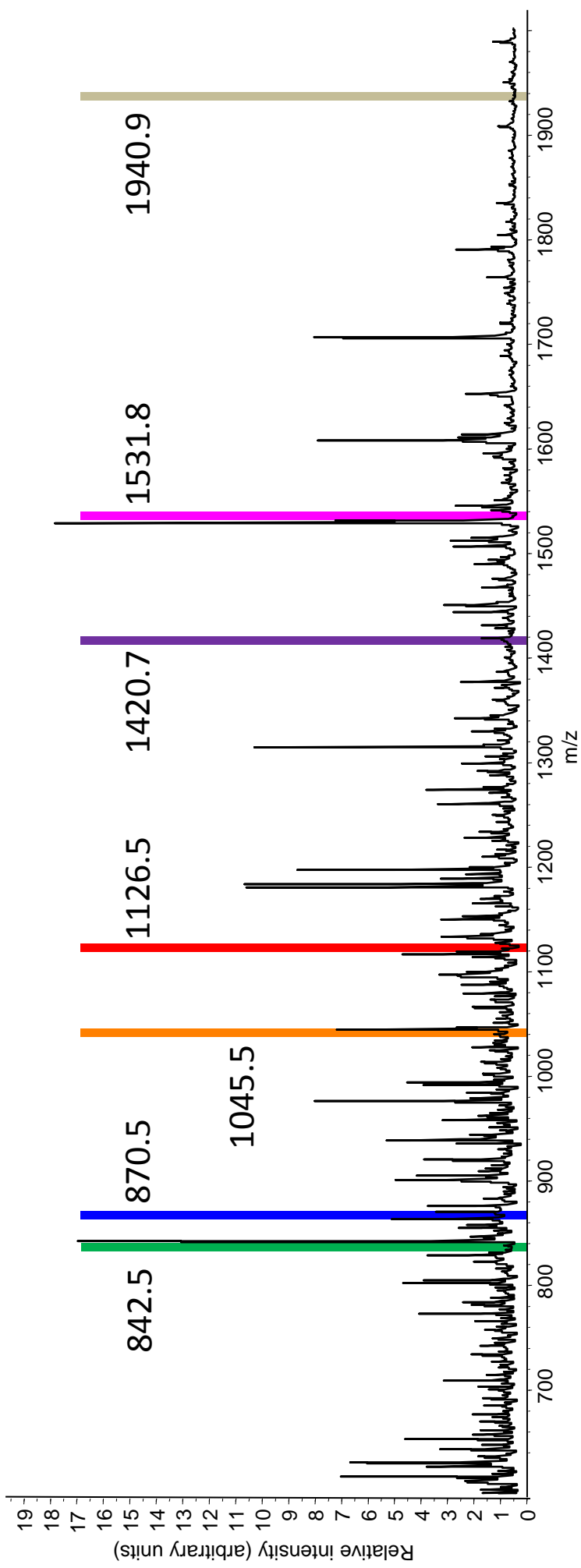


Figure 3.13. Distribution of major modified porcine trypsin autolysis peptides detected by IMS in a section of control uninfected NHL tissue. IMS was carried out as described in Section 2.26 on ethanol-fixed, paraffin wax-embedded NHL tissue sections. The top panel shows the spectrum of mass-peaks summed from all the mass spectra taken from each point across a section of liver tissue (x-axis) and their intensities (y-axis). The intensity of the detected mass peaks were normalised to total ion charge and are shown in arbitrary units. The expected major mass peaks associated with the autolysis of modified porcine trypsin (previously reported by Harris et al. (2002)) are highlighted in various colours.

In the lower panel, the distribution of some of the trypsin autolysis mass peaks over the section of liver are shown as peptide distribution maps. The intensity of colour in each peptide distribution map shows the intensity for the particular mass peak as a percentage of the total intensity. Autolysis peptides that were found in moderate levels (highlighted and mapped in orange) were concentrated around the edges, whereas those peptides found at low levels (all others) were homogeneously distributed and likely to be noise. The resolution of the peptide distribution map is 100 μm between each pixel.

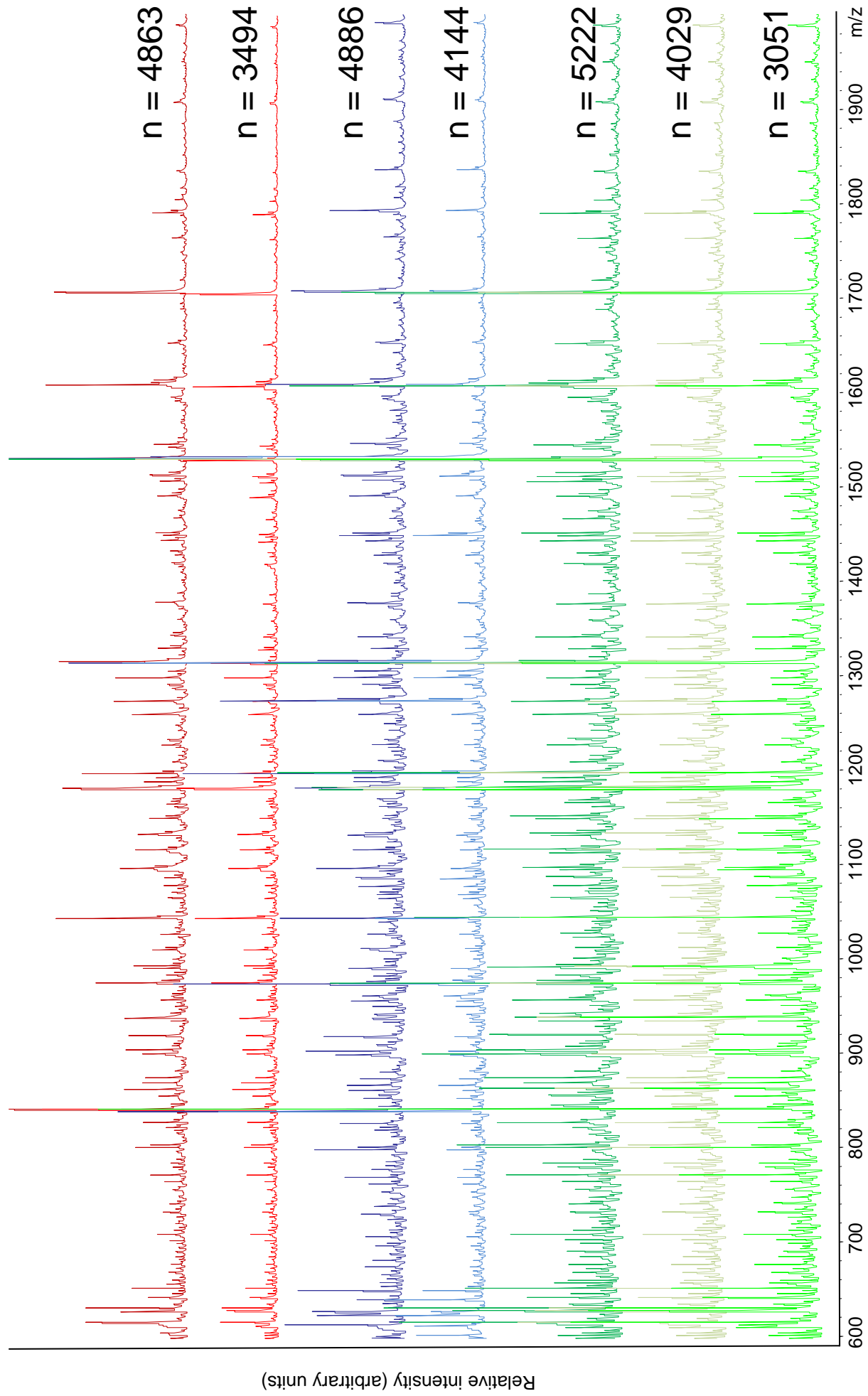
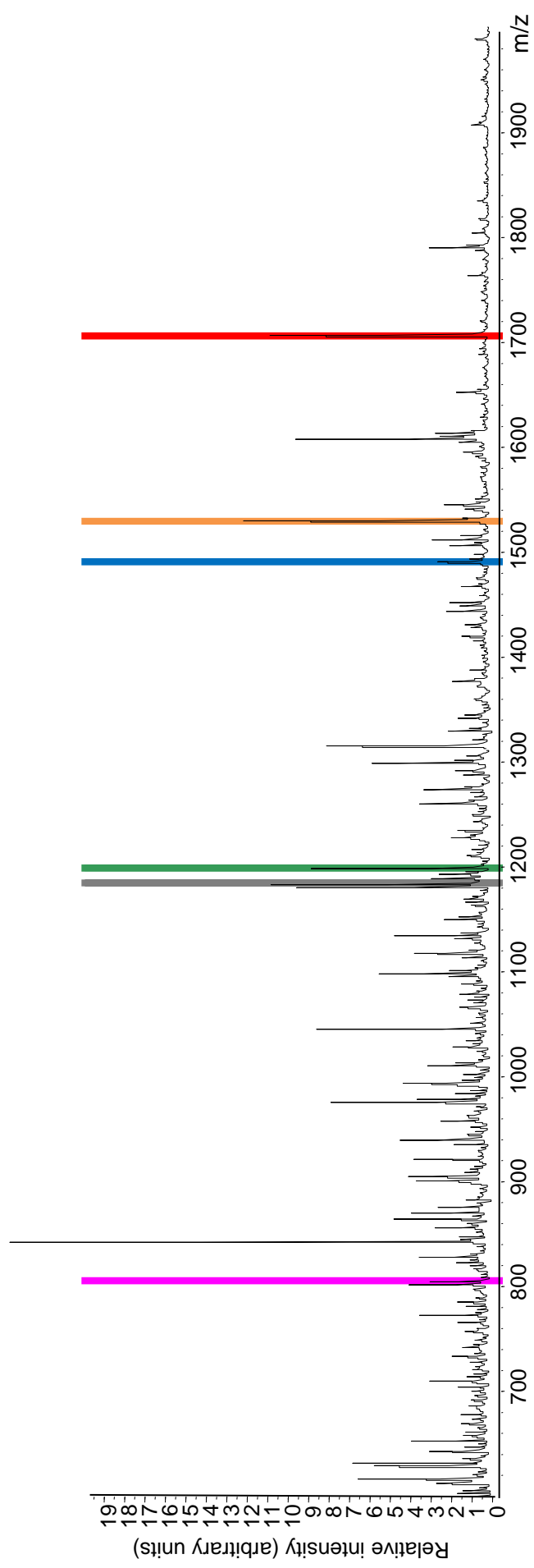
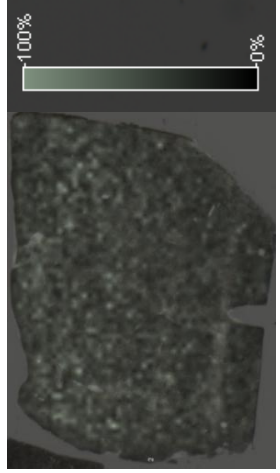


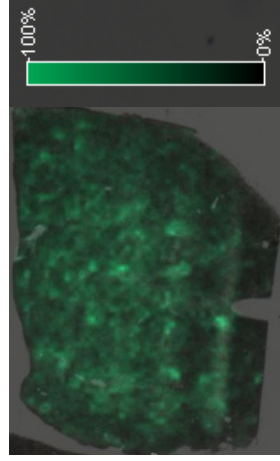
Figure 3.14. Sum spectra of 7 tissue sections of control uninfected NHL tissue. IMS was carried out as described in Section 2.26 on 7 different ethanol-fixed, paraffin wax-embedded liver tissue sections from 3 control uninfected patients. Sum spectra from patients NHL1, NHL2 and NHL3 are shown in different shades of red, blue and green. The spectrum of mass-peaks summed from all the mass spectra taken from each point across a section of liver tissue (x-axis) are plotted against their relative intensities (y-axis). Detected mass peaks were not normalised to total ion charge. The number of separate mass spectra (n) that were summed to produce each sum spectrum for each section are shown. Similar mass peaks were observed in all 7 sum spectra, suggesting similar protein expression profiles.



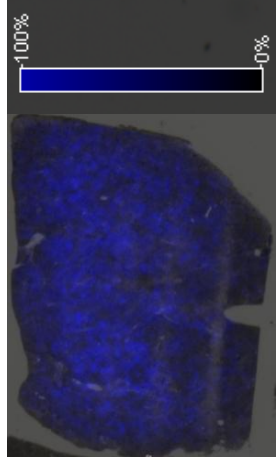
1184.5±0.1



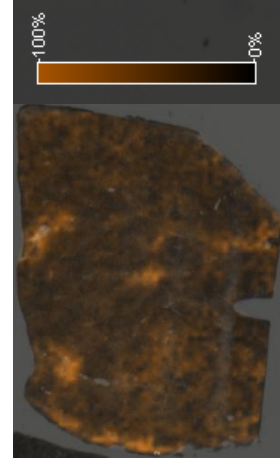
1199.0±0.1



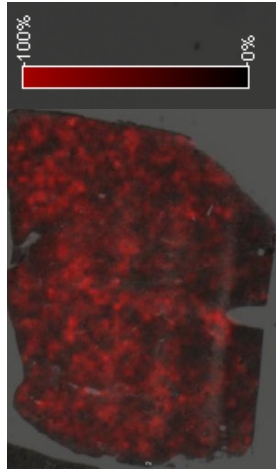
1491.0±0.1



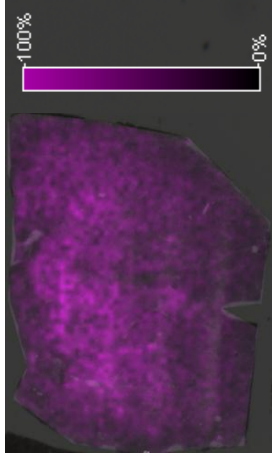
1531.1±0.1



1706.0±0.1



805.5±0.1



1mm

Figure 3.15. Distribution of representative peptide fragments detected by IMS in a section of control uninfected NHL tissue. IMS was carried out as described in Section 2.26 on ethanol-fixed, paraffin wax-embedded liver tissue sections from control uninfected patients. In the top panel, the sum spectrum of a single section of liver tissue is shown. Detected mass peaks were normalised to total ion charge. Six representative mass peaks were selected and the distribution of each mass ($\pm 0.1m/z$) in the lower panels. The intensity of colour in each peptide distribution map shows the intensity for the particular mass peak as a percentage of the total intensity. The resolution of the peptide distribution map is 100 μm between each pixel.

3.2.2.2 Optimisation of High Pressure Liquid Chromatography nano-Electrospray Ionisation Linear Ion Trap Orbitrap Mass Spectrometry (HPLC-nESI-LTQ Orbitrap MS)

To further characterise and identify observed tryptic peptides, collaboration was undertaken with Ms Yin Ying Ho (Adelaide Proteomics Centre) to optimise a HPLC-nESI-LTQ Orbitrap MS peptide identification strategy.

Briefly, foci of HBsAg-positive and -negative hepatocytes were isolated by laser-microdissection. Proteins were extracted using urea and isolated on a VivaSpin 10 kDa filter column (Cole-Parmer). Isolated proteins were digested with 333 ng/mL or 33.3 ng/mL Trypsin (Promega) and eluted. The resultant peptides were analysed by MALDI-TOF/TOF MS to detect potential contaminants (e.g. polyethylene glycol). If no contaminants were observed, 1-5 μ L of peptide eluate was separated by reverse phase liquid chromatography (RP-HPLC). RP-HPLC allows differentiation by organic solvent solubility (i.e. hydrophobicity) and thus decreases the complexity of the sample analysed, allowing more in depth characterisation. The peptides eluted from the RP-HPLC gradient were subsequently analysed by nESI-LTQ Orbitrap MS and MS/MS.

For pilot studies, sections of NHL were mounted onto polymer membrane-coated slides and areas containing 2800-13000 hepatocytes were laser-microdissected. The protein from laser-microdissected tissue was extracted, digested and cleaned using either a VivaSpin 10 kDa centrifugal concentrator (Cole-Parmer) or PepClean C-18 Spin Columns (Thermo Scientific) as described in Section 2.27. Cleaned peptides were analysed by MALDI-TOF/TOF MS to determine whether membrane polymer had contaminated the final peptide eluates. If no polymer contamination was detected, then HPLC-nESI-LTQ Orbitrap MS was used to identify peptides in the eluates.

As the summary in Table 3.6 shows, nESI-LTQ Orbitrap MS was able to identify 14-56 proteins per extract processed by the VivaSpin method. As expected, greater numbers of peptides were detected in extracts of greater numbers of cells. Increased trypsin concentration did not improve peptide detection or protein identification in most cases. Additional masses were not observed as peptide eluate volume was increased, suggesting that maximum sensitivity was reached for these extracts with regard to sample volume.

Table 3.6. A summary of HPLC-nESI-LTQ-Orbitrap MS results from VivaSpin 10 kDA centrifugal concentrator-purified peptides of laser-microdissected hepatocytes from liver sections of HBV-uninfected patients.

Hepatocytes extracted	Trypsin concentration (ng/μL)	Number of peptides detected	Number of proteins identified¹
2800	33.3	26	9
2800	333	794	10
4400	33.3	962	17
4400	333	931	6
11000	33.3	935	14
11000	333	987	13
13000	33.3	1246	15
13000	333	1512	25

¹A full list of proteins detected can be found in Appendix 9.4.1.

Table 3.7. Summary of HPLC-nESI-LTQ-Orbitrap MS results from PepSpin C18 Spin column-purified peptides of laser-microdissected hepatocytes from liver sections of HBV-uninfected patients.

Hepatocytes extracted	Trypsin concentration (ng/μL)	Number of peptides detected	Number of proteins identified¹
660	50	482	11
1300	50	817	15
2300	50	357	6
2300	50	5459	686
2800	50	865	91
4000	50	5081	699

¹A full list of detected proteins can be found in Appendix 9.4.2.

As summarised in Table 3.7, nESI-LTQ Orbitrap MS was able to identify 11-699 proteins per extract processed by the VivaSpin method. In equivalent samples, the Pepclean did indeed detect greater numbers of peptides and proteins. The proteins that were identified by alignment of the peptides to known human proteins by MASCOT search (available at http://www.matrixscience.com/search_form_select.html) are summarised in Appendix 9.4 and included liver-specific proteins, such as liver fatty acid-binding protein, liver 6-phosphofructokinase, liver glycogen phosphorylase and liver carboxylesterase. Furthermore, several metabolic enzymes which are highly expressed in the liver were also identified, such as catalase, aldehyde dehydrogenase, and short chain specific acyl-CoA dehydrogenase. This suggests that the assay was specifically detecting peptides extracted from the laser-microdissected hepatocyte foci.

3.2.2.3 Optimisation of High Pressure Liquid Chromatography Matrix-Assisted Laser-Desorption Ionisation- Time of Flight (HPLC-MALDI-TOF) Tandem Mass Spectrometry

An alternate technique to identify specific mass peaks is the use of HPLC-MALDI-TOF MS/MS. The advantage of this technique is that the ionisation method is similar to that used in the IMS protocol, thus the peptides detected by IMS are more likely to be detected and identified by HPLC-MALDI-TOF MS/MS.

Therefore, 5 μ L of PepClean-purified samples previously analysed by HPLC-nESI-LTQ Orbitrap MS was further analysed by HPLC-MALDI-TOF MS/MS as described in Section 2.28. The results of peptide and protein identification by HPLC-MALDI-TOF MS/MS are summarised in Table 3.8. The majority of the proteins that were identified by HPLC-MALDI-TOF MS/MS (Appendix 9.5) were also identified by HPLC-nESI-LTQ Orbitrap MS, e.g. histone proteins, catalase and heat shock proteins.

A comparison of all of the peptide identification techniques is shown in Figure 3.16. As these results show, HPLC-MALDI-TOF MS/MS was generally less sensitive in detecting peptides compared to HPLC-nESI-LTQ Orbitrap MS analysis of the same samples.

Since HPLC-nESI-LTQ Orbitrap MS identified a similar subset of proteins to and is more

Table 3.8. Summary of HPLC MALDI-TOF MS/MS results from PepSpin C18 Spin column-purified peptides of laser-microdissected hepatocytes from liver sections of HBV-uninfected patients.

Hepatocytes extracted	Trypsin concentration (ng/μL)	Number of peptides detected¹	Number of proteins identified²
2300	50	ND	ND
2800	50	37	16
4000	50	79	40

¹ND = None detected

²A full list of detected proteins can be found in Appendix 9.5.

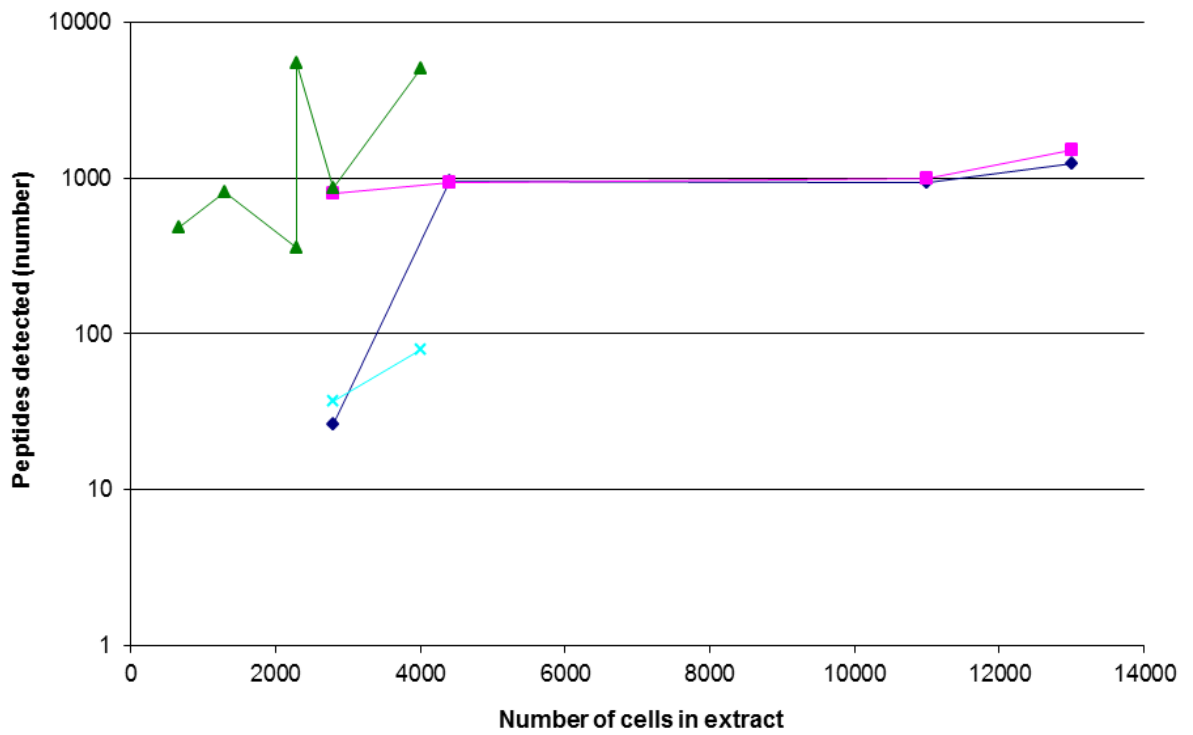
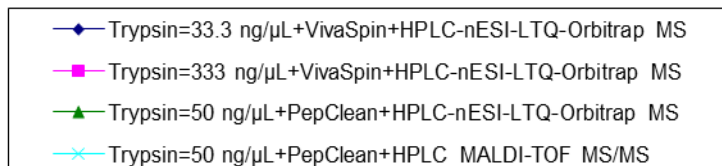
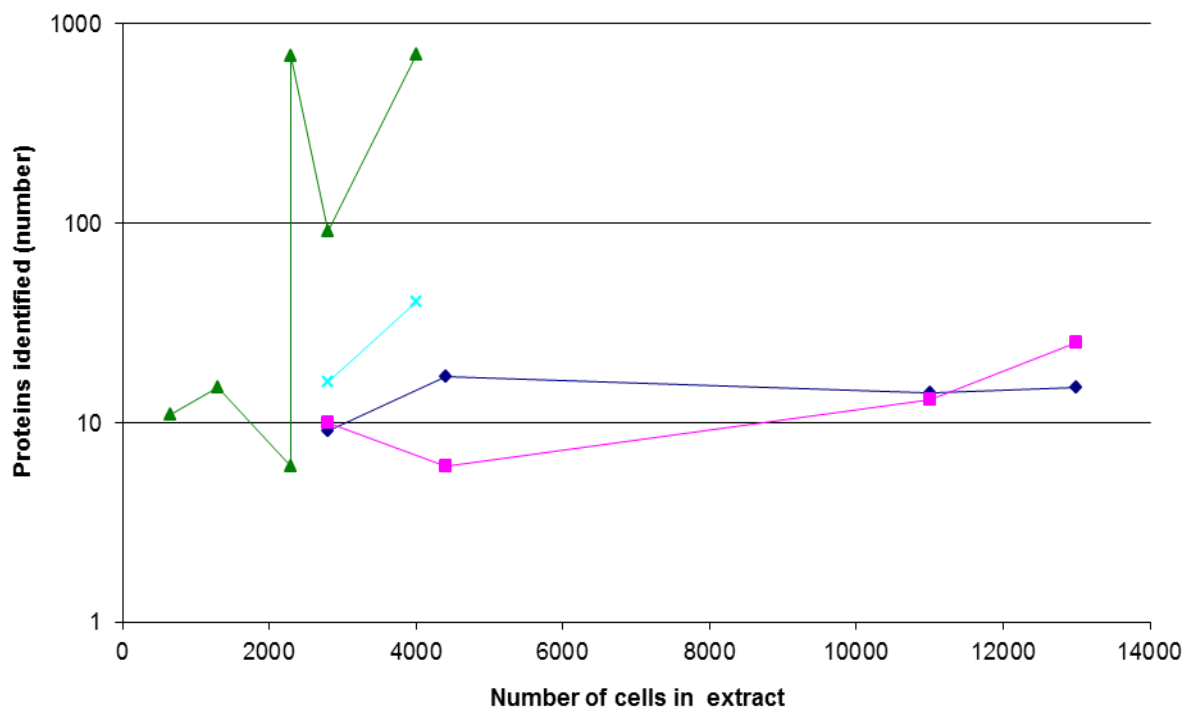
A**B**

Figure 3.16. Peptides detected (3.16A) and proteins identified (3.16B) by various preparations of tryptic digestion of laser-microdissected hepatocytes from NHL sections. Peptides contained within purified protein extracts of laser-microdissected tissues were detected using either: HPLC-nESI-LTQ Orbitrap MS, as described in Section 2.27; or HPLC-MALDI-TOF MS/MS, as described in Section 2.28. The Pepclean peptide purification method (green and light blue) produced greater yield of both peptides detected and proteins identified compared to the Vivaspin method (dark blue and pink) for equivalent numbers of cells extracted. Furthermore, the Orbitrap MS technique (green) was more sensitive in peptides detection and proteins identification compared to MALDI-MS/MS (light blue). Greater concentrations of trypsin (pink) did not affect detection or identification rates compared to lower concentrations (dark blue).

sensitive than HPLC-MALDI-TOF MS/MS, the former technique was used in experimental samples to identify mass peaks detected by IMS.

3.2.2.4 Optimisation of visualising the distribution of known proteins in liver tissue sections

Peptides identified by HPLC-nESI-LTQ Orbitrap MS were matched to their protein by MASCOT analysis and their distribution mapped in FlexImaging Software (Bruker Daltonics) as described in Section 2.29.

Briefly, mass peaks were automatically selected from Gaussian-smoothed sum spectra generated by IMS of three liver tissue section from control uninfected patients. Subsequently, ion maps were generated for each mass peak. Then, the list of identified mass peaks generated by HPLC-nESI-LTQ Orbitrap MS analysis of HBV-negative hepatocytes as described in Section 3.2.2.2 was aligned with the list of mass peaks generated by IMS. Therefore a particular mass peak could be matched with both its distribution within the liver tissue and its sequence. MASCOT analysis of the peptide sequence could then be used to align each sequence to a known protein.

IMS and HPLC-nESI-LTQ Orbitrap MS both independently detected a large number of peptides (944-1214 and up to 5081 peptides, respectively) as shown in Tables 3.7 and 3.8. However, only a minority (15-27%) of mass peaks identified by IMS were able to be matched up with a peptide sequence identified by HPLC-nESI-LTQ Orbitrap MS.

Mass peaks that were aligned to the same protein had similar distributions. For example, the distributions of 4 mass peaks aligned to α and beta subunits of haemoglobin are shown in Figure 3.17. All four mass peaks had a similar distribution. High concentrations of α and beta subunits of haemoglobin were observed in the ion maps that corresponded to blood vessels in H&E-stained adjacent serial sections.

Furthermore, the distributions of 3 different peptides that were aligned to the liver-specific protein Aldolase B were similar and are shown in Figure 3.18. Like many of the mass peaks detected by IMS previously mentioned in Section 3.2.2.1, the distribution of the peptides that were aligned to Aldolase B were generally homogeneously distributed amongst the hepatocytes in the tissue sections.

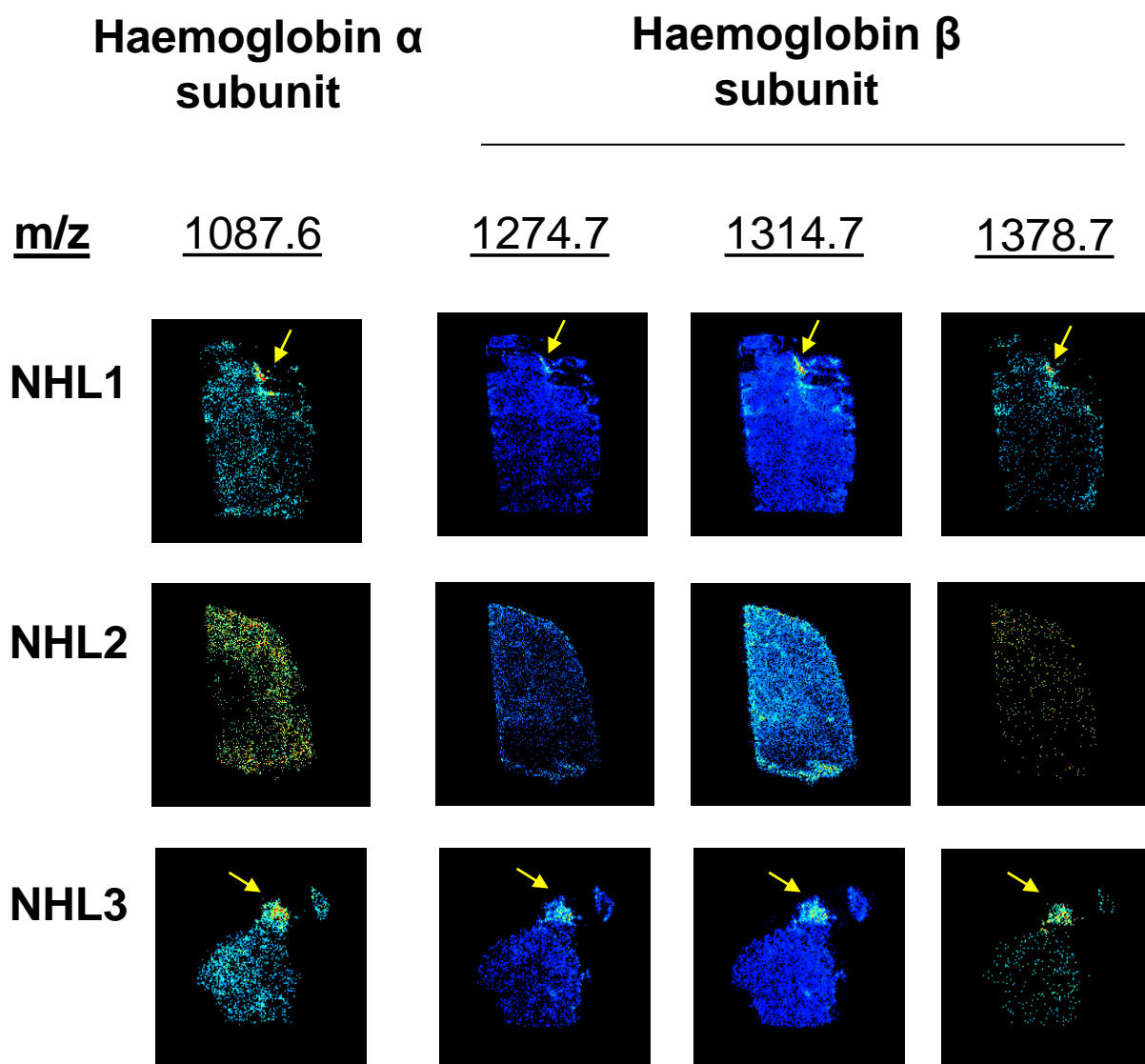


Figure 3.17. Distribution of three peptide fragments that align with α and β subunits of haemoglobin in three NHL sections. IMS was carried out in Section 2.26 on ethanol-fixed, paraffin wax-embedded NHL sections from 3 patients (NHL1-3). Mass peaks were selected and the distribution of each mass (± 0.2 m/z) in the tissues are shown as a percentage of the total intensity. Mass peaks were identified by comparing mass-charge ratios with those attained by HPLC-nESI-LTQ Orbitrap MS as described in Section 2.29. Generally homogeneous distributions throughout the parenchyma of all mass peaks that align with haemoglobin proteins are apparent in all three sections. Concentrations of haemoglobin (yellow arrows) correspond to blood vessels. The resolution of the peptide distribution map is 100 μm between each pixel.

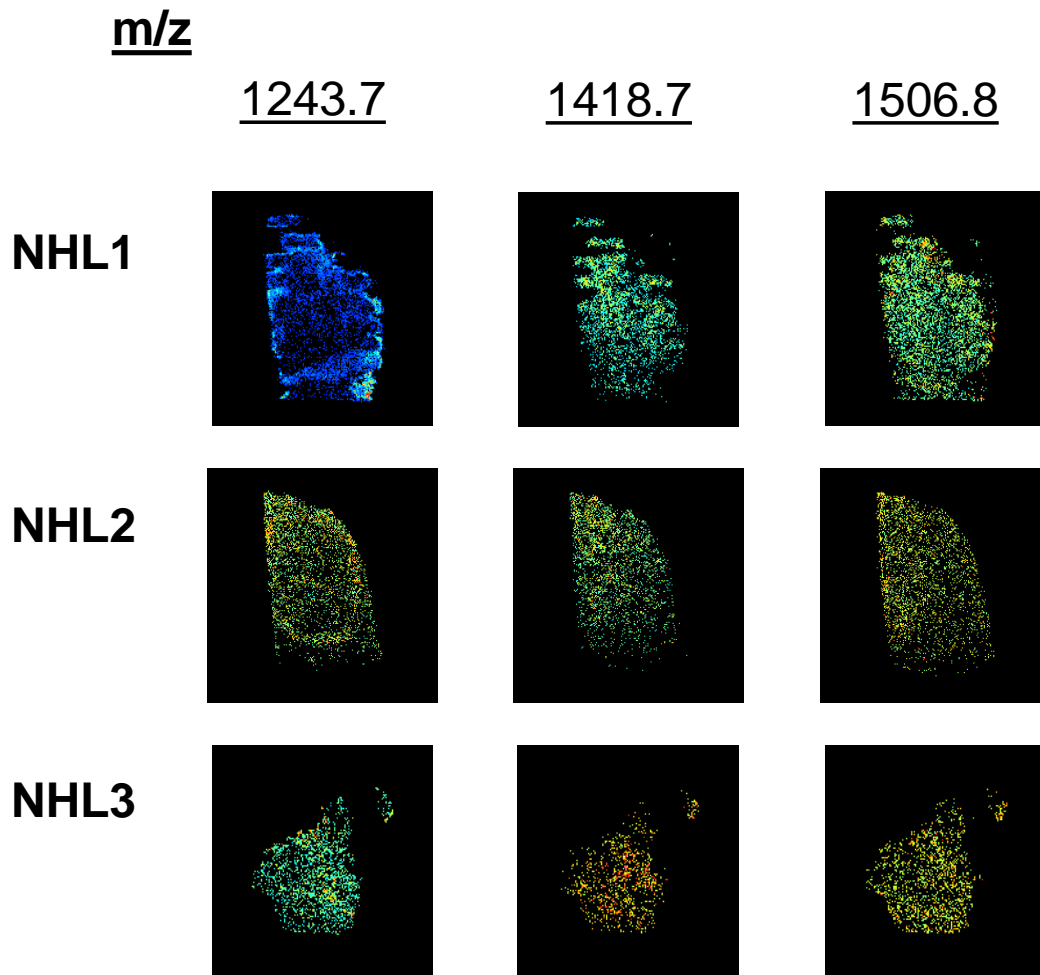


Figure 3.18. Distribution of three peptide fragments that align with Aldolase B, a.k.a. liver-type fructose-bisphosphate aldolase, in three NHL sections. IMS was carried out as described in Section 2.26 on ethanol-fixed, paraffin wax-embedded NHL sections from 3 patients (NHL1-3). Mass peaks were selected and the distribution of each mass (± 0.2 m/z) in the tissues are shown as a percentage of the total intensity. Mass peaks were identified by comparing mass-charge ratios with those attained by HPLC-nESI-LTQ Orbitrap MS as described in Section 2.29. Homogeneous distributions of all mass peaks that align with Aldolase B are apparent in all three sections. The resolution of the peptide distribution map is 100 μm between each pixel.

Thus, this is strong evidence that the mass peaks identified in IMS-generated sum spectra are indeed matched up to proteins accurately using the peak identities generated by HPLC-nESI-LTQ Orbitrap MS.

3.2.3 Optimisation of RNA extraction from frozen liver tissues

RNA expression can be a highly descriptive aspect of cell phenotype. In order to investigate differences in RNA expression between HBsAg-positive and -negative hepatocytes, the quality of RNA in the available patient liver tissues was investigated.

For optimisation experiments, RNA from ~10 mg of snap-frozen liver tissue from patients CYRY and HS was extracted. As a positive control, RNA from ~10mg resected duck liver tissue that had been previously preserved in RNAlater RNA stabilisation solution (Ambion) and snap-frozen by Dr. Georget Reaiche were extracted. Total RNA was extracted from liver tissue using an RNeasy mini RNA isolation kit (Qiagen), as per the manufacturer's instructions. RNA was quantified by optical density spectrometry as described in Section 2.11. Approximately 1 µg of total RNA extract was subjected to agarose gel electrophoresis and analysed by gel electrophoresis, as described in Section 2.12 (Figure 3.19).

A smear of RNA was observed in addition to 2 bands, presumably 18S and 28S ribosomal RNA, suggesting that the RNA in the tissues was intact. As expected due to the preservation of the duck liver tissue by RNAlater during snap-freezing, stronger bands were seen in RNA-preserved duck tissue extracts than human tissue extracts.

To further investigate the RNA integrity, cDNA was produced by incubation of total RNA extract with reverse transcriptase system using poly-T oligoprimers (Promega). A PCR containing GAPDH-specific primers (Toyobo) with sequences 5'-ACCACAGTCCATGCCATCAC-3' and 5'-TCCACCACCCTGTTGCTGTA-3' was set up using the cDNA as a template. Due to post-transcriptional splicing, these primers amplify a 450 bp product from cDNA produced from GAPDH mRNA transcripts, but amplify a 550 bp product in the genomic copy of the GAPDH gene, according to Genbank sequences #AB062273.1 and #NG_007073.2. The expected ~450 bp product was amplified using both control duck liver and HBV-infected liver tissue extracts as template. This suggests that some

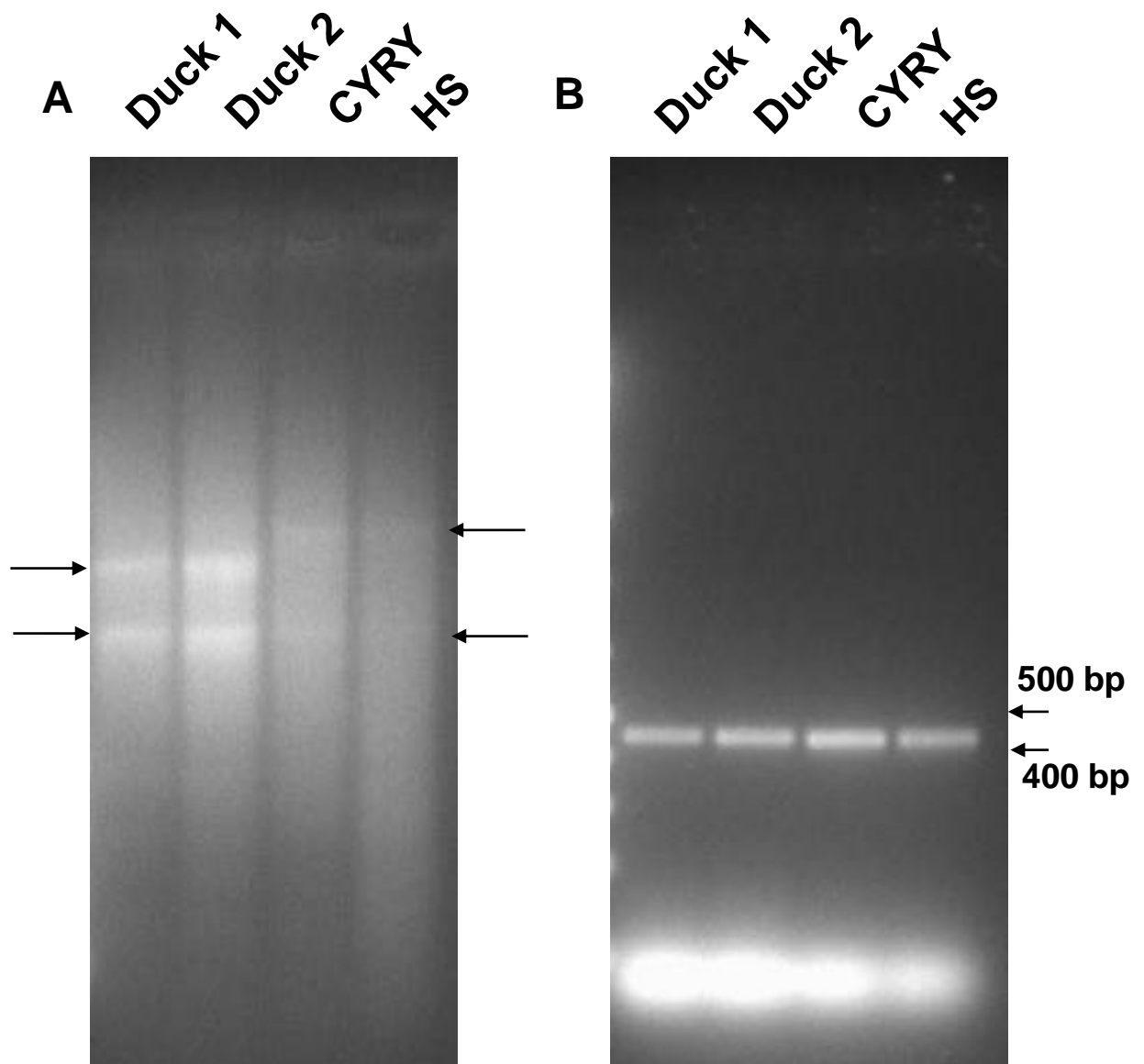


Figure 3.19. Intact RNA was detected in extracts from snap-frozen liver tissue from uninfected ducks and human chronic HBV patients. Approximately 1 μ g of total RNA (3.19A) extracted from RNAlater-treated liver-tissue from two ducks and liver tissue from two HBV-infected patients as described in Section 3.2.3. Analysis of the RNA extracts by gel electrophoresis showed the presence of two distinct bands (arrows), presumably 18S and 28S ribosomal RNA subunits. When GAPDH RNA transcripts were amplified by reverse-transcription (RT) PCR as described in Section 3.2.3 with the RNA extracts used as templates, the expected 450 bp product was produced (3.19B).

proportion of >450 bp mRNA transcript fragments was preserved in human snap-frozen liver tissue from chronic HBV patients.

However, due to limitations in funding and time, the extraction from laser-microdissected tissue, mRNA expression microarray assays and reverse-transcription qPCR of specific mRNA transcripts could not be optimised.

3.3 Discussion

As a result of these experiments, protocols for DNA extraction from liver tissue have been optimised for characterisation of genomic DNA. Also, the optimisation of the invPCR protocol in our laboratory has allowed us to measure the level of clonal proliferation in chronically HBV-infected human liver tissue. Furthermore, the percentage of HBV DNA methylation in a DNA extract can be detected using the developed *AciI* digestion qPCR assay. Preliminary data also shows that imaging MS and nano-LC/Orbitrap MS can be used to examine the protein expression of tissue mounted on slides and laser-microdissected liver tissue, respectively.

3.3.1 DNA extraction

Isolation of DNA was a high priority for optimisation as techniques involved in the characterisation of tissues, the detection of clonal proliferation and the analysis of foci of HBsAg-positive and -negative hepatocytes require isolation of cellular DNA. The current protocols for total DNA extraction isolated cellular DNA from snap-frozen liver tissue, sections of ethanol-fixed paraffin wax-embedded liver tissue and laser-microdissected liver tissue with relatively high yields.

However, the nuclear DNA extraction protocol did not return high yields in our hands (Section 3.2.1.2.3). Also, high molecular weight DNA isolation by gel purification was unsuccessful in completely excluding HBV DNA cccDNA and RI DNA (Section 3.2.1.5). The sensitive nature of PCR and the high-copy number of HBV DNA in tissues and circulating in the bloodstream of the liver still proves an obstacle yet to be resolved.

3.3.2 Detection of virus-cell DNA junctions

To observe hepatocyte clones in HBV-infected liver tissue, unique virus-cell DNA junctions were detected and used as a marker for clonal proliferation. As previously described in

Section 1.5, multiple techniques have been used in the past for detecting integrated HBV genomes; including semi-DOP PCR, Alu-PCR and invPCR.

Attempts to detect integrated HBV DNA by semi-DOP PCR were unsuccessful (Section 3.2.1.3). Semi-DOP PCR had previously been used to amplify sequences of a previously unknown bocavirus subspecies from faecal matter (Arthur, Higgins et al. 2009). While this technique was sensitive enough to amplify the relatively high copy number of virus sequence in faecal extracts of infected patients, semi-DOP PCR was not sensitive enough to detect relatively low (<1,000) copies of DNA sequences in our hands.

Alu-PCR has been used previously to detect integrated virus-DNA junctions, including HBV (Minami, Poussin et al. 1995; Vandegraaff, Kumar et al. 2001; Murakami, Saigo et al. 2005). However, due to high Alu-Alu artefactual amplification at low levels of HBV DNA reported by collaborators (Mason, personal communication), optimisation of Alu-PCR was not pursued.

InvPCR was successfully optimised in our laboratory using a DNA extract from a WHV-infected woodchuck (Section 3.2.1.4). As the extract was taken from a large amount of tissue, a smaller fraction of the total extract was analysed by invPCR. Thus, only a few repeated virus-cell DNA junctions were observed. However, the alignment of the ends of the integrated WHV genome matched the expected left hand junction of the dsDNA form and was consistent with previous studies (Mason, Jilbert et al. 2005).

Furthermore, the *NcoI* invPCR design was able to detect an inverted cellular gene in a control invPCR assay (Section 3.2.1.6). The average efficiency of inversion rate was shown to be ~26%. Thus, any detected clones may be up to 4 times larger than shown by invPCR. In Appendix 9.3, we analysed the theoretical inversion efficiency in the context of previous studies in intra-molecular DNA ligation (Wang and Davidson 1966; Dugaiczky, Boyer et al. 1975). We showed that intra-molecular DNA ligation is preferred to inter-molecular DNA ligation under the conditions of our methods. However, the wide range of calculated inversion rates indicates there are strong inter-patient and inter-sample factors that influence inversion efficiency. These potential factors are discussed in greater detail in Section 5.3.2.

3.3.3 Epigenetic phenotyping

Techniques were developed to compare epigenetic differences in DNA between HBsAg-positive and HBsAg-negative hepatocytes. These experiments were designed to detect differences in methylation of cellular and HBV DNA CpG islands.

Many of the proposed bisulphite-based methylation analyses were unsuccessful. Bisulphite-treatment involves incubation in a highly reducing solution for >1 hr and thus DNA integrity is compromised. Previous studies have shown >80% of DNA can become damaged in this process (Grunau, Clark et al. 2001). The low concentration of DNA isolated from laser-microdissected tissue was therefore not deemed compatible with attempted bisulphite-based analyses of CpG island methylation.

As a result, an assay involving *AciI* digestion coupled with qPCR was developed and optimised. This assay not only detects but also quantifies the level of methylation of cytosine residues in HBV CpG islands 1 and 2. Methylation of cccDNA has previously been shown to decrease levels of HBV protein expression *in vitro* (Vivekanandan, Thomas et al. 2009). Vivekanandan *et al.* (2010) also found an increase in host protein effectors of CpG methylation in HBV infected tissues compared to uninfected controls (Vivekanandan, Daniel et al. 2010). This suggests that methylation is a mechanism for host control of HBV and may also account for the appearance of HBsAg-negative foci. This point is further investigated in Section 7.3.7.

3.3.4 Proteomic phenotyping

Mass spectrometry-based techniques were developed to compare the protein expression of HBsAg-positive and HBsAg-negative hepatocyte foci. Data on the similarities and differences of the protein expression profiles of the hepatocyte foci will allow insights into whether HBsAg-positive and HBsAg-negative hepatocytes 1) are clonal; 2) undergo preneoplastic changes as part of clonal proliferation and 3) show changes in protein expression involved in the progression of a HBsAg-positive to HBsAg-negative phenotype.

Trypsin autolysis mass peaks were observed by IMS to be concentrated around the edges of tissue sections (Section 3.2.2.1). A possible mechanism is that a deposited drop of trypsin straddling the edge of the tissue would contain lower concentrations of cellular protein and thus increase the likelihood of trypsin autolysis. Alternatively, fixation may be more effective at the edges of the tissue section, leading to greater resistance to trypsin digestion. This would lead to greater trypsin autolysis.

Although some trypsin autolysis peaks were observed in preliminary studies using IMS, HPLC-MALDI-TOF MS/MS and HPLC-nESI-LTQ Orbitrap MS assays, the majority of the detected peptides represented cellular proteins. In addition, liver-specific proteins were identified, suggesting these were not protein contaminants.

However, a concerning limitation of the current mass spectrometry-based assays is that levels of protein expression of only the most abundant proteins can be measured. Up to 700 peptides were identified, consistent with previous mass-spectrometry studies on liver tissue (Chaurand, Cornett et al. 2006; Chaurand, Latham et al. 2008; Seeley and Caprioli 2008). However, the human genome contains 20000-25000 genes, so the mass-spectrometry assays are likely to be detecting a small minority of all protein expressed by the cells (International Human Genome Sequencing Consortium, 2004).

There are several potential reasons for the limited identification of cellular proteins.

Firstly, proteins may have trypsin-digested fragments that are unable to be detected by current settings and equipment. Trypsin fragments may either be too short to be confidently mapped to a protein sequence, or too long to be detected and sequenced by TOF/TOF mass spectrometry. Also, the mass spectrometers were operated in reflectron positive ion mode in our assays. Thus, only fragments under a net positive charge were detectable. Also, detection of a particular peptide is preferential for those in greater abundance, so changes in expression in low-abundance proteins may not be detected. Furthermore, since ions must be detected >50 times in a liver section to be included in analyses as part of a noise-reduction filter, proteins expressed in only a few cells may remain unobserved.

Only a minority (15-27%) of mass peaks observed in IMS were able to be matched up with mass peaks detected by HPLC-nESI-LTQ Orbitrap MS, limiting identification of the majority of tryptic peptides. There are multiple potential explanations for this non-overlapping peptide subset. Firstly, the efficiency of trypsin may be different with IMS *in situ* digestion, as compared to in solution digestion, which was performed for HPLC-nESI-LTQ Orbitrap MS. Thus, a different subset of peptides could be generated. Secondly, some peptides may be too hydrophobic or hydrophilic to be efficiently fractionated during the HPLC fractionation process. Also, the different ionisation techniques between IMS (matrix-assisted laser

desorption ionisation) and HPLC-nESI-LTQ Orbitrap MS (nano-electrospray ionisation) may also lead to different subsets of peptides being identified.

Furthermore, even if a particular peptide is detected in both IMS and HPLC-nESI-LTQ Orbitrap MS, only a short peptide chain is used to assign expression of a protein. Thus, the current mass spectrometry assay design cannot differentiate between some post-transcriptional or post-translational gene changes; such as splice variants, various protein isotypes and phosphorylation. For example, a change in a protein isotype expression ratio would not be detected if the sum expression of the isotypes remained constant.

These compromises in the range of proteins detected are in part a consequence of the limited protein yield from tissue isolated by laser-microdissection and by the complexity of tryptic-peptide mixtures produced *in situ* during IMS. These limitations may be overcome by using alternate approaches, such as immunohistochemistry to confirm cellular changes in protein expression or reverse transcription qPCR to determine changes in RNA expression in particular subsets of cells. Indeed, intact RNA was detected in the liver samples available to us, as shown in Section 3.2.3. However, these methods are beyond the scope of the current project.

3.3.5 Summary

In summary, the overall experimental outline is to first characterise tissues using histochemical methods, including the localisation of HBsAg-positive and HBsAg-negative hepatocytes (Chapter 4). In tandem, invPCR will be used to detect virus-cell DNA junctions, which is a measure of clonal proliferation of hepatocytes (Chapter 5). In those tissues that exhibit strong clonal proliferation, immunohistochemical analysis and laser-microdissection will be used to isolate foci of HBsAg-positive and HBsAg-negative hepatocytes (Chapter 7). To determine if the expression of HBsAg is associated with clonal proliferation, these foci will be subject to invPCR.

Other potential phenotypical changes in hepatocyte foci will also be examined using DNA methylation- and protein expression-centred analyses that have been optimised in this Chapter. Results from this work will be complemented by imaging MS, which gives an overall impression of protein expression throughout an entire section of liver (Chapter 7). HPLC-nESI-LTQ Orbitrap MS will be used to identify specific mass peaks detected by imaging MS. These phenotyping analyses will give some insight into the changes that occur

in spatially-distinct hepatocyte subpopulations over the course of chronic HBV infection and, potentially, into the associated liver disease progression.

4 - Characterisation of liver tissues

4.1 Introduction

As mentioned in Chapter 1, chronic HBV infection increases the risk of both HCC and cirrhosis, which generally manifest after decades of chronic HBV infection. Changes in the liver cell population that occur during the development of cirrhosis and HCC have not been fully characterised.

As described in Section 1.10, various lines of evidence support a model where hepatocytes with a heritable survival advantage in the environment of the HBV-infected liver proliferate to form hepatocyte clones in both cirrhotic and non-cirrhotic liver tissues collected from patients with and without HCC (Shafritz, Shouval et al. 1981; Hino, Kitagawa et al. 1984; Chen, Harrison et al. 1986; Chen, Harrison et al. 1988; Esumi, Tanaka et al. 1989; Mason, Liu et al. 2010). Independent studies of HBV-infected liver tissue using Southern blot hybridisation or invPCR have detected integration of HBV DNA into the cell chromosomes. Due to the relatively low sensitivity of Southern blot hybridisation, fragments of virus-cell junctions formed by integrated HBV DNA can only be detected as distinct bands on a Southern blot when many (>1000) hepatocytes contain the same unique virus-cell DNA junction (Shafritz, Shouval et al. 1981; Hino, Kitagawa et al. 1984; Chen, Harrison et al. 1986; Chen, Harrison et al. 1988; Esumi, Tanaka et al. 1989; Mason, Liu et al. 2010).

The invPCR technique, which is more sensitive than Southern blot hybridisation, has also been used to detect multiple copies of unique virus-cell DNA junctions generated by the proliferation of hepatocytes containing integrated HBV DNA (Yang, Mason et al. 1996; Summers, Jilbert et al. 2003; Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). In the woodchuck model, mathematical simulations showed that the extent of hepatocyte turnover during chronic infection was not sufficient to produce the hepatocyte clones with a size of >10000 cells by stochastic clonal proliferation (Mason *et al.* 2005). Thus, data from both Southern blot hybridisation and invPCR strongly suggests that clonal proliferation of hepatocytes with survival advantages occurs during chronic HBV infection.

However, many unknowns remain regarding the clonal proliferation of hepatocytes, such as: the relative extent of clonal proliferation of hepatocytes between cirrhotic and non-cirrhotic tissues; whether the clonal proliferation of hepatocytes contributes to the development of late-stage diseases; the stage of chronic HBV infection at which clonal proliferation of

hepatocytes occurs; and the identity and the phenotype of the specific cell populations that undergo clonal proliferation.

To examine these aspects, liver tissues were collected from patients with chronic HBV infection at various stages of disease. Each liver sample was examined for markers of liver pathology and HBV infection to aid in classifying the samples and to determine the stage in terms of disease progression. Based on liver pathology and markers of HBV infection, patients were assigned to different groups of varying disease progression.

This information was then used in subsequent Chapter to determine whether the progression of liver disease in each group of patients was associated with more extensive clonal proliferation of hepatocytes (Chapter 5) or specific changes in the phenotype of hepatocytes (Chapter 7).

4.2 Experimental outline

In total, liver tissue from 30 HBV-positive patients and 4 HBV-negative patients were obtained from various clinical research centres. All tissues were studied under Royal Adelaide Hospital Human Research Ethics Committee (HREC) approval (Protocol #070301, Appendix 9.4.1). Copies of the ethics approvals for all protocols are provided in Appendix 9.4. The liver tissues were kindly supplied by the following collaborators:

Dr. Nick Shakel and Prof. Geoff MacCaughan (Centenary Institute, Sydney, NSW) generously provided snap-frozen liver tissue from 10 patients with chronic HBV infection and cirrhosis. Patient tissues were collected under NSW Health HREC approval (Protocols #X10-0072 & #HREC/10/RPAH/130, Appendix 9.4.2) as explant material collected prior to liver transplantation.

Emeritus Prof. William Mason (Fox Chase Cancer Centre, Philadelphia, PA, USA) generously provided paraffin wax-embedded ethanol-fixed liver tissue mounted on slides. These tissues were collected under Fox Chase Cancer Center Institutional Review Board Approval (Approvals #31-281 & 08-801, Appendix 9.4.3). The liver tissues were collected from the University of Washington Medical Centre and

included: i) non-tumour tissue of 5 patients that had liver resection for HCC; and ii) liver needle biopsy tissues from 5 patients with chronic active hepatitis. Tissues from the 5 HCC patients have been used in previous studies (Mason *et al.*, 2010).

A/Prof. Hugh Harley (Royal Adelaide Hospital, Adelaide, SA) generously provided liver needle biopsy tissue from 2 patients with chronic active hepatitis. These patients underwent a liver biopsy as part of clinical management of their HBV infection. These tissues were collected under Royal Adelaide Hospital HREC approval (Protocol #070301, Appendix 9.4.1).

Dr. John Chen and Ms. Libby John (South Australian Liver Transplant Unit, Flinders Medical Centre, Adelaide, SA) generously provided snap-frozen liver tissue from 7 decompensated cirrhotic patients with chronic HBV infection. Patient tissues were collected under Southern Clinical HREC approval (Protocol 171.11, Appendix 9.4.4) as excess material recovered on liver resection. For our studies, we sought retrospective approval from each patient to gain access to these liver samples. These amendments to Protocol 171.11 were approved as shown in Appendix 9.4.4.

Dr. Andrew Ruskiewicz (SA Pathology, Adelaide, SA) generously provided snap-frozen resected liver tissue from 3 HBV-negative patients with colon cancer tumours that had metastasised to the liver. These liver tissues were used as negative controls in our assays. Dr. Ruskiewicz also provided snap-frozen resected non-tumour liver tissue from a HBV-positive patient with HCC. These tissues were collected under Royal Adelaide Hospital HREC approval (Protocol #050803, Appendix 9.4.5).

Liver tissue from the 30 HBV-positive patients and 3 HBV-negative patients were fixed, paraffin wax-embedded, sectioned, and then mounted onto slides, as described in Section 2.2. Tissues were stained as described in Section 2.3 by H&E, Haematoxylin Van-Geisen, PAS-D, Gordon and Sweet's Reticulin staining, Perls' Prussian Blue and Fouchet's bile staining (SA Pathology). Liver histology was interpreted by Assoc. Prof. Andrew Clouston (University of Queensland Centre for Clinical Research). The HBV DNA sequence associated with each patient was determined by extraction of liver DNA, PCR amplification (primers sequences in Table 2.2) and Sanger sequencing, as described in Section 2.12.

Furthermore, the extent and distribution of HBV replication was determined by: qPCR assays, using liver DNA extracts to quantify the average number of copies of HBV DNA per liver cell; *in situ* hybridisation of liver sections to determine the distribution of HBV DNA-containing hepatocytes; and staining of liver sections by IHC to detect the distribution of HBsAg- and HBcAg-expressing hepatocytes. These protocols were carried out as described in Sections 2.11, 2.6 and 2.4, respectively. These analyses would allow us to identify foci of hepatocytes that do not support HBV replication or antigen expression and thus determine whether these hepatocytes are more likely to undergo clonal proliferation and/or to contribute to disease progression.

In addition, the susceptibility of oval cells, the putative stem cell population of the liver (Evarts, Nagy et al. 1987), to HBV infection was determined by dual-immunofluorescence detection of HBsAg and cytokeratin-19, as described in Section 2.5. Cytokeratin-19 has previously been reported to be a marker for oval cells (Evarts, Nagy et al. 1987; Lowes, Brennan et al. 1999). This analysis was designed to determine if oval cells were infected with HBV, expressed HBsAg and therefore potentially contained integrated HBV DNA. We ultimately aimed to determine if the cellular clones detected using invPCR were likely to be due either to:1) proliferation of oval cells that had been infected with HBV and contained integrated HBV DNA; or 2) clonal proliferation of terminally-differentiated hepatocytes that had been infected with HBV.

4.3 Results

4.3.1 Patient details and liver histology reports

The sections of liver tissue were mounted, stained and sent to liver pathologist A/Prof. Andrew Clouston for examination. The histological findings and patient details are summarised in Table 4.1.

As shown in Table 4.1, the age of patients at biopsy or resection ranged from 22 to 70 years old, with the median age of 47.3 (SD= \pm 9.8) years. Four of the 30 HBV-positive patients were female.

Table 4.1. Summary of clinical patient data and histological interpretation of liver tissues collected from patients with chronic HBV patients in early stages of HBV infection (4.1A), with HBV-associated cirrhosis (4.1B), fulminant hepatitis (4.1C), and HCC without cirrhosis (4.1D) and HBV-negative controls (4.1E). Patient WN was unable to be characterised due to severe diathermy or drying artefact (4.1F).

A – Early-stage HBV infection

Name ¹	Type ²	Age	Sex	Fibrosis (METAVIR/shak) ³	Interface hepatitis (Ishak A) ³	Centro-lobular dropout (Ishak B) ³	Lobular inflam'n (Ishak C) ³	Portal inflam'n (Ishak D) ³	Steatosis ³	Dysplastic cells (%) ³	Ductular reaction ³	Nodule dropout ³	Cholestasis ³
C ^{*A}	NB [*]	22	M	0/0	0	0	1	1	0	0	0	0	0
GS1 ^B	Slide	<30		0/0	0	0	0	0	0	0	0	0	0
GS2 ^B	Slide	<30		1/1	1	0	1	1	0	0	0	0	0
GS3 ^B	Slide	<30		0/0	1	0	1	1	0	0	0	0	0
GS4 ^B	Slide	<30		2/3	1	0	1	1	0	0	1	0	0
GS5 ^B	Slide	<30		1/1	0	0	1	1	0	0	0	0	0
L ^{*A}	NB [*]	54	M	2/3	1	0	1	1	0	0	0	0	0

B – HBV-positive patients with cirrhosis

CNC	Resect.	52	M	4/6	0	0	1	0	0	0	0	rare	moderate
CYC	Resect.	54	F	4/6	3	0	1	2	0	0	2	0	marked
CYRYC	Resect.	48	M	4/6	3	0	3	1	0	0	3	rare	none
FMC1 ^D	Resect.	67	M	4/6	0	0	0	1	0	0	0	0	0
FMC5 ^P	Resect.	40	M	4/6	0	0	0	0	0	0	1	1	0
FMC6 ^P	Resect.	58	M	4/6	0	0	0	1	0	0	2	1	0
MHC	Resect.	44	M	4/6	3	0	3	3	0	0	2	rare	moderate
NTC	Resect.	48	M	4/6	2	0	1	1	0	0	2	rare	marked
SAOC ^C	Resect.	65	M	4/6	1	0	1	1	0	LCC - 30% SCC - 5%	2	0	none
XAC	Resect.	46	M	4/6	2	0	1	1	0	LCC - 10%	2	0	none

Table 4.1. Continued.

C – HBV-positive patients with fulminant hepatitis

Name ¹	Type ²	Age	Sex	Fibrosis (METAVIR/shak) ³	Interface hepatitis (Ishak A) ³	Centro-lobular dropout (Ishak B) ³	Lobular inflam'n (Ishak C) ³	Portal inflam'n (Ishak D) ³	Steatosis ³	Dysplastic cells (%) ³	Ductular reaction ³	Nodule dropout ³	Cholestasis ³
FMC2 ^D	Resect.	62	M	4/6	2	6	2	2	1	0	3	2	2
FMC3 ^D	Resect.	49	M	4/6	0	6	0	1	0	0	3	2	2
FMC4 ^D	Resect.	68	F	3/4	2	4	3	2	0	0	3	all	2
FMC7 ^D	Resect.	49	F	3/4	2	6	2	2	0	0	3	2	0
HNC	Resect.		F	1/2	4	6	4	4	0	0	3	all	marked
HSC	Resect.	48	M	4/6	1	0	1	1	0	0	3	many	moderate

D – HBV-positive patients without cirrhosis and with concurrent HCC

DGE	Resect.	56	M	2/3	0	0	0	0	0	0	0	0	0
Y2 ^B	Slide		M	2/4	0	0	1	1	0	0	1	0	0
Y3 ^B	Slide		M	2/3	0	0	1	1	0	LCC – 5%	2	0	0
Y4 ^B	Slide		M	1/1	0	0	1	1	0	0	0	0	0
Y5 ^B	Slide		M	3/4	1	0	3	2	1	LCC - 20%	1	0	0
Y6 ^B	Slide		M	3/5	2	0	1	2	0	LCC- 25%	2	0	0

E – HBV-negative control patients

NHL1 ^E	Resect.			0/0	0	0	0	0	0	0	0	0	0
NHL2 ^E	Resect.			0/0	0	0	0	0	0	0	2	0	0
NHL3 ^E	Resect.			0/0	0	0	0	0	0	0	0	0	0

F – Uncharacterised tissues

WNC	Resect.	45	M										
-----	---------	----	---	--	--	--	--	--	--	--	--	--	--

■ = Unknown

Severe diathermy or drying artefact, not interpretable.

¹ Patient liver tissues were sourced from clinical care centre based at: the Royal Adelaide Hospital (A); University of Washington (B); Centenary Institute (C); SA Liver Transplant Unit, Flinders Medical Centre (D); and SA Pathology (E).

² Available tissues were in the form of liver resection (Resect.), needle biopsy (NB) or ethanol-fixed slide-mounted liver tissue (Slide).

³ Liver histology was interpreted by Associate Professor Andrew Clouston (University of Queensland) and, based on the histological findings, HBV positive patients were assigned to 4 groups: Group A - early-stage HBV infection (7 patients); Group B - late-stage HBV infection with cirrhosis (10 patients); Group C - late-stage fulminant hepatitis (6 patients); and Group D - late-stage HBV infection without cirrhosis and with concurrent HCC (6 patients). Histology was also interpreted for HBV-negative patients, assigned to Group E. Histology could not be interpreted for Group F - Patient WN. METAVIR and Ishak scores for fibrosis are scored out of 4 and 6 respectively, with high scores representing greater progression to cirrhosis.

* For some needle biopsy samples, insufficient material was available for histological interpretation. Histological scores for these samples (highlighted in yellow) were assigned using information taken from clinical reports provided by each clinical care centre.

Each tissue was scored using two established liver histological scoring methods: METAVIR and Ishak staging (Ishak, Baptista et al. 1995; Bedossa and Poynard 1996). These systems rank the progression of liver tissues towards cirrhosis with a score out of 4 and 6, respectively. High scores represent greater progression, with respective scores of 4 and 6 in the METAVIR and Ishak systems representing complete and definite cirrhosis.

The activity of the HBV infection was also measured by the Ishak scoring system using inflammation and hepatocyte dropout as markers. Some tissues contained evidence of lymphocytes infiltrating into periportal sinusoids, indicating inflammation, as shown in Figure 4.1. In 6 patients, loss of large regions of hepatocytes and evidence of ductular reaction was observed, as shown in Figure 4.2. This indicated that these patients had evidence of fulminant hepatitis (FH).

Tissues from patients CY and HN contained brown pigmentation of unknown origin mainly localised in the cytoplasm of hepatocytes, shown in Figures 4.3A and 4.3B. The granular pigmentation could potentially be a build-up of iron or bile due to cholestasis. Perls' Prussian Blue iron stain and Fouchet's bile pigment stain were used on formalin-fixed tissues to identify of the brown pigmentation. In the tissue from patient CY, a positive Perls' Prussian Blue stain (Figure 4.3C) suggested that the pigmentation was due to iron deposition. Iron deposition is a common occurrence in chronic HBV patients, having been found in ~37% of patients in late-stage disease (Sebastiani, Tempesta et al. 2012). In the tissue from patient HN, the pigment was negative for both Perls' Prussian Blue stain and Fouchet's bile stain (Figure 4.3D). The identity of the pigmentation in tissue of patient HN remains unknown.

Hepatocytes with small and large cell changes (SCCs and LCCs) were observed in some liver tissue sections, as summarised in Table 4.1. As mentioned in Section 1.7, SSC lesions indicate preneoplastic progression (Watanabe, Okita et al. 1983; Le Bail, Bernard et al. 1997). On the other hand, LCC are less specific but more common in diverse disease states, such as HBV infection, cholestasis and HCC (Le Bail, Bernard et al. 1997; Su, Benner et al. 1997). Thus, the presence of SCC and LCC indicate liver disease in these patients.

Tissue from patient SAAO displayed hepatocytes with both SCC and LCC. Slide-mounted tissue from patients XA, Y3, Y4 and Y6 contained only LCC. In the Y-series of patients, foci of hepatocytes with LCC were localised and are shown in Figure 4.4.

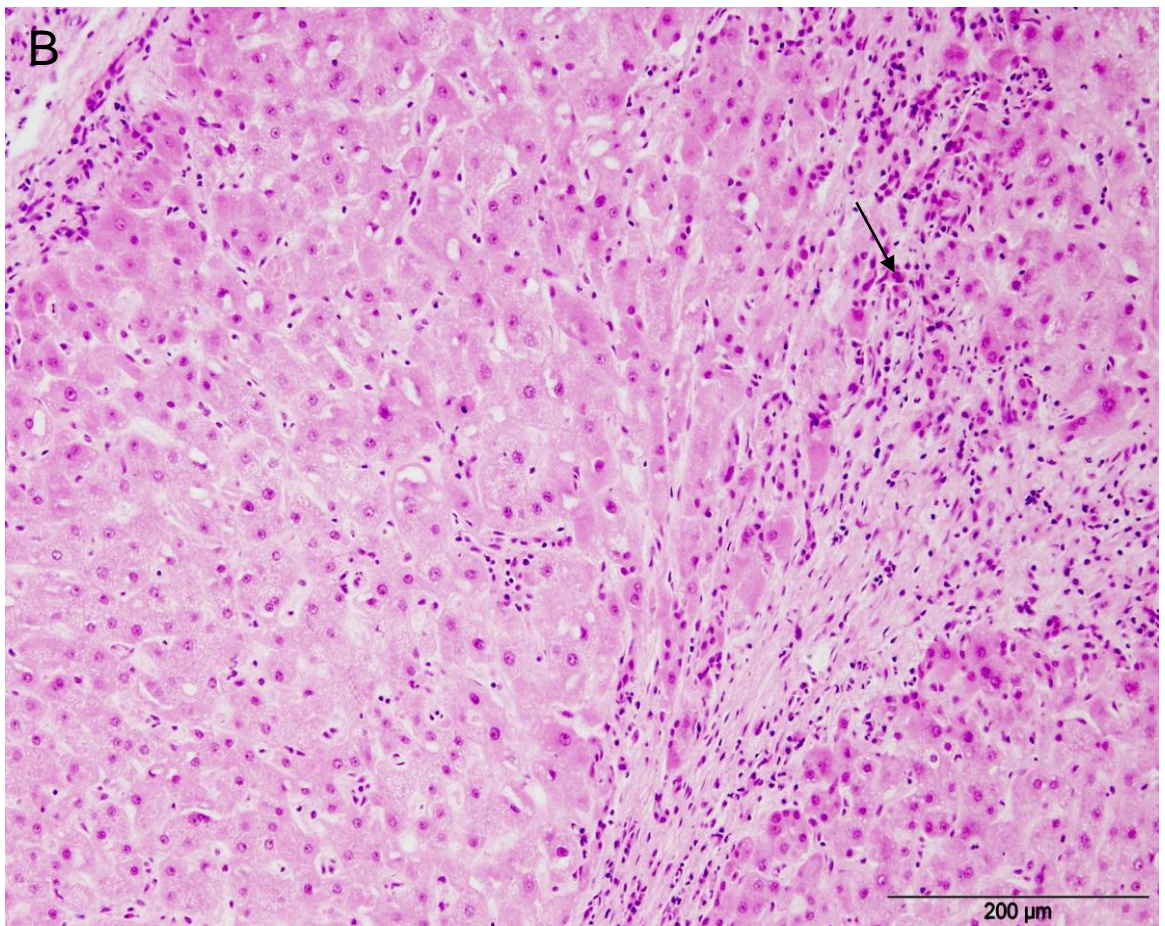
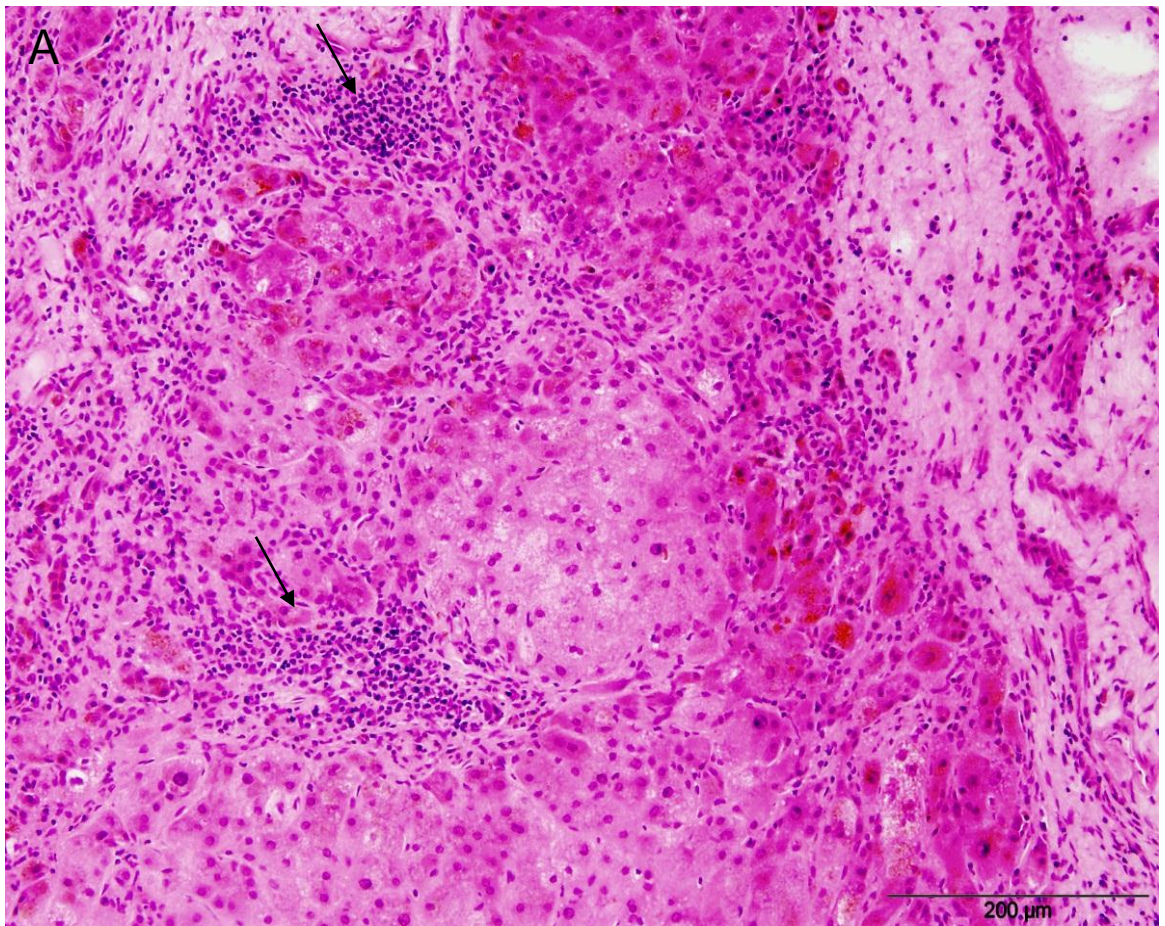


Figure 4.1. Inflammation of the liver. Sections of liver tissue from from patients CY (4.1A) and CYRY (4.1B) showed marked inflammation. Mononuclear infiltration is shown by arrows. Stained with H&E. Magnification 10x. Scale bars = 200 μ m.

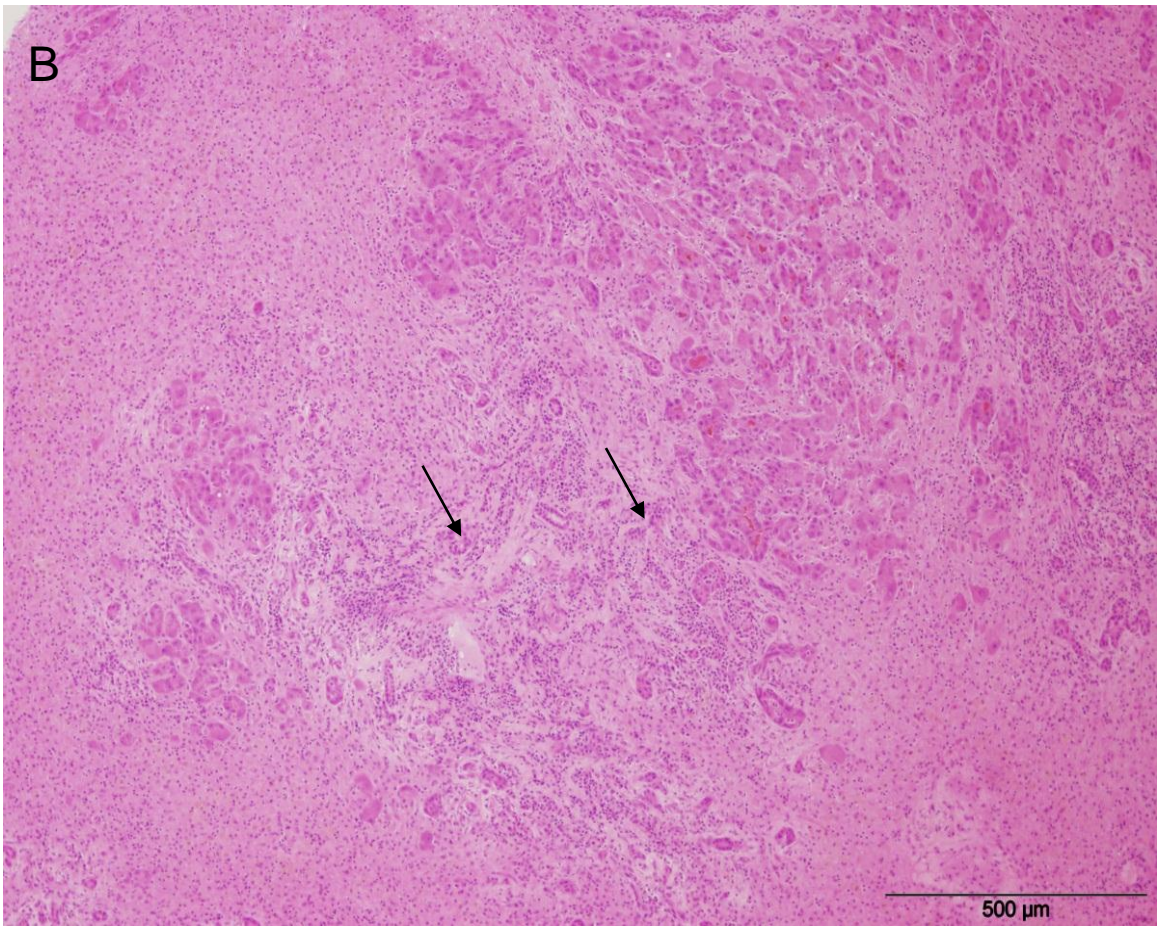
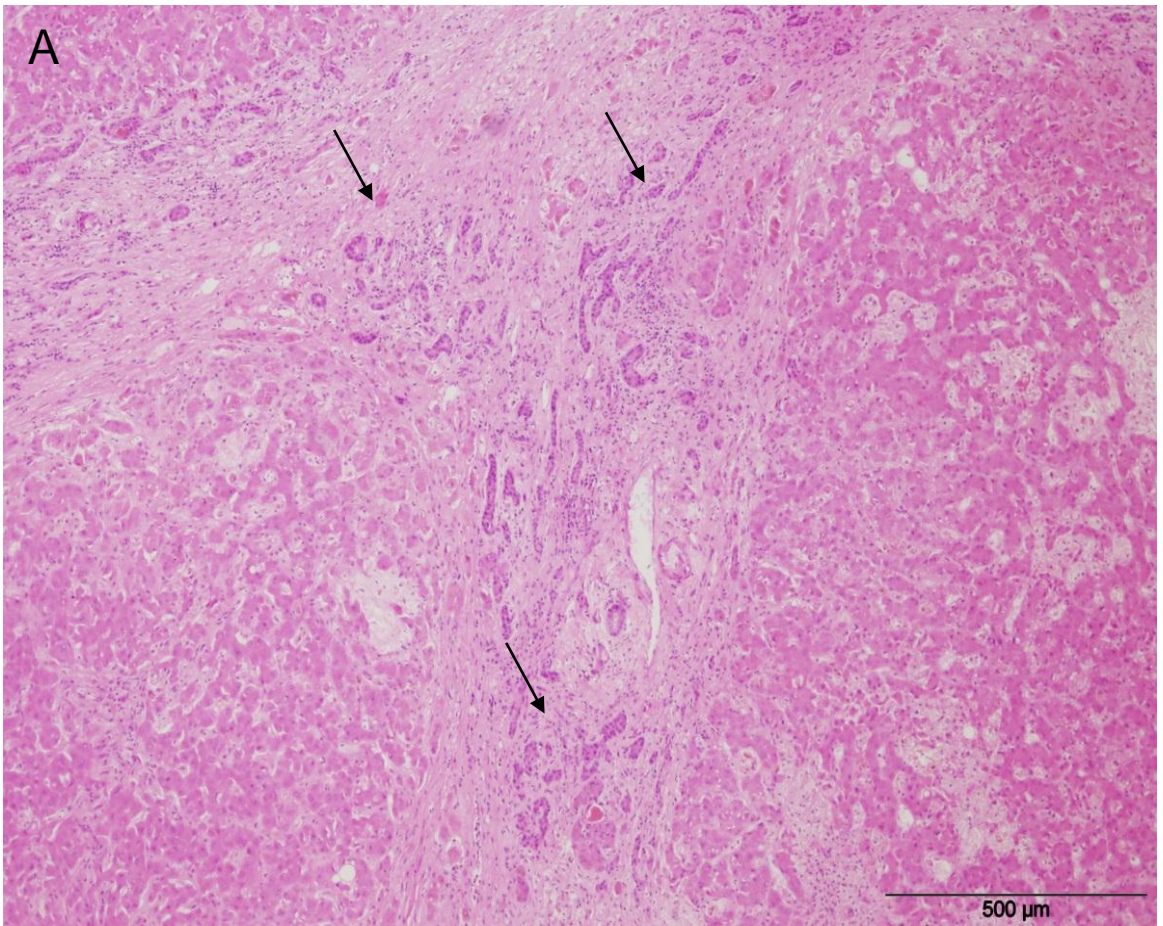
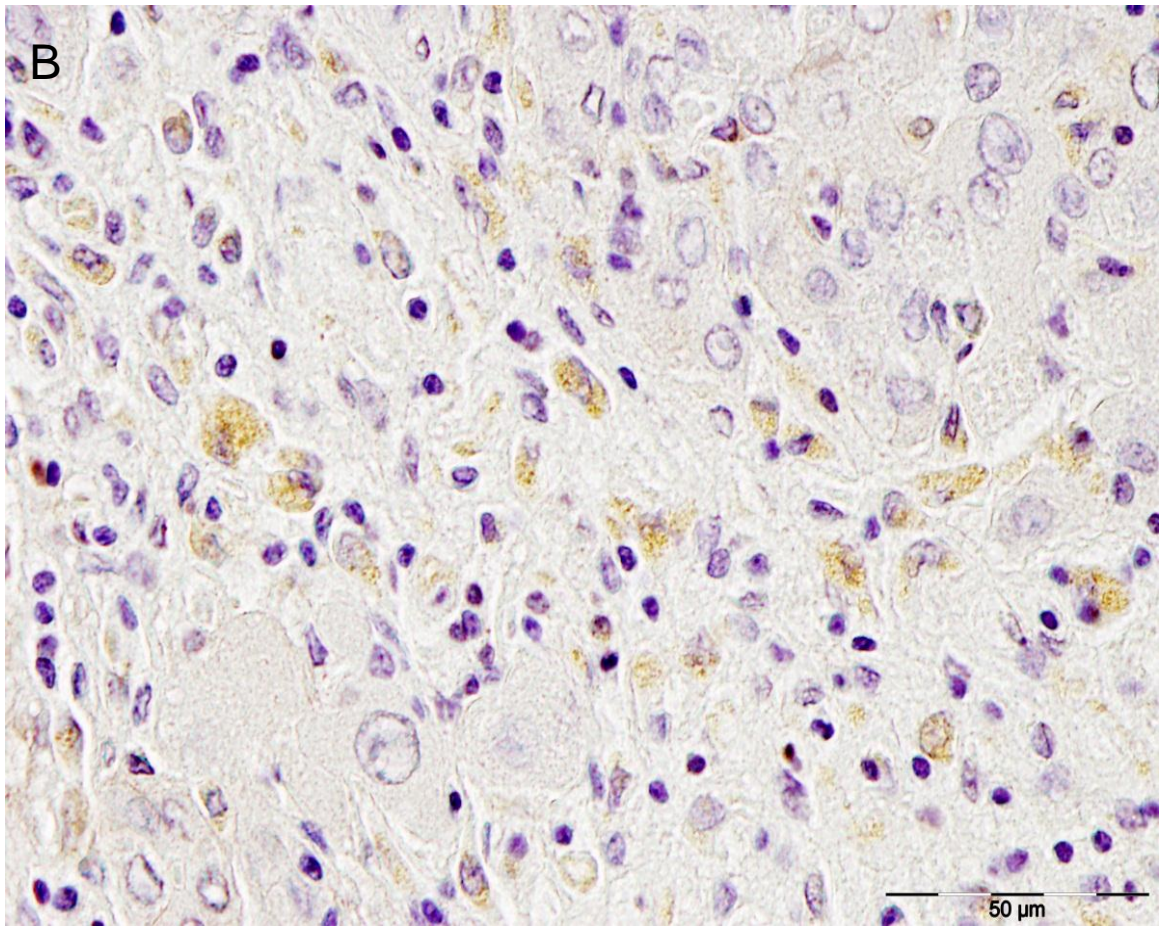
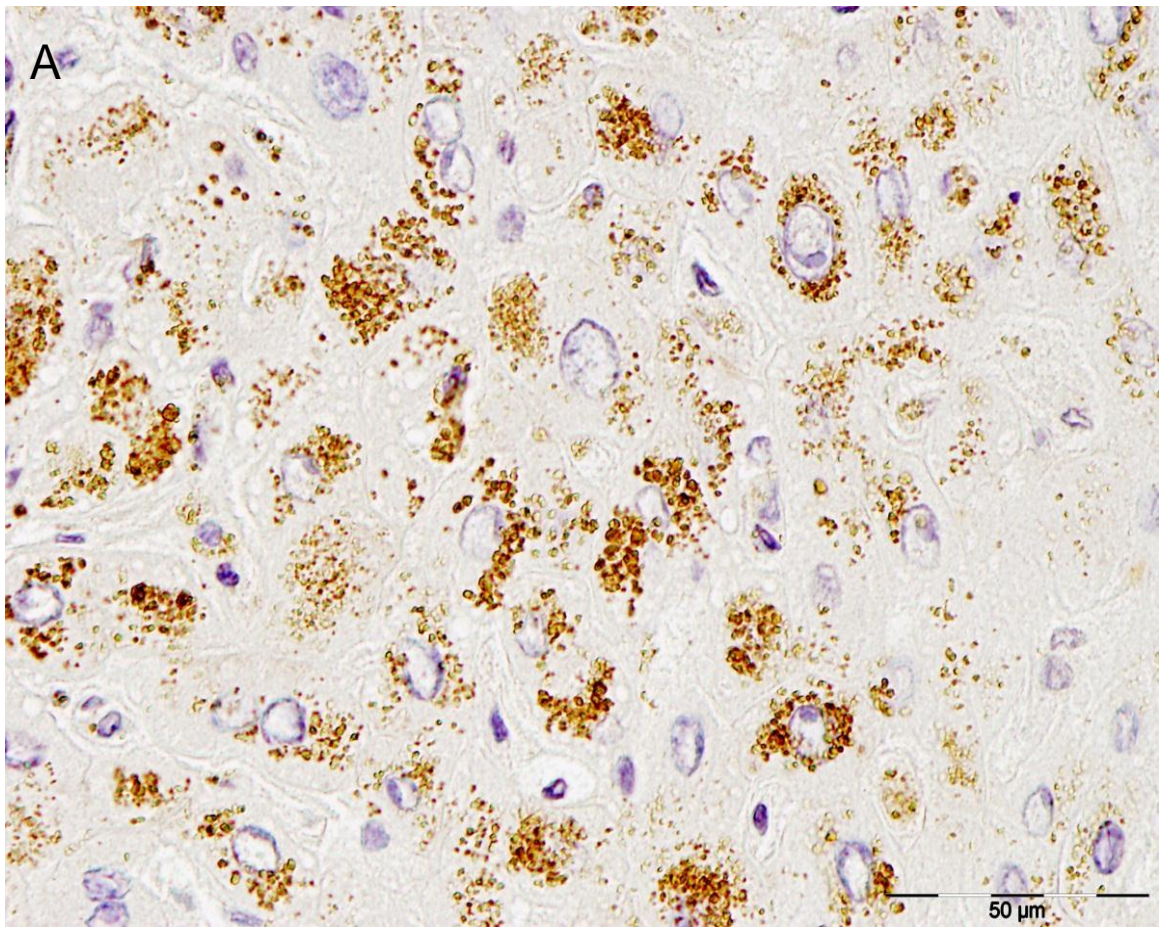


Figure 4.2. Fulminant hepatitis. Arrows indicate bile duct cell proliferation that occurs in association with hepatocyte loss in liver sections from Patients HN (4.2A) and HS (4.2B). Few hepatocytes remain, particularly in Patient HS (4.2B), where most cells are stromal cells. In Patient HN, cirrhosis is also evident. Stained with H&E. Magnification 4x. Scale bars = 500 μ m.



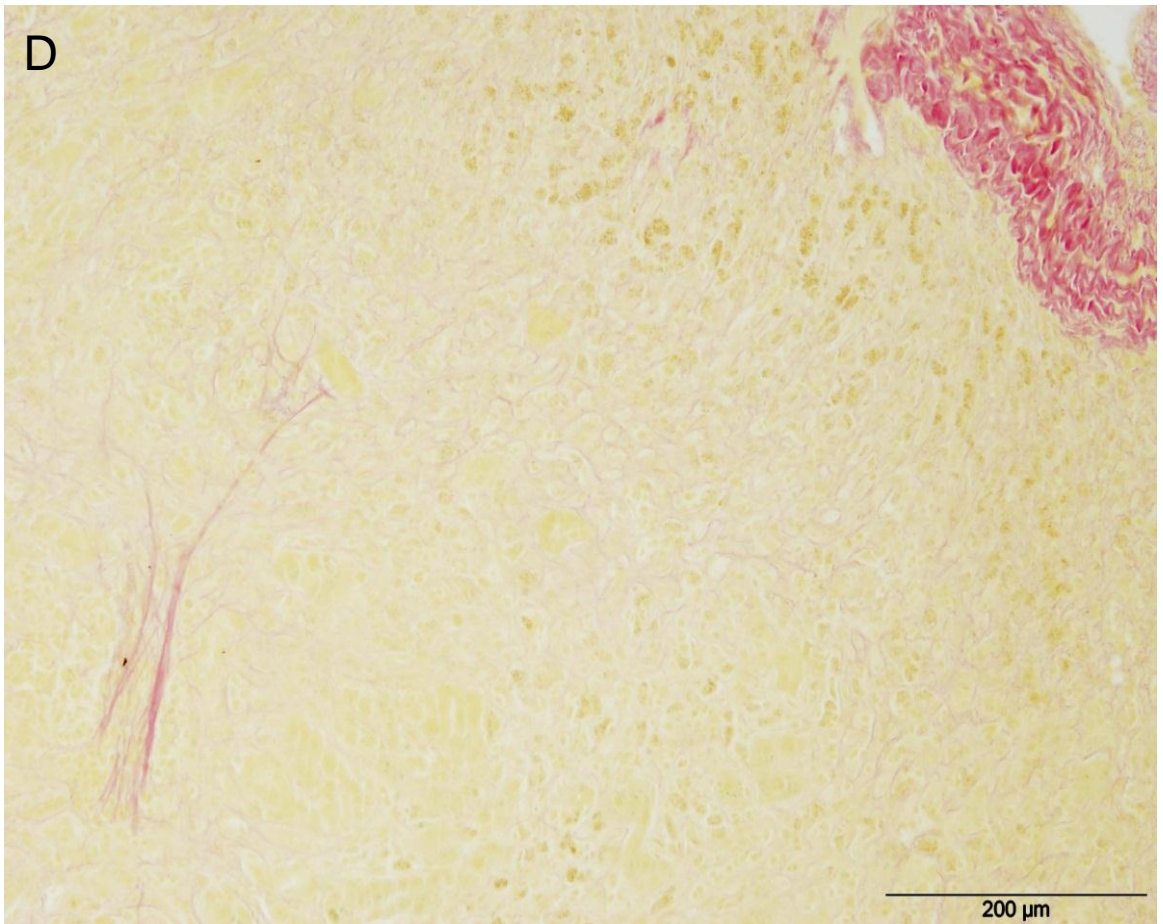
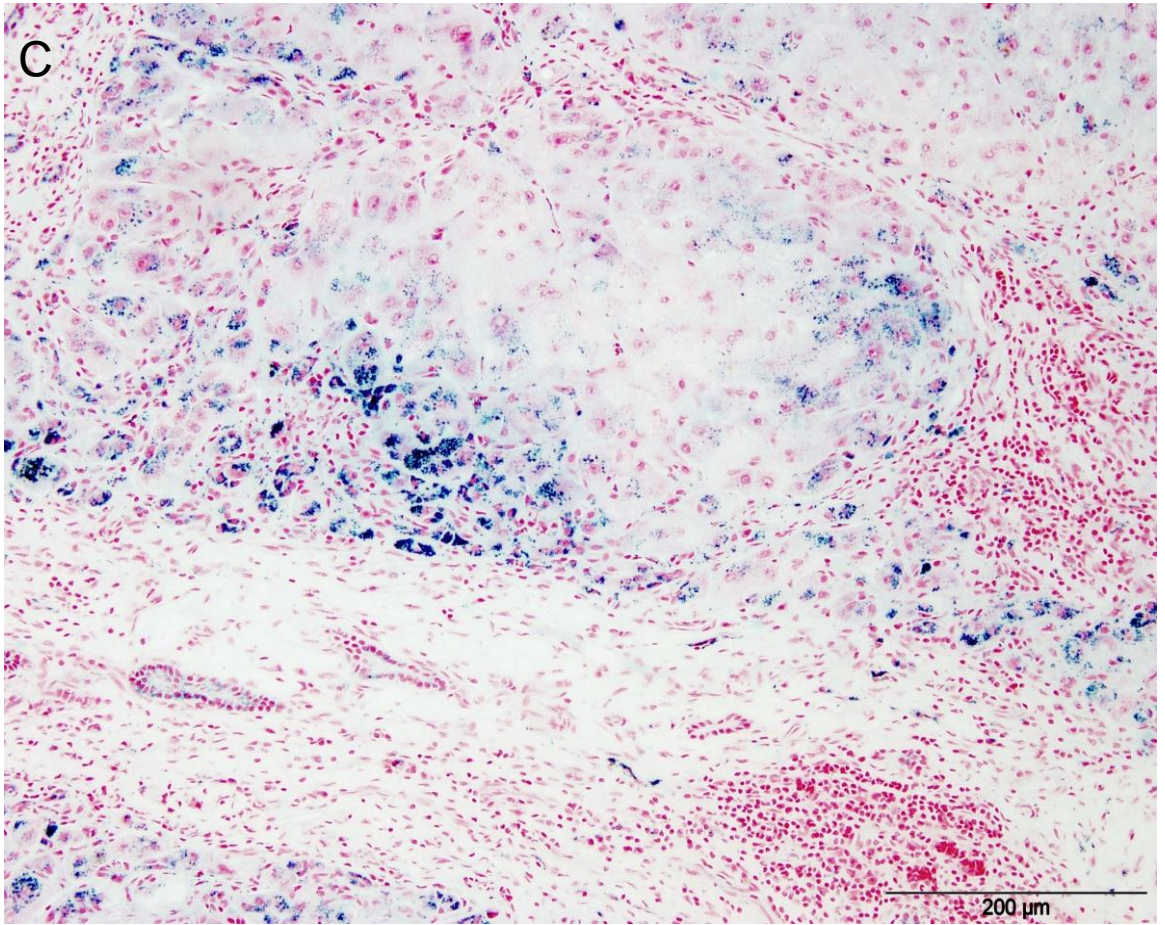


Figure 4.3. Pigmentation of the liver. Liver sections from Patients CY (4.3A) and HN (4.3B), when counterstained with haematoxylin, showed pigmentation of hepatocytes. The pigment present in the liver of Patient CY stained positive (blue) with Perls' Iron staining (4.3C). The pigment present in the liver of Patient HN stained negative with Perl's Iron staining and with Fouchet's bilirubin stain (4.3D). For 4.3A and B: Magnification = 40x. Scale bars = 50 μ m. For 4.3C and D: Magnification = 10x. Scale bars = 200 μ m.

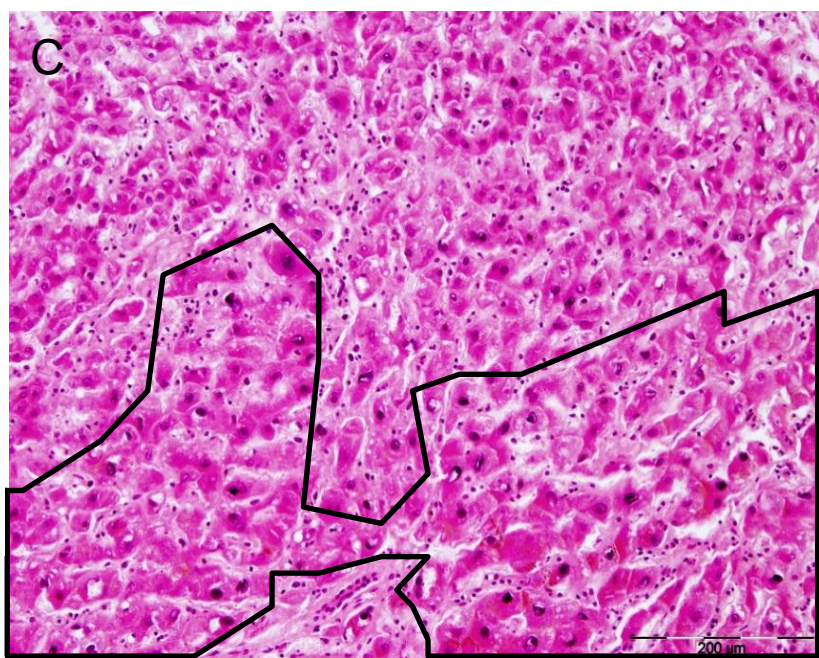
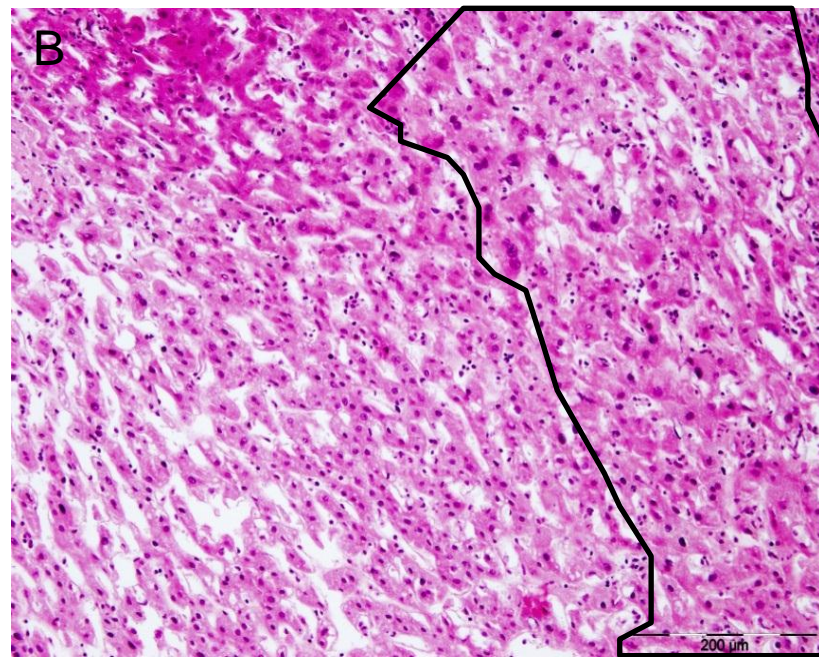
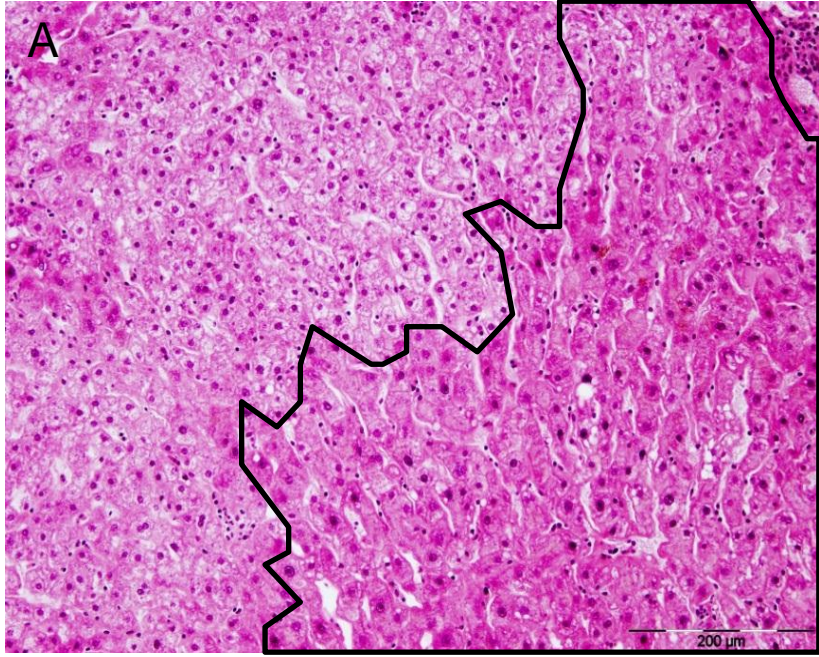


Figure 4.4. Localisation of hepatocytes with large cell changes (LCC) of in the liver tissues of Patients Y3 (A), Y4 (B) and Y6 (C). Areas composed of hepatocytes with LCC were identified by A/Prof. Andrew Clouston and are bordered in black. Sections were stained with H&E. Magnification 10x. Scale bars = 200 μ m.

As a result of this histological analysis and clinical data, the HBV-positive patients were split into four Groups:

- A. Patients without HBV-associated disease, with criteria being low histological activity, and METAVIR and Ishak scores of ≤ 2 and ≤ 3 , respectively, indicating no to rare formation of fibrotic septa (7 patients);
- B. Cirrhotic patients, with criteria being METAVIR and Ishak scores of 4 and 6, respectively (10 patients);
- C. Fulminant hepatitis patients, with criteria being severe ductular reaction (a score of 3 out of 3) and widespread dropout of hepatocyte nodules (6 patients) and;
- D. Non-cirrhotic patients with concurrent HCC, with criteria being a clinical background of HCC, and METAVIR and Ishak scores of ≤ 3 and ≤ 5 , respectively (6 patients).

One patient (Patient WN) could not be classified as severe diathermy or drying artefacts had rendered histology unable to be interpreted.

Virological data for each patient is summarised in Table 4.2. Only 7 patients were treated with antiviral therapy immediately prior to resection or biopsy. About half of the HBV-positive patients (16 out of 30) were known to be serum HBeAg-negative, while only 9 out of 30 patients were known to be HBeAg-positive. Details on HBeAg status were not available for the remaining 5 patients. No information was available for levels of serum HBsAg, though all patients are HBV carriers and thus, by definition, have serum HBsAg.

4.3.2 HBV DNA analysis

The HBV DNA sequence of the predominant strain present in each patient was determined as described in Section 2.12. The HBV DNA sequences for the HBV strains associated with Patients Y2 to Y6 were previously determined by Mason *et al.* (listed as Patients Y1-5 in (Mason et al. 2010)). In all other patients, DNA was extracted from liver tissue and HBV DNA was amplified and then sequenced in the regions upstream from the probable right hand virus-cell junction (nt1672-2026, according to numbering in Genbank Accession #AB241115). Sequence data was aligned and analysed by BLAST and numbering was based on a reference HBV DNA sequence (Genbank Accession #AB241115). HBV genotypes were

Table 4.2. Summary of virological data from patients with chronic HBV infection. Groups are identical to those listed in Table 4.1.

A – Early-stage HBV infection

Name	Antiviral therapy ¹	Serum HBsAg ¹	Serum HBeAg ¹	HBV Genotype ²	Liver HBV DNA/BG (qPCR) ³	%Hepatocytes positive for HBV DNA (ISH) ³	Hepatocytes positive for HBsAg ⁴	Hepatocytes positive for HBeAg ⁴
C		+	-	D	1.6		++++	-
GS1	-	+	+	C	7000		+	+
GS2	-	+	+	C	15000		+	+
GS3	-	+	+	A	5		+	-
GS4	-	+	+	D	1300		+++	+++
GS5	-	+	+	A	240		+	-
L		+	+	C	260		+	+

B – HBV-positive patients with cirrhosis

CN		+	+	D	ND	ND	-	-
CY				C	290	50	+	-
CYRY	-	+	-	A	370	80	+++	+++
FMC1	+	+	-	B	39		-	-
FMC5	+	+	+	D	2		+	-
FMC6	+	+	-	B	830		+	-
MH	-	+	-	A	0.12	ND	++	-
NT	-			A	ND	ND	-	-
SAAO	-	+	-	A	0.035	ND	-	-
XA	-	+	-	C	0.25	ND	++	-

Table 4.2. Continued

C – HBV-positive patients with fulminant hepatitis

Name	Antiviral therapy ¹	Serum HBsAg ¹	Serum HBeAg ¹	HBV Genotype ²	Liver HBV DNA/BG (qPCR) ³	%Hepatocytes positive for HBV DNA (ISH) ³	Hepatocytes positive for HBsAg ⁴	Hepatocytes positive for HBeAg ⁴
FMC2	+	+	-	A	1.4		-	-
FMC3	+	+	-	B	4800		-	-
FMC4	+	+	-	B	160		++	-
FMC7	+	+	-	C	22		+	-
HN		-	-	D	0.33	ND	++	-
HS		+	+	D	2.3	ND	++	-

D – HBV-positive patients without cirrhosis and with concurrent HCC

DG	-	+	-	C	0.13	ND	+	-
Y2	-	+	-	B	1.7 [#]	ND	+	-
Y3	-	+	-	B	2.4 [#]	ND	+	-
Y4	-	+	-	B	2.9 [#]	ND	+++	-
Y5	-	+	-	C	2.6 [#]	ND	+	-
Y6	-	+	-	B	1.4 [#]	ND	+	-

E – HBV-negative control patients

NHL1	-	-	-		ND	ND	-	-
NHL2	-	-	-		ND	ND	-	-
NHL3	-	-	-		ND	ND	-	-

F – Uncharacterised tissues

WN				A	ND	Severe diathermy or drying artefact, not interpretable.		
----	--	--	--	---	----	---	--	--

¹ The use of antiviral therapy immediately prior to resection/biopsy and serum HBsAg/HBeAg status were elucidated for each patient from clinical notes. Unknown values are represented by grey-filled squares.

² Total DNA was extracted from liver tissues and HBV DNA was amplified by specific primers, as described in Section 2.12. Products were sequenced and genotypes were elucidated by BLAST alignment to HBV sequences in the Genbank database.

³ ND = Not detected. Unknown values are represented by grey-filled squares. *In situ* hybridisation was not done on all tissue samples due to restricted amounts of tissue.

⁴ HBsAg and HBcAg was detected by immunohistochemistry. The percentages of antigen-positive hepatocytes was estimated by cell-counts of >5 fields of view at 10x magnification and are represented as follows:

- = 0%; + = <25%; ++ = 25-50%; +++ = 50-75%; ++++ = >75%

Copy numbers of HBV DNA/cell in these tissues were previously published by Mason *et al.* (2010).

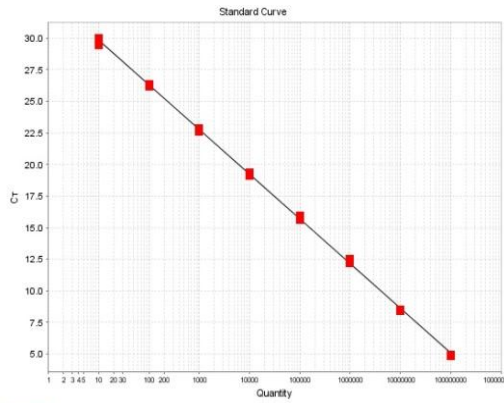
identified for each patient and are summarised in Table 4.2. The HBV DNA sequence determined for each patient is listed in Appendix 9.7.

The level HBV DNA per cell in liver DNA extracts was determined by qPCR as described in Section 2.13. The sequence of the primers used to amplify the single-copy cellular gene β -globin (BG, based on GenBank Accession #L26462) and HBV DNA are listed in Table 2.1. The standard curves and melt curves for the two qPCR assays are shown in Figure 4.5. The standard curves show a sensitivity of ≤ 10 copies per reaction and the melt curves show that single DNA products were specifically amplified in each PCR reaction. The quantity of HBV DNA was normalised using BG, as summarised in Table 4.2. The HBV DNA levels varied amongst patients from undetectable to 1.5×10^4 copies of HBV DNA per copy of BG. Liver tissue from HBeAg-positive patients (mean = 2600 HBV DNA copies/BG; range = Not detectable to 15000) did not contain significantly greater levels of HBV DNA compared to HBeAg-negative patients (mean = 560 HBV DNA copies/BG; range = Not detectable to 4800), after analysis with a Student's T-test ($p > 0.05$). The limit of detection for the qPCR assay was ~ 1 copy of HBV DNA per 200 cells.

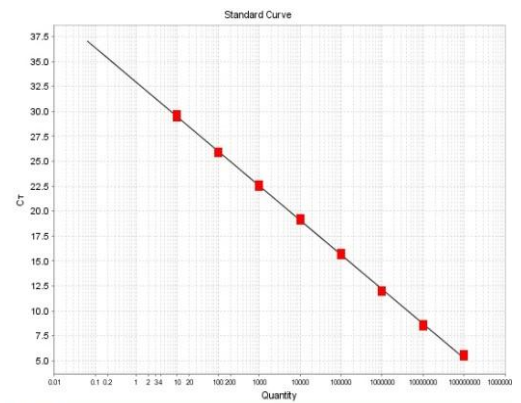
4.3.3 Detection and localisation of hepatocytes containing HBV DNA by *in situ* hybridisation

HBV DNA was localised in ethanol:acetic acid (EAA)-fixed and sectioned liver tissue using *in situ* hybridisation as described in Section 2.6. The DIG UTP-labelled HBV DNA probe is expected to detect all genotypes of HBV as the probe is produced by nick-translation of the entire HBV genome, which would have high homology to at least some regions of all genotypes. However, this technique has relatively low sensitivity with a limit of detection of 40-160 kb of target DNA per cell, i.e. on the order of 10-50 copies of HBV per cell (Walboomers, Melchers et al. 1988; Speel, Ramaekers et al. 1995). HBV cccDNA is tightly wound around histones and therefore may be poorly accessible to DNA probes, decreasing the sensitivity of the *in situ* hybridisation assay to a greater extent (Bock, Schwinn et al. 2001). Thus, only hepatocytes that support HBV replication with high copy numbers of HBV RI DNA, rcDNA and cccDNA genomes will be detected.

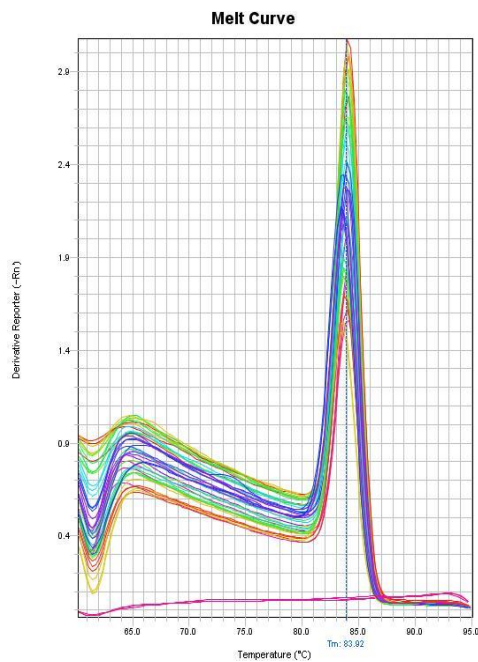
As shown in Table 4.2, the majority of tissues had undetectable levels of HBV DNA when tested using *in situ* hybridisation. This suggested that low levels of HBV replication were



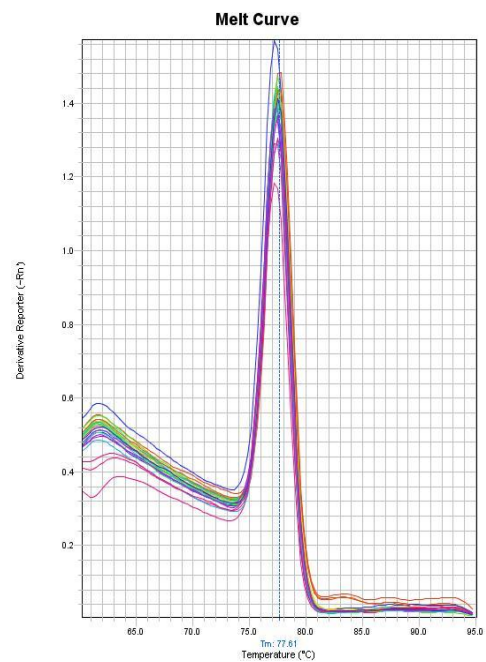
$R^2 > 0.999$
Efficiency=91%



$R^2 = 1$
Efficiency=97%



$T_m = 84^\circ\text{C}$



$T_m = 77.5^\circ\text{C}$

Figure 4.5. β -globin (BG) and HBV DNA were detected by qPCR in total DNA extracts from liver tissue fragments. The sequence of the primers used for amplifying BG and HBV DNA are listed in Table 2.1. Due to the number of samples, two separate qPCR assays were run as described in Section 2.13. Each DNA extract was run in triplicate. The BG gene was subcloned into a pBlueBac vector, as described in Section 2.13.1, to form pBG. pBG and the HBV DNA-containing plasmid pBB4.5HBV1.3 were used to produce the standard curves. As shown in Figures 4.5A and 4.5B, plasmid DNA standards were diluted from 10^1 - 10^8 copies per reaction. The melt curves in Figure 4.5C and 4.5D show that only single products were amplified in all samples.

occurring in these tissues. Only low levels of HBV DNA were detected in these tissues by qPCR. The percentage of hepatocytes positive for HBV DNA by *in situ* hybridisation in tissue sections was consistent with the qPCR DNA results. Also shown in Table 4.2, patient tissues that had higher levels of HBV DNA per cell when tested by qPCR contained a higher percentage of hepatocytes positive by *in situ* hybridisation with significant correlation (analysed by Pearson's correlation analysis, $r^2 > 0.99$).

In the tissues where HBV DNA was detected by *in situ* hybridisation, foci of hepatocytes positive for HBV DNA and hepatocytes negative for HBV DNA were detected, as shown in Figure 4.6. This suggests that there is heterogeneity in the hepatocyte population of patients with chronic HBV infection with respect to ability to support HBV replication, consistent with previous literature (Gowans and Burrell 1985; Gowans, Burrell et al. 1985).

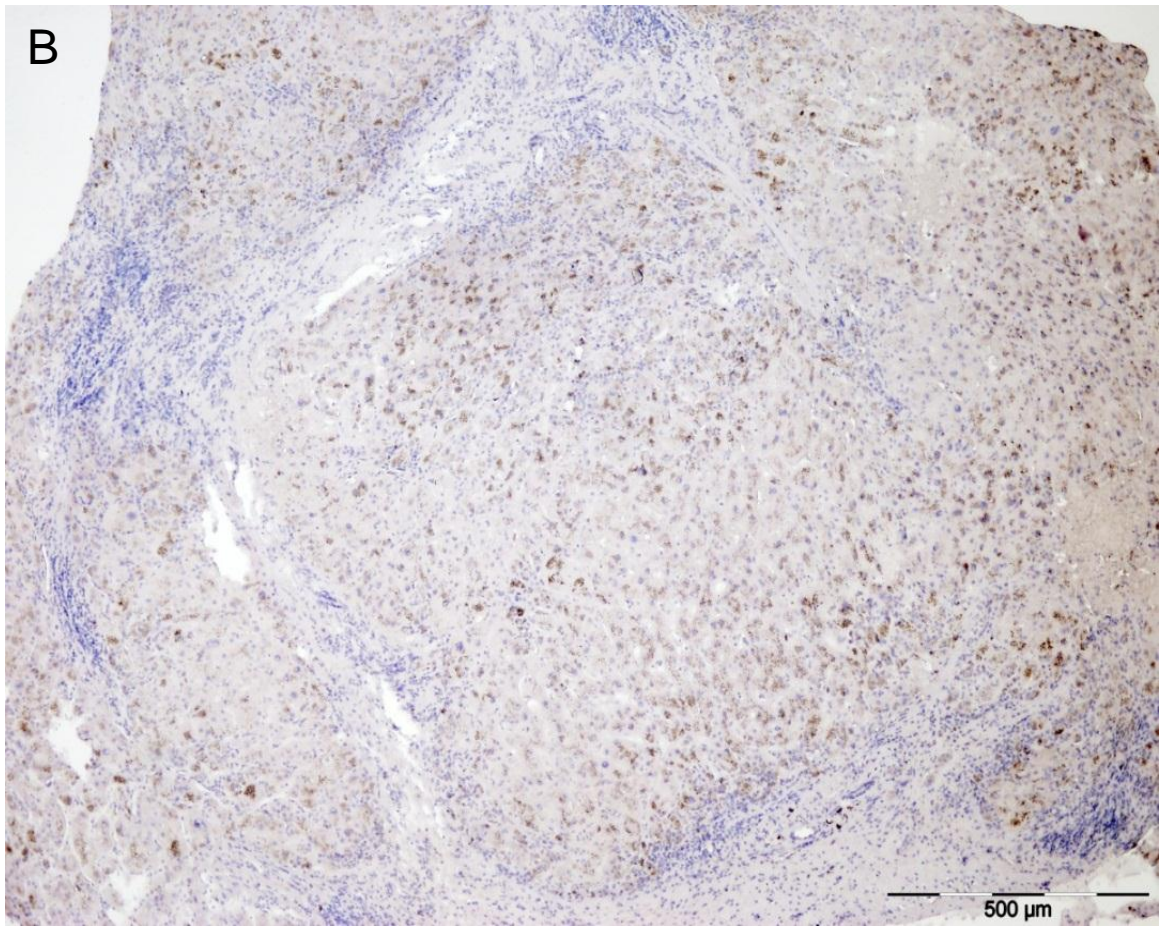
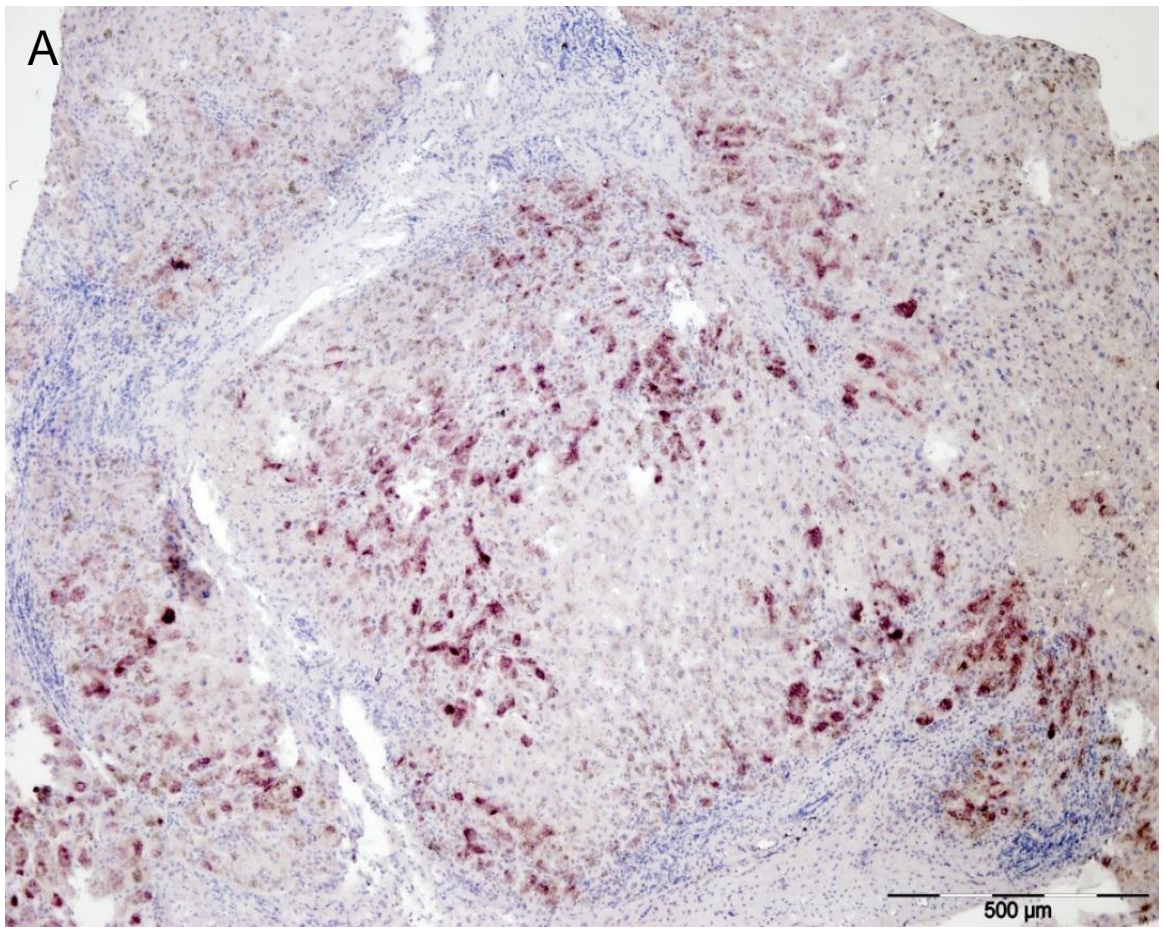
Furthermore, the intracellular distribution of the HBV DNA was generally in the cytoplasm of hepatocytes, and not in the nucleus. This observation is consistent with previous reports of HBV DNA replication occurring in the cell cytoplasm and a failure to detect lower copy numbers of cccDNA in the nucleus by *in situ* hybridisation (Gowans and Burrell 1985; Gowans, Burrell et al. 1985).

4.3.4 Detection of HBV antigens by IHC

The percentage of hepatocytes expressing HBcAg and HBsAg was determined in EAA-fixed patient tissues by IHC and the results are summarised in Table 4.2.

Consistent with the *in situ* hybridisation data, hepatocytes were found to be heterogeneous in their ability to express HBV antigens. Two patterns in HBsAg staining were observed, as shown in Figure 4.7: diffusely scattered hepatocytes with high levels of cytoplasmic HBsAg; and clusters or foci of hepatocytes with high levels of HBsAg. Furthermore, foci of hepatocytes that did not express detectable levels of HBsAg were also observed.

The distribution of HBsAg with antigen-positive hepatocytes was generally cytoplasmic or membranous. However, in HBcAg- positive hepatocytes, staining was observed in the nucleus and cytoplasm. These intracellular distributions of HBV antigens are consistent with previous reports (Gowans and Burrell 1985; Gowans, Burrell et al. 1985).



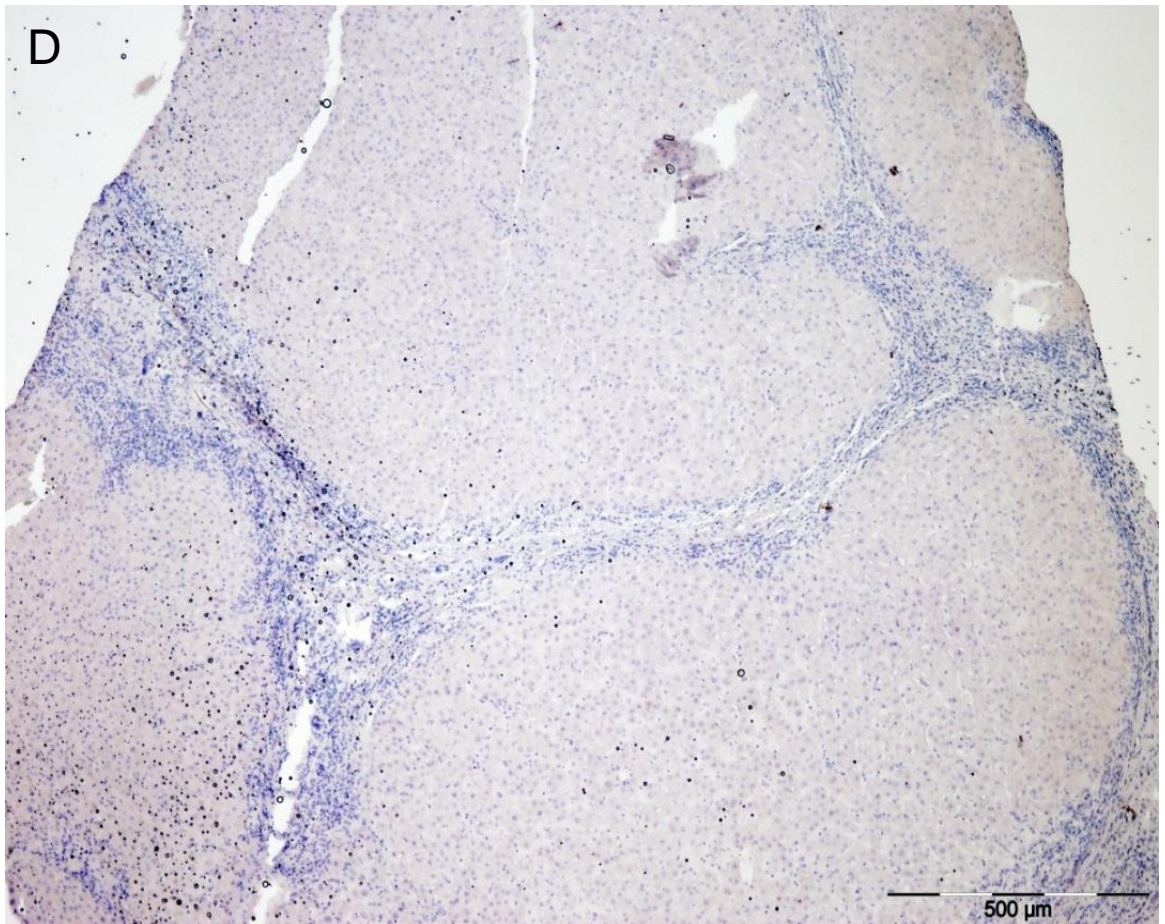
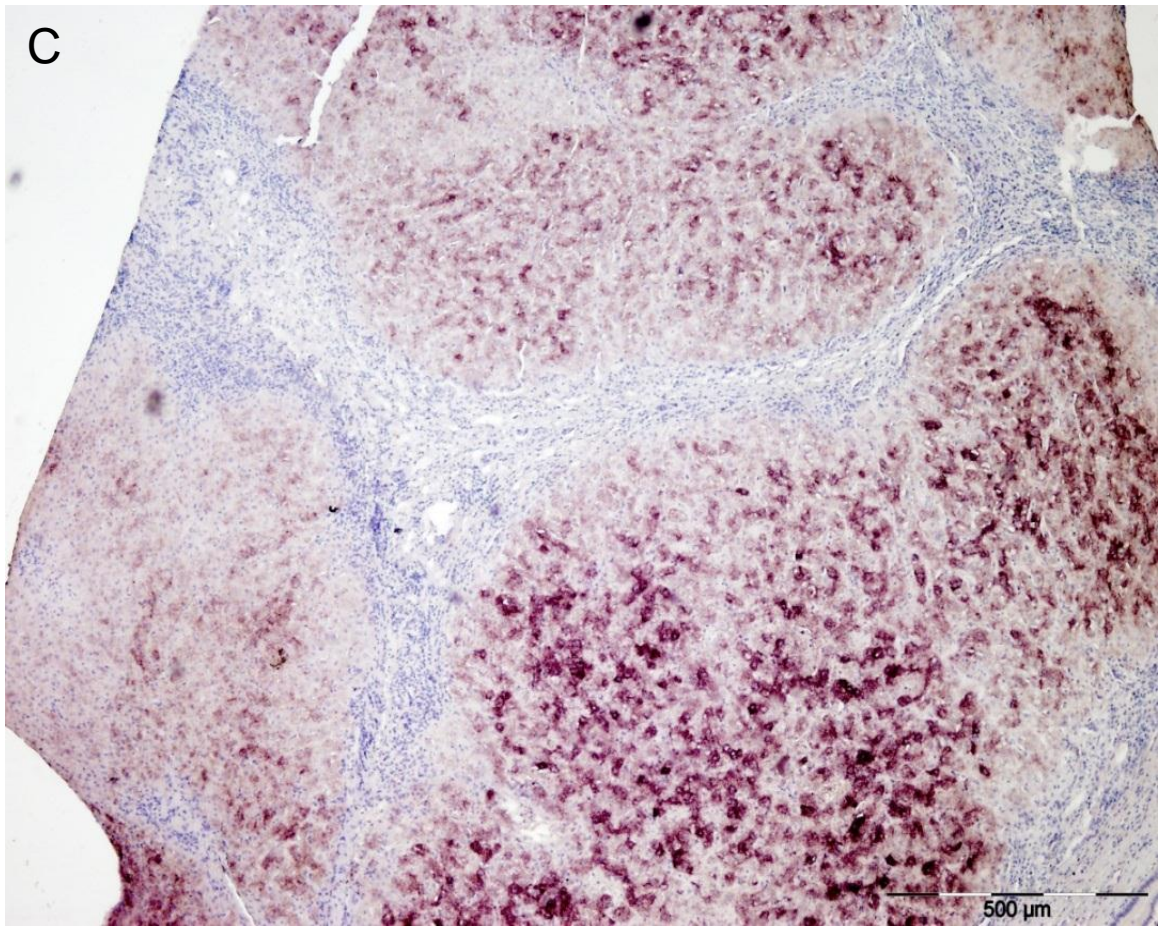


Figure 4.6. Localisation of hepatocytes supporting HBV DNA replication in liver tissue sections. HBV DNA was detected by *in situ* hybridisation using a DIG-UTP-labelled DNA probe as described in Section 2.6. HBV DNA (brown) was detected in the cytoplasm of hepatocytes with the heterogeneous distribution in liver sections from Patients CY (4.6A) and CYRY (4.6C). Hybridisation negative controls (4.6B and 4.6D) were probed using a labelled DIG-UTP-labelled pBlueBac plasmid backbone. The pigmentation evident in the negative control of the section of Patient CY (4.6B) is that previously discussed in Figure 4.3. Magnification 4x. Scale bars = 500 μ m.

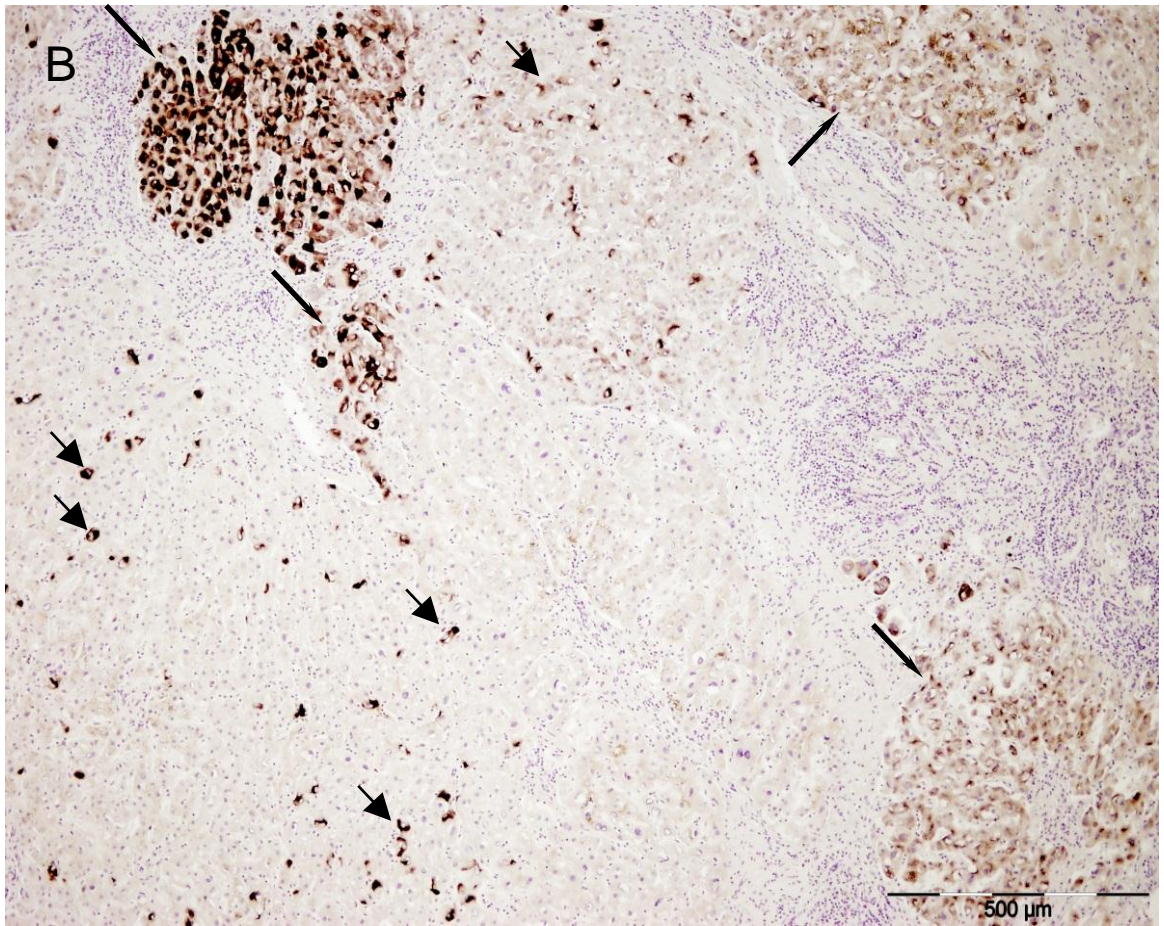
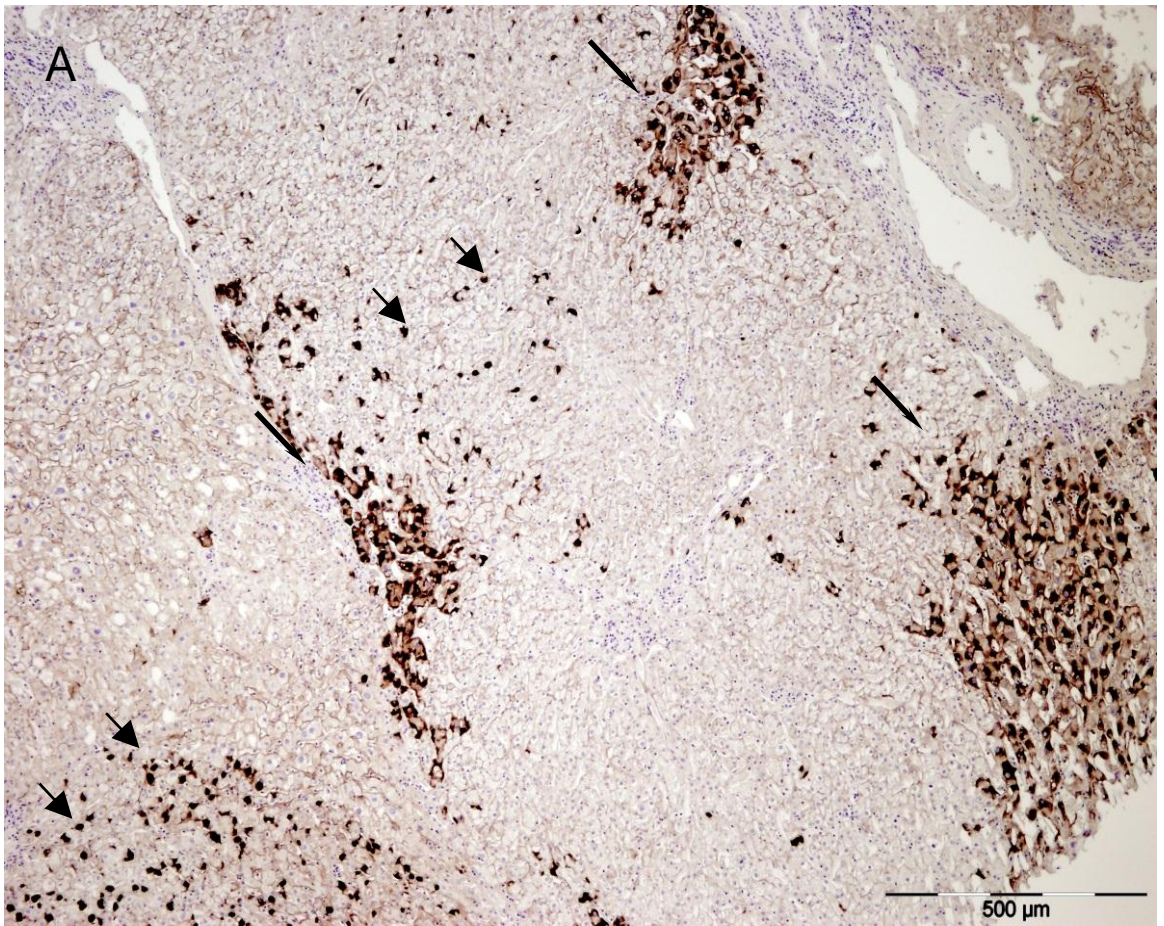


Figure 4.7. Examination of the distribution of HBsAg-positive hepatocytes in liver sections of Patients Y5 (A) and Y6 (B). HBsAg was detected by immunohistochemistry as described in Section 2.4. Two patterns in HBsAg-hepatocytes were observed: diffusely scattered hepatocytes with high levels of HBsAg (short, thick arrows); and clusters or foci of hepatocytes with high levels of HBsAg (long, thin arrows). Magnification 4x. Scale bars = 500 μ m.

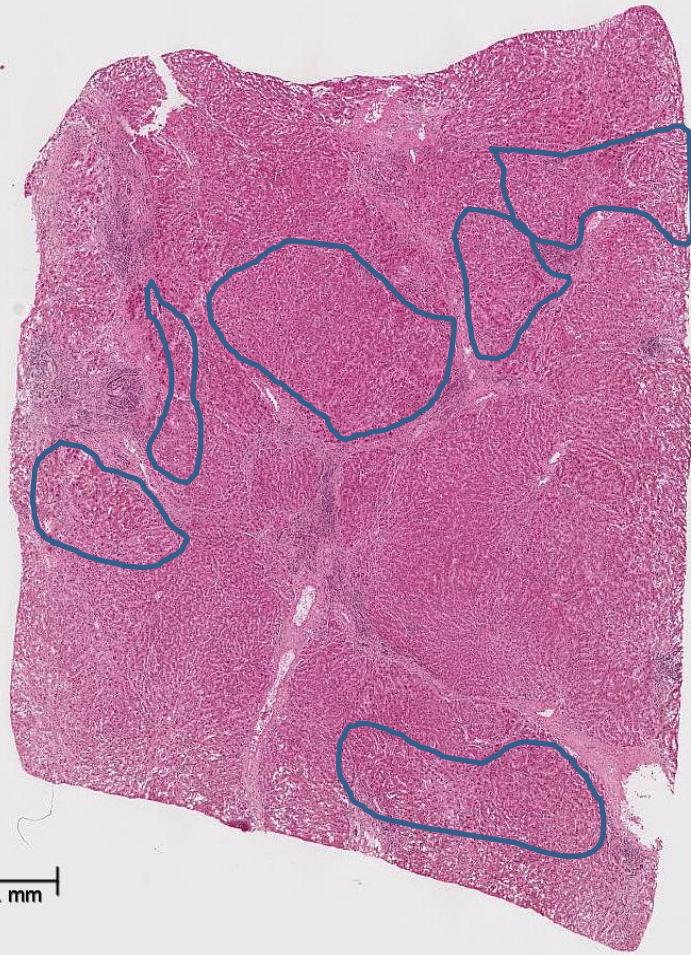
Hepatocytes positive for HBcAg and HBsAg were detected in the liver of 5 and 24 of the 30 patients, respectively. This suggests that staining for HBsAg was more sensitive and/or that HBsAg is produced in infected cells more readily than HBcAg. Previous studies investigating the relative activities of HBV transcriptional promoters in the expression of luciferase determine that the activity of the HBsAg promoter is similar to or stronger than that of the HBcAg promoter (Raney, Milich et al. 1990) . However, the activity of these promoters can be altered by inducible transcription factors (Lin, Li et al. 2003; Ramiere, Scholtes et al. 2008). Thus, due to variable environmental- and patient-specific factors, the relative activity of the HBV ORF promoters remains unknown.

The level of expression may also be different due to integrated HBV DNA acting as a transcriptional template for HBsAg expression. When HBV DNA is integrated in the host-cell genome, the surface ORF is preserved, but the core ORF is separated from its promoter (Bill and Summers 2004). Thus, hepatocytes containing high levels of HBsAg but not HBcAg may be using integrated HBV DNA to transcribe HBsAg.

Alternatively, the HBsAg-specific IHC may be detecting only GGH, which are hepatocytes that have accumulated with high levels of HBsAg. Some studies show that GGH possess a pre-cancerous histological phenotype (Chisari, Filippi et al. 1987; Pasquinelli, Bhavani et al. 1992; Wang, Wu et al. 2003). GGH have been reported to be common in chronic HBV infections, particularly during phases of low HBV replication (Fan, Lu et al. 2001). Indeed, the observed HBsAg-positive hepatocytes were arranged in foci and in patterns consistent with previously described GGH (Wang, Wu et al. 2003). However, as shown in Figure 4.8, hepatocyte foci with either LCC or SCC observed in H&E stained sections were not consistently HBsAg-positive or HBsAg-negative in serial sections stained by immunohistochemistry. This suggests that HBsAg-expression was not necessarily associated with altered hepatocyte histology changes and therefore the observed foci of HBsAg-positive hepatocytes do not necessarily represent GGH, as described in previous studies.

Antiviral therapy and HBeAg status were not significantly associated with the percentages of hepatocytes that contained detectable levels of HBsAg and HBcAg (Student's t-test, $p > 0.05$).

A



B

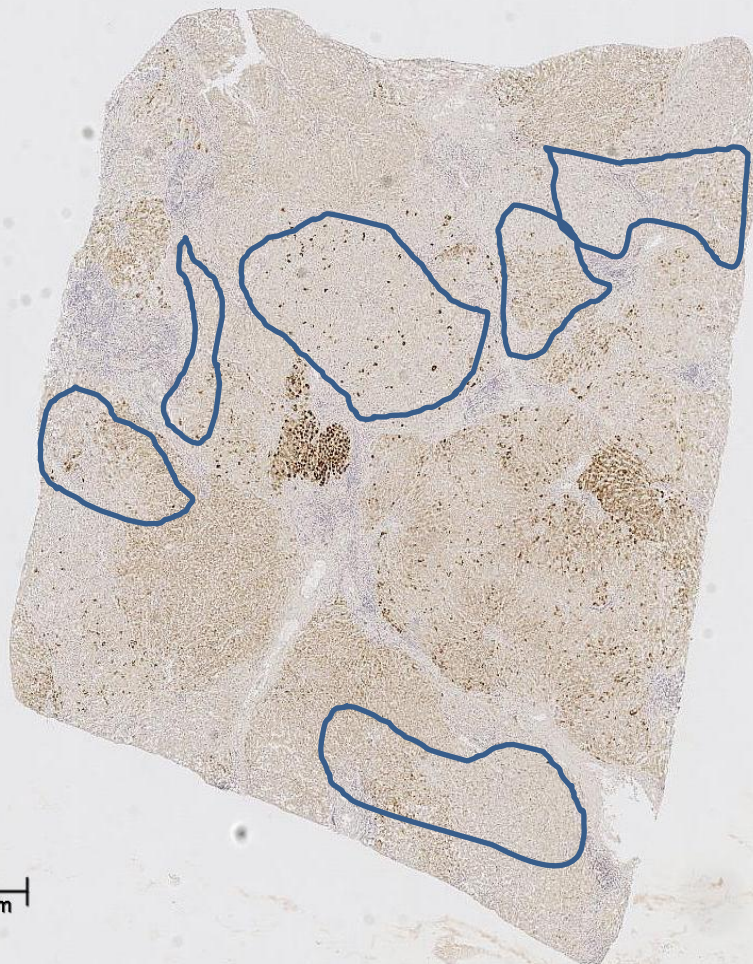


Figure 4.8. Localisation of hepatocytes with large cell change (LCC) in a liver tissue section from Patient Y6. Areas of hepatocytes with LCC in tissues sections stained by H&E (4.8A) were identified by A/Prof. Andrew Clouston (outlined in blue). HBsAg distribution was detected in neighbouring, serial sections (4.8B) by immunohistochemistry. The identified areas of hepatocytes with LCC did not have any obvious correlation with the distribution of HBsAg staining. Magnification 1.25x. Scale bars = 1 mm.

4.3.5 Localisation of HBsAg and Cytokeratin 19 (CK19) by dual immunofluorescence

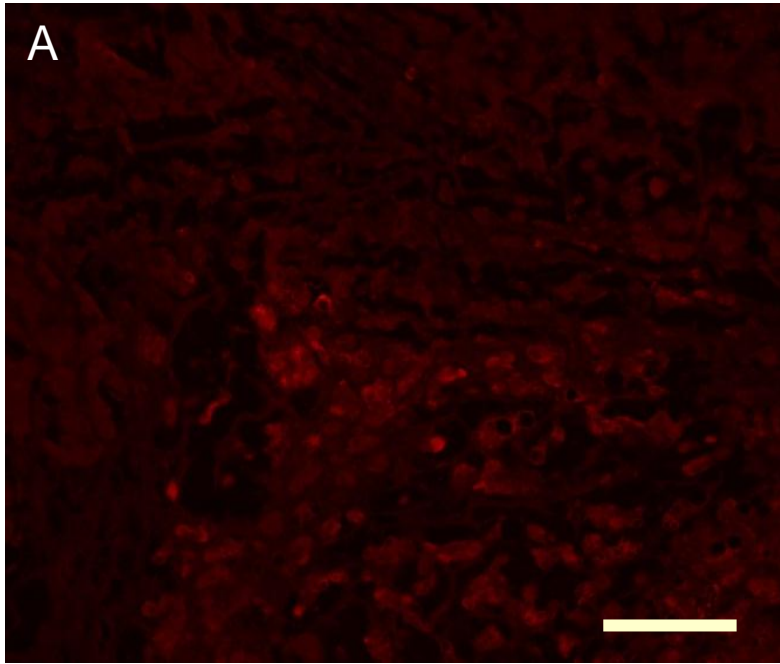
Immunofluorescence analysis was used to determine if oval cells were susceptible to HBV infection. Liver tissues of patients with chronic HBV infection have previously been reported to contain liver cells positive for both HBsAg and CK19, a cell marker associated with bile duct cells and oval cells (Hsia, Thorgeirsson et al. 1994). Oval cells are usually located close to bile duct epithelial cells, strongly express CK19, have scant oval-shaped cytoplasm, round nucleus, and a slightly smaller size than hepatocytes (Evarts, Nagy et al. 1987; Lowes, Brennan et al. 1999).

A standardised HBsAg/CK19 dual immunofluorescence protocol had not been previously developed. In order to undertake dual staining liver tissues from HBV-infected patients were fixed with formalin and used for pilot studies.

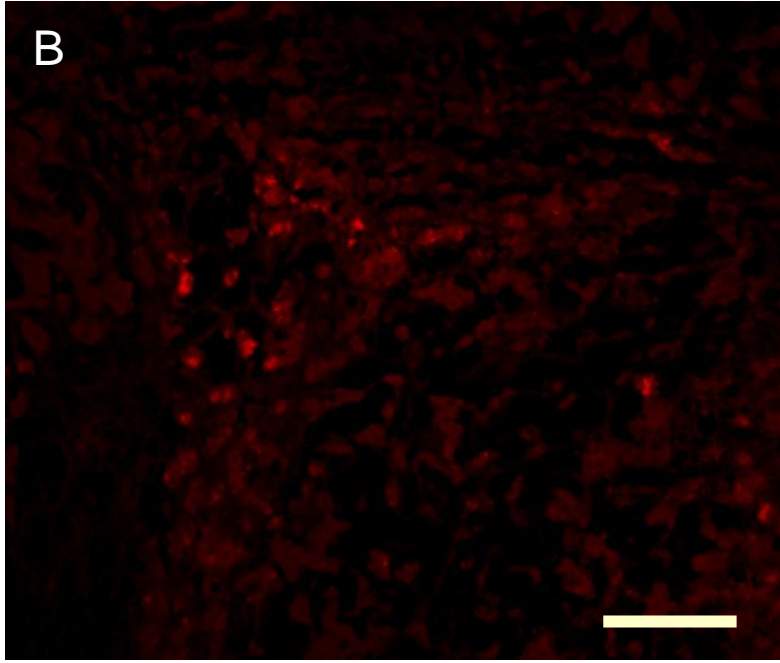
Formalin-fixed tissue was treated with 0.1% pepsin as described in Section 2.5 to allow detection of CK19 with a purified polyclonal rabbit anti-cytokeratin-19 antibody (Novus Biologicals), as per the manufacturer's recommendation. To investigate whether this affected the detection of HBsAg, liver tissue from HBV-infected patient CYRY (previously shown in Section 4.3.4 to demonstrate HBsAg expression) was treated with 0.1% pepsin or 0.1% trypsin to retrieve antigens. Antigen retrieval by enzymatic digestion using 0.1% trypsin had been used previously in our laboratory for immunohistochemical staining and acted as a positive control. PBS treatment was used as a negative control.

For HBsAg staining and detection, all treatments stained the same area of hepatocytes in serial sections, as shown in Figure 4.9. As the background staining was higher in the non-antigen-retrieved sections compared to those treated with enzymatic antigen-retrieval, 0.1% pepsin was used subsequently used to retrieve antigens in dual-stained sections.

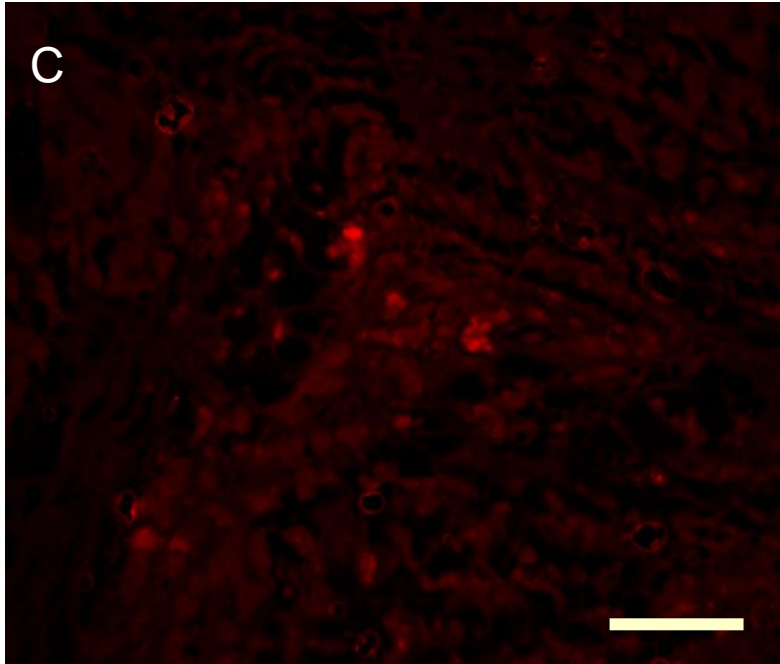
Only liver tissue from patients who had undergone resection could be assayed by dual-stain immunofluorescence due to availability of the amount of liver tissue. Furthermore, liver tissue from patients CN, NT, WN and XA were not assayed due to limited quantities of tissue. A list of the patients assayed and the results of the dual immunofluorescence are summarised in Table 4.3.



0.1%
pepsin



0.1%
trypsin



PBS

Figure 4.9. Immunofluorescence staining for HBsAg in formalin-fixed liver tissue sections from Patient CYRY. HBsAg-positive hepatocytes in serial sections were detected by immunofluorescence as described in Section 2.5 after being digested with pepsin (4.9A) or trypsin (4.9B) for antigen retrieval. PBS treatment (4.9C) was used as a negative control. HBsAg distribution was visualised by fluorescence microscopy using wavelength ranges of 532-587 nm (excitation) and 608-683 nm (emission). The distribution of positive cells remained constant in serial sections, suggesting that pepsin or trypsin treatments were not rendering HBsAg undetectable. Magnification 10x. Scale bars = 150 μm .

Table 4.3. Summary of results from dual immunofluorescence assay detecting Cytokeratin 19 and HBsAg in formalin-fixed liver tissue sections.

I – HBV-positive patients with cirrhosis

Patient	Total number of hepatocytes ¹	Total number CK19-positive cells ²	Total number HBsAg-positive cells ²	Total number dual-positive cells ²
CY	3359	0	438	0
FMC1	2947	0	0	0
FMC5	5479	1988	0	0
FMC6	3880	677	22	22
CYRY	3474	2	80	0
MH	1902	10	155	3
SAAO	3318	0	0	0

II – HBV-positive patients with fulminant hepatitis

FMC2	3435	0	0	0
FMC3	5091	0	61	0
FMC4	3467	127	1744	52
FMC7	5817	1244	1	0
HS	3022	84	134	3
HN	3312	17	0	0

III – HBV-negative patients

NHL1	4813	4	0	0
------	------	---	---	---

¹Total number of hepatocytes from five fields of view were randomly selected from each liver section at 10x magnification. The total number of hepatocytes in each field of view was determined by DAPI staining of nuclei and cell morphology.

²Includes counts of both hepatocytes and non-hepatocytes.

The staining was shown to be specific for HBsAg and CK19 in experimental samples, as dual immunofluorescence on serial tissue sections using the appropriate negative control sera showed only a low background fluorescence compared to the highly specific signal in experimental samples (Figure 4.10)

In all tissues, CK19-positive cells were generally found close to portal tracts and central veins, as shown in Figure 4.11. This is consistent with previous data which reported detection of CK19-positive cells in bile duct epithelia and in small clusters in the periportal spaces, i.e. in Zone 1 of the liver lobule (Chiu, Huang et al. 2007). This suggests that the assay was specifically detecting CK19. In tissues where HBsAg was detected by immunofluorescence, the HBsAg-positive cells appeared to be histologically indistinguishable from hepatocytes, consistent with previous reports (De Meyer, Gong et al. 1997).

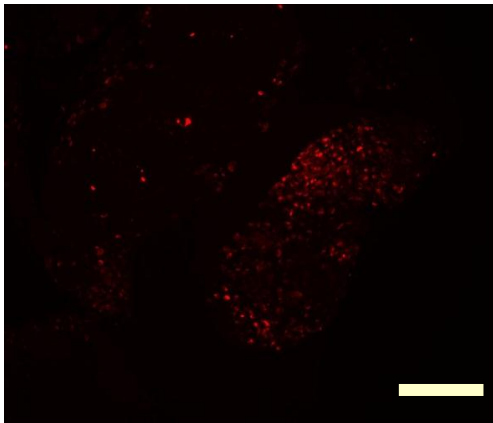
Overall, only a limited number of cells were positive for both HBsAg and CK19 in patient liver tissues. These dual immunofluorescence positive cells (shown in Figure 4.12) did not match the histological phenotype of oval cells; instead, the cells more closely resembled hepatocytes, with large cytoplasm, round nucleus and cuboidal shape. This suggested dual-positive cells were hepatocytes recently differentiated from oval cells, which retained CK19 and were infected with circulating HBV virions. Thus, this finding suggests that any virus-cell DNA junctions detected in the DNA extracts of liver tissues are unlikely to have originated from oval cells, but rather are maintained in mature hepatocytes.

4.4 Discussion

4.4.1 Classification of patient tissues

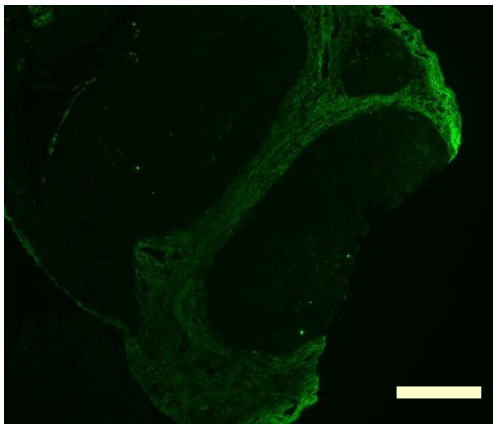
The histological analysis of the liver tissues indicates that the patients are from diverse stages of HBV infection: from early immune tolerant states (Group A) to late-disease states such as cirrhosis (Group B), fulminant hepatitis (Group C) and HCC (Group D). All patients with cirrhosis or fulminant hepatitis had a higher levels of histological activity compared to those in early non-disease states. Increased histological activity measured by Ishak and METAVIR scores has been associated with increased risk of HCC (Rodriguez-Diaz, Rosas-Camargo et al. 2007; Lok 2009; Lok, Seeff et al. 2009).

A

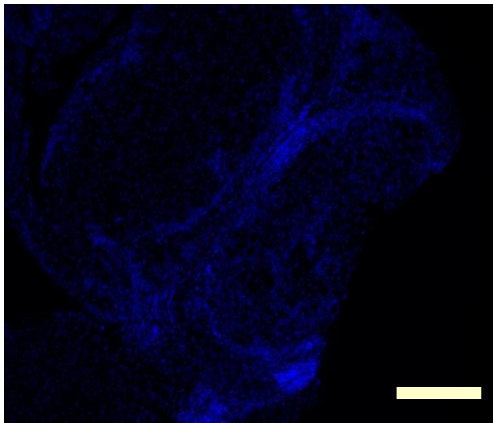
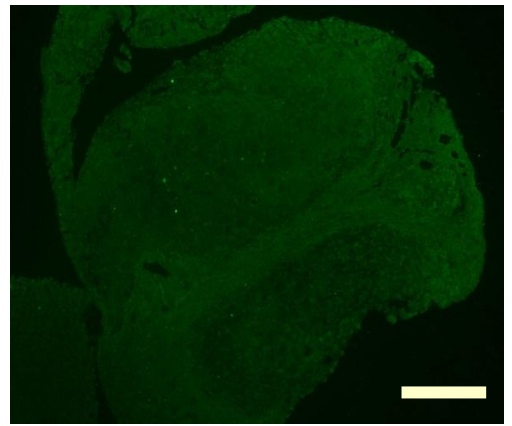


HBsAg
Ex = 532-587
Em = 608-683

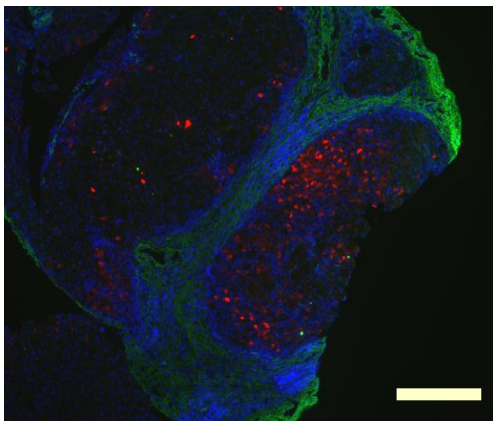
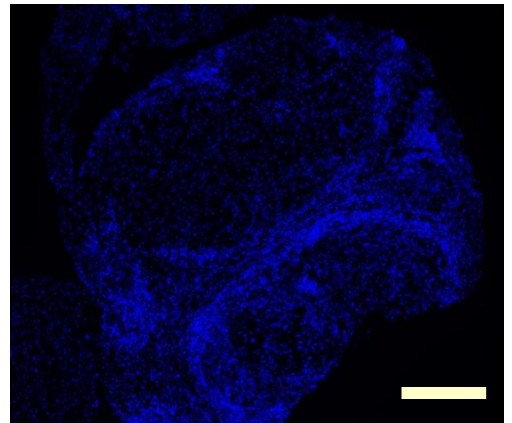
B



CK19
Ex = 470-490
Em = 520-560



DAPI
Ex = 330-380
Em = 435+



Merge

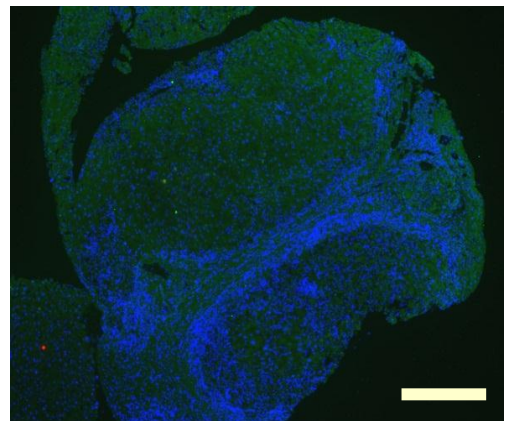
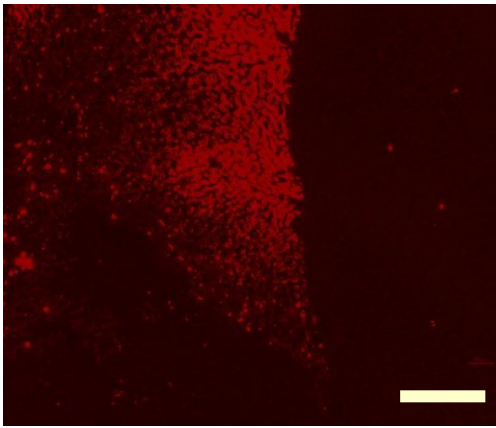


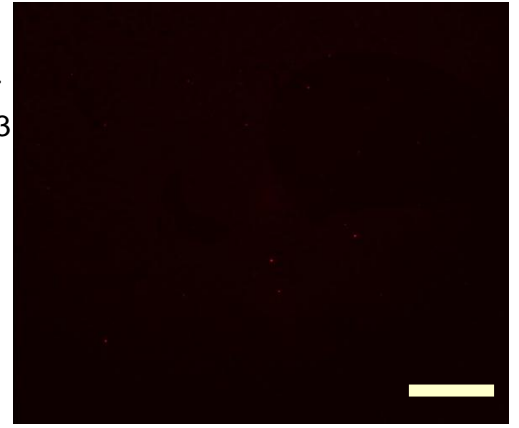
Figure 4.10. Dual immunofluorescence staining for HBsAg and CK19 was highly specific in liver tissue sections from Patient CY. Dual immunofluorescence was conducted as described in Section 2.5 in sections of formalin-fixed liver tissue from Patient CY with monoclonal HBsAg-specific antibodies raised in mouse hybridoma cells and polyclonal CK19-specific antibodies raised in rabbits (4.10A). Fluorescence was detected using the excitation (Ex) and emission (Em) wavelengths shown (in nm). Low levels of background fluorescence was seen in non-specific foetal calf serum and normal rabbit serum negative controls (4.10B), suggesting that the antigen staining is specific. Magnification 4x. Scale bars = 375 μm .

A

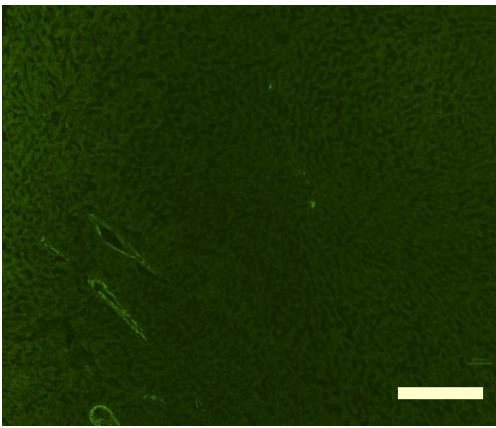


HBsAg
Ex = 532-587
Em = 608-683

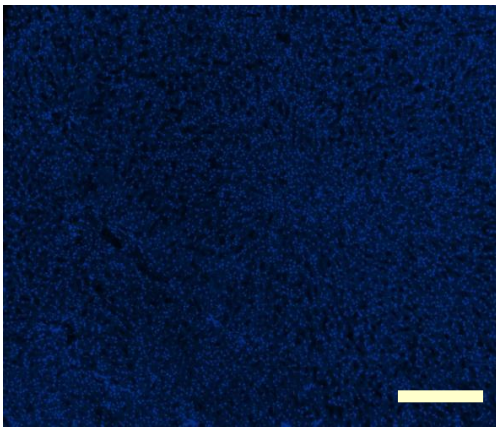
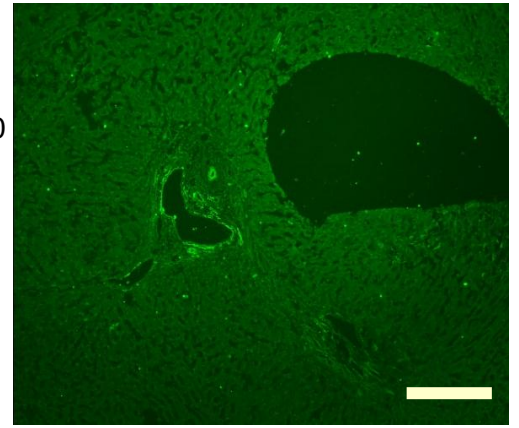
B



CK19
Ex = 470-490
Em = 520-560



DAPI
Ex = 330-380
Em = 435+



Merge

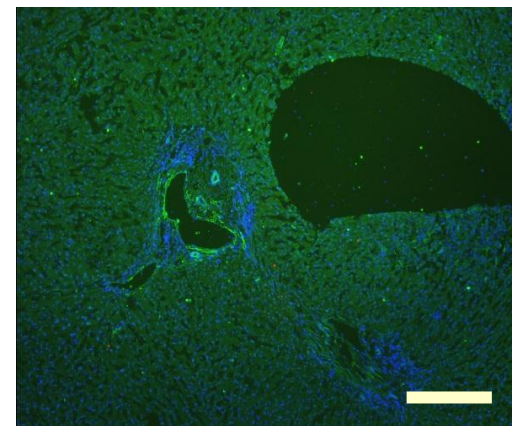
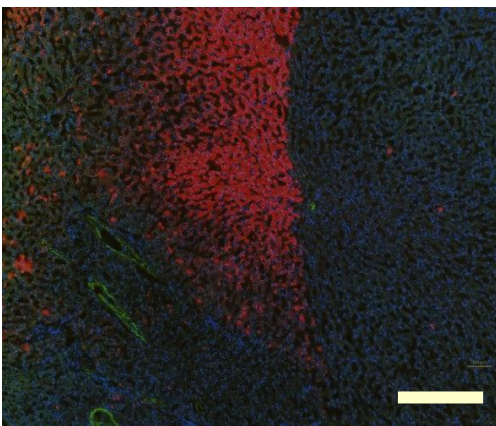
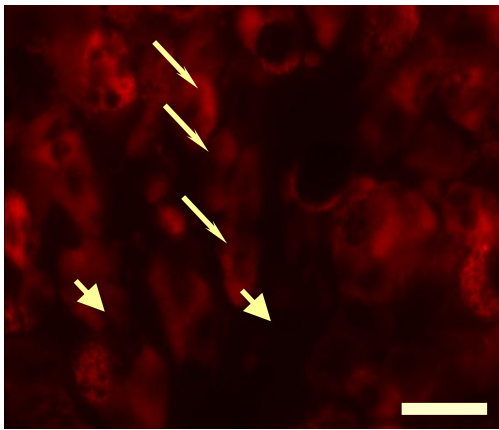
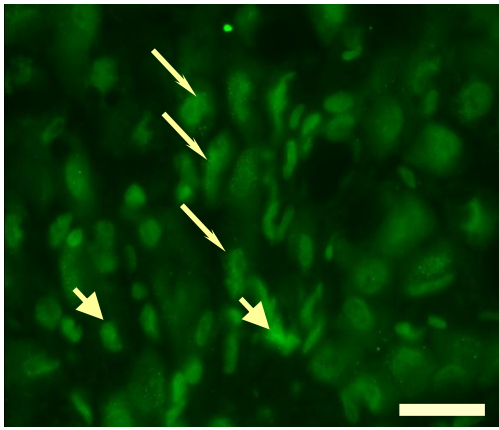
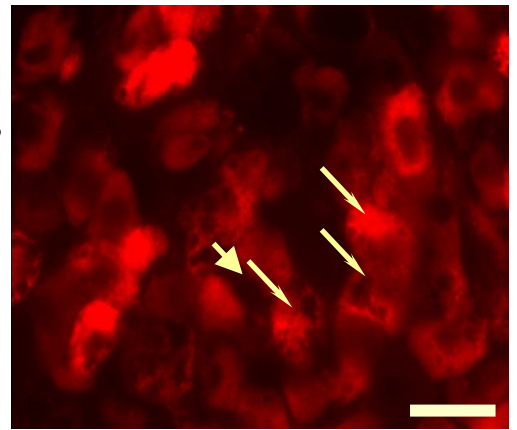


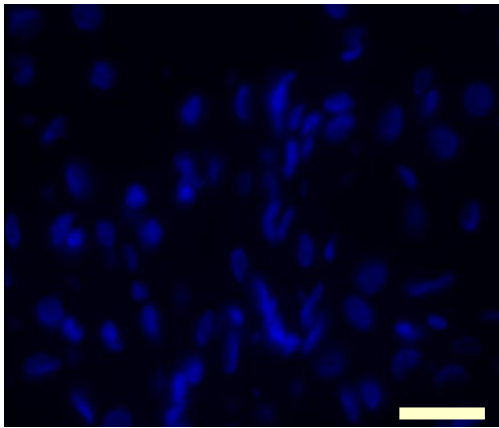
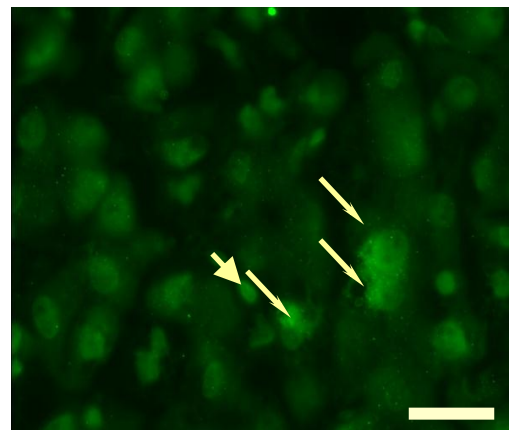
Figure 4.11. Dual immunofluorescence staining for HBsAg and CK19 in HBV-infected and non-infected tissues. Dual immunofluorescence (as described in Section 2.5) was used to visualise the distribution of HBsAg, CK19 and nuclei in tissue from HBV-infected patient FMC4 (4.11A) and non-infected patient NHL1 (4.11B). Nuclei were counterstained using the nuclear-specific stain DAPI (4.11C). Fluorescence was detected using the excitation (Ex) and emission (Em) wavelengths shown (in nm). Staining specific for HBsAg was only observed in liver tissues from HBV-infected Patient FMC4. Also, staining specific for CK19 was observed in peri-portal cells (generally non-hepatocytes) in all tissues, regardless of HBV infection. Magnification 4x. Scale bars = 375 μm .



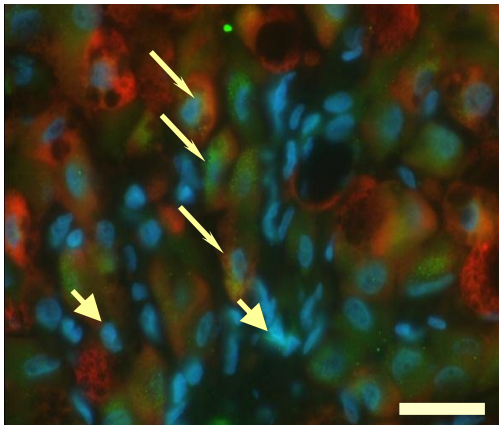
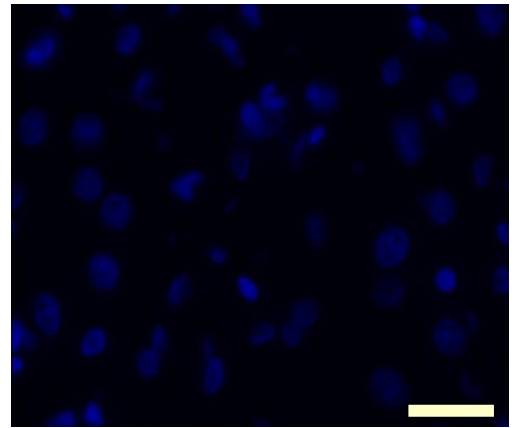
HBsAg
Ex = 532-587
Em = 608-683



CK19
Ex = 470-490
Em = 520-560



DAPI
Ex = 330-380
Em = 435+



Merge

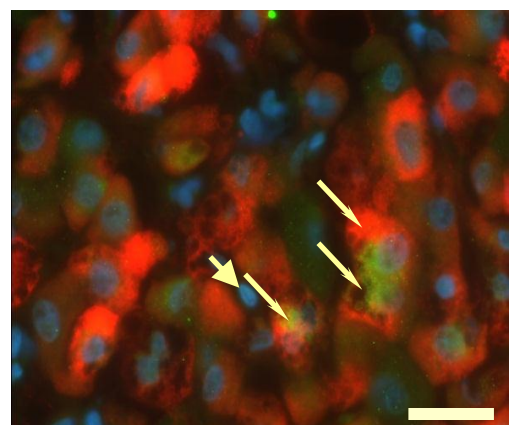


Figure 4.12. Dual immunofluorescence to detect cells positive for both HBsAg and CK19. Two different fields of view are shown of formalin-fixed liver tissue from Patient FMC4 after HBsAg and CK19 were detected by dual immunofluorescence as described in Section 2.5. Nuclei were counterstained using the nuclear-specific stain DAPI. Fluorescence was detected using the excitation (Ex) and emission (Em) wavelengths shown (in nm). Merged images (4.12D) show that dual-positive cells (long, thin arrows) have similar size and morphology to neighboring hepatocytes, as opposed to highly CK19-positive cells (short, thick arrows) with similar morphology to oval cells in previous literature. Magnification 60x. Scale bars = 25 μm .

The HCC cohort (Group D) did not show cirrhosis. Since ~80-90% of patients with HCC have been reported to have concurrent cirrhosis (Fattovich, Bortolotti et al. 2008), this sample of HCC patients is biased. However, this will allow us to measure whether clonal proliferation is associated with HCC development outside the context of cirrhosis. If clonal proliferation is indeed associated with disease progression, then more extensive clonal proliferation of hepatocytes should be observed in tissues from late-stage disease patients (Groups B, C and D) compared to early-stage non-disease patients (Group A). This hypothesis will be examined in Chapter 5 of this Thesis.

4.4.2 Virological aspects of HBV patients

In the current study, no correlation was observed between HBeAg status and HBV DNA concentration or the number of HBV antigen-positive hepatocytes. HBeAg seroconversion represents entry into the immune control phase of HBV infection, wherein many hepatocytes that support virus replication have been destroyed by the immune response (Yim and Lok 2006). In contrast to the current observations, previous data has shown that HBeAg-positive patients have significantly higher HBV titres and the number of HBV-positive hepatocytes compared to HBeAg-negative patients (Maruyama, Iino et al. 1993; Nguyen, Thompson et al. 2010).

This may be a result of the impact of host-virus interactions on HBV antigen expression. For example, during the immune clearance phase, the immune responses may target hepatocytes expressing HBV antigens and may lead to decreases in serum and liver HBV DNA levels, even though HBeAg-seroconversion may not have occurred (Yim and Lok 2006). Also, in the lulls between immune-mediated flares, it takes some period for the newly generated hepatocytes to become infected and produce high levels of HBV (Murray, Wieland et al. 2005). Indeed, all the tissues from HBV-positive patients showed evidence of inflammation or past immune-mediated damage.

Conversely, HBeAg-negative HBV escape mutants can emerge during immune control phase, leading to an increased HBV titre and HBV-expressing hepatocytes despite HBeAg to HBeAb seroconversion (Hoofnagle 2009). The presence of HBeAg HBV mutants was not tested, but HBeAg-negative HBV escape mutants may explain the high levels of HBV DNA in the liver of some of the HBeAg-negative patients.

Also, HBV DNA, HBsAg and HBcAg were detected heterogeneously across liver sections of patients with chronic HBV infection, consistent with previous studies (Gowans and Burrell 1985; Gowans, Burrell et al. 1985). The heterogeneity supports our hypothesis that there are many hepatocyte sub-populations in the liver of patients with chronic HBV infections. The focal nature of the hepatocytes also suggests that the hepatocytes are clonal in nature. To test this hypothesis further, specific foci of hepatocytes were isolated by laser-microdissection and tested using invPCR to determine whether the isolated cells are clonal. The results of this work are described in Chapter 7.

4.4.3 HBV antigen expression in cells other than hepatocytes

We also wished to determine whether any potential clonal proliferation due to a growth advantage occurs on the level of hepatocytes or on the level of stem cells, i.e. oval cells. The invPCR assay only detects clones of cells previously infected by HBV and containing integrated HBV DNA. Therefore, we investigated the susceptibility of oval cells to HBV infection.

In our study, no cell types other than hepatocytes were observed to be HBsAg-positive. This result is contrary to previous findings. Hsia *et al.* (1994) took serial sections of HCC and neighbouring liver tissue and alternately stained for HBsAg and CK19 (Hsia, Thorgeirsson et al. 1994). In this study, cells co-expressing both CK19 and HBsAg were reported, suggesting that oval cells are susceptible to HBV infection.

A possible explanation for this observation is that the exact same cells in the serial sections were not stained with HBsAg and CK19. Dual immunofluorescence staining of tissue sections in the current study ensured that the two antigens could be detected in the same section, and therefore the same cells. Only very rare cells (0.16%) were positive for both CK19 and HBsAg. Dual-positive cells did not appear to display a classical stem cell histological phenotype, instead displaying a hepatocyte-like phenotype. Furthermore, all HBV DNA positive cells detected by *in situ* hybridisation resembled hepatocytes. This suggests that oval cells do not support HBV antigen expression.

Where the block in the HBV replication cycle occurs in oval cells occurs is unknown. However, previous studies have found that HBV DNA transfected into non-hepatocyte liver cells had significantly lower expression compared to hepatocytes (Aragona, Burk et al. 1996; Ott, Ma et al. 1999). This suggests that there is a post-entry block to HBV expression in oval cells.

While not conclusive, this data is consistent with the hypothesis that CK19-positive oval cells are not susceptible to HBV infection. However, due to issues of sensitivity for both HBV DNA and HBV protein detection, we cannot exclude that oval cells may be infected by HBV and contain integrated HBV DNA at low levels that are undetectable using our assays. What is apparent from our studies is that oval cells, which have undergone differentiation into transitional hepatocytes, do support HBV infection and expression. Furthermore, data described in Section 7.3.1-2 supports these findings by showing significantly lower levels of HBV DNA in DNA extracts of non-hepatocellular cells compared to DNA extracts of hepatocytes. Together, these findings suggest that any clonal proliferation detected by using virus-cell DNA junctions in the DNA extracts of liver tissues are unlikely to have been due to proliferation of oval cells, but reflects the clonal proliferation of only hepatocytes.

4.4.4 Summary

In conclusion, analysis of clinical data and histological findings indicates that the tissues are from patients in early and late stages of chronic HBV infection. Also, both HBsAg-expression and HBV DNA replication were shown to be heterogeneous within the hepatocyte population of a given liver. Histological and immunofluorescence data suggests that oval cells are not vulnerable to HBV infection, but rather transitional and terminally differentiated hepatocytes are the main targets and acting reservoirs for HBV infection in the liver. Thus, the source of multiple copies of virus-cell junctions is likely to be limited to hepatocyte proliferation.

In the following Chapters, the extent of clonal proliferation of hepatocytes occurring in each of these patient cohorts will be determined. The characterisation of these patient tissues will contextualise the results of these studies by providing a link between HBV-associated pathogenesis and clonal proliferation of hepatocytes.

5 - Detection of hepatocyte clones in HBV-infected liver tissues

5.1 Introduction

As described in Section 1.2, the 3 main phases of chronic HBV infection include immune tolerance, immune clearance and immune control. Throughout these phases, complex virus-host interactions occur in the livers of patients with chronic HBV infection, include but are not limited to: HBV infection of hepatocytes; immune-mediated cell death of HBV-infected hepatocytes; natural selection and evolution HBV quasispecies; and phenotypical changes in the hepatocyte population. The interactions between the liver cell and the HBV populations over the decades of chronic HBV infection can eventually culminate in liver cirrhosis or HCC. While many HBV-associated changes in the liver cell population have been characterised, each of these mechanisms alone are not sufficient in explaining the emergence of HCC, as described in Section 1.8.

Results from independent studies have supported the hypothesis that clonal proliferation of hepatocytes occurs during HBV and other hepadnavirus infections (Hino, Kitagawa et al. 1984; Tanaka, Esumi et al. 1988; Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). Evidence suggests that there are hepatocytes with survival advantage in the context of the HBV-infected liver that undergo clonal proliferation. Whether clonal proliferation is connected to HBV-associated pathogenesis is currently unknown.

To measure clonal proliferation, previous studies have used integrated HBV DNA as a marker for hepatocytes lineages (Hino, Kitagawa et al. 1984; Tanaka, Esumi et al. 1988; Esumi, Tanaka et al. 1989; Yasui, Hino et al. 1992; Mason, Jilbert et al. 2005; Mason, Liu et al. 2010). Since integration of HBV dsDNA occurs by non-homologous end-joining (NHEJ), the integration site in the host genome is random and thus forms a unique virus-cell DNA junction for each integration event (Bill and Summers 2004). This provides a marker for a lineage of hepatocytes derived from the single cell in which the integration of dsDNA occurred. Thus, the extent of clonal proliferation is measured by determining the number of cells in a DNA extract of liver that contain the same unique virus-cell DNA junction.

Using integrated virus DNA as a marker for hepatocyte proliferation, large hepatocyte clones ($>10^3$ cells) have been found by Southern blot hybridisation in tumour tissue of HBV-infected patients with HCC (Hino, Kitagawa et al. 1984) and in the surrounding non-tumorous cirrhotic tissue containing hepatocytes (Tanaka, Esumi et al. 1988).

InvPCR has also been used in previous studies to detect virus-cell DNA junctions in liver tissues collected from a limited number of non-cirrhotic human patients (Mason, Liu et al. 2010). Hepatocyte clones have also been found using this method in animal models of HBV infection, such as HBV-infected chimpanzees and WHV-infected woodchucks (Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). In these studies, hepatocyte clones of >10000 cells were detected in chronically-infected liver tissues from humans, chimpanzees and woodchucks.

Liver turnover produces stochastic clonal proliferation of hepatocytes. However, mathematical modelling of the rates of hepatocyte turnover that occurred in WHV-infected liver suggested that stochastic clonal proliferation of hepatocytes was not sufficient to explain the large hepatocyte clones that were detected (Mason, Jilbert et al. 2005). Clones of >10000 were predicted by computer models to be formed stochastically after ~124 liver turnovers. Proliferating cell nuclear antigen staining indices of liver biopsies suggested only 6-12 turnovers had occurred over the lifetime of the woodchucks. Thus, some hepatocytes containing virus-cell DNA junctions must necessarily have had some survival advantage in the context of the virus-infected liver.

Mason *et al.* (2005) proposed that hepatocytes that gain a hyper-responsiveness to growth signals, lose the ability to express virus antigens, fail to display virus antigens on MHC or otherwise resist immune-mediated cell death have a survival advantage and thus are more likely to clonally proliferate (Mason, Jilbert et al. 2005). This suggests that clonal proliferation is an active promoter of pathogenesis; some survival advantages, such as hyper-responsiveness to growth signals or failure to express intrinsic antigens on MHC, may contribute to carcinogenesis. This is supported by the observation of greater numbers of FAH in WHV antigen-negative hepatocytes compared to WHV antigen-positive hepatocytes (Xu, Yamamoto et al. 2007). Also, only a minority of HCC tumour tissue has observed HBV antigen expression (Xuan, Xin et al. 2007). Thus, hepatocytes with loss of HBV virus expression may not only have a survival advantage, but also be more likely to contain preneoplastic changes.

Studies on clonal proliferation of hepatocytes during chronic HBV infections have generally focused on liver tissue of patients with HBV-associated late-stage disease. To determine

whether clonal proliferation is associated with the pathogenesis of HBV infection, we aimed to detect the extent of clonal proliferation of hepatocytes in liver tissue from patients in various stages of HBV infection and investigate potential causes of the observed clonal proliferation.

The mechanism of changing the cell phenotype in a way that promotes survival is currently unknown. One potential explanation is that insertional mutagenesis caused by the integration of HBV DNA could contribute to altering cellular phenotype. Such a mechanism would also bias assays that detect hepatocyte clonal proliferation using integrated HBV DNA as a marker.

Indeed, WHV integration into the oncogene N-myc2 is commonly observed in the HCC tumour tissues (Wei, Fourel et al. 1992; Wei, Ponzetto et al. 1992; Bruni, D'Ugo et al. 1999). HBV DNA integrations into oncogenes have been observed has also been observed (Paterlini-Brechot, Saigo et al. 2003; Brechot 2004). On the other hand, other studies have found that the majority of HBV DNA integration events appear completely random with respect to chromosomes and distance from genes (Matsubara and Tokino 1990; Takada, Gotoh et al. 1990; Wang, Lau et al. 2004; Mason, Low et al. 2009; Mason, Liu et al. 2010). We therefore aimed to use bioinformatics analysis to determine whether integrated HBV DNA was a phenotypically-neutral marker for hepatocyte lineages.

We hypothesise that hepatocyte clonal proliferation occurs in patients with late-stage disease, but occurs to a lesser extent in the liver of patients in earlier stages of HBV infection. Also, we hypothesise that the clonal proliferation occurs due to hepatocytes with survival advantage(s), rather than stochastically. Furthermore, we hypothesise that the survival advantages are not due to insertional mutagenesis by integrated HBV DNA.

5.2 *Experimental outline*

Hepatocyte clones were detected in liver tissues that were characterised in Chapter 4. In brief, the circulating HBV DNA from each patient was sequenced to allow for the design of invPCR assays. Based on histological analysis, the patients were split into 4 groups: HBV-positive non-cirrhotic patients, HBV-positive cirrhotic patients, HBV-positive HCC patients and

HBV-negative control patients. The clinical details of the patients from which these liver tissues were taken are reproduced in Table 5.1.

Total DNA was extracted from: i) 5 mg fragments of frozen, resected liver tissue (if available), as detailed in Section 2.7; and ii) 5 µm sections of paraffin wax-embedded ethanol-fixed liver tissue, as detailed in Section 2.8.

The invPCR technique was carried out as outlined in Section 2.14, and is summarised in Figures 5.1 and 5.2. Briefly, the HBV dsDNA form has been reported to integrate into random sites in the host cell genome by NHEJ (Bill and Summers 2004). A restriction enzyme (RE) (either *NcoI* or *DpnII*) was used to excise the right-hand virus-cell DNA junction by digesting in a known site in the HBV dsDNA genome and an unknown downstream site in the host cell genome. Next, T4 DNA ligase was used to circularise the DNA fragments containing the excised right-hand virus-cell DNA junction. A second RE was used to linearise and invert the circularised piece of DNA. This flanks the unknown cellular DNA sequence with two known HBV DNA sequences. In order to decrease amplification of false positive products formed by cccDNA, the products were digested with a third RE to limit PCR extension past the first RE cut (Figure 5.2).

Primers specific for the HBV DNA sequence were used to amplify the unknown cellular sequence using nested PCR. Virus-cell DNA junctions were then sequenced by Sanger sequencing (Section 2.12.3). The copy number of each unique virus-cell DNA junction was enumerated by end-point dilution of template across a 96-well plate (outlined in Section 2.14) and then analysis by most probable number (MPN) method with a freely available MPN calculator (Curiale 2004). Due to the possibility of artefactual products being formed by ligation of dsDNA to random fragments of cellular DNA, the clones sizes were calculated only for virus-cell DNA junctions that were detected greater than once.

Furthermore, bioinformatics analysis was used to calculate the theoretical maximum percentages of virus-cell DNA junctions detectable using the *NcoI* and *DpnII* invPCR assays. This analysis was performed in an attempt to quantify some of the inaccuracies associated with determining clonal proliferation using the invPCR used.

Table 5.1. An abridged summary of known patient data, histological interpretation and virological data of liver tissues from early stage, cirrhotic, fulminant hepatitis, HCC HBV-positive patients and HBV-negative patients. Full summaries of patient data can be found in Tables 4.1. and 4.2.

A – Early-stage HBV infection

Name ¹	Type ²	Age	HBV Genotype	HBV DNA/BG (qPCR) ³	Serum HBsAg	Serum HBeAg	Metavir score ⁴	Liver HBsAg ⁵	Liver HBcAg ⁵
C ^A	NB	22	D	1.6	+	-	0	++++	-
GS1 ^B	Slide	<30	C	7000	+	+	0	+	+
GS2 ^B	Slide	<30	C	15000	+	+	1	+	+
GS3 ^B	Slide	<30	A	5	+	+	0	+	-
GS4 ^B	Slide	<30	D	1300	+	+	2	+++	+++
GS5 ^B	Slide	<30	A	240	+	+	1	+	-
L ^A	NB	54	C	260	+	+	2	+	+

B – HBV-positive patients with cirrhosis

CN ^C	Resect.	52	D	ND	+	+	4	-	-
CY ^C	Resect.	54	C	290	■	■	4	+	-
CYRY ^C	Resect.	48	A	370	+	-	4	+++	+++
FMC1 ^D	Resect.	67	B	39	+	-	4	-	-
FMC5 ^D	Resect.	40	D	2	+	+	4	+	-
FMC6 ^D	Resect.	58	B	830	+	-	4	+	-
MH ^C	Resect.	44	A	0.12	+	-	4	++	-
NT ^C	Resect.	48	A	ND	■	■	4	-	-
SAAO ^C	Resect.	65	A	0.035	+	-	4	-	-
XA ^C	Resect.	46	C	0.25	+	-	4	++	-

C – HBV-positive patients with fulminant hepatitis

FMC2 ^D	Resect.	62	A	1.4	+	■	4	-	-
FMC3 ^D	Resect.	49	B	4800	+	-	4	-	-
FMC4 ^D	Resect.	68	B	160	+	-	3	++	-
FMC7 ^D	Resect.	49	C	22	+	■	3	+	-
HN ^A	Resect.	■	D	0.33	-	-	1	++	-
HA ^C	Resect.	48	D	2.3	+	+	4	++	-

ND = not detected

■ = Unknown

Table 5.1. Cont.

D – HBV-positive patients without cirrhosis and with concurrent HCC

Name ¹	Type ²	Age	HBV Genotype	HBV DNA/BG (qPCR) ³	Serum HBsAg	Serum HBeAg	Metavir score ⁴	Liver HBsAg ⁵	Liver HBcAg ⁵
DG ^E	Resect.	56	C	0.13	+	-	2	+	-
Y2 ^B	Slide		B	1.7 ^S	+	-	2	+	-
Y3 ^B	Slide		B	2.4 ^S	+	-	2	+	-
Y4 ^B	Slide		B	2.9 ^S	+	-	1	+++	-
Y5 ^B	Slide		C	2.6 ^S	+	-	3	+	-
Y6 ^B	Slide		B	1.4 ^S	+	-	3	+	-

E – HBV-negative control patients

NHL1 ^E	Resect.			ND	-	-	0	-	-
NHL2 ^E	Resect.			ND	-	-	0	-	-
NHL3 ^E	Resect.			ND	-	-	0	-	-

F – Uncharacterised tissues

WN ^A	Resect.	45	A	ND			Not interpretable.		
-----------------	---------	----	---	----	--	--	--------------------	--	--

ND = not detected

■ = Unknown

Liver histology was interpreted by A/Prof. Andrew Clouston (University of Queensland) and, based on the histological findings, HBV positive patients were assigned to 4 groups: Group A - early-stage HBV infection (7 patients); Group B - late-stage HBV infection with cirrhosis (10 patients); Group C - late-stage fulminant hepatitis (6 patients); and Group D - late-stage HBV infection without cirrhosis and with concurrent HCC (6 patients). Histology was also interpreted for HBV-negative patients, assigned to Group E. Histology could not be interpreted for Group F - Patient WN.

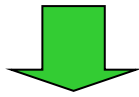
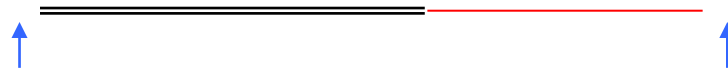
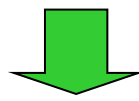
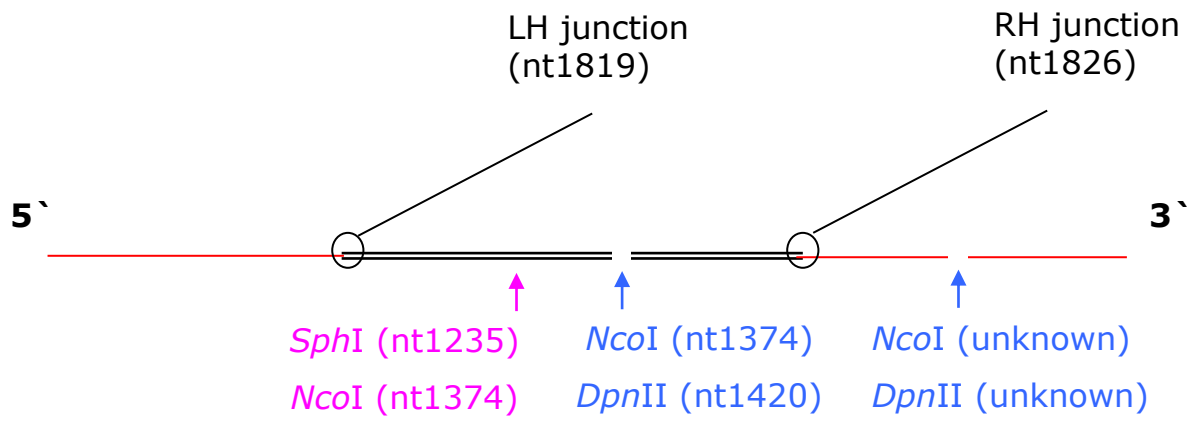
¹ Patient liver tissues were sourced from clinical care centres based in: the Royal Adelaide Hospital (A); University of Washington (B); Centenary Institute (C); Flinders Medical Centre (D); and SA pathology (E).

² Available tissues were in the form of liver resection (Resect.), needle biopsy (NB) or ethanol-fixed slide-mounted liver tissue (Slide).

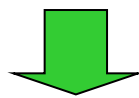
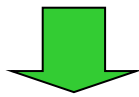
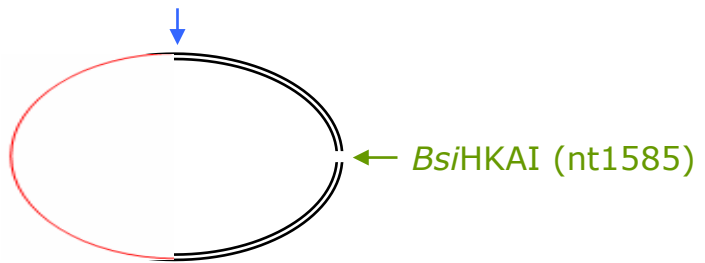
³ Copy numbers of HBV DNA per beta-globin as previously determined by qPCR in Section 4.3.2. Values indicated with “^S” were previously published by Mason *et al.* (2010).

⁴ For some needle biopsy samples, insufficient material was available for histological interpretation. Histological scores for these samples (highlighted in yellow) were assigned using information taken from clinical reports provided by each clinical care centre.

⁵ Liver HBsAg and HBcAg levels were detected by immunohistochemistry in Section 4.3.4. The percentage of antigen-positive hepatocytes are shown: “-”=0%; “+”=<25%; “++”=25-50%; “+++”=50-75%; “++++”=>75%.



Add T4 ligase



Nested PCR

Figure 5.1. Inversion of DNA extracted from HBV-infected human liver. These designs have been used previously to detect virus-cell DNA junctions in DNA extracts from liver tissue of chimpanzees and human patients chronically infected with HBV (Mason *et al.* 2010; Mason *et al.* 2009).

Virus-cell DNA junctions were detected in human liver tissue using invPCR as described in Section 2.14. Firstly, the right-hand virus-cell DNA junction formed by the integration of HBV DNA into cellular DNA was excised with digestion with either *NcoI* or *DpnII*. The excised product was diluted and circularised using T4 ligase. The circularised product was cleaved with *BsiHKAI* to form an inverted product, with known viral sequences surrounding unknown cellular sequences. The inverted products were also cleaved with either *SphI* or *NcoI* to reduce aberrant amplification of HBV cccDNA (Figure 5.2). Inverted products were then amplified by nested PCR.

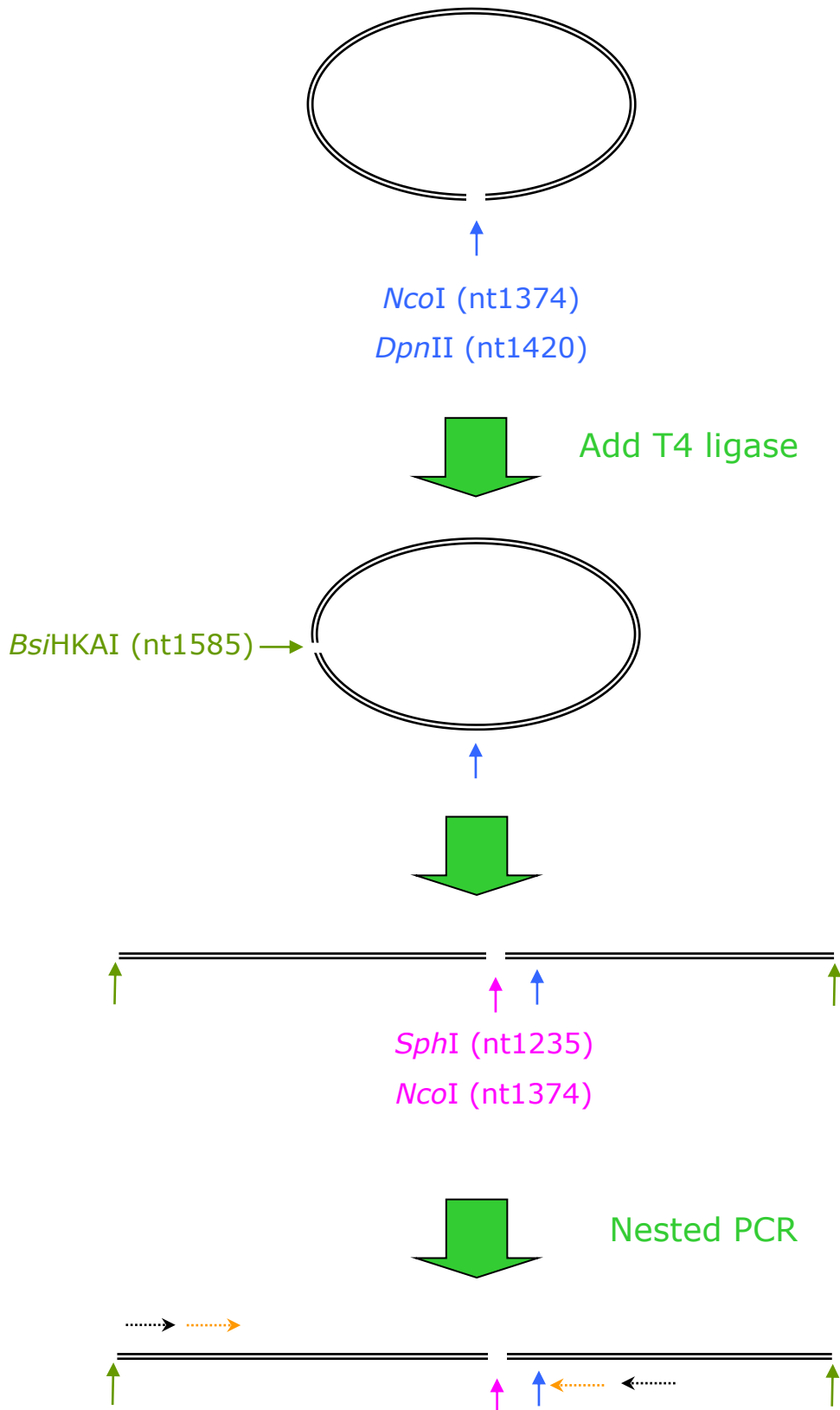


Figure 5.2. Inclusion of a third restriction enzyme during inversion decreases amplification of cccDNA. cccDNA (black double lines) can potentially be amplified during invPCR. If the first cut (blue arrows) is rejoined during the ligation reaction, then products could potentially be amplified by nested PCR. However, if a third restriction enzyme (pink arrows) is used to digest the inverted products, then nested primers (black and orange dashed arrows) are unable to amplify the cccDNA-derived inversion product.

A qPCR assay (carried out as described in Section 2.13) using primers that spanned selected unique virus-cell DNA junctions was used to independently determine the number of copies of each virus-cell DNA junction. This independent verification was performed to check and validate the invPCR assay for determining the extent of clonal proliferation of hepatocytes.

The virus-cell integrations were subjected to bioinformatics analysis to determine whether HBV DNA integration itself was likely to be altering the cellular phenotype, by, for example, insertional mutagenesis into known oncogenes or by increasing the expression of downstream growth factors by providing a HBV-derived promoter.

5.3 Results

5.3.1 Primer design and RE selection for invPCR

The sequence of the region on the HBV DNA genome upstream of the likely right hand end virus-cell DNA integration junction (Section 4.3.2) was used to design inversion protocols and primers required for invPCR. All HBV DNA sequences are shown in full in Appendix 9.6.

The invPCR designs used are summarised in Table 5.2. The *NcoI* invPCR protocol, previously designed by Mason *et al.* (2009), was compatible with HBV DNA sequences associated with all patients. In addition, only some HBV DNA sequences from patients were compatible with the *DpnII* invPCR protocol, previously designed by Mason *et al.* (2010); namely, CN, FMC3, GS4, HN, HS, NT, SAAO, XA, Y2, Y4 and Y6. In the other patients, extra *DpnII* sites in the HBV DNA sequence upstream of the expected right-hand junction would have prevented the detection of virus-cell DNA junctions by invPCR, as shown in Figure 5.3.

Furthermore, the primers used by Mason *et al.* (2010) were shown to be applicable to the current liver DNA extracts for the *NcoI* invPCR protocol. However, for the *DpnII* invPCR protocol, new reverse outer primers were designed due to the shift of the invPCR RE sites and sequence variation in HBV sequences, as shown in Table 2.2 5.2.

Table 5.2. InvPCR designs (previously developed by Mason *et al.* (2009)) used to detect integrated HBV DNA in DNA extracts of liver tissue from patients with chronic HBV infections.

1 st RE ¹	2 nd RE ¹	3 rd RE ¹	Forward Outer primer ²	Reverse Outer primer ²	Forward Inner primer ²	Reverse Inner primer ²	Patients					
<i>Nco</i> I	<i>Bst</i> HKAI	<i>Sph</i> I	1380	1383	1381	1416	HN					
					1385	1382	C, CY, FMC1, FMC3, FMC5, FMC7, GS1, GS4, Y2, Y3, Y4, Y6, XA					
							1397	Y5				
					1414	1382	FMC2, FMC4, FMC6, GS2, L, MH					
					1395	1381	1416	GS3, GS5, HS, WN				
						1385	1382	CYRY, HS				
						1385	1416	WN				
					1383 + 1395	1385	1382	CN, NT, SAAO, DG				
					<i>Dpn</i> II	<i>Bst</i> HKAI	<i>Nco</i> I	1380	<i>Dpn</i> II A ³	1385	1383	CN, HN, SAAO, GS4
									<i>Dpn</i> II B ³	1385	1383	XA
<i>Dpn</i> II C ³	1385	1383	Y2, Y4, Y6, FMC3									

¹ RE = Restriction enzyme.

² Primer sequences are listed in full in Table 2.4B. All primers are designed by Mason *et al.* (2010), unless otherwise indicated.

³ New primer sequence designed for this study.

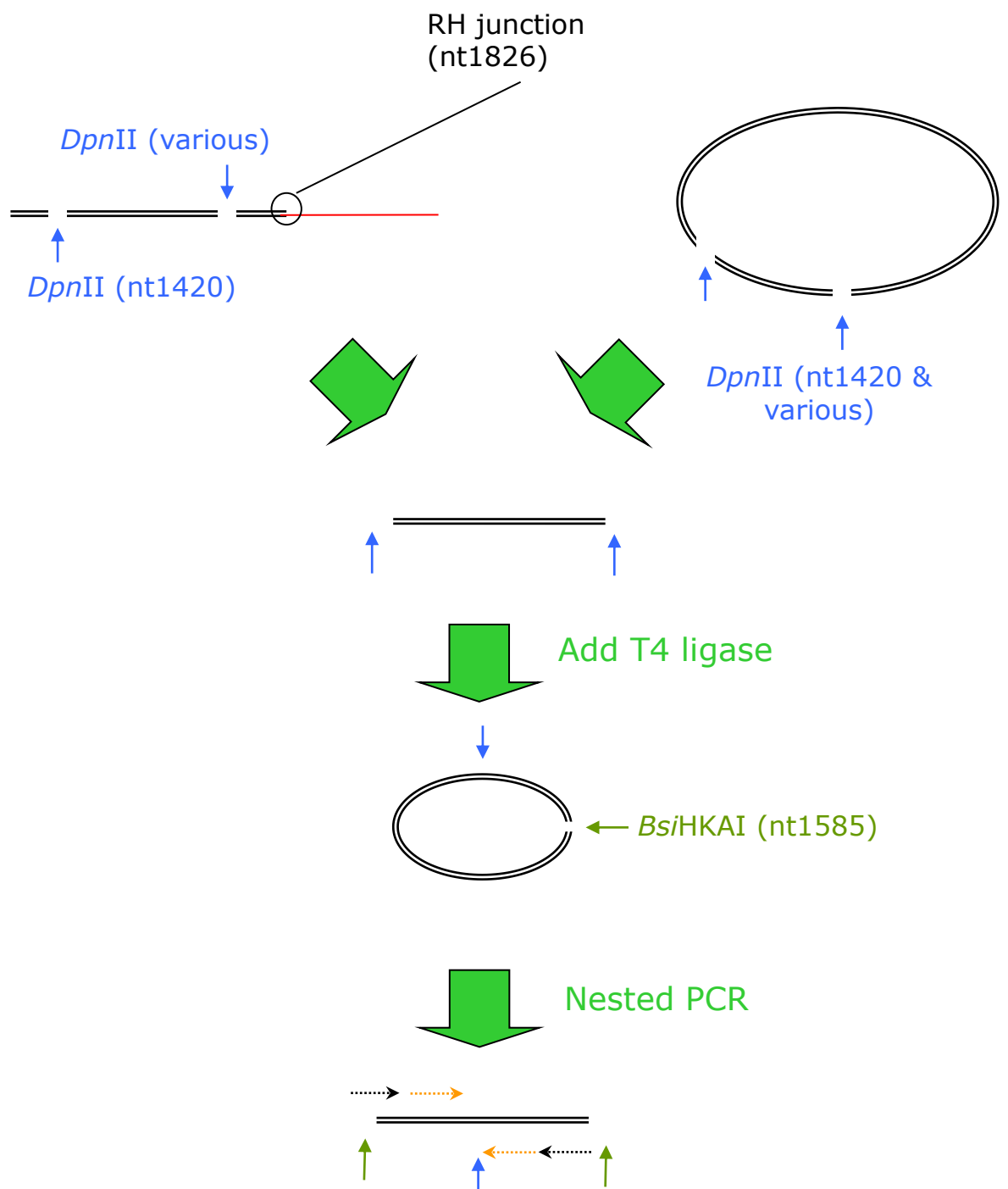


Figure 5.3. HBV strains incompatible with the *DpnII* invPCR assay have extra *DpnII* sites, which allow production of false artefacts. The sequences of HBV DNA associated contained extra *DpnII* sites upstream of the expected right hand junction. Following *DpnII* digestion, both dsDNA and cccDNA are able to produce a small HBV DNA-derived fragment. This small fragment can be inverted and act as a template for nested PCR. Thus, DNA extracts with HBV sequences containing these extra *DpnII* were not able to be analysed by the *DpnII* invPCR design.

5.3.2 Calculation of the theoretical maximum percentage of virus-cell DNA junctions observable by invPCR

We intended to use the invPCR designs outlined in Section 5.3.1 to detect virus-cell DNA junctions. The copy number of unique virus-cell DNA junctions represents the hepatocyte clone sizes. However, not all virus-cell DNA junctions may be detected by the invPCR assay. This may affect the accuracy in calculating the extent of clonal proliferation that the measured liver tissues have undergone. Therefore, we aimed to quantify some of the limitations of the invPCR assay by calculating the maximum percentage of virus-cell DNA junctions that could be observed by the current invPCR designs.

Multiple factors during the invPCR protocol impact on the detection of virus-cell DNA junctions, including: the purity of extracted DNA, the efficiency of RE digestions, the efficiency of the circularisation of the excised virus-cell DNA junctions, and the sensitivity of the nested PCR. The efficiency of inversion and nested PCR in the total liver DNA extracts was studied in Section 3.2.1.3 using a ~400 bp segment of the cellular gene procollagenase 1 alpha. On average, ~23% of the cellular DNA segments were inverted. Whether this efficiency changes with the length of the excised virus-cell DNA junction is not known.

Losses of DNA during extraction or inversion may cause virus-cell DNA junctions to be lost and therefore left undetected. Also, the interference of HBV RI DNA could affect the inversion reactions or sensitivity of the nested PCR by providing extra DNA forms that can be ligated or be sites of PCR primer annealing. Furthermore, while sequencing of HBV DNA will detect the dominant circulating strain, HBV quasi-species emerge due to mutations produced by the error-prone HBV reverse transcriptase at a rate of 1 mutation per 10^4 - 10^5 nucleotides (Park, Kim et al. 2003). Quasi-species that are mutated in sites for DNA restriction or PCR primer binding may also escape detection by invPCR.

Many of these factors are difficult to measure and may vary from patient-to-patient, or for specific virus-cell DNA junctions within a patient sample. Even if a system were perfected where these factors were insignificant, the invPCR assay is intrinsically limited in the percentage of virus-cell DNA junctions it can detect. This may be due to the REs used in the invPCR assay and the distribution of the RE sites in the human DNA genome.

For a virus-cell DNA junction to be detected, the first restriction site must be found within the cellular DNA downstream of the right-hand junction. However, as shown in Figure 5.4, the virus-cell DNA junction is not detected by invPCR if the recognition sequence of the second or third RE is located in the cellular DNA between the first restriction site and the virus-cell DNA junction. Amplification by nested PCR would be blocked by the digestion of the inverted virus-cell DNA junction template. So, the expected percentage of virus-cell DNA junctions detectable by an invPCR assay is a ratio of the frequency of the first RE site and the sum of the frequencies of the first, second and third RE sites within the cellular genome.

The human genome is not a random sequence; restriction sites are not uniformly distributed due to non-random GC content and dinucleotide frequencies found in human DNA sequences. The GC content of the human genome is ~41% compared to the theoretical ideal of 50% (Cohen, Dagan et al. 2005). Furthermore, some dinucleotide sequences have frequencies divergent from their theoretical values, independent of GC content (Nussinov 1980; Shioiri and Takahata 2001). This leads to RE site frequencies divergent from those expected from a random DNA sequence (Fedorova, Dizadex et al. 2001; Shioiri and Takahata 2001).

The expected percentage of observable virus-cell DNA junctions for each design was calculated taking into account GC content and dinucleotide frequencies. Firstly, the expected mean frequency each RE site of the enzymes used in each invPCR design was calculated using an algorithm developed by Bastie-Sigeac and Lucotte (Bastie-Sigeac and Lucotte 1983). The frequency (F) of a 4-nucleotide RE site, WXYZ, (e.g. GATC for *DpnI*) is calculated by:

$$F(W-X-Y-Z) = (W\%)(X\%)(Y\%)(Z\%)([W-X])([X-Y])([Y-X])$$

where, W%, X%, Y%, and Z% are the relative frequencies of the respective nucleotides in the human genome; and [W-X], [X-Y], and [Y-X] are dinucleotide sequence preferences (defined as the ratio of the observed frequency of each dinucleotide sequence compared to its expected frequency).

The frequency of 6-nucleotide RE sites, UVWXYZ, (e.g. CCATGG for *NcoI*) was similarly calculated by:

$$F(U-V-W-X-Y-Z) = (U\%)(V\%)(W\%)(X\%)(Y\%)(Z\%)([U-V])([V-W])([W-X])([X-Y])([Y-X])$$

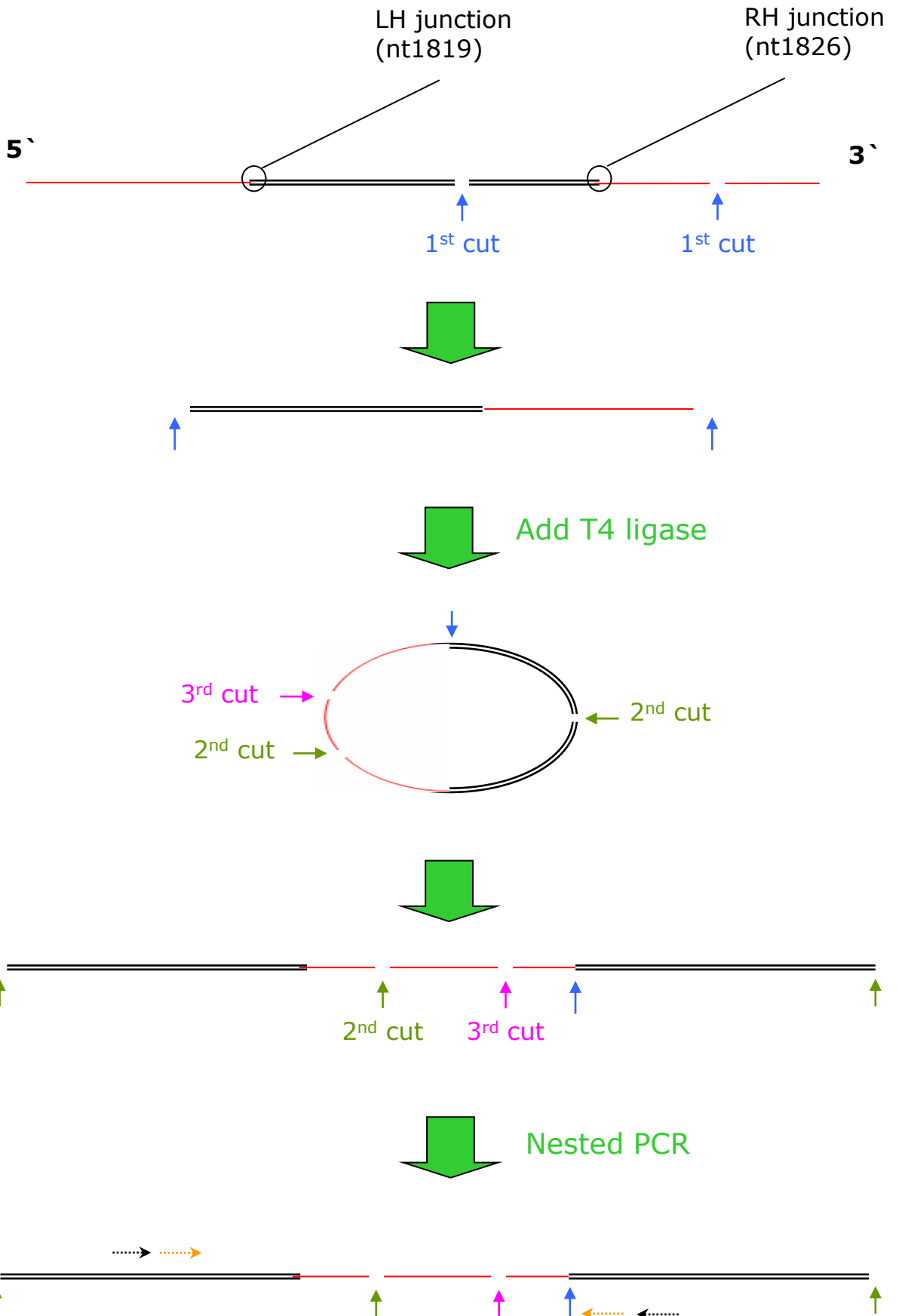


Figure 5.4. If 2nd or 3rd cut restriction enzyme sites are present in the cellular DNA of the virus-cell DNA junction, it is not detected by invPCR. If there are 2nd or 3rd cut restriction enzyme sites (green and pink arrows, respectively) in the cellular DNA portion (red lines) of the inverted virus-cell DNA junction, the inverted product is digested. This digestion stops the amplification of the virus-cell DNA junction by nested primers (black and orange dashed arrows).

The human genome GC content of 41% (Cohen, Dagan et al. 2005) was used in these calculations. Thus, nucleotides A and T have a relative frequency of 0.295, and nucleotides G and C have a relative frequency of 0.205.

The values for each dinucleotide preference in human genomic sequences were observed empirically by previous studies with concordant values (Nussinov 1981; Bulmer 1987; Nakashima, Nishikawa et al. 1997). Bulmer (1987) found differences in dinucleotide preferences between intron and exon sequences (Table 5.3). Dinucleotide preferences have been attributed to biological restrictions and the chemistry of base mutations. For example, exons sequences are restricted by the fecundity of the protein which they encode and thus presumably stray from expected frequencies to a lesser extent. The decreased frequency of CG dinucleotides with mirrored increased frequency of TG has been attributed to 1) the ease of which methylated cytosine residues (present in CpG islands) are converted to thymidine bases and 2) because CG dinucleotides are often used by transcription machinery as markers for coding sequences (Ohno 1988). Different frequencies of amino acids also restrict the frequencies of dinucleotide sequences (Hanai and Wada 1988).

Therefore, in the equation, the values 0.295 and 0.205 were used for relative frequencies of A and T and, G and C, respectively. Also, the dinucleotide sequence preferences outlined in Table 5.2 were used as inputs for the equations. In the case of degenerate restriction sites, the frequency was calculated as the sum of the frequencies of all recognisable sequences. The calculated frequencies of REs are summarised in Table 5.4.

The theoretical maximum observable percentage of virus-cell DNA junctions was calculated as the frequency of the first RE site divided by the sum of the frequencies of the first, second and third RE sites. The maximum percentage of total virus-cell DNA junctions theoretically observable by the *NcoI* and *DpnII* invPCR designs are 14/23/17 % (intergenic/intron/exon) and 68/75/71%, respectively.

These results suggest that there should not be a great difference in detecting virus-cell DNA junctions in intergenic, intron or exon sequences. However, the *DpnI* invPCR assay should detect about 3 times the number of virus-cell DNA junctions compared to the *NcoI* invPCR assay. This greater percentage is mainly due to the higher frequency of the 4-nucleotide RE

Table 5.3. Dinucleotide preferences in human intergenic regions (5.3A), introns (5.3B) and exons (5.3C).

A

		3'			
		A	C	G	T
5'	A	0.89	1.12	1.25	0.89
	C	1.25	0.96	0.2	1.25
	G	0.92	0.97	0.96	1.12
	T	1.01	0.92	1.25	0.89

B

		3'			
		A	C	G	T
5'	A	1.06	0.82	1.25	0.87
	C	1.22	1.20	0.30	1.29
	G	1.06	0.91	1.19	0.81
	T	0.73	1.07	1.18	1.07

C

		3'			
		A	C	G	T
5'	A	1.02	0.91	1.19	0.84
	C	1.19	1.11	0.50	1.34
	G	1.08	1.00	1.00	0.86
	T	0.52	0.97	1.46	0.94

The preferences for each dinucleotide sequence is defined as the observed frequency divided by the expected frequency, given the GC% content of human sequences. The value for each dinucleotide sequence (X-Y) is shown in the intersecting box of the X row and Y column. Data tables were adapted from Bulmer (1987).

Table 5.4. Theoretical frequency of restriction enzyme sites for *NcoI* (5.4A) and *DpnII* invPCR designs (5.4B).

A

Restriction enzyme	Recognition site (5' -> 3') ¹	F based on GC%	Genetic area	Total dinucl. pref. ²	Frequency (bp ⁻¹) ³	Period (bp)
1) <i>NcoI</i>	CCATGG	0.000154	Intergenic	1.28	0.000197	5070
			Intron	1.79	0.000276	3620
			Exon	1.62	0.000249	4020
2) <i>BsiHKAI</i>	GWGCWC	0.000154	Intergenic	6.30	0.000970	1030
			Intron	5.00	0.000770	1300
			Exon	6.05	0.000932	1070
3) <i>SphI</i>	GCATGC	0.000154	Intergenic	1.31	0.000202	4960
			Intron	1.03	0.000159	6290
			Exon	1.46	0.000225	4440

Detection fraction (inter.) = $F1/(F1+F2+F3) = 0.000197/(0.000197 + 0.000970 + 0.000202) = 14\%$
 Detection fraction (intron) = $F1/(F1+F2+F3) = 0.000276/(0.000276 + 0.000770 + 0.000159) = 23\%$
 Detection fraction (exon) = $F1/(F1+F2+F3) = 0.000249/(0.000249 + 0.000932 + 0.000225) = 17\%$

B

Restriction enzyme	Recognition site (5' -> 3') ¹	F based on GC%	Genetic area	Total dinucl. pref. ²	Frequency (bp ⁻¹) ³	Period (bp)
1) <i>DpnII</i>	GATC	0.00321	Intergenic	0.75	0.00242	414
			Intron	0.99	0.00318	314
			Exon	0.88	0.00283	353
2) <i>BsiHKAI</i>	GWGCWC	0.000154	Intergenic	6.30	0.000970	1030
			Intron	5.00	0.000770	1300
			Exon	6.05	0.000932	1070
3) <i>NcoI</i>	CCATGG	0.000154	Intergenic	1.28	0.000197	5070
			Intron	1.79	0.000276	3620
			Exon	1.62	0.000249	4020

Detection fraction (inter.) = $F1/(F1+F2+F3) = 0.00242/(0.00242 + 0.000970 + 0.000197) = 68\%$
 Detection fraction (intron) = $F1/(F1+F2+F3) = 0.00318/(0.00318 + 0.000770 + 0.000276) = 75\%$
 Detection fraction (exon) = $F1/(F1+F2+F3) = 0.00283/(0.00283 + 0.000932 + 0.000249) = 71\%$

¹ In the recognition site column, W represents a degenerate base that could be A or T.

² The calculated preference for the nucleotide sequence of the restriction enzyme site as described in Table 5.3 and Section 5.3.2.

³ The frequency (F) of restriction enzyme sites was calculated by multiplying the frequency based on the GC% content of the human genome by the total preference of the sequence based on the dinucleotide sequences. For degenerate sequences of *Bsi*HKAI, the total frequency was the F based on GC% for all four possible sequences was multiplied by the sum of the dinucleotide preferences of all four possible sequences.

recognition site for *DpnI* compared to the frequency of the 6-nucleotide RE recognition site for *NcoI*.

5.3.3 Large hepatocyte clones were observed in DNA extracts from 5 mg liver fragments by invPCR

Virus-cell DNA junctions were detected in human liver tissues by the *NcoI* and *DpnII* invPCR assays. DNA was extracted from ~5 mg fragments of resected liver tissue from 20 chronic HBV patients; of which 2 were early stage HBV patients, 10 were cirrhotic patients, 6 were fulminant hepatitis patients and 1 had concurrent HCC. DNA from 5 fragments of liver tissue from each patient was analysed by invPCR. Most probable number statistical methods, as described in Section 2.14, were used to determine the copy number of unique virus-cell DNA junctions from invPCR results and therefore hepatocyte clone sizes. The HBV DNA integration sites detected by invPCR were identified using a BLAST search for sequence homology to the human genome.

The virus-cell DNA junctions detected in DNA extracts from each of the patient tissues are summarised in Table 5.5. The virus-cell DNA junctions are listed fully along with their sequences in Appendix 9.8. Due to the high level of uncertainty in calculated clone sizes represented by virus-cell DNA junctions detected only once, these were excluded in further analyses involving hepatocyte clone size.

Repeated virus-cell DNA junctions were found in liver tissue from 9 out of 20 patients assayed (Table 5.5). The copy numbers of virus-cell DNA junctions were then used to predict the most probable sizes of hepatocyte clones, which ranged from 40 to 20000 cells. Three of these repeated virus-cell DNA junctions (assigned the identifiers XA v, CY iv, and HN iv) indicated clones composed of >10000 cells. Identical virus-cell DNA junctions were detected in multiple fragments of tissues from CY and XA patients, indicating high levels of clonal proliferation (Figure 5.5).

When the virus-cell DNA junctions found in each of the patient cohorts were compared, significantly larger clones were found in DNA extracts of tissues from cirrhotic and fulminant hepatitis patients compared to early stage and the HCC patient DG (Figure 5.6). However, low numbers of patients in each cohort and of detected virus-cell DNA junctions limited

Table 5.5. Clonal proliferation was observed in DNA extracted from ~5mg liver tissue fragments

Patient ¹	Total DNA extracted from fragments ² (µg)	Clones sizes found by <i>NcoI</i> invPCR ³ (MPN±95%CI)	Clones sizes found by <i>DpnII</i> invPCR ⁴ (MPN±95%CI)
C	10	xvi) 670 (310-1500) xvii) 93 (13-660)* xviii) 93 (13-660)* xx) 93 (13-660)* xv) 90 (12-660)* ix) 41 (15-110) xi) 31 (10-98) v) 10 (1.4-73)* xii) 10 (1.4-73)* xiii) 10 (1.5-73)* xiv) 10 (1.5-73)*	NT
L	8	ND	NT

CN	28	ND	ND
CYRY	23	ii) 41(5.8-290)* iii) 41(5.8-290)* i) 40(5.5-290)* v) 40(5.5-290)*	NT
CY	27	iv) 2700(1600-4500) iv) 3600(2200-5900) iv) 4500(2800-7300) iv) 800(360-1800)	NT
FMC1	33	i) 67 (9.5-470)*	NT
FMC5	46	i) 1200 (700-2100) ii) 67 (9.5-470)*	NT
FMC6	36	iii) 1200 (670-2100) i) 360 (150-870) i) 690 (650-1400) ii) 67 (9.5-470)*	NT
MH	39	i) 3300(2000-5500)	NT
NT	73	ND	ND
SAAO	88	i) 510(270-950) xvi) 83(21-330) x) 40(5.5-290)* i) 40(5.5-290)* viii) 40(5.5-290)*	xi) 220(93-540) ix) 170(61-460) iii) 130(43-400) iv) 130(40-390) ii) 83(21-330) xiii) 82(20-330) xiv) 41(5.8-290)* vii) 41(5.8-290)* v) 40(5.5-290)* xv) 40(5.5-290)* xii) 40(5.5-290)*

Table 5.5. Cont.

Patient ¹	Total DNA extracted from fragments ² (µg)	Clones sizes found by <i>NcoI</i> invPCR ³ (MPN±95%CI)	Clones sizes found by <i>DpnII</i> invPCR ⁴ (MPN±95%CI)
XA	51	v) 4500(2800-7200) v) 3600(2200-5900) v) 980(590-1600) v) 41(5.8-290) v) 40(5.5-290) x) 40(5.5-290)*	v) 6100(2500-15000) iii) 6000(2400-15000) iv) 5800(2400-14000) ii) 2300(590-9200) i) 40(5.5-290)* viii) 40(5.5-290)*

FMC2	34	ND	NT
FMC3	81	i) 510 (240-1100) ii) 410 (100-1700) iii) 67 (9.5-470)*	ii) 890 (330-2400) iv) 67 (9.5-470)* v) 67 (9.5-470)*
FMC4	61	i) 67 (9.5-470)*	NT
FMC7	29	ND	NT
HN	23	i) 40(5.5-290)* ii) 40(5.5-290)*	iv) 20000(4900-79000) iii) 41(5.8-290)*
HS	25	ND	ND

DG	136	i) 610(340-1100) vii) 240(60-1000) ii) 40(5.5-290)* iii) 40(5.5-290)* iv) 40(5.5-290)* v) 40(5.5-290)* vi) 40(5.5-290)*	NT
----	-----	---	----

WN	34	i) 40(5.5-290)* ii) 40(5.5-290)* iii) 40(5.5-290)* iv) 40(5.5-290)* v) 40(5.5-290)* vi) 40(5.5-290)* vii) 40(5.5-290)* viii) 40(5.5-290)*	NT
----	----	--	----

¹ invPCR was used to detect virus-cell DNA junctions in DNA extracts from 5 mg liver tissue fragments from patients for which resection or needle biopsy tissue was available.

² The amount of extracted DNA from the tissues from each patient was approximated by optical densitometry at 280nm. A tenth of the total DNA extracted was used for each invPCR design.

³ Clone sizes were determined by enumeration and detection of unique virus-cell DNA junctions using invPCR in conjunction with the MPN method, calculated by MPN calculator (Curiale, 2000). Virus-cell DNA junctions that were observed only in once in an extract are indicated with “*” and were not included in further analyses due to high uncertainty in clone sizes. Patient DNA extracts in which no virus-cell DNA junctions were detected are represented by ND. Coloured numbers represent virus-cell DNA junctions observed in multiple tissue fragments. Each virus-cell junction was given a unique identifier (i, ii, iii, and so on) and can be matched up with DNA sequences and other details fully listed in Appendix 9.8.

⁴ The *DpnII* invPCR design was not compatible with HBV DNA sequences found in some patients (shown in grey) and therefore not tested (NT).

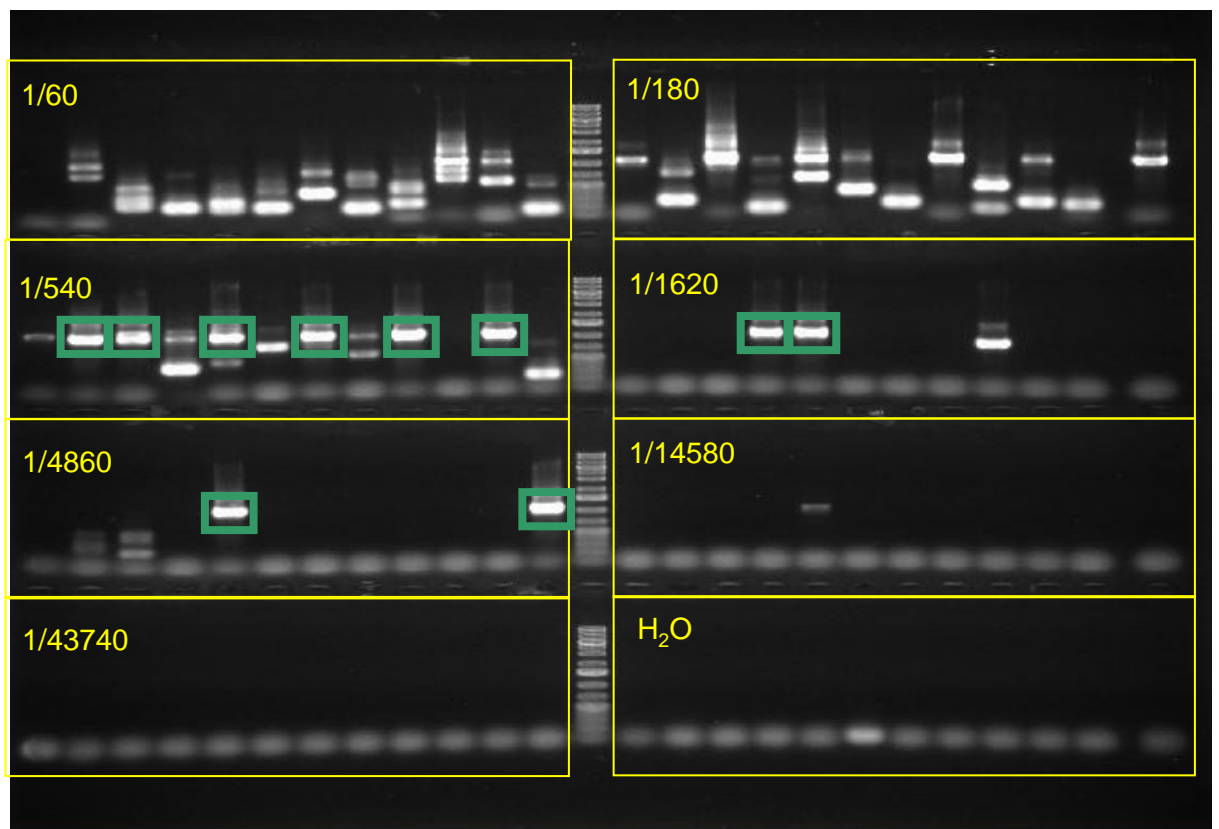
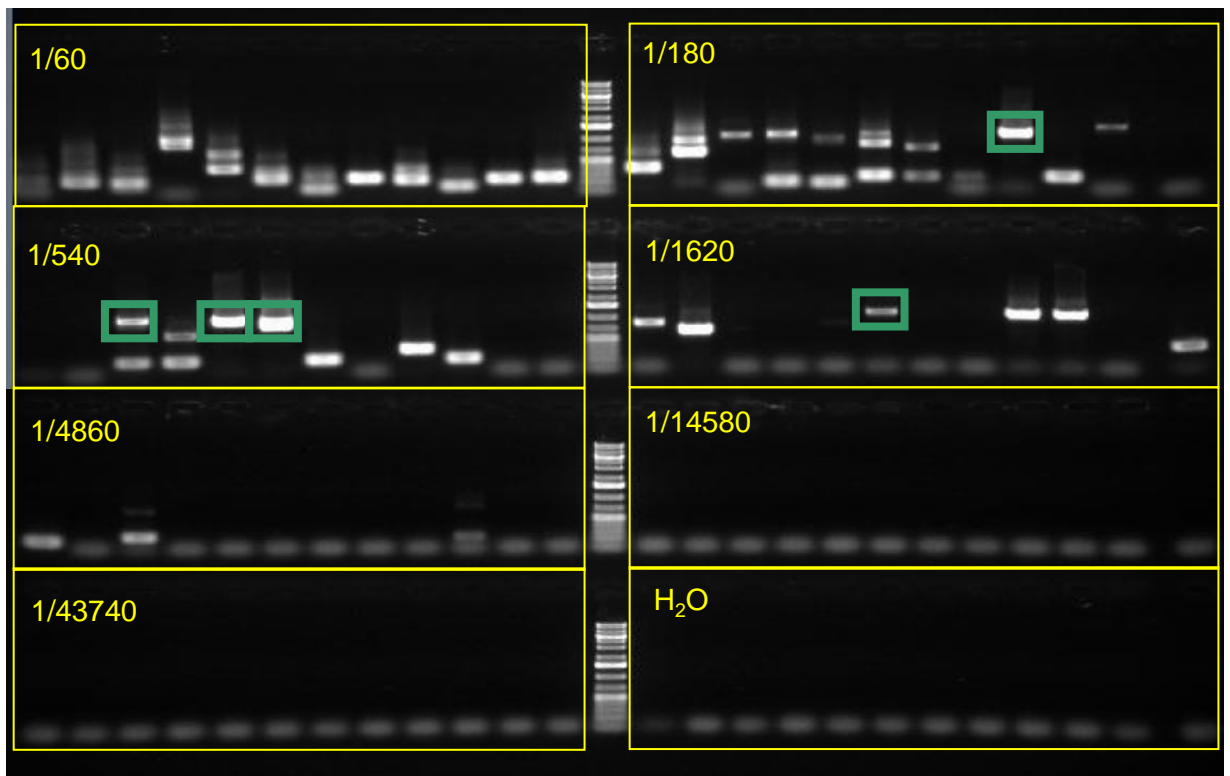


Figure 5.5. Virus-cell DNA junction CY iv was observed in two DNA extracts from 5 mg liver tissue fragments of Patient CY. Virus-cell DNA junctions in liver DNA extracts from Patient CY were detected by invPCR as described in Section 2.14. Dilutions of the template DNA with respect to the original extract are listed in each yellow box. Sanger sequencing of DNA products confirmed that reoccurring products (green boxes) were identical inverted sequences containing the hepatocyte clone CY iv.

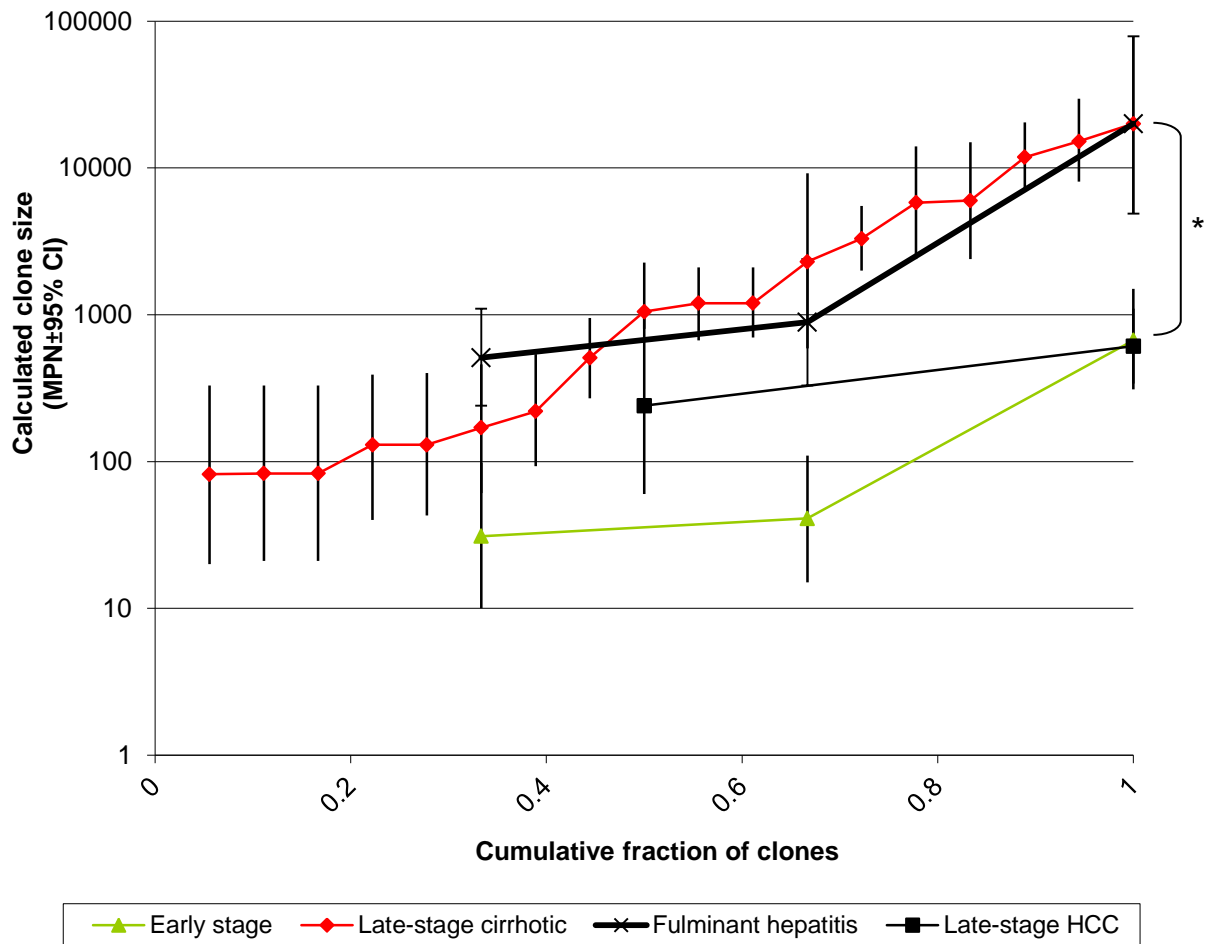


Figure 5.6. Hepatocyte clones detected by invPCR of total DNA extracts from 5 mg HBV-infected tissue fragments of early-stage , cirrhotic, fulminant hepatitis and HCC patients. Only clone sizes calculated using virus-cell DNA junctions detected multiple times are represented in this graph. Clones observed by invPCR as described in Section 2.14 in DNA extracts of cirrhotic patients were significantly larger than of both early stage patients and the HCC patient DG (* $p < 0.05$, Student's t-test of square root transformed values). The clones that are displayed in the graph are summarised by patient extract in Table 5.5 and fully listed in Appendix 9.8.

interpretation and could not conclusively show association between disease stage and clonal proliferation.

5.3.4 Characterisation of large hepatocyte clones found in patients XA and CY

The virus-cell DNA junctions that represented large hepatocyte clones (>10000 cells) found in multiple pieces of liver tissue from patients CY and XA, designated as CY iv and XA v respectively, were analysed in greater depth. These clones were large enough that they could be characterised further by qPCR and PCR-based methods.

The virus-cell DNA junction CY iv was integrated into the ORF of hypothetical protein LOC85379 (aka., KIAA1671) on chromosome 22. The virus-cell DNA junction XA v was integrated into the open reading frame (ORF) of ciliary neurotrophic factor (CNTF) on chromosome 11. Upstream and downstream PCR primers were designed using Primer 3 that bound to each virus-cell DNA junction and to a sequence downstream of the junction in the cellular DNA as shown in Figure 5.7A and 5.7B respectively. The sequence of each upstream and downstream primer are listed in Table 5.6.

Using these primers, the virus-cell DNA junctions could be amplified by PCR from the same uninverted DNA extracts as used in invPCR assays (Figure 5.8). This confirmed that observed virus-cell DNA junctions were not artefacts of the inversion reaction, but had been produced by integration of HBV DNA into the host cell genome.

Next, to independently confirm the copy number of the CY iv and XA v virus-cell DNA junctions in the liver DNA extracts of Patients CY and XA, a qPCR assay was developed by first cloning the invPCR products into a pCR2.1-TOPO vector using a TOPO TA cloning kit (Invitrogen).

The inserted sequences in the resultant plasmids, designated pTOPOCYiv and pTOPOXAv respectively, were confirmed by both RE analysis and sequencing. First, the plasmids digested with *EcoRI* to liberate the inserted sequence. Digested fragments were separated by agarose gel electrophoresis and inserts of the expected sizes were observed, as shown in Figure 5.9. M13 sequences flanked the insertion site in the pCR2.1-TOPO vector. Thus, the

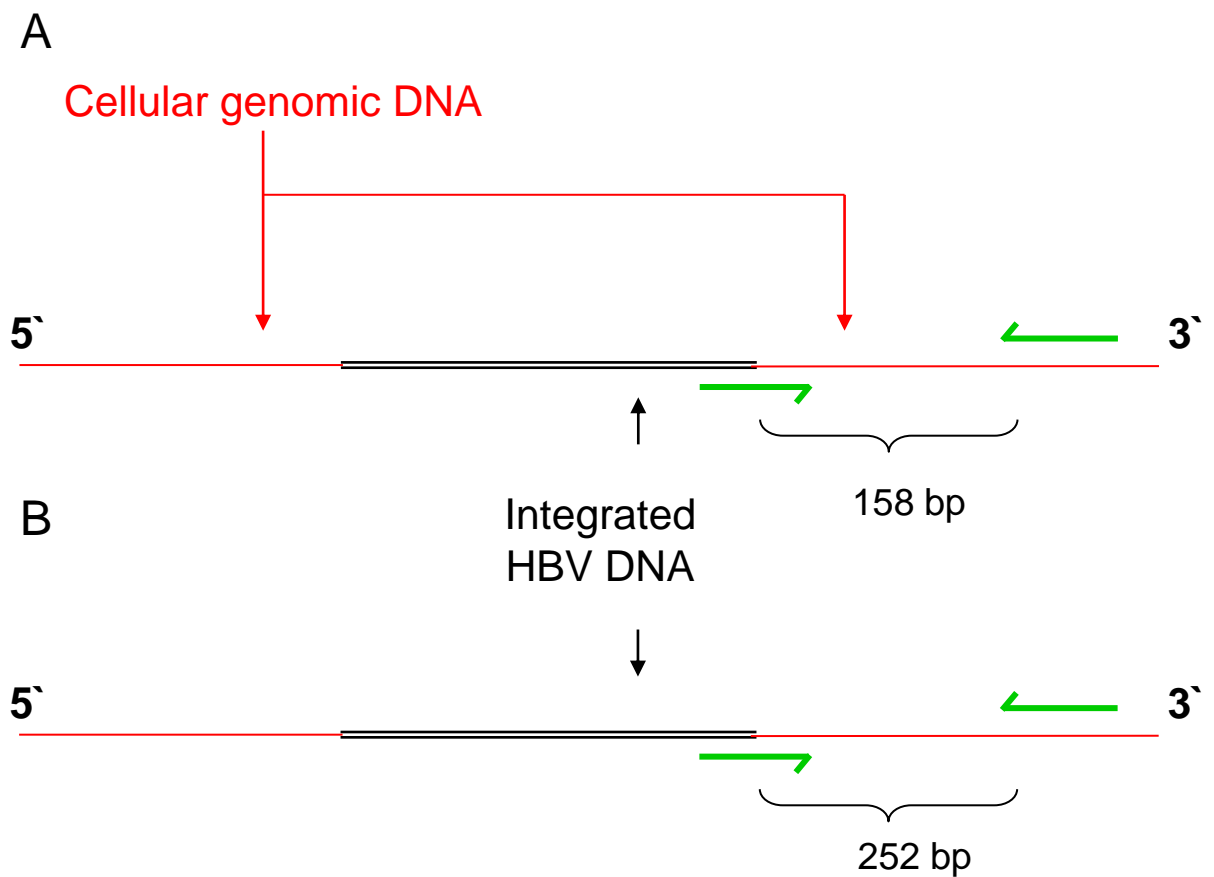


Figure 5.7. Primers were designed to amplify specific virus-cell DNA junctions in DNA extracts from patient CY (5.7A) and XA (5.7B). Schematic figure of positions of primers to detect the specific virus-cell DNA junctions, CY iv and XA v, using qPCR as described in Section 2.18. Full details of these virus-cell DNA junctions are described in Table 5.6. The black line (-) represents the integrated HBV DNA, the red line (-) represents the flanking cellular DNA, and the green arrows (\rightarrow) represent the primers annealing to complementary sequences.

Table 5.6. Sequences of the primers used to amplify the specific virus-cell DNA junctions, CY iv and XA v, in DNA extracts from patients CY and XA.

Primer name	Sequence (5' -> 3')	Tm (°C) ¹	Length (nt)
XA clone F	agcgccatgcggaagata	50	18
XA clone R	gtggaagtgtcaaattccttgagc	53	23
CY clone F	caccatgcgcttcttgctt	50	18
CY clone R	aaaaggtgactgagagcgtca	52	21

¹ Melting temperature (Tm) calculated by BioMath oligonucleotide calculator (Promega, <http://www.promega.com/techserv/tools/biomath/calc11.htm>) using salt-adjusted, base-stacking calculations.

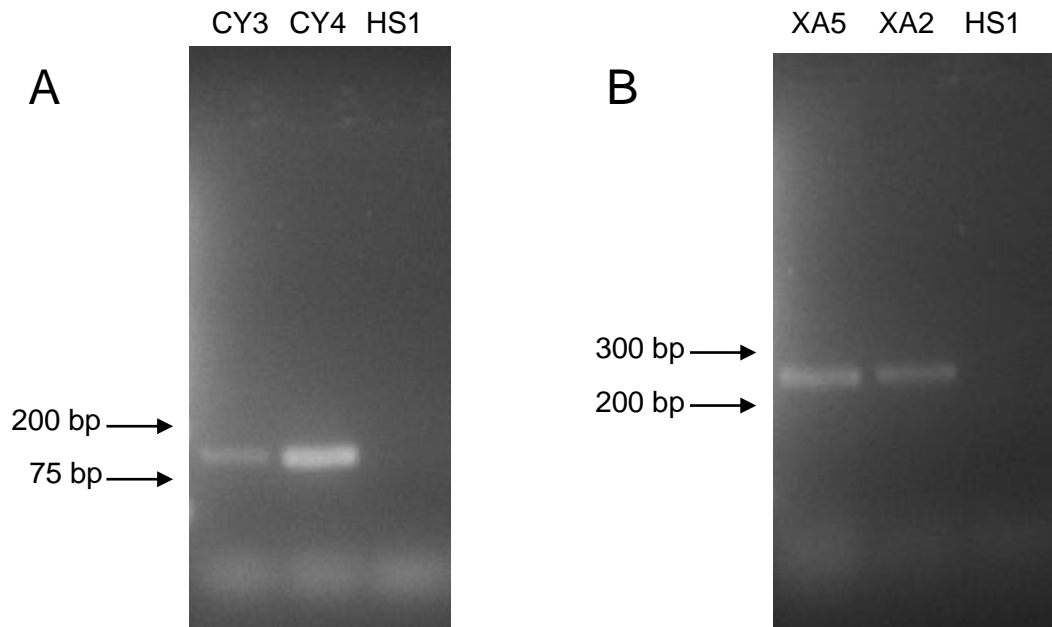


Figure 5.8. The specific virus-cell DNA junctions, CY iv and XA v, were amplified by PCR using uninverted DNA extracts from patient CY (5.8A) and XA (5.8B) as a template. Two out of the five DNA extracts of separate liver fragments from each patient (3rd and 4th for Patient CY; 5th and 3rd for Patient XA) and a DNA extract from an unrelated patient HS were used as a template for PCR to amplify specific virus-cell junctions as described in Section 2.18. Expected products at 158 bp and 252 bp respectively were amplified in uninverted extracts from patients CY (5.8A) and XA (5.8B) using primers described in Table 5.7. No products are visible in the negative controls containing DNA extracts from an unrelated infected patient.

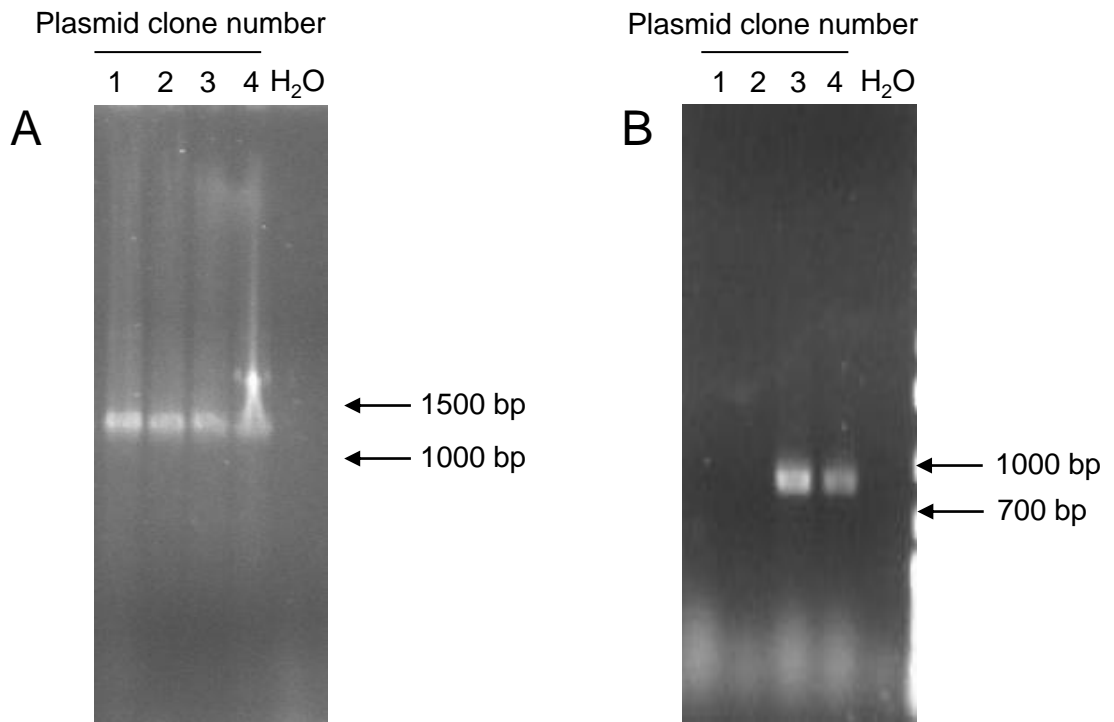


Figure 5.9. PCR of pCR2.1TOPO subcloned with the virus-cell DNA junctions, CY iv (pTOPOCYiv, 5.9A) and XA v (pTOPOXAv, 5.9B). Plasmids containing virus-cell DNA junctions were created, as described in Section 2.13.1. DNA extracts from 4 transformed colonies were used as a PCR template with M13 primers to amplify inserted sequences. Expected products of ~1200 bp and ~850 bp for CY iv and XA v were amplified by PCR respectively, confirming successful subcloning.

DNA inserts in pTOPOCYiv and pTOPOXAv were PCR amplified using M13-specific primers and the resultant products were sequenced. The expected insertion sequences were confirmed after alignment with the sequence of each original virus-cell DNA junction, as shown in Figure 5.10.

The plasmids pTOPOCYiv and pTOPOXAv were then diluted from 10^8 to 10^1 copies and used to generate standard curves for each qPCR assay as shown in Figures 5.11A and 5.12A. The standard curves showed good agreement within triplicates and efficient amplification.

Using the pTOPOCYiv and pTOPOXAv standard curves, qPCR was used to analyse the copy number of each virus-cell DNA junctions in uninverted DNA extracts from patients CY and XA. Single peaks were observed in each melt curve with melt temperatures of 82.5 and 78.9°C, respectively (Figures 5.11B and 5.12B). This is concordant with the respective melting temperatures of 83.8 and 82.4°C calculated using the online oligonucleotide calculator OligoCalc, developed by Kibbe (Kibbe 2007). This indicated specific amplification of the virus-cell DNA junctions in each DNA extract.

The copy numbers of each virus-cell DNA junction detected by qPCR were larger with those detected previously by invPCR (Table 5.7 and Figure 5.13). In total of all 5 DNA extracts from 5mg tissue fragments, the copy numbers of CY iv and XA v detected by qPCR were 33800 and 31680 respectively. These were significantly higher than the totals detected by invPCR - 11840 and 15180, respectively. The hepatocyte clone sizes detected by invPCR were, on average, 37% of the clone sizes detected by qPCR in the respective DNA extracts.

As previously described in Section 5.3.2, multiple factors during the invPCR protocol could limit the detection of virus-cell DNA junctions by invPCR. The factors that could have led to this decrease in copy number include: the efficiency of RE digestions, the efficiency of the circularisation of the excised virus-cell DNA junctions, and the sensitivity of the nested PCR, the interference of HBV RI DNA.

Thus, clones found in cirrhotic patients CY and XA were confirmed to be repeated in DNA extracts from multiple liver tissue fragments. This suggests that clones are larger than the size

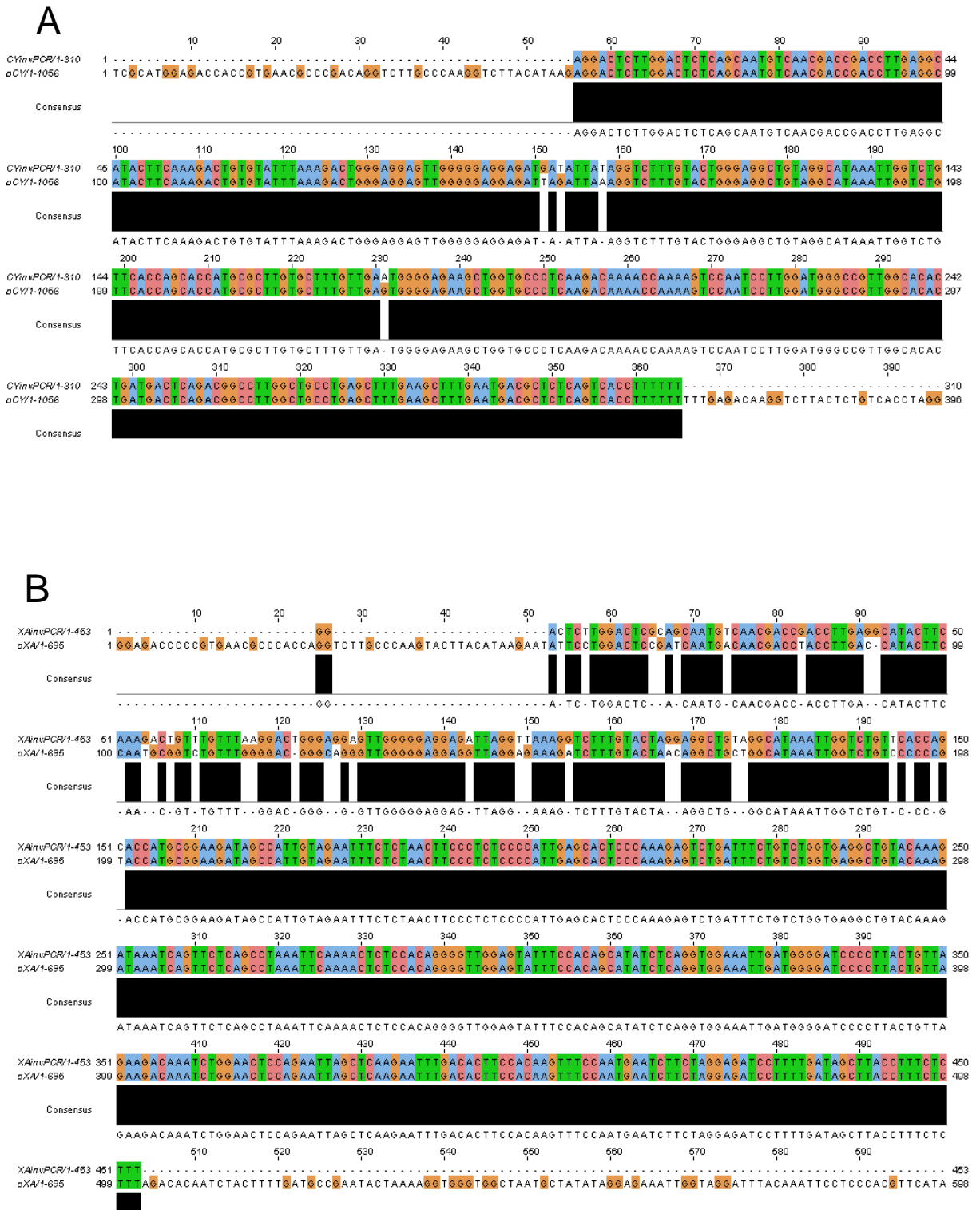


Figure 5.10. Sequence alignment the virus-cell DNA junctions, CY iv (5.10A) and XA v (5.10B), amplified by invPCR and PCR-amplified products from pTOPOCYiv and pTOPOXAv. For each set of aligned sequences, the top sequence is the sequence of products amplified by invPCR, while the bottom sequence is that of subcloned plasmids. Images produced using Jalview.

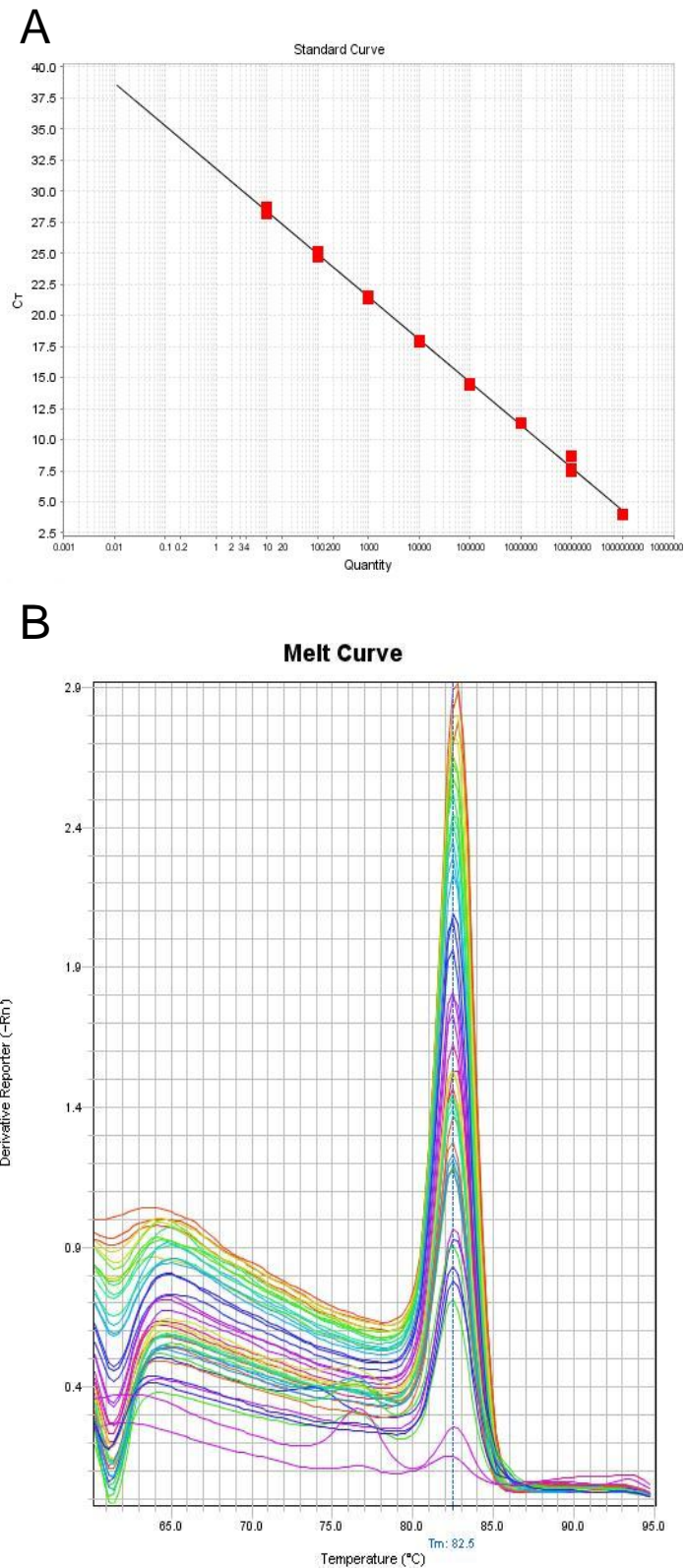


Figure 5.11. Standard curve (5.12A) and melt curve (5.12B) of qPCR analysis of virus-cell DNA junction CY iv. qPCR analysis of DNA extracts to quantify virus-cell DNA junction CY iv was conducted as described in Section 2.13 in triplicate. The standard curve was produced by amplification and quantification of pTOPOCYiv at copy numbers from 10^8 to 10^1 per reaction. Melt curve includes standards and experimental samples from five DNA extracts from separate 5 mg liver tissue fragments. Results of the qPCR are shown in Table 5.8.

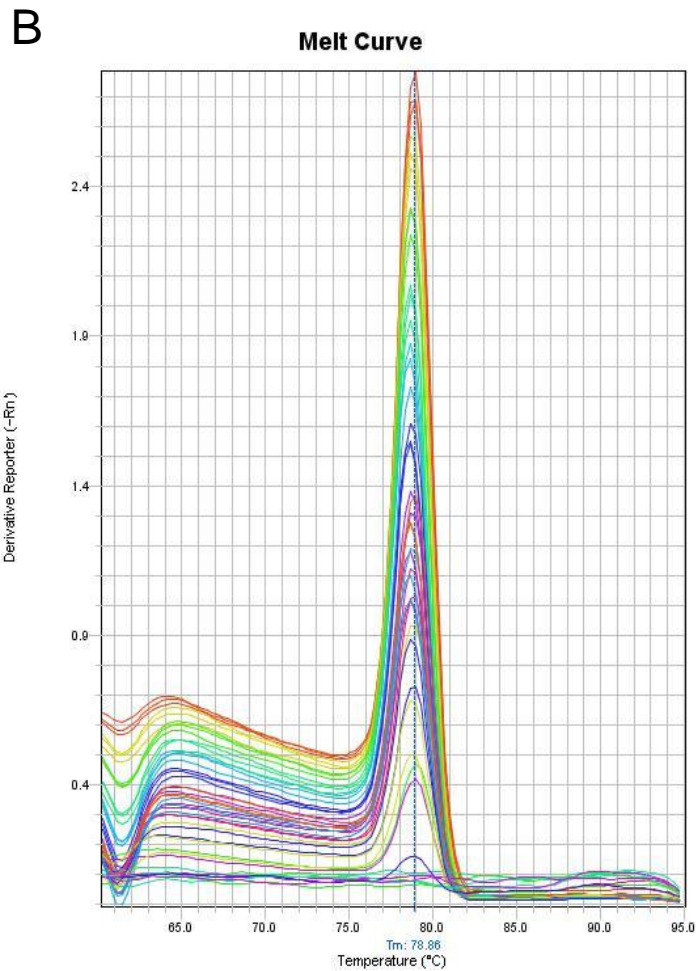
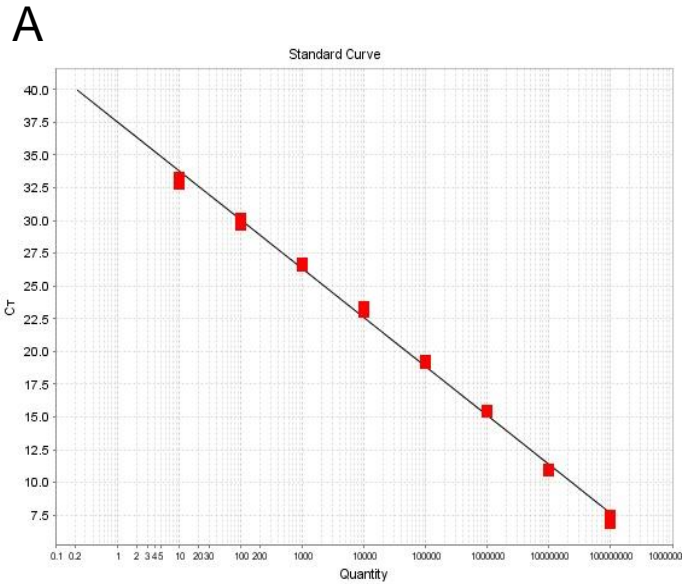


Figure 5.12. Standard curve (5.12A) and melt curve (5.12B) of qPCR analysis of virus-cell DNA junction XA v. qPCR analysis of DNA extracts to quantify virus-cell DNA junction XA v was conducted as described in Section 2.13 in triplicate. The standard curve was produced by amplification and quantification of pTOPOXAv at copy numbers from 10^8 to 10^1 . Melt curve includes standards and experimental samples from four DNA extracts from separate 5 mg liver tissue fragments. Results of the qPCR are shown in Table 5.8. qPCR analysis was conducted in triplicate.

Table 5.7. Comparison of copy number of CY iv and XA v as measured by invPCR and qPCR. These results are graphed in Figure 5.13.

Tissue extract¹	Clone sizes detected by invPCR (MPN (95% CI))²	Clone sizes detected by qPCR (Mean (95% CI))
CY1	3600 (2200-5900)	4880 (4840-4920)
CY2	4500 (2800-7300)	13700 (13700-13800)
CY3	240 (59-990)	980 (940-1020)
CY4	2700 (1600-4500)	10600 (10500-10600)
CY5	800 (360-1800)	3640 (3580-3700)
Total	11840 (7019-20490)	33800 (33560-34040)

XA2	6100 (2500-15000)	26300 (26200-27400)
XA3	4500 (2800-7200)	2000 (1970-2050)
XA4	3600 (2200-5900)	ND
XA5	980 (590-1600)	3380 (3340-3430)
Total	15180 (8090-29700)	31680 (31510-32800)

¹ Total DNA was extracted from five separate 5 mg liver tissue fragments from each patient.

² invPCR clone sizes as previously shown in Table 5.6A.

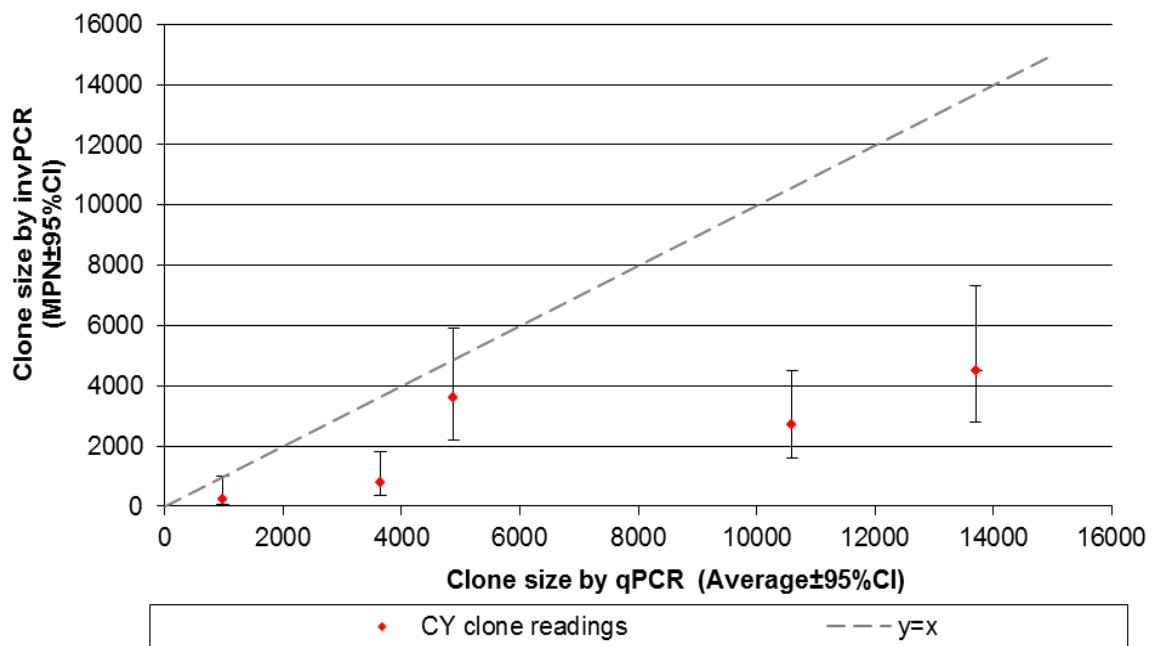
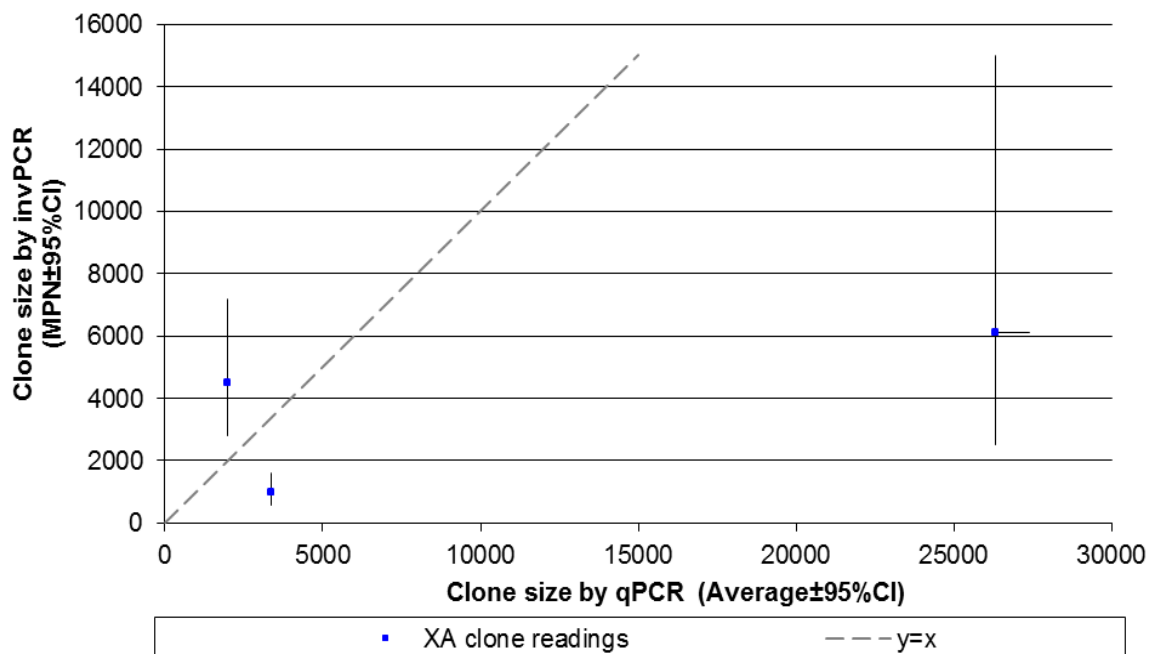
A**B**

Figure 5.13. Comparison of copy number of CY iv (5.13A) and XA v (5.13B) as measured by invPCR and qPCR. The dashed grey lines represent a 1:1 correspondence of copy numbers detected by invPCR and qPCR. For the majority of measurements, the copy numbers detected by qPCR were larger than the respective copy numbers detected by invPCR. On average, the invPCR clone sizes were 36% of their respective qPCR clone sizes.

predicted by invPCR and qPCR assays as portions of the large clones may remain in the unextracted portion of tissue.

Attempts were made to characterise the left-hand junction at the other end of the integrated HBV DNA found in patient XA. Assumptions made in the design were that: i) the entire length of HBV dsDNA was inserted into the cellular genome; ii) the HBV dsDNA had integrated by NHEJ and was integrated into the cellular DNA at approximately the same as the right-hand end and iii) there were no deletions of >1000 bp in the cellular DNA. Primers specific for HBV DNA 500 bp downstream from the expected left-hand junction and specific for regions 500 or 1000 bp upstream in the cellular DNA in the CNTF ORF were designed, as shown in Figure 5.14.

No PCR products could be amplified using these primers. This suggested that at least one of our assumptions was wrong and that chromosomal rearrangement or a large deletion of either the HBV DNA or cellular gene had occurred during integration.

5.3.5 InvPCR analysis of DNA extracted from sections of ethanol-fixed liver tissue

Having determined that hepatocyte clones could be detected in 5 mg fragments of liver tissue, experiments were designed to identify hepatocyte clones histologically and determine the spatial arrangement of the hepatocyte clones within the liver. Pieces of frozen liver tissue ~0.5 cm³ in size from livers of patients that were identified as having multiple identical virus-cell DNA junctions by invPCR were cut, fixed in 70% ethanol and embedded in paraffin wax, as described in Section 2.8. Ethanol-fixed, slide-mounted liver tissue sections from Patients Y2 to Y6 were also used in this analysis. Liver samples from Patients Y2 to Y6 were not analysed as 5 mg samples, as only slide-mounted liver tissue was available to us.

DNA was extracted from two 5 µm sections of each paraffin wax-embedded tissue and analysed by invPCR, as described in Section 2.14. The *NcoI* invPCR protocols were used on all tissues. The *DpnII* invPCR protocol was only compatible with HBV sequences amplified from with patients FMC3, GS4, SAAO, XA, Y2, Y4 and Y6. All virus-cell DNA junctions detected in slide-mounted liver tissue by invPCR are summarised by patient in Table 5.8 and their sequences are listed in Appendix 9.8. Representative gel electrophoresis of invPCR products from cirrhotic and non-cirrhotic patients is shown in Figure 5.15.

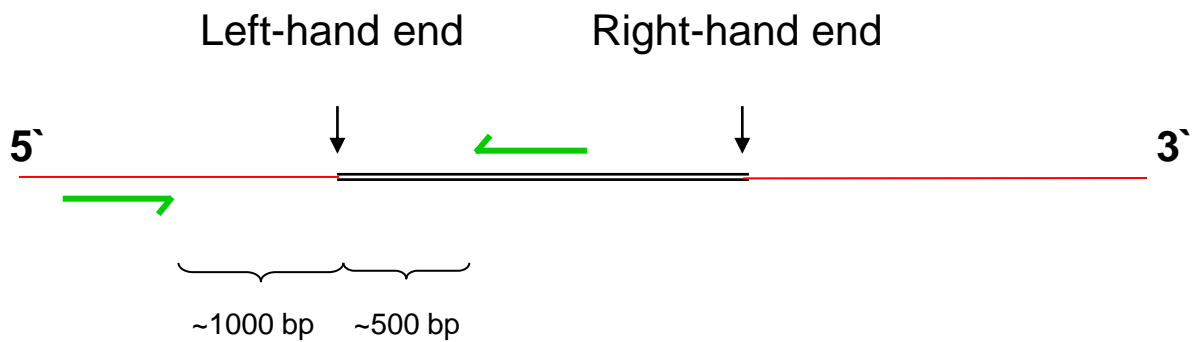
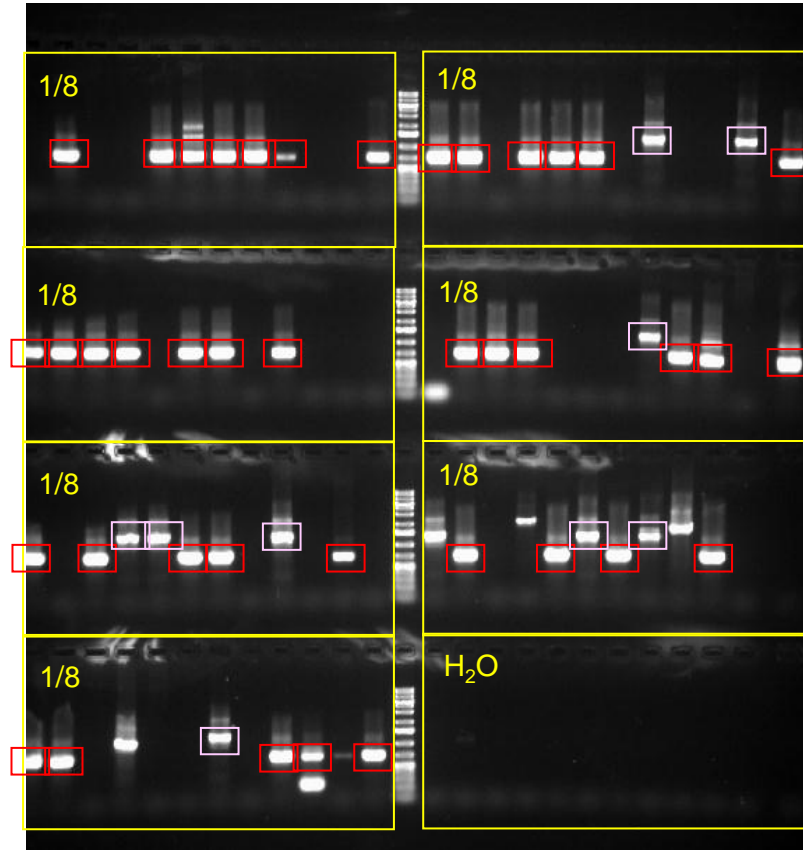


Figure 5.14. Primers were designed to amplify cellular sequences upstream of the expected left-hand end of the virus-cell DNA junction XA v. Attempts to characterise the left-hand virus-cell DNA junction were undertaken as described in Section 5.3.4. Above is a schematic figure of positions of primers to detect cellular DNA sequences upstream of the observed virus-cell DNA junction, assuming NHEJ-mediated integration of the entire genome into the cellular DNA. The black line (-) represents the integrated HBV DNA, the red line (-) represents the flanking cellular DNA, and the green arrows (\rightarrow) represent the primers annealing to complementary sequences.

A



B

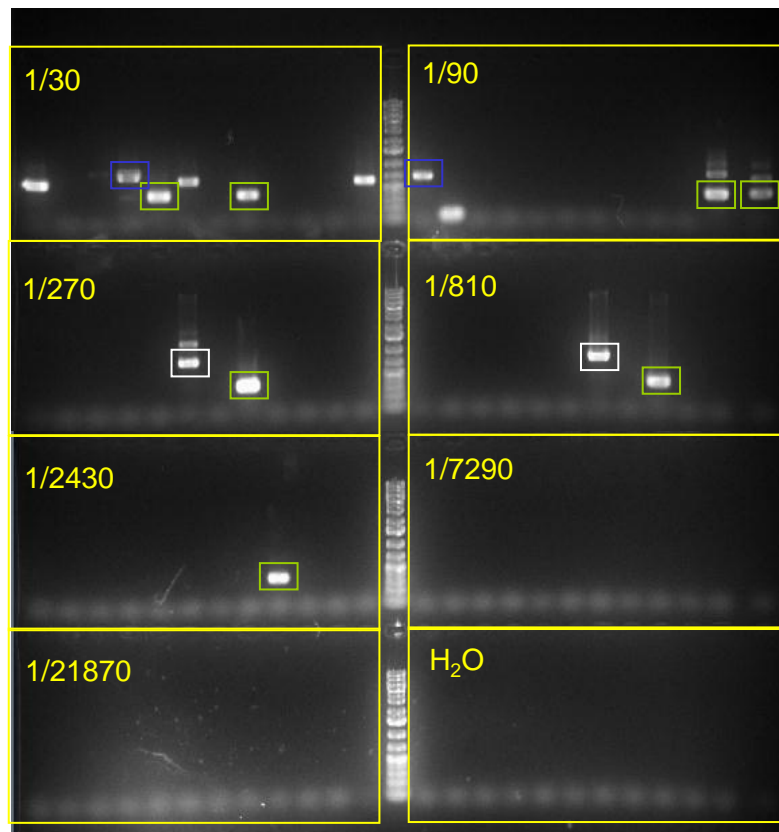


Figure 5.15. Representative invPCR results of DNA extracts from slide mounted liver tissue from cirrhotic Patient XA (5.15A) and non-cirrhotic HCC Patient Y4 (5.15B). Integrated HBV DNA was detected in DNA extracted from slide-mounted sections of liver tissues using the *NcoI* invPCR design as described in Section 2.14. Dilutions of the inverted DNA as a fraction of the original DNA extract are shown. Inverted DNA from patient XA (5.12A) was not serially diluted prior to nested PCR, as such high numbers of invPCR products were not anticipated. invPCR products that were confirmed to be derived from virus-cell DNA junctions using Sanger sequencing (boxed bands). Virus-cell DNA junctions with the same sequence are highlighted in the same colours. The following clones are represented: XA v (red), XA xi (pink), Y4 xxv (green), Y4 xxvii (white), and Y4 xxvi (blue).

Table 5.8. Clonal proliferation was observed by invPCR of DNA extracted from slide-mounted liver tissue from early-stage, cirrhotic, fulminant hepatitis and HCC patients.

A – Early-stage HBV infection

Patient ¹	Secti- ons (#)	NcoI invPCR design		DpnII invPCR design ⁴	
		Clones sizes in slide (95%CI) ²	Min 3D clone size (95%CI) ³	Clones sizes in slide (95%CI)	Min 3D clone size (95%CI)
C	2	viii) 9.5 (4.3-21) i) 4.6 (1.5-14) iv) 3 (0.76-12) v) 3 (0.76-12) vi) 3 (0.76-12) vii) 3 (0.76-12) ii) 1.5(0.21-11)* iii) 1.5(0.21-11)*	48(15-157) 16(3-85) 8(1-68) 8(1-68) 8(1-68) 8(1-68) 3(1-60)* 3(1-60)*	NT	NT
GS1	1	i) 41(5.8-290)*	428(22-8058)*	NT	NT
GS2	1	ND	ND	NT	NT
GS3	1	ND	ND	NT	NT
GS4	1	ND	ND	ND	ND
GS5	1	ND	ND	NT	NT
L	2	iii) 78(33-190) i) 15(2.1-110)* ii) 15(2.1-110)* iv) 15(2.1-110)*	1124(309-4273) 146(8-2998)* 146(8-2998)* 146(8-2998)*	NT	NT

B – HBV-positive patients with cirrhosis

CY	4	i) 16(2.3-110)* ii) 16(2.3-110)* iii) 16(2.3-110)* iv) 16(2.3-110)*	104(5.7-1882)* 104(5.7-1882)* 104(5.7-1882)* 104(5.7-1882)*	NT	NT
CYRY	1	iii) 32(8.1-130) iv) 16(2.3-110)* vi) 16(2.3-110)*	295(38-2419) 104(5.7-1882)* 104(5.7-1882)*	NT	NT
FMC5	2	iv) 330(110-1000) iii) 320(100-1000) vi) 310(99-980) v) 100(15-720)* vii) 100(15-720)*	9782(1882-51599) 9340(1632-51599) 8906(1607-50059) 1692(95-31524)* 1692(95-31524)*		
FMC6	2	ND	ND	NT	NT
MH	4	ND	ND	NT	NT
SAAO	1	ND	ND	vi) 8(1.1-57)*	37(1.8-702)*
XA	1	v) 420(300-540) xi) 24(7.9-76)	14045(8478-20475) 192(36-1081)	NT	NT

C – HBV-positive patients with fulminant hepatitis

FMC3	2	ND	ND	ND	ND
HN	2	ND	ND	NT	NT

Table 5.8. Cont.

D – HBV-positive patients without cirrhosis and with concurrent HCC

Patient	Sections (#)	NcoI invPCR design		DpnII invPCR design ³	
		Clones sizes in slide (95%CI) ¹	Min 3D clone size (95%CI) ²	Clones sizes in slide (95%CI)	Min 3D clone size (95%CI)
DG	4	xii) 3100(1900-4900) viii) 280(150-530) x) 270(140-510) iv) 250(130-490) xiv) 200(97-410) xv) 60(18-200) xi) 43(11-170) xvi) 43(11-170) xiii) 21(2.9-150)*	281635(135137-559678) 7645(2998-19909) 7239(2703-18793) 6449(2419-17698) 4615(1559-13546) 758(125-4615) 460(60-3617) 460(60-3617) 157(8-2998)*	NT	NT
Y2	1	xiv) 63(20-200) iii) 21(2.9-150)* x) 21(2.9-150)* xiii) 21(2.9-150)* ii) 20(2.8-150)* iv) 20(2.8-150)* v) 20(2.8-150)* vi) 20(2.8-150)* vii) 20(2.8-150)* xv) 20(2.8-150)* xvi) 20(2.8-150)*	816(146-4615) 157(8-2998)* 157(8-2998)* 157(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)*	viii) 120(30-500) i) 20(2.8-150)* xii) 20(2.8-150)*	2145(268-18243) 146(8-2998)* 146(8-2998)*
Y3	1	i) 20(2.8-150)* ii) 20(2.8-150)* iii) 20(2.8-150)* iv) 20(2.8-150)*	146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)*	NT	NT
Y4	1	xxv) 150(71-330) xxvii) 42(10-170) xxi) 40(9.8-160) xxvi) 40(9.8-160)	2998(976-9782) 444(52-3617) 413(50-3302) 413(50-3302)	x) 20(2.8-150)* xiii) 20(2.8-150)* xix) 20(2.8-150)* xx) 20(2.8-150)* xxiii) 20(2.8-150)* xxiv) 20(2.8-150)* xxviii) 20(2.8-150)*	146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)* 146(8-2998)*
Y5	1	vi) 86(32-230) iv) 63(20-200) i) 40(9.9-170) vii) 21(2.9-150)* viii) 21(2.9-150)* v) 20(2.8-150)*	1301(295-5692) 816(146-4615) 413(50-3302) 157(8-2998)* 157(8-2998)* 146(8-2998)*	NT	NT
Y6	1	iii) 550(270-1100) ii) 400(320-680)	21046(7239-59530) 13054(9340-28934)	i) 20(2.8-150)	146(8-2998)

¹ InvPCR of slide-mounted tissue from cirrhotic patients was only carried out if the particular patient was positive for multiple identical virus-cell DNA junctions after invPCR of DNA extracted from liver fragments.

² Clone sizes were determined by enumeration and detection of unique virus-cell DNA junctions using invPCR in conjunction with the MPN method, calculated by MPN calculator (Curiale, 2000). Virus-cell DNA junctions that were observed only in once in an extract are indicated with “*” and were not included in further analyses due to high uncertainty in clone sizes. Patient extracts in which no virus-cell DNA junctions were detected are represented by ND. Each virus-cell junction was given a unique identifier (i, ii, iii, and so on) and can be matched up with DNA sequences and other details fully listed in Appendix 9.8.

³ Minimum clone sizes in 3 dimensions were calculated as outlined in Section 2.22.

⁴ The *DpnII* invPCR design was only used on DNA extracts from some patients, as DNA extracts from other patients were incompatible with the *DpnII* invPCR design and were not tested (NT).

The size of clones found on tissue sections were converted to minimum sizes in 3 dimensions as outlined in Section 2.22. In this calculation it was assumed that: i) the hepatocyte clone was spherical; and ii) that the section had cut through the centre of the sphere, or the equator. This should underestimate the size of the hepatocyte clones, as the calculated clones would be smaller than the actual clone if it was sectioned through either side of the equator.

A comparison of the detected hepatocyte clones in the early stage, cirrhotic and HCC cohorts is shown in Figure 5.16. Significantly larger clones ($p > 0.001$, Student's t-test on square root transformed values) were observed in the cirrhotic and non-cirrhotic HCC cohorts than the early-stage cohort. Contrary to the results from invPCR of DNA taken from 5 mg liver tissue fragments, there was no significant difference in clone sizes between cirrhotic and non-cirrhotic patients. Thus, there is an observable association between disease progression and clonal proliferation.

In the early-stage patient cohort, repeated virus-cell DNA junctions were observed in 2 out of 7 patients (Table 5.8). After MPN analysis, these virus-cell DNA junctions were found to represent small and numerous hepatocyte clones, ranging from 8-1124 cells.

In the cirrhotic patient cohort, repeated virus-cell DNA junctions were observed in 3 of the 8 patients (Table 5.8). The clones detected in patient XA were especially large; one of which, XA v, contained a virus-cell DNA junctions that was identical to that detected in the 5 mg tissue fragments.

In the non-cirrhotic HCC patient cohort, repeated virus-cell DNA junctions were detected by invPCR in 5 out of 6 patients (Table 5.8). This detection rate is similar to previous studies. Clonal proliferation in non-cirrhotic HBV-infected humans and chimpanzees was detected in most tissues analysed (Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). Mason *et al.* (2009) showed comparable results to the current findings, detecting 50 unique clones in a total of ~35 mg of liver tissue from 5 chronically-infected chimpanzees without cirrhosis.

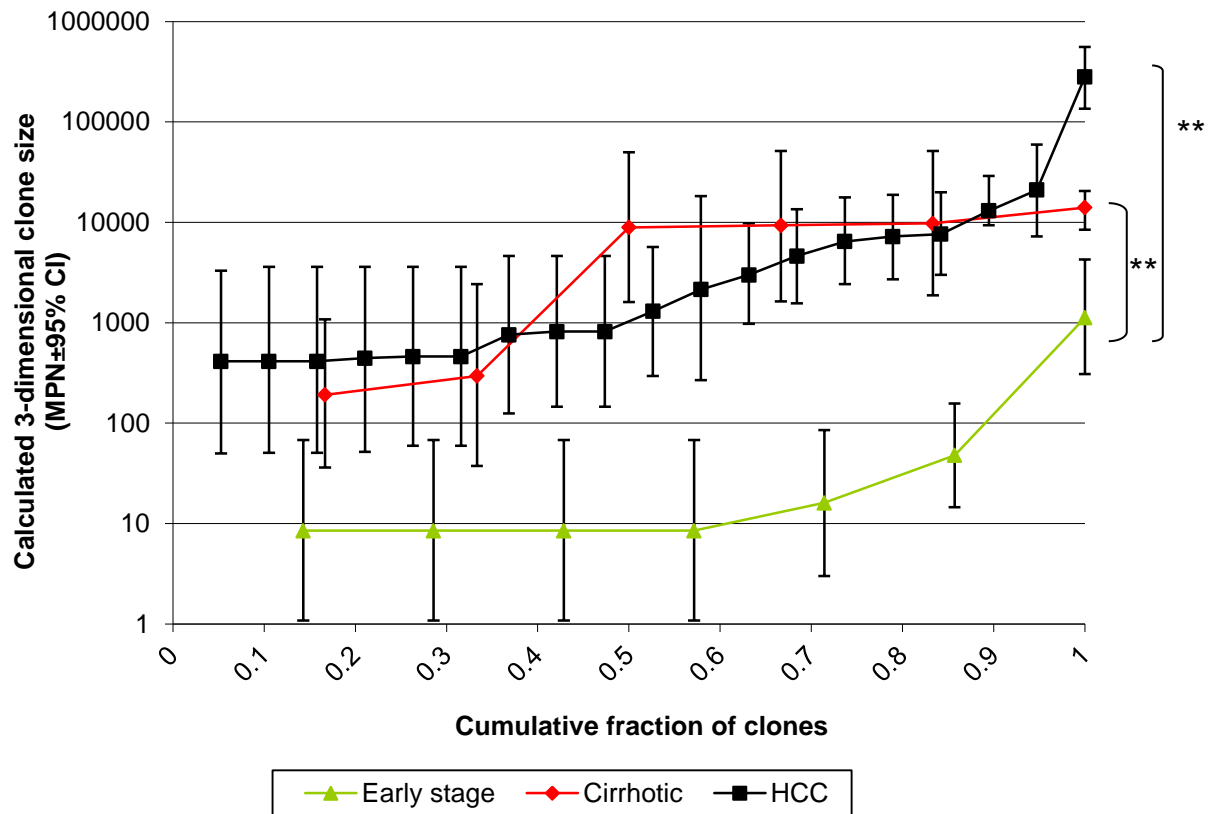


Figure 5.16. Clonal distribution of hepatocyte clones found in DNA extracted from 5 μ m liver tissue sections of early-stage, cirrhotic and non-cirrhotic HCC chronic HBV patients. Virus-cell DNA junctions were detected by invPCR as described in Section 2.14 and the clone sizes represented by the detected virus-cell DNA junctions were estimated as described in Section 2.22. Only clone sizes calculated using virus-cell junctions detected multiple times are represented in this graph. Clones sizes were significantly smaller in early stage patients compared to cirrhotic and non-cirrhotic HCC patients (** $p < 0.001$, Student's t-test on square root transformed values). No significant differences in hepatocyte clones sizes were observed between cirrhotic patients and non-cirrhotic HCC patient extracts. The clones that are displayed in the graph are summarised by patient extract in Table 5.9 and fully listed in Table 5.10.

Though similar numbers of slide-mounted tissue sections were analysed by invPCR from both cirrhotic and non-cirrhotic tissues, >5 times more unique virus-cell DNA junctions were detected in the non-cirrhotic tissues. The reason behind this difference in virus-cell DNA junction detection between non-cirrhotic and cirrhotic liver tissues is not known. Potentially changes in hepatocyte density due to the increased fibrotic tissue and hepatocyte hyperplasia may lead to fewer hepatocytes from cirrhotic tissues being analysed by invPCR. Further investigation is required to determine the mechanism for the paucity of observed virus-cell DNA junctions in cirrhotic tissues.

5.3.6 Predominantly unique subsets of virus-cell DNA junctions were observed using the *DpnII* and the *NcoI* invPCR designs

In tissues from patients CN, FMC3, GS4, HN, HS, NT, SAAO, XA, Y2, Y4 and Y6, the HBV DNA sequence allowed the use of both the *NcoI* and *DpnII* invPCR designs. The *NcoI* and *DpnII* invPCR designs detected similar numbers of unique virus-cell DNA junctions in same tissues; 31 and 32 virus-cell DNA junctions, respectively. There was little overlap in the previously described virus-cell DNA junctions detected by each design; only two clones (FMC3 ii and XA v) were detected by both designs. Thus, completely different sets of clones were being detected in the same DNA extract by the *NcoI* and *DpnII* invPCR designs, suggesting that the majority of HBV DNA integration events are not detected by invPCR.

5.3.7 Observed integration sites of HBV DNA are clustered around the expected right hand end of dsDNA form

The ends of the observed integrated HBV dsDNA were mapped with respect to the distance from the single *EcoRI* site of the HBV genome in Figure 5.17. Almost all of the observed (both unique and repeated) virus-cell DNA junctions occurred upstream of the expected right hand end of the HBV dsDNA form at nt 1832 (Staprans, Loeb et al. 1991; Yang and Summers 1999). Greater than 70% of all integrations occurred within 50 bp of nt 1832. The clustering was seen with both *DpnII* and *NcoI* invPCR designs.

Similar distributions were seen in previous studies using these invPCR designs on liver DNA from HBV-infected humans and chimpanzees (Mason, Low et al. 2009; Mason, Liu et al. 2010). Similarly, in pilot studies (Section 3.2.1.3) and previous studies of WHV and DHBV DNA integration using a different invPCR designs, >70% of all virus-cell DNA junctions

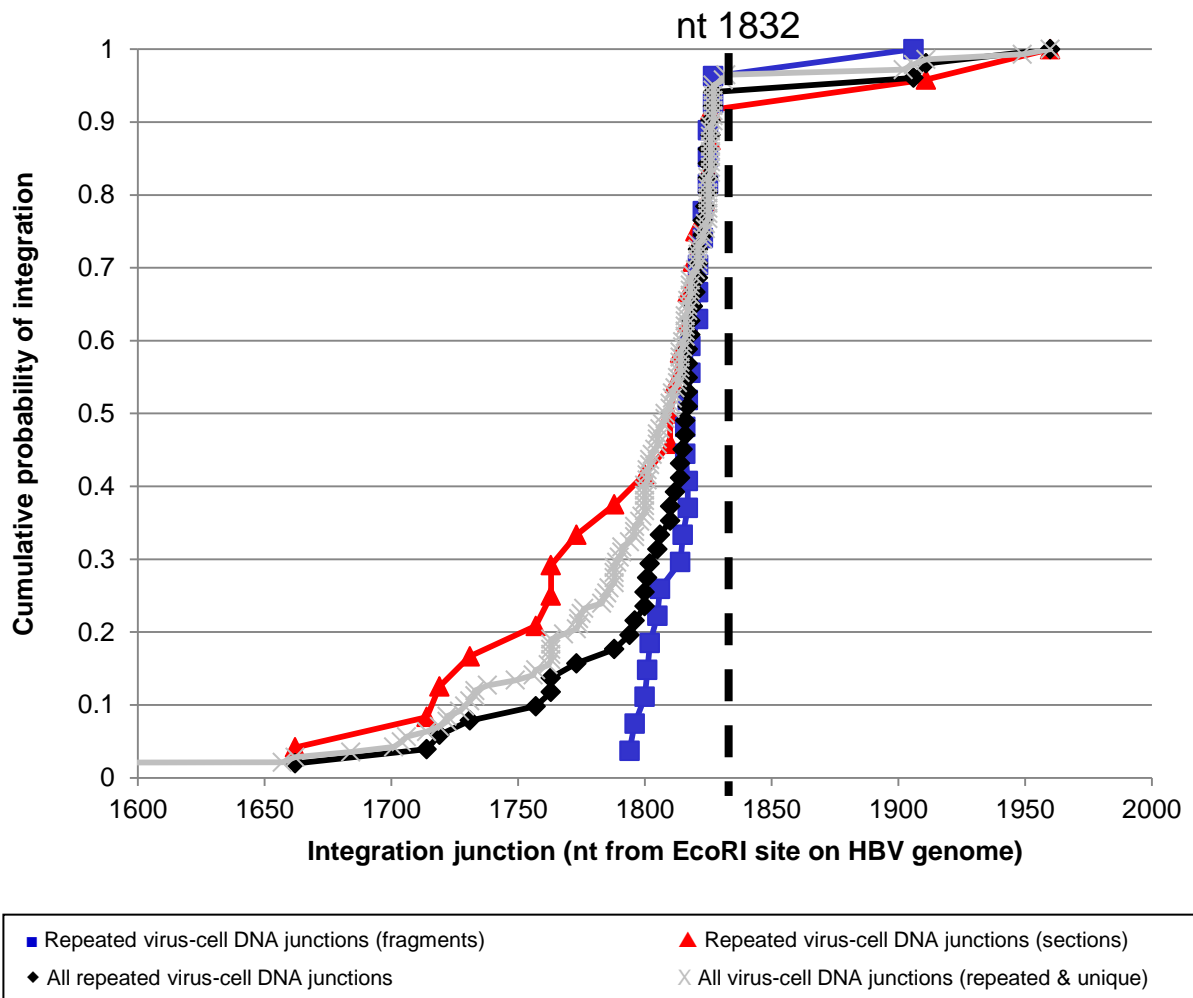


Figure 5.17. Integration sites of HBV DNA into the host cell chromosome matches the expected right hand end of the dsDNA form. Virus-cell DNA junctions detected multiple times by invPCR of DNA extracted from liver tissue fragments and liver tissue sections were aligned with the HBV genome (Genbank accession #AB241115). The cumulative probability of integration was plotted against the site of integration with respect to the HBV DNA sequence. In all cases, the integration site rose and clustered before the expected right hand end of the double-stranded linear DNA form of HBV located at nt 1832. This pattern held when virus-cell DNA junctions that had been only detected once were also analysed (grey crosses).

were detected <100 bp of the expected left-hand and right-hand ends of the viral DNA in liver DNA extracts from chronically infected woodchucks and ducks respectively (Yang and Summers 1999; Summers and Mason 2004; Mason, Low et al. 2009). This supports the hypothesis that integration occurs through NHEJ of the dsDNA form of HBV into the host cell chromosome (Bill and Summers 2004).

5.3.8 Observed integration sites are random within the cellular DNA with respect to chromosomes and genes and do not affect clone size

Virus-cell DNA junctions and the cellular genes in which they were integrated are listed in Appendix 9.8. While the majority (70%) of integration events occurred in intergenic sequences, 35 out of 120 integration events were observed in cellular ORFs. As ~75% of the human genome is composed of intergenic regions (Venter, Adams et al. 2001), the distribution of virus-cell DNA junctions suggests that the integration of HBV DNA into the cellular DNA does not affect the cell phenotype.

Integration events were dispersed throughout the entire cellular genome as shown in Figure 5.18. The mean frequency of observed integrations was 0.0354 ± 0.017 integrations per million bases (normalised by the length of each chromosome (NCBI 2011)). Preferences for chromosomes were tested using Grubb's test for outliers (aka. maximum normed residual test) using GraphPad online calculator (freely available at <http://www.graphpad.com/quickcalcs/Grubbs1.cfm>). No significant outlying chromosome was detected ($p > 0.05$). There was also no obvious correlation between the chromosome into which integration occurred and clone size (Figure 5.19).

In cases where the integration occurred in or <50 kb from an ORF, the levels of mRNA expression of each gene reported in the liver and in liver tumours was determined by consulting the online transcriptome database Unigene, located at <http://www.ncbi.nlm.nih.gov/unigene>. Unigene gives an estimated expression of a particular gene in a tissue by returning a value in transcripts per million. Expression of genes that were located adjacent to the integration sites were either not different between liver and liver tumour tissue cDNA libraries or were not observed at all in liver cDNA libraries.

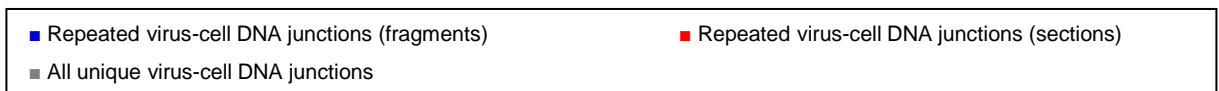
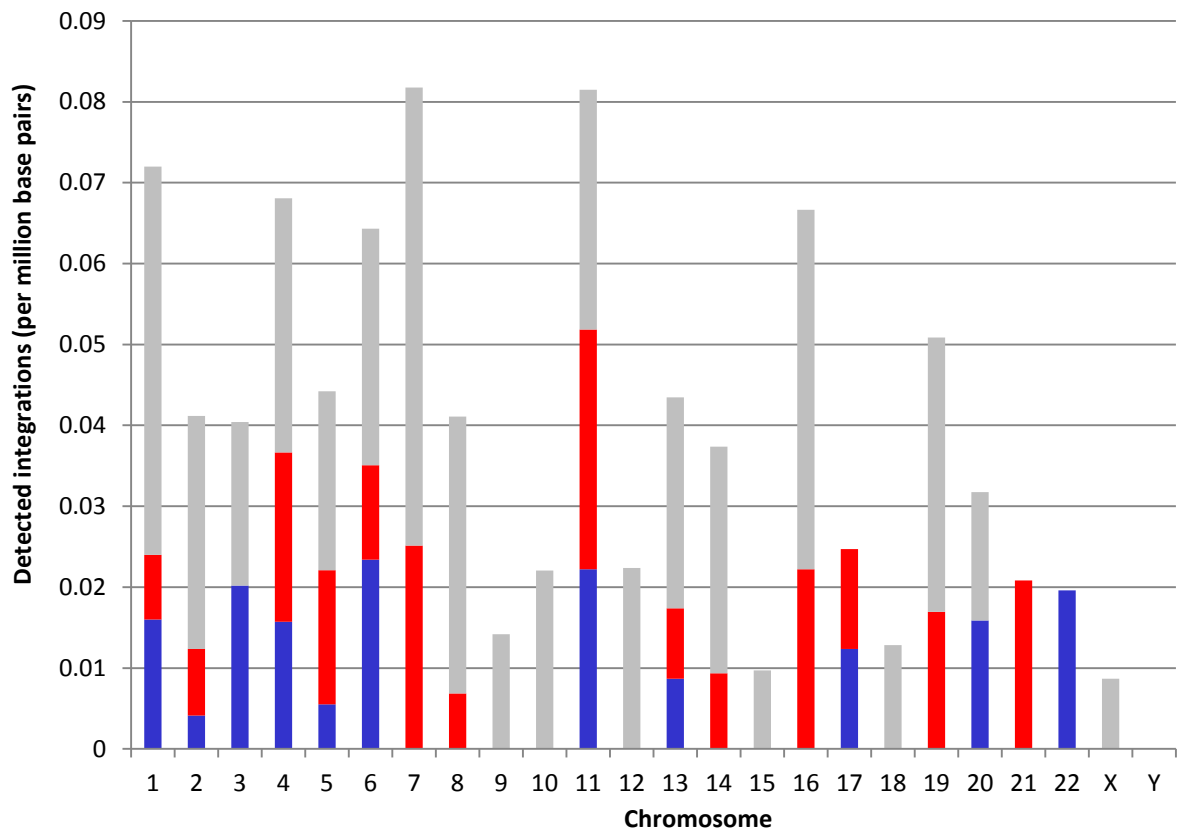


Figure 5.18. Host cell chromosomes are equally likely to contain detectable HBV DNA integrations. Repeated virus-cell DNA junctions detected by invPCR of DNA extracted from liver tissue fragments (blue) and liver tissue sections (red) were aligned with the human genome using BLAST analysis. The frequency of detectable integrations into each chromosome was normalised to the length of each chromosome and analysed by Grubb's outlier test. Assuming a normal distribution, the frequency of integration in all chromosomes did not significantly differ from each other ($p > 0.05$). These relationships held when virus-cell DNA junctions that had been only detected once were also analysed (grey).

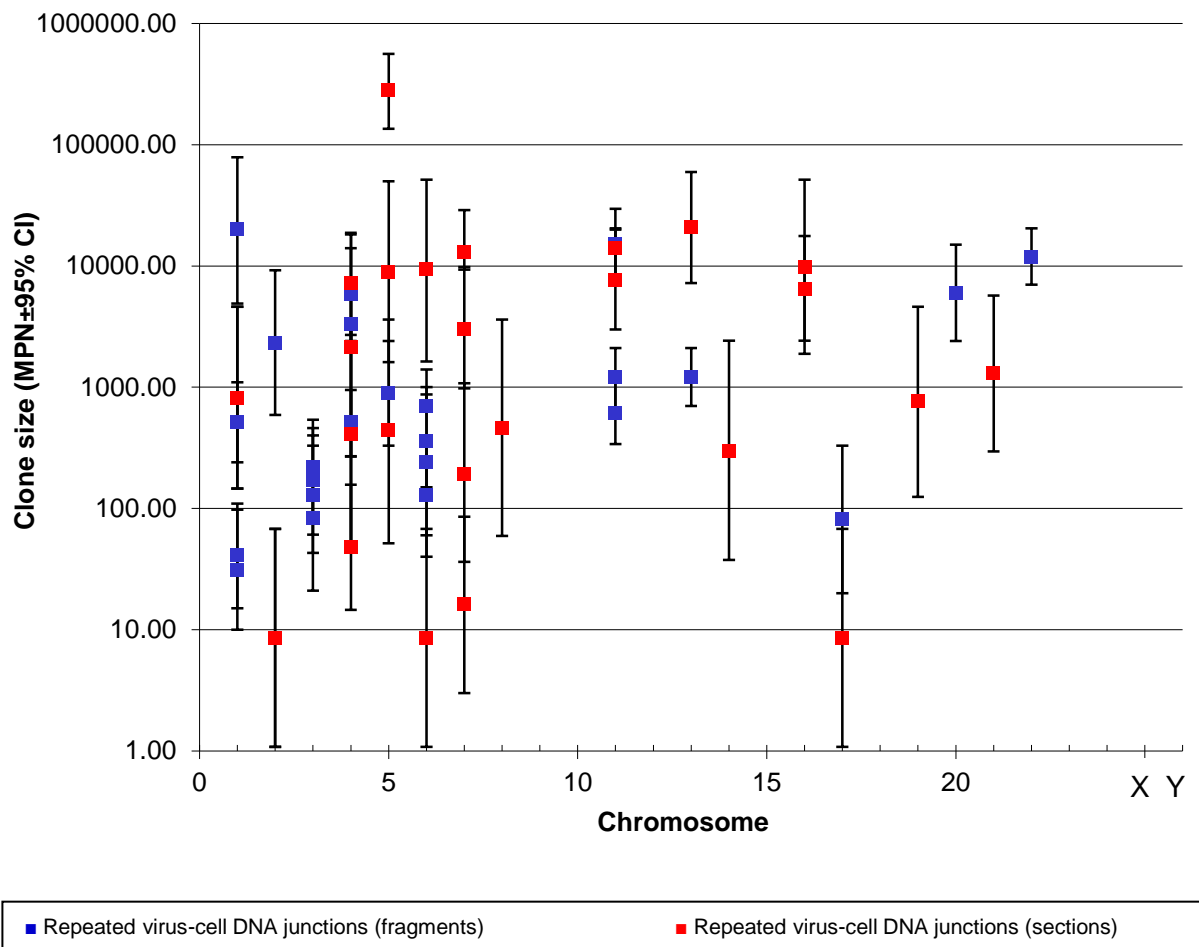


Figure 5.19. Integration of HBV DNA into particular chromosomes are not obviously associated with calculated clone size in 5mg fragments (blue) or slide-mounted tissue (red). Virus-cell DNA junctions were aligned to the human genome using BLAST analysis to determine the chromosome into which DNA integration had occurred. The integration sites were graphed against: either 1) clone sizes calculated using MPN statistical analysis as mentioned in Section 2.14 (5 mg fragments, blue); or 2) clone sizes in three dimensions as determined by MPN statistical analysis as mentioned in Section 2.14, in conjunction with 3-dimensional size estimation as mentioned in Section 2.22 (tissue sections, red). No association between clone size and distance from S/MAR was observed. Only clones represented by repeated virus-cell DNA junctions are shown.

5.3.9 Observed integration sites (but not clone sizes) are weakly associated with structural/matrix attachment regions (S/MAR)

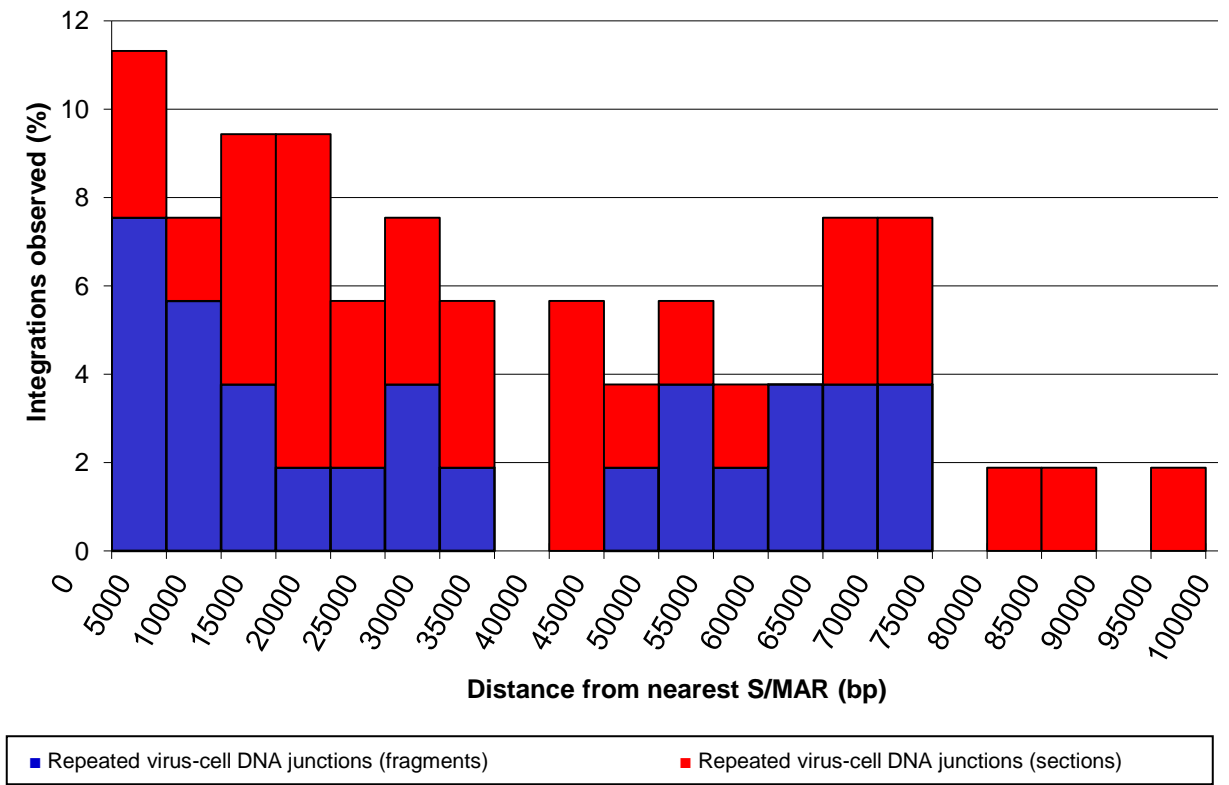
As described in Chapter 1, S/MAR are regions in cellular chromatin that are reported to modulate the activity of enhancers *in cis* and to interact with DNA-binding, transcription-associated proteins, such as DNA topoisomerase II (Boulikas 1993; Bode, Stengert-Iber et al. 1996). While there is no consensus sequence for S/MAR, there are conserved motifs, such as AT-rich regions, topoisomerase II sites and unwinding elements, which can be predicted as being likely to act as S/MAR by *in silico* analysis (Singh, Kramer et al. 1997; Frisch, Frech et al. 2002). S/MAR occur more frequently in non-coding regions and occur on average every 50-200 kb in the human genome (Pienta, Getzenberg et al. 1991; Bode, Stengert-Iber et al. 1996). HBV DNA integration in S/MAR may alter protein expression through disruption of enhancer elements in the cellular DNA.

All observed integration events were analysed for their proximity to S/MAR using MAR-Wiz (<http://www.futuresoft.org/MarFinder/>), the latest iteration of the MAR-Finder algorithm developed by Singh *et al.* (Singh, Kramer et al. 1997). Cellular sequences predicted by BLAST analysis 100000 bp upstream and downstream of the virus-cell DNA junction were analysed by MAR-Wiz.

Generally, the further away from S/MAR, the fewer virus-cell DNA junctions were observed (Figure 5.20). When all (both repeated and those detected only once) virus-cell DNA junctions were analysed by Grubb's outlier test, a significantly greater proportion of virus-cell DNA junctions were found within 5000 bp of S/MAR ($p < 0.01$, Figure 5.20B). No integration events were found > 95000 bp away from a predicted S/MAR. This suggests that S/MAR may represent sites of preferred DNA integration.

The proximity of the integration event to the nearest S/MAR was plotted against the clone size determined by invPCR for each virus-cell DNA junction found in DNA extracts from 5 mg fragments of liver (Figure 5.21) and slide-mounted liver tissue sections (Figure 5.22). No strong correlation was found in either case. This suggests that HBV DNA integration does not induce clonal proliferation of hepatocytes through the insertional mutagenesis of S/MAR.

A



B

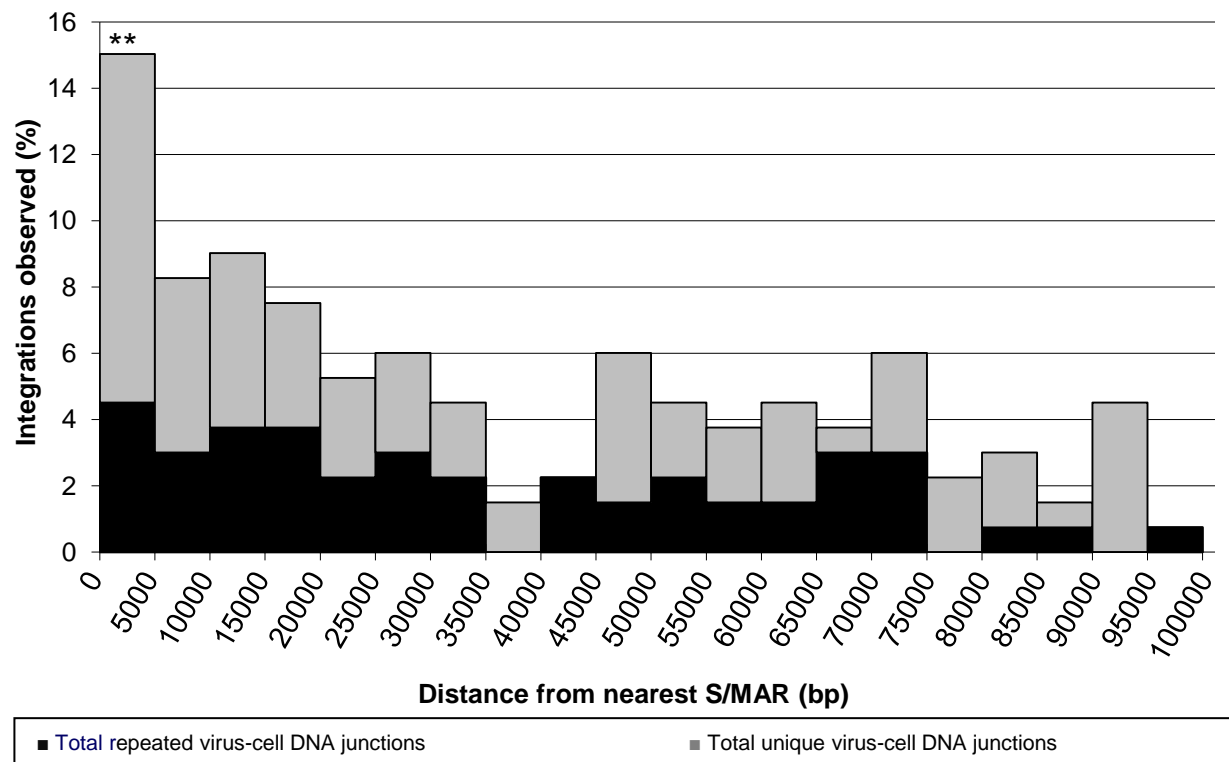


Figure 5.20. Integration sites are more likely to occur close to predicted S/MAR. Expected cellular sequences 100 kb either side of each virus-cell DNA junction detected was examined for S/MAR using freely-available, online program MAR-Wiz (<http://www.futuresoft.org/MarFinder/>). The minimum distance from a predicted S/MAR was recorded and graphed above. Analysis of repeated virus-cell DNA junctions (Figure 5.20A) detected in DNA extracted from liver tissue fragments (blue) and liver tissue sections (red) showed that a greater number of integrations were observed within 5000 bp of predicted S/MAR. This relationship became stronger when virus-cell DNA junctions that had been only detected once (grey) were also analysed (Figure 5.20B). In this case, a significantly greater number of integrations were observed within 5000 bp of predicted S/MAR ($p < 0.01$, using Grubb's outlier test).

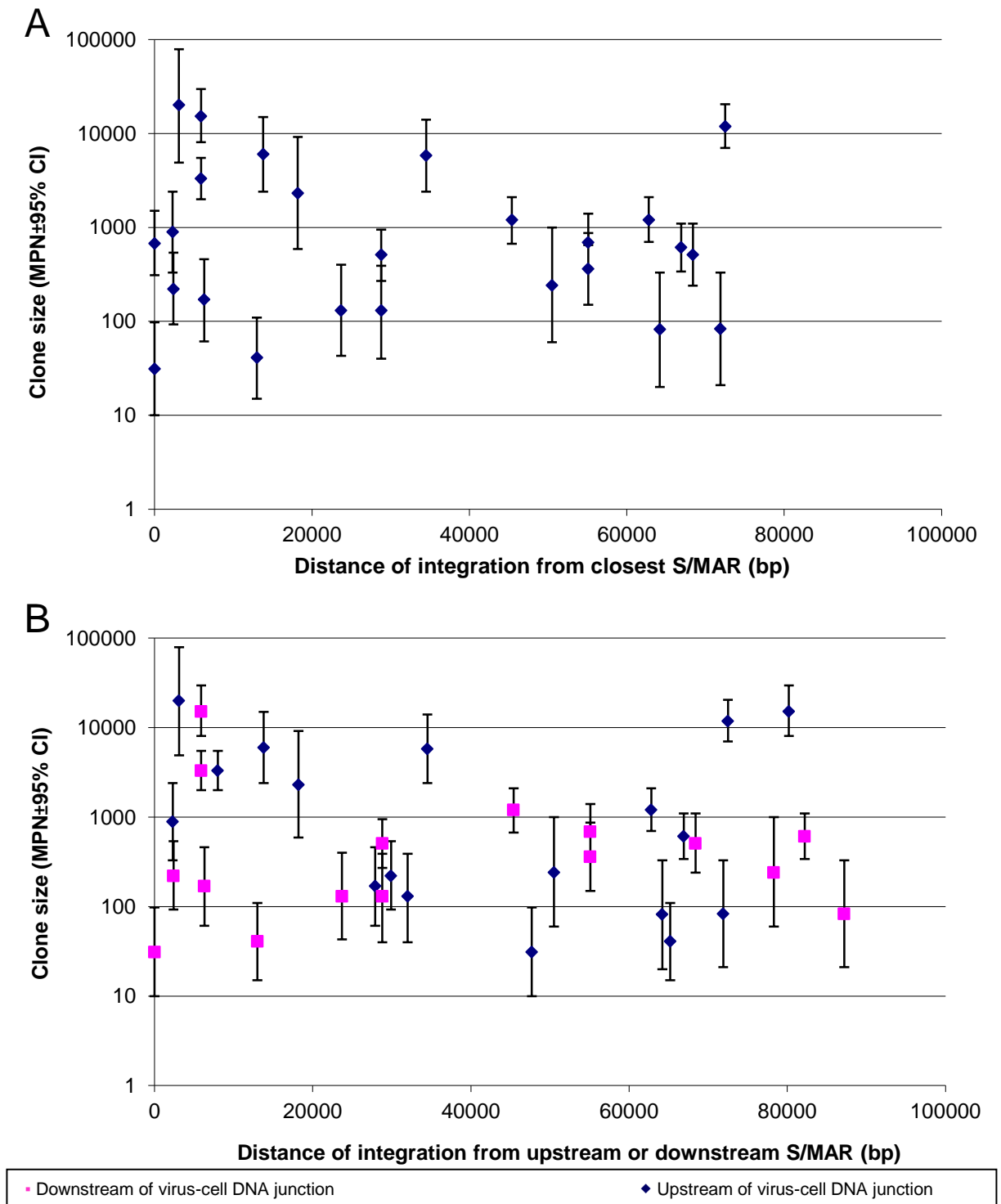


Figure 5.21. No association was found between clone size and proximity of virus-cell DNA junctions to S/MAR in integrations observed in ~5 mg tissue fragments. Clone sizes were determined by invPCR detection of virus-cell DNA junctions and the MPN statistical analysis as described in Section 2.14. Clone sizes were graphed against the distance from the closest predicted S/MAR (5.19A) or both upstream and downstream distances from S/MAR (5.19B). For the latter, distances greater than 100 kb were not included. No association between clone size and distance from S/MAR was observed. Only clones represented by repeated virus-cell DNA junctions are shown.

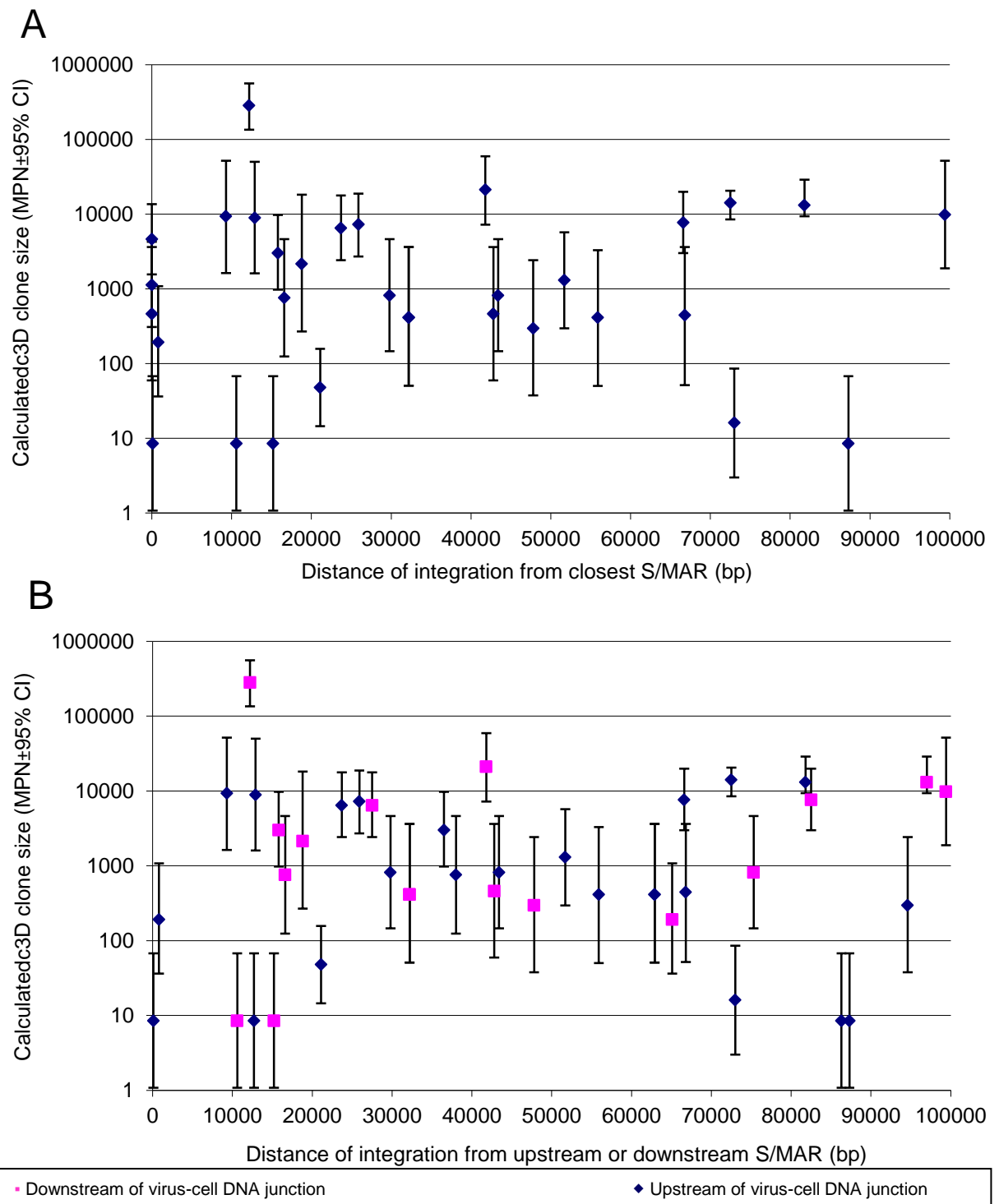


Figure 5.22. No association was found between clone size and proximity of virus-cell DNA junctions to S/MAR in integrations observed in slide-mounted tissue from cirrhotic and non-cirrhotic patients. Clones were detected in slide-mounted liver tissue. Each clone size in three dimensions was determined by invPCR, dilution of virus-cell DNA junctions and the MPN method as mentioned in Sections 2.14 and 2.22. Clone sizes were graphed against the distance from the closest predicted S/MAR (5.20A) or both upstream and downstream distances from S/MAR (5.20B). For the latter, distances greater than 100 kb were not included. No association between clone size and distance from S/MAR was observed. Only clones represented by repeated virus-cell DNA junctions are shown.

5.4 Discussion

This data strongly suggests that the invPCR assay is underestimating both the size and number of the hepatocyte clones present in the HBV-infected liver. During optimisation, the invPCR assay showed an efficiency of ~25% in a defined cellular region. If this efficiency can be extended to our experimental samples, then the hepatocyte clones are likely to be even larger than detected by invPCR. This is consistent with the observations from this study where invPCR clones were observed as 37% smaller on average compared to an independent qPCR assay on the same DNA extracts.

Even if HBV DNA integration had occurred in some hepatocytes, they may not necessarily be detected by invPCR. The virus-cell DNA junctions observed by *NcoI* and *DpnII* invPCR designs consisted of 2 completely different subsets. This strongly suggests that there are additional virus-cell DNA junctions that could not be detected by either invPCR design. Furthermore, our calculations showed that the *NcoI* and *DpnII* invPCR designs could both theoretically detect a maximum of only ~15-20% and ~70-75% all virus-cell DNA junctions, respectively. However, the invPCR designs detected a similar number of unique virus cell junctions in the same patients. Therefore, the data suggests that mechanisms other than RE site frequency were limiting the detection of virus-cell DNA junctions.

For example, the restriction site frequency calculations do not take into account that excised virus-cell DNA junctions of a large size may not circularise efficiently during the inversion reaction. All invPCR products observed were <2 kb in length. However, according to the RE site frequency calculations in Section 5.3.2, the average size of the expected products being 4 kb for the *NcoI* invPCR design. Furthermore, the RE site frequency analysis does not take into account loss of DNA during the inversion reaction. Also, while whole genome studies of HBV infected tissue have shown that the majority of HBV integration occurs in the region of the HBV DNA genome corresponding to the ends of the dsDNA form, some integrations were also detected across the remainder of the HBV DNA genome (Jiang, Yang et al. 2012; Jiang, Jhunhunwala et al. 2012; Sung, Zheng et al. 2012). Such integration events would not be detected by our invPCR design.

Thus, our detection frequency of virus-cell DNA junctions is likely to be <15%, the theoretical maximum of the *NcoI* invPCR design.

Also, the invPCR assay cannot observe all hepatocyte clones. To be detected, the hepatocyte must have a virus-cell DNA junction. However, HBV DNA integration does not occur in all cells. In the livers of ducks infected with DHBV, virus-cell DNA junctions were detected at a frequency of ~1 per 10000 hepatocytes after 6 days of DHBV infection (Yang and Summers 1999). The current data is consistent with this observed rare integration rate as the liver tissue DNA extracts from some patients were consistently negative for invPCR products, despite most patients being long-term HBV carriers with multiple opportunities for integration of HBV DNA and for clonal proliferation. It follows that invPCR does not detect all the hepatocyte clones present in the HBV-infected liver.

So, hepatocyte clones are expected to be much larger and more abundant than has been shown by invPCR. Even with these restrictions in hepatocyte clone detection, evidence for clonal proliferation was observed in the liver tissues of chronic HBV patients (Tables 5.5 and 5.9).

While significantly larger clone sizes were observed in the liver tissue fragments of cirrhotic patients compared to non-cirrhotic HCC patients, no such difference was seen between the two cohorts in clone sizes detected in liver tissue sections. This is probably due to the low sample number of large tissues fragments from HCC patients. If the clone sizes with the previously published results of invPCR analysis on DNA extracted tissue fragments of non-cirrhotic patients Y2 to Y6 are included (Mason *et al.* 2010), then the clone size difference between cirrhotic non-HCC and non-cirrhotic HCC cohorts becomes insignificant (Figure 5.23). This suggests that the hepatocytes of patients with late-stage disease have undergone similar amounts of clonal proliferation, despite the manifestation of the disease (either HCC or cirrhosis).

Interestingly, significantly larger hepatocyte clones were detected in liver tissue from HBV patients with late-stage disease compared to HBV patients in early-stage disease (Section 5.3.5). This suggests that disease progression is associated with clonal proliferation of hepatocytes. The causal relationships and mechanisms responsible for the difference in clonal proliferation however cannot be shown by the studies described in this Chapter. Further work (described in Chapter 7) has sought to determine the phenotype of the cells that have

Table 5.9. Full summary of invPCR results.

Histopathology type ¹		Early stage	Cirrhotic	Fulminant hepatitis	HCC
Number of patients		7	10	6	6
5mg liver fragments	# Patients /total tissue mass analysed ² (mg)	2/20	10/250	6/150	1/25
	# virus-cell DNA junctions detected (repeated)	11(3)	35(17)	10(3)	7(2)
	Clone size range (MPN (95% CI)) ³	10(1.5-73) to 670(310-1500)	40(5.5-290) to 15261(8102-30280)	40(5.5-290) to 20000(4900-79000)	40(5.5-290) to 610(340-1100)
Slide-mounted liver tissue	# Patients /Total tissue mass analysed ⁴ (mg)	7/~18	7/~30	2/~8	6/~18
	# virus-cell DNA junctions detected (repeated)	13(7)	15(6)	NA	46(19)
	Clone size range (MPN (95% CI))	3(1-60) to 1124(309-4273)	37(1.8-702) to 14045 (8478-20475)	NA	146(8-2998) to 281635(135137-559678)

¹ Histology was interpreted by A/Prof. Andrew Clouston (University of Queensland). HBV-positive patients could be grouped based on pathohistology into 4 groups: early-stage HBV infection (7 patients), cirrhotic HBV infection (10 patients), fulminant hepatitis (6 patients) and non-cirrhotic HBV infection with concurrent HCC (6 patients).

² Five 5 mg fragments from HBV patients were analysed by invPCR only where resection or needle biopsy tissue was available.

³ Clone sizes were determined by enumeration and detection of unique virus-cell DNA junctions using invPCR in conjunction with the MPN method, calculated by MPN calculator (Curiale, 2000).

⁴ Assuming ~2 mg of tissue per section.

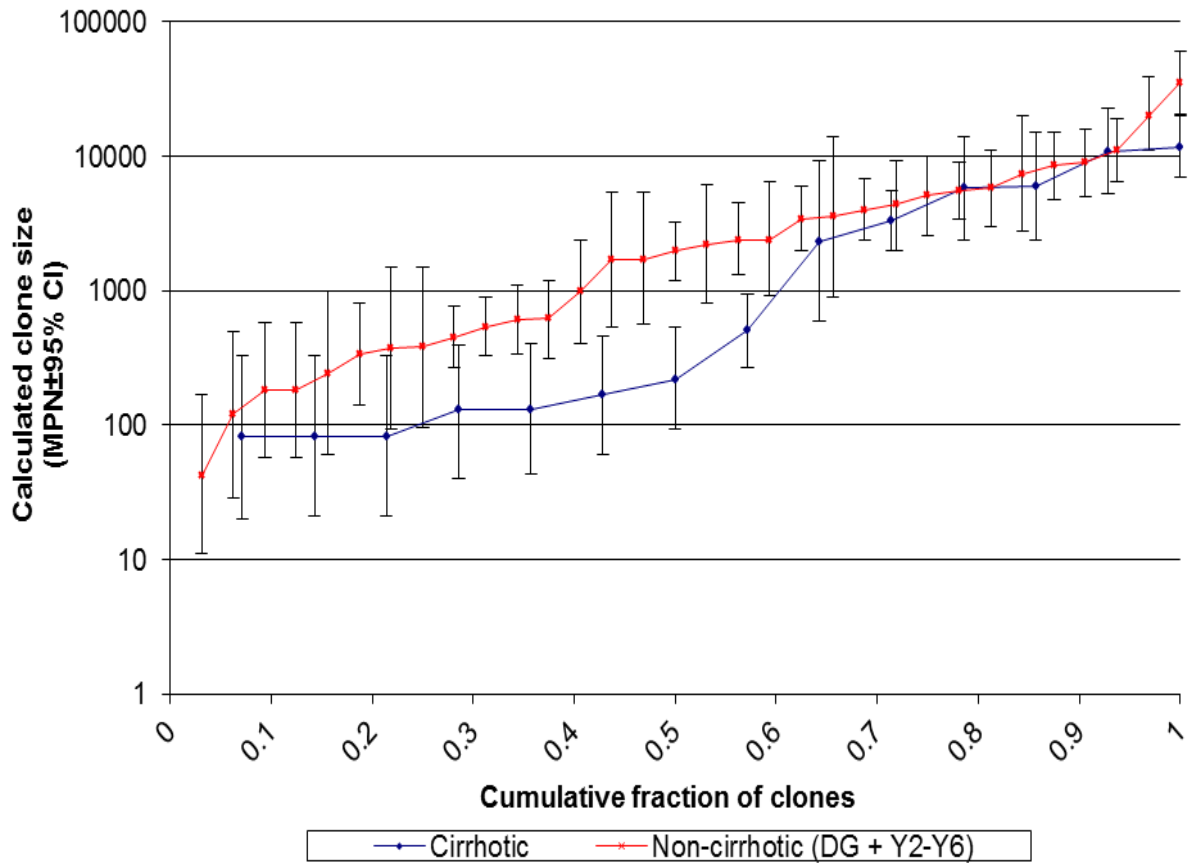


Figure 5.23. No significant difference in clone sizes observed in DNA extracts of liver tissue fragments of cirrhotic and non-cirrhotic HCC patients. Only clone sizes calculated using virus-cell junctions detected multiple times are represented in this graph. Data from invPCR of detected in tissue fragments of non-cirrhotic HCC Patients Y2 to Y6 from a previously published study (Mason *et al.* 2010) was used to calculate clone sizes by the MPN method. When added to the clones sizes detected in Patient DG of the current study, clone sizes observed in non-cirrhotic HCC patients (red) were not significantly different compared to cirrhotic patients (blue).

undergone clonal proliferation, thereby investigating the mechanism behind the observed clonal proliferation and its relationship to disease progression.

Also, clonal proliferation was still detected in early-stage patients (Sections 5.3.3 and 5.3.5), indicating that clones of hepatocytes can form before the long-term liver turnover associated with the continual flares of immune clearance of HBV-infected hepatocytes. Future studies of clonal proliferation liver tissue collected from patients in the immune tolerance phase of HBV infection may pinpoint the time at which large hepatocyte clones form.

There are numerous potential explanations for the apparent clonal proliferation of hepatocytes observed in the livers of HBV-infected patients in all cohorts.

5.4.1 PCR-amplified DNA artefacts

The observed virus-cell DNA junctions may simply be the presence of DNA artefacts by either non-specific ligation or by recombination during PCR (Judo, Wedel et al. 1998; Manivasakam and Schiestl 1998). These artefacts could be amplified by the invPCR assay and falsely interpreted as large hepatocyte clones. However, this hypothesis is not supported by our data. DNA extracts from some patients were persistently negative for invPCR products, suggesting that contamination artefacts are unlikely. Also, the same virus-cell DNA junction was never found in extracts from separate patients.

Also, the repeated virus-cell DNA junctions CY iv and XA v were observed in independent invPCR, PCR and qPCR assays. Since both invPCR and qPCR of uninverted DNA extracts detected the virus-cell DNA junctions, it is highly unlikely that the observed virus-cell DNA junctions are artefacts of the ligation reaction in the invPCR assay.

The possibly exists that some virus-cell DNA junctions were created by aberrant ligation of free dsDNA to random fragments of cellular DNA present in the DNA extracts. However, such ligation products would not be repeated more than once and, for this reason, clones size analysis was limited to those clones calculated from repeated virus-cell DNA junctions. While non-specific ligation or recombination during PCR may have occurred, it is unlikely that the same virus-cell DNA junction would be created by these random mechanisms.

Even so, virus-cell DNA junctions that were only detected once had similar distributions to repeated virus-cell DNA junctions with respect to the integration site in the integrated HBV DNA sequence, distance from S/MAR in the cellular sequence, and the chromosome into which HBV DNA had integrated (Sections 5.7-5.9). This suggests that the virus-cell DNA junctions detected once represent actual HBV DNA integration events.

5.4.2 *Changes in phenotype directly caused by HBV DNA integration*

Integration of HBV DNA into the host cell genome may potentially alter the phenotype of the hepatocyte directly, leading to the observed clonal proliferation by altering growth or survival pathways. The current data and analysis models show that this interpretation is unlikely.

Firstly, HBV DNA integration into gene regulatory structures is unlikely to have caused altered host cell phenotypes. Viral genome integration near and into S/MAR has been observed in virus-associated carcinomas, including both WHV-associated and HBV-associated HCC (Bruni, D'Ugo et al. 1999; Shera, Shera et al. 2001; Bruni, D'Ugo et al. 2003). However, the current study found no correlation between proximity to predicted S/MAR and clone size. This suggests that altering expression via insertion near S/MAR is not the cause of clonal proliferation.

S/MAR analysis suggests a site preference for HBV DNA integration. Shera *et al.* observed that the average distance from a S/MAR to a randomly chosen site in the host cell genome was ~40 kb (Shera, Shera et al. 2001). The current data shows a significant preference for virus-cell DNA junctions to occur <5kb from predicted S/MAR. This is consistent with previous studies (Pineau, Marchio et al. 1998). While reasons for this distribution are unknown, the nature of S/MAR as an unravelling site for DNA transcription and replication may be involved. Unpaired regions in the DNA genome, including S/MAR, are more susceptible to DNA breaks, the substrate for HBV DNA integration by NHEJ (Legault, Tremblay et al. 1997; Pineau, Marchio et al. 1998; Liu, Ribocco-Lutkiewicz et al. 2003; Bill and Summers 2004). Also, dysfunction in S/MAR has been proposed to induce illegitimate recombination in eukaryotes (Sperry, Blasquez et al. 1989). Whether the underlying mechanism for this observed integration site preference is due to different nucleotide

frequencies, and therefore RE site frequencies, is still unknown and requires further investigation.

If insertional mutagenesis was causing the clonal proliferation observed, then more virus-cell DNA junctions should be found in ORF coding sequences. However, most (~70%) of the detected virus-cell DNA junctions occurred in intergenic regions outside ORFs. This data is consistent with previous studies into the sites of HBV integration into the host cell genome (Nagaya, Nakamura et al. 1987; Mason, Low et al. 2009; Mason, Liu et al. 2010). Current bioinformatics analysis predicted that virus-cell DNA junctions there would not be large observational bias in intergenic, intronic or exonic regions (Table 5.4). The cellular integration sites are consistent with the overall composition of the human genome, in which 75% of sequences are intergenic (Venter, Adams et al. 2001). This suggests that it is unlikely that the changes in cell phenotype are caused by insertional mutagenesis directly.

In cases where HBV DNA had integrated <50 kb from host cell genes, the online bioinformatics tools, Unigene and EST, were used to determine gene mRNA levels in the liver. According to these cDNA libraries, the mRNA encoding these genes either had low to no expression in liver tissue or were not altered between normal liver and cancerous liver tissue. While the online cDNA libraries do not contain transcriptomes of all liver cancer phenotypes, this data suggests the host cell would not be altered to a cancerous phenotype even if observed integrations were interfering with mRNA expression.

The data does not rule out the role of HBV DNA integration causing chromosomal deletions and chromosomal instability, thereby altering cell phenotype. Insertion of virus genomes into host cell chromosomes has been reported to cause large chromosomal deletions in cellular DNA and increase the probability of the loss or gain of chromosomes via chromosomal instability (Ohshima, Ohgami et al. 1998; Pett, Alazawi et al. 2004). There are multiple conflicting studies both rejecting (Yaginuma, Kobayashi et al. 1985; Takada, Gotoh et al. 1990; Kimbi, Kramvis et al. 2005) and supporting (Rogler, Sherman et al. 1985; Hino, Tabata et al. 1991; Pineau, Marchio et al. 1998; Brechot, Gozuacik et al. 2000) chromosomal deletion and rearrangement during HBV DNA integration. Furthermore, chromosomal deletion and rearrangement and instability have been suggested as a mechanism for HCC initiation and progression (Hino, Tabata et al. 1991; Levy, Renard et al. 2002; Feitelson and Lee 2007). In this study, cellular DNA could not be amplified upstream of the observed virus-cell DNA

junction XA v. This opens up the possibility of a >500 bp deletion or complete rearrangement of cellular sequences as a result of HBV DNA integration. Further experimentation on greater numbers of HBV DNA integration events is required to confirm or overturn this hypothesis.

Also, HBV DNA integration may produce chimeric proteins that alter the phenotype of the host cell. There have been numerous accounts of integration of HBV DNA in HCC producing new chimeric proteins that are expressed in HCC cells (Wang, Zindy et al. 1992; Saigo, Yoshida et al. 2008). However, there was no consistency of integration sites with respect to specific cellular gene, the distance from known ORFs, or chromosome. Thus, it is unlikely that a chimeric protein formed by a particular detected virus-cell DNA junction is responsible for a potential survival advantage associated with clonal proliferation. However, it is still unknown whether there are common undetected virus-cell DNA junctions that may code for chimeric proteins and altering hepatocyte phenotype.

5.4.3 Clonal proliferation due to stem cell activation and proliferation

Another potential mechanism to explain the detected virus-cell DNA junctions is that stem-cells containing detectable virus-cell DNA junctions become activated and undergo high proliferation. During high levels of HBV-associated cell death, stem cells are induced to divide to replace hepatocytes that have been killed by the immune response (Roskams, Desmet et al. 2007). These high levels of cell death are seen in cirrhosis, especially in regenerative nodules (Yasui, Hino et al. 1992; Crawford 2007; Roskams, Desmet et al. 2007). The proliferation of oval cells and newly differentiated hepatocytes both containing detectable virus-cell DNA junctions may cause multiple copies of unique virus-cell DNA junctions in cirrhotic patients.

However, evidence has been provided in other Sections showing that HBV probably does not infect the stem-cell population in the liver. Dual immunofluorescence staining suggested that oval cells were not susceptible to HBV infection (Section 4.3.5). Also, virus-cell DNA junctions were not detected by invPCR of DNA extracts from non-hepatocellular cells isolated by laser-microdissection (Section 7.3.1). Furthermore, only low levels of total HBV DNA were detected in this cellular population by qPCR, nor any virus-cell DNA junctions detected by invPCR (Section 7.3.1). Therefore, our data suggests that stem cells are unlikely

to contain virus-cell DNA junctions and thus are unlikely to be a cause of the observed clonal proliferation.

5.4.4 Increased liver turnover over the course of infection increases stochastic clonal proliferation or clonal proliferation of hepatocytes with survival advantage

Disease progression may increase the liver turnover, causing stochastic clonal proliferation. This is potentially the cause of the observed clonal proliferation. However, the observed clone sizes of >10000 cells require high levels of liver turnover to occur stochastically. Previously, >10000 cells clones were detected in woodchucks chronically-infected with WHV that could not be explained by stochastic clonal proliferation caused by the limited liver turnover in the woodchucks' lifetime (Mason, Jilbert et al. 2005).

Hepatocyte clones with a heritable survival advantage increase in size at a greater rate than hepatocyte clones without a survival advantage during stochastic liver turnover. According to computer simulations outlined in Section 6, the observed clone sizes of >10000 cells are highly unlikely to be produced stochastically.

The potential survival advantages that may be attained by large hepatocyte clones are almost infinite, due to the changing microenvironment of the HBV-infected liver and the associated selection pressures. However, the current bioinformatics studies show that these phenotypical changes are unlikely to be caused directly by integration of HBV DNA. The observed HBV DNA integration sites show no preference for particular genes or coding regions and only display a weak preference for S/MAR. Furthermore, virus-cell DNA junctions were detected in intergenic regions at a frequency consistent with the intergenic content of the human genome (~75%). Thus, all data suggest that virus-cell DNA junctions act as a neutral marker for hepatocyte lineages.

The ability to detect clones in slide-mounted liver tissue sections allows further analysis of the clonal proliferation of HBsAg-positive and -negative hepatocytes. In following Sections, laser-capture microdissection was used to isolate foci of HBsAg-positive and -negative hepatocytes. DNA extracts from these foci were analysed by invPCR to determine whether they are clonal. Phenotype analyses into protein expression and methylation of HBV DNA

gene promoters will be examined to determine potential mechanisms for clonal proliferation (Section 7).

In conclusion, multiple clones of >10000 cells have consistently been detected by both invPCR and qPCR. Clones have been detected in DNA extracts from ethanol-fixed, slide-mounted liver tissue sections and snap-frozen fragments of liver tissue. Together, this data suggests that clonal proliferation of hepatocytes with a survival advantage occurs within the chronically HBV-infected liver. Furthermore, we have shown that the sites of HBV integration have an undetectable effect on clonal proliferation.

Results presented in subsequent Chapters of this Thesis build on the results of the current study. In Chapter 6, computer simulations have been used to show that hepatocyte clones of >10000 cells are highly unlikely to be formed through random cell death and regeneration. Also, in Chapter 7, hepatocyte clones were localised in slide tissue and the histology, epigenetic status and protein expression of those hepatocytes were analysed to investigate mechanisms for clonal proliferation of hepatocytes.

6 - Computer simulation of hepatocyte proliferation during chronic HBV infection

6.1 Introduction

Chronic HBV infection is associated with cirrhosis and HCC. HBV-infected patients have a 100-fold increased risk of HCC compared to uninfected individuals (Beasley 1988). This late stage disease generally occurs decades after the initial HBV infection (Goodman and Terracciano 2007; Tan, Yeh et al. 2008). Therefore, characterisation of the changes that occur in the hepatocyte population between the initial infection and late-stage disease are of great importance.

In Chapter 5 of this thesis and in previous studies (Mason, Liu et al. 2010), hepatocyte clones of >10000 cells have been detected using invPCR in the livers of patients chronically infected with HBV. As explained in Section 5.4.4, it is unknown whether these hepatocyte clones were formed by clonal proliferation due to random liver turnover (stochastic turnover), or whether the clones were formed by hepatocytes with a survival or growth advantage. If clones contain cells with a survival or growth advantage, then the hepatocytes will undergo clonal proliferation and the clone will increase in size at a greater rate.

Mason and Litwin have previously developed a Monte Carlo-based mathematical model to predict the size of hepatocyte clones that may occur through stochastic turnover (Summers, Jilbert et al. 2003; Mason, Jilbert et al. 2005). Clones of >10000 cells were detected using invPCR in the liver of 2.4-year-old woodchucks infected at birth with WHV (Mason, Jilbert et al. 2005). Based on proliferating cell nuclear antigen (PCNA) staining indices, estimates of the total amount of hepatocyte turnover that had occurred in the woodchucks since birth ranged from 6-12 turnovers of the entire hepatocyte population. Mathematical modelling with the Monte Carlo model showed that this level of hepatocyte proliferation would not have been sufficient to produce such large hepatocyte clones through stochastic liver turnover alone.

However, the particular model developed by Mason and Litwin is limited as it does not simulate the geometry of the liver. The geometry of the liver may be important in the development of large clones if hepatocytes can only replace neighbouring cells that have died. For example, hepatocytes that die and are located in the centre of large clones can be replaced only by hepatocytes belonging to the same clone. Thus, large clones are less likely to shrink and greater clone sizes may be formed.

Also, by modelling the geometry of the liver, the effects of non-random cell death on clonal proliferation of hepatocytes may be determined. Non-random cell death is known to occur

during chronic HBV infection via infiltration of immune cells resulting in areas of interface hepatitis where cell death is concentrated at the edge of portal tracts (Popper, Paronetto et al. 1965; Theise, Bodenheimer Jnr. et al. 2007). Furthermore, HBV-associated liver cirrhosis may lead to focal cell death via the formation of regenerative nodules. Focal cell death increases the turnover of hepatocytes in a local area. Since greater levels of stochastic turnover is expected based on mathematical models to increase the size of hepatocyte clones, increasing cell death in a particular area may lead to the formation of large hepatocyte clones (Summers, Jilbert et al. 2003; Mason, Jilbert et al. 2005).

Finally, the Mason/Litwin model also does not simulate the growth of the liver during normal human development. In the liver of a newborn, the initial number of hepatocytes (and therefore clones) is lower than that of a fully-developed adult (Shepard, Shi et al. 1988; Heymsfield, Gallagher et al. 2007). Thus, large hepatocyte clones may be formed by the growth of the liver during development from a newborn infant to an adult.

Therefore, we aimed to determine whether the hepatocyte clones observed by invPCR in Chapter 5 could have been formed by random, stochastic liver turnover. To do this, a novel computer simulation of a liver undergoing stochastic liver turnover was developed. The results from this simulation will indicate whether it is likely that the observed hepatocyte clones were formed by stochastic turnover or by clonal proliferation of hepatocytes with a survival or growth advantage.

6.1.1 Assumptions

Several assumptions were made in the development of this model:

6.1.1.1 The number of hepatocytes in the liver increases by 10-fold during development from a neonate to an adult.

The average newborn weighs ~3.5 kg and grows into a full-grown adult weighing ~70 kg (Yoshizumi, Gondolesi et al. 2003). Additionally, the average liver weight by percentage of body weight decreases from ~5% at birth (~175 g) to ~2.5% in fully developed adults (~1.75 kg) (Shepard, Shi et al. 1988; Heymsfield, Gallagher et al. 2007). Therefore, the average liver is expected to increase 10-fold in weight during development. Liver growth during development occurs at the lobule level; both the size and number of lobules increase by division of hepatocytes, rather than through hyperplasia of cells (Roskams, Desmet et al.

2007). Hence, the number of hepatocytes in the liver increases 10-fold from a newborn to an adult.

6.1.1.2 The number of hepatocytes in the adult liver remains constant.

The growth and size of the liver is tightly regulated (Fausto 2000; Koniaris, McKillop et al. 2003). Hepatocytes in the normal liver have a very low proliferation index, in the order of 0.05%/day (Mancini, Marucci et al. 1994). Livers that have undergone injury, on the other hand, regenerate quickly to replace lost cells and restore hepatocyte number (Roskams, Desmet et al. 2007). Studies of the response of the liver to partial hepatectomy performed on animals and humans have shown that the remaining liver cells regenerate to restore function and mass of the liver (Zajicek, Arber et al. 1991; Nagy, Kiss et al. 1998; Sun, Chow et al. 1999; Michalopoulos 2010). This regeneration is dependent on host size, as a small liver transplanted into a large host proliferates to an optimal size (Van Thiel, Gavaler et al. 1987). Conversely, a large liver transplanted into a small host does not proliferate as it is already at a functional size for the host (Fausto and Riehle 2005). Accordingly, in the current model, the rate of hepatocyte division has been modelled to be equal to the rate of hepatocyte death, thus the number of hepatocytes in the adult liver remains constant.

6.1.1.3 All hepatocytes are infected early after the onset of HBV infection.

The percentage of hepatocytes supporting HBV antigen expression in chronic HBV patients late in infection can range from 50 to 15% (Gowans and Burrell 1985; Mason, Liu et al. 2010). In experimentally HBV-infected chimpanzees and WHV-infected woodchucks, the number of virus-infected hepatocytes almost all of the hepatocytes support HBV-replication early in HBV infection before decreasing to 50-15% (Mason, Cullen et al. 1998; Mason, Jilbert et al. 2005; Murray, Wieland et al. 2005; Xu, Yamamoto et al. 2007; Mason, Low et al. 2009). These studies indicate that initially all hepatocytes can be infected with HBV. Therefore, this model assumes that all hepatocytes may be targeted for immune-mediated cell-death.

6.1.1.4 Survival advantages are heritable.

It is currently unknown if a survival or growth advantage is heritable. This model assumes that a survival or growth advantage arises from genetic mutations or epigenetic changes, which are heritable. However, possible survival or growth advantages could be non-heritable

changes, such as paracrine activity from surrounding cells or polyphenic changes during hepatocyte development. The potential changes associated with survival advantage in the context of the HBV-infected liver are discussed in greater detail in Chapter 7.

6.1.2 Program description

The created program (source code is shown in Appendix 9.7.3) is outlined in Figure 6.1. First, a 3-dimensional cell array (LiverArray) of a user-defined size is created to represent a portion of the liver.

In the simulations to determine the effect of liver growth on clonal proliferation, a growth phase was invoked: only one tenth of the LiverArray was assigned a unique identifier (ID) that acts a marker of individual hepatocyte clones. This represents a neonatal liver that is estimated to be a tenth of the size of the final adult liver (Shepard, Shi et al. 1988; Heymsfield, Gallagher et al. 2007). During the phase of liver growth a cell that has an ID assigned to it (Cell A) is randomly chosen to replace a cell without an ID assigned to it (Cell B). The ID from Cell B is then assigned as the ID of Cell A, representing a mitotic event where the liver grows in size. This process is repeated until all of the cells in the LiverArray are assigned an ID.

In all other simulations, all cells in the LiverArray were assigned unique IDs just before the liver cell death phase of the program. At the start of the liver cell death phase, a number of randomly chosen cells in the LiverArray are given a user-defined survival advantage (SA). SA is defined as the probability that a cell survives when chosen randomly for cell death. Then, a random cell in the array (C_{die}) is nominated for cell death. The survival advantage of the cells was implemented by comparing the SA of C_{die} (SA_{die}) against a uniformly-distributed random number rnd , for $0 < rnd \leq 1$.

If $rnd < SA_{die}$, C_{die} survives cell death and so another random cell from LiverArray is nominated. Since the distribution of rnd is uniform, SA_{die} is directly proportional to probability that C_{die} will survive cell death. So, if $SA_{die} = 1$, C_{die} will always survive cell death; whereas if $SA_{die} = 0$, C_{die} will always succumb to cell death; and if $SA_{die} = 0.5$, C_{die} will succumb to cell death 50% of the time whenever the cell is nominated for cell death.

If $rnd > SA_{die}$, C_{die} succumbs to cell death and another cell (C_{live}) is chosen to replace C_{die} . C_{live} is randomly chosen from a pool of cells surrounding C_{die} . The size of this pool is user-

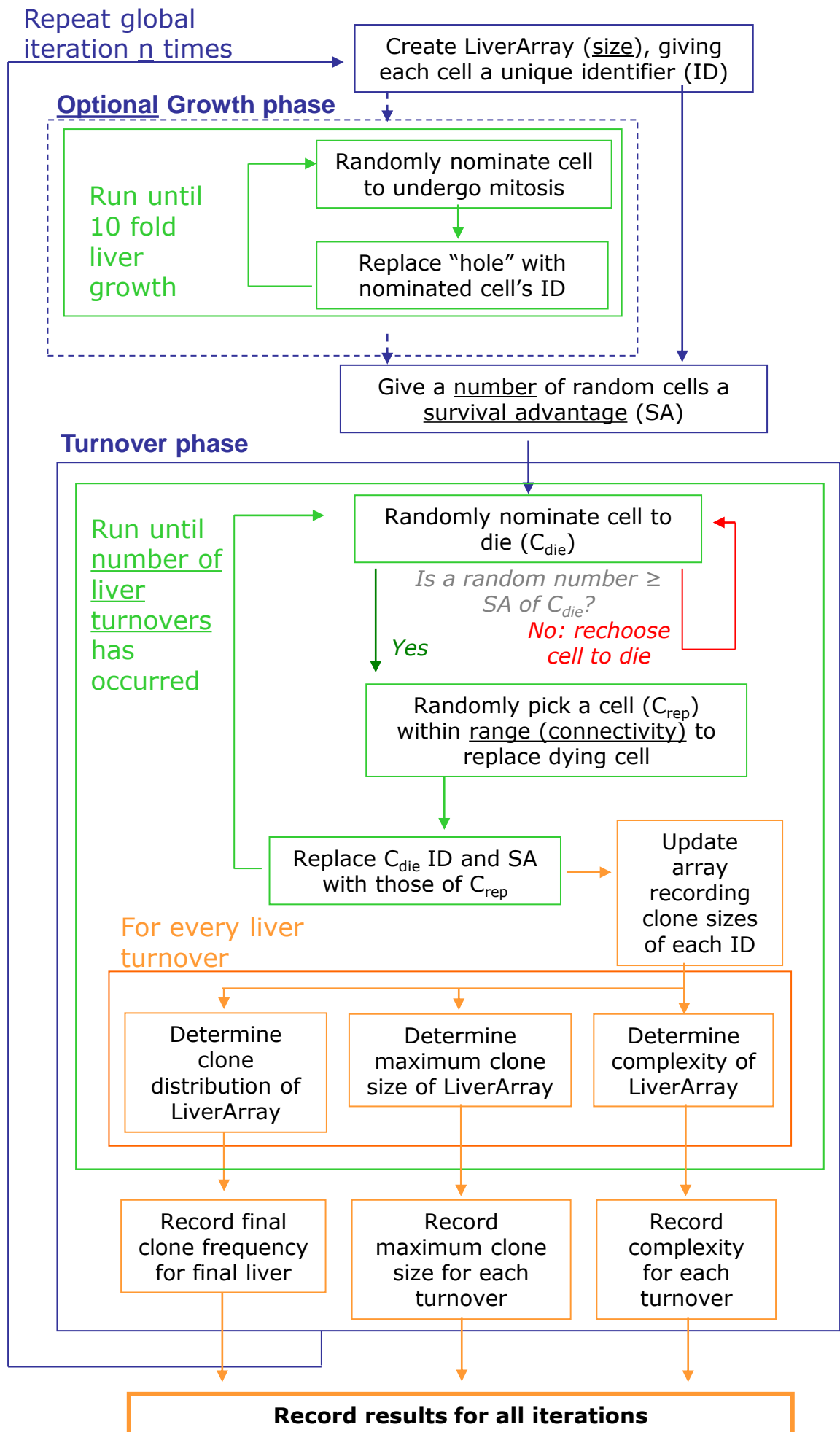


Figure 6.1 Flowchart outlining the created predictive computer simulation. Underlined words represent user-defined variables. A single cycle of the blue boxes represents a single global iteration of the program. Recording events are boxed in gold. Further description of the program is provided in Section 6.1.2.

defined by the connectivity value and further discussed in Section 6.1.3.2. SA_{die} and the ID of C_{die} are then replaced by SA_{live} and the ID of C_{live} .

This completes a single liver cell turnover. A turnover of a whole liver is defined as the number of cell turnovers that equals the user-defined number of cells in the LiverArray. The program randomly selects another C_{die} and repeats this process for a user-defined number of liver turnovers.

By measuring the number of copies of each cell ID, three outputs were recorded during the simulation: 1) the complexity of the LiverArray; 2) the maximum clone size in the LiverArray; and 3) the clone size frequency in the LiverArray.

Complexity is a measure of diversity within the liver used previously by Mason *et al.* in originally published models by Mason and Litwin (Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). Complexity is defined as the number of unique hepatocyte clones divided by the total number of cells in the LiverArray. It is therefore a measure of the fraction of initial cell lineages that remain in the LiverArray and is analogous to the measure of species richness index in the field of ecology (Magurran 2004). The complexity of the LiverArray was recorded after each turnover of the entire LiverArray.

The maximum hepatocyte clone size is defined as the largest hepatocyte clone observed in the entire simulated LiverArray. The maximum clone size was recorded after each turnover of the entire LiverArray.

The clone size frequency is the number of clones composed of 1 cell, 2 cells, 3 cells..., etc. The clone size frequencies of the LiverArray were only recorded at the end of a user-defined number of liver turnovers.

.

6.1.3 Parameters

The simulation was run with changes in several parameters.

6.1.3.1 LiverArray size

Due to limitations in available computing power, the $\sim 2 \times 10^{11}$ cells present in the average human liver (Shepard, Shi et al. 1988; Heymsfield, Gallagher et al. 2007; Roskams, Desmet et

al. 2007) were unable to be fully simulated. The effect of altering the size of the LiverArray from 50 cells cubed (equivalent to 1.25×10^5 cells) to 250 cells cubed (equivalent to 1.56×10^7 cells) on clone sizes was measured. These sizes represent the number of cells present in 1.5 mg to 200 mg of liver, respectively.

6.1.3.2 Connectivity

Connectivity determines the size of the pool of cells that surround C_{die} that can be selected as the cell that replaces C_{die} , and is therefore chosen as C_{live} . For connectivity = 1, the location of C_{live} is no more than 1 cell away from C_{die} on each axis, meaning that only neighbouring hepatocytes divide to replace lost hepatocytes.

Some studies have suggested that dying hepatocytes are replaced by neighbouring hepatocytes during chronic HBV infection, indicating a connectivity of 1. The distribution of PCNA, a marker of cellular proliferation, was analysed by immunohistochemistry in liver tissues from patients with acute and chronic viral hepatitis (Kawakita, Seki et al. 1992). The majority of the PCNA-positive hepatocytes were observed adjacent to areas of necrosis. PCNA-positive hepatocytes were also observed in periportal areas with interface hepatitis where mononuclear cells had congregated and presumably induced cell death in the HBV-infected hepatocytes. Furthermore, after experimental hepatectomy of $\frac{1}{4}$ of the liver in fully grown pigs, a greater number of hepatocytes adjacent to the surgical site were observed to have incorporated the thymidine-analogue bromodeoxyuridine (BrDU) than those distant from the surgical site (Sun, Chow et al. 1999).

On the other hand, other studies have shown that dying hepatocytes can also be replaced by hepatocytes distant from the site of cell death. Bralet *et al.* showed that after experimental hepatectomy of $\frac{2}{3}$ of the liver of adult rats, hepatocytes proliferation was seen throughout the entire remaining host liver (Bralet, Branchereau et al. 1994). The liver of the rats was perfused with a retroviral vector containing an *E. coli* β -galactosidase gene (*lacZ*) at 1 day post-resection. Only dividing cells were able to be transduced. Sections of liver were taken and analysed for β -galactosidase activity. The distribution of LacZ-positive cells was scattered throughout the liver lobules. This suggests that all hepatocytes are able to undergo increased proliferation during severe parenchymal damage. Therefore, there may be instances during chronic HBV infection when severe parenchymal damage occurs and connectivity will be >1 , with cells in more distant positions in the liver dividing to take the place of hepatocytes that have been killed.

As the true nature of hepatocyte replacement may be a mixture of both mechanisms, effects of changing connectivity from 1 (i.e. dying cells are replaced by neighbouring cells) to 250 (i.e. dying cells are replaced by any cell in the liver) on maximum clone size and clone size frequency were simulated.

6.1.3.3 Focal cell death

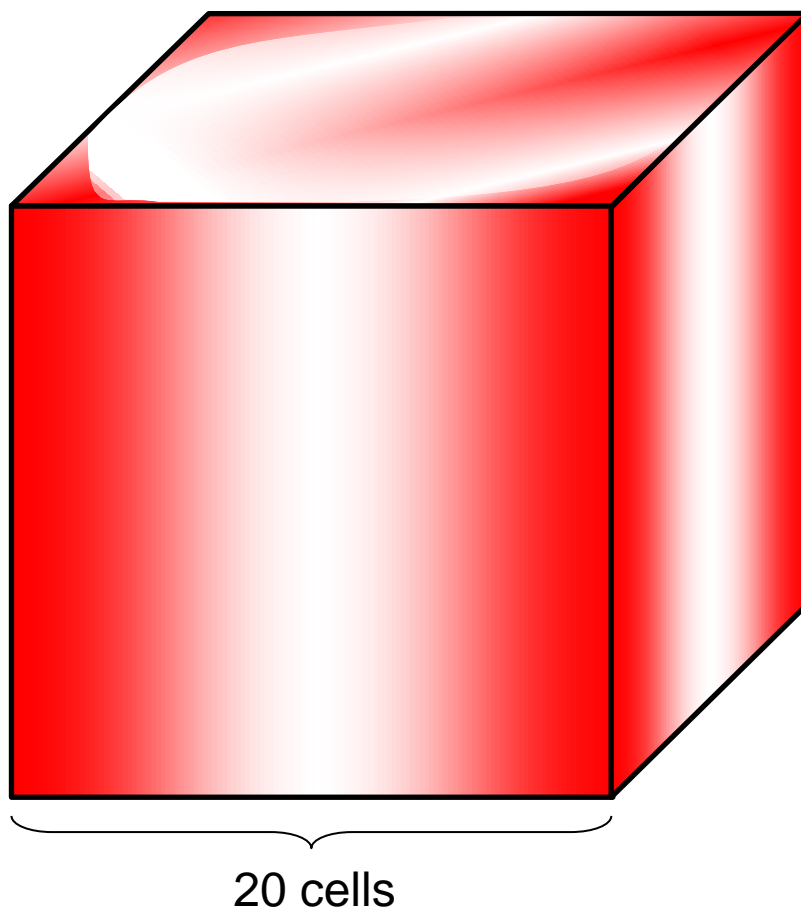
All hepatocytes in the liver can be killed by immune-mediated cell death during HBV infection. In liver sections taken from patients with chronic HBV infection, hepatocyte cell death can occur across all zones of the liver during so-called “spotty” or “confluent” necrosis (Liaw, Chu et al. 1982).

However, certain zones can be favoured over others for cell death over the period of chronic HBV infection. Hepatocytes near periportal regions of interface hepatitis may be targeted for cell death mediated by infiltrating lymphocytes. Such focal death has been referred to as “piecemeal necrosis” or “interface necrosis” (Popper, Paronetto et al. 1965; Theise, Bodenheimer Jnr. et al. 2007). Furthermore, following experimental 2/3 hepatectomy of the rat liver, regrowth occurred preferentially in periportal and mediolobular regions (Bralet, Branchereau et al. 1994). Furthermore, hepatotoxins, such as alcohol and paracetamol, can induce peri-portal cell-death (Powell, Kosyk et al. 2006; Dadarkar, Fonseca et al. 2010).

The effect of focal cell death was modelled by simulating liver lobules in the LiverArray and associating a probability of cell death dependent on the distance of the cell from the portal tracts. In the liver proper, the structure of lobules in 2-dimensional sections resembles a hexagon, with portal tracts at the vertices draining blood into the central vein in the centre, with an average radius of ~20 hepatocytes (Matsumoto and Kawakami 1982; Roskams, Desmet et al. 2007). This hexagonal structure was difficult to implement in the cube-based LiverArray. Two models were used to approximate the structure of the liver lobules, as shown in Figure 6.2.

In the first model, the lobular structure was simulated by splitting the LiverArray into cubes with sides of 20 cells, the average radius of a liver lobule, as shown in Figure 6.2A. The vertical edges of each cube represent the portal tracts.

A



B

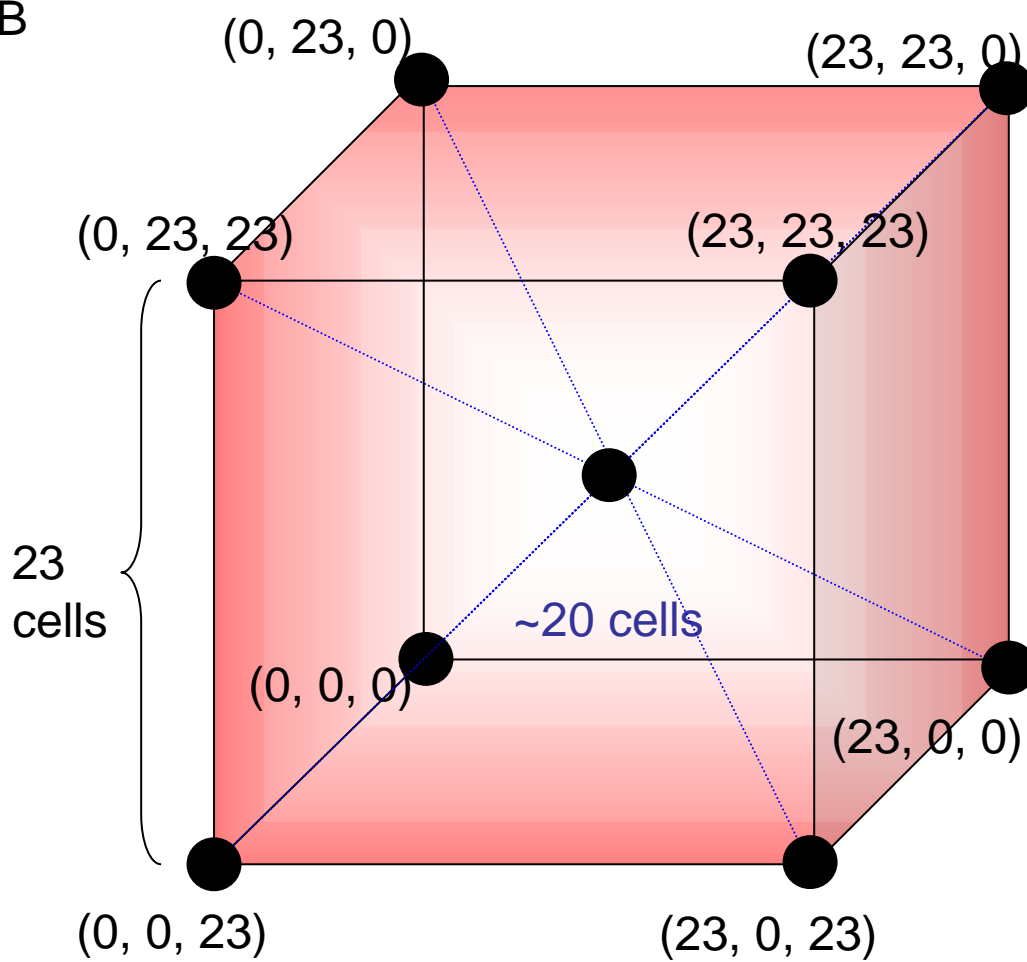


Figure 6.2 Two models of a simulated liver lobule. In the first model of the liver lobule, shown in Figure 6.2A, LiverArray was split up into cubes with sides of 20 cells, the approximate radius of a human liver lobule. The portal tracts were simulated to be running vertically along the vertical edges of the cubes. The cells closer to the vertices have a greater chance of being killed (red) than those further away (white).

In the second model, shown in Figure 6.2B, the LiverArray was split up into cubes of 23^3 cells. Portal tracts were simulated on the vertices of cubes. The cells closer to the vertices have a greater chance of being killed (red) than those further away (white). The distance from the vertices to the centre of the cube (blue dotted lines) is about 20 cell diameters.

In the second model, the lobular structure was approximated by splitting the LiverArray into cubes of 23^3 cells (2.2×10^4 cells), as shown in Figure 6.2B. The vertices of each cube represented the portal tracts. The distance from the vertices to the centre of the cube is ~ 20 cells.

The probability of survival for a given cell was programmed as directly proportional to either (i) the square of the distance or (ii) the distance from the closest simulated portal tract, as shown in Figure 6.3. These two options were simulated because the relationship between distance from the portal tract and cell death is unknown. For example, if indeed focal cell death is simply due to contact with infiltrating cells or to diffusion of a hepatotoxin from the portal tract to the central vein, then one would expect the cell death to follow the concentration gradient and be directly proportional to the square of the distance from the portal tract. However, since hepatocytes in the liver are known to heterogeneously express toxin metabolising enzymes (Lindros 1997), hepatocytes closer to the portal tract may be more resistant to the effects of hepatotoxins, thereby resulting in a more linear distribution of cell death across the lobule.

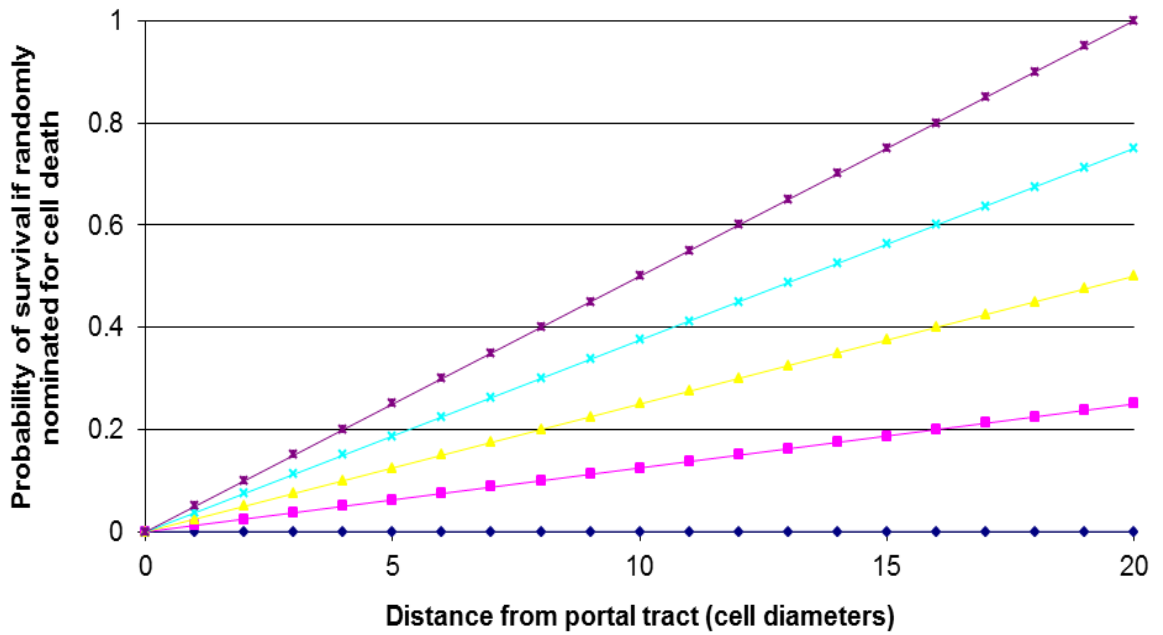
For each cell, the distance from the closest portal tract ($PTdist$) was calculated as a fraction of the distance from the centre of the cube. $PTdist$ or $PTdist^2$ of each cell as stored in the array. When nominated for cell death, $PTdist$ or $PTdist^2$ was compared to a uniformly distributed random number for $0 \leq rnd < 1$.

In the case that $PTdist$ or $PTdist^2 \leq rnd$, another cell is randomly chosen to die. Otherwise if $PTdist$ or $PTdist^2 > rnd$, the cell is killed and replaced by a neighbouring cell. The ratio of random cell death due to natural hepatocyte turnover and distance-dependent cell-death was implemented multiplying $PTdist$ or $PTdist^2$ by a fractional constant.

6.1.3.4 Hepatocytes with a survival or growth advantage

It is unknown what factors lead to a survival or growth advantage for individual cells in the hepatocyte population. The extent of a survival or growth advantage is also unknown. The range of potential survival advantages, including evasion of the immune response, cessation to HBV antigen expression and hypersensitivity to growth signals, are discussed further in Section 7.4. The simulation used in this Chapter was run with an arbitrarily small number of cells (1 out of every 100000 cells) with survival advantages ranging from 2 - 50% (i.e. they are killed 1.02 to 1.5 times less often).

A – Cell survival directly proportional to distance from PT



B – Cell survival directly proportional to the square of the distance from PT

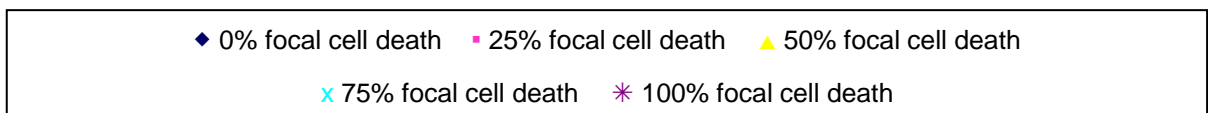
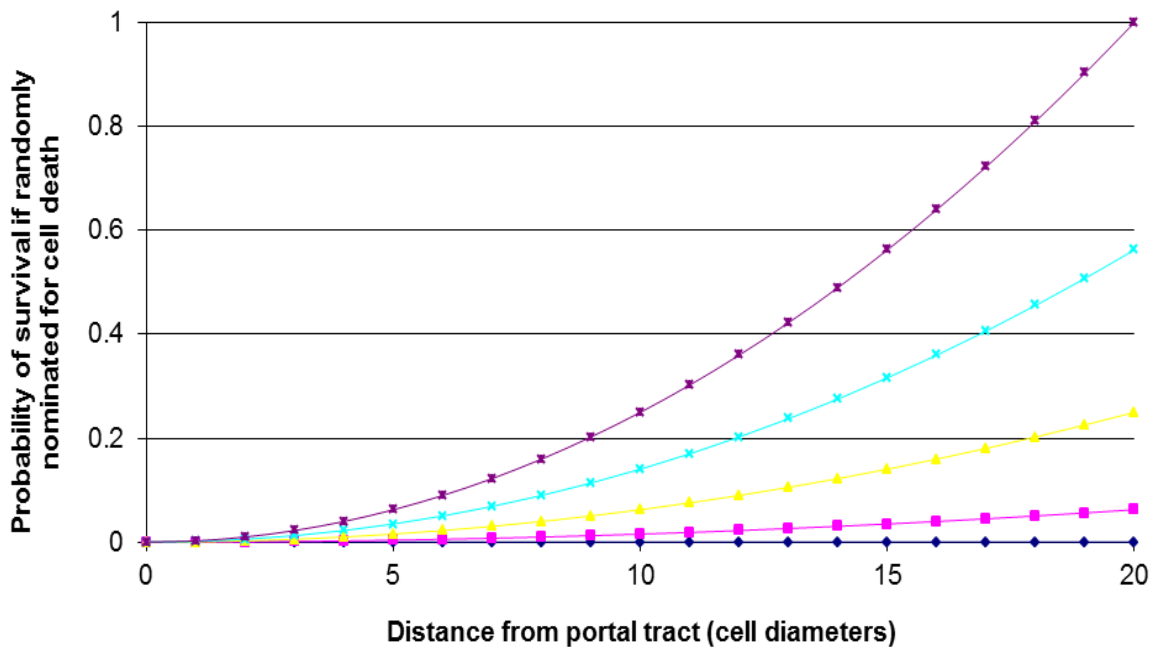


Figure 6.3 Programmed focal cell death increases the chance of survival as cells are located further from portal tracts. In the simulation of focal cell death, survival of a cell was programmed to be directly proportional to either the distance of the cell from a simulated portal tract (6.3A) or the square of the distance of the cell from a simulated portal tract (6.3B).

6.1.3.5 Stem cell populations

As mentioned in Section 1.8.3, the susceptibility of oval cells, the bipotent stem cell population of the liver, to HBV infection is controversial (Evarts, Nagy et al. 1987; Hsia, Thorgeirsson et al. 1994; Lowes, Brennan et al. 1999).

Some evidence provided in other Sections in this Thesis support the hypothesis that stem cell populations do not support HBV infection and replication. Firstly, only a small fraction of cells were positive for HBsAg and the oval cell marker cytokeratin 19 by immunofluorescence (Section 4.3.5). Those that were dual-stained did not match the general phenotypic features of oval cells that include: scant oval-shaped cytoplasm, round nucleus, and a slightly smaller size than hepatocytes (Evarts, Nagy et al. 1987; Lowes, Brennan et al. 1999). This suggests that the HBsAg-positive cells were transitional hepatocytes and that the stem-cell population was not infected. Also, non-hepatocytes were laser-dissected and assayed by a HBV-specific PCR and invPCR (Section 7.3.1). These assays detected very low HBV DNA levels (conceivably from contaminating HBV particles in the central veins) and no detectable virus-cell DNA junctions. This supports the hypothesis that stem cells appear not to support HBV infection and therefore are invisible to the invPCR assay used to detect hepatocyte clones.

However, the possibility that a stem cell population can be infected with HBV and can contain detectable virus-cell DNA junctions was included in the simulation. To implement this in LiverArray, randomly distributed cells were given a survival advantage of 1. These immortal cells do not pass on their survival advantage. This models an oval cell population producing hepatocytes that are subject to immune-mediated cell death and natural liver turnover.

6.2 Results

6.2.1 The current model replicates the results of a previous stochastic simulation

Dr. William Mason kindly allowed access to the program previously used by Mason *et al.* to analyse clonal proliferation of hepatocytes in chronically hepadnavirus-infected humans, woodchucks and chimpanzees (Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). As shown in Figure 6.4, the Mason/Litwin model and the current model

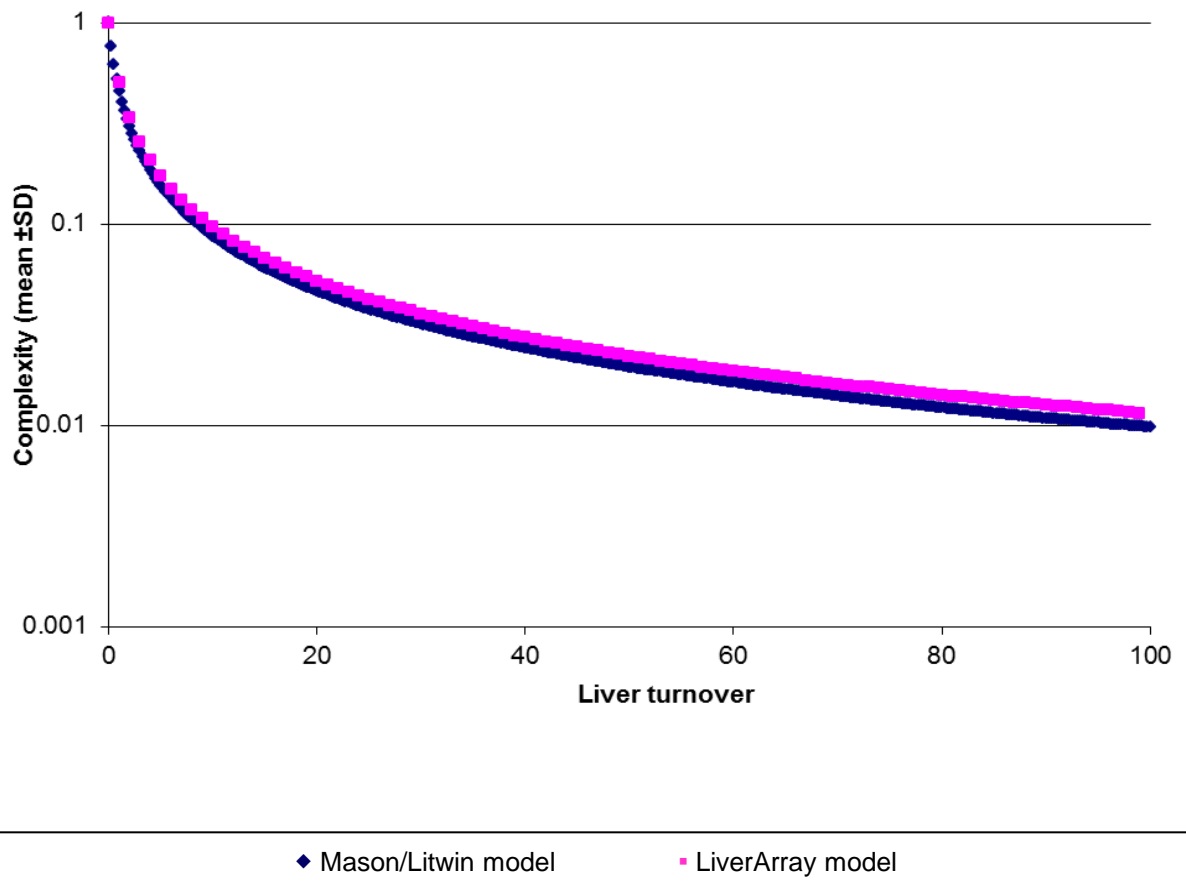


Figure 6.4 The current LiverArray model gives similar results to a previous simulation (mean±SD, n=10). The complexity of a simulation developed by Mason and Litwin (Mason *et al.* 2005) (blue) was compared to the complexity of the current LiverArray model (pink). For the Mason model, a 10^6 cell liver undergoing 100 liver turnovers was simulated. For the current LiverArray model, a 200^3 liver undergoing 100 liver turnovers was simulated with complexity = 1. Error bars are not visible at this scale.

produced very similar liver complexities over 200 liver turnovers. Both models show a gradual decline in complexity asymptotically approaching 0 as the number of liver turnovers increase. Thus, the current model is consistent with the Monte-Carlo-based simulation previously used by Mason *et al.* However, the current model had the added benefits of allowing us to study the geometry of the simulated liver and to measure absolute hepatocyte clone sizes and clone frequencies with respect to liver turnover.

6.2.2 The effect of LiverArray size on complexity, maximum clone size and clone frequency

As mentioned in Section 6.1.2, the complexity is a measure of the fraction of initial cell lineages that remain in the LiverArray. When run using different LiverArray sizes, ranging from 50^3 , 100^3 , 150^3 , 200^3 , 250^3 cells, the measured complexity did not significantly vary, as shown in Figure 6.5A. Thus, the decline in initial cell lineages did not vary between simulations with different LiverArray sizes, suggesting that the distribution of the cells selected for cell death is the same in all simulations.

The maximum clone sizes with respect to liver turnover increased as the size of LiverArray increased, as shown in Figure 6.5B. Furthermore, the graph of clonal frequency moves vertically with increasing size of the LiverArray, as shown in Figure 6.5C. These results are expected as increasing the number of cells (and, therefore, clones) increases the probability that larger clones may be formed by stochastic turnover.

A LiverArray composed of 200^3 cells was chosen for subsequent simulations: both due to practicality, as simulations would run in relatively short periods on desktop computers; and for statistical robustness, as the cell number is equivalent to ~100 mg of liver, 20 times the amount analysed by invPCR in previous chapters.

6.2.3 Connectivity weakly affects clonal proliferation

The distance of the replacing cell was changed from directly adjacent the dying cell (connectivity = 1) to any cell within LiverArray (connectivity = 200).

As shown in Figure 6.6A, the complexity of the LiverArray was slightly decreased as connectivity was changed from 1 to 200. Also, as shown in Figure 6.6B, simulations of a

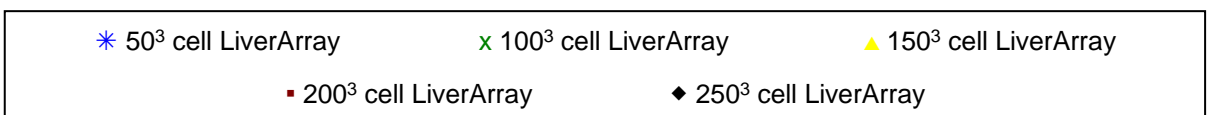
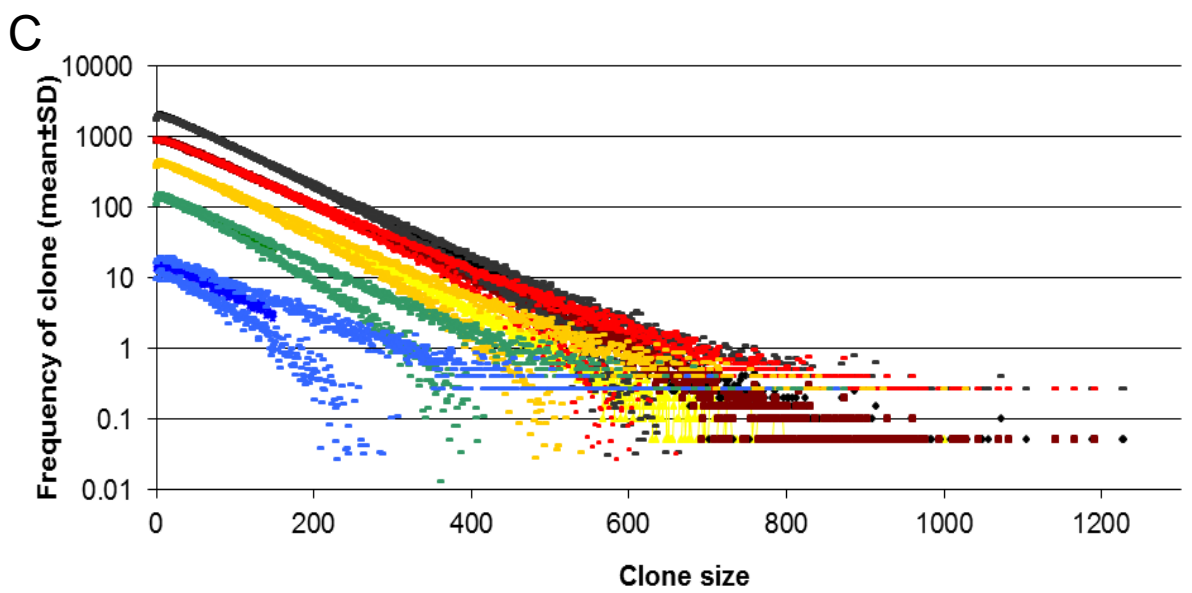
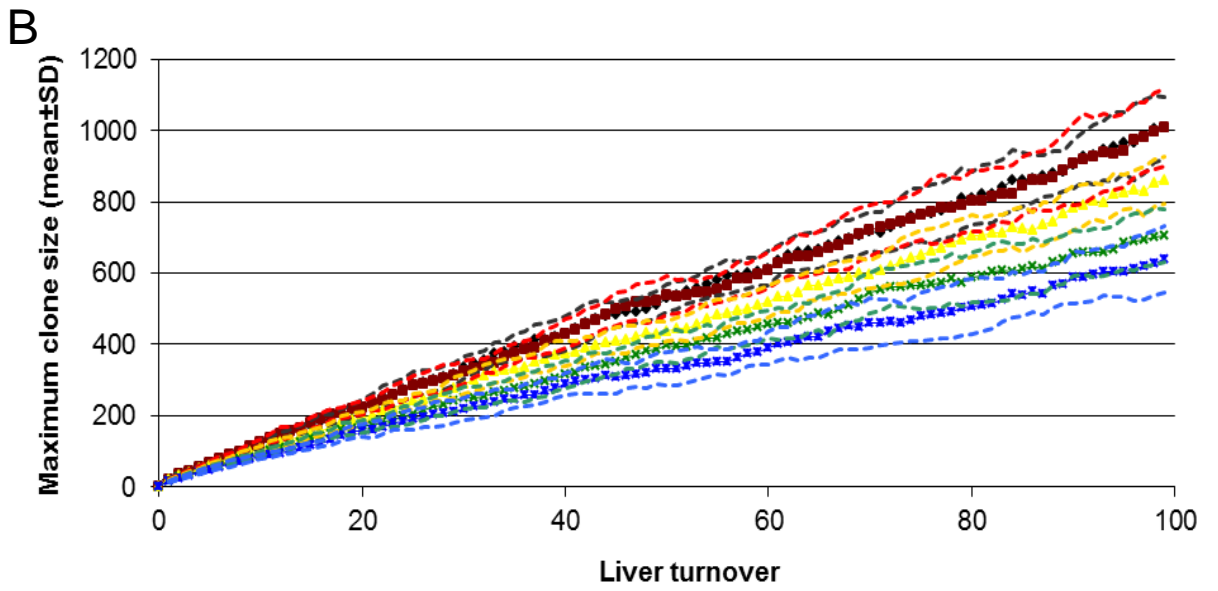
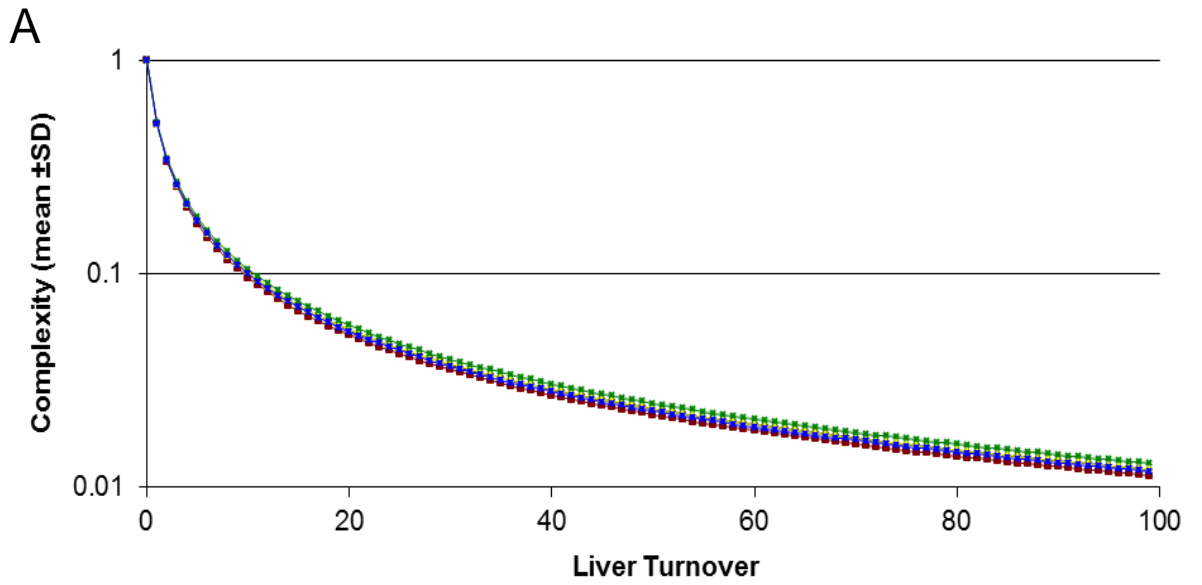


Figure 6.5. The effect of LiverArray size on clonal proliferation (mean±SD, n=20). LiverArrays composed of 50^3 (blue), 100^3 (green), 150^3 (yellow), 200^3 (red), and 250^3 (black) were simulated. Standard deviations for each LiverArray size data series are shown as shaded dashes. Different sizes of LiverArray retain identical levels of complexity with respect to liver turnover (6.5A) (Error bars are not visible at this scale). In addition, increasing size expectedly increases both the maximum clone size during liver turnover (6.5B) and shifts the final clonal frequency upwards (6.5C).

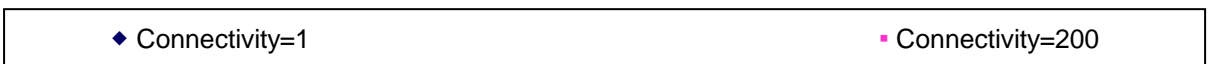
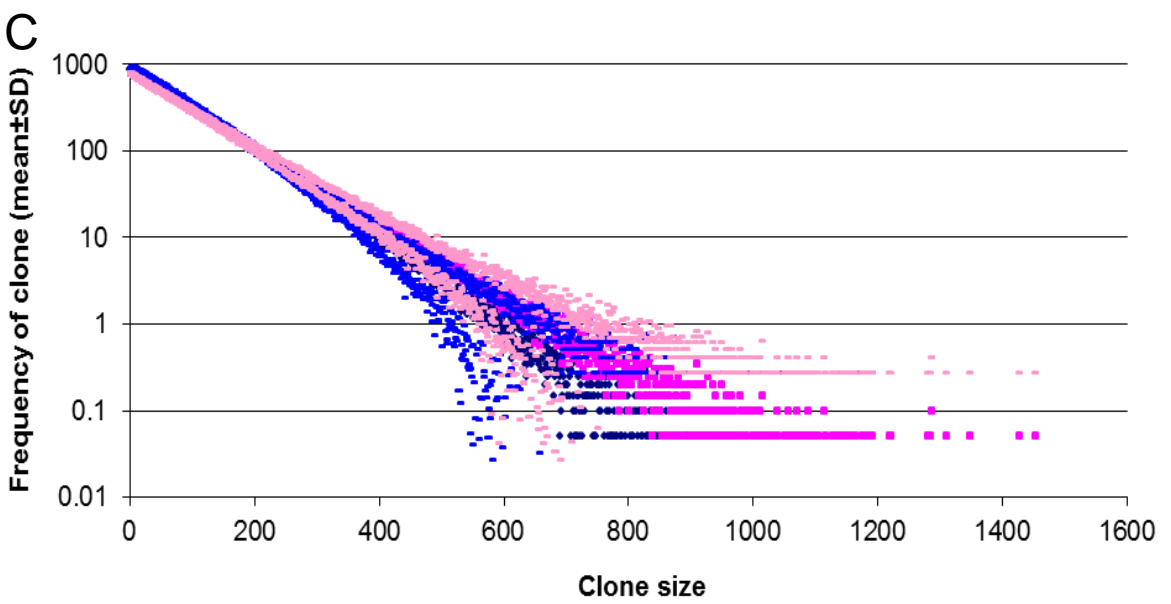
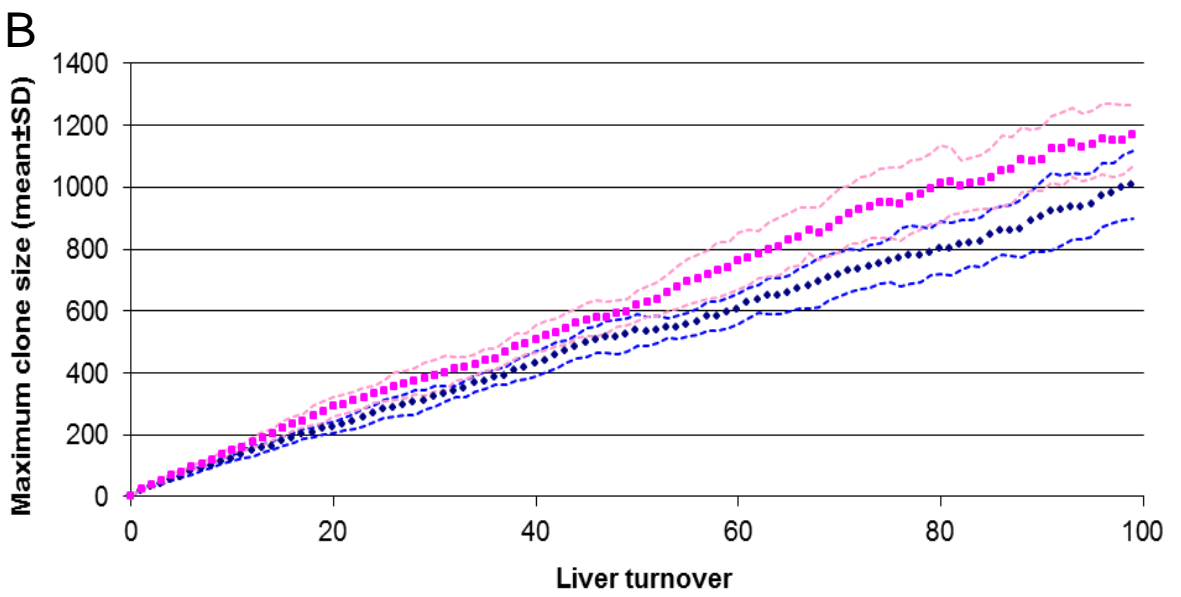
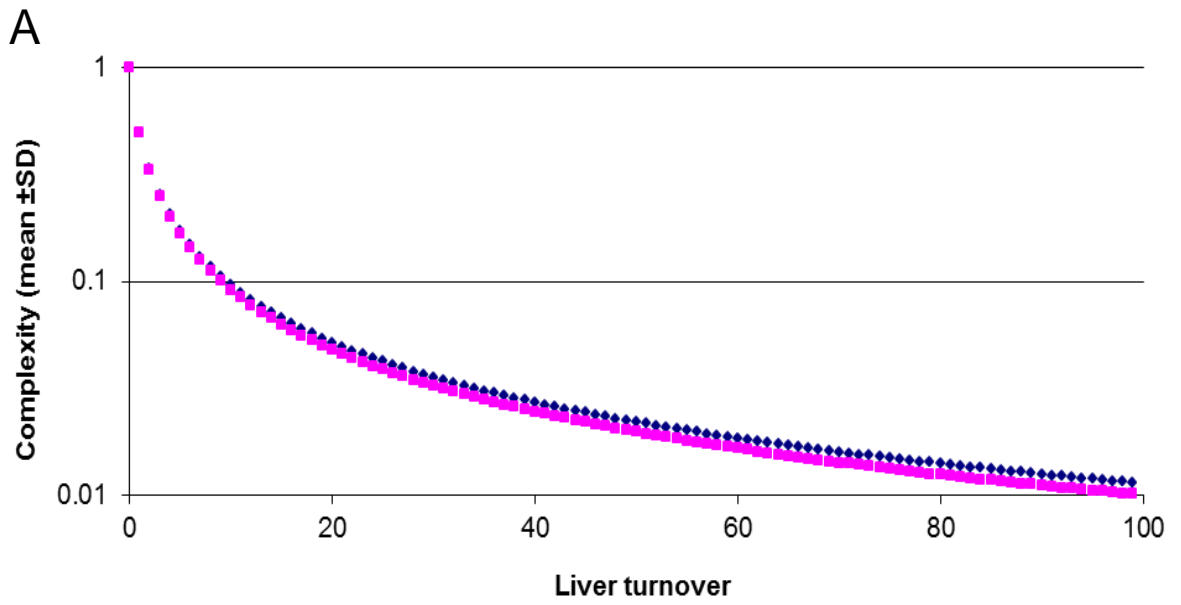


Figure 6.6 Changes in connectivity have a small effect on clonal proliferation (mean±SD, n=20). We changed the distance of the replacing cell from adjacent to the dying cell (connectivity = 1, blue) to any cell within the liver (connectivity = 200, pink). Standard deviations for each data series are shown as shaded dashes. Upon increasing connectivity, complexity slightly decreased (6.6A), maximum clone size slightly increased (6.6B) and clonal frequency after 100 liver turnovers was shifted downwards (6.6C).

connectivity of 1 produced slightly smaller maximum clone sizes than those with a connectivity of 200.

This may be due to the formation of spherical clones in the LiverArray where connectivity = 1. Hepatocytes in the middle of a spherical clone are guaranteed to be replaced by hepatocytes of the same clone and are thus “protected” against replacement by the progeny cells of other clones. This “wastes” a cell turnover without having changed the clone frequency or the maximum clone size. Thus, for connectivity = 1, maximum clone sizes are smaller than those of higher connectivity conditions with respect to liver turnover.

Furthermore, the lowered complexity of LiverArray with respect to liver turnover can also be explained by this phenomenon. As previously mentioned, complexity is a measure of the loss of unique hepatocyte clones. The formation of larger clones comes at the expense of the destruction of smaller clones. Thus, the smaller clone sizes in simulations with connectivity = 1 lead to the greater complexity.

However, on the whole, these differences were minor. As such, a connectivity value of 1 was used in most of the simulations because this represents a more likely model of what occurs in most instances of hepatocyte repopulation during the low to moderate cell death seen in chronic HBV infection, as previously described.

6.2.4 Stochastic clonal proliferation is highly predictable during liver turnover

The complexity, maximum clone size and clone frequency were measured in LiverArray composed of 200^3 cells with connectivity = 1. The LiverArray was modelled for up to 2500 liver turnovers. As shown in Figure 6.7, the greater the number of liver turnovers, the greater the maximum clone size found in LiverArray. Using the results of the simulation, the average maximum clone sizes ($\text{MaxS}_{\text{Clone}}$) was shown to be directly proportional to the number of liver turnovers since infection (t) governed by the equation:

$$\text{MaxS}_{\text{Clone}} \approx 7.64t + 1 \quad (r^2 = 0.9983) \quad (1)$$

The number of clones composed of >10000 cells at each time point were measured since integration of HBV DNA can occur throughout the duration of chronic infection and plotted in Figure 6.8. The number of >10000 cell clones remained at zero until >1000 liver turnovers. Even so, the number of clones >10000 cells remained relatively low as a percentage of the

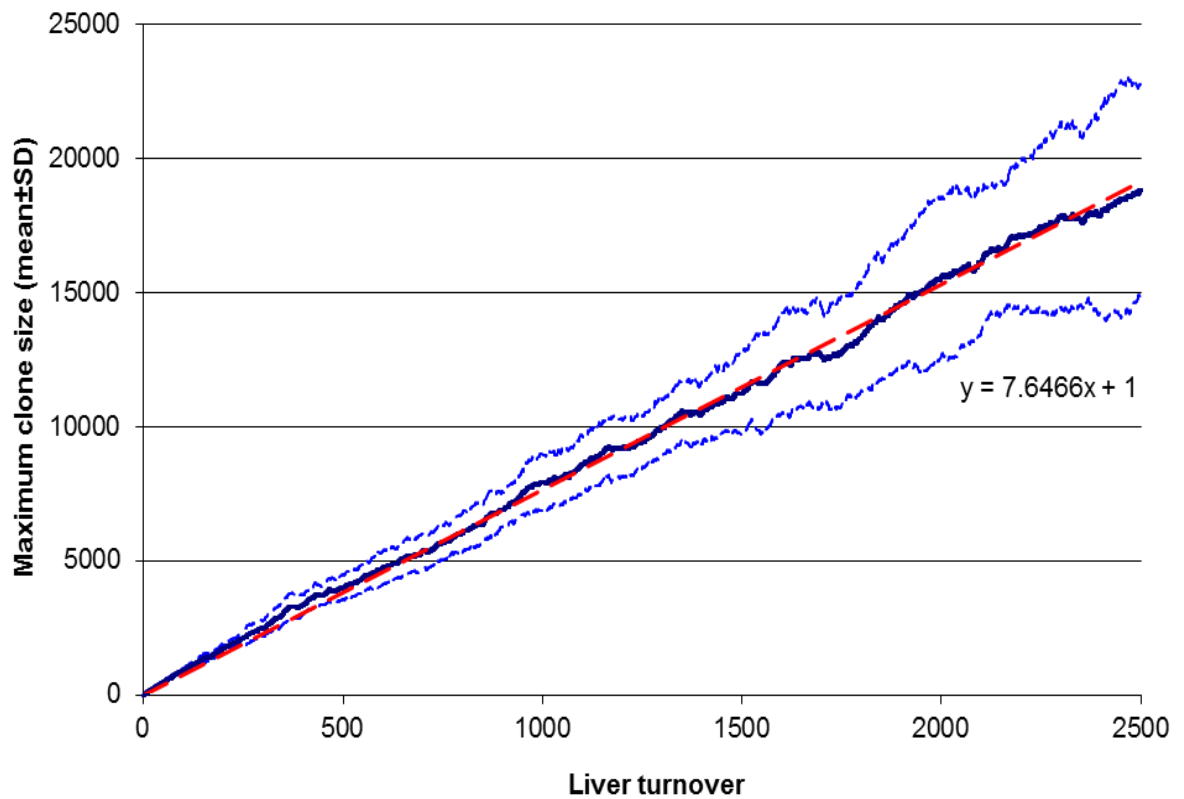


Figure 6.7 Maximum clone size increases as liver turnover increases (mean±SD, n=22). The largest clone size in a 200^3 cell LiverArray was measured at each liver turnover for 2500 liver turnovers (connectivity = 1). Standard deviation is shown as light blue lines. Line of best fit (red line) was produced by Excel 2005 least squares algorithm ($r^2 = 0.9983$).

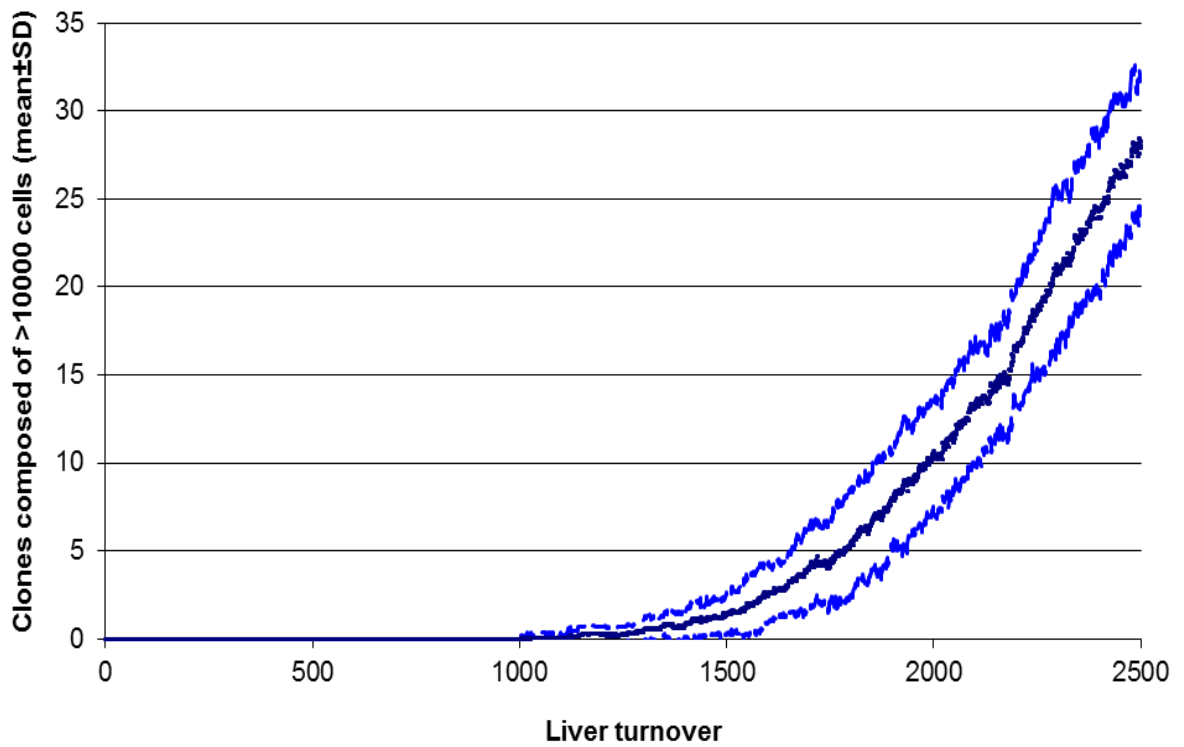


Figure 6.8 Number of clones >10000 cells remains relatively low (mean±SD, n=22). The number of clones composed of >10000 cells were measured in a simulation of a liver comprising of 200^3 cells undergoing 2500 random liver turnovers. Standard deviation is shown as light blue lines. This remains at 0 until ~1000 liver turnovers and then steadily increases.

total clones in LiverArray. For example, even after ~1750 turnovers only 5 (or 0.0000625%) of the original clones are composed of >10000 cells. To detect a single clone of >10000 cells in this instance, on average, 20% of all hepatocyte clones would have to be detected.

The clone frequency of the liver was also measured at particular points of liver turnover and plotted in Figure 6.9A. At each time point, the log of the frequency of a clone of a particular size (F_{clone}) had an extremely strong ($r^2 > 0.91$) linear relationship with clone size (S_{clone}), such that:

$$\log(F_{clone}) = m(S_{clone}) + c \quad (2)$$

where, m is the slope of the graph and c is the y-intercept.

c decreased and m became less negative as the liver turned over, i.e. cell clones grew larger at the expense of smaller clones. The changes in c and m that occurred over liver turnover were graphed to find linear relationships and therefore a predictive formula for clone frequency with respect to liver turnover.

As shown in Figure 6.9B, the increase in $1/-m$ is directly proportional to liver turnovers since infection (t). The line of best fit was such that:

$$1/-m = 2.01t \quad (r^2 = 0.999) \quad (3)$$

As shown in Figure 6.9C, c decreased at a linear rate with respect to the log of the liver turnovers since infection, such that:

$$c = -2.00\log t + 7.00 \quad (r^2 = 0.997) \quad (4)$$

Substituting into equation (2), the average clonal frequencies of a 200^3 cell liver can be predicted by the following equation:

$$\log(F_{clone}) = 7.00 - (S_{clone}/2.01t) - 2.00\log t \quad (5)$$

for $t > 0$.

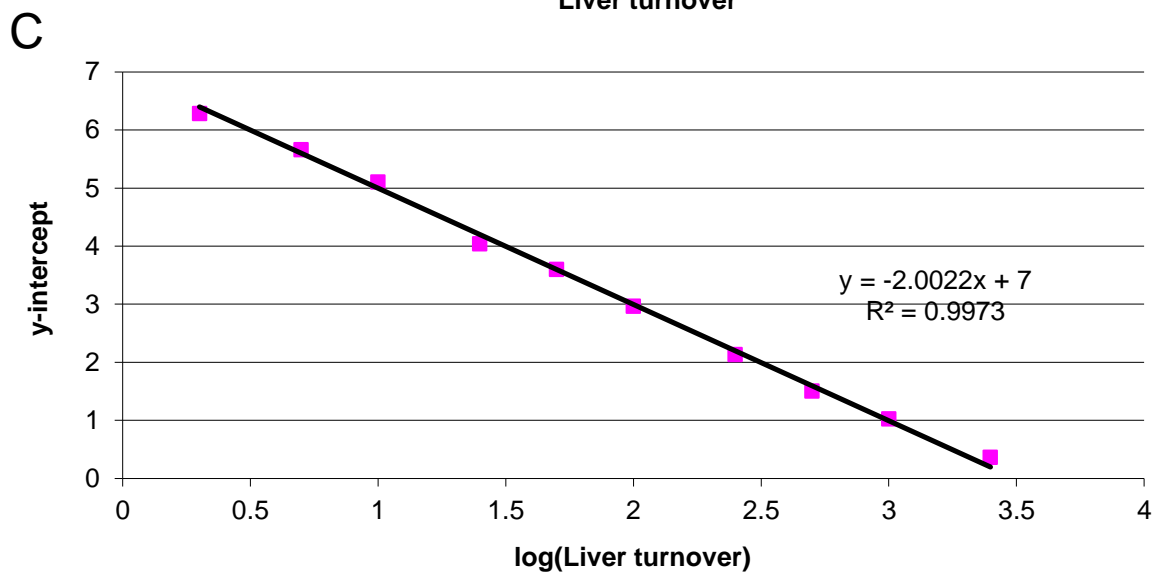
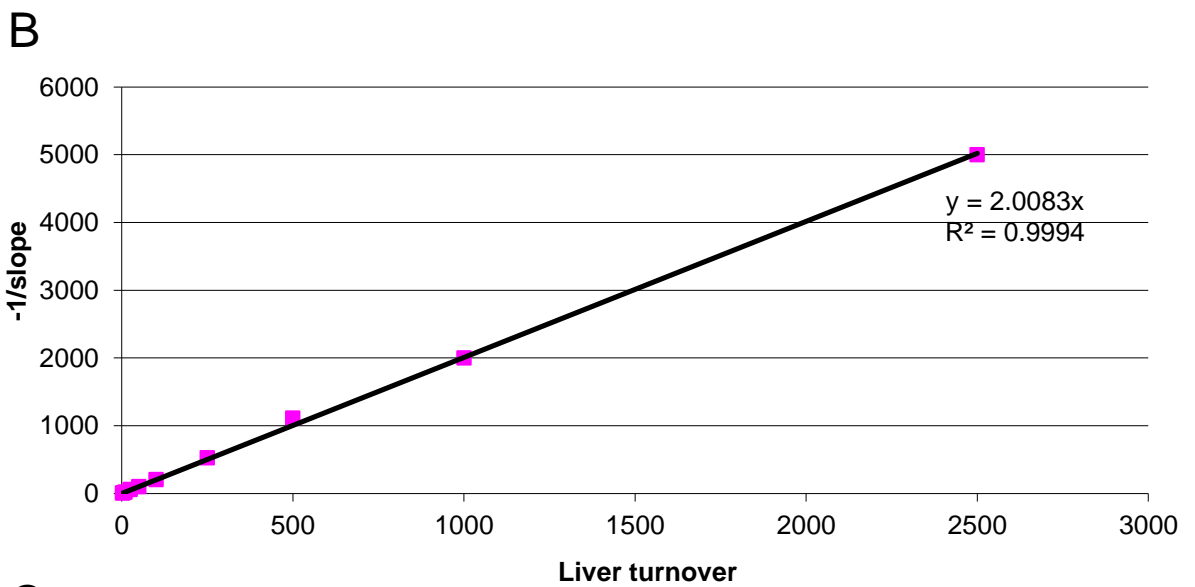
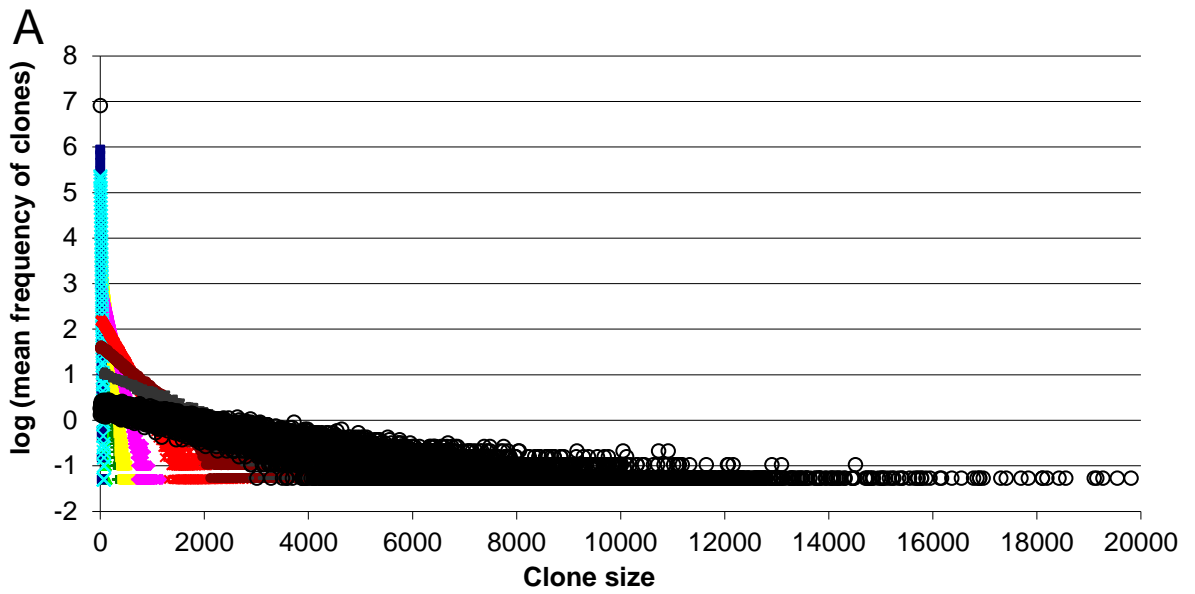


Figure 6.9 Clonal frequency over liver turnover is highly predictable (mean±SD, n=20). Figure 6.9A shows a very strong linear correlation between the log of clonal frequency and clonal size at every time point up to 2500 liver turnovers (TO). Trend-lines and error bars were omitted for clarity.

The slope of the line of best fit for the clonal frequency vs. size graph became less negative as LiverArray underwent cell turnover. The reciprocal of the negative of the slope was found to be directly proportional to the liver turnovers with a high degree of correlation ($r^2 > 0.99$) (6.9B). Additionally, the y-intercept of the line of best fit for the clonal frequency vs. size graph (6.9C) was inversely proportional to log of liver turnovers ($r^2 > 0.99$).

Given this equation, the approximate clonal frequencies contained within a LiverArray can be predicted without simulating the entire course of stochastic turnover.

6.2.5 Clonal proliferation due to liver growth during development is similar to that of 10 liver turnovers

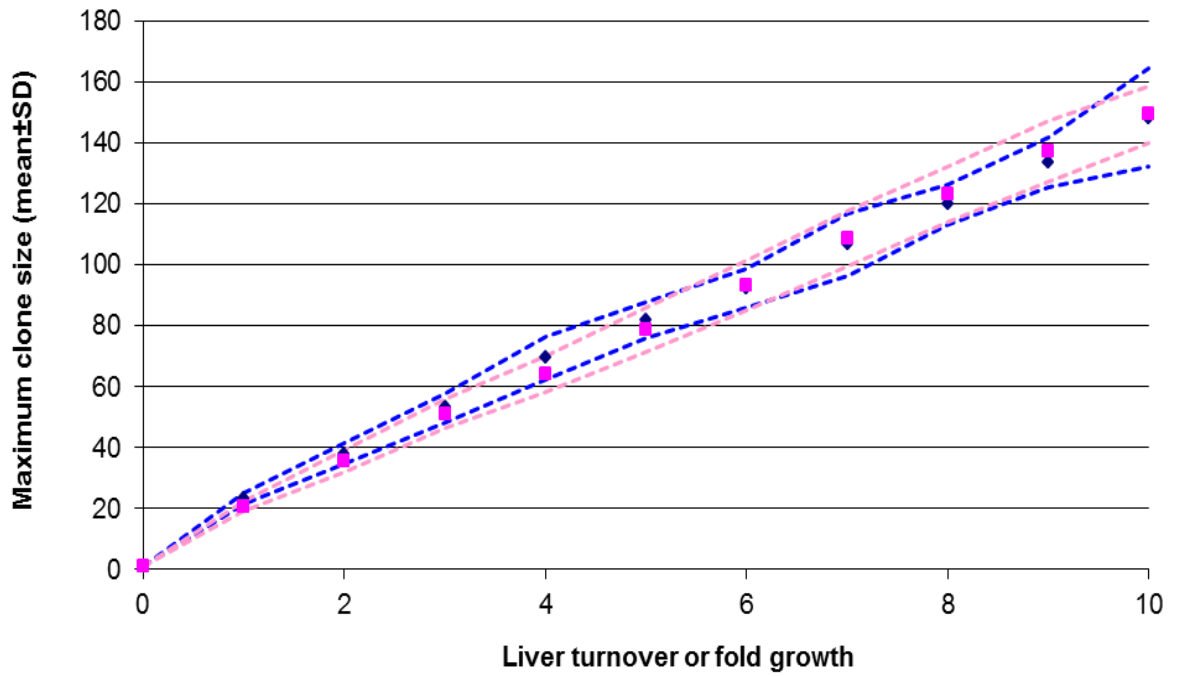
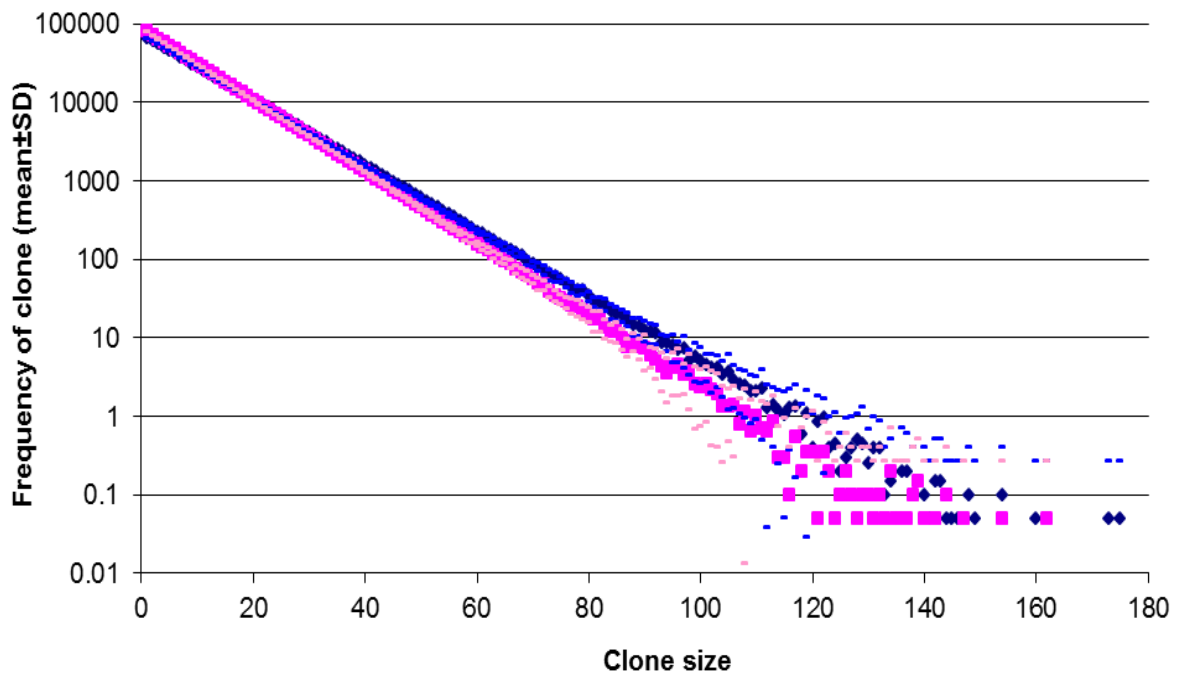
The effect of liver growth on clonal proliferation was simulated as explained in Section 6.1.2. During normal development, the liver mass increases ~10-fold (Shepard, Shi et al. 1988; Heymsfield, Gallagher et al. 2007). To simulate this, cells in LiverArray were chosen at random to proliferate until the liver grew 10-fold. Concurrently, a simulation of a liver of the same final size undergoing 10 liver turnovers (connectivity = 200) was run. As shown in Figure 6.10, there was no significant difference in maximum clone size or clonal frequencies between liver turnover and liver growth.

6.2.6 Focal liver cell death inhibits clonal proliferation

Simulations were run where the probability of cell death was inversely proportional to the distance from portal tracts (d) and d^2 for both distributions of focal cell death, as described in Figure 6.2. The ratio between distance-dependent focal cell death and random cell death was set at 0, 25, 50, 75 and 100%; where 0% means that all cell death is due to random cell death, and 100% means that all cell death is due to the distance from a portal tract as shown in Figure 6.3.

No obvious differences with respect to clonal proliferation were detected between the two distributions. For both, as the percentage of distance-dependent focal cell death was increased, complexity plateaued at a slightly higher level, shown in Figures 6.11A and 6.11D. While the maximum clone size increased slightly more quickly at early stages of liver turnover, eventually growth of maximum clone size was inhibited dependent on the percentage of focal cell death, as shown in Figures 6.11B and 6.11E. Predictably, the slope of the clonal frequency increased as the percentage of focal cell death increased, as shown in Figures 6.11C and 6.11F.

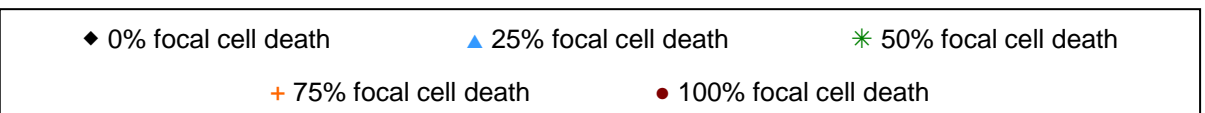
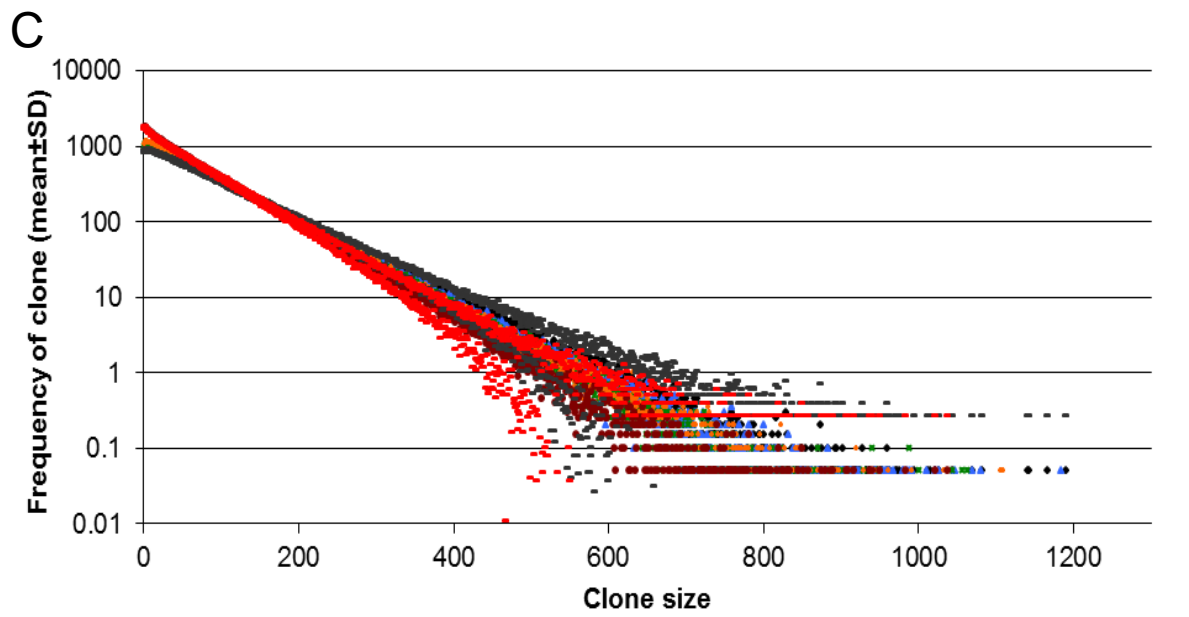
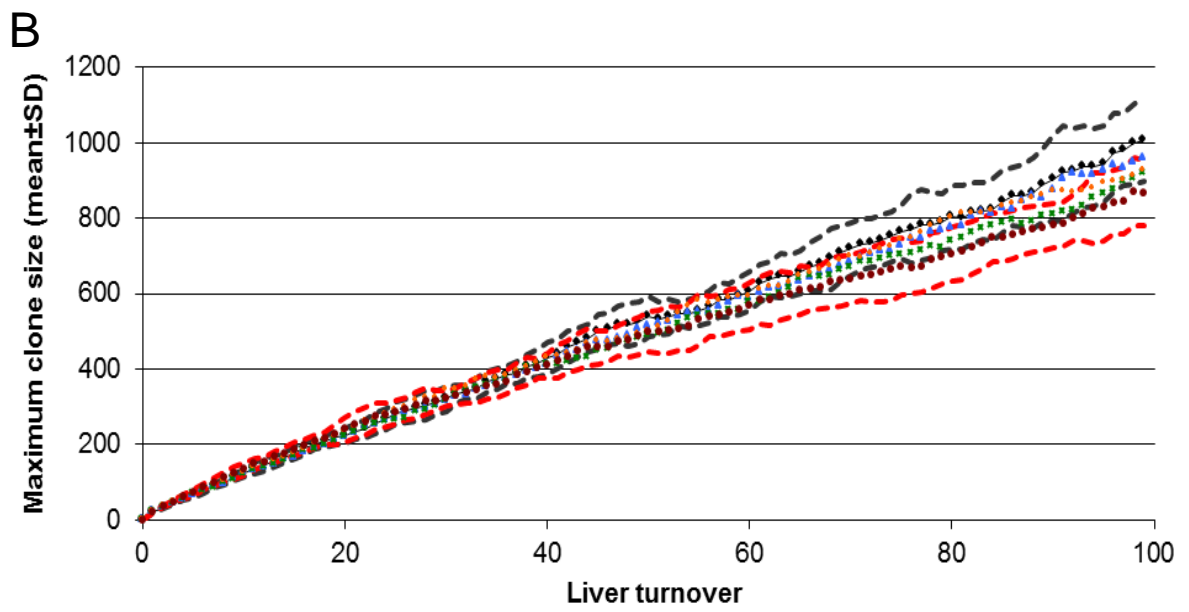
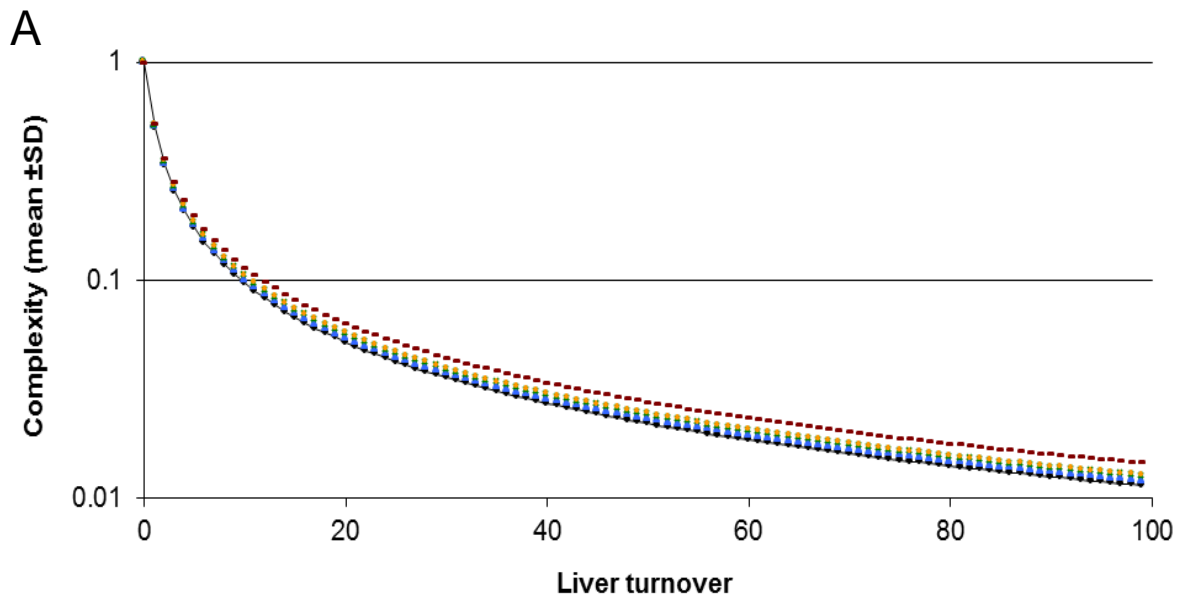
To visualise what mechanisms were creating this change in clonal proliferation, a 2-dimensional visualisation of LiverArray was created in VisualBasic (source code is shown in Appendix 9.7.4). The visualisation is a 100x100 cell array that mimics a smaller version of

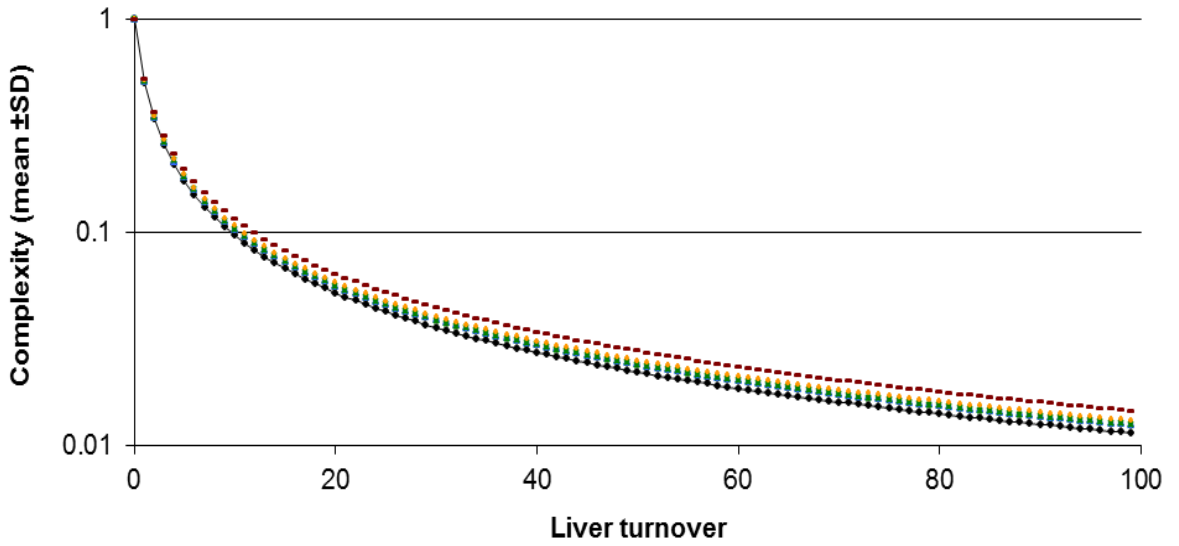
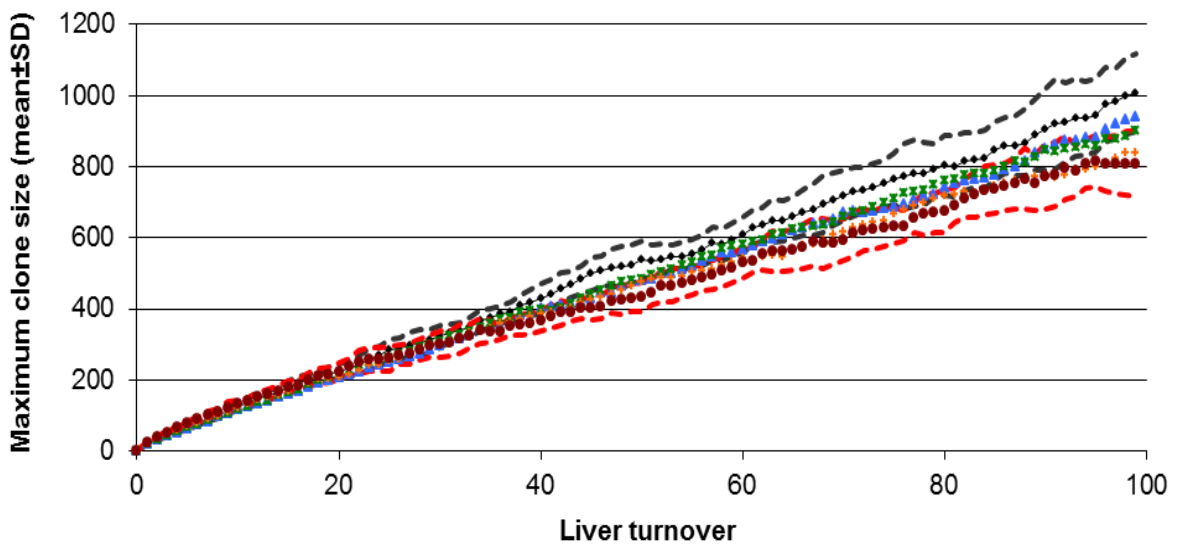
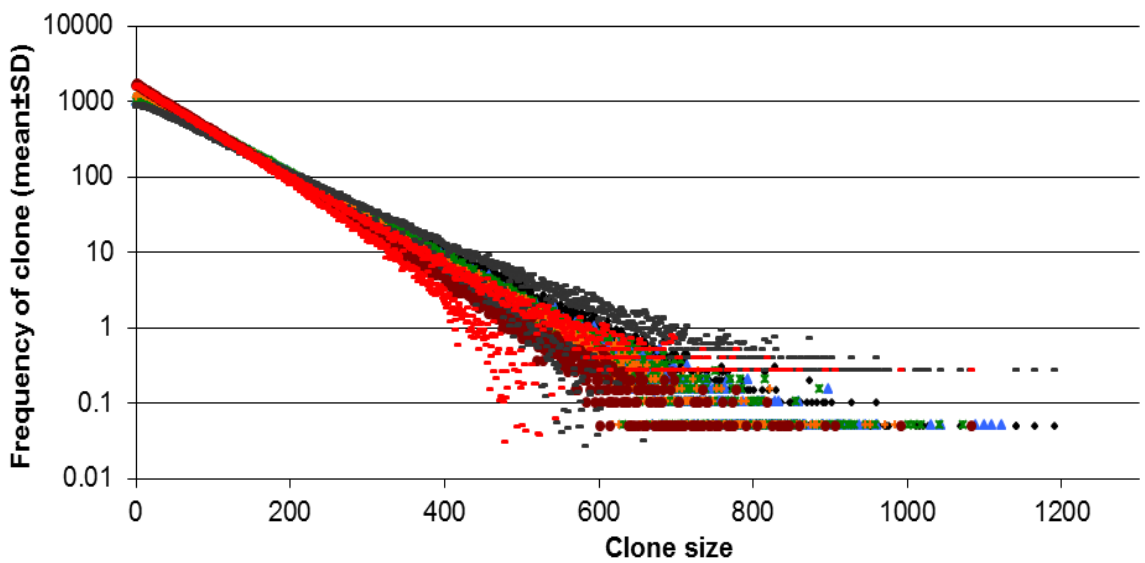
A**B**

◆ Random cell turnover (no growth)

■ Liver growth

Figure 6.10 Growth during development produces similar clonal proliferation to 10 liver turnovers (mean \pm SD, n=20). Simulations where LiverArray grew 10-fold were concurrently run with simulations where a fully grown LiverArray underwent 10 liver turnovers. Standard deviations for each data series are shown as shaded dashes. No significant differences in maximum clone size (6.10A) or clone frequency (6.10B) were observed between growth and turnover simulations.



D**Π****Π**

◆ 0% focal cell death

▲ 25% focal cell death

* 50% focal cell death

+ 75% focal cell death

● 100% focal cell death

Figure 6.11 Focal cell death slightly increases complexity and decreases maximum clone size (mean, n=20). Two different distributions of focal cell death described in Figures 6.2A and 6.2B were simulated and the results from each distribution were summarised in Figures 6.11A-C and 6.11D-F, respectively. Focal cell death was simulated by altering the probability of cell survival to be directly proportional to the square of the distance from portal tract, as described in Figure 6.3B. Standard deviations for only 0% (grey) and 100% (red) focal cell death are shown as shaded dashes for clarity.

For both models, simulations of 100 liver turnovers were carried out and complexity (6.11A and 6.11D) and maximum clone size (6.11B and 6.11E) in LiverArray were measured at each liver turnover. Small increases in complexity and decreases in maximum clone sizes were seen as the percentage of focal cell death increased. Furthermore, the clonal frequency graph (6.11C and 6.11F) became more steep as the percentage of focal cell death increased. The clonal proliferation observed in the two different distributions did not appear to differ significantly from one another.

Similar results were seen in simulations where cell survival is directly proportional to the distance from portal tract as described in Figure 6.3A (Data not shown).

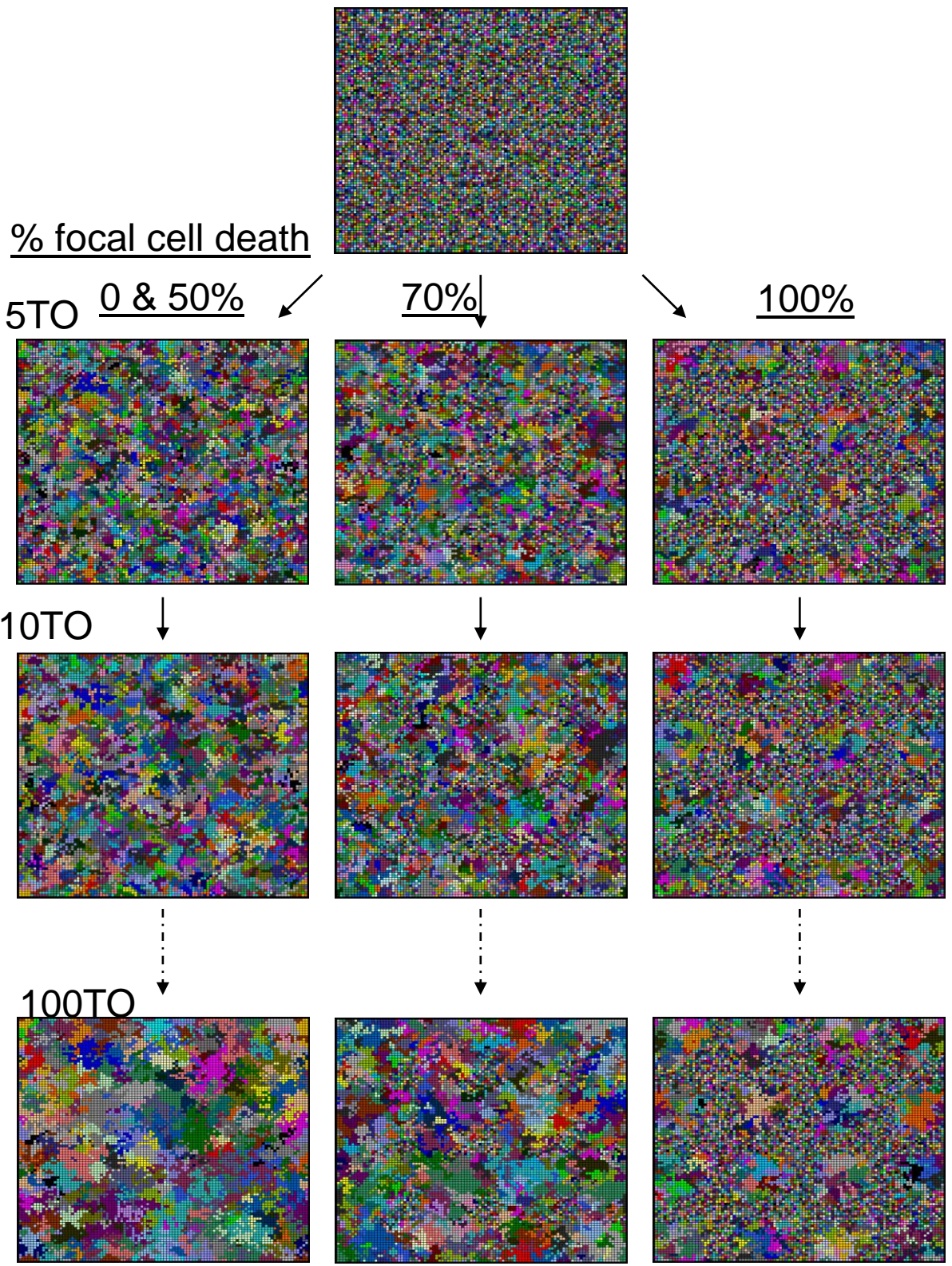


Figure 6.12 Visualisation of focal cell death. A 100^2 cell visualisation of LiverArray undergoing focal cell death was created in VisualBasic. At the start of each simulation, each cell was given a unique identifier, represented by a random colour. The simulation was run with various percentages (0, 50, 75 and 100%) of focal cell death, as detailed in Figure 6.3B. As liver turnover (TO) occurs, large clones form, represented by swaths of homogenous colour. In cases of high focal cell death, regions of low turnover can be seen as regions of “static”, where many small clones are congregated. These act as bottlenecks to clonal proliferation.

Similar results were seen in simulations where cell survival is directly proportional to square of the distance from portal tract as described in Figure 6.2B (Data not shown).

the 3-dimensional LiverArray. Visualisation of the simulation were run with various levels of focal cell death and shown in Figure 6.12.

While focal cell death increased the amount of liver turnover near the simulated portal tracts, the areas away from the simulated portal tracts undergo less liver turnover. Essentially, the proportion of the LiverArray undergoing turnover shrinks when focal cell death is implemented. As previously shown in Figure 6.5, a smaller LiverArray size decreases maximum clone size and shifts the clonal frequency graph downwards.

Additionally, hepatocytes located between focal areas of death that do not undergo as much turnover may act to bottleneck or provide complete barriers to clonal proliferation. Growth of a clone is directly proportional to the surface area facing hepatocytes undergoing turnover. By limiting the surface area of a clone that neighbours another clone undergoing turnover, clonal proliferation is slowed.

6.2.7 Clonal proliferation is exquisitely sensitive to the introduction of cells with survival advantage

Eighty cells (or 1 out of every 100000 cells) in LiverArray were randomly given a range of survival advantages and the maximum clone sizes were plotted in Figure 6.13. LiverArrays containing cells with a 25% and 10% survival advantage produced clones of >10000 cells within 50 and 90 liver turnovers, respectively. Interestingly, even a very slight survival advantage of 5% and 2% in such a small number of cells produced clones of >10000 cells within 150 and 350 liver turnovers, respectively. Thus, the amount of liver turnovers required to produce large clones is greatly reduced in livers containing hepatocytes with a survival advantage.

The exponential growth of maximum clone size seen in simulations containing cells with a survival advantage may be due to synergism between the resistance to killing and the increasing surface area of the clone composed of cells with a survival advantage. Greater in depth mathematical analysis may be able to explain these results more definitively, but this is beyond the limits of the current study.

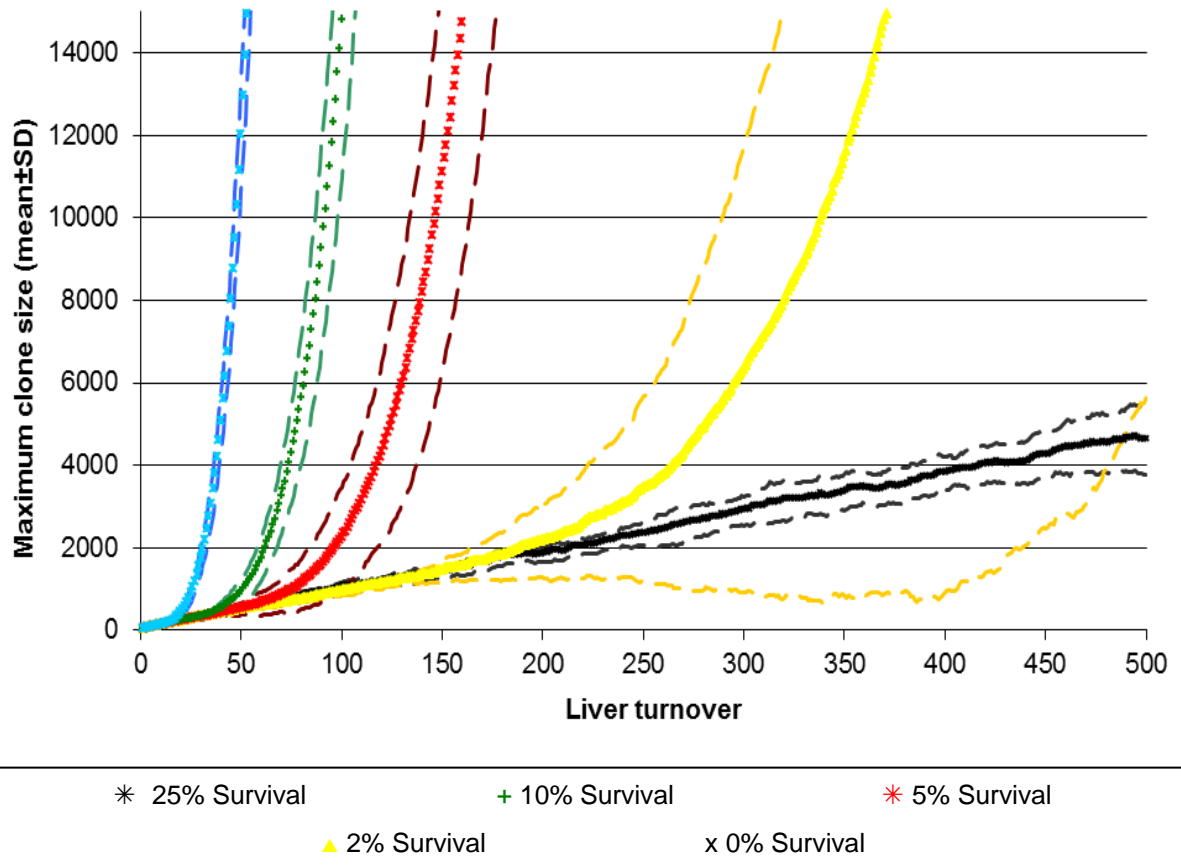
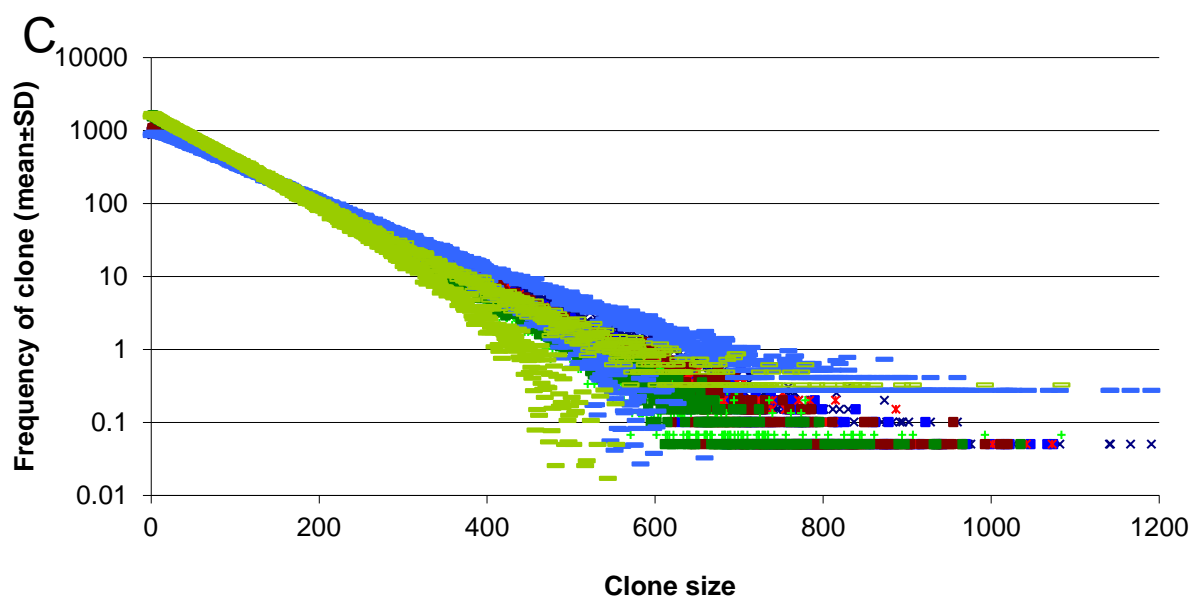
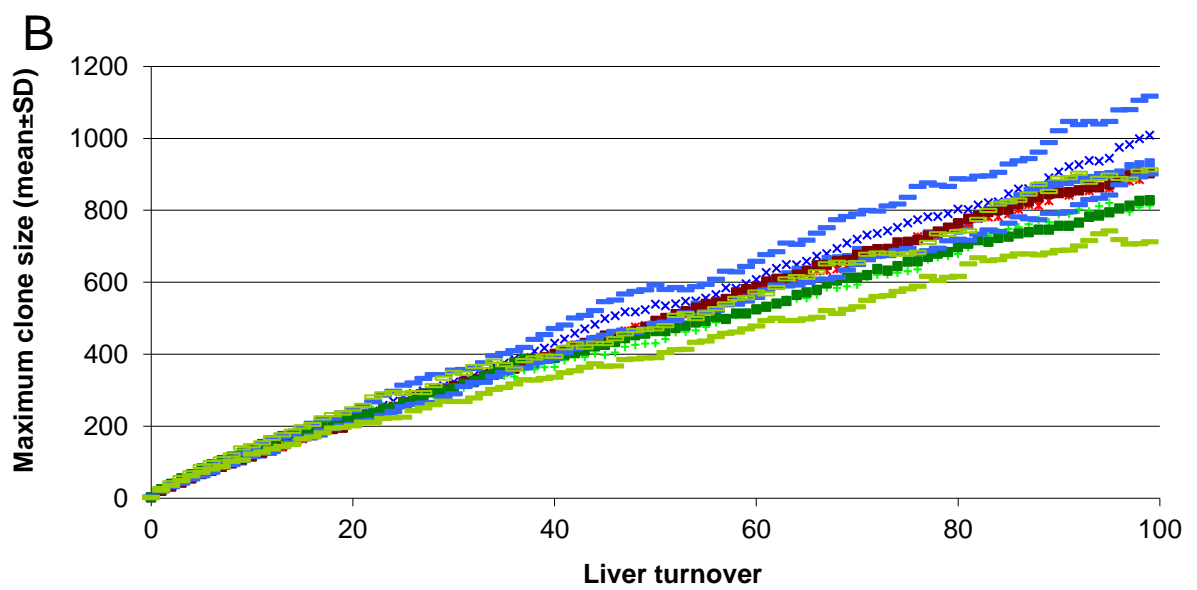
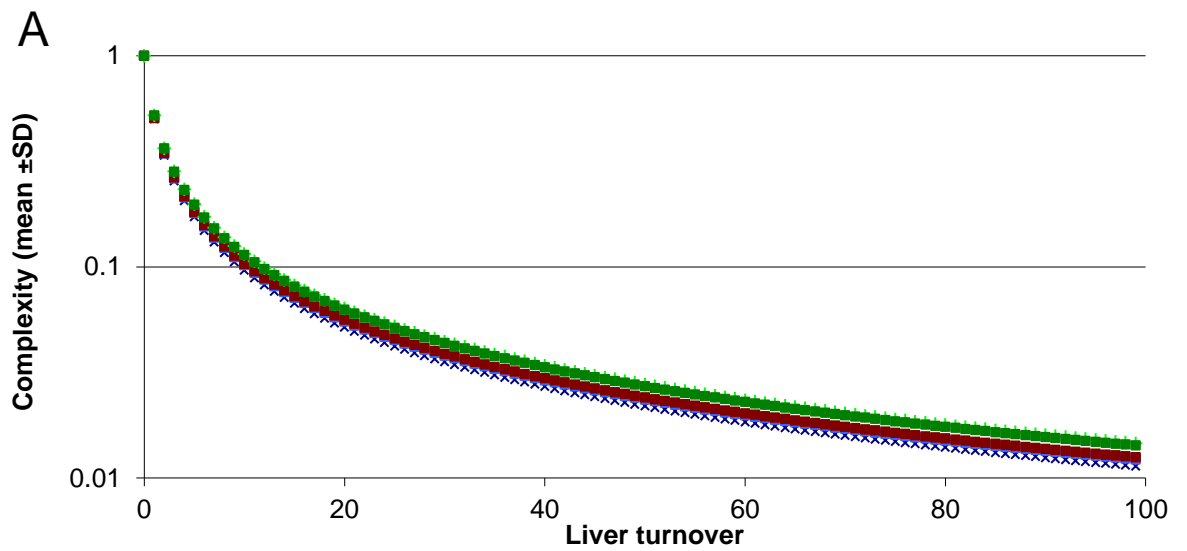


Figure 6.13 Maximum clone size is exquisitely sensitive to cells with survival advantage (mean, n=20). Maximum clone size was measured over 500 liver turnovers in LiverArrays containing 80 randomly chosen cells with 25% (blue), 10% (green), 5% (red), 2% (yellow) or 0% (black) survival advantage. Standard deviations for each data series are shown as shaded dashes. The large standard deviation seen in the 2% survival advantage data series is due to the extinction of all the hepatocytes with a survival advantage in some simulations, leading to a significantly lower maximum clone size. In simulations containing cells with a survival advantage, an average maximum clone size of 10000 cells is reached under such conditions within 40, 100, 150 and 350 liver turnovers respectively.



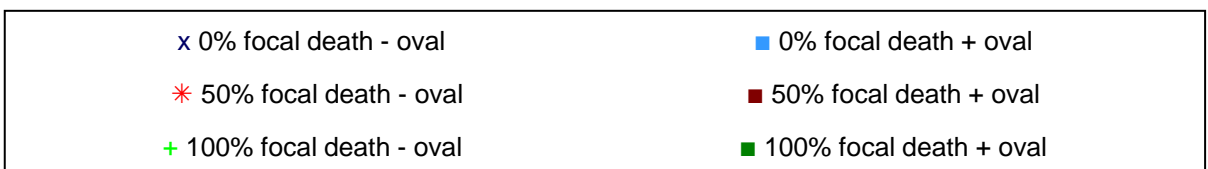
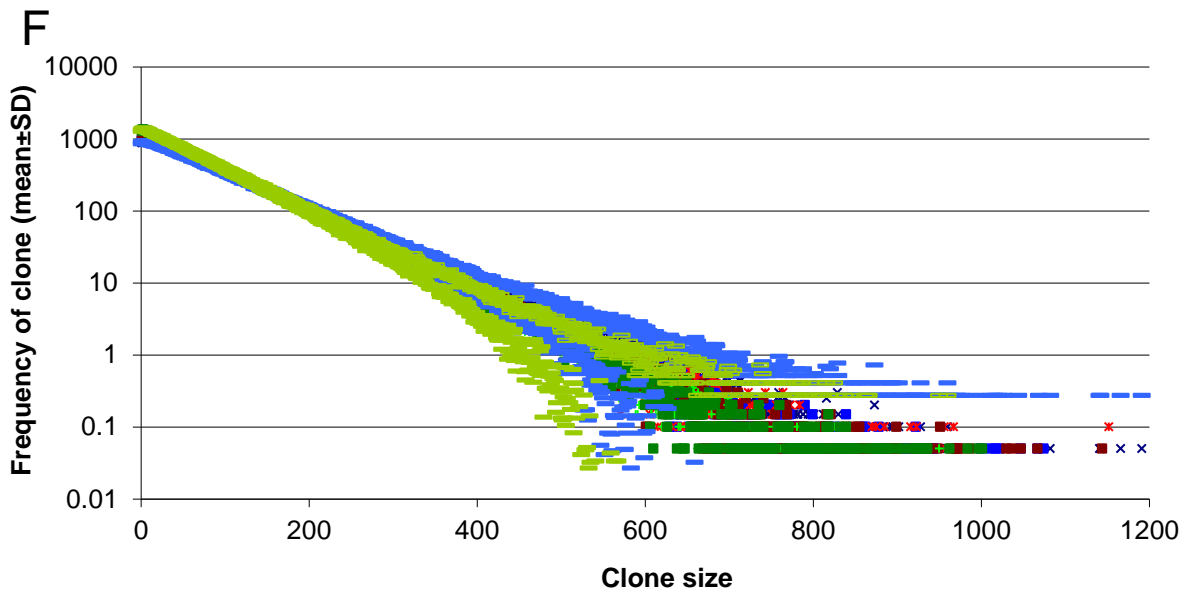
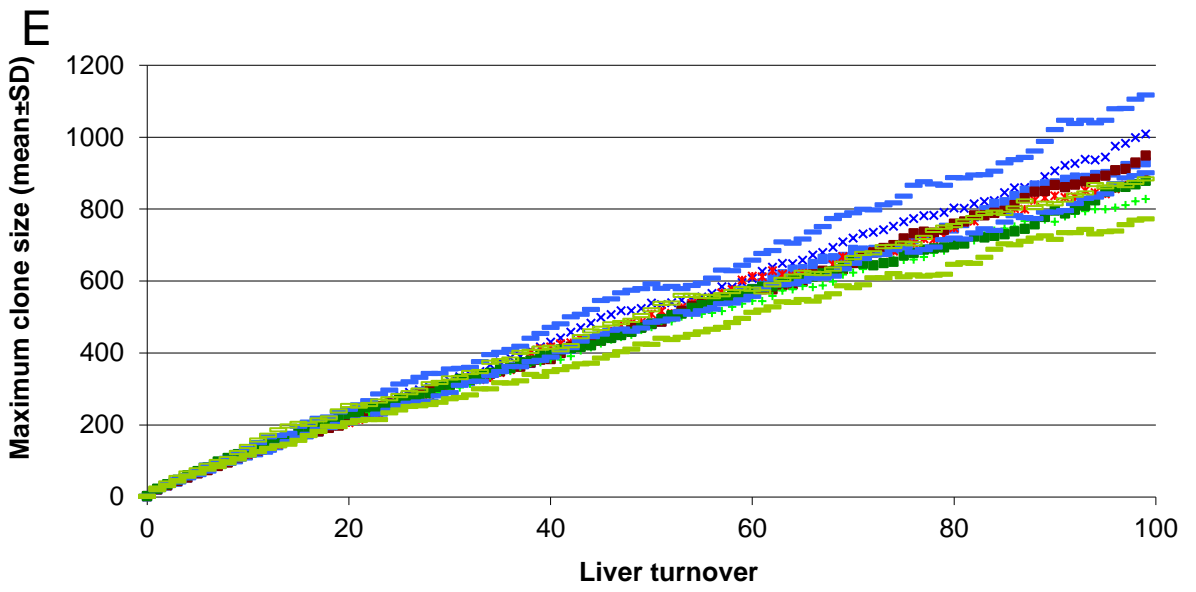
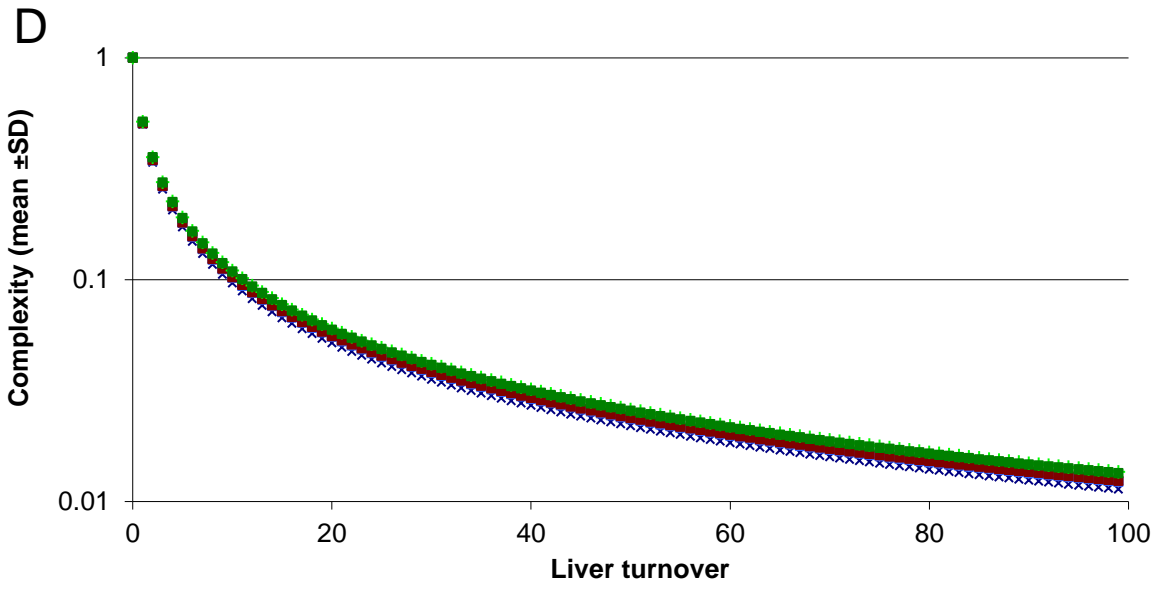


Figure 6.14 Oval cells have undetectable effects on clonal proliferation. Complexity and maximum clone size were measured over 100 liver turnovers in 200^3 cell LiverArrays containing 800 simulated oval cells. The clonal frequency of the final liver was also measured. In cases of cell survival is directly proportional to either the distance from portal tract (6.14A-C) or the square of the distance from portal tract (6.14D-F), no significant differences in complexity, maximum clone size or clonal frequency was observed between the simulations with and without oval cells. Standard deviations were only displayed for 0% focal cell death (blue dashes) and 100% focal cell death (green dashes) for clarity.

Similar results were seen in simulations where the distribution of focal cell death was simulated as described in Figure 6.2A (Data not shown).

6.2.8 The effect of oval cells is undetectable

Arbitrarily, eight hundred cells (or 1 out of every 10000 cells) were randomly designated as stem cells. In the program, the designated stem cells are immortal themselves and thus have a survival advantage of 1. When an oval cell is chosen to replace a neighbouring dying cell, its progeny cells do not share the same immortality as they differentiate into HBV-susceptible hepatocytes and therefore have a survival advantage of 0.

As shown in Figure 6.14, differences could not be detected in random cell-death and focal cell-death settings up to 100 liver turnovers between simulations that did and did not contain oval cells. This observation was seen in the two different distributions of focal cell death (as described in Figure 6.2) and the two different relationships between the distance from portal tracts and the probability of cell death (as described in Figure 6.3).

6.2.9 Estimation of maximum liver turnover in patients with chronic HBV infection

The amount of liver turnover that individual patients have undergone since HBV infection is unknown. Using PCNA staining indices of liver tissue sections, the reported average daily hepatocyte turnover in uninfected patients is ~0.05% (Mancini, Marucci et al. 1994). At this rate, it would take >5000 years to achieve 1000 liver turnovers.

Of course, the liver turnover rate is variable. Liver tissue from patients with chronic HBeAg-positive HBV infection have been found to contain $0.2 \pm 0.147\%$ hepatocytes positive for the proliferation marker Ki67 (Farinati, Cardin et al. 1996), corresponding to turnover of ~0.2% of hepatocytes per day (Gerdes, Lemke et al. 1984). At the height of HBV immune-clearance phase, PCNA studies have shown that 6% (Kawakita, Seki et al. 1992) to 43% (Wolf and Michalopoulos 1992) of hepatocytes are PCNA-positive. As cells accumulate nuclear PCNA in S-phase (Bravo and Macdonald-Bravo 1987; Morris and Mathews 1989), which last ~8 hours in hepatocytes (Schultze, Kellerer et al. 1978; Loyer, Glaise et al. 1994), this may equate to a daily hepatocyte turnover of 9-65%. Intermittent flares of cell-death detected by elevations of serum ALT levels are common in patients in immune-clearance and immune-control phase. The frequency of these flares in patients in immune-clearance and immune-control phase is ~0.268 and ~0.103 per year, respectively (Liaw, Tai et al. 1987). Generally, these flares last for less than a month (Liaw, Yang et al. 1985).

Higher estimates of all these measurements taken into account are:

- A background hepatocyte turnover rate of 0.5%/day;
- A frequency of flares of 0.27/year;
- Hepatocyte turnover during flares of 65%/day;
- Flares lasting a full month.

Using these higher estimates, an approximate upper limit for liver turnover was calculated to be 730 liver turnovers for a hypothetical 100-year-old patient. In the current simulations of 200^3 hepatocytes as shown in Figure 6.7, a clone size of 10000 cells lies outside the 99.9% confidence interval of the simulated maximum clone size at 730 liver turnovers. This strongly suggests that the observed hepatocyte clones were not produced by stochastic liver turnover.

6.3 Discussion

In Chapter 5, clone sizes of >10000 cells were detected using both invPCR and qPCR assays independently, consistent with the previous literature (Mason, Jilbert et al. 2005; Mason, Low et al. 2009; Mason, Liu et al. 2010). Data from the current simulation suggests that these hepatocyte clones could not have been produced by stochastic liver turnover.

The robustness of the current simulation has been shown by replicating the results of a previous simulation used to simulate clonal proliferation in the liver. The current simulation expands on the Mason/Litwin model by taking the geometrical relationships between cells into account and quantifying not only complexity, but also clone sizes and frequency.

According to the simulations, >1000 liver turnovers would need to occur after infection to stochastically produce an observable clone of >10000 cells, if 100% of the hepatocyte clones are detected. However, experimental and bioinformatics analysis shown in Section 5.3.2 suggested that much less than 20% of all hepatocyte clones were detected by invPCR. At a hepatocyte clone detection rate of 20%, the turnovers required to produce the observed hepatocyte clones by stochastic liver turnover is, on average, ~1750 liver turnovers. Our calculations using conservative figures from previous data estimate an upper limit of ~730 liver turnovers in 100 years, falling short of the requisite minimum of ~1750 turnovers. This contradiction suggests that the hepatocytes in the observed large clones were not produced by stochastic liver turnover.

Increases in liver turnover could have been induced by mechanisms other than the natural liver turnover and immune-mediated cell death against HBV-infected hepatocytes. For example, hepatotoxins, such as alcohol and paracetamol, could have accelerated hepatocyte death over the lifetime of the patients (Hall 2007; Dadarkar, Fonseca et al. 2010). However, hepatotoxins induce cell-death concentrated in peri-portal hepatocytes due to the concentration gradient of toxins between peri-portal and peri-central areas (Popper, Paronetto et al. 1965; Powell, Kosyk et al. 2006; Theise, Bodenheimer Jnr. et al. 2007; Dadarkar, Fonseca et al. 2010). The current simulation indicates in Section 6.2.6 that while growth of clones is increased due to an increased liver turnover, clone growth is slower with focal cell-death than an equivalent amount of stochastic cell death.

Focal cell death probably inhibits clonal proliferation through at least two mechanisms: 1) by decreasing the effective size of LiverArray undergoing liver turnover and; 2) by limiting the surface area of hepatocyte clones that neighbours hepatocytes undergoing turnover. In the programmed distribution of focal cell death, these inhibitory mechanisms outweigh enhancement of clonal proliferation by the increase of liver turnover in a given area. However, we have used an approximate distribution of the focal cell death that may occur in the HBV-infected liver. Further studies may reveal alternate distributions of focal cell death that enhance clonal proliferation rather than inhibit it.

The effect of liver growth during development was minimal, as the modelled growth of the liver produced clonal frequencies and maximum clone sizes almost indistinguishable from 10 liver turnovers. Also, our simulations suggest that connectivity (i.e. the distance between a dying hepatocyte and the hepatocyte that undergoes compensatory mitosis) does not have a strong effect on the clonal proliferation of hepatocytes.

Furthermore, our simulation shows that even a slight survival or growth advantage in a small subset of hepatocytes produces clones much more quickly than random cell death. This suggests that even subtle changes in cell phenotype can convey a survival or growth advantage. In summary, apart from clonal proliferation of hepatocytes with a survival advantage, few alternate hypotheses remain that may explain the large hepatocyte clones detected by invPCR in the liver tissue of patients with chronic HBV infection.

There are some limitations with the LiverArray simulation. Firstly, the simulation does not model ongoing integration. The rate of ongoing integration of HBV DNA into chromosomal

DNA during chronic HBV infection is not known and is likely to depend on host-virus interactions, which are variable among patients. Despite this, the simulation can still differentiate between stochastic liver turnover and clonal proliferation by survival or growth advantage models. Clone frequency would be affected by adding greater numbers of small hepatocyte clones. However, ongoing integration would only increase complexity and would not alter the maximum clone sizes observed in the liver.

Another limitation is that the LiverArray simulation does not model normal liver structures, such as hepatocyte plates, sinusoids and epithelial lining. This simulation models the liver as a solid cube composed entirely of hepatocytes, neglecting the fine structure seen in normal liver. Hepatocytes within the liver lobules are arranged in plates that radiate from portal veins (Roskams, Desmet et al. 2007). The space between these walls (sinusoids) contains epithelial cells, collagen and other potential barriers to hepatocyte repopulation by surrounding hepatocytes. This may increase connectivity along plates compared to between plates. While simulations have not been run with such a limited value, the current data shows that connectivity plays a relatively weak role in clonal proliferation, suggesting clone sizes would not be greatly affected by changes in connectivity.

Since the structure of the liver is not a simple network of repeated identical units, potential future applications of the computer model might include additional considerations, for example, heterogeneity within the hepatocyte population (Jungermann and Katz 1989; Jungermann and Kietzmann 1996), differences in cellular ploidy (Toyoda, Bregerie et al. 2005), and dynamic changes in immune responses and levels of HBV replication (Murray, Wieland et al. 2005; Li, Wang et al. 2011). The inclusion of these features may give rise to complex relationships and lead to non-linear relationships between liver turnover and clonal proliferation of hepatocytes. For example, it may allow modelling of the outgrowth of large hepatocyte clones arising in areas of focal cell death.

Additionally, the current computer simulation does not model abnormal liver structure, such as fibrosis, cirrhosis, and regenerative nodules. During the course of chronic liver disease, alteration in liver microstructure can occur including fibrosis, cirrhosis, regenerative nodules, and hepatic necrosis (Horney and Galambos 1977; Liaw, Chu et al. 1982; Giannelli, Quaranta et al. 2003; Canbay, Friedman et al. 2004; Lavanchy 2004; Villeneuve 2005; Kim, Oh et al. 2009; Lok 2009). However, the current simulation shows that focal death only decreases hepatocyte clone sizes throughout the liver. Furthermore, large hepatocytes clones have been

observed in non-cirrhotic tissues in the current and previous studies (Mason, Liu et al. 2010), so fibrosis and cirrhosis are evidently not required for the formation of large hepatocyte clones.

Potentially, oval cells as long-lived cells, may act as a reservoir for a particular hepatocyte clone. However, when oval cells were simulated, no effect on clonal proliferation was observed. This emphasises the requirement for survival advantages to be heritable in order for large hepatocyte clones to be produced. However, data from the current simulation does not rule out the role of oval cells in clonal proliferation. Some mechanisms not yet explored are oval cell proliferation, oval cell migration into highly proliferative areas and oval cells conferring a temporary survival advantage during differentiation into hepatocytes (Factor, Radaeva et al. 1994; Grisham 1994; Lowes, Brennan et al. 1999; Paku, Schnur et al. 2001; Fellous, Islam et al. 2009). The data on these phenomena are scarce and further research is needed to provide an accurate simulation.

In summary, the current simulation has modelled the effects of liver growth during development, of different connectivities, of stem cell populations, and of focal liver death on clonal proliferation of hepatocytes in a ~100 mg fragment of liver. Overall, the geometrical alterations that were implemented did not increase the rate of clonal proliferation and, in some cases, the rate of clonal proliferation decreased. Furthermore, the average clonal frequency and maximum clone sizes produced by stochastic liver turnover can be predicted with a developed equation. The data shows that it is unlikely that the observed >10000 cell clones detected in the liver of patients with chronic HBV could have developed stochastically, given current estimates of liver turnover. Also, the simulation showed that rare hepatocytes with even a slight survival advantage could quickly outgrow hepatocytes without a survival advantage.

Thus, there is strong evidence that the observed hepatocyte clones described in Chapter 5 have not come about due to stochastic liver turnover. We have also presented evidence showing that the clonal proliferation of hepatocytes with a heritable survival or growth advantage is a plausible explanation for the observed clonal proliferation. In Chapter 7, hepatocytes that have undergone clonal proliferation have been isolated and various aspects of their cell phenotypes have been investigated to determine potential mechanisms for the observed clonal proliferation.

7 - Characterisation of hepatocyte foci

7.1 Introduction

Diseases associated with chronic HBV infection, e.g. HCC and cirrhosis, generally occur after many decades (Goodman and Terracciano 2007; Tan, Yeh et al. 2008). During this period, changes in the hepatocyte population occur as a result of virus-host interactions in the HBV-infected liver, which increases the risk of HCC by 100-fold compared to uninfected patients (Beasley 1988). While a wide range of changes that occur during the progression to HCC have been reported (described in Section 1.8), how these changes arise remains unknown. We therefore studied the changes that occur in the hepatocytes of patients with chronic HBV infection in an attempt to determine their cause and role in disease progression.

Changes in the HBV-infected liver cell population have been shown to lead to the formation of large hepatocyte clones. Clones composed of >10000 cells have been detected using invPCR and confirmed using qPCR in the liver tissues of patients with chronic HBV infection, as detailed in Chapter 5 and as previously reported (Mason, Liu et al. 2010). Using the mathematical modelling described in Chapter 6, our data strongly suggests that the cells in these large clones must necessarily have a survival advantage. Furthermore, as described in Chapter 5, clonal proliferation of hepatocytes has been associated with HBV disease progression.

The phenotype(s) of hepatocytes that have undergone clonal proliferation are unknown. In this Chapter, the following aspects were investigated:

1. Whether cell populations other than hepatocytes contain detectable virus-cell DNA junctions. If stem cells or cells located in fibrotic tissue contain detectable virus-cell DNA junctions, then the interpretation of the clonal proliferation of hepatocytes observed using invPCR in Chapter 5 would be complicated.
2. Whether the hepatocytes that form the observed clones have normal histology. Both large cell changes (LCC) and small cell changes (SCC) had been observed in the sections of liver tissue taken from patients with chronic HBV infection (Chapter 4). We investigated whether these histological changes are associated with the clonal proliferation of hepatocytes.
3. Whether the hepatocytes express HBV antigens or contain high levels of HBV RNA.

4. Whether the HBV DNA contained within the hepatocytes is altered, by changes such as methylation or mutations.
5. And lastly, whether the hepatocytes that have undergone clonal proliferation have altered cellular protein expression.

Examination of these factors may give insight into the mechanisms by which hepatocytes undergo clonal proliferation and may provide insight into the progression of HBV-associated liver disease.

7.1.1 Detectable virus-cell DNA junctions in cells other than hepatocytes

While HBV primarily infects hepatocytes, previous literature has suggested that other cell populations, such as hepatocyte progenitors or oval cells, can also support virus replication (Hsia, Thorgeirsson et al. 1994). If these cell populations can be infected by HBV and thus contain integrated HBV DNA, then the large clones detected by invPCR may be due to the proliferation of hepatocyte progenitor cells or oval cells. Furthermore, if other cells in fibrotic tissue, such as stellate cells, can support HBV replication, then the observed clonal proliferation may be due to excessive fibrogenesis in those patients with fibrosis and cirrhosis.

However, data shown in Chapter 4 indicated that no cells other than hepatocytes support HBsAg antigen expression. Explicitly, oval cells were not observed to support the expression of HBsAg. This data suggests that oval cells are not susceptible to infection and therefore the multiple copies of virus-cell DNA junctions observed by invPCR are the result of the clonal proliferation of hepatocytes. We therefore hypothesise that virus-cell DNA junctions will not be detectable in isolated cell populations other than hepatocytes.

7.1.2 Histological changes in hepatocytes that have undergone clonal proliferation

We aim to investigate histological changes, such as LCC, which are associated with HBV-associated liver disease as described in Section 1.7, as sources of clonal proliferation. Previous studies have suggested that hepatocytes with LCC represent a senescent cell population (Lee, Tsamandas et al. 1997; Ikeda, Sasaki et al. 2009). We have previously observed hepatocytes with LCC arranged in foci (Section 4.3.1). Potentially these foci of hepatocytes with LCC may be composed of hepatocytes that have previously undergone clonal proliferation, and are now halted by senescence.

On the other hand, the level of clonal proliferation detected by invPCR in liver tissues was not correlated with the amount of hepatocytes with LCC or SCC observed in sections of patient liver tissues (Chapter 4 and 5). This suggests that the clonal proliferation that is detected by invPCR is not associated to histological changes in hepatocytes. We therefore hypothesise that hepatocytes with altered histology are not the main sources of the observed clonal proliferation.

7.1.3 Association between the expression of HBV DNA and clonal proliferation of hepatocytes

Other changes in the liver cell population that occur in chronic HBV infection may reveal the mechanism behind the observed clonal proliferation of hepatocytes. For example, heterogeneity in HBsAg and HBcAg expression within the lobules of the liver has been observed in liver tissues of patients with chronic HBV infection (Section 4.3.4). As previously hypothesised by Mason *et al.*, the loss of HBsAg and HBcAg expression may allow the hepatocyte to escape immune-mediated cell death and therefore have a survival advantage and undergo clonal proliferation (Mason, Jilbert *et al.* 2005). We therefore hypothesise that a greater extent of clonal proliferation will be observed in HBsAg-positive hepatocytes compared to HBsAg-negative hepatocytes.

7.1.4 Mechanisms of heterogeneous expression of HBV antigens

The causes of heterogeneity in HBsAg- and HBcAg-expression during chronic HBV infection are a focus of active research. It has been reported that HBV antigen expression may be silenced by CpG methylation of HBV DNA via cellular DNA methyltransferase proteins (Vivekanandan, Thomas *et al.* 2008; Vivekanandan, Daniel *et al.* 2010). Alterations in DNA methylation enzymes may induce methylation of cellular promoters and provide a mechanism for HBV-associated disease progression. It has been reported that altered methylation of cellular promoters occurs early in HCC development and has been observed in peri-tumour, normal-appearing hepatocytes (Su, Zhao *et al.* 2008; Tischoff and Tannapfe 2008).

Mutations in the PreS region of the HBV genome have previously been associated with altered distribution and accumulation of HBsAg in hepatocytes leading to the definition of ground-glass hepatocytes (GGH) (Wang, Wu *et al.* 2003; Wang, Huang *et al.* 2006). HBV

DNA with deletion or mutations in the PreS1 and PreS2 regions of the HBsAg gene have been isolated from type 1 GGH (diffuse, intensely HBsAg-positive hepatocytes) and type 2 GGH (large clusters of HBsAg-positive hepatocytes), respectively (Wang, Wu et al. 2003). Indeed, we observed these two different distributions of HBsAg using immunohistochemistry in liver tissue sections of patients with chronic HBV infection (Section 4.3.4).

We therefore hypothesise that either methylation of HBV promoters or mutations in PreS regions of the HBV genome could contribute to the heterogeneity of HBsAg-positive hepatocytes in the livers of patients with chronic HBV and thus a potential driver for the clonal proliferation of hepatocytes observed by invPCR.

7.1.5 Altered cellular protein expression in hepatocytes

The changes in cellular protein expression that may cause the cessation of HBsAg expression have not been investigated. However, microarray-based studies have detected overall changes in RNA expression in liver tissue from HBV-infected patients compared to uninfected controls (Kim, Ye et al. 2004; Yuen, Yuan et al. 2005; Ieta, Ojima et al. 2007). Mass-spectrometry studies have also shown difference in liver protein expression between HBV-infected patients and uninfected controls (Han, Lee et al. 2010; Le Faouder, Laouirem et al. 2011). Whether some of these genes regulate the expression of HBsAg or contribute to clonal proliferation of hepatocytes is unknown.

We hypothesise that there are changes in the expression of cellular proteins in hepatocyte subpopulations within the liver and that these hepatocytes with altered cellular protein expression have altered expression of HBV antigens and undergo clonal proliferation.

7.1.6 Aims and hypotheses

Overall, we aim to determine the phenotype of the hepatocytes that undergo clonal proliferation, in particular the levels of HBsAg expression. Furthermore, we aim to investigate mechanisms for the development of HBsAg-negative hepatocytes by analysing alterations in the methylation of HBV DNA, deletions in HBV DNA, and cellular protein expression of HBsAg-positive and -negative hepatocyte populations in liver tissue of chronically HBV-infected patients.

We hypothesise that the hepatocytes that have lost the ability to express HBsAg escape immune-mediated cell death, and therefore undergo clonal proliferation. Furthermore, we hypothesise that the cause of the heterogeneity in HBsAg-expression of the chronically HBV-infected hepatocyte population is hepatocyte-mediated through either hypermethylation of HBV DNA by host cell DNA-methyltransferase proteins, or changes in expression of cellular proteins that control susceptibility to HBV infection or replication within the hepatocytes, which in turn affect disease progression.

7.2 *Experimental Outline*

The in-depth analyses described in this Chapter were carried out on ethanol-fixed and paraffin wax-embedded liver tissue sections that were shown to be positive by invPCR for repeated virus-cell DNA junctions in Section 5.6.5, and shown to contain HBsAg-positive and -negative foci in Section 4.3.4: namely liver tissue sections from patients XA, Y2, Y3, Y4, Y5, and Y6. A short summary of clinical information from each of these patients is provided in Table 7.1.

To determine which subsets of cells were detected in invPCR assays outlined in Chapter 5, laser-microdissection (as described in Section 2.20) was used to isolate non-hepatocytes, such as cells that make up fibrotic tissue, central veins and portal tracts. These were first identified in neighbouring serial sections stained with H&E and then the respective areas were microdissected from unstained liver sections mounted on membrane-coated slides. In the same way, hepatocytes with LCC (identified by A/Prof. Andrew Clouston) were isolated. DNA was extracted from the isolated tissues and subjected to invPCR, as described in Section 2.21.

Also, laser-microdissection was used to isolate foci of HBsAg-positive and -negative hepatocytes located by using serial sections from the same ethanol-fixed paraffin-embedded block and stained by immunohistochemistry for HBsAg as a guide. DNA extracts from the hepatocyte foci isolated by laser-microdissection were analysed by invPCR, as described in Section 2.21, to determine whether loss of HBsAg expression is associated with more extensive clonal proliferation. The observed virus-cell DNA junctions were subjected to bioinformatics analysis to determine whether differences in HBsAg-expression could be linked to the site of HBV DNA integration in a particular gene.

Table 7.1. An abridged summary of markers of HBV infection in serum and liver in patients with chronic HBV infection.

Name ¹	HBV Genotype	HBV DNA/BG (qPCR) ²	METAVIR score ³	Serum HBsAg	Serum HBeAg	Liver HBsAg ⁴	Liver HBcAg ⁴	Min 3D clone sizes detected in slide tissue (95%CI) ⁵
Y2 ^A	B	1.7 [#]	2	+	-	+	-	816(146-4615) 157(8-2998) 157(8-2998) 157(8-2998) 146(8-2998) 146(8-2998) 146(8-2998) 146(8-2998) 146(8-2998) 146(8-2998)
Y3 ^A	B	2.4 [#]	2	+	-	+	-	146(8-2998) 146(8-2998) 146(8-2998) 146(8-2998)
Y4 ^A	B	2.9 [#]	1	+	-	+++	-	2998 (976-9782) 444(52-3617) 413(50-3302)
Y5 ^A	C	2.6 [#]	3	+	-	+	-	1301 (295-5692) 816(146-4615) 413(50-3302) 157(8-2998) 157(8-2998) 146(8-2998)
Y6 ^A	B	1.4 [#]	3	+	-	+	-	21046 (7239-59530) 13054 (9340-28934)
XA ^B	C	0.25	4	+	-	++	-	14045 (8478-20475) 192(36-1081)

NHL1 ^C		ND	0	-	-	-	-	ND
NHL2 ^C		ND	0	-	-	-	-	ND
NHL3 ^C		ND	0	-	-	-	-	ND

ND = not detected

■ = Unknown

¹ Patient liver tissues were sourced from the University of Washington (^A), Centenary Institute (^B), and SA Pathology (^C). Full summaries of patient data are shown in Tables 4.1. and 4.2. Full summaries of hepatocyte clones detected in liver tissue sections are shown in Table 5.9.

² Total DNA extracts from tissue sections were analysed by qPCR as described in Section 2.11. A full discussion of the results is shown in Section 4.3.2.

Copy numbers of HBV DNA/cell in these tissues were previously published by Mason *et al.* (2010).

³ Histology was interpreted by Associate Professor Andrew Clouston (University of Queensland). METAVIR scores are out of 4. A full discussion of the results is shown in Section 4.3.1.

⁴ The percentage of HBsAg- and HBcAg-positive hepatocytes were detected by immunohistochemistry as described in Section 2.4. A full discussion of the results is shown in Section 4.3.4. The legend is as follows:

- = 0%; + = <25%; ++ = 25-50%; +++ = 50-75%; ++++ = >75%

⁵ Clone sizes previously detected in DNA extracts of entire liver tissue sections using the *NcoI* invPCR design. Full methods and analysis are shown in Section 5.3.5. Minimum clone sizes in 3 dimensions were calculated as outlined in Section 2.22.

Levels of HBV DNA in hepatocyte foci were quantified by qPCR as described in Section 2.13 to determine whether differences in HBV replication were also associated with differences in HBsAg and HBcAg expression and/or the detection of virus-cell DNA junctions.

The mechanisms involved in the heterogeneous HBsAg-expression were also investigated. Firstly, the PreS regions of HBV DNA extracted from foci of the HBsAg-positive and -negative hepatocytes were sequenced to determine whether PreS mutations were causing changes in HBsAg expression. Also, the methylation state of HBV DNA was analysed by a combined restriction enzyme and qPCR approach, as described in Section 2.24 and optimised in Section 3.2.1.8.

To determine whether changes in cellular protein expression were involved in either clonal proliferation or the expression of HBsAg, HBsAg-positive and -negative foci of hepatocytes were isolated from ethanol-fixed paraffin wax-embedded sections of liver by laser-microdissection. Proteins from isolated cells were digested with trypsin and analysed using High Pressure Liquid Chromatography nano-Electrospray Ionisation Linear Ion Trap Orbitrap Mass Spectrometry (HPLC-nESI-LTQ Orbitrap MS), as described in Section 2.27 and optimised in Section 3.2.2.2. This technique was used to identify cellular proteins that may be expressed uniquely in foci of either HBsAg-positive or HBsAg-negative hepatocytes.

Imaging mass-spectrometry was used to determine the distribution of peptides derived from trypsin-digested cellular proteins across a section of liver tissue. Sections of liver tissue were analysed as described in Section 2.26 and optimised in Section 3.2.2.1. The mass peaks that were distributed across the tissue were identified by matching up mass charge ratios (m/z) to those of peptides sequences identified by HPLC-nESI-LTQ MS, as described in Section 2.29.

7.3 Results

7.3.1 Isolation of liver cell subpopulations by laser-microdissection

Liver cell populations were isolated as described in Section 2.20 by laser-microdissection of ethanol-fixed paraffin wax-embedded liver tissue sections from 7 patients previously shown to be positive for integrated HBV DNA by invPCR (Table 7.1).

Hepatocyte foci were identified by first visualising HBsAg-positive and -negative hepatocyte foci in ethanol-fixed, paraffin wax-embedded sections immunostained for HBsAg (as described in Section 4.3.4). The corresponding areas were then identified in serial sections of liver tissue mounted on polyethylene naphthalate membrane-coated glass slides and isolated by laser-microdissection. The hepatocyte foci that were isolated for each tissue section are shown in Figure 7.1. Total DNA was extracted from the hepatocyte foci as previously described in Section 3.2.1.2.4.

Non-hepatocyte cells that remained from the isolation of hepatocytes, such as areas of inflammation, fibrotic tissue, central veins and portal tracts, were also isolated and total DNA was extracted from these isolated cells for further analysis.

7.3.2 InvPCR and qPCR analysis of non-hepatocyte cell populations for HBV DNA

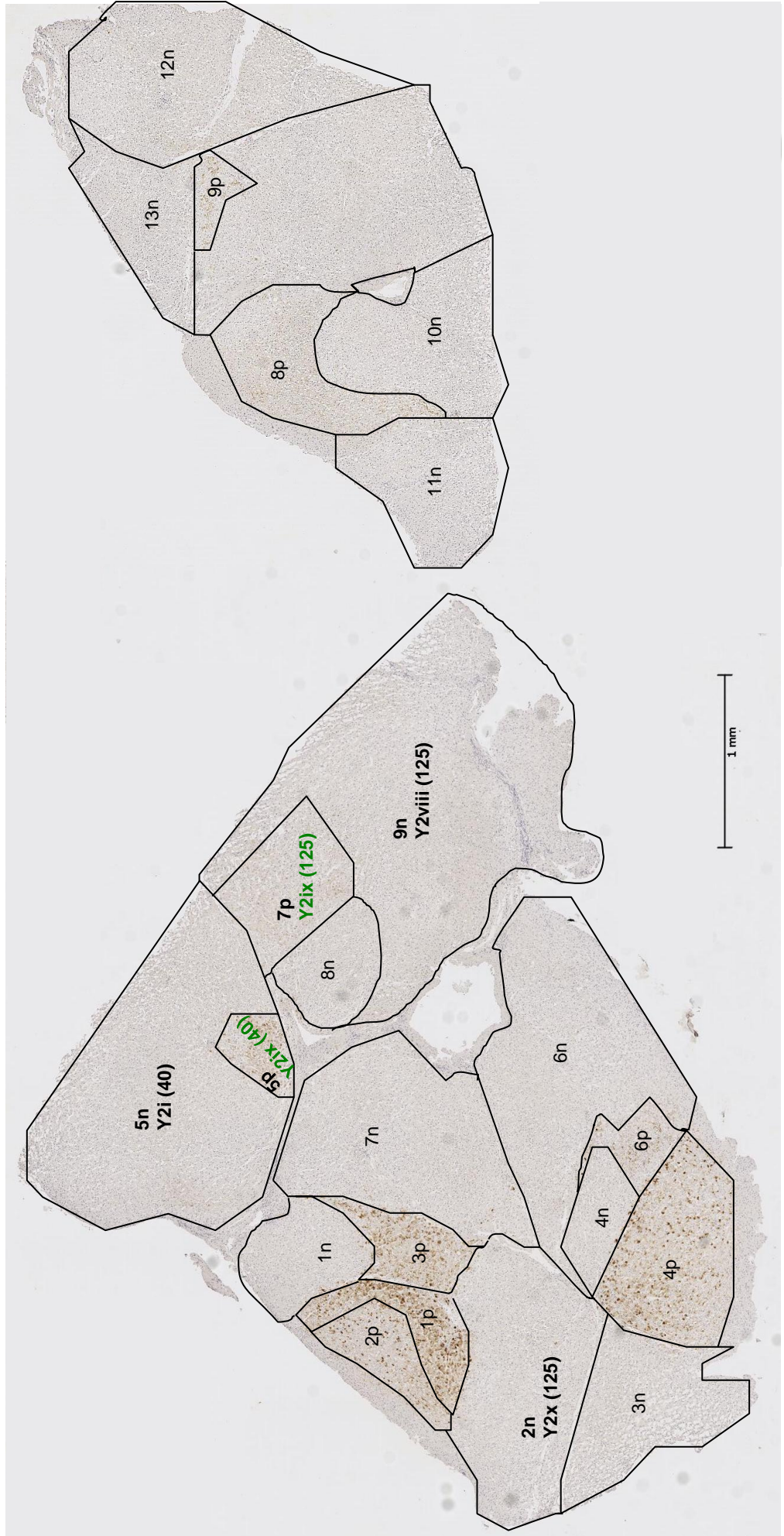
DNA from the non-hepatocyte cell populations from each patient liver tissue was isolated and subjected to invPCR to determine whether those cells contained detectable virus-cell DNA junctions. Total DNA extracts of these isolated cell populations were analysed by invPCR using both *NcoI* and *DpnI* designs as described in Section 2.21. No virus-cell DNA junctions were detected in these extracts.

DNA extracted from the non-hepatocyte cells were also analysed by qPCR to quantify HBV DNA present in the non-hepatocyte cell population compared to hepatocyte populations of the same section. A qPCR assay detecting levels of total HBV DNA normalised to the cellular beta-globin gene (BG) was performed, as described in Section 2.13.

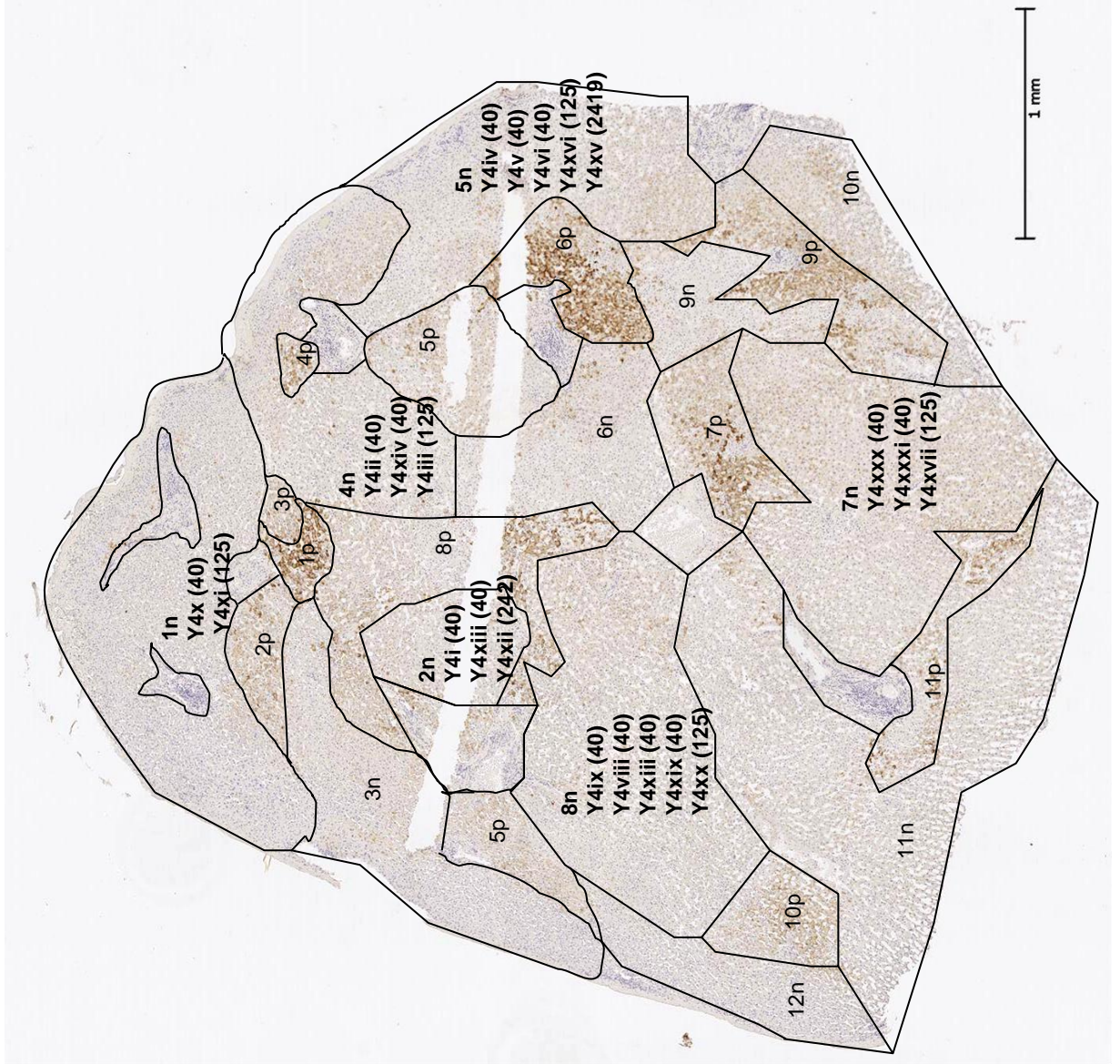
As shown in the standard curves and melt curves of Figure 7.2, the qPCR assay was sensitive to ≤ 10 copies of either BG (Figure 7.2A and 7.2C) or HBV DNA (Figure 7.2B and 7.2D) per reaction and specifically amplified single products. The qPCR results are summarised in Figure 7.3.

DNA extracts from areas containing only hepatocytes contained on average 1.4 copies of HBV DNA per copy of BG, while DNA extracts from areas containing non-hepatocytes contained on average 0.19 copies of HBV DNA per copy of BG. After analysis using a one-

A



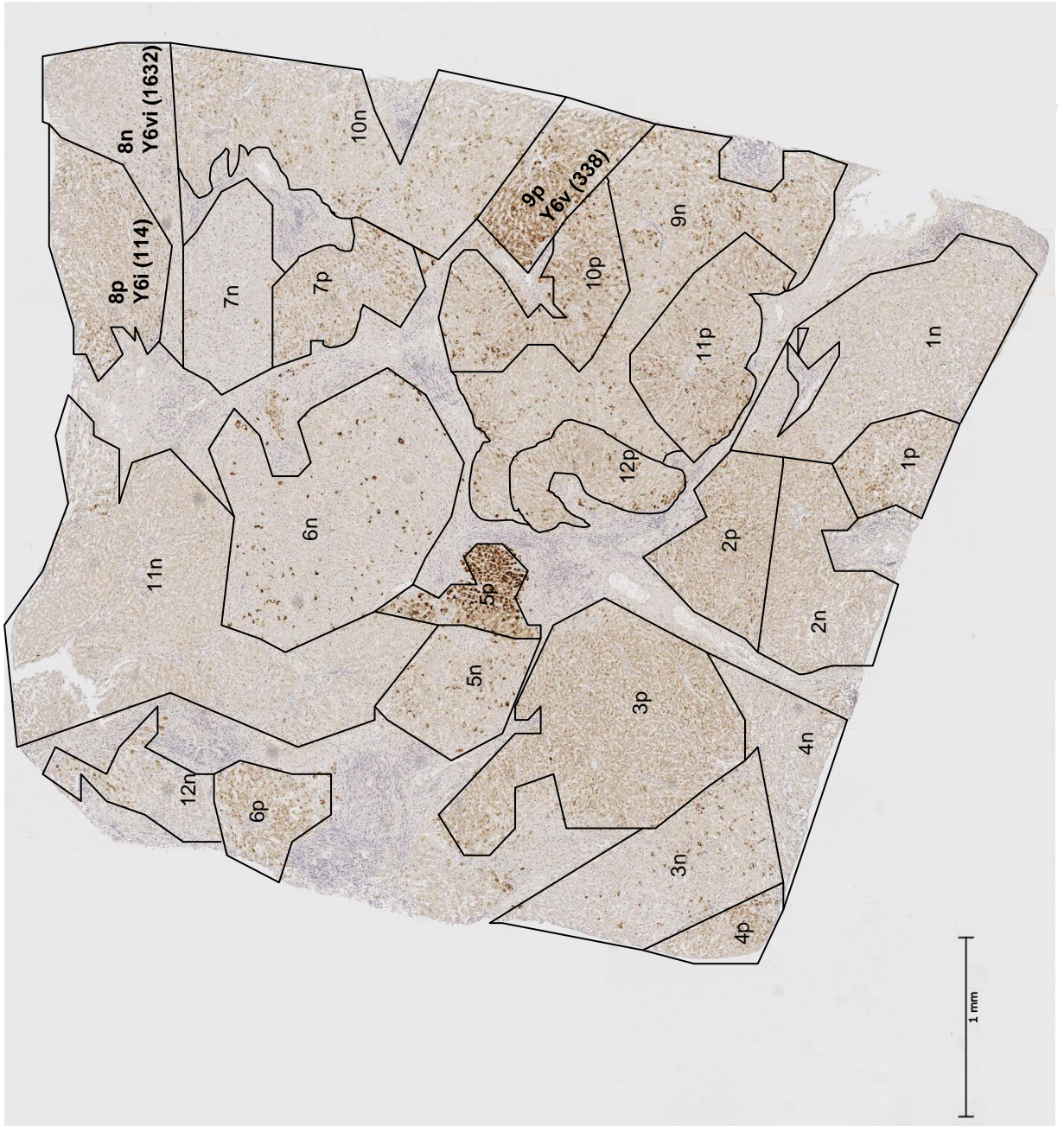
B





C

D



E

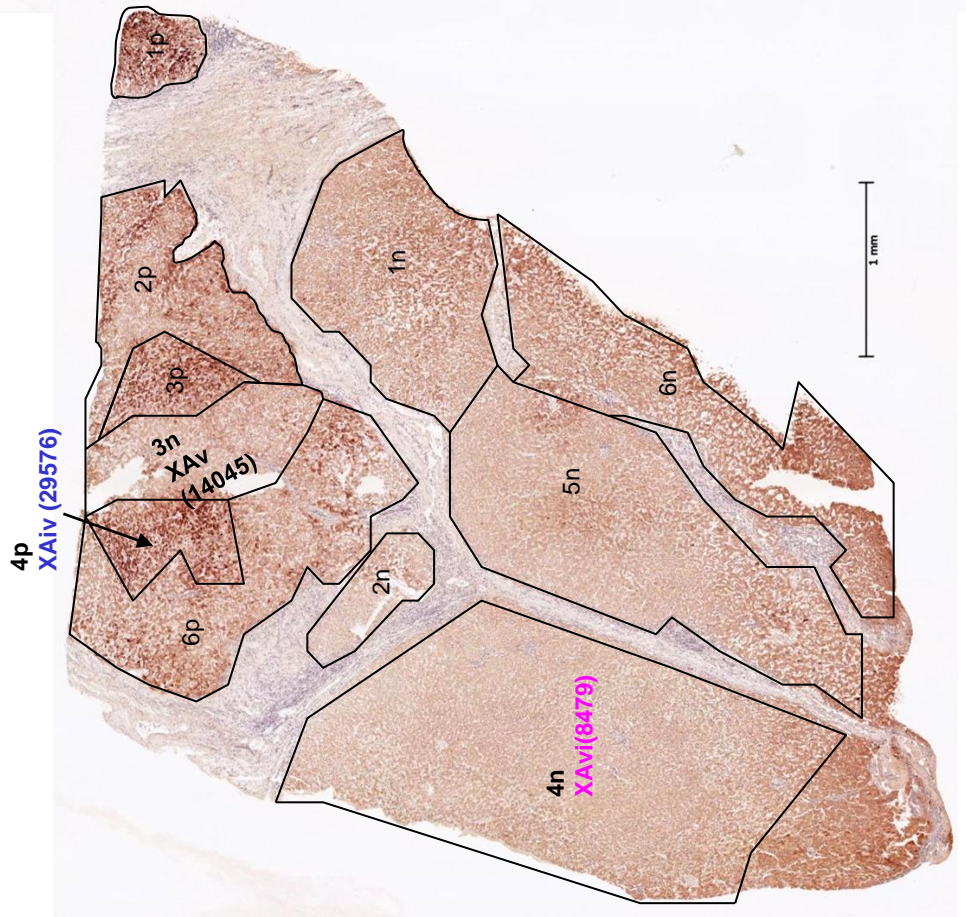
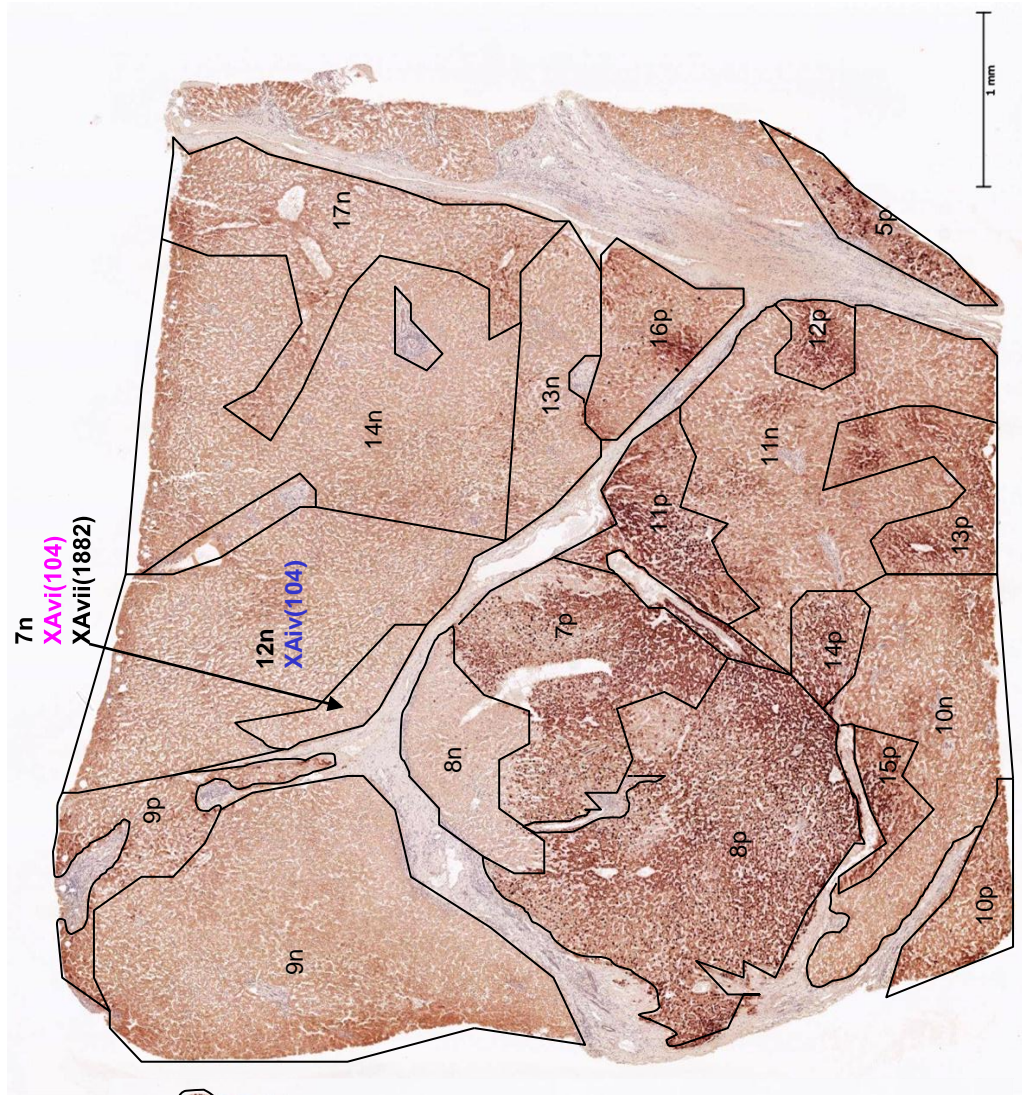
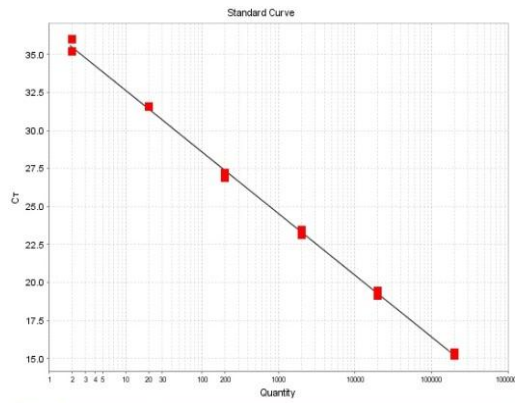


Figure 7.1. Virus-cell junctions found in hepatocyte foci isolated by laser microdissection of tissue sections from Patient Y2 (7.1A), Y4 (7.1B), Y5 (7.1C), Y6 (7.1D) and XA (7.1E). Laser-microdissected foci of hepatocytes were superimposed on the HBsAg-stained sections, generated as previously discussed in Section 4.3.4, and outlined in black. Foci were classified as HBsAg-positive (suffixed with “p”) or HBsAg-negative (suffixed with “n”) and isolated by laser microdissection, as described in Section 2.20. DNA was extracted as described in Section 2.21 and virus-cell DNA junctions were detected in DNA extracts using invPCR as described in Section 2.22. The foci in which virus-cell junctions were detected by invPCR are labelled in bolded letters along with the unique identifier for the virus-cell junction that was detected (Appendix 9.8). Virus-cell junctions detected in multiple laser-dissected foci are highlighted in the same colour.

Beta globin

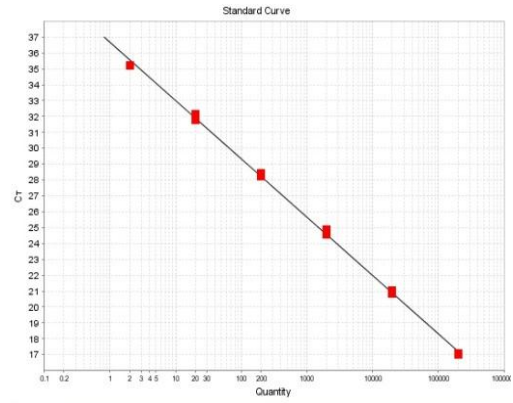
A



$R^2 > 0.999$
Efficiency = 77%

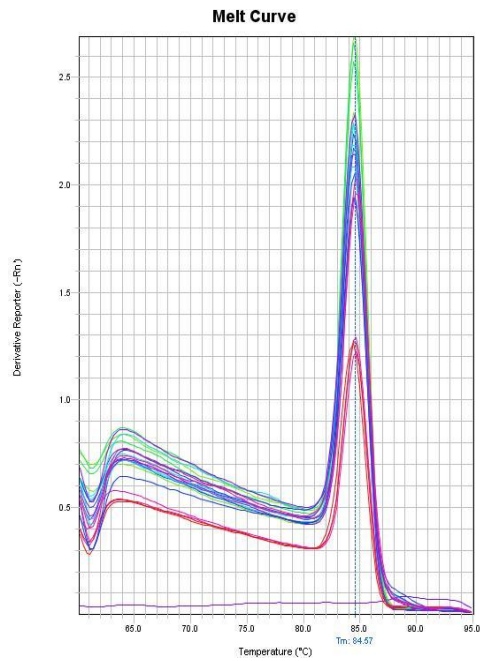
HBV DNA

B



$R^2 > 0.999$
Efficiency = 87%

C



D

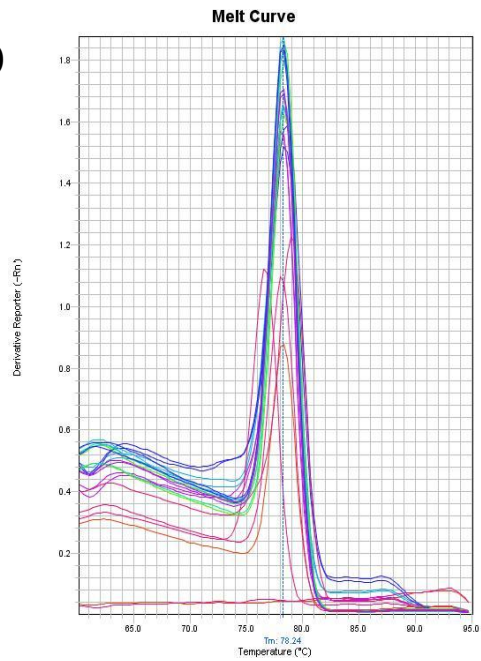


Figure 7.2. Standard curves (A and B) and melt curves (C and D) of qPCR assays detecting copy numbers of BG (A and C) and total HBV DNA (B and D) in DNA extracts of non-hepatocyte cells isolated by laser microdissection from liver tissue sections of HCC patients without cirrhosis. qPCR was carried out as described in Section 2.13 to determine the copy number of HBV DNA in DNA extracts of cells isolated by laser microdissection, as described in Section 2.20. The sequence of the primers used for amplifying BG and HBV DNA are listed in Table 2.1. The BG gene was subcloned into a pBlueBac vector, as described in Section 2.13.1, to form pBG. pBG and the HBV DNA-containing plasmid pBB4.5HBV1.3 were used to produce the standard curves, shown in Figures 7.2A & 7.2B respectively. Plasmid DNA standards were diluted from 10^1 - 10^8 copies per reaction. The melt curves in Figure 7.2C and 7.2D show that only single products were amplified in all samples.

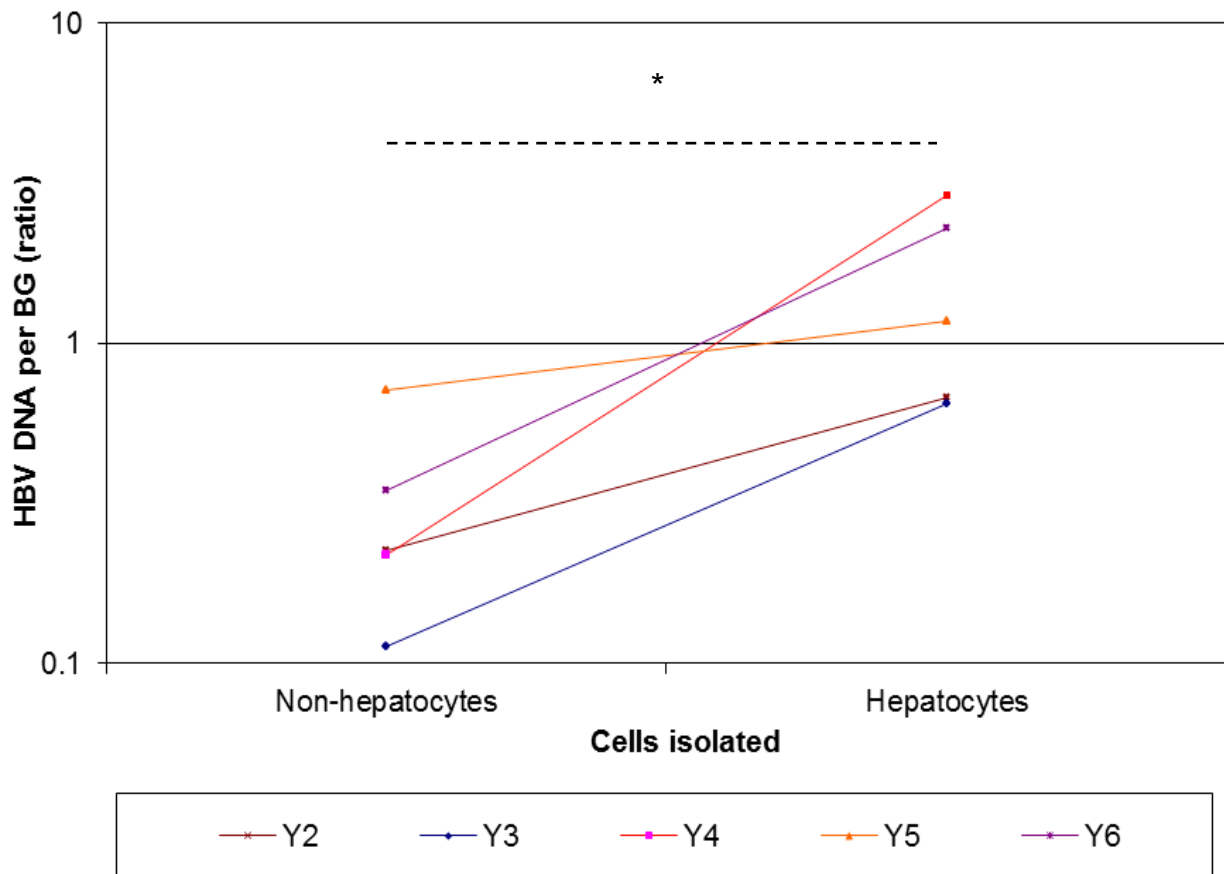


Figure 7.3. Significantly lower levels of HBV DNA were detected by qPCR in DNA extracts of non-hepatocyte cells isolated by laser microdissection from liver tissue sections. Hepatocytes and non-hepatocytes were isolated by laser microdissection as described in Section 2.20 in liver tissue sections of Patients Y2-Y6. Total DNA was isolated (as described in Section 2.21) from the two pools from each patient and levels of HBV DNA were quantified in each extract by qPCR as described in Section 2.13. HBV DNA copy numbers were normalised to copies of BG, a single-copy cellular gene. Significantly greater copies of HBV DNA were detected in hepatocytes compared to paired non-hepatocytes from the same tissues after analysis with a paired student's t-test ($p < 0.05$).

tailed paired t-test, non-hepatocytes were shown to contain significantly lower levels of HBV DNA compared to hepatocytes isolated from the same tissue with a p value of <0.05. It is possible that the small amounts of circulating HBV virions within central veins contributed to the presence of HBV DNA in the DNA extracts of the non-hepatocyte cell populations.

7.3.3 Virus-cell DNA junctions were not frequently detected in hepatocytes with LCC

Foci of hepatocytes with LCC were identified by A/Prof. Andrew Clouston in H&E-stained liver sections from 3 out of the 7 liver tissue sections, namely tissues from Patients Y3, Y5 and Y6 (Figure 4.4). These foci of hepatocytes with LCC were then isolated from serial sections by laser microdissection. Total DNA was extracted from the isolated cells and analysed by invPCR. Only a single unique virus-cell DNA junction (Y5 ix) was detected in DNA extracts of 15 foci of hepatocytes with LCC. However, in hepatocytes without LCC, a total of 7 virus-cell DNA junctions were detected in hepatocyte foci isolated from the three liver tissue sections. This indicates that the majority of the observed virus-cell DNA junctions were not present in hepatocytes with LCC.

7.3.4 Virus-cell DNA junctions were 2 times more likely to be detected in foci of HBsAg-negative hepatocytes compared to foci of HBsAg-positive hepatocytes

To determine whether differences in levels of HBsAg-expression was indicative of different rates of clonal proliferation, individual HBsAg-positive and -negative foci of hepatocytes were identified and isolated by laser-microdissection from patients Y2, Y4, Y5, Y6 and XA. Each focus isolated by laser-microdissection is shown in Figure 7.1 outlined in black.

Total DNA was extracted from each focus isolated by laser-microdissection and was analysed by invPCR. The number of foci isolated from each liver tissue and the virus-cell DNA junctions that were detected by invPCR are summarised in Table 7.2 and are listed in full in Appendix 9.8. The foci in which virus-cell junctions were detected by invPCR are labelled in Figure 7.1 using bold letters along with the unique identifier for the virus-cell junction that was detected (Appendix 9.8). Virus-cell junctions detected in multiple foci are highlighted in the same colour.

Table 7.2. Virus-cell DNA junctions were detected in foci of HBsAg-positive and -negative hepatocytes isolated by laser microdissection of liver tissue sections from patients with chronic HBV infection.

Patient ²	HBsAg-positive foci of hepatocytes ¹			HBsAg-negative foci of hepatocytes ¹		
	Foci isolated	Cells isolated ³	Minimum 3D clone sizes (95% CI) ⁴	Foci isolated	Cells isolated ³	Minimum 3D clone sizes (95% CI) ⁴
XA	10	7052	iv) 29575 (6067-59530)	13	14017	v) 14045 (18243-20476) vi) 8479 (3617-19349) vii) 1883 (560-6450) iv) 104 (6-2145)
Y2	9	1996	viii) 40 (2-739)*	14	9085	i) 2145 (268-18243) ix) 125 (15-956) viii) 125 (15-956) x) 40 (2-739)*
Y3	13	3319	ND	27	16221	ND
Y4	11	1652	ND	11	6248	xv) 2419 (875-7239) xii) 242 (43-1301) x) 146 (8-2998) xiii) 146 (8-2998) xix) 146 (8-2998) xx) 146 (8-2998) xi) 125 (15-956) xvi) 125 (15-956) xvii) 125 (15-956) iii) 125 (15-956) i) 40 (2-739)* ii) 40 (2-739)* iv) 40 (2-739)* v) 40 (2-739)* ix) 40 (2-739)* vi) 40 (2-739)* viii) 40 (2-739)* xiv) 40 (2-739)* xxx) 40 (2-739)* xxxi) 40 (2-739)*
Y5	11	7437	iii) 666 (125-3617) i) 413 (51-3617)	10	5100	vi) 1301 (295-5692) ix) 815 (146-4615) ii) 114 (6-2144)
Y6	11	3653	v) 338 (42-2703) i) 114 (6-2144)	12	8064	vi) 1632 (460-6450)

¹ InvPCR was used to detect virus-cell DNA junctions in DNA extracts of HBsAg-positive and -negative foci of hepatocytes isolated by laser microdissection of liver tissue sections from chronic HBV patients as described in Sections 2.20-22. HBsAg status was determined by immunohistochemistry of serial liver tissue sections as described in Section 2.4.

² Only patients previously found to contain both repeated virus-cell DNA junctions in slide-mounted tissue and HBsAg-positive and -negative foci were used in this analysis.

³ The area of micro-dissected foci were measured by ImageJ and the approximate number of nuclei isolated was estimated by using the calculated nuclear concentration from Section 3.2.1.6 H. Both *NcoI* and *DpnII* invPCR designs were used in patients XA, Y2, Y4 and Y6 and results were pooled. Patient extracts in which no virus-cell DNA junctions were detected are represented by ND.

⁴ Clone sizes were determined by dilution of inverted DNA templates in conjunction with the MPN method. Each virus-cell junction was given a unique identifier (i, ii, iii, and so on) and can be matched up with DNA sequences and other details fully listed in Appendix 9.8.

In cases where the same virus-cell DNA junction was detected in DNA extracts of both whole tissue sections and hepatocytes isolated by laser microdissection, the larger of the two measurements was taken. Three dimensional clone sizes were calculated as previously described in Section 2.22, using the calculated nuclear concentration from Section 3.2.1.6. Virus-cell DNA junctions detected only once are indicated by “*” and were not included in further analyses.

No virus-cell DNA junctions were detected in HBsAg-positive or -negative hepatocyte foci isolated by laser-microdissection from tissue of Patient Y3.

The majority of virus-cell DNA junctions were detected in foci of HBsAg-negative hepatocytes, as shown in Figure 7.4. While foci of HBsAg-positive hepatocytes comprise of 30% of all the cells isolated from the liver sections, the number of virus-cell DNA junctions found in foci of HBsAg-positive hepatocytes is only 15%. In other words, virus-cell DNA junctions were detected twice as often in foci of HBsAg-negative hepatocytes compared to foci of HBsAg-positive hepatocytes after normalisation of the number of cells analysed.

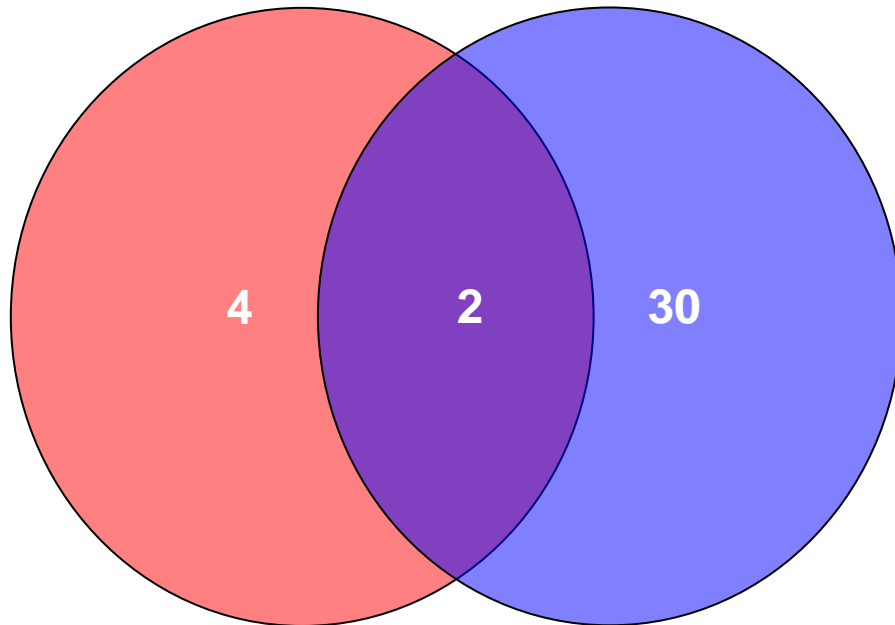
Hepatocyte clone sizes were calculated as outlined in Section 2.22. No significant difference ($p > 0.05$, by one-tailed student's t-test) in clone size was observed in hepatocyte clones found in the foci of HBsAg-positive hepatocytes as compared to those found in foci of HBsAg-negative hepatocytes (Figure 7.5). This suggests that the decreased levels of HBsAg expression do not confer a survival advantage that leads to larger hepatocyte clones.

7.3.5 Detected HBV DNA integrations display no preference for particular genes or proximity to structural and matrix associated regions (S/MAR)

The ends of the integrated HBV DNA detected using invPCR were mapped and graphed in Figure 7.6. Greater than 80% of virus-cell DNA junctions detected in tissue isolated by laser microdissection occurred within 100 bp from the expected right-hand junction at nt 1832, numbered with respect to the *EcoRI* site on the HBV genome (Genbank Accession #AB241115). This is consistent with the virus-cell DNA junctions detected in liver tissue fragments and liver tissue sections that have been described in Section 5.3.7 and in previous literature (Mason, Low et al. 2009; Mason, Liu et al. 2010), where $>70\%$ of all integrations of HBV DNA occurred <50 bp from nt 1832 with respect to the *EcoRI* site on the HBV genome.

As shown in Appendix 9.8, each virus-cell DNA junction was aligned to the human genome using BLAST analysis (Altschul, Gish et al. 1990) and the site of integration of HBV DNA into the host-cell chromosome was determined. None of the cellular ORF into which integration had occurred was differentially expressed between normal human liver and HCC tissue, according to the online Unigene expression database located at <http://www.ncbi.nlm.nih.gov/unigene> (Boguski and Schuler 1995). This is consistent with the

Virus-cell DNA junctions found in
HBsAg-negative hepatocytes



Virus-cell DNA junctions found in
HBsAg-positive hepatocytes

Figure 7.4. The majority of virus-cell DNA junctions were detected exclusively in HBsAg-positive or -negative foci of hepatocytes. The number of unique virus-cell DNA junctions detected by invPCR in foci of HBsAg-positive (red), -negative (blue) or in both HBsAg-positive and HBsAg-negative (purple) hepatocytes isolated by laser microdissection are shown. Only 2 unique virus-cell DNA junctions were detected in both HBsAg-positive and -negative hepatocyte foci.

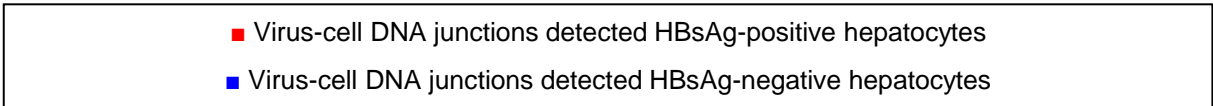
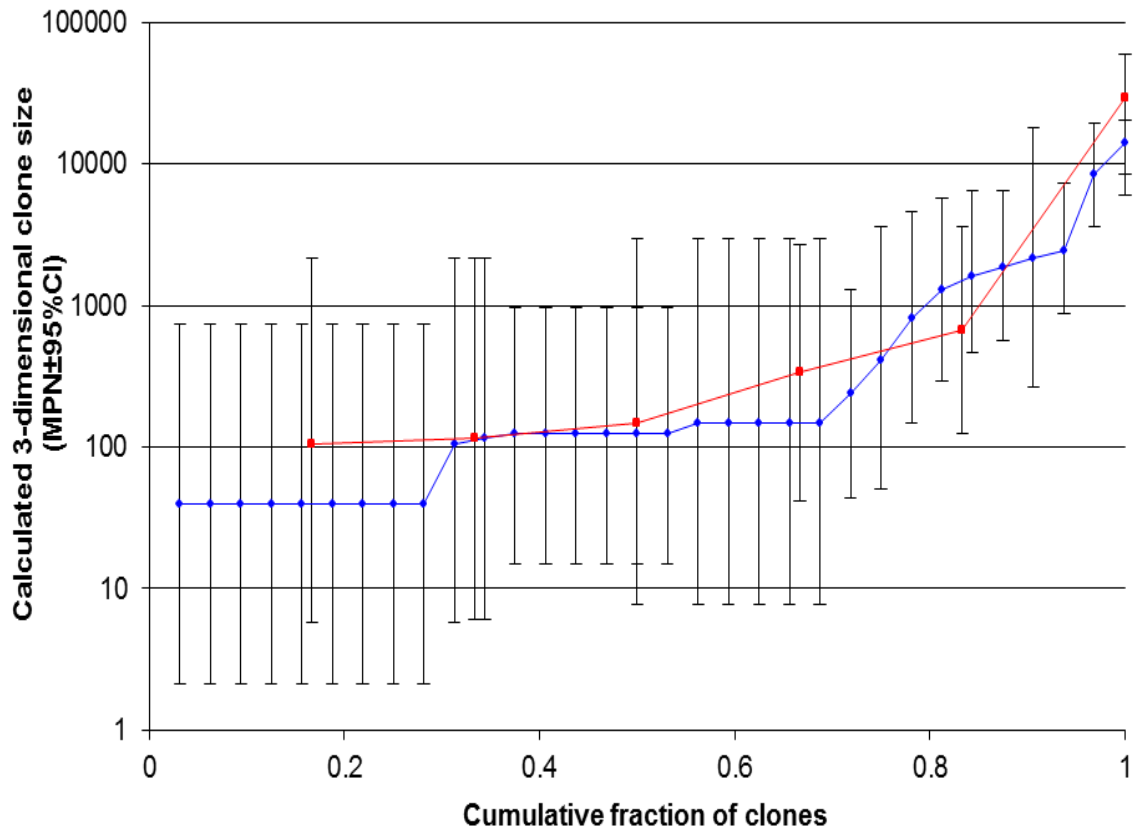


Figure 7.5. Sizes of hepatocyte clones detected in HBsAg-positive foci of hepatocytes were not significantly different from those found in foci of HBsAg-negative hepatocytes. Foci of hepatocytes that were HBsAg-positive or -negative were isolated by laser microdissection (as shown in Figure 7.1) and DNA extracted from the isolated foci were analysed by invPCR as described in Sections 2.21-2.22. Three dimensional clone sizes were estimated by MPN statistical analysis as mentioned in Section 2.14, in conjunction with 3-dimensional size estimation as mentioned in Section 2.22.. No significant differences in the calculated 3-dimensional clone sizes represented by virus-cell DNA junctions detected in HBsAg-positive (red) or -negative (blue) hepatocytes ($p > 0.05$, Student's t-test of square root transformed values). The clones that are displayed in the graph are summarised by patient extract in Table 7.2 and fully listed in Appendix 9.8.

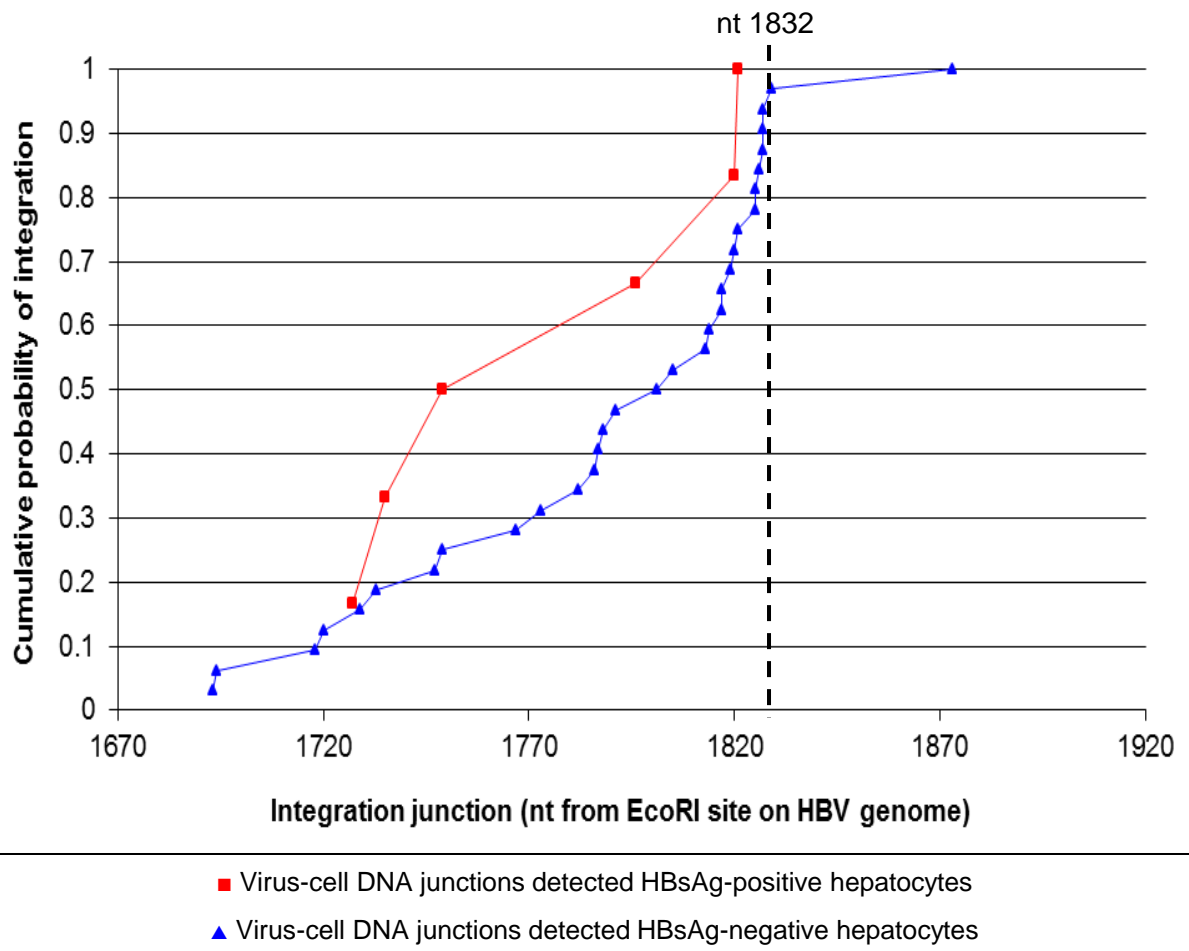


Figure 7.6. No significant difference in integration site with respect to HBV DNA genome was observed between virus-cell DNA junctions found in HBsAg-positive and -negative hepatocytes. Virus-cell DNA junctions detected by invPCR as described in Section 2.22 on DNA extracted from hepatocyte foci isolated by laser microdissection were aligned with the HBV genome (Genbank accession #AB241115). The cumulative probability of integration was plotted against the site of integration with respect to the HBV DNA sequence. In both cases, the integration site rose and clustered before the expected right hand end of the dsDNA form of HBV located at nt 1832. No significant difference in site preferences were observed between the virus-cell DNA junctions detected in HBsAg-positive hepatocytes compared to those detected in HBsAg-negative hepatocytes (analysed with Student's t-test, $p > 0.05$)

hypothesis that insertional mutagenesis of cellular oncogenes due to integration of HBV DNA was not a contributing factor in clonal proliferation of hepatocytes.

The distance of each virus-cell DNA junction from S/MARs was also determined using MAR-Finder software (<http://genomecluster.secs.oakland.edu/MarWiz/>) and is graphed in Figure 7.7. No significant preference for S/MAR proximity was observed by Grubb's outlier test ($p > 0.05$). Furthermore, no significant difference in calculated clone size was associated with proximity to S/MARs, as shown in Figure 7.8, or integration into particular chromosomes, as shown in Figure 7.9. These observations agree with the properties of virus-cell DNA junctions detected in liver tissue fragments and liver tissue sections, as described in Sections 5.3.8-5.3.9.

Furthermore, as shown in Appendix 9.8 and Figures 7.6-7.7, there were no significant differences between the virus-cell DNA junctions found in foci of HBsAg-positive and -negative hepatocytes with respect to ORF insertion, integration site relative to the HBV genome, or proximity to S/MAR. This data is consistent with the virus-cell DNA junctions observed in Chapter 5 and suggests that integration of HBV DNA into the host cell chromosome occurs randomly, as supported by previous studies (Bill and Summers 2004; Mason, Low et al. 2009; Mason, Liu et al. 2010).

As shown in Figure 7.10, a significant albeit small preference (analysed by Grubb's outlier test, $p > 0.05$) for chromosome 4 was observed in virus-cell DNA junctions detected in HBsAg-positive foci of hepatocytes. Virus-cell DNA junctions found in chromosome 4 had been formed by integration into widely differing areas of the chromosome (Appendix 9.8). This suggested that the overrepresented HBV DNA integration into chromosome 4 was a chance event, rather than a DNA-break hotspot. This, added to the small number of virus-cell DNA junctions found in HBsAg-positive hepatocytes, indicates that any significant preference should be confirmed with greater numbers of samples.

7.3.6 HBsAg-positive hepatocytes contain two times more HBV DNA genomes per cell compared to HBsAg-negative hepatocytes

Differing levels of HBV DNA replication could lead to the observed differences in detection of virus-cell DNA junctions between foci of HBsAg-positive and -negative hepatocytes. For

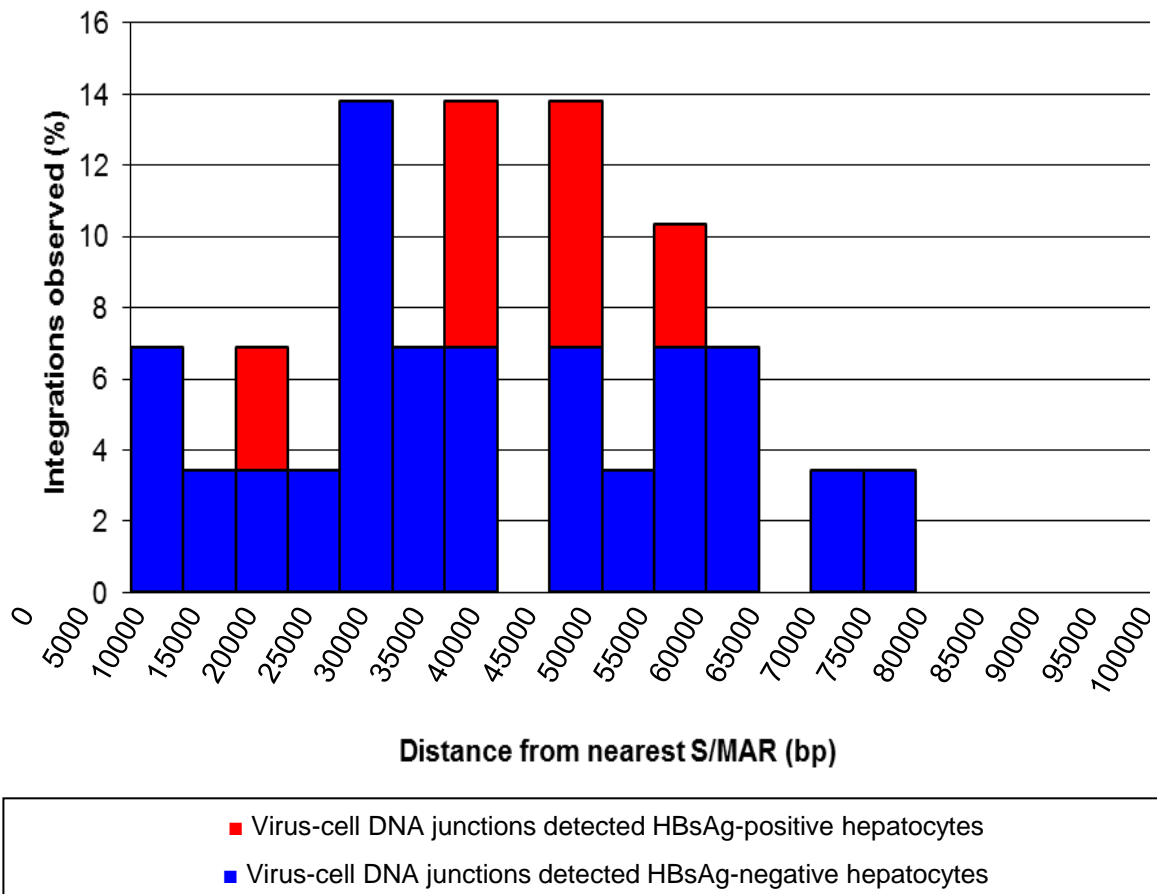
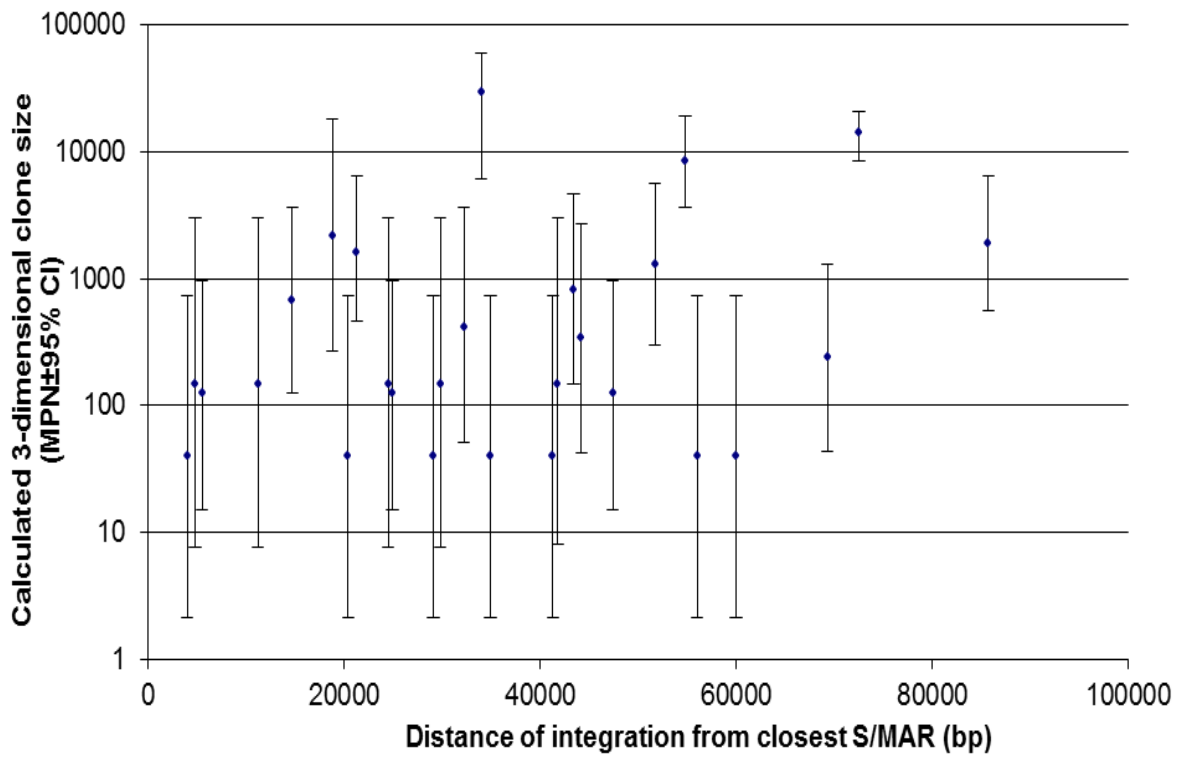
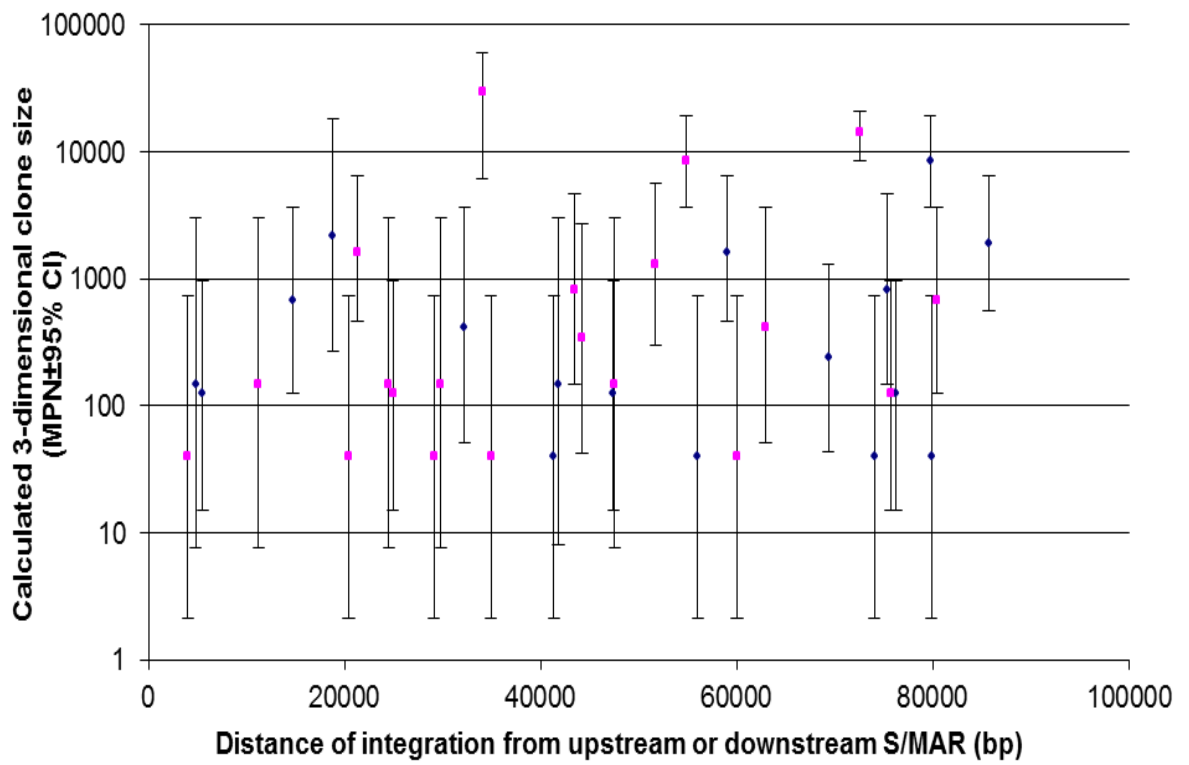


Figure 7.7. No difference in proximity of integration sites to S/MARs was observed between virus-cell DNA junctions found in foci of HBsAg-positive and -negative hepatocytes. Expected cellular sequences 100 kb either side of each virus-cell DNA junction detected was examined for S/MAR as previously described in Section 5.3.9 using freely-available, online program MAR-Wiz (<http://www.futuresoft.org/MarFinder/>). The minimum distance from a predicted S/MAR was recorded and graphed above. Analysis of virus-cell DNA junctions detected by invPCR in DNA extracted from foci of HBsAg-positive (red) and -negative (blue) hepatocytes showed no preference for particular distances from S/MAR ($p > 0.05$ by Grubb's outlier test). No significant difference was seen in the distances of the integration events from S/MAR between virus-cell DNA junctions detected in HBsAg-positive hepatocytes compared to those observed in HBsAg-negative hepatocytes ($p > 0.05$ by Student's t-test).

A**B**

■ Upstream of virus-cell DNA junction

◆ Downstream of virus-cell DNA junction

Figure 7.8. No association was found between clone size and proximity of virus-cell DNA junctions to S/MARs in virus-cell DNA junctions observed in foci of hepatocytes isolated by laser microdissection of liver tissue section from Patients Y2-Y6 and XA. Minimum clone sizes in three dimensions represented by each virus-cell DNA junction was calculated by MPN statistical analysis as described in Section 2.14, in conjunction with 3-dimensional size estimation as described in Section 2.22. Calculated clone sizes were graphed against the distance of each virus-cell DNA junction from the closest predicted S/MAR (7.8A) or both upstream and downstream distances from S/MARs (7.8B). For the latter, distances greater than 100 kb were not included. No obvious association in either case was seen.

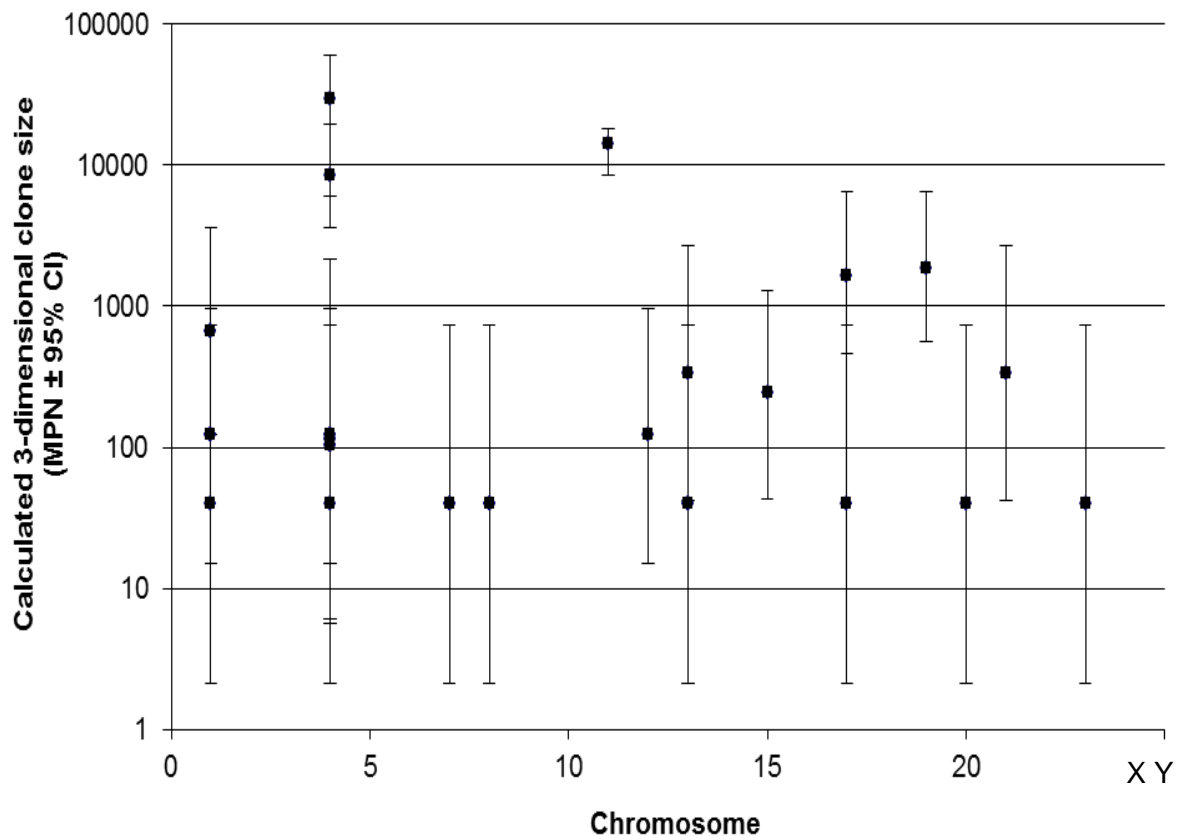


Figure 7.9. No association was found between clone size and integrated chromosome in virus-cell DNA junctions detected by invPCR in foci of hepatocytes isolated by laser microdissection from cirrhotic and non-cirrhotic patients. Virus-cell DNA junctions were aligned to the human genome using BLAST analysis to determine the chromosome into which DNA integration had occurred. The integration sites were graphed against calculated clone sizes in three dimensions as determined by MPN statistical analysis as mentioned in Section 2.14, in conjunction with 3-dimensional size estimation as mentioned in Section 2.22. No association between clone size and distance from S/MAR was observed.

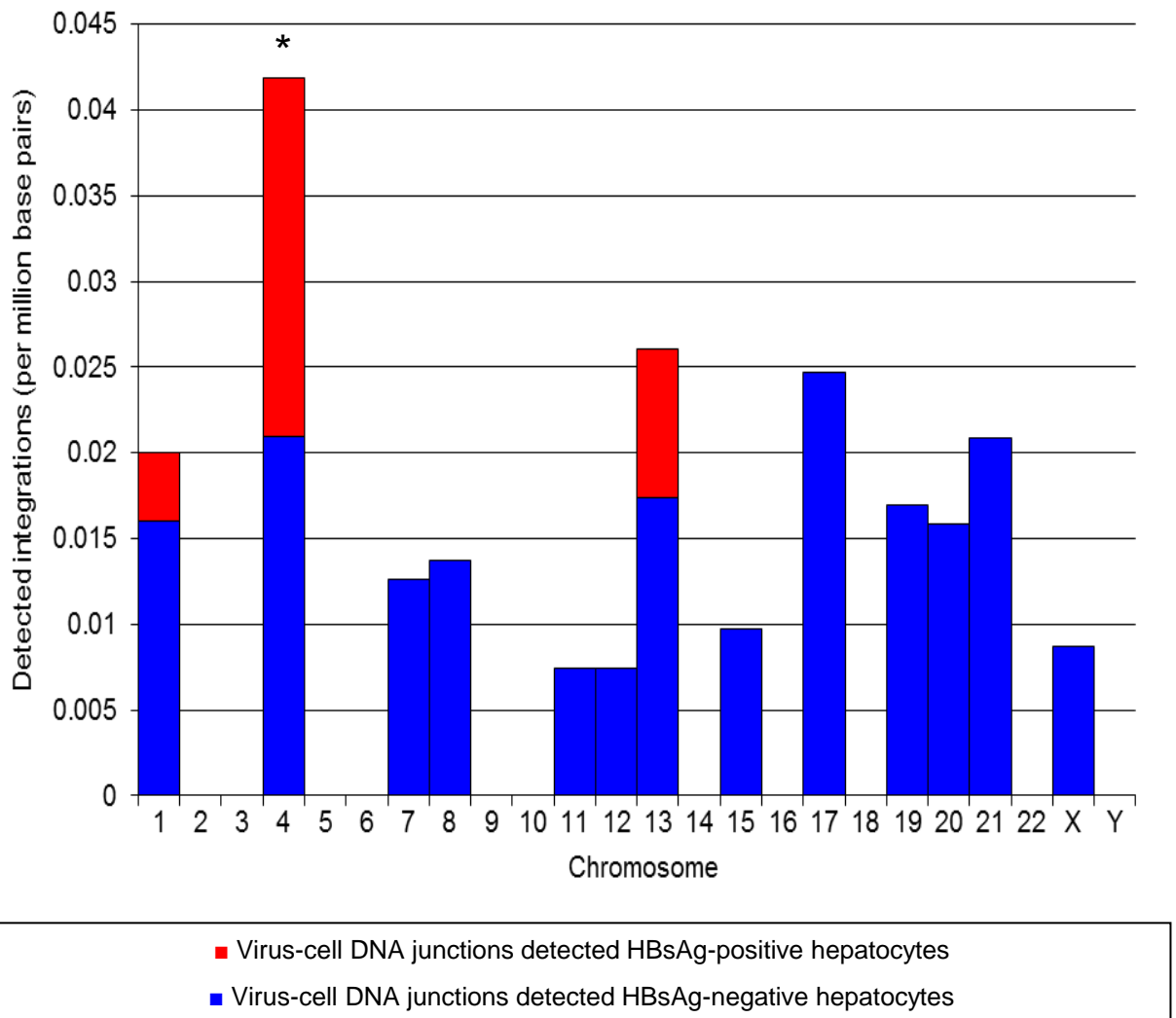


Figure 7.10. Integration into chromosome 4 was a significant outlier in virus-cell DNA junctions detected in HBsAg-positive hepatocytes. Virus-cell DNA junctions detected by invPCR of DNA extracted from foci of hepatocytes positive (blue) or negative (red) for HBsAg were aligned with the human genome using BLAST analysis as described in Section 2.20-2.22. The frequency of detectable integrations into each chromosome was normalised to the length of each chromosome and analysed by Grubb’s outlier test.

Assuming a normal distribution, the frequency of integration in all chromosomes did not significantly differ from each other in virus-cell DNA junctions found in HBsAg-negative hepatocytes ($p > 0.05$ by Grubb’s outlier test). However, in virus-cell DNA junctions found in HBsAg-positive hepatocytes, integrations occurred with significantly greater frequency in chromosome 4 ($*p < 0.05$ by Grubb’s outlier test).

example, if HBsAg-positive hepatocytes have lower levels of HBV RI DNA, then lower levels of dsDNA would be present and therefore lower levels of HBV DNA integration would be likely to occur. This would explain the fewer detectable virus-cell DNA junctions detected in HBsAg-positive hepatocytes. Therefore, the total HBV DNA levels in foci of HBsAg-positive and -negative hepatocytes were determined by PCR and qPCR.

Furthermore, differences in levels of HBV DNA replication between foci of HBsAg-positive and -negative hepatocytes may give an indication of the mechanism behind the heterogeneity of HBsAg expression. If levels of HBV DNA are lower in HBsAg-positive hepatocytes, then the HBsAg expression may be due to integrated, non-replication competent HBV DNA. If levels of HBV DNA are higher in HBsAg-positive hepatocytes, then this suggests that some hepatocytes are truly more susceptible to infection or HBV-replication than others.

To determine the presence of HBV DNA in foci of HBsAg-positive and -negative hepatocytes, total DNA was extracted from foci of hepatocytes isolated by laser-microdissection as described in Section 2.20 and analysed using PCR specific for HBV DNA, as described in Section 2.12. HBV DNA could be amplified by PCR in DNA extracts from all isolated foci, as shown in Figure 7.11.

Also, paired foci of HBsAg-positive hepatocytes and neighbouring HBsAg-negative hepatocytes were isolated by laser-microdissection. The paired samples are summarised in Table 7.3. Total DNA was extracted from the dissected tissues and the levels of HBV DNA (normalised to BG gene copy number) were quantified by qPCR. As shown in the standard and melt curves of Figure 7.12, the qPCR assay was sensitive to ≤ 10 copies of either HBV DNA or BG per reaction and specifically amplified single products. As shown in Figure 7.13, HBV levels were significantly higher ($p < 0.05$) in foci of HBsAg-positive hepatocytes compared to paired foci of HBsAg-negative hepatocytes by an average of 2-fold. This suggests that HBsAg-positive hepatocytes are more susceptible to HBV infection or replication than HBsAg-negative hepatocytes.

The mechanism of this decreased virus replication in HBsAg-negative hepatocytes is unknown, but could be due to: 1) virus-specific factors, such as mutations or epigenetic alterations in the HBV genome; or 2) host-specific factors, such as the loss or downregulation

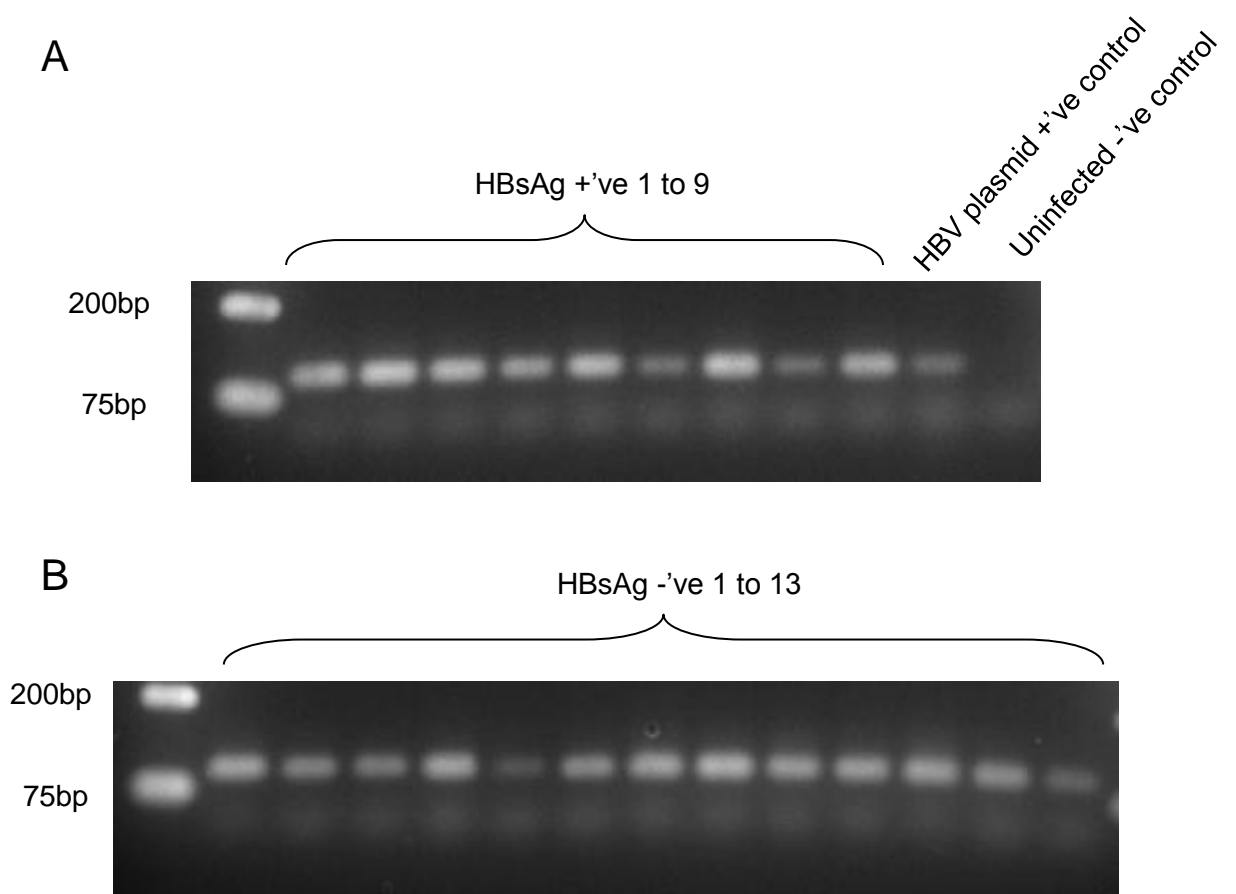


Figure 7.11. HBV DNA was detected in all hepatocyte foci. HBV DNA was detected by PCR using the primers described in Table 2.3 in DNA extracts from HBsAg-positive (7.11A) and -negative (7.11B) hepatocyte foci of patient XA. The isolated hepatocyte foci are shown in Figure 7.1E. PCR conditions used were as described in Section 2.12.2. The expected 99bp product was amplified from DNA extracts from all foci.

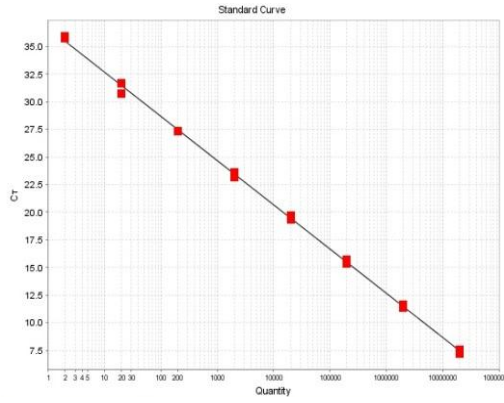
Table 7.3. Paired foci of hepatocytes analysed for number of HBV DNA copies per cell by qPCR

Paired foci of hepatocytes	HBsAg-positive focus¹	HBV DNA/BG (qPCR)²	HBsAg-negative focus¹	HBV DNA/BG (qPCR)²
Y2 A	3p	0.59	1n	0.40
Y2 B	4p	1.54	10n	0.18
Y3 A	13p	1.17	15n	0.88
Y3 B	7p	0.52	17n	0.30
Y3 C	8p	0.73	26n	0.49
Y5 A	1p	0.45	1n	0.19
Y5 B	5p	0.91	5n	0.31
Y6 A	7p	4.03	7n	0.40
XA A	1p	4.32	1m	1.72
XA B	2p	3.26	2m	2.09
XA C	3p	8.13	3m	4.26
XA D	4p	1.23	4m	2.84

¹ HBsAg-positive and -negative hepatocytes were isolated by laser microdissection as described in Section 2.20 in liver tissue sections of Patients Y2-Y6. Focus identifiers are those used to identify regions in Figure 7.1.

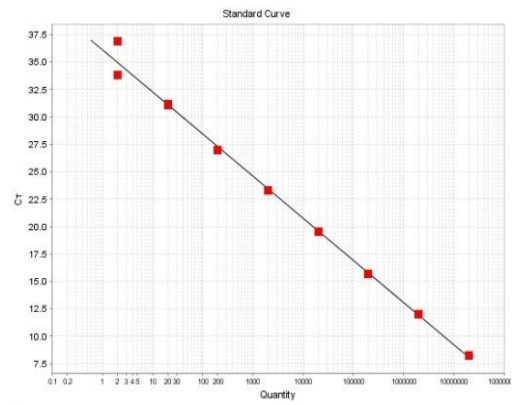
² Total DNA was isolated (as described in Section 2.21) from the two pools of isolated foci of hepatocytes from each patient and levels of HBV DNA were quantified in each extract by qPCR as described in Section 2.13. HBV DNA copy numbers were normalised to copies of BG, a single-copy cellular gene. Significantly greater copies of HBV DNA were detected in HBsAg-positive hepatocytes compared to paired HBsAg-negative hepatocytes from the same tissues after analysis with a student's t-test ($p < 0.05$). Results are graphed in Figure 7.13.

A Beta globin



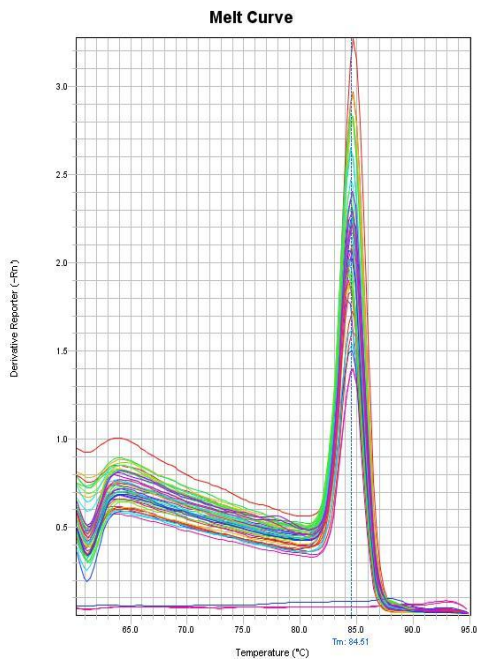
$R^2 > 0.99$
Efficiency = 78%

B HBV DNA



$R^2 > 0.99$
Efficiency = 82%

C



D

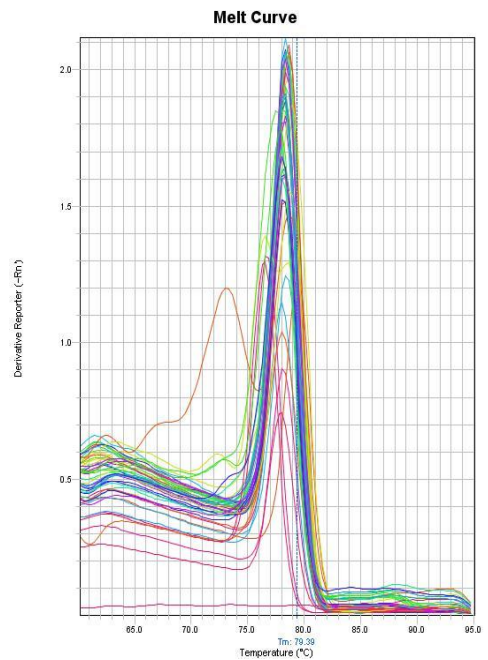


Figure 7.12. Standard curves (7.12A and 7.12B) and melt curves (7.12C and 7.12D) of qPCR assays detecting copy numbers of BG (7.12A and 7.12C) and total HBV DNA (7.12B and 7.12D). qPCR was carried out as described in Section 2.13 to determine the copy number of HBV DNA in DNA of HBsAg-positive and –negative hepatocyte foci isolated by laser microdissection, as described in Section 2.20 The sequence of the primers used for amplifying BG and HBV DNA are listed in Table 2.1. The BG gene was subcloned into a pBlueBac vector, as described in Section 2.13.1, to form pBG. pBG and the HBV DNA-containing plasmid pBB4.5HBV1.3 were used to produce the standard curves, shown in Figures 7.12A & 7.12B respectively. Plasmid DNA standards were diluted from 2×10^7 copies per reaction. The melt curves in Figure 7.12C and 7.12D show that only single products were amplified in all samples.

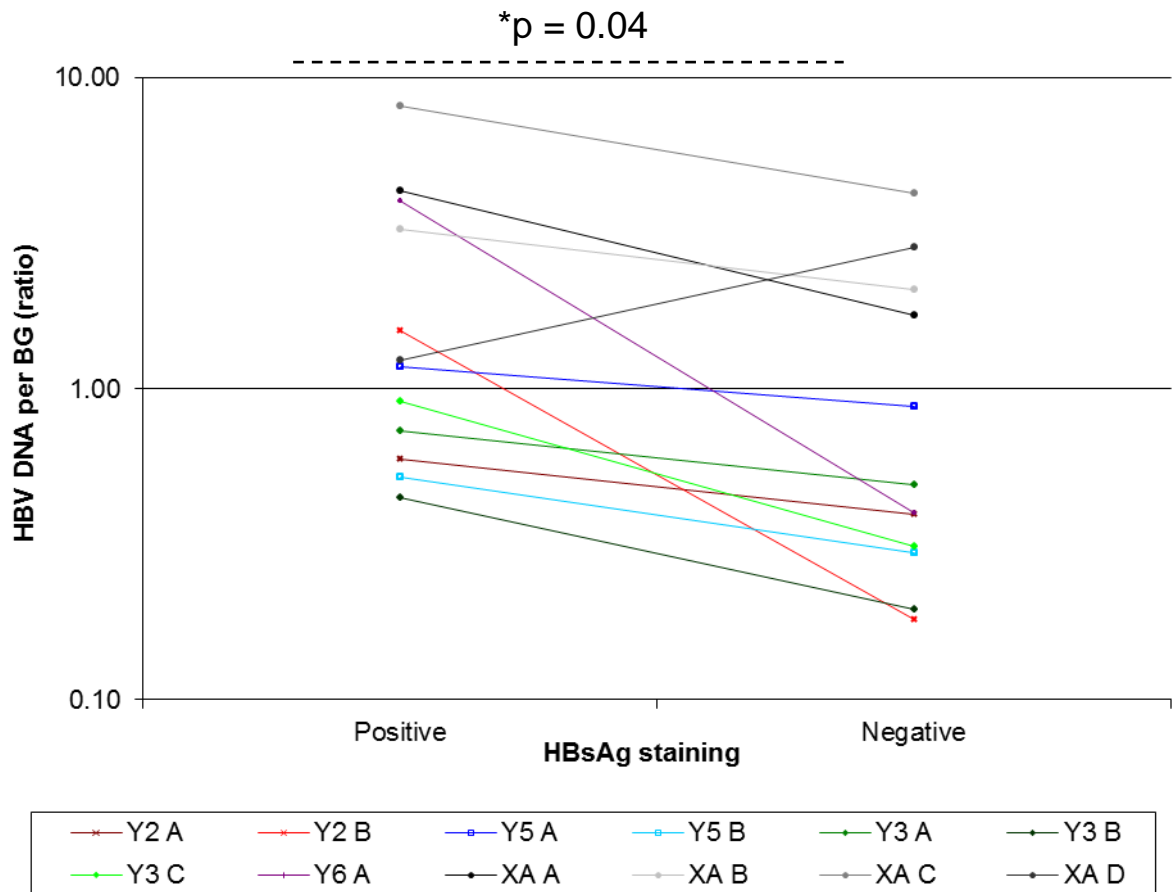


Figure 7.13. Twice as many HBV copies per cell were detected by qPCR on average in DNA extracts of HBsAg-positive hepatocyte foci compared to paired HBsAg-negative hepatocyte foci. Paired, neighbouring foci of HBsAg-positive and -negative hepatocytes were isolated by laser microdissection as described in Section 2.20 from the liver tissue sections of Patients Y2-Y6 and XA (listed in Table 7.3). Total DNA was isolated from each focus as described in Section 2.21 and levels of HBV DNA were quantified in each extract by qPCR as described in Section 2.13. HBV DNA copy numbers were normalised to copies of BG, a single-copy cellular gene. Significantly greater copies of HBV DNA were detected in HBsAg-positive hepatocyte foci compared to HBsAg-negative hepatocyte foci after analysis with a paired student's t-test ($p < 0.05$). The average difference in the HBV copy number per cell between each pair of foci was two-fold.

of surface receptor, enhancer or binding factor required by HBV to infect the host cell. These aspects were investigated in greater detail in Sections 7.3.7-7.3.11.

7.3.7 HBV PreS1/S2 mutants were predominantly detected in HBsAg-positive hepatocytes

Other differences in the phenotypes of HBsAg-positive and -negative hepatocytes were investigated to determine the cause of the heterogeneity in HBV antigen expression. During characterisation of the patient tissues, HBsAg antigen expression in the liver tissues of chronic HBV patients as described in Section 4.3.4 displayed a similar distribution to previous studies (Wang, Wu et al. 2003; Wang, Huang et al. 2006). In these studies, deletion mutants in the PreS region were described as the cause of the patterning of HBsAg accumulation and GGH phenotype. Therefore, the PreS region of HBV DNA in HBsAg-positive and -negative hepatocytes was sequenced to determine whether PreS mutations were responsible for the difference in levels of HBsAg expression.

DNA extracts of the foci of HBsAg-positive and -negative hepatocytes isolated in Section 7.3.1 were analysed for mutations in the PreS region of the HBV genome. The PreS sequences of the HBV DNA contained in the DNA extracts were amplified with hemi-nested primers as described in Table 2.8, and then sequenced as described in Section 2.23.

As shown in Table 7.4, four unique PreS deletion mutants were detected in both HBsAg-positive and HBsAg-negative foci of hepatocytes. The four different PreS mutants had deletions in nt 476-524, nt 489-492, nt 496-544 and nt 492-540 (based on numbering as described in Genbank accession #EU093912.1), respectively. PreS mutants were not detected in the majority of HBV-positive tissues, or in any foci of HBsAg-negative hepatocytes. Moreover, all of the HBsAg-positive foci also contained wild-type HBV. This suggests that PreS mutations are not sufficient to explain HBsAg-positive and negative hepatocytes.

7.3.8 No significant difference in methylation of HBV DNA was observed between HBsAg-positive and -negative foci of hepatocytes

HBV DNA methylation could potentially account for the difference in HBsAg expression in different foci of hepatocytes of the same patient. Previous studies have identified 3 CpG

Table 7.4. PreS deletion mutants detected in total DNA extracted from foci of HBsAg-positive and -negative hepatocytes isolated by laser microdissection.

Focus ID ¹	Number of PreS mutants / number of PreS regions sequenced	Deleted region ²
Y2 1p	0/5	
Y2 2p	0/3	
Y2 4p	0/9	
Y2 5p	0/3	
Y2 1m	0/3	
Y2 2m	0/6	
Y2 3m	0/5	
Y2 4m	0/3	
Y4 1p	0/1	
Y4 6p	0/7	
Y4 7p	0/4	
Y4 9p	0/1	
Y4 1m	0/8	
Y4 5m	0/3	
Y4 7m	0/8	
Y5 1p	1/4	476-524
Y5 5p ³	10(7+3)/11	(7) 489-492
		(3) 496-544
Y5 8p	1/1	496-544
Y5 1m	0/1	
Y5 2m	1/3	496-544
Y5 4m	0/12	
Y5 5m	1/3	492-540

¹ Focus identifiers are those used to identify regions in Figure 7.1.

² Based on the sequence numbering as described in Genbank accession #EU093912.1

³ Two distinct preS mutants were detected in the DNA extracted from sample Y5 5p; one was detected 7 times out of 11 sequences and the other 3 times out of 11 sequences.

islands in the HBV DNA genome (Vivekanandan, Thomas et al. 2008). In particular, CpG island 1 and 2 influence the expression of the HBsAg and HBx ORFs respectively. Previous studies have found that HBV DNA can be methylated in liver tissues and that this is associated with down-regulated HBV antigen expression (Vivekanandan, Thomas et al. 2008; Vivekanandan, Thomas et al. 2009). Therefore, the extent of HBV DNA methylation was compared between foci of HBsAg-positive and -negative hepatocytes, using a novel combined restriction enzyme and qPCR assay that had been previously optimised in Section 3.2.1.8.

For Patients Y2, Y3, Y4, Y5, Y6 and XA, all of the foci of HBsAg-positive and -negative hepatocytes in each liver tissue section were isolated by laser-microdissection and pooled into two groups. As shown in the standard and melt curves of Figure 7.14, single products were amplified by qPCR assays specific for HBV DNA, CpG island 1 and CpG island 2 sequences, with and without digestion of the DNA extracts with *AciI*. The extent of methylation of HBV DNA was then compared between the DNA extracts isolated from foci of HBsAg-positive and -negative hepatocytes in each tissue. As shown in Figure 7.15, no significant differences in methylation of *AciI* sites in either HBV CpG islands 1 or 2 were detected between foci of HBsAg-positive and -negative hepatocytes for any of the patients tested. This suggests that HBV DNA methylation was not the cause of the differences in HBsAg-expression in hepatocytes.

7.3.9 Cellular protein expression was altered in foci of HBsAg-positive hepatocytes compared to foci of HBsAg-negative hepatocytes

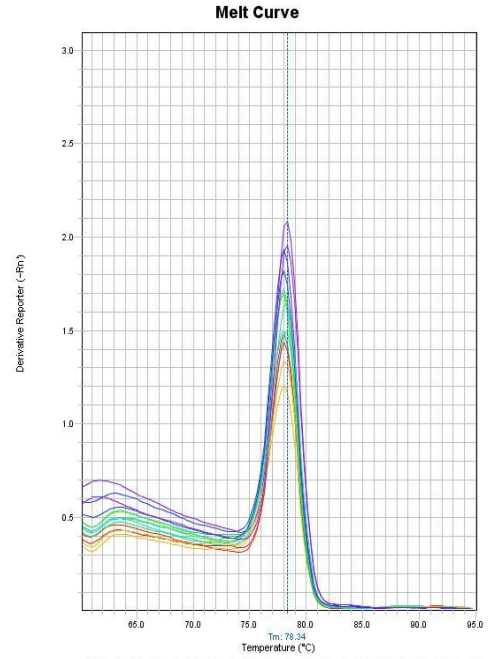
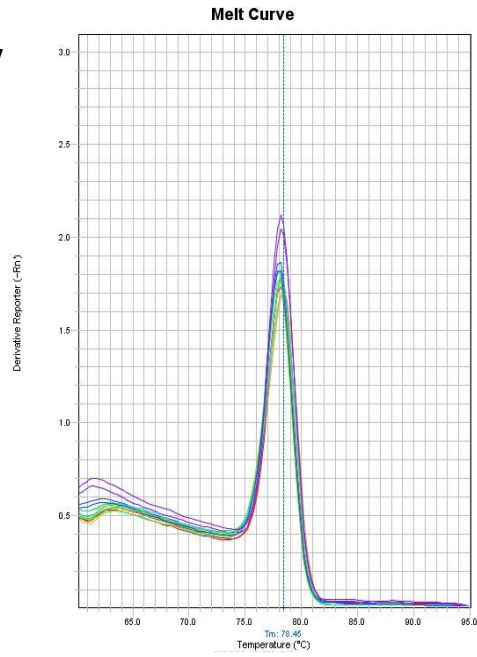
The heterogeneous HBsAg expression in the hepatocyte population may be altered by differences in cellular protein expression. We therefore used HPLC-nESI-LTQ Orbitrap MS to analyse differences in cellular protein expression between foci of HBsAg-positive and -negative hepatocytes. Mass-spectrometry analysis was done in collaboration with Mr. Johan Gustafsson, Ms. Yin Ying Ho and A/Prof. Peter Hoffman (Adelaide Proteomics Centre, University of Adelaide).

Neighbouring foci of HBsAg-positive and -negative hepatocytes (listed in Table 7.5) that were approximately the same size were isolated by laser-microdissection from liver tissue sections from Patients Y2, Y3, Y4, Y5 and Y6. The isolated foci of hepatocytes were lysed

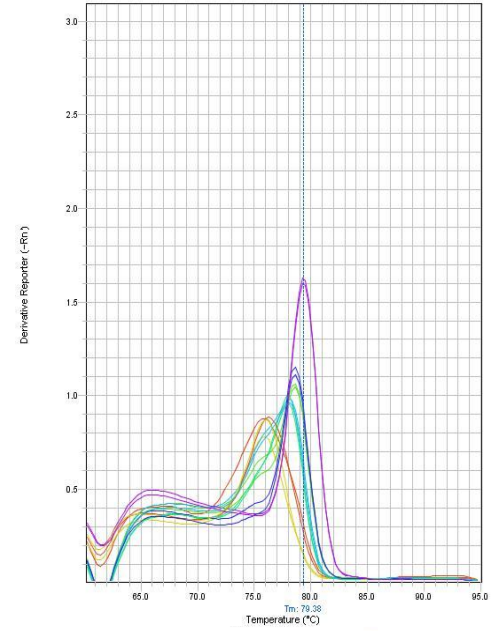
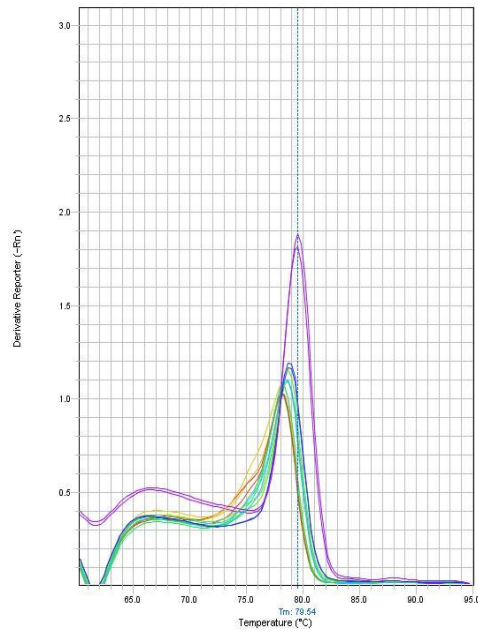
Mock-cut

AcI-cut

Total HBV DNA



CpG island 1



CpG island 2

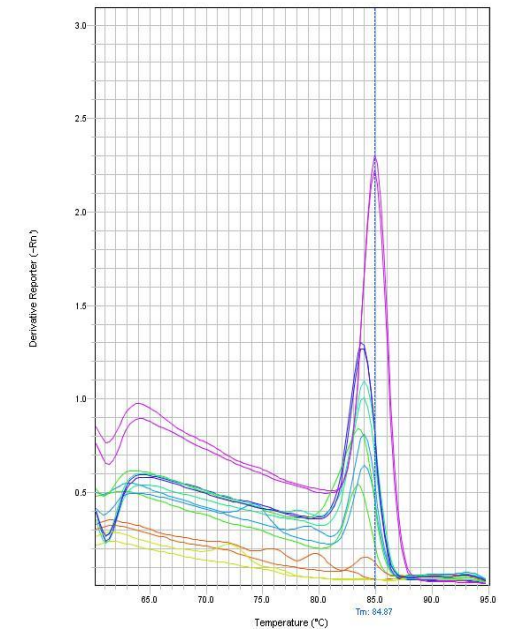
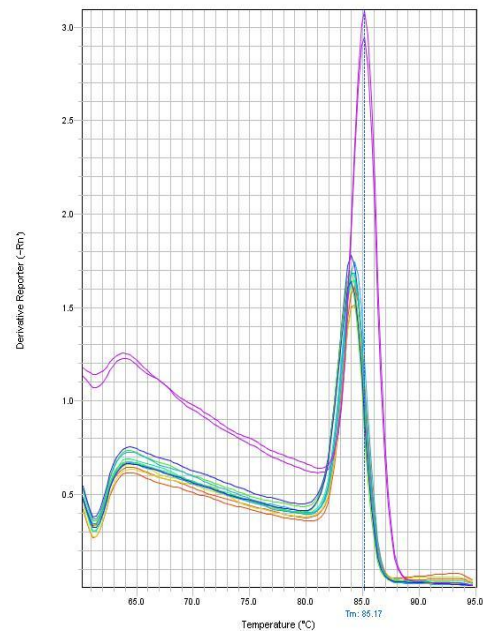


Figure 7.14. Melt curves for qPCR assays specific for the HBV surface gene (measuring total HBV DNA), HBV CpG islands 1 and 2 in mock-digested and *AciI* digested DNA extracts from pooled HBsAg-positive and -negative hepatocytes from HBV-infected patients. The extent of methylation of HBV CpG islands 1 and 2 were analysed (as described and optimised in Section 3.2.1.8) in total DNA extracts of paired HBsAg-positive and -negative hepatocyte foci isolated as described in Section 2.20 from sections of liver tissue of patients with chronic HBV infection. Single products were amplified in all assays, showing that the qPCR assay was specific.

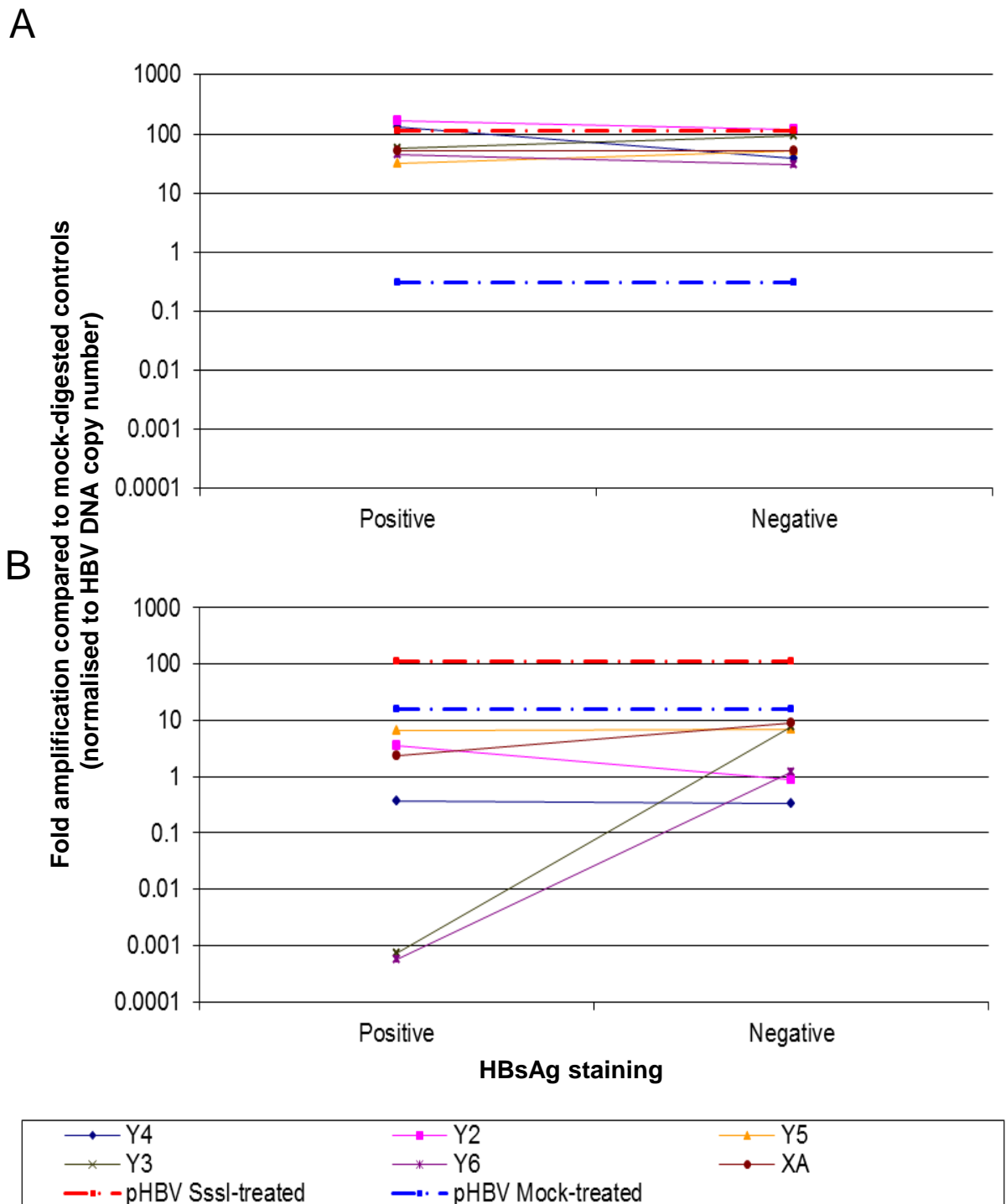


Figure 7.15. No significant difference in HBV DNA methylation of CpG island 1 (7.15A) or CpG island 2 (7.15B) was detected between HBsAg-positive and -negative foci of hepatocytes pooled from cirrhotic and non-cirrhotic liver tissues. The extent of methylation of HBV CpG islands 1 and 2 were analysed (as described and optimised in Section 3.2.1.8) in total DNA extracts of paired HBsAg-positive and -negative hepatocyte foci isolated from sections of liver tissue from patients with chronic HBV infection. The paired hepatocyte foci are described in Table 7.3 and their locations in the section are shown in Figure 7.3. SssI- and mock-treated samples of plasmids containing a 1.3-genome length HBV DNA genome were used as positive and negative controls respectively. No significant difference in HBV DNA methylation in either CpG islands 1 and 2 between HBsAg-positive and HBsAg-negative hepatocyte foci.

Table 7.5. Paired foci of hepatocytes analysed by HPLC-nESI-LTQ Orbitrap MS

Pair name	HBsAg-positive focus¹	Number of unique mass peaks detected	HBsAg-negative focus¹	Number of unique mass peaks detected
Y2A	5p	2022	2n	1951
Y5 A	1p	1698	4n	734
Y6 A	8p	1181	8n	1244
Y6 B	9p	1978	10n	306
XA A	4p	1184	3n	1246
XA B	6p	1700	12n	1735

¹ Focus identifiers are those used to identify regions in Figure 7.1.

with urea and the cellular proteins were digested with trypsin as described in Section 2.27. The digests were then analysed by HPLC-nESI-LTQ Orbitrap MS to precisely determine the mass-charge value, and therefore the sequence, of the tryptic peptides as described in Section 2.27. The lists of identified peptides were then compared between neighbouring foci of HBsAg-positive and -negative hepatocytes.

The numbers of peptides detected in protein extracts of each focus of HBsAg-positive or -negative hepatocytes are shown in Table 7.5. When each pair of hepatocyte foci was compared, as shown in Figure 7.16, the majority of the peptides detected were found in extracts of both foci of HBsAg-positive and HBsAg-negative hepatocytes.

To determine whether any cellular proteins are associated with either HBsAg-positive or HBsAg-negative hepatocytes, all peptides detected in extracts of both foci of HBsAg-positive and HBsAg-negative hepatocytes were combined into two lists. To decrease spurious associations, peptides detected in <2, <3, <4, <5 or <6 extracts were excluded from the lists before they were compared. The results of the comparisons between the peptides detected in the extracts from foci of HBsAg-positive and HBsAg-negative hepatocytes are shown in Figure 7.17. In most cases, the majority of peptides were found in protein extracts of foci of HBsAg-positive and HBsAg-negative hepatocytes. Also, 450-550 peptides were consistently detected only in foci of HBsAg-positive hepatocytes, while only few (generally <100) were detected exclusively in foci of HBsAg-negative hepatocytes.

We aligned the peptide sequences that were found in the majority of extracts, i.e. >3 extracts, to their protein using MASCOT database queries. To improve robustness and avoid false positives, we excluded any proteins to which only a single peptide was aligned. The protein lists for extracts from HBsAg-positive and -negative hepatocytes were then compared and the results are summarised in Figure 7.17. From this analysis, 18 proteins were exclusively detected in extracts of foci of HBsAg-positive hepatocytes and are listed in Table 7.6. No proteins were exclusively detected in extracts of foci of HBsAg-negative hepatocytes. This suggests that: 1) cellular expression or upregulation of these genes predisposes the hepatocytes to expression of HBsAg; 2) down-regulation of particular proteins causes the loss or reduction in HBV infection, HBV replication, and/or HBsAg expression; and/or 3) increased expression of HBsAg causes expression or upregulation of these genes.

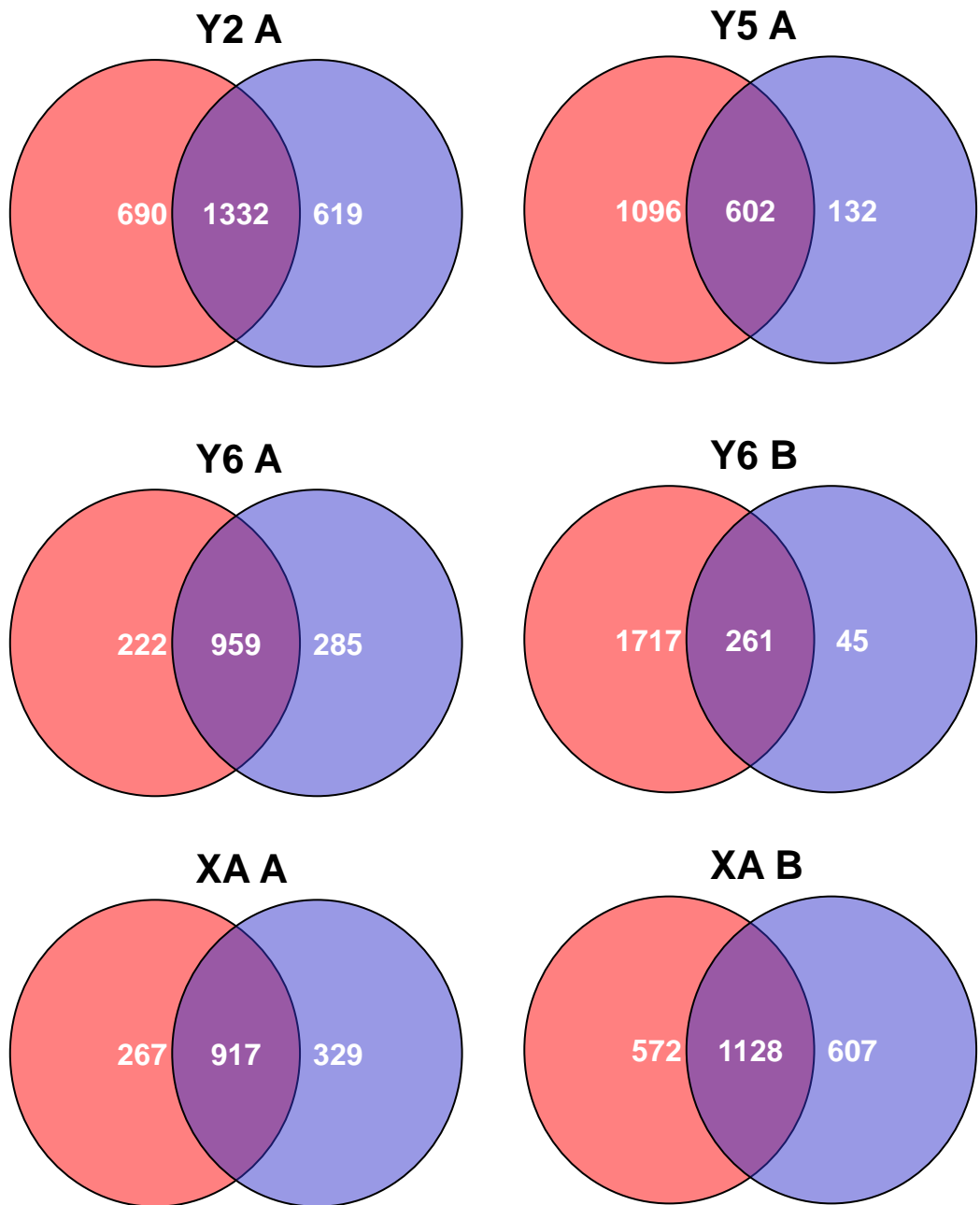


Figure 7.16. Unique mass peaks found in proteins extracts from paired HBsAg-positive (red) and -negative (blue) hepatocyte foci. Protein was extracted from isolated hepatocyte foci (listed in Table 7.5) and subjected to HPLC-nESI-LTQ Orbitrap MS as described in Section 2.27. High levels of overlap was seen in the mass peaks detected in HBsAg-positive and -negative hepatocyte foci.

Peptides repeated in:

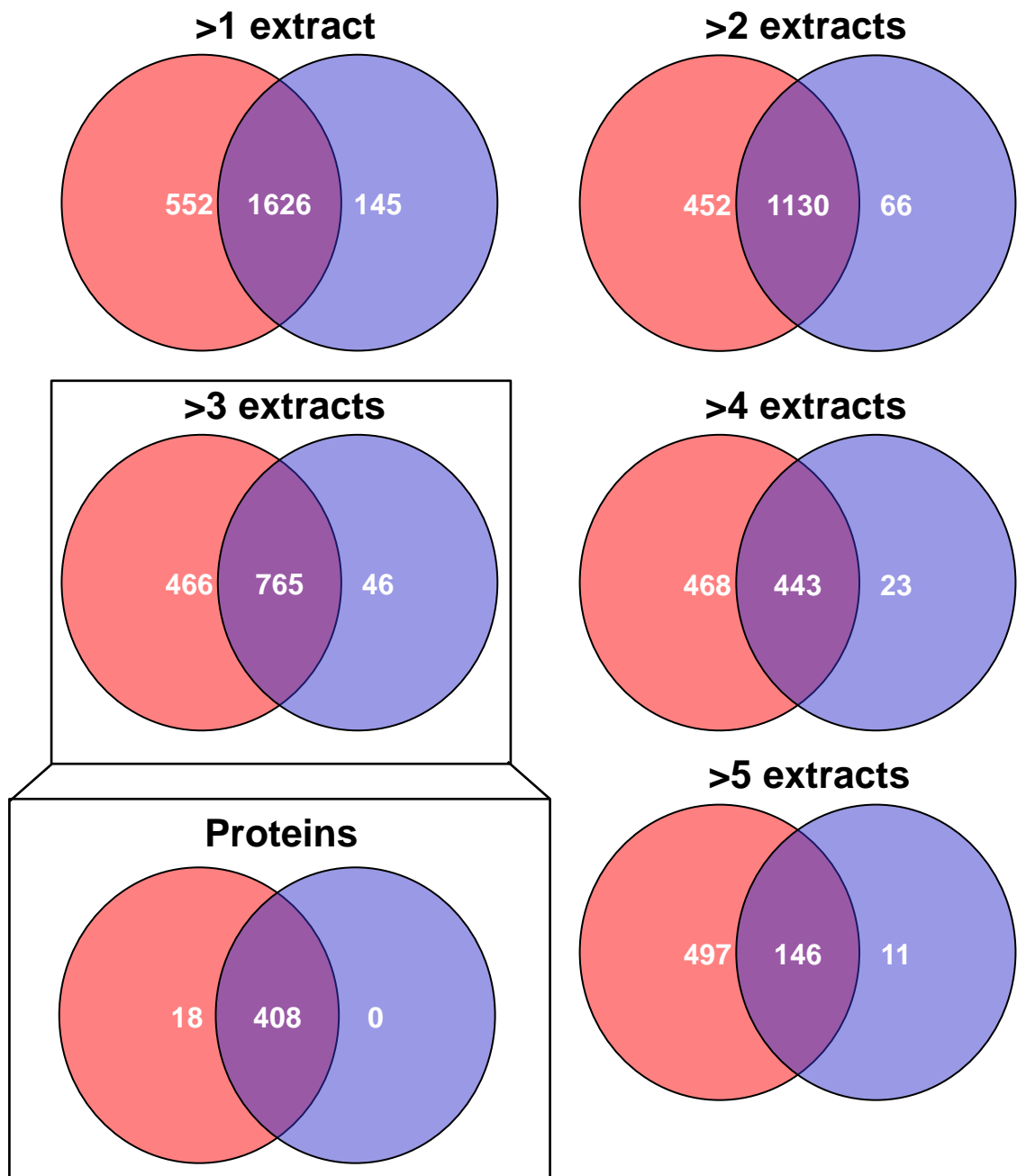


Figure 7.17. Total unique mass peaks found in protein extracts from all pairs of HBsAg-positive (red) and -negative (blue) hepatocyte foci. The number of mass peaks found commonly in protein extracts by HPLC-nESI-LTQ Orbitrap MS as described in Section 2.27. For example, for the “>2 extracts” diagram, the numbers represent peptides that were repeated only in extracts of HBsAg-positive hepatocytes in >2 extracts (red), only in extracts of HBsAg-negative hepatocytes in >2 extracts (blue) and in both HBsAg-negative and -positive hepatocytes in >2 extracts of both (purple).

The peptides found in >3 extracts, i.e. in the majority of the extracts analysed, were aligned with their cellular proteins through a MASCOT search as described in Section 2.27 and grouped into expression in HBsAg-positive hepatocytes only (red), HBsAg-negative hepatocytes only (blue) and both (purple). Proteins with only a single peptide aligned were then excluded to increase confidence. The majority of proteins were not unique to either HBsAg-positive or -negative hepatocytes. Those proteins that were uniquely detected in HBsAg-positive hepatocytes are listed in Table 7.7.

Table 7.6. Proteins detected by HPLC-nESI-LTQ Orbitrap MS in >3 extracts of HBsAg-positive hepatocyte foci, but not detected in HBsAg-negative hepatocyte foci

Protein accession ¹	Protein description	m/z	Peptide sequence
ACON_HUMAN	Aconitate hydratase, mitochondrial GN=ACO2 PE=1 SV=2	1666.759	WVIGDENYGEGRSSR
		1752.868	DINQEVYNFLATAGAK
		1952	AKDINQEVYNFLATAGAK
		2779.439	NDANPETHAFVTSPEIVTALAIAGTLK
ADH1G_HUMAN	Alcohol dehydrogenase 1C GN=ADH1C PE=1 SV=2	1075.566	INEGFDLLR
		1568.845	KPFSIEEVEVAPPK
AL1B1_HUMAN	Aldehyde dehydrogenase X, mitochondrial GN=ALDH1B1 PE=1 SV=3	1774.841	TFVEESIYNEFLER
		1816.035	VAEQTPLSALYLASLIK
ALAT1_HUMAN	Alanine aminotransferase 1 GN=GPT PE=1 SV=3	1279.667	AWALDVAELHR
		2570.323	DGGIPADPNNVFLSTGASDAIVTVLK
CNDP2_HUMAN	Cytosolic non-specific dipeptidase GN=CNDP2 PE=1 SV=2	1232.64	TVFGVEPDLTR
		1733.884	EGGSIPVTLTFQEATGK
COX2_HUMAN	Cytochrome c oxidase subunit 2 GN=MT-CO2 PE=1 SV=1	1105.686	VVLPIEAPIR
		2242.149	MMITSQDVLHSHWAVPTLGLK
DHB13_HUMAN	17-beta-hydroxysteroid dehydrogenase 13 GN=HSD17B13 PE=2 SV=1	1454.777	LWPVLETDEVVR
		1705.948	SVAGEIVLITGAGHGIGR
		1834.043	KSVAGEIVLITGAGHGIGR
EF2_HUMAN ^A	Elongation factor 2 GN=EEF2 PE=1 SV=4	1443.761	EGIPALDNFLDKL
		1741.831	YLAEKYEWVDAEAR
		1798.889	AYLPVNESFGFTADLR
ENOB_HUMAN	Beta-enolase GN=ENO3 PE=1 SV=4	1883.953	LAMQEFMILPVGASSFK
		2742.335	DATNVGDEGGFAPNILENNEALELLK
HUTI_HUMAN	Probable imidazolonepropionase GN=AMDHD1 PE=2 SV=2	1341.787	SLAVLEGASLVVGK
		1993.001	MSMPEALAAATINAAYALGK
MAAI_HUMAN	Maleylacetoacetate isomerase GN=GSTZ1 PE=1 SV=3	2109.187	MISDLIAGGIQPLQNLVSLK
		2125.182	MISDLIAGGIQPLQNLVSLK
NNMT_HUMAN ^B	Nicotinamide N-methyltransferase GN=NNMT PE=1 SV=1	1985.138	NLGSLKPGGFLVIMDALK
		2021.979	EIVVTDYSDQNQLQELEK
SARDH_HUMAN ^C	Sarcosine dehydrogenase, mitochondrial GN=SARDH PE=1 SV=1	1052.566	FYLVGLDAR
		1876.863	GPAPVLEYDYGYGAYGSR
SPRE_HUMAN	Sepiapterin reductase GN=SPR PE=1 SV=1	1512.851	LLLINNAGSLGDVSK
		1891.053	VPADLGAEAGLQQLLGALR
SPTA2_HUMAN	Spectrin alpha chain, brain GN=SPTAN1 PE=1 SV=3	1857.943	DLAALGDKVNSLGETAER
		1950.928	SSLSSAQADFNQLAELDR
		1971.058	IAALQAFADQLIAAGHYAK
UDB10_HUMAN	UDP-glucuronosyltransferase 2B10 GN=UGT2B10 PE=1 SV=1	1463.769	TEFENIIMQLVK
		1479.764	TEFENIIMQLVK
VIME_HUMAN ^D	Vimentin GN=VIM PE=1 SV=4	1168.707	ILLAELEQLK
		1532.845	KVESLQEEIAFLK
		1569.888	ISLPLPNFSSLNLR
		1660.94	KVESLQEEIAFLKK
WDR1_HUMAN ^E	WD repeat-containing protein 1 GN=WDR1 PE=1 SV=4	2201.954	EMEENFAVEAANYQDTIGR
		1617.768	YAPSGFYIASGDVSGK
		1718.815	FATASADGQIYIDGK

¹ Proteins included in this list were aligned to multiple mass peaks detected by HPLC-nESI-LTQ Orbitrap MS in extracts of >3 different HBsAg-positive hepatocyte foci and were not detected in >3 different HBsAg-negative hepatocyte foci. Those shown in red have been reported to be up-regulated in HCC. Those shown in blue have been reported to be down-regulated in HCC.

^A Up-regulated in HCC according to Li *et al.* (2008).

^B Down-regulated in HCC according to Lee *et al.* (2005).

^C Down-regulated in HCC according to Lim *et al.* (2002).

^D Up-regulated in HCC according to Sun *et al.* (2010).

^E Up-regulated in HCC according to Wang *et al.* (2011).

The proteins expressed exclusively in foci of HBsAg-positive hepatocytes did not appear to match a pre-neoplastic phenotype. As shown in Table 7.6, the majority of proteins have not been reported to be altered between HCC and non-HCC tissues. Furthermore, some proteins were previously reported to be up-regulated in HCC, whereas others have been reported to be down-regulated (Lim, Park et al. 2002; Lee, Chen et al. 2005; Li, Chen et al. 2008; Sun, Poon et al. 2010; Wang, Sun et al. 2011). This suggests that a preneoplastic phenotype is not necessarily associated with foci of HBsAg-positive hepatocytes.

7.3.10 Altered cellular protein expression due to zonal heterogeneity was not observed in liver tissue sections by IMS

Survival advantages may manifest as changes in cellular protein expression. In order to identify potential cellular protein candidates for survival advantages, IMS and HPLC-nESI-LTQ Orbitrap MS were used as described in Section 2.26-2.29 to determine the distribution of cellular protein expression in HBV non-infected controls and HBV-infected tissues. Mass-spectrometry analysis was done in collaboration with Mr. Johan Gustafsson, Ms. Yin Ying Ho and A/Prof. Peter Hoffman (Adelaide Proteomics Centre).

Briefly, rehydrated sections of liver tissue from HBV non-infected controls and HBV-infected patients were digested with trypsin to produce tryptic peptides. As described in Section 2.26, a matrix of 7 mg/mL α -cyano-4-hydroxycinnamic acid was deposited on the tissue sections. Matrix-assisted laser-desorption mass spectrometry (MALDI-MS) was carried out *in situ* at 100 μ m intervals in a raster pattern. This results in a mass spectrum for each interval on the tissue. For each tissue section, mass peaks were automatically picked using in-house flexAnalysis scripts from section-wide sum spectra and a map of the distribution of each mass peak was digitally produced. Sections of liver tissue from non-infected patients were used as controls and compared to the distributions of mass peaks detected in the non-tumour liver tissues from HBV-infected patients with HCC.

In order to identify mass peaks, the mass-peak list of peptides identified by HPLC-nESI-LTQ Orbitrap MS from each liver tissue section produced in Section 7.3.8 was aligned with a mass peak list produced by IMS analysis. The subsequent mass peak distribution maps and matched peptide identities were filtered by the following criteria:

1. For each patient, IMS was conducted on serial liver tissue sections in triplicate. If a mass peak was detected in the sum spectra of <2 of the triplicate IMS runs, then the mass peak was discarded.
2. If peptide identities that were assigned to a single mass peak conflicted with another from a separate IMS run, then that mass peak was discarded from our analysis.
3. If a mass peak was detected in the IMS sum spectra of i) <2 of the 3 NHL control tissue samples, and ii) <3 of the 5 HBV-infected tissue samples, then the mass peak was discarded from our analysis.
4. If a mass peak was unable to be aligned with peptide identities determined by HPLC-nESI-LTQ Orbitrap MS from >1 independent runs, then the mass peak was discarded from our analysis.

After this stringent filtration, 143 peptides that were present in the majority of tissue sections were confidently detected by IMS and identified by HPLC-nESI-LTQ Orbitrap MS. These peptides could be aligned with 118 unique cellular proteins.

Previous studies have shown heterogeneous expression of proteins between peri-central and peri-portal hepatocytes, such as in the case of fructose 1,6-bisphosphatase and cytochrome P450 (Puschel and Jungermann 1994; Eilers, Modaressi et al. 1995; Jungermann and Kietzmann 1996). These two genes are known to be more highly expressed at both an mRNA and functional activity levels in the peri-portal and peri-central hepatocytes, respectively. The heterogeneity is assumed to be a by-product of the zonal heterogeneity within the liver as fructose 1,6-bisphosphatase and cytochrome P450 have roles in the metabolism of nutrients and toxins respectively, which form concentration gradients from the peri-portal to the peri-central regions of the liver lobules (Jungermann and Katz 1989).

When mass-peaks associated with these gene products were analysed by IMS, as shown in Figure 7.18, no heterogeneity was observed. This suggests that IMS may be relatively insensitive to changes in cellular protein expression. This also suggests that any changes in cellular protein expression detected by IMS between cell populations are greater than those present due to zonal heterogeneity that occurs due to the position of hepatocytes in liver lobules.

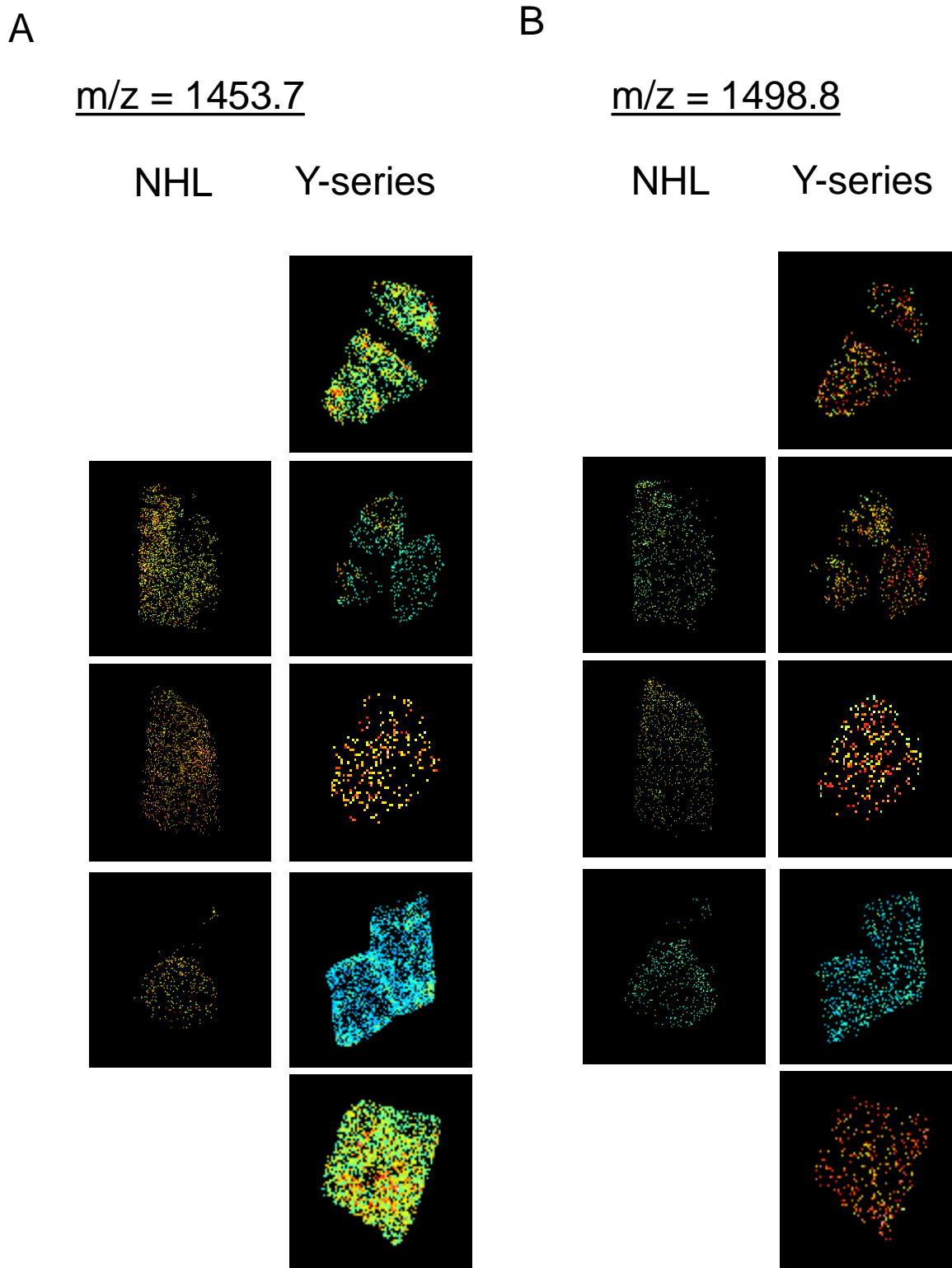


Figure 7.18. The distribution of tryptic peptides aligned with fructose 1,6-bisphosphatase (7.18A) and cytochrome P450 2D6 (7.18B) in liver tissue sections of **HBV-negative control patients and chronic HBV patients**. The distribution of proteins shown in previous studies to be expressed differently in peri-portal and peri-central hepatocytes were detected by IMS and HPLC-nESI-LTQ Orbitrap MS as described in Sections 2.26, 2.27 and 2.29. No obvious heterogeneity was seen in their distributions by IMS. The resolution of the peptide distribution map is 100 μm between each pixel.

7.3.11 Cellular protein distribution is heterogeneous in liver tissue sections from patients with chronic HBV infection compared to HBV non-infected controls

We sought to further characterise and to determine the distribution of the peptides detected by mass spectrometry analysis. Of the 143 peptides identified by matching up peptide lists produced by IMS to peptide lists produced by HPLC-nESI-LTQ Orbitrap MS, 104, 8 and, 31 peptides showed increased, decreased and no change respectively in observed expression in the tissues of HBV-infected patients compared to HBV non-infected controls. When analysed by MASCOT database queries, the peptide sequences could be assigned to 84, 8 and, 26 unique proteins, respectively. Multiple peptides could be assigned to the same protein in 22 of the identified proteins, which are listed in Table 7.7. The ion maps showing the distributions of the 22 proteins detected in each patient tissue are shown in Appendix 9.10.

The peptides that were aligned to the same protein had similar distributions. For example, the distributions of three peptides aligned with haemoglobin subunit alpha, as shown in Figure 7.19, all showed areas of high expression in particular areas. When compared to H&E-stained sections of the same tissues, these areas of high expression aligned with blood vessels. Furthermore, a similar distribution is seen in peptides aligned with haemoglobin subunit beta, as shown in Figure 7.20. This finding provides an accurate internal control and suggests that the peptide identities assigned to mass peaks are accurate.

The majority of the identified proteins were upregulated in HBV-positive liver tissues. As shown in Figure 7.21, heterogeneous distribution was observed in some of the identified proteins, including: pyruvate carboxylase; tubulin beta-2C chain; thiosulfate sulphur-transferase; and selenium-binding protein 1, according to alignments with MASCOT database queries. The distributions of peptides assigned to these proteins did not coincide with the distributions of peptides assigned to haemoglobin subunits alpha and beta. Also, when compared to H&E-stained sections of the same tissues, the foci of high expression coincided with hepatocytes, which appeared to be histologically normal. In uninfected control tissues, these proteins were either homogeneously distributed or not detected by IMS.

This revealed that there is indeed heterogeneity in the phenotype(s) of the hepatocyte population and that those altered hepatocytes are arranged in foci, suggesting clonal

Table 7.7. List of cellular proteins identified by HPLC-nESI-LTQ Orbitrap MS and observed by multiple peptides by IMS in liver tissue sections of HBV-negative and -positive patients. Ion maps of each mass peak are shown in Appendix 9.10.

Protein accession	Protein description	m/z	Sequence	Expression in HBV-infected tissues
ACOX_HUMAN	Peroxisomal acyl-coenzyme A oxidase OS=Homo sapiens GN=ACOX1 PE=1 SV=3	1716.9	SKEVAWNLTSDVLR	Up
		1556.8	LTNILDGGAQNTALR	Up
ACSL1_HUMAN	Long-chain-fatty-acid--CoA ligase 1 OS=Homo sapiens GN=ACSL1 PE=1 SV=1	1351.7	VLQPTVFPVVR	No change
		1709.9	IIVVMDAYGSELSVER	Up
ACTN4_HUMAN	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	1421.7	GVEEWLLNEIR	Up
		1429.7	TINEVENQILTR	Up
		1537.8	FAIQDISVEETSAK	Up
ALBU_HUMAN	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	1149.6	LVNEVTEFAK	No change
		1511.8	VPQVSTPTLVEVSR	No change
		1467.7	RHPDYSVVLRLR	No change
		1639.9	DVFLGMFLYEYAR	No change
ALDH2_HUMAN	Aldehyde dehydrogenase, mitochondrial OS=Homo sapiens GN=ALDH2 PE=1 SV=2	1527.8	ANNSTYGLAAAVFTK	No change
		1385.7	LGPALATGNVVMK	Up
ALDOB_HUMAN	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2	1243.7	ALQASALAAWGGK	Up
		1490.8	GILAADESVGTMGNR	Up
		1458.7	ELSEIAQSIVANGK	Up
		1916	KYTPEQVAMATVTALHR	Up
ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1	1171.6	VVDALGNAIDGK	Up
		1316.7	TSIAIDTIINQK	Up
		2104.2	GMSLNLEPDNVGVVVFVGNPK	Up
ATPB_HUMAN	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3	1457.7	TVLIMELINNVAK	Up
		1650.8	LVLEVAQHLGESTVR	Up
CPSM_HUMAN	Carbamoyl-phosphate synthase [ammonia], mitochondrial OS=Homo sapiens GN=CPS1 PE=1 SV=2	1306.7	AADTIGYPVMIR	Up
		1607.8	VLGTSVESIMATEDR	Up
		1419.7	AVNTLNEALEFAK	Up
DOPD_HUMAN	D-dopachrome decarboxylase OS=Homo sapiens GN=DDT PE=1 SV=3	1499.8	PFLELDTNLPANR	Up
		2194.2	PFLELDTNLPANRVPAGLEK	Up
ECHA_HUMAN	Trifunctional enzyme subunit alpha, mitochondrial OS=Homo sapiens GN=HADHA PE=1 SV=2	1318.7	MQLLEIITTEK	Up
		1521.8	DLNSDMDSILASLK	Up
		1626.8	TIEYLEEVAITFAK	Up
ENOA_HUMAN	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2	1406.7	GNPTVEVDLFTSK	Up
		1804.9	AAVPSGASTGIYEALRLR	Up
		2033.1	FTASAGIQVVGDDTLVTNPK	Up

Table 7.7. Cont.

Protein accession	Protein description	m/z	Sequence	Expression in HBV-infected tissues
HBA_HUMAN	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=3	1071.5	MFLSFPTTK	No change
		1087.5	MFLSFPTTK	No change
		1529.7	VGAHAGEYGAEALER	No change
HBB_HUMAN	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2	1274.7	LLVVYPWTQR	No change
		1314.7	VNVDEVGGEALGR	No change
HS90B_HUMAN	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4	1349.7	TLTLVDTGIGMTK	Up
		1544.7	ELISNASDALDKIR	Up
PDIA3_HUMAN	Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4	1619.9	DLLIAYYDVDYEK	Up
		1370.7	ELSDFISYLQR	Up
PYC_HUMAN	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2	1547.8	AEAEAQAEELSFPR	Up
		1654.9	DFTATFGPLDSLNTK	Up
RPN_HUMAN	Dolichyl-diphosphooligosaccharide-- protein glycosyltransferase OS=Homo sapiens GN=RPN1 PE=1 SV=1	1301.7	ALTSEIALLQSR	Up
		1372.7	SIVEEIEDLVAR	Up
SBP1_HUMAN	Selenium-binding protein 1 OS=Homo sapiens GN=SELENBP1 PE=1 SV=2	1313.7	AVFWIEFVMR	Up
		1906.1	NTGTEAPDYLATVDVDPK	Up
SPYA_HUMAN	Serine--pyruvate aminotransferase OS=Homo sapiens GN=AGXT PE=1 SV=1	1494.7	IMAAGGLQMIGSMSK	Up
		2505.4	ALLKPLSIPNQLLLGPGPSNLPPR	Up
TBB2C_HUMAN	Tubulin beta-2C chain OS=Homo sapiens GN=TUBB2C PE=1 SV=1	1143.6	LAVNMVPFPR	Up
		1696.9	NSSYFVEWIPNNVK	Up
THTR_HUMAN	Thiosulfate sulfurtransferase OS=Homo sapiens GN=TST PE=1 SV=4	1452.7	TYEQVLENLESK	Up
		2010	FLGTEPEPDAVGLDSGHIR	Up

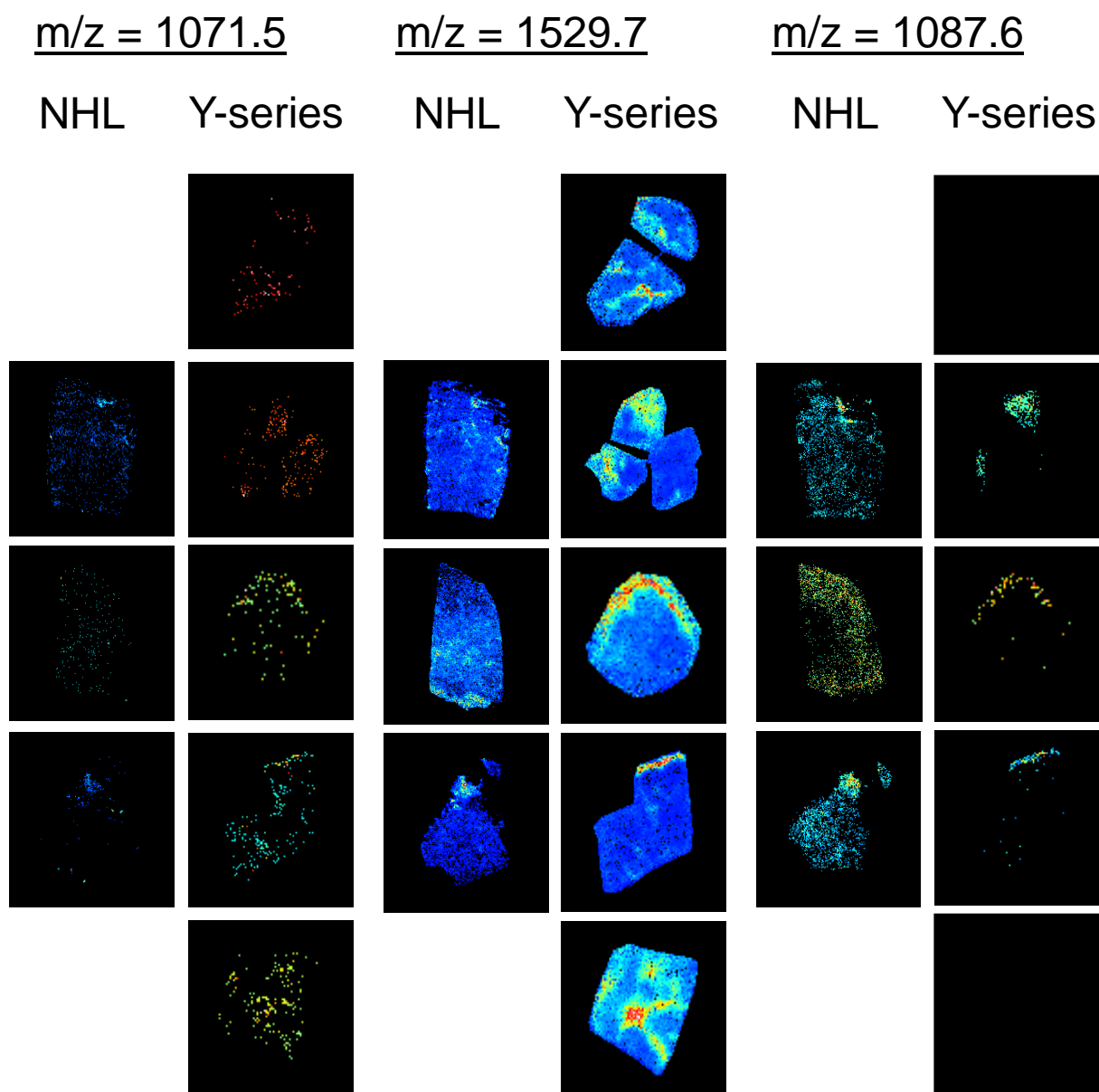


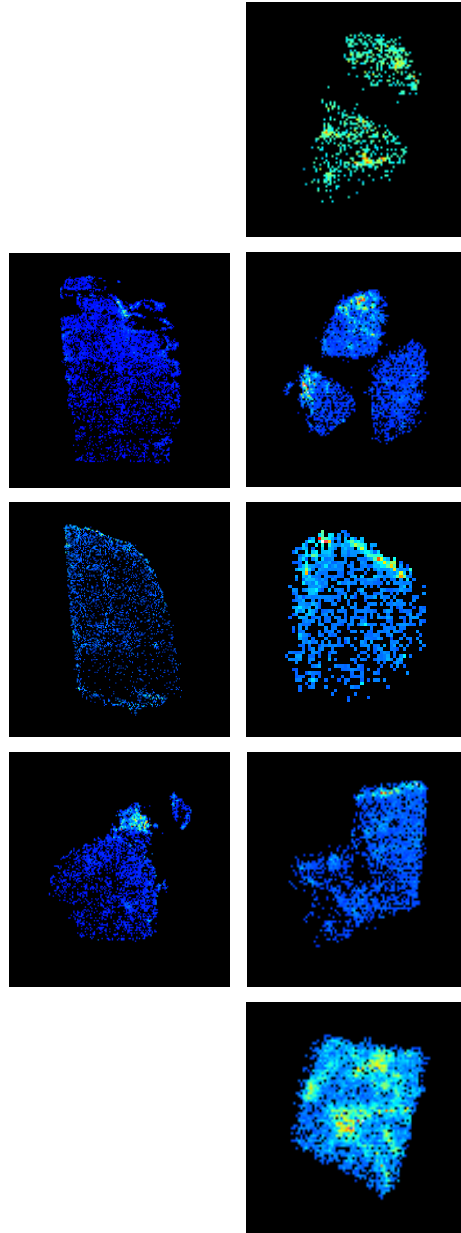
Figure 7.19. The distribution of three tryptic peptides of haemoglobin subunit alpha as identified by HPLC-nESI-LTQ Orbitrap MS and detected by IMS in liver tissue sections of HBV-negative control patients (NHL) and chronic HBV patients (Y-series). As a positive control, mass peaks associated with tryptic peptides of haemoglobin subunit alpha were identified using HPLC-nESI-LTQ Orbitrap MS as described in Section 2.27. The distribution of the peaks over the liver tissue sections was visualised using IMS as described in Sections 2.26 and 2.29. The distributions of the three mass peaks were similar in each section of liver. The resolution of the peptide distribution map is 100 μm between each pixel.

A

 $m/z = 1274.7$

NHL

Y-series



B

 $m/z = 1314.7$

NHL

Y-series

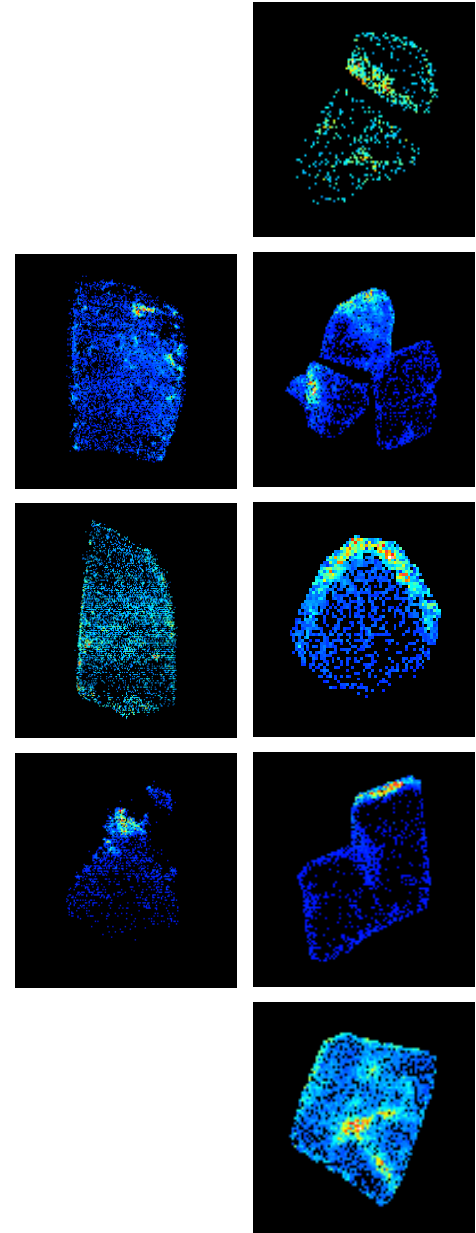


Figure 7.20. The distribution of two tryptic peptides of haemoglobin subunit beta as identified by HPLC-nESI-LTQ Orbitrap MS and detected by IMS in liver tissue sections of HBV-negative control patients (NHL) and chronic HBV patients (Y-series). As a positive control, mass peaks associated with tryptic peptides of haemoglobin subunit beta were identified using HPLC-nESI-LTQ Orbitrap MS as described in Section 2.27. The distribution of the peaks over the liver tissue sections was visualised using IMS as described in Sections 2.26 and 2.29. The distributions of the two mass peaks were similar in each section of liver and similar to the distributions of mass peaks associated with haemoglobin subunit alpha, shown in Figure 7.19. This suggests that IMS is accurately detecting the distribution of cellular proteins *in situ*. The resolution of the peptide distribution map is 100 μm between each pixel.

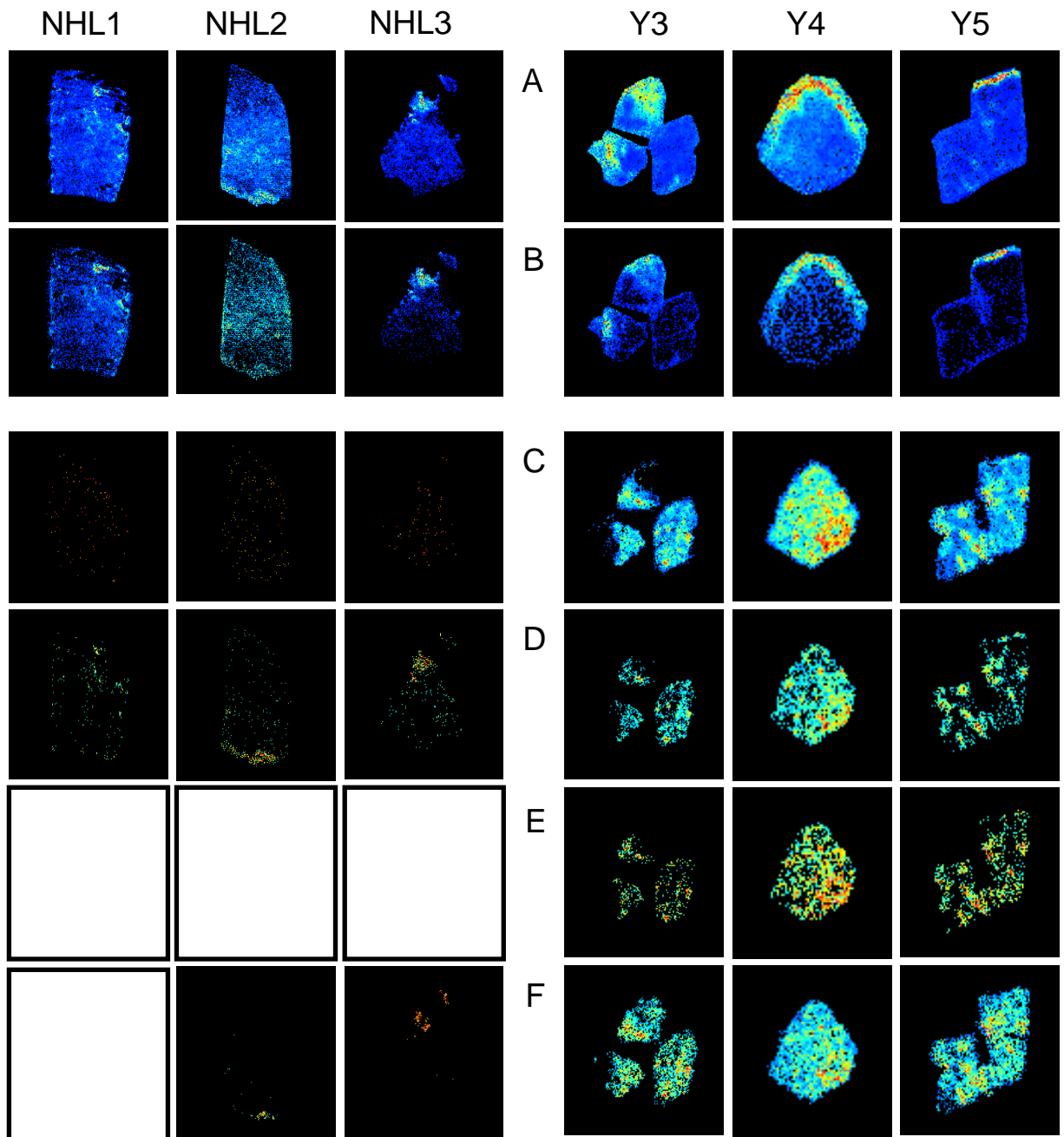


Figure 7.21. The heterogeneous distribution of tryptic peptides aligned with haemoglobin alpha ($m/z = 1529.7$, 7.21A), haemoglobin beta ($m/z = 1314.7$, 7.21B), pyruvate carboxylase ($m/z = 1654.9$, 7.21C), tubulin beta-2C chain ($m/z = 1696.9$, 7.21D), thiosulfate sulfurtransferase ($m/z = 2010.0$, 7.21E) and selenium-binding protein 1 ($m/z = 1906.1$, 7.21F) in liver tissue sections of HBV-negative control patients and chronic HBV patients. Mass peaks associated with tryptic peptides of cellular proteins were identified using HPLC-nESI-LTQ Orbitrap MS as described in Section 2.27. The distribution of the peaks over the liver tissue sections was visualised using IMS as described in Sections 2.26 and 2.29. Empty boxes indicate that mass peaks were not automatically selected by software analysis of IMS sum spectra. The distribution of some mass peaks representing cellular proteins (7.21C-F) are shown compared to the distribution of haemoglobin subunits (7.21A and 7.21B). Heterogeneous distribution of the selected cellular proteins was observed in liver tissues from patients chronically infected with HBV, but not uninfected controls. The distributions of the selected cellular proteins do not match the distribution of haemoglobin, suggesting that there is heterogeneous expression of proteins within the hepatocyte populations within the livers of HBV-infected patients. The resolution of the peptide distribution map is $100 \mu\text{m}$ between each pixel.

proliferation. The heterogeneity of expression of these proteins was only detected in HBV-infected liver tissue sections and not in HBV non-infected controls. Together, these data therefore suggest that chronic HBV infection induces the clonal growth of hepatocytes with altered phenotype.

However, the foci of hepatocytes with altered cellular protein expression did not appear to coincide with foci of HBsAg-positive hepatocytes or with hepatocyte clones detected by invPCR. Also, the peptides that were aligned to proteins exclusively in HBsAg-positive hepatocytes listed in Table 7.6 could not be readily detected by IMS. Only two of the proteins on the list in Table 7.6, a probable imidazolonepropionase ($m/z = 1342.7$) and vimentin ($m/z = 1619.9$), were confidently aligned to other mass-peaks detected by IMS; the distributions of which are shown in Figure 7.22. Although the distributions of these peaks appeared heterogeneous in some tissue sections, the sensitivity for the mass-peaks was not strong enough to determine this with much certainty. This is consistent with the low sensitivity of IMS observed in Section 7.3.9.

7.4 Discussion

7.4.1 Identification of the cell populations that had undergone the clonal proliferation detected by invPCR

Our data strongly supports the assumption of previous studies (Shafritz, Shouval et al. 1981; Hino, Kitagawa et al. 1984; Chen, Harrison et al. 1986; Chen, Harrison et al. 1988; Esumi, Tanaka et al. 1989; Mason, Liu et al. 2010) that histologically-normal hepatocytes are the cells that undergo clonal proliferation detected using integrated HBV DNA as proliferation markers. Hepatocytes were the only cell types that harboured detectable levels of virus-cell DNA junctions. Laser-microdissection was used to isolate specific cell populations other than histologically-normal hepatocytes to determine whether they contained the observed virus-cell DNA junctions. No virus-cell DNA junctions were detectable by invPCR in areas that contained fibrous tissue, inflammatory cells, portal tracts or central veins. Also, only a single unique virus-cell DNA junction was found in histologically-altered hepatocytes with LLC, suggesting that hepatocytes with normal-appearing histology are the main reservoirs of the observed virus-cell DNA junctions.

m/z = 1342.7

m/z = 1619.9

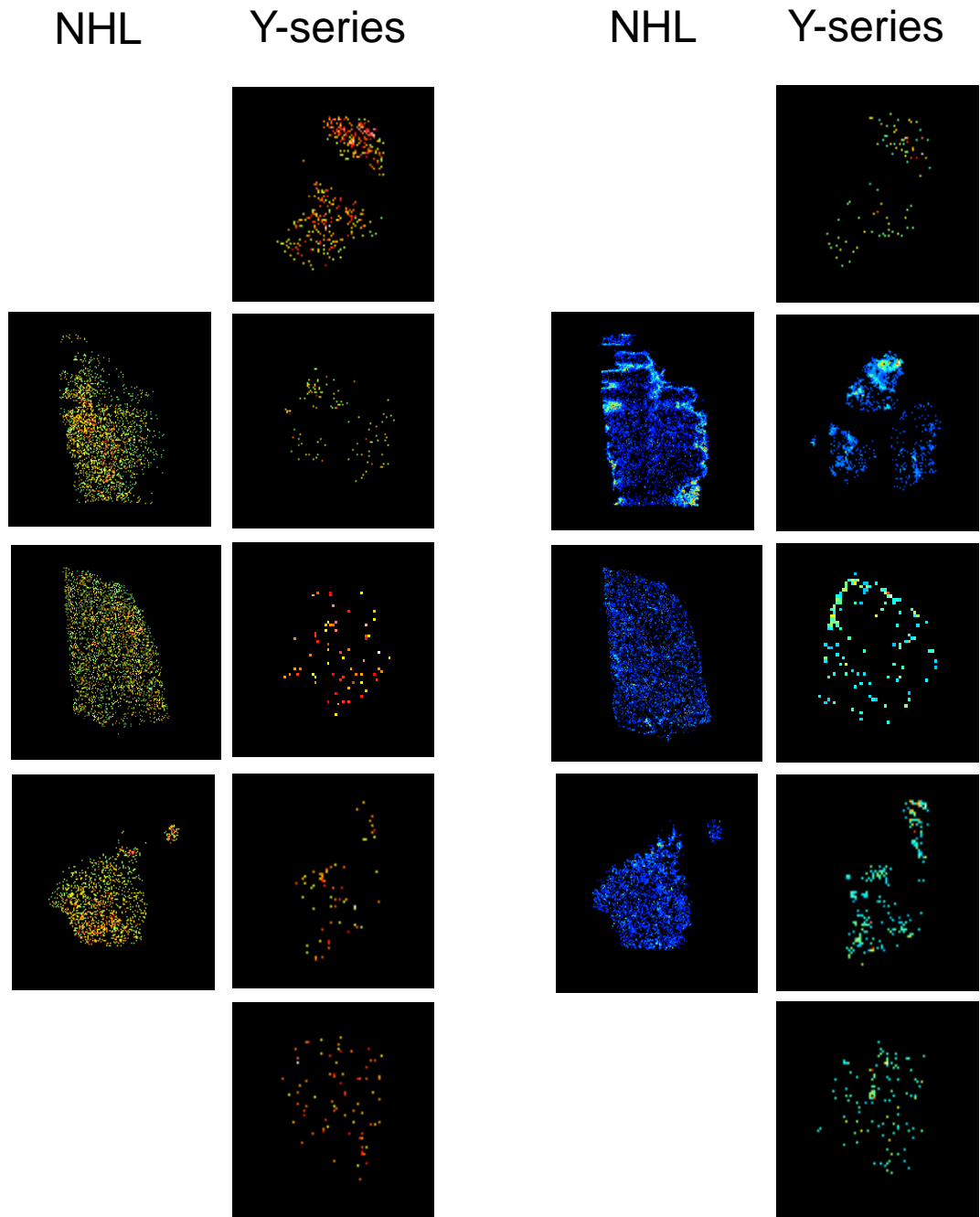


Figure 7.22. The distribution of two tryptic peptides of a probable imidazolonepropionase (7.22A, $m/z = 1342.7$) and vimentin (7.22B, $m/z = 1619.9$) as identified by HPLC-nESI-LTQ Orbitrap MS and detected by IMS in liver tissue sections of HBV-negative control patients (NHL) and chronic HBV patients (Y-series). Protein extracts from HBsAg-positive and -negative hepatocyte foci were analysed by HPLC-nESI-LTQ Orbitrap MS as described in Section 2.27. Two of the mass peaks ($m/z = 1342.7$ and 1617.9), which were detected in the majority of only HBsAg-positive hepatocyte foci (as shown in Table 7.6), were also detected by IMS analysis as described in Section 2.26. The resolution of the peptide distribution map is $100 \mu\text{m}$ between each pixel.

7.4.2 HBsAg expression and clonal proliferation

In this study, we hypothesised that hepatocytes that lack expression of HBV antigens, such as HBsAg, escape the selective pressure of HBV-specific immune responses that occur during chronic HBV infection. We therefore used invPCR to detect virus-cell junctions in foci of HBsAg-positive and -negative hepatocytes to determine their relative extents of clonal proliferation.

The virus-cell DNA junctions detected by invPCR were likely to be neutral markers for hepatocyte lineages and unlikely to be altering the phenotype of the hepatocytes. As shown in Section 7.3.5, HBV DNA integration showed no preferences for particular genes or genetic exons or introns, which was consistent with virus-cell DNA junctions found in previous Chapters (Section 5.3.8). Results of experiments detecting the integration site relative to the HBV genome and the proximity to S/MARs agreed with data described in Section 5.3.9 and in previous literature (Mason *et al.*, 2010).

As shown in Section 7.3.6, virus-cell DNA junctions were detected twice as frequently in HBsAg-negative compared to -positive hepatocytes after normalising for number of cells analysed. This variation in the number of observable virus-cell DNA junction was not likely to be due to differences in HBV dsDNA production, since HBsAg-negative hepatocytes contained only half the amount of HBV DNA per cell compared to HBsAg-positive hepatocytes (Section 7.3.6). Although virus-cell DNA junctions were detected twice as often in HBsAg-negative hepatocytes, no significant differences were observed in preferences of integrations into genes, proximity to S/MARs or integration sites with respect to the HBV genome in virus-cell DNA junctions detected in foci of HBsAg-positive hepatocytes compared to HBsAg-negative hepatocytes.

The reason for the difference in observed number of virus-cell DNA junctions is unknown, thus further investigation into the determinants of virus-cell DNA junction detection is required. In Section 5.3.6, we showed that different invPCR designs detected approximately the same number of virus-cell DNA junctions. This observation was found to contradict the RE site frequency analysis, which suggested that there should be a discrepancy in the number of observed virus-cell DNA junctions. This indicates that restriction site frequency is unlikely to be a limiting factor in virus-cell DNA junction detection. Additional factors that may lead

to the differences in the observable virus-cell DNA junctions between foci of HBsAg-positive and -negative hepatocytes were described in Section 5.3.2, including: the efficiency of restriction enzyme digestion; the circularisation efficiency of excised virus-cell DNA junctions and/or; the interference of HBV RI DNA.

Also, when comparing adjacent hepatocyte foci, it was apparent that the majority of virus-cell DNA junctions were detected exclusively in foci of either the HBsAg-positive or HBsAg-negative hepatocytes, not both. This suggested that foci of HBsAg-positive and -negative hepatocytes were composed of non-overlapping clones. An assay to detect specific virus-cell DNA junctions *in situ* would be able to more finely localise the hepatocyte clones and confirm whether this supposition is true.

When clone sizes detected in foci of HBsAg-positive and -negative hepatocytes were compared, our hypothesis that hepatocytes that lose expression of HBV antigens and escape the selective pressure of HBV-specific immune responses was not supported. As described in Section 7.3.4, no significant difference in the size of the observed hepatocyte clones between foci of HBsAg-positive and HBsAg-negative hepatocyte was detected.

However, this does not completely dismiss the hypothesis of the cessation of HBV antigen expression being a major growth advantage in the context of the HBV-infected liver. It is possible that the expression of other HBV antigens, such as HBcAg and Pol, could affect the clonal proliferation of hepatocytes. Indeed, both HBcAg- and Pol-specific CTL have been detected in the serum of HBV-infected patients (Penna, Artini et al. 1996; Mizukoshi, Sidney et al. 2004; Vassilopoulos, Rapti et al. 2008).

In Section 4.3.4, HBcAg was not detected by immunohistochemistry in the liver sections of the majority of patients with late-stage disease. Thus, any selection based on HBcAg expression may occur early in HBV infection and/or be due to such a strong selection pressure that the immunohistochemistry assay is not sensitive enough to detect a potentially lower level of HBcAg expression in the remaining hepatocytes. Indeed, several studies have shown a direct relationship between the detection of HBcAg expression in the cytoplasm of hepatocytes and the level of HBV DNA replication (Gowans, Burrell et al. 1985; Hsu, Su et al. 1987; Sari, Dere et al. 2011).

The distribution of Pol has been detected in liver cell lines by immunofluorescence (Cao and Tavis 2004) and is known to be a target of immune-mediated cell death (Mizukoshi, Sidney et al. 2004). Therefore, future studies may be conducted to determine whether the expression of HBV Pol is associated with clonal proliferation of hepatocytes.

7.4.3 Heterogeneous expression of HBsAg

Phenotypical features of HBsAg-positive and -negative hepatocytes were investigated to determine a mechanism for the observed heterogeneous HBsAg expression. In our studies, neither PreS mutation nor HBV DNA methylation were sufficient explanations for the emergence of HBsAg-positive and -negative hepatocytes. As circulating HBV virions were still present and thus all hepatocytes were still exposed to infection, the differences in HBsAg expression were presumably determined by hepatocyte-specific factors. Therefore, the sections of liver tissue were analysed by mass-spectrometry to determine whether expression of particular cellular proteins in hepatocytes were associated with HBsAg-expression.

Some differences in cellular protein expression were detected by HPLC-nESI-LTQ Orbitrap MS analyses as shown in Section 7.3.8. However, these differences could not be confirmed by IMS. As previously mentioned in Section 3.3.4, there are limitations in the detection and identification of peptides in the mass-spectrometry techniques used, in terms of both specificity and sensitivity. Furthermore, heterogeneous expression of proteins known to be differentially expressed between peri-central and peri-portal hepatocytes, such as fructose 1,6-bisphosphatase and cytochrome P450 (Puschel and Jungermann 1994; Eilers, Modaressi et al. 1995; Jungermann and Kietzmann 1996), was not observed in liver sections from either HBV-infected patients or uninfected control patients, as shown in Figure 7.18. This is consistent with the previous observation that IMS is not sensitive enough to detect some subtle changes in protein expression. Thus, the differences in expression of cellular proteins seen in the hepatocytes in Section 7.3.10 are much greater than those differences of protein expression associated with zonal heterogeneity. Due to these factors, there may be additional differences in protein expression between HBsAg-positive and -negative hepatocytes that remain undetected by the designed mass-spectrometry assays.

The cause of the heterogeneous HBsAg-expression seen remains unknown. However, the proteins that were detected exclusively in HBsAg-positive hepatocytes in the current study may provide a lead. The 18 proteins listed in Table 7.6 are candidates for future studies to investigate whether altered expression of particular cellular proteins causes altered expression or accumulation of HBsAg. Generally, the observed proteins with altered expression were enzymes involved with metabolism or catabolism. None of the proteins were reported to localise to the cell membrane. This suggests that differences in HBsAg-expression in the different foci of hepatocyte were not due to alterations in the expression of the yet-unknown HBV receptor. However, this may be due to detection bias, as membrane-bound proteins are more difficult to detect by mass-spectrometry (Santoni, Molloy et al. 2000).

7.4.4 Heterogeneous expression of cellular proteins

The yield of protein from the small mass of hepatocytes isolated by laser-microdissection was insufficient for in-depth quantitative proteomic studies, such as 2D gel electrophoresis. As such, more sensitive but limited mass-spectrometry assays were used. In previous studies, differences in protein expression between uninfected liver tissues, liver tissues with HBV-associated HCC and cirrhosis had been observed using IMS (Han, Lee et al. 2010; Le Faouder, Laouirem et al. 2011).

Our IMS studies were limited to proteins that had been detected in the majority of patient tissues and with multiple aligned peptides to increase our confidence of the distribution of cellular proteins. This limited our analysis to only 22 proteins, a tiny fraction of the total number of cellular proteins expressed by hepatocytes.

Despite this limitation, the majority of analysed cellular proteins appeared to be differentially expressed when compared in HBV-infected tissues and HBV-negative control tissues. It is unknown whether these increases in expression are due to: i) HBV directly altering the expression of cellular genes via viral components or ER stress; ii) an immunological response by the hepatocytes; iii) altered expression due to factors associated with HBV-specific disease, such as liver damage or cytokine expression by the immune system; or iv) due to selection and subsequent clonal proliferation of hepatocytes with altered expression of cellular proteins.

Some of the detected cellular proteins were heterogeneously distributed, suggesting that differences in hepatocyte phenotype occur in neighbouring foci of hepatocytes. There were no detected changes in protein expression obviously associated with HBsAg-expression or clonal proliferation detected by invPCR. This does not rule out that those hepatocytes with altered expression are not clonal. As discussed in Sections 7.4.2 and 5.3.2, clonal proliferation can only be determined by the invPCR assay if the particular hepatocytes contain detectable virus-cell DNA junctions.

While many mass peaks showed a heterogeneous distribution within the HBV-infected liver sections, no single mass peak detected by IMS appeared to be associated with clonal proliferation detected by invPCR. This observation may be an outcome of several, non-exclusive factors:

Firstly, there may be multiple survival advantages that are sustainable in the microenvironment of the HBV-infected liver. Thus, a single change in protein expression would not be expected for all hepatocyte clones composed of cells with a survival advantage. Future investigation of the effect of changing expression of the several candidate proteins identified by the mass-spectrometry analyses on hepatocyte proliferation or survival could provide insight into the range of survival advantages that hepatocytes possess.

Secondly, the invPCR assay does not give a complete picture of the amount of clonal proliferation that has occurred in the hepatocyte population. As described in Section 5.3.2., invPCR cannot detect all hepatocyte clones or even all virus-cell DNA junctions. Thus, there may be large hepatocyte clones that are missed since they do not contain a detectable virus-cell DNA junction. Additionally, proliferation of the hepatocyte clones may have occurred prior to HBV DNA integration. In this case, while invPCR may have detected only a small hepatocyte clone, the actual proliferation that the hepatocyte clone has undergone is much greater. Therefore, the invPCR assays are unlikely to detect the full extent of the clonal proliferation of hepatocytes, making association with changes in protein expression difficult. To detect a larger proportion of hepatocyte clones, future studies may use a panel of methods, such as shifts in microsatellite lengths, methylation of particular cellular gene promoters, loss of heterozygosity or mutations in mitochondrial genes (Galipeau, Li et al. 2007; Nomoto, Kinoshita et al. 2007; Fellous, Islam et al. 2009).

Also, the changes in protein expression that convey a survival advantage in the hepatocytes may be transient or occur in only specific circumstances. These could include proteins involved in responses during cell mitosis, oxidative stress, inflammation or interactions with immune-effector cells. In the absence of these situations, protein expression levels may not be different from baseline.

Furthermore, the changes in expression may not be absolute, rather relative to other cellular proteins. For example, the expression ratio of the signalling proteins Bcl-2 and Bax determine whether a cell undergoes apoptosis following cytokine deprivation in an IL-3-dependent tumour cell line (Oltvai, Milliman et al. 1993). This is due to the promotion of apoptosis by homodimers of Bax, and the promotion of cell survival by heterodimers of Bcl-2 and Bax (St Clair, Anderson et al. 1997). Similar mechanisms could be altering the phenotype of hepatocytes in the liver of patients with chronic HBV. Mathematical clustering and principal component analyses could potentially be used to determine whether altered ratios of protein expression are associated with clonal proliferation of hepatocytes (Ove Gustafsson, personal communication).

Finally, the altered protein expression may not be revealed due to the limitations in the detection of cellular peptides by the implemented mass-spectrometry techniques. These limitations have previously been discussed in Section 3.3.4. To overcome these limitations with mass spectrometry, various supplementary techniques could be used in future experiments.

For example, comparing 2D protein gel electrophoresis fingerprints of HBsAg-positive hepatocytes to HBsAg-negative hepatocyte extracts may be used to show the differences in proteins, if greater amounts of tissue could be isolated efficiently. Subsequently post-transcriptional and -translational changes in particular proteins, such as transcription factors, can be determined by more specific approaches, such as Western blot, glycosylation and phosphorylation assays.

Furthermore, RNA could be extracted from isolated hepatocytes to determine changes in transcription that are associated with clonal proliferation. Preliminary data suggests that the

patient tissues contain intact RNA, as shown in Section 3.2.3. If the extraction from hepatocytes isolated from ethanol-fixed paraffin wax-embedded tissue by laser-microdissection was optimised, subtractive RNA hybridisation or RNA expression microarray technologies could be used to detect differences in genome-wide RNA expression between foci of HBsAg-positive and -negative hepatocytes.

7.4.5 Summary

In conclusion, the observed clonal proliferation of hepatocytes occurs in hepatocytes with normal histology and independent of HBsAg expression. Expression of particular cellular proteins was associated with disease progression, despite no visible changes in hepatocyte histology. Although clonal proliferation has also been shown to be associated with disease progression in Section 5.3.5, the relationship between changes in cellular protein expression and clonal proliferation remains unknown. Furthermore, the mechanism behind the development of heterogeneous HBsAg and cellular protein expression remains unknown. However, the techniques in these studies provide a stepping-stone for future studies to determine the changes that occur in the hepatocyte population during chronic HBV infection and their role in disease progression.

8 - Concluding remarks

8.1 Summary

During chronic HBV infection, infected hepatocytes are killed by immune cells and replaced by the division of the surviving hepatocytes. While the rate of hepatocyte turnover is generally low, it still is higher than in the healthy liver (Farinati, Cardin et al. 1996). Moreover, since the liver is self-replenishing, any hepatocyte turnover will cause the liver to become more clonal over time. In our model for progression of liver disease during chronic HBV infection, some hepatocytes may have changes that confer a heritable survival advantage in the environment of the HBV-infected liver, e.g. the ability to evade immune-mediated cell death. Hepatocytes with survival advantage therefore undergo selective clonal proliferation. Clonal proliferation of histologically-normal cells has previously been associated with progression of cancer in various aetiologies, including HCC (Solt, Medline et al. 1977; Chen, Rabinovitch et al. 2005; Merlo, Shah et al. 2010). However, the existing studies on the role HBV infection in the clonal proliferation of hepatocytes have not fully characterised the involvement of clonal proliferation in progression to HCC (Hino, Kitagawa et al. 1984; Fowler, Greenfield et al. 1986; Esumi, Tanaka et al. 1989; Yasui, Hino et al. 1992; Mason, Low et al. 2009; Gong, Li et al. 2010; Mason, Liu et al. 2010).

Therefore, this project aimed to i) determine the extent of clonal proliferation of hepatocytes in liver tissues from patients with chronic HBV infection, ii) identify in liver tissue sections the subsets of cells that have undergone clonal proliferation, and iii) characterise those hepatocytes to determine the possible mechanisms behind the clonal proliferation.

Initially, we hypothesised that clonal proliferation resulted from the loss of HBV antigen expression, allowing hepatocytes to avoid immune-mediated attack. Thus, we expected hepatocytes in HBV-antigen negative foci to have undergone greater clonally proliferation than those in HBV-antigen positive foci. Furthermore, we hypothesised that the changes in gene expression and DNA methylation in HBV-antigen negative foci of hepatocytes include those previously reported to be associated with HCC initiation and progression.

To address these questions, we studied liver tissue of 30 patients in various stages of chronic HBV infection (Chapter 4). We showed that clonal proliferation does indeed occur in the

livers of patients in all stages of chronic HBV infection (Chapter 5). Computer simulations of a liver undergoing random liver turnover showed that some of the large hepatocyte clones that were detected were unlikely to be produced stochastically, i.e. without a survival advantage (Chapter 6), confirming our hypothesis that the emergence of large hepatocyte clones required a survival advantage.

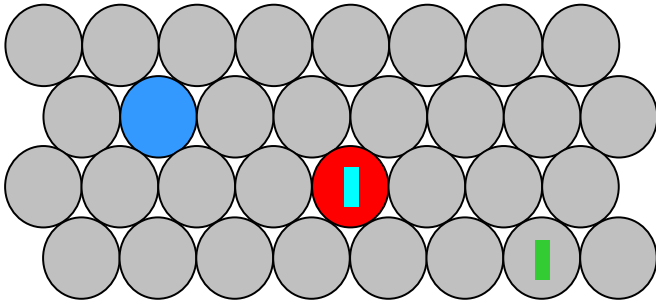
Using laser-microdissection to isolate specific subsets of hepatocytes, we detected no difference in clonal proliferation between foci of HBsAg-positive and -negative hepatocytes (Chapter 7), contrary to the hypothesis that the survival advantage is due to loss of HBV protein expression. We were unable to evaluate the role of HBcAg, which carries major targets for cytotoxic T-cells, since its expression was not detected in the majority of the tissues that were available to us.

Also, while differences in expression of cellular proteins were detected between foci of HBsAg-positive and -negative hepatocytes, these changes were not linked in any obvious way to HCC initiation or progression, contrary to our last hypothesis. However, using IMS, we detected foci of hepatocytes with altered cellular expression of HCC-associated proteins.

Thus, our data supports the model described in Figure 8.1, whereby rare hepatocytes with altered phenotypes are selected during chronic HBV infection. Hepatocytes that have altered phenotypes may have a survival advantage in the context of the HBV-infected liver. This would allow the hepatocytes to clonally proliferate to a greater extent than hepatocytes lacking a survival advantage.

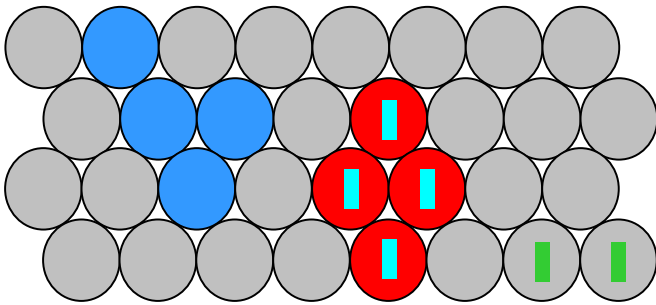
We hypothesise that the potential mechanisms for the clonal proliferation associated with disease progression observed in Chapter 5 may include: 1) that chronic HBV infection causes clonal proliferation of hepatocytes with a survival advantage thereby reducing the diversity of hepatocytes so that the effect of random DNA mutation is amplified; 2) that HBV infection produces a microenvironment that selects for a survival advantage that also causes preneoplastic changes through a common regulator; or 3) that HBV infection provides an

A



Selection by HBV infection

B

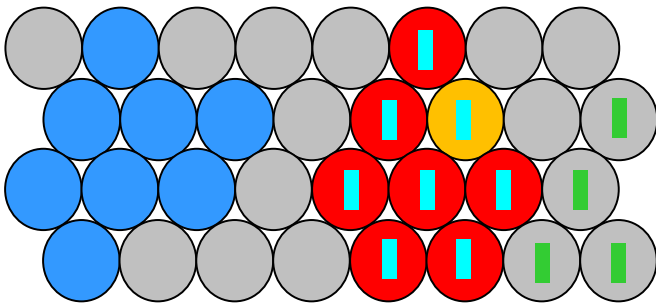


Selection for preneoplastic features?

Restricting cellular diversity?

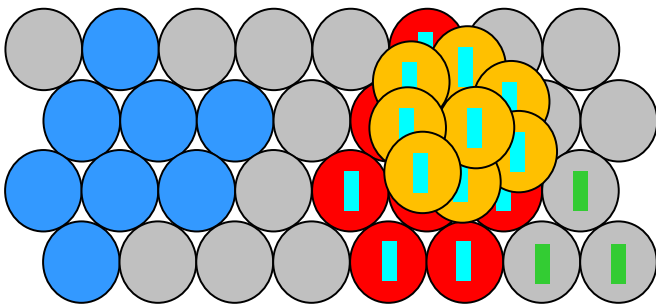
Survival advantages and preneoplastic changes governed by common regulators?

C



Preneoplasia progresses to HCC

D



Legend


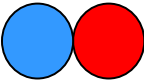


-  Infected hepatocytes
-  Infected hepatocytes with survival advantage
-  Unique virus-cell junctions
-  Preneoplastic hepatocytes

Figure 8.1. Hypothetical model for the clonal proliferation of hepatocytes and disease progression during chronic HBV infection. The figure represents a section of liver containing HBV-infected hepatocytes (circles). At the start of infection (Figure 8.1A), there are rare hepatocytes with altered phenotypes that result in a survival advantage (red and blue circles). Furthermore, some hepatocytes contain a unique virus-cell junction caused by the integration of HBV DNA (light blue and green rectangles).

As the HBV infection progresses and the liver undergoes turnover (Figure 8.1B), hepatocytes with altered phenotype clonally proliferate through compensatory mitosis to replace neighbouring hepatocytes killed by immune-mediated cell death associated with HBV infection. Through yet unknown mechanisms, hepatocytes with preneoplastic changes (yellow circles) are selected for (Figure 8.1C). Eventually, hepatocytes with preneoplastic changes progress to HCC (Figure 8.1D).

environment that actively selects for a preneoplastic cell phenotype. Further work is required to determine the relative roles of these hypothesised mechanisms on disease progression.

8.2 *Future work*

The current study has raised further questions regarding the role of clonal proliferation in HBV-associated disease progression. The pathways that should be explored include pinpointing the stage of HBV infection when clonal proliferation of altered hepatocytes begins, elucidating the phenotypic changes that cause clonal proliferation and determining which of these survival advantages influence progression to HCC.

The evolution of the hepatocyte population may be different during different stages of chronic HBV infection and between individual patients. As such, the single biopsies from HBV-infected patients studied in this thesis may be insufficient to determine the complex changes that occur during HBV-associated disease progression. Furthermore, chronic infection is often considered to have four sequential phases: immune tolerance; immune clearance; immune control; and reactivation. Most of our samples were from patients in the latter three stages.

Potentially, experiments using the techniques described in this study with a highly controlled HBV infection model, such as the humanised liver mouse model, and with multiple biopsies of the same individuals over the course of chronic HBV infection could allow us to accurately determine the changes in the hepatocyte population.

While we approximated the position of hepatocyte clones in the current study by isolating foci of hepatocytes by laser-microdissection, the exact cells that have undergone clonal proliferation were not directly observed. Single molecule detection techniques, such as *in situ* PCR (Long and Komminoth 1997) or a modified *in situ* proximity ligation assay (Fredriksson, Gullberg et al. 2002), could potentially be used to localise single copies of known virus-cell junctions in liver tissue sections. In this way, hepatocyte clones in liver sections could be finely localised and further characterised. These techniques could also be used to investigate our assumption that hepatocyte clones are approximately spherical. Furthermore, other measures of clonal proliferation, e.g. mutations in the mitochondrial DNA (as described in

Section 1.6), could be used to detect hepatocyte clones that may not have been detected by the invPCR technique. Furthermore, whole genome studies from small numbers (10-20) of cells could be used to supplement this data by using DNA mutations within the cellular genome as a measure of evolutionary distance (Peters, Kermani et al. 2012).

In this study, investigations into potential survival advantages were limited to observations of protein distribution and epigenetic factors. None of the changes in cellular protein expression detected by IMS were associated with the clonal proliferation observed by invPCR in foci of HBsAg-positive or -negative hepatocytes (Chapter 7). Though they can be powerful tools, the mass spectrometry-based approaches used in this study are somewhat limited, as previously discussed (Sections 3.3.4 and 7.4). Increasing the sensitivity of these mass spectrometry assays may reveal candidates for the survival advantages responsible for the observed clonal proliferation. If some protein candidates are found by this method, their altered levels of expression could be confirmed using immunohistochemistry, reverse transcription qPCR, and/or functional assays.

As shown in Figure 8.2, multiple, interacting factors can influence the level of functional proteins present in the cell and therefore cell phenotype. Thus, future studies could investigate the role of other cellular aspects that may confer survival advantages. This would widen the search for factors which might influence survival, and may confirm pathways which allow survival advantage by showing alterations at multiple levels of expression.

Chromosomal instability (CIN), i.e. the loss or gain of whole chromosomes or large chromosomal fragments, has been associated with the carcinogenic process. CIN could potentially alter the cellular protein levels by changing the copy number of many genes. In this study, we were unable to rule out that the occurrence of chromosomal instability was due to HBV DNA integration. In Section 5.4.2, using PCR primers specific to the expected region upstream of the HBV DNA integration event assuming that the entire HBV genome had been integrated with <1 kb deletion of the host cell genome, no product could be amplified. This suggests that major rearrangement of the chromosome had occurred during integration of the HBV genome. Chromosome copy number analyses, such as analysis of whole-genome

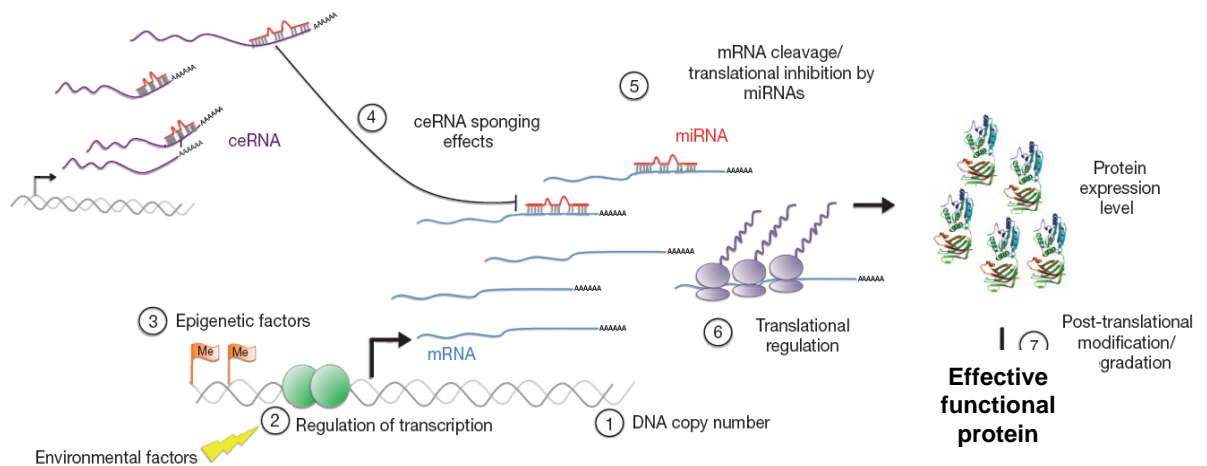


Figure 8.2. Multiple regulatory factors that affect the levels of functional protein in liver cells. Pre-transcriptional changes, such as change in chromosome number (1), regulation of transcription based on external stimuli (2) or epigenetic factors (3) can affect the mRNA copy number of the transcribed gene. Post-transcriptional factors, such as degradation by miRNA (5), decrease in miRNA levels by competing endogenous RNA (ceRNA) transcripts (4) and translational regulation (6), can all affect the number of mRNA molecules translated into protein. Multiple post-translational modifications (7), such as phosphorylation, proteolysis, and glycosylation, can affect the function of the final protein. Thus, the level of effective functional protein levels is sensitive to multiple regulatory elements. Figure reproduced with permission from Berger *et al.* (2010).

shotgun sequence data (Sudmant, Kitzman et al. 2010), could provide some insight into mechanisms responsible for clonal proliferation.

Also, small deletion mutations in the host-cell genome, termed minimally deleted regions (MDR), have previously been associated with the progression to HCC (Jou, Lee et al. 2004). Potentially, these same MDR could induce clonal proliferation in the HBV-infected hepatocyte population. Even small hepatocyte foci (10-20) isolated by laser-microdissection could be assayed by whole genome sequencing for evidence of these changes in host-cell DNA (Peters, Kermani et al. 2012).

RNA microarray studies were not performed in these studies. However, if high quality RNA could be extracted from the isolated hepatocytes that had undergone clonal proliferation, microarray studies for mRNA and miRNA could be used to determine transcriptional changes in hepatocytes that had undergone clonal proliferation compared to those that had not.

miRNAs are short (~22 nt) RNA transcripts that silence gene expression in a post-transcriptional manner by binding to complimentary mRNA sequences in the 3' untranslated region and inducing mRNA degradation (reviewed in (He and Hannon 2004)). Changes in miRNA expression have been associated with HCC development and HBV infection *in vitro* (Zhang *et al.*, 2010). Since a single miRNA can have several mRNA targets (John, Enright et al. 2004; John, Sander et al. 2006), changes in miRNA expression during chronic HBV infection may affect protein networks associated with hepatocarcinogenesis.

Potential complications in the identification of some survival advantages may arise from the possibility that survival advantages are context-specific. Since the microenvironment of the HBV-infected liver is likely to change during different stages of HBV infection (Hui, Leung et al. 2007), cellular changes that confer a survival advantage in one stage of infection may be useless in another stage. Also, the expression of genes expressing proteins change in response to external stimuli, such as growth signals, death signals, and nutrient levels; all of which change over time. Survival advantages that are only active in specific or transient conditions may be missed by these assays.

Furthermore, linking the clonal proliferation of hepatocytes with altered phenotype is difficult using the methods of the present study, in which fixed liver tissues were used to isolate hepatocytes, identify hepatocyte clones and characterise cellular phenotypes. The fixed cells were unable to be cultured to determine complete phenotype and interactions with other cells. To confirm whether a particular altered phenotype enhances the ability of the hepatocytes to survive and proliferate, animal or *in vitro* models of HBV infection could be used. It is possible that siRNA or retroviral vectors could be used to change levels of protein expression in hepatocyte cell lines or animal hepatocytes, allowing investigation of potential survival advantages. Potential studies include: cytotoxicity assays to determine sensitivity to immune-mediated cell death; BrdU incorporation or immunohistochemistry for PCNA to determine whether altered cells have increased rates of proliferation; and stimulation with mitogens to determine whether the cells have a hypersensitivity to growth signals.

Detecting the link between clonal proliferation and disease has implications on clinical treatment of chronic HBV infection. If clonal proliferation does indeed induce progression to HCC, then anti-viral treatment should be carried out prior to the stage of infection when most of the clonal proliferation occurs in order to decrease the incidence of late-stage disease. A potential route for treatment could be altering the microenvironment of the liver to dissuade the selection for certain pre-neoplastic changes (Merlo, Pepper et al. 2006).

8.3 *Impact on other fields of study*

The approaches used in this study could potentially be used broadly:

The invPCR technique can potentially be used to investigate the recurrence of HCC after surgical resection, or determining whether multiple HCC tumours represent metastases. If the virus-cell junctions in individual tumours in a given patient are identical, then this suggests that either a certain hepatocyte sub-population with pre-neoplastic changes exists within the liver or that the primary tumour has metastasised. Furthermore, if particular peptides could be associated with preneoplastic changes, then mass-spectrometry analyses could provide a quick and simple identification of hepatocyte sub-populations in liver tissue resections that are at risk of causing tumour formation.

Some methods from the current study could also be used to examine the initiation and progression of HCV-associated carcinogenesis. Carcinogenesis HCV infection may be different to that in HBV infection. However, the hepatocyte population in chronic HCV infections must still adhere to evolutionary principles. Thus, investigations into the HCV-associated carcinogenic process in terms of clonal proliferation of hepatocytes with growth advantage are warranted.

If greater sensitivity could be achieved in detecting peptides by imaging mass-spectrometry (IMS), then the distribution of endogenous proteins can be mapped in liver sections. While it is known that some metabolic enzymes are heterogeneously distributed in a liver lobule depending on distance from the portal tract (Jungermann 1988; Eilers, Modaressi et al. 1995; Jungermann and Kietzmann 1996; Lindros 1997), a more comprehensive list could be created using IMS. In this way, the changes in endogenous protein distribution in the liver could also be observed in various disease states, which could be used to investigate the mechanisms of pathogenesis, course of disease and effect of potential treatments.

8.4 Conclusion

In conclusion, our data suggests that more extensive clonal proliferation of hepatocytes has occurred in the livers of patients with late-stage disease compared to those from patients in early-stage immune tolerance. Furthermore, our results strongly suggest that clonal proliferation of hepatocytes with a survival advantage has occurred in the livers of chronic HBV patients with late-stage disease. The mechanisms that allow these survival advantages remain unclear. Contrary to our hypothesis, loss of HBsAg expression does not appear to be significantly associated with clonal proliferation, suggesting that the major selective pressure during chronic HBV infection is not immune-mediated cell targeting of HBsAg-expressing hepatocytes. While altered cellular protein expression were observed between hepatocytes in HBsAg-positive and HBsAg-negative foci, the altered phenotype did not appear to be a preneoplastic phenotype. This study has provided a stepping stone in determining the complex interactions that occur within the hepatocyte population over the decades of a chronic HBV infection. Further investigations may provide insight into chronic HBV infection-associated pathogenesis and potential strategies to more efficiently prevent end-stage disease.

Chapter 9 - Appendices

1.1 Sequence of pBB4.5HBV1.3 (9211 bp)


AGCTGGTACCCGGGTCTAGAATTCGAAGCTTGGAGTCGACAACCTGTTTATTGTCAGCTTATAATGGTTACAAATAAAGCAATA
GCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTTGCCAAACTCATCAATGTATCTTATCAT
CTCTCGACTCTGCTGAAGAGGAGGAAATTTCTCCTTGAAGTTTCCCTGGTGTTCAAAGTAAAGGAGTTTGCACCAGACGCACCT
CTGTTCACTGGTCCGGCGTATTTAAAACACGATACATTGTTATTAGTACATTTATTAAGCGCTAGATTCTGTGCGTTGTTGATT
TACAGACAATTTGTGTACGTATTTAATAATTCATAAATTTATAATCTTTAGGGTGGTATGTTAGAGCGAAAATCAAATGAT
TTTCAGCGTCTTTATATCTGAATTTAAATATTTAAATCCTCAATAGATTTGTAAAAATAGGTTTTCGATTAGTTTCAAACAAGGGT
TGTTTTTCCGAACCGATGGCTGGACTATCTAATGGATTTTCGCTCAACGCCACAAAACCTTGCCAAATCTTGTAGCAGCAATCT
AGCTTTGTGATATTCGTTTGGCTTTTGGTTTTGTAAATAAAGGTTGCAGCTCGTCAAATATTATAGCCTTTTGTATTTCTTT
CATCACTGTGCTTAGTGTACAATTGACTCGACGTTAAACACGTTAAATAAAGCCTGGACATATTTAACCATCGGGCGTGTAGCT
TTATTAGGCCGATTTATCGTCTGTCGCCAACCCCTCGTCTTAGAAGTTGCTTCCGAAGACGATTTTGCATAGCCACACGACG
CCTATTAATTTGTGTCGGCTAACACGTCGCGCATCAAATTTGTAGTTGAGCTTTTTTGAATTTATTTCTGATTGCGGGCGTTTTT
GGCGGGTTTTCAATCTAACTGTGCCCGATTTTAAATTCAGACAACACGTTAGAAAGCGATGGTGCAGCGCGTGGTAAACATTTCA
GACGGCAAATCTACTAATGGCGCGGTGGTGGAGCTGATGATAAATCTACCATCGGTGGAGGCGCAGGCGGGGCTGGCGGCGG
AGGCGGAGGCGGAGGTGGTGGCGGTGATGCAGACGGCGGTTTAGGCTCAAATTTGTCTCTTTCAGGCAACACAGTCGGCACCTC
AECTATTGTACTGGTTTCGGGCGTATGGTGCACCTCTCAGTACAATCTGCTGATGCCGCATAGTTAAGCCAGCCCCGACACC
CGCCAAACCCCGTGACGCGCCCTGACGGGCTTGTCTGCTCCGCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGC
TGATGTGTGACAGGTTTTTACCCTCATCACCGAAACGCGCGAGACGAAAGGGCCCTGTGATACGCTATTTTTATAGGTTAA
TGTCATGATAAATATGGTTTTCTTAGACGTCAGGTGGCACTTTTTCGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCT
AAATACATTTCAAATATGTATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAAATAATTTGAAAAAGGAAGATGATGAGT
ATTCAACATTTCCGTGTGCGCCCTATTCCCTTTTTTGGCGCATTTTGCCTTCCTGTTTTTGTCTACCCGAAACGCTGGTGAA
AGTAAAGATCTGAAGATCAGTTGGGTGCACGAGTGGTTACATCGAATGGATCTCAACAGCGGTAGAGATCTTTGAGAGTT
TTCGCCCCGAAGAACGTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACCCGGG
CAAGAGCAACTCGGTGCGCGCATACTACTATTCTCAGAATGACTTGGTTGAGTACTCACAGTACAGAAAAGCATCTTACGGA
TGGCATGACAGTAAGAGAATTATGCAGTGTGCCATAACCATGAGTATAACACTGCGGCCAATTTACTTCTGACAACGATCG
GAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCTTGTATCGTTGGGAACCGGAGCTGAAT
GAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAAT
ACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGAGGCGGGATAAAGTTGCAGGACCCTTCTGCGCTCGGCCCTTC
CGCTGGCTGGTTTTATTTGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCGCGGTATCATTTGAGCACTGGGGCCAGATGGT
AAGCCCTCCCGTATCGTAGTTATCTACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGG
TGCTCACTGATTAAGCATTGGTAACTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAAT
TTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACTGAGTTTTTGTTCCTACTGAGCGTCA
GACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAACTGCTGCTTGCAAAACAAAAAACCACC
GCTACCAGCGGTGGTTTTGTTTCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTACAGCAGAGCGCAGATAC
CAAATAGCTTCTCTAGTGTAGCCGTAGTTAGCCACTTCAAGACTCTGAGCACCCTACATACCTCGCTCTGCTA
ATCTGTTACCAGTGGCTGTGCCAGTGGCGATAGCTGTCTTACCAGGTTGGACTCAAGACGATAGTACCAGGATAAAGGC
GCAGCGGTGCGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAATGAGATACCTACAGC
GTGAGCTATGAGAAAAGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGTCGGAACAGGAGAG
CGCAGGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCTCGGCCTTTTGTGCTGGC
CTTTGTCTACATGTTCTTCCGTGTTATCCCTGATTTCTGTGGATAACCGTATTACCGCCTTTGAGTGAAGCTGATACCCGCT
CGCCGACCGAAGCAGCGAGCGAGCTCAGTGCAGGAGGATACCTGACCACTCTGCTCATGACCTGATGACCTGACCTGAC
ATGCAGAGGATGATGCTCGTGACGGTTAACGCCCTGAATCAGCAACGCGCTTGCCTTACAGCAGCAGACCAATTTTCAATCC
GCACCTCGCGGAAACCGACATCGCAGGCTTCTGCTTCAATCAGCGTGCCTGCGCGGTGTGCAGTTCAACCACCGCACGATAG
AGATTCCGGATTTCCGGCTCCACAGTTTCCGGTTTTTGCAGTTTACAGCTAGTGTGACGCGATCGGTATAACCACCAGCTC
ATCGATAATTTACCCGCCGAAAGCGCGGTGCCGTGGCGACCTGCGTTTTACCCCTGCCATAAAGAACTGTTACCCGTAGGT
AGTCACGCAACTCGCCGCACATCTGAACCTCAGCTCCAGTACAGCGCGGCTGAAATCATATTAAGCGAGTGGCAACATGG
AAATCGTGTATTTGTGTAGTGGTTTTATGCAGCAACGAGACGTCAGGAAATGCCGCTCATCCGCCACATATCTGATCTTC
CAGATAACTGCCGTCACTCCAAGCAGCACCATCACCGCGAGGCGTTTTTCTCCGCGCGTAAAAAATGCGCTCAGGTCAAAT
CAGACGGCAAACGACTGTCTTGGCCGTAACCGACCCAGCGCCGTTTGCACCACAGATGAAACGCCGAGTTAACGCCATCAAAA
ATAATTCGCGTCTGGCTTCTGTAGCCAGCTTTTATCAACATTTAAATGTGAGCGAGTAAACAACCCGTCGGATTTCTCCGTGGG
AACAAACGGCGGATTTGACCGTAATGGGATAGGTACGTTGGTGTAGATGGGCGCATCGTAACCGTGATCTGCCAGTTTGGAGG
GGACGACGACGATTCGGCCTCAGGAAGATCGCACTCCAGCCAGCTTTCGGCACCGCTTCTGGTGGCGGAAACCGGCAAAAG
CGCATTTCGCCATTCAGGCTGCCAACTGTTGGGAAGGGCGATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAAGG
GGATGTGCTGCAAGCGATTAAGTTGGGTAACGCCAGGTTTTTCCAGTCACGAGCTGTAAACAGCAGGATCTATCATTT
TTAGCAGTATTCTAATTGCAGCTGCTCTTTGATACAATAATTTTACGACGACGATGCGAGCTTTTATTCAACCGAGCGTGC
ATGTTTGAATCGTGCAAGCGTTATCAATTTTTTCAATATCGTATTGTTGCACATCAACAGGCTGGACACCAGTTGAACTCGC
CGCAGTTTTTGGCGAAGTTGGACCCGCCGCGCATCCAATGCAAACCTTCCGACATTTCTGTTGCCTACGAACGATTGATCTTT
GTCCATTGATCGAAGCGAGTGCTTCCGACTTTTTCGTGTCCAGTGTGGCTTGATATCATGGAGATAATAAAATGATAACCAT
CTCGCAAATAAATAAGTATTTTACTGTTTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAATATTCGGGATTTTACATACCG
TCCACCATCGGGCTGTCTAGCGGATCCGAGCTCGCCGGGACTCTGCTACTGAGGAAAGGACTCAGAAGGCAAAAACGAGA
GTAACCTCCAGCTAGCTCCAAATCTTTATAAAGGTCGATGTCCATGCCCCAAAGCCACCAAGGCAAGCTTGGAGGCTTGA
ACAGTAGGACATGAACAAGAGATGATTAGGCAGAGGTGAAAAAGTTGCATGGTGTGGTGGCGCAGACCAATTTATGCCATACAG
CCTCCTAATACAAAGACCTTTAATCTAATCTCCTCCCAACTCCTCCAGTCTTTAAACAAACAGTCTTTGAAGTATGCCTC
AAGGTCCGTGCTTGCATTTACAGAGAGTCCAAGATCCTCTTATGTAAGACCTTGGCAAGAATGGTTGGGCTTTTACGGTGG
TCTCCATGCGAGCTGCAGAGGTGAAGCGAAGTGCACACGGTCCGCGAGATGAGAAGGCACAGACGGGAGTCCGCGTAAAGAG
AGGTGCGCCCGTGGTGGTGGTGAACGGCAGACGGAAAGGGGACGAGAGTCCCAAGCGACCCCGAGAAGGTCGTCCGCG
GGATTACGCGCCGACGGACGTAACAAGGACGTCGCCGCGAGGATCCAGTTGGCAGCAGCTAGCAGCTAGCAGCTAGGATACGA
TGTATATTGCGGGAGAGAACAACAGAGTTATCCGTCCCGAGAATGTTTGTCTCCAGACTGCTGCGAGCAAAAACAAGCGGCTA
GGAGTTCCGCGATGATGGATCGGCAGACGAGCCAGAAAGGTTCCACGCTATGCGCTGATGGCCATGACCAAGCCCCAGCCAGTG
GGGTTGCGTACGAAACACTTGGCACAGACCTGGCCGTTGCCGGCAACGGGTAAGGTTTCAAGTATTGTTTACACAGAAA


```

GGCCTTGTAAGTTGGCGAGAAAGTAAAAGCCTGCTTCGACTGAATACATGCATACAAAGGCATTAAAGCAGGATAACCACATT
GTGTAAAAGGGGCGAGCAAACCCAAAAGACCCACAATACGTTGACAGACTTTCCAATCAATAGGCCGTGTTAACAGGAAGTTTT
CTAAAGCATTCTTTGATTTTCTGTATGATGTGATCTTGTGGCAATGACCCATAACATCCAATGACATAGCCCATGAAATGTAA
AGAGTAACCCCATCTTTTTGTTTTGTTAGGGTTTTAAATGTATACCCAAAGACAGAAGAAAATTGGTAAACAGCGGTAAAAGGG
ACTCAAGATGCTGTACAGACTTGGCCCCAATACCACATCATCCATATAACTGAAAGCCAAACAGTGGGGGAAAGCCCTACGA
ACCCTGAACAAATGGCACTAGTAAACTGAGCCAGGAGAAACGGGCTGAGGCCCACTCCCATAGGAATTTTCCGAAAGCCCAG
GATGATGGGATGGGAATACAGGTGCAATTTCCGTCCGAAGGTTTGGTACAGCAACAGGAGGGATACATAGAGGTTCCCTTGAGC
AGGATCGTGCAGGTTCTGCATGGTCCCCTGCTGGTAGTTGAAGATCCTGGAATTAGAGGACAAACGGGCAACATACCTTGAT
AGTCCAGAAGAACAACAAGAAGATGAGGCATAGCAGCAGGATGAAGAGGAAGATGATAAAACGCCGACAGACACATCCAGCGA
TAACCAGGACAAGTTGGAGGACAGGAGGTTGGTGAAGTATTGGAGGTTGGGGACTGCGAATTTTGGCCAAGACACACGGTAGT
TCCCCCTAGAAAATTGAGAGAAGTCCACCACGAGTCTAGACTCTGCGGTATTGTGAGGATTCTTGTCAACAAGAAAAACCCG
CCTGTAACACGAGCAGGGGTCCTAGGAATCCTGATGTGATGTTCCATGTTCCAGCGCAGGGTCCCAATCCTCGAGAAGATT
GACGATATGGGAGAGACAGTAGTCGGAACAGGGTTACTGTTCCGAACTGGAGCCACCAGCAGGGAAAATACAGGCCTCTCAC
TCTGGGATCTTGACAGGTTTGGTGAAGGTTGGGAGTTCACATGCATGGCCTGAGGATGAGTGTCTCAAAGGTGGAGACA
CGGGGTAGGCTGCCTTCCCTGACTGGCGATTGGTAGAGGCAGGAGGCGGATTTGCTGGCAACGTTTCTAGTATGCCCTGAGCC
TGAGGGCTCCACCCAAAAGGCCCTCCGTGTGGTGGGTGAATCCAGCCGAATGCTCCAGCTCCTACCTTGTGGCGTCTGG
CCAGGTGTCTTGTGGGATTGAAGTCCCAATCTGGATTTGCGGTGTTGCTCTGAAGGCTGGATCCAACCTGGTGGTCCGGAA
AGAATCCAGAGGATTGCTGGTGGAAAGATTTCGCCCATGCTGTAGCTCTTGTCCCAAGAATATGGTGACCCACAAAATGA
GCGCTATGTGTAGTTTCTCTCTGTATAATATACCCGCCTTCCATAGAGTGTGTAATAATGTCTAGTTTGGAAAGTAATGAT
TAACTAAATATCTGGATAATAAGGTTTAATACCTTATCCAATGGCAAATATTTGGTAAACATTTGGATAAAAACCTAGCTGGC
ATAATCAATTGCAGTCTTCTTTCTCATTGACTGTGAGTGGGCCACAAATTGTTACATTTTTTGATAATGTCTTGGTGTAA
ATGTATATTAGGAAAAGAGGGTGTTTTCCAATGAGGGTTAAAGACAGGTACAGTAGAAGAATAAAGCCCCGTAAAGTTTCCCA
CCTTATGAGTCCAAGGAATACTAACATTGAGATTCCCAGATTGAGATCTTCTGCGACGCGCGGATTGAGACCTTCGCTGCG
AGGCAGGGGAGTTCTTCTTAGGGACCTGCCTCGTCTAACAACAGTAGTCTCCGGAAGTGTGATAAGATAGGGGCAT
TTGGTGGTCTATAAGCTGGAGGAGTGCGAATCCCACTCCAAAAGACACCAAATACTCTATAACTGTTTCTTCCAAAAGTG
AGACAAGAAATGTGAAACCACAATAGTTGCCTGAATTTTAGGCCATATTAGTGTGACATAACTGACTACTAGGTCCTGGGA
TATTGGATCTTCCAAATTACCACCACCCAGGTAGCTAGAGTCATTAGTTCCCCCAGCACAGTATTGCTTGCCTGAGTGCAG
TATGGTGAGGTGAACAATGCTCAGGAGACTTAAGGCTTCCCGATACAGAGCTGAGGCGTTATCTAGAAGATCTCGTACTGAA
GGAAAGAAGTCAAGGCAAAAACGAGAGTAATCCACAGTAGTCCAAATCTTTATAAGGATCAATGTCCATGCCCAAAG
CCACCCAAAGGCACAGCTTGGAGGCTTGAACAGTAGGACATGAACAAGAGATGATTAGGCAGAGGTGAAAAGTTGCATGGTGC
TGGTGGCAGACCAATTTATGCCTACAGCCTCCTAATACAAAGACCTTTAATCTAATCTCCTCCCCAACTCCTCCCAGTCTT
TAAACAAACAGTCTTTGAAGTATGCCTCAAGGTCGGTTCGTTGACATTACAGAGAGTCCAAGAGTCCCTTATGTAAGACCTTG
GGCAAGAATGGTTGGGCTTTCACGGTGGTCTCCATGCGACGTGCAGAGGTGAAGCGAAGTGCACACGGTCCGGCAGATGAGAA
GGCACAGACGGGAGTCCGCGTAAAGAGAGGTGCGCCCGTGGTCCGTCGAAACGGCAGACGGAGAGGGGACGAGAGAGTCC
CAAGCGACCCCGAGAAGGGTTCGTCGCGGGATTACGCGCCGACGGGACGTAAACAAGGACGTCCCGCGCAGGATCCAGTTGG
CAGCACAGCCTAGCAGCCATGGATACGATGTATTTGCGGGAGAGAACAACAGAGTTATCCGTCCGAGAATGTTTGTCTCA
GACCTGCTGCGAGCAAAAACAAGCGGCTAGGAGTTCCGCGATGATGGATCGGCAGACGAGCCAGAAAGGTTCCACGCATGCGCTG
ATGGCCCATGACCAAGCCCCAGCCAGTGGGGGTTGCGTCAGCAACACTTGGCACAGACCTGGCCGTTGCCGGGCAACGGGGT
AAAGGTTACAGGTATTGTTTACACAGAAAGGCCCTGTAAGTTGGCGAGAAAGTAAAAGCCTGCTTCGACTGAATACATGCATAC
AAAGGCATTAAAGCAGGATAACCACATTGTGTAAGGGGCGAGCAAAACCCAAAAGACCCACAATACGTTGACAGACTTTCCA
ATCAATAGGCCGTAAACAGGAAGTTTTCTAAAGCATTCTTTGATTTTCTGTATGATGTGATCTTGTGGCAATGACCCAT

```

Where,  is the HBV 1.3mer insert (genotype D) Genbank number AF305422

 is the pTT expression vector

 is the pBlueBac4.5 expression vector

9.2 Melting temperature of DIG-labelled hybridisation probe

Previous studies have shown the melting temperature of DNA:DNA hybrids formed between DNA probes and specific sequences is shown by the following equation (McConaughy, Laird et al. 1969; Persson, Perricaudet et al. 1979);

$$T_m = 16.6 \log [Na^+] + 0.41[\%G + C] + 81.5 - 0.72[\%FA]$$

Where, T_m is the melting temperature of the probe in °C

$[Na^+]$ is the Sodium ion concentration in Mol/L

$[\%G+C]$ is the GC content of the probe expressed as a percentage

And $[\%FA]$ is the formamide concentration in the final probe mix expressed as a percentage

For the DIG-labelled probe in the hybridisation mix during the hybridisation step,

$$[Na^+] = 0.825M, [\%G+C] = 50\%, \text{ and } [\%FA] = 50\%$$

$$\therefore T_m = 16.6 \log(0.825) + 0.41(50) + 81.5 - 0.72(50)$$

$$\therefore T_m = 64.6^\circ C$$

Hybridisation occurred at 34°C, i.e. at $T_m - 30.6^\circ C$

For the high stringency wash in 0.1x SSC, 0.1% SDS

$$[Na^+] = 0.0165M, [\%G+C] = 50\%, \text{ and } [\%FA] = 0\%$$

$$\therefore T_m = 16.6 \log(0.0165) + 0.41(50) + 81.5 - 0.72(0)$$

$$\therefore T_m = 72.4^\circ C$$

Washes were conducted at 58°C, i.e. at $T_m - 14.4^\circ C$

9.3 Efficiency of self-ligation of excised virus-cell junctions in the invPCR protocol

The probability of ring closure of linear DNA during ligation is governed by the ratio of factors j and i (Dugaiczky, Boyer et al. 1975), where j is the effective concentration of one end of a DNA molecule in the neighbourhood of the other end of the same molecule in ends/mL and i is the total concentration of DNA ends within a given solution in ends/mL. The probability of circularisation can be shown as simply j/i . When $j/i = 1$, equal products of circular and linear ligations are expected. When $j/i > 1$, self-ligated circular forms are favoured and when $j/i < 1$, intermolecular linear ligation is preferred. Dugaiczky *et al.* (1975) demonstrated that covalently closed circular structures of DNA could be found only when $j/i < 2-3$ (Dugaiczky, Boyer et al. 1975).

9.3.1. Assumptions:

- All excised virus-cell junctions are 1000bp (1)
- 1.75ug of DNA is present in the ligation reaction (2)
- Cellular DNA has a random distribution of bases (3)
- First cut restriction enzyme is a 4 base cutter (4)
- 0.4mL ligation volume (5)

9.3.2. Calculation of factor j

From Dugaiczky *et al.* (Dugaiczky, Boyer et al. 1975), $j = j_{\lambda} \left(\frac{MW_{\lambda}}{MW_{vc}} \right)^{3/2}$ (6)

Where, j is the effective concentration of one end of a DNA molecule in the neighbourhood of the other end of the same molecule in ends/mL

j_{λ} is the factor j for phage λ DNA, calculated and confirmed through experiment to be 3.6×10^{11} ends/mL (Wang and Davidson 1966)

MW_{λ} is the molecular weight of phage λ DNA, shown to be 30.8×10^6 g/mol (Wang and Davidson 1966)

And, MW_{vc} is the molecular weight of the DNA to be ligated in g/mol.

$$MW \approx (l \times 607.4) + 157.9 \quad (7)$$

Where, l is length of DNA in nucleotides

$$\begin{aligned} \text{From (1) \& (7), } MW_{vc} &\approx (1000 \times 607.4) + 157.9 \\ &= 6.08 \times 10^5 \text{ g/mol} \end{aligned} \quad (8)$$

$$\begin{aligned} \text{From (6) \& (8), } j &= 3.6 \times 10^{11} \left(\frac{30.8 \times 10^6}{6.08 \times 10^5} \right)^{3/2} \\ &= 1.30 \times 10^{14} \text{ ends/mL} \end{aligned} \quad (9)$$

9.3.3. Calculation of factor i

From Dugaiczky *et al.* (Dugaiczky, Boyer *et al.* 1975),

$$i = 2N_0 M \times 10^{-3} \text{ ends/mL} \quad (10)$$

Where, i is the total concentration of DNA ends within a given solution in ends/mL

N_0 is Avogadro's constant = $6.022 \times 10^{23} \text{ mol}^{-1}$

And, M is the molar concentration of all DNA molecules within the solution in mol/L

$$M = F/v \quad (11)$$

Where, F is the number of DNA fragments in solution in moles

And, v is the ligation volume in L. From (5), $v = 0.0004\text{L}$

Assume that the cellular DNA in solution is a single strand that is cut by the restriction enzyme.

$$F \approx cf/N_o \quad (12)$$

Where, c is the length of DNA in ligation solution in nucleotides

And, f is the frequency the restriction enzyme cuts in cuts per nucleotide. From (3) and (4), $f = 1/(4^4)$

$$c = (m/MW_{nt}) \times N_o$$

Where, m is the mass of DNA in solution. From (2), $m = 1.75\mu\text{g} = 1.75 \times 10^{-6} \text{g}$

MW_{nt} is the molecular weight of a single base in g/mol (on average, $MW_{nt} = 607.4 \text{g/mol}$)

$$\begin{aligned} c &= (1.75 \times 10^{-6} / 607.4) \times 6.022 \times 10^{23} \\ &= 1.73 \times 10^{15} \text{nt} \end{aligned} \quad (13)$$

From (12) & (13),

$$\begin{aligned} F &\approx 1.73 \times 10^{15} \times (1/(4^4)) / 6.022 \times 10^{23} \\ &= 1.12 \times 10^{-11} \text{mol} \end{aligned} \quad (14)$$

From (11) & (14),

$$\begin{aligned} M &= 1.12 \times 10^{-11} / 0.0004 \\ &= 2.81 \times 10^{-8} \text{mol/L} \end{aligned} \quad (15)$$

From (10) & (15),

$$\begin{aligned} i &= 2 \times 6.022 \times 10^{23} \times 2.81 \times 10^{-8} \times 10^{-3} \\ &= 3.38 \times 10^{13} \text{ends/mL} \end{aligned} \quad (16)$$

9.3.4. Calculation of j/i

From (9) & (16),

$$\begin{aligned}j/i &= 1.30 \times 10^{14} / 3.38 \times 10^{13} \\ &= 3.84 \\ &> 1\end{aligned}$$

[QED]

Therefore, with current assumptions, circularisation of DNA fragments is preferable to linear ligation during the inversion reaction.

Appendix 9.4. Copies of Ethics Approval forms
Appendix 9.4.1. Royal Adelaide Hospital HREC approval for Protocol #070301



Government of South Australia
Central Northern Adelaide
Health Service

**ROYAL ADELAIDE
HOSPITAL**

North Terrace,
Adelaide, SA 5000
Tel: +61 8 8222 4000
Fax: +61 8 8222 5939
ABN 80 230 154 545
www.rah.sa.gov.au

Research Ethics Committee

Level 3, Hanson Institute
Tel: (08) 8222 4139
Fax: (08) 8222 3035
Email: hodea@mail.rah.sa.gov.au

5 March 2007

Dr Allison Jilbert
Head, Hepatitis Virus Research,
Infectious Diseases Laboratories
IMVS

Dear Dr Jilbert,

Re: "Clonal proliferation of hepatocytes in progression of liver disease in chronic hepatitis B virus infection." Patient Information Sheet & Consent Form.
RAH PROTOCOL NO: 070301.

I am pleased to advise that Research Ethics Committee approval has been given to the above project. Please quote the RAH Protocol Number allocated to your study on all future correspondence.

Research Ethics Committee deliberations are guided by the NHMRC National Statement on Ethical Conduct in Research Involving Humans.

The general conditions of approval follow:

- Adequate record-keeping is important. If the project involves signed consent, you should retain the completed consent forms which relate to this project and a list of all those participating in the project, to enable contact with them in the future if necessary. The duration of record retention for all research data is 15 years.
- You must notify the Research Ethics Committee of any events which might warrant review of the approval or which warrant new information being presented to research participants, including:
 - (a) serious or unexpected adverse events which warrant protocol change or notification to research participants,
 - (b) changes to the protocol,
 - (c) premature termination of the study,
 - (d) completion of the study with a study completion summary.
- The Committee must be notified within 72 hours of any serious adverse event occurring at this site.
- Approval is ongoing, subject to satisfactory annual review.

Yours sincerely,

pe

Dr M James
CHAIRMAN
RESEARCH ETHICS COMMITTEE

Appendix 9.4.2. NSW Health HREC approvals for Protocols #X10-0072 and #HREC/10/RPAH/130

ADDRESS FOR ALL CORRESPONDENCE
RESEARCH DEVELOPMENT OFFICE
LEVEL 3, BUILDING 92
ROYAL PRINCE ALFRED HOSPITAL
CAMPERDOWN NSW 2050

TELEPHONE: (02) 9515 6766
FACSIMILE: (02) 9515 7176
EMAIL: lesley.townsend@email.cs.nsw.gov.au
REFERENCE: X10-0072 & HREC/10/RPAH/130

SYDNEY SOUTH WEST
AREA HEALTH SERVICE
NSW HEALTH

18 May 2010

Dr N Shackel
AW Morrow Gastroenterology & Liver Centre
Level 9, Building 75
Royal Prince Alfred Hospital

Dear Dr Shackel,

Re: Protocol No X10-0072 & HREC/10/RPAH/130 - "Pathogenesis of human liver disease"

The Executive of the Ethics Review Committee, at its meeting of 1 April 2010, considered your undated correspondence (received on 22 March 2010) and subsequently Ms E Prakoso's correspondence of 18 May 2010. In accordance with the decision made by the Ethics Review Committee, at its meeting of 10 March 2010, ethical approval is now granted.

This approval includes the following:

- Information for Participants (Version 3, 17 May 2010)
- Consent Form (Version 2, 22 March 2010)

You are asked to note the following:

- This approval is valid for four years, and the Committee requires that you furnish it with annual reports on the study's progress beginning in April 2011.
- This human research ethics committee (HREC) has been accredited by the NSW Department of Health as a lead HREC under the model for single ethical and scientific review and is constituted and operates in accordance with the National Health and Medical Research Council's *National Statement on Ethical Conduct in Human Research* and the *CPMP/ICH Note for Guidance on Good Clinical Practice*.

Appendix 9.4.2. Cont.

- You must immediately report anything which might warrant review of ethical approval of the project in the specified format, including unforeseen events that might affect continued ethical acceptability of the project.
- You must notify the HREC of proposed changes to the research protocol or conduct of the research in the specified format.
- You must notify the HREC and other participating sites, giving reasons, if the project is discontinued at a site before the expected date of completion.
- Where appropriate, the Committee recommends that you consult with your Medical Defence Union to ensure that you are adequately covered for the purposes of conducting this study.

Should you have any queries about the Committee's consideration of your project, please contact me. The Committee's Terms of Reference, Standard Operating Procedures, membership and standard forms are available from the Sydney South West Area Health Service website.

You are reminded that this letter constitutes ethical approval only. You must NOT commence this research project at ANY site until you have submitted a Site Specific Assessment Form to the Research Governance Officer and received separate authorisation from the Chief Executive or delegate of that site.

A copy of this letter must be forwarded to all site investigators for submission to the relevant Research Governance Officer.

The Ethics Review Committee wishes you every success in your research.

Yours sincerely,

Lesley Townsend
Executive Officer
Ethics Review Committee (RPAH Zone)

HERC\EXCOR\10-04

Appendix 9.4.3. Fox Chase Cancer Center Institutional Review Board approvals for Protocols #31-281 and #08-801

UW Medicine
SCHOOL OF MEDICINE

January 20, 2009

Dr. William S. Mason
Fox Chase Cancer Center
333 Cottman Avenue
Philadelphia, PA 19111

Dear Bill:

This letter is to confirm our collaboration to study causes of liver cancer in patients chronically infected with HBV. For the liver samples you will receive, there are two specific conditions:

- (1) The samples you will receive, collected under IRB #31281 were not taken specifically for your research program.
- (2) I will not disclose the identity or characteristics of the subjects to you for your research.

Regards,

Matthew M. Yeh, MD, PhD

Department of Pathology
Box 357470
University of Washington
Seattle, WA 98195-7470

Appendix 9.4.3. Cont.

**FOX CHASE
CANCER CENTER
IRB NOTICE OF EXEMPT STATUS REVIEW**

*Institutional Review Board
Fox Chase Cancer Center
50 Huntingdon Pike
Rockledge, PA 19046*

To: William Mason, PhD

Re: IRB #08-801

Title: 118642: Clonal Expansion of Hepatocytes and Liver Cancer in Individuals Chronically Infected with Hepatitis B virus (HBV)

This is to inform you that the IRB has reviewed the following:

		Date
Type of action	Continuing Review	
Type of review	Exempt Status	06/09/2011
Exempt Status expiration		06/08/2012
Grant number		
Protocol version		
Informed consent document		
Investigator's Brochure		
Other		

Comments:

- This study continues to meet the criteria for exemption as outlined in 45 CFR 46.101(b)(4).
- Continuing review of exempt status is required prior to 06/08/2012
- If the study changes in such a way that exempt status would no longer apply, the modifications must be submitted to the IRB office prior to implementing the changes.

CONDITIONS OF EXEMPT STATUS

1. No participants may be involved in any study activity prior to the IRB determination and documentation of exempt status or after the exempt status expiration date. (45 CFR 46.101(b)(1-6) or 21 CFR 56.104(a-d)).
2. Changes, amendments, and updates to the protocol and/or informed consent document(s) must be submitted through the RRC for review and determination by the IRB prior to the activation of the changes. This includes any change of investigator.
4. Only informed consent documents with a valid electronic approval or stamp may be presented to participants. All informed consent documents signed by participants enrolled in the study must be retained on file. The IRB Internal Audit Committee conducts periodic audits of protocol records, and consent documentation is part of these audits.
5. FCCC IRB policies require review of an exempt study not less than once per 12-month period. In order to ensure timely processing of your review, the IRB suggests submitting your application 90 days prior to expiration. The continuing review application for exempt studies is available on the IRB website and should be completed and forwarded directly to the IRB Office.

Carolyn Fang, PhD, Acting IRB Chairperson

06/09/2011
Date

Appendix 9.4.3. Cont.

FOX CHASE
CANCER CENTER
Institutional Review Board
Continuing Review Form for Exempt Studies

IRB # 08-801

Title: Clonal expansion of hepatocytes in cirrhotic and non-cirrhotic liver of patients chronically infected with hepatitis B virus

PI Name: William Mason

Date of Last Review: 9/17/2009

Date Submitted: June 2, 2011

Previously this study qualified for exemption under CFR 46.101b 4 (IRB will complete).

Has this study changed in any significant way since last reviewed?

YES

NO x

If yes, please attach description of changes.

Principal Investigator Signature

June 2, 2011

Date

IRB Chairperson Signature

6/9/11

Date

RECEIVED
JUN 06 2011
Fox Chase Cancer Center
Institutional Review Board

Appendix 9.4.4. Southern Clinical HREC approval for Protocol #171.11 and amendments

Dear Allison

*This is a formal correspondence from the **Southern Adelaide Clinical Human Research Ethics Committee**. Whilst this official title of the committee has changed the committee is still properly constituted under AHEC requirements with the registration number EC00188. This committee operates in accordance with the "National Statement on Ethical Conduct in Human Research (2007)." This department only uses email correspondence for all documents unless prior arrangements have been made with the manager. No hard copy correspondence will be issued.*

Application Number: 171.11

Title: Clonal proliferation of hepatocytes and progression of liver disease in chronic hepatitis B virus infection

Chief investigator: A/Prof Allison R. Jilbert

The Issue: The Southern Adelaide Health Service / Flinders University Human Research Ethics Committee (SAFUHREC) have reviewed and approved the above application. Your project may now commence. The approval extends to the following document:

- General Research Application Form version 2 (including PISCF).
- Supporting letter from Dr John Chen, dated 30 March 2011.
- Approval from the Royal Adelaide Hospital Human Research Ethics Committee, dated 05 March 2007.

Approval Period: 13 April 2011 to 13 April 2014

Please retain a copy of this approval for your records.

TERMS AND CONDITIONS OF ETHICAL APPROVAL

Final ethical approval is granted subject to the researcher agreeing to meet the following terms and conditions:

1. Compliance with the *National Statement on Ethical Conduct in Human Research (2007)* & the *Australian Code for the Responsible Conduct of Research (2007)*
2. To immediately report to FCREC anything that may change the ethical or scientific integrity of the project.
3. To regularly review the FCREC website and comply with all submission requirements as they change from time to time.
4. Submit an annual report on each anniversary of the date of final approval and in the correct template from the FCREC website
5. Confidentiality of research participants **MUST** be maintained at all times.
6. A copy of the signed consent form must be given to the participant unless the project is an audit
7. Any reports or publications derived from the research should be submitted to the Committee at the completion of the project.
8. Report Significant Adverse events (SAE's) as per SAE requirements available at our website.
9. The researchers agree to use electronic format for all correspondence with this department.
10. All requests for access to medical records at any SAHS site must be accompanied by this approval email.

Kind regards,
Monika

Monika Malik
A/Administrative Services Officer
Southern Adelaide Clinical Human Research Ethics Committee
SA Health
Room 2A221 - Inside Human Resources
Flinders Medical Centre, Bedford Park SA 5042

Tel: 08 8204 6453
Fax: 08 8204 4586
Email: research.ethics@health.sa.gov.au

Appendix 9.4.5. Royal Adelaide Hospital HREC approval for Protocol #050803



Government of South Australia
Central Northern Adelaide
Health Service

ROYAL ADELAIDE HOSPITAL

North Terrace,
Adelaide, SA 5000
Tel: +61 8 8222 4000
Fax: +61 8 8222 5939
ABN: 80 230 154 545
www.rahs.sa.gov.au

Research Ethics Committee

Level 3, Hanson Institute
Tel: (08) 8222 4139
Fax: (08) 8222 3035

16 August 2005

Dr A Ruskiewicz
Gastroenterology Research Laboratory
Division of Tissue Pathology
IMVS

Dear Dr Ruskiewicz,

**Re: "Liver and Pancreas Tissue Bank for prospective collection of Pancreatic and Liver Tissues at Royal Adelaide Hospital." Patient Information Sheet & Consent Form, Version 1.0 (15 August 2005).
RAH PROTOCOL No: 050803**

I am writing to advise that Research Ethics Committee approval has been given to the above project.

Research Ethics Committee deliberations are guided by the NHMRC National Statement on Ethical Conduct in Research Involving Humans.

The general conditions of approval follow:

- Adequate record-keeping is important. If the project involves signed consent, you should retain the completed consent forms which relate to this project and a list of all those participating in the project, to enable contact with them in the future if necessary. The duration of record retention for all research data is 15 years.
- You must notify the Research Ethics Committee of any events which might warrant review of the approval or which warrant new information being presented to research participants, including:
 - (a) serious or unexpected adverse events which warrant protocol change or notification to research participants,
 - (b) changes to the protocol,
 - (c) premature termination of the study.
- The Committee must be notified within 72 hours of any serious adverse event occurring at this site.
- Approval is ongoing, subject to satisfactory annual review. An annual review form will be forwarded to you at the appropriate time.

Yours sincerely,



Dr M James
CHAIRMAN
RESEARCH ETHICS COMMITTEE

Appendix 9.5. Proteins detected by HPLC-nESI-LTQ-Orbitrap XL MS of trypsin-digested, lysates of foci of hepatocytes isolated from NHL tissue sections

As described in Section 3.2.2.2, different sized foci of hepatocytes were isolated from NHL sections by laser-microdissection, as described in Section 2.20. Proteins were extracted from each focus, digested with trypsin and purified by either VivaSpin (Section 9.5.1) or Pepclean (Section 9.5.2) methods, as described in Section 2.27. Peptides were then detected and sequenced using HPLC-nESI-LTQ-Orbitrap XL MS and matched to corresponding proteins by MASCOT alignment, as described in Section 2.27. The number of cells in each focus was calculated from the area of the focus, as described in Section 3.2.1.1.

Appendix 9.5.1. VivaSpin-purified lysates of hepatocytes

Appendix 9.5.1.A. 2788 cells digested with 3.33 ng/μL trypsin

Accession ¹	Coverage (%)	# Peptides	Description
HPS3_HUMAN	0.70	4	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
CAPR1_HUMAN	0.71	1	Caprin-1 OS=Homo sapiens GN=CAPRIN1 PE=1 SV=2
RPA1_HUMAN	0.29	1	DNA-directed RNA polymerase I subunit RPA1 OS=Homo sapiens GN=POLR1A PE=1 SV=2
EIF3D_HUMAN	0.91	1	Eukaryotic translation initiation factor 3 subunit D OS=Homo sapiens GN=EIF3D PE=1 SV=1
ALG11_HUMAN	1.42	4	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
DDX27_HUMAN	0.75	1	Probable ATP-dependent RNA helicase DDX27 OS=Homo sapiens GN=DDX27 PE=1 SV=2
ERCC5_HUMAN	0.42	1	DNA repair protein complementing XP-G cells OS=Homo sapiens GN=ERCC5 PE=1 SV=3
CI025_HUMAN	3.24	1	Uncharacterized protein C9orf25 OS=Homo sapiens GN=C9orf25 PE=1 SV=3
NTNG2_HUMAN	0.94	1	Netrin-G2 OS=Homo sapiens GN=NTNG2 PE=2 SV=2

Appendix 9.5.1.B. 2788 cells digested with 333 ng/μL trypsin

CYC_GORGO	10.48	1	Cytochrome c OS=Gorilla gorilla gorilla GN=CYCS PE=3 SV=3
HPS3_HUMAN	0.70	3	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
NARG2_HUMAN	0.61	1	NMDA receptor-regulated protein 2 OS=Homo sapiens GN=NARG2 PE=2 SV=2
K1967_HUMAN	0.87	3	Protein KIAA1967 OS=Homo sapiens GN=KIAA1967 PE=1 SV=2
HEAT1_HUMAN	0.28	1	HEAT repeat-containing protein 1 OS=Homo sapiens GN=HEATR1 PE=1 SV=3
NU214_HUMAN	0.33	1	Nuclear pore complex protein Nup214 OS=Homo sapiens GN=NUP214 PE=1 SV=2
ALG11_HUMAN	1.42	3	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
CAMP1_HUMAN	0.44	1	Calmodulin-regulated spectrin-associated protein 1 OS=Homo sapiens GN=CAMSAP1 PE=1 SV=2
PAG16_HUMAN	6.79	1	Group XVI phospholipase A2 OS=Homo sapiens GN=PLA2G16 PE=1 SV=2
OSGP2_HUMAN	1.93	1	Probable O-sialoglycoprotein endopeptidase 2 OS=Homo sapiens GN=OSGEPL1 PE=2 SV=2

Appendix 9.5.1.C. 4356 cells digested with 3.33 ng/μL

Accession ¹	Coverage (%)	# Peptides	Description
AVIL_HUMAN	0.85	1	Advillin OS=Homo sapiens GN=AVIL PE=1 SV=3
ALG11_HUMAN	1.42	5	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
CAD10_HUMAN	0.76	1	Cadherin-10 OS=Homo sapiens GN=CDH10 PE=1 SV=2
TESK2_HUMAN	1.23	2	Dual specificity testis-specific protein kinase 2 OS=Homo sapiens GN=TESK2 PE=2 SV=1
DYH8_HUMAN	0.13	1	Dynein heavy chain 8, axonemal OS=Homo sapiens GN=DNAH8 PE=1 SV=2
DRP2_HUMAN	0.63	1	Dystrophin-related protein 2 OS=Homo sapiens GN=DRP2 PE=2 SV=2
EDF1_BOVIN	5.41	1	Endothelial differentiation-related factor 1 OS=Bos taurus GN=EDF1 PE=2 SV=1
ETV3_HUMAN	1.37	1	ETS translocation variant 3 OS=Homo sapiens GN=ETV3 PE=1 SV=2
X3CL1_HUMAN	1.51	1	Fractalkine OS=Homo sapiens GN=CX3CL1 PE=1 SV=1
HPS3_HUMAN	0.70	5	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
IGLO5_HUMAN	1.49	1	IgLON family member 5 OS=Homo sapiens GN=IGLON5 PE=2 SV=4
O7E24_HUMAN	1.77	1	Olfactory receptor 7E24 OS=Homo sapiens GN=OR7E24 PE=2 SV=1
S1A7A_HUMAN	5.94	1	Protein S100-A7A OS=Homo sapiens GN=S100A7A PE=2 SV=3
DVL1L_HUMAN	0.90	1	Segment polarity protein dishevelled homolog DVL-1-like OS=Homo sapiens GN=DVL1L1 PE=1 SV=1
CHK2_HUMAN	1.29	1	Serine/threonine-protein kinase Chk2 OS=Homo sapiens GN=CHEK2 PE=1 SV=1
THUM2_HUMAN	1.06	1	THUMP domain-containing protein 2 OS=Homo sapiens GN=THUMP2 PE=2 SV=1
VTI1B_HUMAN	4.31	2	Vesicle transport through interaction with t-SNAREs homolog 1B OS=Homo sapiens GN=VTI1B PE=1 SV=3

Appendix 9.5.1.D. 4356 cells digested with 333 ng/μL

CPNE9_HUMAN	1.63	2	Copine-9 OS=Homo sapiens GN=CPNE9 PE=1 SV=3
JAM2_HUMAN	2.35	1	Junctional adhesion molecule B OS=Homo sapiens GN=JAM2 PE=1 SV=1
HPS3_HUMAN	0.70	4	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
BOLL_HUMAN	2.83	1	Protein boule-like OS=Homo sapiens GN=BOLL PE=1 SV=2
ALG11_HUMAN	1.42	5	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
PKHH2_HUMAN	0.60	1	Pleckstrin homology domain-containing family H member 2 OS=Homo sapiens GN=PLEKHH2 PE=2 SV=2

Appendix 9.5.1.E. 10946 cells digested with 3.33 ng/μL

Accession ¹	Coverage (%)	# Peptides	Description
ZNF14_HUMAN	1.09	1	Zinc finger protein 14 OS=Homo sapiens GN=ZNF14 PE=2 SV=3
CYC_GORGO	10.48	1	Cytochrome c OS=Gorilla gorilla gorilla GN=CYCS PE=3 SV=3
RU2B_HUMAN	3.56	1	U2 small nuclear ribonucleoprotein B" OS=Homo sapiens GN=SNRPB2 PE=1 SV=1
LV401_HUMAN	4.72	1	Ig lambda chain V-IV region Bau OS=Homo sapiens PE=1 SV=1
HELLS_HUMAN	0.84	1	Lymphoid-specific helicase OS=Homo sapiens GN=HELLS PE=1 SV=1
HPS3_HUMAN	0.70	4	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
POMT2_HUMAN	1.07	1	Protein O-mannosyl-transferase 2 OS=Homo sapiens GN=POMT2 PE=1 SV=2
SCAPE_HUMAN	0.50	1	S phase cyclin A-associated protein in the endoplasmic reticulum OS=Homo sapiens GN=SCAPER PE=1 SV=1
MP2K2_HUMAN	2.00	3	Dual specificity mitogen-activated protein kinase kinase 2 OS=Homo sapiens GN=MAP2K2 PE=1 SV=1
MED23_HUMAN	1.10	3	Mediator of RNA polymerase II transcription subunit 23 OS=Homo sapiens GN=MED23 PE=1 SV=2
ALG11_HUMAN	1.42	4	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
AN30A_HUMAN	1.07	1	Ankyrin repeat domain-containing protein 30A OS=Homo sapiens GN=ANKRD30A PE=1 SV=3
OR2B2_HUMAN	1.96	1	Olfactory receptor 2B2 OS=Homo sapiens GN=OR2B2 PE=2 SV=1
MARHA_HUMAN	0.87	1	Probable E3 ubiquitin-protein ligase MARCH10 OS=Homo sapiens GN=MARCH10 PE=1 SV=3

Appendix 9.5.1.F. 10946 cells digested with 0.333 ng/μL

K2C1_HUMAN	1.71	1	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
CYC_GORGO	10.48	1	Cytochrome c OS=Gorilla gorilla gorilla GN=CYCS PE=3 SV=3
HPS3_HUMAN	0.70	6	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
CO3A1_HUMAN	0.82	1	Collagen alpha-1(III) chain OS=Homo sapiens GN=COL3A1 PE=1 SV=4
RPN1_HUMAN	2.64	3	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 OS=Homo sapiens GN=RPN1 PE=1 SV=1
AFF2_HUMAN	0.53	3	AF4/FMR2 family member 2 OS=Homo sapiens GN=AFF2 PE=1 SV=4
DESP_HUMAN	0.24	1	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3
JAM2_HUMAN	2.35	1	Junctional adhesion molecule B OS=Homo sapiens GN=JAM2 PE=1 SV=1
S9A10_HUMAN	0.59	3	Sodium/hydrogen exchanger 10 OS=Homo sapiens GN=SLC9A10 PE=2 SV=2
BOLL_HUMAN	2.83	1	Protein boule-like OS=Homo sapiens GN=BOLL PE=1 SV=2
RBP1_HUMAN	1.68	1	RalA-binding protein 1 OS=Homo sapiens GN=RALBP1 PE=1 SV=3
ALG11_HUMAN	1.42	6	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
TEX15_HUMAN	0.25	5	Testis-expressed sequence 15 protein OS=Homo sapiens GN=TEX15 PE=1 SV=2

Appendix 9.5.1.G. 13172 cells digested with 3.33 ng/μL

Accession ¹	Coverage (%)	# Peptides	Description
NANO2_HUMAN	3.62	1	Nanos homolog 2 OS=Homo sapiens GN=NANOS2 PE=2 SV=1
HPS3_HUMAN	0.70	4	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
S9A10_HUMAN	0.59	1	Sodium/hydrogen exchanger 10 OS=Homo sapiens GN=SLC9A10 PE=2 SV=2
GRIK4_HUMAN	1.15	2	Glutamate receptor, ionotropic kainate 4 OS=Homo sapiens GN=GRIK4 PE=2 SV=2
BBS7_HUMAN	0.84	1	Bardet-Biedl syndrome 7 protein OS=Homo sapiens GN=BBS7 PE=1 SV=2
ALG11_HUMAN	1.42	4	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
CE290_HUMAN	0.24	1	Centrosomal protein of 290 kDa OS=Homo sapiens GN=CEP290 PE=1 SV=2
CK041_HUMAN	0.32	1	UPF0606 protein C11orf41 OS=Homo sapiens GN=C11orf41 PE=2 SV=2
DOT1L_HUMAN	0.81	2	Histone-lysine N-methyltransferase, H3 lysine-79 specific OS=Homo sapiens GN=DOT1L PE=1 SV=2
GAS6_HUMAN	0.83	1	Growth arrest-specific protein 6 OS=Homo sapiens GN=GAS6 PE=1 SV=2
K1529_HUMAN	1.15	1	Uncharacterized protein KIAA1529 OS=Homo sapiens GN=KIAA1529 PE=1 SV=2
R3HD2_HUMAN	0.61	1	R3H domain-containing protein 2 OS=Homo sapiens GN=R3HDM2 PE=1 SV=3
TJAP1_HUMAN	1.26	1	Tight junction-associated protein 1 OS=Homo sapiens GN=TJAP1 PE=1 SV=1
CHD9_HUMAN	0.38	2	Chromodomain-helicase-DNA-binding protein 9 OS=Homo sapiens GN=CHD9 PE=1 SV=2
NAV3_HUMAN	0.75	3	Neuron navigator 3 OS=Homo sapiens GN=NAV3 PE=1 SV=3

Appendix 9.5.1.H 13172 cells digested with 333 ng/μL

CD2L5_HUMAN	0.86	2	Cell division cycle 2-like protein kinase 5 OS=Homo sapiens GN=CDC2L5 PE=1 SV=2
INM02_HUMAN	5.34	1	UPF0510 protein INM02 OS=Homo sapiens GN=C19orf63 PE=1 SV=1
NR2E3_HUMAN	2.20	4	Photoreceptor-specific nuclear receptor OS=Homo sapiens GN=NR2E3 PE=1 SV=1
LTN1_HUMAN	0.34	1	E3 ubiquitin-protein ligase listerin OS=Homo sapiens GN=LTN1 PE=1 SV=6
SYEP_HUMAN	0.46	1	Bifunctional aminoacyl-tRNA synthetase OS=Homo sapiens GN=EPRS PE=1 SV=5
HPS3_HUMAN	0.70	6	Hermansky-Pudlak syndrome 3 protein OS=Homo sapiens GN=HPS3 PE=1 SV=1
LICH_HUMAN	1.50	1	Lysosomal acid lipase/cholesteryl ester hydrolase OS=Homo sapiens GN=LIPA PE=1 SV=2
WDFY3_HUMAN	0.40	3	WD repeat and FYVE domain-containing protein 3 OS=Homo sapiens GN=WDFY3 PE=1 SV=2
CCD80_HUMAN	0.84	1	Coiled-coil domain-containing protein 80 OS=Homo sapiens GN=CCDC80 PE=1 SV=1

Appendix 9.5.1.H cont.

Accession ¹	Coverage (%)	# Peptides	Description
BBS10_HUMAN	0.97	2	Bardet-Biedl syndrome 10 protein OS=Homo sapiens GN=BBS10 PE=1 SV=2
CNNM4_HUMAN	1.68	2	Metal transporter CNNM4 OS=Homo sapiens GN=CNNM4 PE=1 SV=3
KCNQ4_HUMAN	1.58	1	Potassium voltage-gated channel subfamily KQT member 4 OS=Homo sapiens GN=KCNQ4 PE=1 SV=2
ITK_HUMAN	1.13	1	Tyrosine-protein kinase ITK/TSK OS=Homo sapiens GN=ITK PE=1 SV=1
CSMD2_HUMAN	0.26	1	CUB and sushi domain-containing protein 2 OS=Homo sapiens GN=CSMD2 PE=1 SV=2
ZBT45_HUMAN	1.76	3	Zinc finger and BTB domain-containing protein 45 OS=Homo sapiens GN=ZBTB45 PE=2 SV=1
AL2SA_HUMAN	0.96	1	Amyotrophic lateral sclerosis 2 chromosomal region candidate gene 11 protein OS=Homo sapiens GN=ALS2CR11 PE=2 SV=1
ALG11_HUMAN	1.42	6	Asparagine-linked glycosylation protein 11 homolog OS=Homo sapiens GN=ALG11 PE=1 SV=2
ARBK1_HUMAN	1.31	3	Beta-adrenergic receptor kinase 1 OS=Homo sapiens GN=ADRBK1 PE=1 SV=2
CBPC4_HUMAN	0.56	3	Cytosolic carboxypeptidase 4 OS=Homo sapiens GN=AGBL1 PE=2 SV=2
HPSE2_HUMAN	2.53	2	Heparanase-2 OS=Homo sapiens GN=HPSE2 PE=2 SV=3
K1529_HUMAN	1.15	1	Uncharacterized protein KIAA1529 OS=Homo sapiens GN=KIAA1529 PE=1 SV=2
MA2A2_HUMAN	0.79	3	Alpha-mannosidase 2x OS=Homo sapiens GN=MAN2A2 PE=2 SV=2
PAR15_HUMAN	1.37	4	Poly [ADP-ribose] polymerase 15 OS=Homo sapiens GN=PARP15 PE=1 SV=1
SG269_HUMAN	0.40	1	Tyrosine-protein kinase SgK269 OS=Homo sapiens GN=SGK269 PE=1 SV=4
RIC8A_HUMAN	2.82	2	Synembryn-A OS=Homo sapiens GN=RIC8A PE=1 SV=3

¹MASCOT search parameters were as follows: precursor mass range = 350-5000 Da; protein database = SwissProt; maximum missed cleavages = 2; precursor mass tolerance = 0.5 Da; fragment mass tolerance = 0.8 Da; dynamic modifications = oxidation; static modifications = carbamidomethylation.

Appendix 9.5.2. Pepclean-purified lysates of hepatocytes. All extracts were digested with 50 ng/ μ L trypsin.

Appendix 9.5.2.A. Extract of 660 cells

Accession ¹	Coverage (%)	# Peptides	Description
K2C1_HUMAN	3.42	2	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
ALBU_HUMAN	2.46	1	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2
PHF3_HUMAN	0.44	1	PHD finger protein 3 OS=Homo sapiens GN=PHF3 PE=1 SV=3
K1C10_HUMAN	2.05	1	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6
H13_HUMAN	4.07	1	Histone H1.3 OS=Homo sapiens GN=HIST1H1D PE=1 SV=2
EFHA1_HUMAN	2.30	1	EF-hand domain-containing family member A1 OS=Homo sapiens GN=EFHA1 PE=1 SV=2
MBNL1_HUMAN	2.06	1	Muscleblind-like protein 1 OS=Homo sapiens GN=MBNL1 PE=1 SV=2
NMD3A_HUMAN	0.63	1	Glutamate [NMDA] receptor subunit 3A OS=Homo sapiens GN=GRIN3A PE=1 SV=2

Appendix 9.5.2.B. Extract of 1500 cells

K2C1B_HUMAN	2.08	1	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=3
C1174_HUMAN	1.15	1	Uncharacterized protein C9orf174 OS=Homo sapiens GN=C9orf174 PE=1 SV=2
MUCB_HUMAN	4.09	1	Ig mu heavy chain disease protein OS=Homo sapiens PE=1 SV=1
ALBU_HUMAN	2.46	1	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2
CAD23_HUMAN	0.42	1	Cadherin-23 OS=Homo sapiens GN=CDH23 PE=1 SV=1
NO66_HUMAN	1.72	1	Lysine-specific demethylase NO66 OS=Homo sapiens GN=NO66 PE=1 SV=2
KIBRA_HUMAN	1.26	1	Protein KIBRA OS=Homo sapiens GN=WWC1 PE=1 SV=1
HIP1R_HUMAN	1.03	1	Huntingtin-interacting protein 1-related protein OS=Homo sapiens GN=HIP1R PE=1 SV=2
H13_HUMAN	4.07	1	Histone H1.3 OS=Homo sapiens GN=HIST1H1D PE=1 SV=2
MA2C1_HUMAN	0.67	1	Alpha-mannosidase 2C1 OS=Homo sapiens GN=MAN2C1 PE=1 SV=1
NRX1A_HUMAN	0.54	1	Neurexin-1-alpha OS=Homo sapiens GN=NRXN1 PE=1 SV=1
KDM8_HUMAN	1.68	1	Lysine-specific demethylase 8 OS=Homo sapiens GN=JMJD5 PE=1 SV=1
S2542_HUMAN	2.52	1	Solute carrier family 25 member 42 OS=Homo sapiens GN=SLC25A42 PE=1 SV=2

Appendix 9.5.2.C. Extract of 2300 cells

K1C10_HUMAN	3.60	2	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6
K2C1_HUMAN	4.66	3	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
MUCB_HUMAN	4.09	1	Ig mu heavy chain disease protein OS=Homo sapiens PE=1 SV=1

Appendix 9.5.2.D. Extract of 2300 cells

List too long to print, See .xls file attached in electronic version of this thesis

Appendix 9.5.2.E. Extract of 2800 cells

Accession ¹	Coverage (%)	# Peptides	Description
HBB_HUMAN	29.93	4	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2
HBA_HUMAN	25.35	3	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2
UK114_HUMAN	18.98	2	Ribonuclease UK114 OS=Homo sapiens GN=HRSP12 PE=1 SV=1
ALDOB_HUMAN	15.11	4	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2
CH10_HUMAN	13.73	1	10 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPE1 PE=1 SV=2
ISK4_HUMAN	12.79	1	Serine protease inhibitor Kazal-type 4 OS=Homo sapiens GN=SPINK4 PE=2 SV=1
ATPA_HUMAN	11.39	5	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1
PRDX1_HUMAN	10.55	2	Peroxiredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1
THIL_HUMAN	10.07	3	Acetyl-CoA acetyltransferase, mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1
MIF_HUMAN	9.57	1	Macrophage migration inhibitory factor OS=Homo sapiens GN=MIF PE=1 SV=4
ALDH2_HUMAN	9.48	4	Aldehyde dehydrogenase, mitochondrial OS=Homo sapiens GN=ALDH2 PE=1 SV=2
CPSM_HUMAN	8.20	10	Carbamoyl-phosphate synthase [ammonia], mitochondrial OS=Homo sapiens GN=CPS1 PE=1 SV=2
CBR1_HUMAN	7.94	2	Carbonyl reductase [NADPH] 1 OS=Homo sapiens GN=CBR1 PE=1 SV=3
MDHM_HUMAN	7.40	2	Malate dehydrogenase, mitochondrial OS=Homo sapiens GN=MDH2 PE=1 SV=3
ATPB_HUMAN	7.37	3	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3
CATA_HUMAN	7.21	3	Catalase OS=Homo sapiens GN=CAT PE=1 SV=3
RL18_HUMAN	6.91	1	60S ribosomal protein L18 OS=Homo sapiens GN=RPL18 PE=1 SV=2
DECR_HUMAN	6.87	2	2,4-dienoyl-CoA reductase, mitochondrial OS=Homo sapiens GN=DECR1 PE=1 SV=1
ACADS_HUMAN	6.07	2	Short-chain specific acyl-CoA dehydrogenase, mitochondrial OS=Homo sapiens GN=ACADS PE=1 SV=1
AATM_HUMAN	6.05	2	Aspartate aminotransferase, mitochondrial OS=Homo sapiens GN=GOT2 PE=1 SV=3
CA150_HUMAN	5.93	1	Putative uncharacterized protein C1orf150 OS=Homo sapiens GN=C1orf150 PE=2 SV=1
K1C18_HUMAN	5.81	2	Keratin, type I cytoskeletal 18 OS=Homo sapiens GN=KRT18 PE=1 SV=2
ALBU_HUMAN	5.42	3	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2
SPIN3_HUMAN	5.04	1	Spindlin-3 OS=Homo sapiens GN=SPIN3 PE=2 SV=1
DHB4_HUMAN	5.03	3	Peroxisomal multifunctional enzyme type 2 OS=Homo sapiens GN=HSD17B4 PE=1 SV=3
AK1C3_HUMAN	4.95	2	Aldo-keto reductase family 1 member C3 OS=Homo sapiens GN=AKR1C3 PE=1 SV=4

Appendix 9.4.2.E. Cont.

Accession ¹	Coverage (%)	# Peptides	Description
SERA_HUMAN	4.88	2	D-3-phosphoglycerate dehydrogenase OS=Homo sapiens GN=PHGDH PE=1 SV=4
USE1_HUMAN	4.63	1	Vesicle transport protein USE1 OS=Homo sapiens GN=USE1 PE=1 SV=2
AK1A1_HUMAN	4.62	1	Alcohol dehydrogenase [NADP+] OS=Homo sapiens GN=AKR1A1 PE=1 SV=3
EST1_HUMAN	4.59	2	Liver carboxylesterase 1 OS=Homo sapiens GN=CES1 PE=1 SV=2
G3P_HUMAN	4.48	1	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3
DHSO_HUMAN	4.48	1	Sorbitol dehydrogenase OS=Homo sapiens GN=SORD PE=1 SV=4
RL7_HUMAN	4.44	1	60S ribosomal protein L7 OS=Homo sapiens GN=RPL7 PE=1 SV=1
RGN_HUMAN	4.35	1	Regucalcin OS=Homo sapiens GN=RGN PE=1 SV=1
ECHM_HUMAN	4.14	1	Enoyl-CoA hydratase, mitochondrial OS=Homo sapiens GN=ECHS1 PE=1 SV=4
DCXR_HUMAN	4.10	1	L-xylulose reductase OS=Homo sapiens GN=DCXR PE=1 SV=2
MUCB_HUMAN	4.09	1	Ig mu heavy chain disease protein OS=Homo sapiens PE=1 SV=1
GSTA1_HUMAN	4.05	1	Glutathione S-transferase A1 OS=Homo sapiens GN=GSTA1 PE=1 SV=3
HAOX1_HUMAN	4.05	1	Hydroxyacid oxidase 1 OS=Homo sapiens GN=HAO1 PE=1 SV=1
SPYA_HUMAN	3.83	1	Serine--pyruvate aminotransferase OS=Homo sapiens GN=AGXT PE=1 SV=1
METK1_HUMAN	3.80	1	S-adenosylmethionine synthase isoform type-1 OS=Homo sapiens GN=MAT1A PE=1 SV=2
PCKGM_HUMAN	3.75	2	Phosphoenolpyruvate carboxykinase [GTP], mitochondrial OS=Homo sapiens GN=PCK2 PE=1 SV=3
HSP77_HUMAN	3.54	1	Putative heat shock 70 kDa protein 7 OS=Homo sapiens GN=HSPA7 PE=5 SV=2
TM38B_HUMAN	3.44	1	Trimeric intracellular cation channel type B OS=Homo sapiens GN=TMEM38B PE=1 SV=1
OTC_HUMAN	3.39	1	Ornithine carbamoyltransferase, mitochondrial OS=Homo sapiens GN=OTC PE=1 SV=3
TBB5_HUMAN	3.38	1	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2
AL1L1_HUMAN	3.10	2	Aldehyde dehydrogenase family 1 member L1 OS=Homo sapiens GN=ALDH1L1 PE=1 SV=2
BDH_HUMAN	2.92	1	D-beta-hydroxybutyrate dehydrogenase, mitochondrial OS=Homo sapiens GN=BDH1 PE=1 SV=3

Appendix 9.4.2.E. Cont.

Accession ¹	Coverage (%)	# Peptides	Description
ADH4_HUMAN	2.63	1	Alcohol dehydrogenase 4 OS=Homo sapiens GN=ADH4 PE=1 SV=5
EBP2_HUMAN	2.61	1	Probable rRNA-processing protein EBP2 OS=Homo sapiens GN=EBNA1BP2 PE=1 SV=2
FILA2_HUMAN	2.51	1	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1
EF1A1_HUMAN	2.38	1	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1
SBP1_HUMAN	2.33	1	Selenium-binding protein 1 OS=Homo sapiens GN=SELENBP1 PE=1 SV=2
AL4A1_HUMAN	2.31	1	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial OS=Homo sapiens GN=ALDH4A1 PE=1 SV=3
KPYR_HUMAN	2.26	1	Pyruvate kinase isozymes R/L OS=Homo sapiens GN=PKLR PE=1 SV=2
K2C1B_HUMAN	2.08	1	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=1 SV=3
TRAP1_HUMAN	1.99	1	Heat shock protein 75 kDa, mitochondrial OS=Homo sapiens GN=TRAP1 PE=1 SV=3
MMSA_HUMAN	1.50	1	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial OS=Homo sapiens GN=ALDH6A1 PE=1 SV=2
RPGF5_HUMAN	1.38	1	Rap guanine nucleotide exchange factor 5 OS=Homo sapiens GN=RAPGEF5 PE=1 SV=1
ACTN1_HUMAN	1.35	1	Alpha-actinin-1 OS=Homo sapiens GN=ACTN1 PE=1 SV=2
FAS_HUMAN	1.31	3	Fatty acid synthase OS=Homo sapiens GN=FASN PE=1 SV=3
DPOLQ_HUMAN	0.31	1	DNA polymerase theta OS=Homo sapiens GN=POLQ PE=1 SV=2
HYDIN_HUMAN	0.25	1	Hydrocephalus-inducing protein homolog OS=Homo sapiens GN=HYDIN PE=1 SV=3
TRRAP_HUMAN	0.23	1	Transformation/transcription domain-associated protein OS=Homo sapiens GN=TRRAP PE=1 SV=3
MLL3_HUMAN	0.16	1	Histone-lysine N-methyltransferase MLL3 OS=Homo sapiens GN=MLL3 PE=1 SV=3

Appendix 9.5.2.F. Extract of 4000 cells

List too long to print, See .xls file attached in electronic version of this thesis

¹MASCOT search parameters were as follows: precursor mass range = 350-5000 Da; protein database = SwissProt; maximum missed cleavages = 2; precursor mass tolerance = 0.5 Da; fragment mass tolerance = 0.8 Da; dynamic modifications = oxidation; static modifications = carbamidomethylation.

Appendix 9.6. Proteins detected by HPLC MALDI-TOF MS/MS of trypsin-digested, lysates of foci of hepatocytes isolated from NHL tissue sections

As described in Section 3.2.2.3, different sized foci of hepatocytes were isolated from NHL tissue sections by laser-microdissection, as described in Section 2.20. Proteins were extracted from each focus, digested with trypsin and purified by the Peppclean method, as described in Section 2.28. Peptides were then detected and sequenced using HPLC MALDI-TOF MS/MS and matched to corresponding proteins by MASCOT alignment, as described in Section 2.28. The number of cells in each focus was calculated from the area of the focus, as described in Section 3.2.1.1.

Appendix 9.6.A. Extract of 2800 cells

Accession ¹	Coverage (%)	# Peptides	Description
HS90B_HUMAN	14.1	12	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4
HS90A_HUMAN	6.3	2	Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5
EF2_HUMAN	5.5	4	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4
ALBU_HUMAN	3.6	2	Serum albumin precursor - Homo sapiens (Human)
ACTB_HUMAN	10.7	3	Actin, cytoplasmic 1 - Homo sapiens (Human)
HBA_HUMAN	10.6	1	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2
HBB_HUMAN	15.6	2	Hemoglobin subunit beta - Homo sapiens (Human)
SFPQ_HUMAN ²	3.3	2	Splicing factor, proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2
CATA_HUMAN	5.7	2	Catalase - Homo sapiens (Human)
H33_HUMAN ²	14.7	2	Histone H3.3 OS=Homo sapiens GN=H3F3A PE=1 SV=2
TRY2_HUMAN	8.1	1	Trypsin-2 OS=Homo sapiens GN=PRSS2 PE=1 SV=1
THIL_HUMAN	3	1	Acetyl-CoA acetyltransferase, mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1
K2C1_HUMAN	2.2	1	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
GELS_HUMAN ²	1.4	1	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1
H2AV_HUMAN ²	7	1	Histone H2A.V OS=Homo sapiens GN=H2AFV PE=1 SV=3
UK114_HUMAN	11.7	1	Ribonuclease UK114 OS=Homo sapiens GN=HRSP12 PE=1 SV=1

Appendix 9.6.B. Extract of 4000 cells

HS90B_HUMAN	22.7	14	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4
HS90A_HUMAN	11.2	4	Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5
HBB_HUMAN	42.9	5	Hemoglobin subunit beta - Homo sapiens (Human)
ACTB_HUMAN	18.1	5	Actin, cytoplasmic 1 - Homo sapiens (Human)
CPSM_HUMAN	2.9	4	Carbamoyl-phosphate synthase [ammonia], mitochondrial OS=Homo sapiens GN=CPS1 PE=1 SV=2
CATA_HUMAN	8.2	3	Catalase - Homo sapiens (Human)

Appendix 9.6.B cont.

Accession ¹	Coverage (%)	# Peptides	Description
K22E_HUMAN	5.8	3	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2
CYB5_HUMAN	25.4	2	Cytochrome b5 OS=Homo sapiens GN=CYB5A PE=1 SV=2
HBA_HUMAN	16.9	2	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2
K2C1_HUMAN	5.6	2	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
ALBU_HUMAN	3.6	2	Serum albumin precursor - Homo sapiens (Human)
THIL_HUMAN	7.3	2	Acetyl-CoA acetyltransferase, mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1
H4_HUMAN	19.4	2	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2
DHAK_HUMAN	3.8	1	Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase (cyclizing) OS=Homo sapiens GN=DAK PE=1 SV=2
CH60_HUMAN	4.4	1	60 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPD1 PE=1 SV=2
H33_HUMAN ²	14.7	2	Histone H3.3 OS=Homo sapiens GN=H3F3A PE=1 SV=2
UK114_HUMAN	11.7	1	Ribonuclease UK114 OS=Homo sapiens GN=HRSP12 PE=1 SV=1
BHMT2_HUMAN	2.8	1	Betaine--homocysteine S-methyltransferase 2 OS=Homo sapiens GN=BHMT2 PE=1 SV=1
H2AV_HUMAN ²	7.0	1	Histone H2A.V OS=Homo sapiens GN=H2AFV PE=1 SV=3
K2C8_HUMAN	4.3	1	Keratin, type II cytoskeletal 8 - Homo sapiens (Human)
AATM_HUMAN	3.3	1	Aspartate aminotransferase, mitochondrial OS=Homo sapiens GN=GOT2 PE=1 SV=3
AL1A1_HUMAN	3.8	2	Retinal dehydrogenase 1 - Homo sapiens (Human)
OTC_HUMAN	4	1	Ornithine carbamoyltransferase, mitochondrial OS=Homo sapiens GN=OTC PE=1 SV=3
THTR_HUMAN	6.4	1	Thiosulfate sulfurtransferase OS=Homo sapiens GN=TST PE=1 SV=4
DCXR_HUMAN	4.5	1	L-xylulose reductase OS=Homo sapiens GN=DCXR PE=1 SV=2
HSP77_HUMAN	3.5	1	Putative heat shock 70 kDa protein 7 OS=Homo sapiens GN=HSPA7 PE=5 SV=2
PYC_HUMAN	1.2	1	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2
UD2B4_HUMAN	2.7	1	UDP-glucuronosyltransferase 2B4 OS=Homo sapiens GN=UGT2B4 PE=1 SV=2
CBR1_HUMAN	5.8	1	Carbonyl reductase [NADPH] 1 OS=Homo sapiens GN=CBR1 PE=1 SV=3
SFPQ_HUMAN ²	3.3	1	Splicing factor, proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2
TRY3_HUMAN ²	6.6	1	Trypsin-3 OS=Homo sapiens GN=PRSS3 PE=1 SV=2
EF2_HUMAN	2.1	1	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4
PDIA3_HUMAN	3.0	1	Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4
AK1C2_HUMAN	2.5	1	Aldo-keto reductase family 1 member C2 - Homo sapiens (Human)

Appendix 9.6.B cont.

Accession ¹	Coverage (%)	# Peptides	Description
GSTA1_HUMAN	3.2	1	Glutathione S-transferase A1 OS=Homo sapiens GN=GSTA1 PE=1 SV=3
ACOT1_HUMAN	2.6	1	Acyl-coenzyme A thioesterase 1 - Homo sapiens (Human)
PDIA1_HUMAN	3.1	1	Protein disulfide-isomerase OS=Homo sapiens GN=P4HB PE=1 SV=3
EST1_HUMAN	2.6	1	Liver carboxylesterase 1 OS=Homo sapiens GN=CES1 PE=1 SV=2
ENPL_HUMAN	1.4	1	Endoplasmin OS=Homo sapiens GN=HSP90B1 PE=1 SV=1
K1C18_HUMAN	2.8	1	Keratin, type I cytoskeletal 18 OS=Homo sapiens GN=KRT18 PE=1 SV=2

¹ MASCOT search parameters were as follows: protein database = SwissProt; maximum missed cleavages = 2; MS tolerance = 30 ppm; MS/MS tolerance = 0.8 Da; dynamic modifications = oxidation; static modifications = carbamidomethylation.

² Proteins that were not also represented in the subset of proteins identified by HPLC-nESI-LTQ-Orbitrap XL analysis are highlighted in yellow.

9.7 The HBV DNA sequence detected in each patient with chronic HBV infection

Total DNA was extracted as described in Section 2.7 from liver tissue fragments of patients with chronic HBV infection (listed in Table 4.1). The HBV DNA contained in the DNA extracts were then amplified and sequenced as described in Section 2.12 using primers listed in Table 2.2. Sequences were then aligned with the closest related HBV strain in the nucleotide BLAST database (<http://blast.ncbi.nlm.nih.gov>) and the HBV DNA sequences were numbered accordingly.

9.7.1 Patient C

1

```
CTCCAGCACATTCACCAAGCTCTGCTAGATCCAGAGTGTAGAGGCTTACTTTCTGCTGGTGGCTCCAGTTCGGAAACAG
TAAACCCTGTTCGACTACTGCCTCTCCCATATCGTCAATCTTCTCGAGGACTGGGGACCCTGCACCGAATATGGAGAGCACC
ACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGTTTTCTTGTGTGACAAGAATCCTCACAATACCGCAGAGTCT
AGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGAGCTACCCACGTGCTGGCCAAAATTTGCAGTCCCAACCTCCAAT
CACTCACCAACCTCCTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCTCTCATCCT
GCTGCTATGCCTCATCTTCTTGTGGTTCTTCTGGACTGTCAAGGTATGTTGCCGTTTGTCTCTAATTCCAGGATCTTCAA
CCACCAGCACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCT
TCGGACGGAAATGCACCTGTATTCCTATCCCATCATCTTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTT
CTCTGGCTCAGTTTACTAGTGCCATTTGTTTCACTGGTTCGTAGGGCTTTCCTCCACTGTTTGGCTTTCAGTTATATGGATGA
TGTGGTATGGGGGCAAGTCTGTACAGCATCTTGTAGTCCCTTTTACCCTGTTACCAATTTCTTTTGTCTTTGGGTATAC
ATTTAAACCCTAACAAAACAAAAGATGGGGTTACTCGTTACAATTCATGGGCTATGTTATTGGATGTTATGGGTCAATGGCA
CAAGATCACATCACACAGAAAATCAAAGAATGTTTTAGAAAACCTCCTGTCAACAGACCTATTGATTGGAAAGTCTGTCAACG
TATTGTGGGTCTTTTGGGTTTTGTGCCCTTTTACACAATGTGGTTATCCTGCTTAAATGCCCTGTATGCATGTATTCAT
CTAAGCAGGCTTTCACCTTCTCGCCAACCTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTACCCCGTTGCCGGCAA
CGGCCAGGTCTGTGCCAAGTGTGTTGCTGACGCAACCCCACTGGCTGGGGCTTGGTTCATGGGCCATCAGCGCATGCGTGGAAAC
CTTTCGGCTCCTCTGCCGATCCATACTGCCGAACCTCCTAGCCGCTTGTGTTGCTCGCAGCAGGTCTGGAGCAAACATTCCTCG
GGACGGATAACTCTGTTGTTCTCTCCCGCAAATATACATCATTTCCATGGCTGCTAGGCTGTGCTGCCAAGTGGATCCTGCGC
GGGACGTCTTTGTTTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCCCTTCTCGGGGCGCTTGGGACTCTCTCGTCCCT
TCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCGCTGTGTGCTTCTCATCTGCCGGACC
GTGTGCACTTCGTTTACCTCTGCAGTTCGCATGGAGACCACCGTGAACGCCCACCCTTCTTGCCAAGGTCTTACATAAGA
GGACTCTTGGACTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTGTTAAAGACTGGGAGGAGTTG
GGGAGGAGATTAGATTAAGGTCCTTGTATTAGGAGGCTGTAGGCATAAATGGTCTGCGCACCAGCACCATGCAACTTTTT
CACCTCTGCCTAATCATCTCTTGTTCATGTCTACTGTTCAAGCCTCCAAGCTGTGCTTGGGTGGCTTTAGGGCATGGACAT
TGATCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTGCTGACTTCTTTCCTCAGTACGAGATCTTC
TAGATACCGCCTCAGCTCTATATCGGGAAGCCTTAGAGTCTCCTGAGCATTGTTACCTCACCTACCATACTGCACTCAGGCAAGCA
ATCTGTGCTGGGGGAACTAATGACTCTAGCCACCTGGGTGGGTGTAATTTGGAAGATCCAATATCTAGGGACCTAGTAGT
CAGTTATGTTAACTCATATGGGCTAAAGTTAGACAGCTATTGTGGTTTTCACTTCTTGTATCACTTTTGGGAGAGAAA
CGGTTATAGAGTATTTGGTGTCTTTCGGAGTTTGGATTGCACTCCTCCAGCTTATAGACCACCAAATGCCCTATCTTATCA
ACACTTCCGGAGACTACTGTTATTAGACGACGAGGCGAGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAGGNTCA
ATCGCCCGTTCGCAAGATCTCAATCTCGGGAACCTCAATGTTAGTATTCCTTGGACTCATAAGGTGGGAAACTTTACGGGG
CTTTATTCTTCTACTATTCTGTCCTTAAACCCTTAGGAAAACACCCCTTTTTTCCATAATATACATCAAGACATTAT
CAAAAATGTGAACAATTTGTAGGCCACTCACAGTCAATGAGAAAAGAAGACTGCAATTGATTATGCCTGCTAGGTTTTATC
CAAATGTTACCAAATATTTGCCATTGGATAAGGTTAATTAACCTTATTATCCAGAACATCTAGTTAATCACTTCCAAACC
AGACATTAATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAAACAACATAGCGCCTCATTGTTGGGTACC
ATATTCTTGGGAACAAGAGCTACAGCATGGGGCAGAATCTTCCACCAGCAATCCTCTGGGATTCTTCCCGACCACCAAGTTG
GATCCAGCCTTCAGAGCAAACACCGCAAATCCAGATTTGGGACTTCAATCCCAACAAGGACACTTGCCAGACGCCAACAAAGGT
AGGAGCTGGAGATTCGGGCTGGGATTCAACCCACCGCACGGAGGCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTAC
AAACCTTGCCAGCAAATCCGCCTCCTGCCTCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCTCTGTCTCCACCTTTGAGA
AACACTCATCCTCAGGCCATGCAGTGGAA
```

3193

9.7.2 Patient CN

1844

```
TCTTGTTCATGTCTACTGTCAAGCCTCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTG
GAGTACTGTGGAGTTACTCTCGTTTTTGCCTTGTGACTTCTTCTCCTCAGTACGAGATCTTCTAGATACCGCCTCAGCTCTG
TATCGGGAAGCCTTAGAGTCTCCTGAGCATTGTTACCTCACCATACTGCACTCAGGCAAGCAATCTTGTCTGGGGGAACT
AATGACTCAGTACCTGGGTGGTGTAAATTTGGAAGATCCAGCGCTAGAGACCTAGTAGTCACTTATGTCAACACTAATA
TGGGCTTAAAGTTCAGGCAACTCTGTGGTTTTACATTTCTTGTCTCACTTTTGGGAAAGAACAGTTATAGAGTATTTGGTG
TCTTTCGGAGTGTGGATTGCACTCCTCCAGCTTATAGACCACCAAATGCCCTATCCTATCAACACTTCCGGAGACTACTGT
TGTTAGACGACGAGGAGGTCCTTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAGGTCTCAATCGCCGCTCGCAGAAGAT
CTCAATCTCGGGAATCTCAATGTTAGTATTCCTTGGACTCATAAGGTGGGGAACCTTACTGGGCTTTATTCTTCTACTGTACC
```

TGCTTTAATCCTCATTGGAAAACACCATCTTTTCTTAATATACATTTACACCAAGACATTATCAAAAAATGTGAACAGTTTG
TAGGCCACTCACAGTTAATGAGAAAAGAGATTGCAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTA
CCATTGGATAAGGGTATTAACCTTATTATCCAGAACATCTAGTTAATCATTACTTCCAAACTAGACACTATTTACACACTCT
ATGGAAGGCGGTATATTATATAAGAGAGAAAACACATAGCGCCTCATT
2810

9.7.3 Patient CY

1
CAGCACATTCCACCAAGATATGTTAGATCCCAGAGTGAGGGGCATATACTTTCTGATGGTGGCTCCAGTTCCGGAACAGTAA
ACCTGTTCCGACTACTGCCTCTCCCATATCGTCAATCTTCTCGAGGACTGGGGACCTGCACCGAATATGGAGAGCACCACA
TCAGGATTCCTAGACCCTGCTCGTGTACAGCGGGGTTTTCTTGTGACAAGAATCCTCACAAATACCACAGAGTCTAGA
CTCGTGGTGGACTTCTCTCAATTTCTAGGGGGAGCACCCACGTGCTGCGCCAAAATTTGCAGTCCCAACCTCCAATCACT
CACCACCTCTGTCTCCAATTTGTCTGGTTATCGCTGGATGTGCTGCGGCGTTTTATCATATTCCTCTTCATCTGCTG
CTATGCCTCATCTTCTTGTGGTTCTTCTGGACTACCAAGGTATGTGCCCGTTTGTCTACTTCCAGGAACATCAACTAC
CAGCACGGGACCATGCAAGACCTGCACGATTCTGCTCAAGGAACCTCTATGTTTCCCTCTTGTGCTGTACAAAACCTTCGG
ACGGAACCTGCCTGTATTCCCATCCCATCATCTTGGGCTTTCGCAAGATTCCATGGGAGTGGGCTCAGTCCGTTTCTCC
TGGCTCAGTTTACAGTGCATTTGTTTCACTGGTTCGTAGGGCTTCCCCACTGTTTGGCTTTCAGTTATATGGATGATTTG
GTATTGGGGGCCAAGTCTGTACACATCTTGAATCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTT
AAACCTACTAAAATAAACCTGGGGCTACTCCCTTAACTTCAATGGGCTATGTAATTGGAAGTTGGGGTACCTTACCACAGG
AACATATTGTACAAAATAAACCAATGTTTTCGGAACTTCCATAAATAGACCTATTGATTGGAAGTTTGTCAACGAATT
GTGGCCTTCTGGGCTTTGCCGCTCCCTTTACACAATGTGGTTACCCAGCATTAAAGCCTTTGTAGCATGTATACAAGCTAA
ACAGGCTTTCACTTTCTCGCCAACTTACAAAACCTTTCTGTGTAAACAATATCTGAACCTTTACCCCGTTGCTCGGCAACGGT
CAGGCTTTGCCAAGTGTGGTGTGACGCAACCCCACTGGTGGGGCTTGGCCATAGGCCATAGCCGATCGCTGGAACTTT
GTGCTCCTCTCCGATCCATACTGCGGAACCTTAGCAGCTTGTTTTGTGCTCGCAGCCGGTCTGGAGCGAAACTTATCGGCAC
CGACAACCTGTGTCTCTCGGAAATACACCTCCTTTCCATGGCTGTAGGCTGTGCTGCCAACTGGATCTGCGCGGGA
CGTCTTTGTCTACGTCCTGCTGCGGCTGAATCCGCGGACGACCCGCTCTCGGGCGCTTTGGGGCTTACCCTCCCTTCTT
CGTCTGCCGTTCCGGCCGACCACGGGGCGACCTCTTTACCGGCTCTCCCGCTGTGTGCCTTCTCATCTGCCGGACCGTGT
GCACTTCGCTTCACTCTGCACGTCGCATGGAGACCACCGTGAACGCCACCAGGCTTGCCTCAAGGCTTTACATAAGAGGAC
TCTTGGACTCTCAGCAATGTCAACGACCGACCTTGGGACACTTCAAAGACTGTGATTTAAAGACTGGGAGGAGTTGGGG
AGGAGACTAGATTAATGATCTTTGTACTGGGAGGCTGTAGGCATAAATTGGTCTGTTCAACAGCACCATGCAACTTTTTAC
TCTGCCTGATCATCTCATGTTTCACTGTCACCTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATTGAC
CCGTATAAAGAAATTTGGAGCTTCTGTGGAGTTACTCTTTTTTGCCTTCTGACTTTTTTCTTCTGTTCGTGATCTCTCGA
CACCGCCTCTGCTCTGTATCGGGAGGCTTAGAGTCTCCGGAACATTGTTACCTACCATAACAGCACTAAGGCAAGCTATTC
TGTGTTGGGGTGAAGTGTGATGAATCTGGCCACCTGGGTGGGAAGTAAATTTGGACGACCCAGCATCCAGAGAATTAGTAGTAAGC
TATGTCAATGTTAATATGGGCCAAAAATAACAACCTATTGTGGTTTCACTTTCTGTCTTACTTTTGGAAAGAGAAACTGT
TCTTGAATTTGGTGTCTTTGGAGTGTGGATTGCACTCTTACCGCTTACAGACCACCAATGCCCTATCTTATCAACACT
TCCGGAACCTACTGTTGTTAGACGACGAGGCAAGTCCCTTAGAAGAAAGAAACTCCCTCACCTCGCAGACGAAGGCTCAATC
GCCGCTAGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAGTATCCCTTGGACTCATAAGGTGGGCAACTTTACTGGACTT
TATCTTCTACTGTACTGTCTTTAATCCTGATTGGCAAACCTCCCTCTTTTCTCACATTCAATTTGAAAGAGGATATATCAA
TAGATGCCAACAAATATGTAGGCCCTCTTACAGTTAACGAAAAAGGAGATAAAAATTGATTATGCTGCTAGGTTCTATCCTA
ACCGTACCAAATATTTGCCCTTAAATAAAGGCATTAACCTTATTATCCTGAACACATAGTTAATCATTACTTCCAAACTAGG
CATTATTTACATACTCTGTGGAAGGCGGTATTTATATAAGAGAGAAACTACTCTCAGCGCCTCATTTTGTGGGTACCATA
TTCTTGGGAACAAGACTATATCATGGGAGGTC
2856
3126
TCCTGCCTCCAACAATCGGCAGTCAAGGGAGACAGCAAACCTACAATATCTCCACCTCTAAGAGACAGTCATCCTC
AGGCCATGCAGTGGAAATC

9.7.4 Patient CYRY

1
TTCCACTGCCTTCCACCAAGCTCTGCAGGATCCCAGAGTCAAGGGTCTGTATTTTCTGCTGGTGGCTCCAGTTCAGGAACAG
TAAACCCTGCTCCGAATATTGCCCTCTCACATCTCGTCAATCTCCGCGAGGACTGGGGACCTGTGACGAACATGGAGAACATC
ACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGGTTTTCTTGTGACAAGAATCCTCACAAATACCGCAGAGTCT
AGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGATCACCCGTGTGCTTGGCCAAAATTCGCAGTCCCAACCTCCAATC
ACTCACCAACCTCCTGTCTCCTCAACTTGTCTGTTATCGCTCGATGTGCTGCGGCGTTTTATCATATTCCTCTCATCCTG
CTGCTATGCCTCATCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCAGGATCAACAAC
AACCAGTACGGGACCATGCAAAACCTGCACGACTCCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACCTA
CGGATGGAAATTGCACCTGTATCCCATCCCATCGTCTGGGCTTTCGCAAAATACCTATGGGAGTGGGCTCAGTCCGTTTC
TCTTGGCTCAGTTTACTAGTGCCATTTGTTTCACTGGTTCGTAGGGCTTTCCTCCACTGTTTGGCTTTCAGCTATATGGATGAT
GTGGTATTTGGGGCCAAAGTCTGTACAGCATCGTGGGCTTTATACCGCTGTACCAATTTTCTTTTGTCTCTGGGTATACA
TTTAAACCTTAAACAAAAGATGGGGTATTCCCTAACTTCAATGGGTTACATAATTGGAAGTTGGGGAACCTTGGCCAC
AGGATCATATTGTACAAAAGATCAAACTCTTGTAGAAAACCTCTGTTAACAGGCTTATGATTGGAAGTATGTTCAAGA
ATTGTGGGTCTTTTGGGCTTGTGCTCCATTTACACAATGTGGATATCCTGCCTTAAATGCCTTTGTATGCATGTATACAAGC
TAAACAGGCTTTCCTTCTCGCCAACTTACAAGGCTTTCTAAGTAAACAGTACATGAACCTTTACCCGTTGCTCGGCAAC
GGCTGGTCTGTGCCAAGTGTGCTGACGCAACCCCACTGGCTGGGGCTTGGCCATAGGCCATCAGCGCATGCGTGGAAAC
TTTGTGGCTCCTCTGCCGATCCATACTGCGGAACCTTAGCCAGCTTACTGCTCGCAGCCGGTCTGGAGCAAAGCTCATCGG
AATGACAATCTGTGCTCTCGCGGAAATACATCGTTTCCATGGCTGCTAGGCTGTACTGCCAACTGGATCCTTCCGCG
GGAGTCTTTGCTACGTTCCCGTGGGCTGAATCCCGCGGACGACCCCTCTCGGGCCGTTGGGACTCTCTCGTCCCT
TCTCCGCTGCGGTTCCAGCCGACCAGGGGCGCACCTCTTTTACCGGCTCTCCCGCTGTGTGCCTTCTCATCTGCCGCTC
GTGTGACTTCCGCTTCACTCTGCACGTTGCATGGAGACCACCGTGAACGCCCATCAGATCCTGCCAAGGCTTTACATAAGA
GGACTCTTGGACTCCAGCAATGTCAACGACCGACCTTGGGCTACTTCAAAGACTGTGTGTTTAAAGACTGGGAGGAGCTG
GGGAGGAGATTAGGTTAATGATCTTTGATTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTT

CACCTCTGCCTAATCATCTCTTGTACATGTCCCACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACAT
TGACCCCTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCCTCCGTCAGAGATCTCC
TAGACACCGCCTCAGCTCTGTATCGAGAAGCCTTAGAGTCTCCTGAGCATTGCTCACCTCACCATACTGCACTCAGGCAAGCC
ATTCTCTGCTGGGTGGAATTGATGACTCTAGCTACCTGGGTGGGTAATAATTTGGAAGATCCAGCATCCAGGGATCTAGTAGT
CAATTATGTTAATACTAACATGGGTTTAAAGATCAGGCAACTATTGTGGTTTCATATATCTTGCCTTACTTTTGGAAAGAGAGA
CTGTACTTGAATATTTGGTCTCTTTCGGAGTGTGGATTTCGCACTCCTCCAGCCTATAGACCACCAATGCCCTATCTTATCA
ACACTTCCGAAACTACTGTTGTTAGACGACGGGACCGAGGCAGGTCCTTAGAAGAAGAACTCCCTCGCCTCGCAGACGCAG
ATCTCAATCGCCGCTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAGTATTCCCTGGACTATAAGGTGGGAACTTT
ACGGGGCTTTTATTCCTCTACAGTACCTATCTTTAATCCTGAATGGCAAACCTCCTTCTTCTTAAGATTTCATTTACAGAGGA
CATTATTAATAGGTGTCAACAATTTGTGGGCCCTCTCACTGTAATGAAAAGAGGAGATTGAAATTAATTATGCCTGCTAGAT
TCTATCCTACCCACACTAAATATTTGCCCTTAGACAAAGGAATTAACCTTATTTATCCGGATCAGGTAGTTAATCATTACTTC
CAAACCAGACATTTTACATACTCTTTGGAAGCTGGTATTCTATATAAGAGGGAAACCACACGTAGCGCATCATTTTGGCG
GTCACCATATTTTGGGAACAAGACTACAGCATGGGAGGTTGGTCATCAAACTCGCAAAGGCATGGGGACGAATCTTTCT
GTTCCCAACCTCTGGGATTCTTTCCGATCATCAGTTGGACCTGCATTTCGGAGCAACTCAAACAATCCAGATTGGGACTT
CAACCATTCAAGGACCACTGGCCAACAGCCAACCAGGTAGGAGTGGGAGCATTTCGGGCCAGGGCTCACCCCTCCACACGGCGG
TATTTTGGGGTGGAGCCCTCAGGCTCAGGGCATATTGACCACAGTGTCAACAATTCCTCCTCCTGCCCTCCACCAATCGGCAGT
CAGGAAGGCAGCTACTCCATCTCTCCACCTTAAGAGACAGTCACTCCTCAGGCCATGCAGTGGAA
3221

9.7.5 Patient DG

1649

CTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTGTATTTAAAGACT
GGGAGGAGTTGGGGGAGGAGATTAGGTTAATGATCTTTGTACTGGGAGGCTGTAGGCATAAATTTGGTCTGTTCCACCAGACCA
TGCAACTTTTTACCTCTGCCTAATCATCTCATGTTTCATGTCCACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGG
GGCATGGACATTTGACCCGTATAAAGAATTTGGAGCTTCTGCGGAGTTACTCTCTTTTTTGCCTTCTGACTTCTTTCCGTCTAT
TCGGGATCTCCTCGACACCGCCTCTGCTCTGTATCGGGAGGCTTAGAGTCTCCGGAACATTGTTCTCCTCACCATAACAGC
TAAGGCAAGCTATTTCTGTGTTGGGTGAGTTGATGAATCTGGCCACTGGGTGGGAAGTAATTTGGAAGACCCAGCATCCAGG
GAATTAGTAGTAAGCTATGTCAATGTTAACATGGGCCATAAATCAGACAATATTGTGGTTTACATTTCTGTCTTACTTT
TGGAAAGAGAACTGTTCTAGAGTATTTGGTGTCTTTTGGAGTGTGGATTTCGCACTCCTCCACCTTACAGACCACCAATGCC
CTATCTTATCAACTTCCGGAACCTACTGTTGTTAGACGACGAGGACGACCCCTAGAAGAAGAACTCCCTCGCCTCGCAGA
CGAAGGTCTCAATCGCCGCTCGCAGAAAGATCTCAATCTCGGGAATCTCAATGTTAGTATC
2456

9.7.6 Patient FMC1

1297

GCTCGCAGCAGGTTTGGGGCAAACTCATCGGGACTGACAATTTCTGTCGTCTCCTCCGCAAGTATACATCATTTCCATGGCT
GCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCCCT
CCCGGGCCGCTTGGGGCTTTACCGCCGCTTCTCCGCTGTTGTTCCGACCGACCAGGGGCGCACCTCTCTTTACGCGGAC
TCCCCGTCTGTGCCTTCTCACCTGCCGACCGTGTGCACTTCGCTTACCTCTGCACGTGCGATGGAAACCACCGTGAACGAC
CACCGGAACCTGCCAAAGGTCTTGCATAAGAGGACTCTTGGACTTTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAA
AGACTGTGTGTTTACTGAGTGGGAGGAGTTGGGGGAGGAGATGAGGTTAATGATCT
1767

9.7.7 Patient FMC2

1296

TGCTCGCAGGCGGCTTGGAGCAAAGCTCATCGGAACTGACAATTTCTGTCGTCTCCTCCGCAAGTATACATCGTTTCCATGG
CTGCTAGGCTGTGCTGCCAACTGGATCCTTCCGCGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCC
CTCTCGGGCCGCTTGGGACTCTATCGTCCCTTCTCCGTCTGCCGTTCCAGCCGACCACGGGGCGCACCTCTCTTTACGCGG
TCTCCCCGTCTGTGCCTTCTCATCTGCCGGTCCGTGTGCACTTCGATTACCTCTGCACGTTGCATGGAGACCACCATGAACG
CCCATCAGATCCTGCCAAAGGTCTTATATAAGAGGACTCTTGGACTCCAGCAATGTCAACGACCGACCTTGAGGCCTACTTC
AAAGACTGTGTGTTTAAAGACTGGGAGGAACTGGGGGAGGAGATTAGGTTAATGATTTTTGTATTAGGAGGCTGTA
1785

9.7.8 Patient FMC3

1297

GCTCGCAGCAGGTTTGGAGCAAAGCTTATTGGGACTGACAATTTCTGTCGTCTATCCGCAAGTATACATCGTTTCCATGGCT
GCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCCCT
CCCGGGCCGCTTGGGGCTTACCGCCGCTCCTCCGTCTGCCGTTACCGCCGACCGGGGCGCACCTCTCTTTACGCGGAC
TCCCCGTCTGTGCCTTCTCATCTGCCGACCGTGTGCACTTCGATTACCTCTGCACGTTGCATGGAGACCACCGTGAACGCC
CACCGGAACCTGCCAAAGGTCTTACATAAGAGGACTATTGGACTTTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAA
AGACTGTGTGTTTACTGAGTGGGAGGAGTTGGGGGAGGAGATGAGGTTAAAGGTCT
1768

9.7.9 Patient FMC4

1297

GCTCGCAGCAGGCTCTGGAGCAAACCTTATCGGGACTGACAATTTCTGTCGTGCTCTCCCGCAAATATACATCCTTTCCATGGCT
GCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCCTCGGCGCTGAATCCC CGGGACGACCCCT
CCCCGGGCGCTTGGGGCTCTACCGACCGCTCCTCCGTCTTCCGTACCGACCGACCACGGGGCGCACCTCTCTTTACCGGGAC
TCCCCGTCTGTGCCTTCTCACCTGCCGGACCGTGTGCACTTCGCTTACCTTTGCACGTGCGGTGGAAACCACCGTGAACGCC
CACTGGAACCTGACCAAGGTCTTACATAAGAGGACTCTTGGACTTTCAGCAATGTCAACGAACGACCTTGAGTCATACTTCAA
AGACTGTGTGTTTACTGAGTGGGAGGAGTTGGGGGAGGAGAT
1753

9.7.10 Patient FMC5

1296
TGCTCGCAGCAGGCTCTGGAGCAAAGCTGCTTGGGACTGACAACCTCTGTTGTTCTCTCCCGCAAATATACATCGTTCATGGC
TGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCCTCTGCCCCTGAATCCTGCGGACGACCCCT
TCTCGGGGCGCTTGGGGATCTATCGTCCCCTTCTCTGCCTGCCGTTCGACCGACCACGGGGCGCACCTCTCTTTACCGGGA
CTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTACCTCTGCACGTGCGATGGAGACCACCGTGAACGC
CCACCAAATATTGCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCTGCAATGTCAACGACCGACCTTGAGGCATACTTCA
AAGACTGTTTGTAAAGTATGGGAGGAGTTGGGGGAGGAGATTAGATTAAGGTCT
1766

9.7.11 Patient FMC6

1298
TGCTCGCAGGCCGGTCTGGGGCAAACCTCATCGGGACTGACAATTTCTGTCGTGCTCTCCCGCAAATATACATCATTTCCATGG
CTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCCTCGGCGCTGAATCCC CGGGACGACCC
CTCCCGGGCGCTTGGGGCTCTACCGCCCGCTTCTCCGCTGTGTACCGACCGACCACGGGGCGCACCTCTCTTTACCGGG
ACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTACCTCTGCACGTGCGATGGAAACCACCGTGAACG
CCCCTGGAACCTGCCAAGGTCTTGTATAAGAGGACTCTTGGACTTTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCA
AAAGACTGTGTGTTTACTGAGTGGGAGGAGTTGGGGGAGGAGTTAGGTTAATGATCTTTGTAGGGGGAGGCTGTAGGCATAA
ATTGG
1797

9.7.12 Patient FMC7

1297
GCTCGCAGGCCGGTCTGGAGCAAAGCTCATCGGCACCGACAACCTCTGTCGTCTTTCTCGGAAATATACCTCCTTTCCATGGC
TGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCCTCGGCGCTGAATCCC CGGGACGACCCG
TCCC CGGGCGCTTGGGGATCTACCGTCCCCTTCTTCTGCTGCGGTTCGCGCCGACCACGGGGCGAACCTCTCTTTACCGGGT
CTCCCCGTCTGTGCCTTCTCATCTGCCGGTCCGTGTGCACTTCGCTTACCTCTGCACGTGCGATGGAGACCACCGTGAACGC
CGCCAGGTCTTGGCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCA
AAGACTGTGTATTTAAAGACTGGGAGGAGTTGGGGGAGGAGACTAGGTTAATGATCTTTGTAGGGGGAGGCTGTAGGCATAA
1773

9.7.13 Patient GS1

3
AACTCCAGCACATTCACCAAGCTCTGCTAGATCCCAGAGTGAGGGCCCTATACTTTCTGCTGGTGGCTCCAGTTCAGGAAC
AGTAAACCCTGTTCGACTACTGCCTCTCCCATACGTCAATCTTCTCGAGGACTGGGGACCTGCACCGAATATGGAGAGCA
CCACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGGTTTTTCTTGTGACAAGAATCCTCACAATACCACAGAGT
CTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTCTGGCCAAAATTTGCAGTCCCCAACCTCCAA
TCACTCACCACCTCTTGTCTCCAATTTGTCTGGTTATCGCTGGATGTGCTGCGGCGTTTTATCATCTTCTCTCATCC
TGCTGCTATGCCTCATCTTCTTGTGGTTCTTCTGGACTACCAAGGTATGTTGCCGTTTTGTCTTACTTCCAGGAACATCA
ACTACCAGCACGGGACCATGCAAGACCTGCACGACTCCTGCTCAAGGAACCTCTATGTTTCCCCTCTGTTGCTGTACAAAACC
TTCGGACGGAAATGCACTTGTATFCCCATCCCATCATCTTGGGCTTTCGCAAGATTCCTATGGGAGTGGGCCTCAGTCCGTT
TCTCCTGGCTCAGTTTACTAGTGCCATTTGTTTCACTGGTTTCGTAGGGCTTTCCCCACTGTTTGGCTTTCAGTTATATGGATG
ATGTGGTATTGGGGGCCAAGTCTGTACAACATCTGAATCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATA
CATTTAAACCCTACTAAAACCAACGTTGGGGTACTCCCTTAACTTCATGGGATATGTAATTGGAAGTTGGGGTACCTTACC
ACAGGAACATATGTACACAAAATCAACAATGTTTTCGGAAACTTCCATAAAATAGACCTATTGATTGGAAAGTATGTCAAC
GCATTTGTTGGGCTTCTGGGCTTTGCCGCTCCCTTACACAATGTGGTTACCCAGCGTTAATGCCTTTGTATGCATGTATACAA
GCTAAGCAGGCTTTCACCTTCTCGCCAACTTACAAGGCCTTCTGTGTAACAATACTGAACCTTACCCCGTTGCTCGGCA
ACGGTCAGGCTTTGCCAAGTGTGTTGCTGACGCAACCCCACTGGTTGGGGCTTGGCCATAGGCCATCAGCGCATGCGTGGAA
CCTTTGTGGCTCCTCTGCCGATCCATACTGCCGAACTCCTAGCAGCTTGTGTTTGTGTCGACGCGGCTGAGAGCGAACCTTCTC
GGCACCGACAACCTCTGTTGTCTCTCTCGGAAATACACCTCCTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCG
CGGGACGTCCTTTGTCTACGTCCCCTCGGCGCTGAATCCC CGGGACGACCCGCTCTCGGGTCTGTTGGGACTTTACCGTCCCC
TTCTTCTGCTGCGGTTCCGACCGACCACGGGGCGCACCTCTCTTTACCGCGGCTCTCCCCTGCTGTGCCTTCTCATCTGCCGGC
CGTGTGCACTTCGCTTACCTCTGCACGTGCGATGGAGACCACCGTGAACGCCCGTCAAGTCTTGGCCAAGGTCTTACATAAG
AGGACTCTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTGTATTTAAAGACTGGGAGGAGTT
GGGGGAGGAGATTAGATTAAGGTCTTTGTACTGGGAGGCTGTAGGCATAAAATGGTCTGTTTACCAGCACCATGCAACTTTT
TCACCTCTGCCTAATCATCTCATGTTTATGTCCCACTGTTTCAAGCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACA
TTGACCCGTATAAAGAATTTGGAGCTTCTGTGGAGTACTCTCTTTTTTGCCTTCTGACTTCTTCCCTCTATTCTGATCTC

CTCGACACCGCCTCTGCTCTATATCGGGAGGCCTTAGAGTCTCCGGAACATTGTTACCTCACCATACAGCACTAAGGCAAGC
TATCCTGTGTTGGGGTGAAGTTGATGAATCTGGCCACCTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGGGAATTAGTAG
TAAGCTATGTCAATGTTAATATGGGCCTAAAAATCAGACAACCTATTGTGGTTTCACATTTCTGTCTTACTTTTGGGAAGAGAA
ACTGTTCTTGAGTATTTGGTGTCTTTGGAGTGTGGATTCCGACTCCTCCCGCTTACAGACCACCAATGCCCTATCTTATC
AACACTTCCGGAACCTACTGTTGTAGACGACGAGGCAGGTCCCTAGAGAAGAACTCCCTCGCCTCGCAGACGAAGGTCTC
AATCGCCGCGTTCGAGAAAGATCTCAATCTCGGGAATCTCAATGTTAGTATCCCTTGGACTCATAAGGTGGGAACTTTACTGG
GCTTTATTCTTCTACTGTACCTGTCTTTAATCTGACTGGCAAACCTCCCTCTTTTCTCACATTCATTTGAAAGAGGATATTA
TCAATAGATGTCAACAATATGTGGGCCCTCTTACAGTTAACGAAAAAGGAGATTAAAATTGATTATGCCTGACTAGGTCTAT
CCTAACCATACCAAAATTTGGCCCTTAGATAAAAGGCATCAAACCTTATTATCCTGAACATATAGTTAATCATTACTTCCAAAC
TAGGCATTATTTACATACTCTGTGGAAGGCGGGTATTTTATATAAGAGAGAACTACTCGCAGCGCCTCATTTTGTGGGTAC
CATATTCTTGGGAACAAGAGCTACAGCATGGGAGGTGGTCTTCCAAACCTCGACAAGGCATGGGGACGAATCTTTCTGTTC
CAATCCTCTGGGATTCTTTCCCGATCACCAGTTGGACCTTGCAATTCAGAGCCAATCAAACAATCCAGATTGGGACTTCAACC
CCAACAAGGATCCATGGCCAGAGGCACACCAGGTAGGAGTGGGATCCTTCCGGCCAGGGTTCACCTCCACCACACGGCCTT
TTGGGGTGGAGCCCTCAGGCTCAGGGCATATTGACAACAGTACCAGCAGCGCCTCCTCCTGCCTCCACCAATCGGCAGTCAGG
AAGACACCTACTCCCATCTCTCCACCTCTAAGAGAAAGTATCCTCAGGCCATGCAGTGG
3213

9.7.14 Patient GS2

1299

TGCTCGCAGCCGGTCTGGAGCAAACCTTATCGGCACCGACAACCTCTGTTGTCTCTCTCGGAAATACACCTCCTTTCCATGGCT
GCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACCTCTTTGTCTACGTCCCGTCGGCGCTGAATCCCGCGGACGCCAT
CTCGGGGCCGTTTGGGACTCTACCGTCCCCTTCTTCGTCTGCCGTTCCGACCACCGGGGCGACCTCTCTTTACCGGGT
TCCCCGTCTGTGCCTTCTCATCTGCCGACCGTGTGCACTTCGCTTCACTCTGCACGTGCGATGGAAACCACCGTGAACGCC
CGCCAGGTCTTGCCAAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCACCTTGGGCATACTTCAA
AGACTGTGTATTTAAAGACTGGGAGGAGTTGGGGAGGAGATTAGTTAATACTGGGAGGCTGTAGGCATAAATTGGTCTGTT
CACCAGCACCATGCAACTTTTTACCTCTGCCTAATCATCTCATGTTTCAAGCCTCAAGCTGTGCCTTG
GGTGGCTTTGGGGCATGGACATTGACCCATATAAAGAATTTGGAGCTTCTGTGGAGTTACTCTCTTTTTTGCCTTCTGACTTC
TTTCTTCTATTCGGGATCTCCTCGACACCGCCTCTGCTTTGTATCGGGAGGCCCTTAGAGTCTCCGGAACATTGTTACCTCA
CCATACAGCCTAAGGCAAGCTATCTGTGTTGGGGTGAAGTTGATGAATCTGGCCACCTGGGTGGGAAGTAATTTGGAAGACC
CAGCATCCAGGGAATTGGTAGTAAGCTATGTCAATGTTAACATGGGCCTAAAAATCAGACAACCTATTGTGGTTTTCACATTTCC
TGCTTACTTTTGGGAAGAGAACTGTTCTAGAGTATTTGGTGTCTTTTGGAGTGTGGATTTCGACTCCTACCGCTTACAGACC
ACCAATGCCCTTATCTTATCAACACTTCCGGAACCTACTGTTGTTAGACGACGAGGCAGGTCCCTAGAAGAAGAACTCCCT
CGCCTCGCAGACGAAGGTCTCAATCGCCGCTCGCAGAAGATCT
2431

9.7.15 Patient GS3

1174

TGCCAAGTGTGTTGCTGATGCAACCCCACTGGCTGGGGCTTGGCCATCGGCCATCAGCGCATGCGTGGAACCTTTTGTGGCTCC
TCTGCCGATCCATACTGCCGAACCTTAGCCGCTTGTTTTGTCTCGCAGCCGGTCTGGAGCAAACCTCATCGGGACTGATAATT
CTGTGCTCCTTTCTCGGAAATATACATCCTTTCCATGGCTGCTAGGNTGTNCTGCCAACTGGANCTCGCGGGACGTCCTTTG
TTTACGTCCCCTCAGCGCTGAATCCCAGGACGACCTTCCGCGGGTCTGCTGGGACTCTATCGTCCCCTTCTCCGTCTGCCG
TACCGTCCGACCAGGGGCGCACCTCTCTTTACGCGGTCTCCCCGTCTGTGCCTTCTCATCTGCCGGTCCGTGTGCACTTCGC
TTCACCTCTGCACGTTGCATGGAGACCACCGTGAACGCCATCAGATCCTGCCCAAGGTCTTACATAAGAGGACTCTTGGACT
CCAGCAATGTCAACGACCACCTTGGAGCCTACTTCAAAGACTGTGTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGA
1752

9.7.16 Patient GS4

1175

GCCAAGTGTGTTGCTGACGCAACCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTGGCTCCT
CTGCCGATCCATACTGCCGAACCTTAGCCGCTTGTTTTGTCTCGCAGCAGGTCTGGAGCAAACCTTCTCGGGACAGATAACTC
TGTGTCTCCTTCCCGCAAATATACATCGTTTCCATGGCTGCTAGGNTGTGCTGCCAACTGGATNNGCCGCGGACCTCCTTTG
TTTACGTCCCCTCAGCGCTGAATCCCAGGACGACCTTCTCGGGGCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCG
TTTCCAGCCACTACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGC
TTCACCTCTGCACGTCGATGGAGACCACCGTGAACGCCACCAATCTTGGCCAAAGGTCTTACATAAGAGGACTCTTGGACT
CTCTGTGATGTCAACGACCACCTTGGAGCATACTTCAAAGACTGTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGACTA
GATTAATTATTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTACCTCTGCCT
1836

9.7.17 Patient GS5

1175

GCCAAGTGTGTTGCTGACRCAACCCCACTGGCTGGGGCTTGGCCATAGGCCATCAGCGCATGCGTGGAACCTTTTGTGGCTCCT
CTGCCGATCCATACTGCCGAACCTTAGCCGCTTGTTTTGTCTCGCAGCCGGTCTGGAGCAAAGCTCATCGGAACCTGACAATTC
TGTCGCTCCTCTCACGAAATATACATCGTTTCCATGGCTGCTAGGCTGACTGCCAACTGGATCCTTCCGGGACGTCCTTTG
TTTACGTCCCCTCAGCGCTGAATCCCAGGACGACCTTCTCGGGGCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCG
TTCCAGCCGACCAGGGGCGCACCTCTCTTTACGCGGTCTCCCCGTCTGTGCCTTCTCATCTGCCGGTCCGTGTGCACTTCGC

TTACCTCTGCACGTTGCATGGAGACCACCGTGAACGCCATCAGATCCTGCCAAGGCTTTACATAAAGAGGACTCTTGGACT
CCCAGCAATGTCAACGACCGACCTTGAGGCCTACTTCAAAGACTGTGTGTTTAAAGGACTGGGAGGAGCTGGGGGAGGATTA
GGTTAAAGGCTTTTGTATTAGGAGGCTGTAGGCATAAATGGTCTGCGCACCAGCACCATGCAACTTTTTTACCTCTGCCTAA
TCATCTCTGTACATGTCCAACGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATTGACCCTTATAAA
GAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCCTTCGGTTCAGANATCTCCTAGACACCGCTC
AGCTCTGTATCGGGAAGCCTTAGAGTCTNCTGANCAATTGCTCACCTCACCATACTGCACTCAGGCAAGCCNTTCTCTGCTGGG
GG
2089

9.7.18 Patient HN

1844

TCTTGTTTCATGTCTACTGTTCAAGCCTCAAAGCTGTGCCTTGGGTGGCTTTAGGACATGGACATTGATCCTTATAAAGAATT
TGGAGCTAGTGTGGAGTTACTCTCGTTTTTGCCTTCCGACTTCTTTCCCTTCAGTACGAGATCTTCTAGATACCGCTCAGCTC
TATATCGGGAAGCCTTAGAGTCTCCTGAGCATTGCTCAGCTCACCATACTGCACTCAGGCAAGCAATTCTCTGCTGGGGGAC
CTTATGACTCTAGCCACCTGGGTGGGTGTAATTTGGAAGATCCAGTATCCAGGGACCTAGTAGTCAGTTATGTTAACTAA
TATGGGCCTAAAGTTCAGGCAACTATTGTGGTTTACATTTCTGTCTCACTTTTGAAGAGAAACGGTTCGTAGAGTATTTGG
TGCTTTTCGGAGTGTGGATTGCGACTCCTCCAGCTTATAGACCACCTAATGCCCTATCTTATCAACACTTCCGGAGACTACT
GTTATTAGACGACGAGGCAGGTCCTTAGAAGAAGAACTCCCTCGCTCGCAGACGAAGGTCTCAATCGCCGCGTCGCAGAAG
ATCTCAATCTCGGGAATCTCAATGTTAATATTCCTTGGACTCATAAGGTGGGAACTTTACGGGGCTTTATTCTTCTACTGTT
CCTATCTTTAACCTCATTGGAACACCCCTCTTTTCTAATATACATTTACACCAAGACATTATCAAAAAATGTGAACAATT
TGTAGGCCACTCACAGTCAATGAGAAAAGAAGACTGCAATTGATTATGCCTGCTAGGTTTTATCCAAATGTTACCAAAATATT
TGCCATTGGATAAAGGGTATTAAACCTTATTATCCAGAACATTTAGTTAATCATTACTTCCAAACCAGACATTATTTACACAC
TCTATGGAAGGCGGGTATATTATATAAGAGAGAAAACAACCGTAGCGCCTCATTTTGTGGTCA
2811

9.7.19 Patient HS

1854

GTCTACTGTTTACGCTCCAAGCTGTGCCTTGGGTGGCTTTAGGACATGGACATCGACCCTTATAAAGAATTTGGAGCTACTG
TGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCCTTCCGTACGAGATCTTCTAGACACCGCTCAGCTTTGTATCGGGAT
GCCTTAGAGTCTCCTGAGCATTGTTACCTCACCATACTGCTCTCAGGCAAGCAATTCTTTGCTGGGGGATCTAATGACTCT
AGCTACCTGGGTGGGTGGTAATTTGCAAGATCAAGCATCTAGGGAAGTACTAGTCAAGCAATATGGCCCTAA
AGTTCAGACAATTATTGTGGTTTACATTTCTGTCTCACTTTTGAAGAGAAACGGTTCATAGAGTATTTGGTGTCTTTTCGGA
TGTTGGATTGCGACTCCTCCAGCTTATAGACCACCAATGCCCTATCTTATCAACACTTCCGGAGACTACTGTTGTTAGACG
TCGAGGCAGGTCCTTAGAAGAAGAACTCCCTCGCTCGCAGACGAAGGTCTCAATCGCCGCGTCGCAGAAGATCTCAATCTC
GGGGATCTCAATGTTAATATTCCTTGGACTCATAAGGTGGGAACTTTACGGGGCTTTATTCTTCTACTGTACCTGTCTTTAA
CCCACATTGGAACACCCCTCTTTTCTAATATACATTTACACCAAGACATTATCAAAAAATGTGAACAATTTGTAGGCCAC
TCACAGTCAATGAGAAAAGAAGACTGCAATTAATTATGCCTGCTAGGTTTTATCCAAAGGGTACCAAAATATTTGCCATTAGAT
AAAGGGTATCAACCTTATTATCCAGAACATGTAGTTAATCATTACTTCCAAGCCAGACATTATTTACACACTCTATGGAAGG
CGGGTATATTATATAAGAGAGAAAACAACACATAGCGCCTCATTTTGTG
2814

9.7.20 Patient L

1

CTCCACAACCTTCCACCAAACCTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAG
TAAACCTGTTCGACTACTGTCTCACATATCGTCAATCTTCTCGAGGATTGGGGACCTGCGCTGAACATGGAGAACATC
ACATCAGGATTCCCTAGGACCCCTGCTCGTGTACAGGCGGGTTTTTCTTGTGACAAGAATCCTCACAATACCACAGAGTCT
AGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCAACCTCCAATC
ACTACTAACCTCTTGTCTCCAATTTGTCTGGTTATCGCTGGATGTGTCTGCGCGTTTTATCATCTTCTCTTTCATCTCTG
CTGCTATGCCATCTTCTTGTGGTTCTTCTGGACTGCCAAGGTATGTTGCCGTTTGTCTCTACTTCCAGGAACATCAAC
TACCAGCACGGGACCATGCAAAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTTTCCCTCTTGTGCTGTACAAAACCTT
CGGACGGAAATTCGACTTGTATTTCCATCCCATCATCTTGGGCTTTTCGCAAGATTCCTATGGGAGTGGGCCTCAGTCCGTTT
TCCTGGCTCAGTTTACTAGTGCATTTGTTTCAAGTGGTTCGAGGCTTTCCCCACTGTTTGGCTTTCAGTTATATGGATGAT
GTGGTATTGGGGGCAAGTCTGTACAACATCTTGAATCCCTTTTACCCTATTACCAATTTTCTTTTGTCTTTGGGTATA
TTTAAACCCTAATAAAACCTAAACGTTGGGGCTACTCCCTTAACTTTCATGGGATATGTAATGGAAGTTGGGGTACCTTGGCA
CAGGACCATATGTACAAAAAATCAACAATGTTTTCGGAACTTCCATATAAATAGACCTATTGATTGGAAGTATGTCAAAG
AATTGTGGGCTTCTAGGCTTTGCGCTCCCTTACACAATGTGGTTACCCAGCATTAATGCCTTTGTATGCATGTATACAAG
CTAAGCAGGCTTTCACTTTCTCGCAACTTACAAGGCCTTTCTGTGTCAACAATATATGCACCTTTACCCCGTTGCTCGGCAA
CGGTCAAGTCTTTGCCAAGTGTGTTGCTGACGCAACCCCACTGGTGGGGCTTGGCCATGGGCGATCAGCGCATCGGTGGAAC
CTTTGTGGCTCCTTCCGATCCATACTGCGGAACCTTCTAGCAGTGTGTTTGGCTCGCAGCAGTTCGGAGCAACCTTATCG
GAACCGACAACCTGTGTTGCTCTCTCGGAAGTACACCTCCTTTCCATGGCTGCTAGGCTGTGCTGCAACTGGATCCTGCGC
GGGACGTCCTTTGTCTACGTCCCGTGGCGCTGAATCCGCGGACGACCCGCTCTCGGGGCGTTTTGGGACTCTATCGTCCCT
TCTTCTGCTGCGGTTCCGGCCGACCAGGGGCGCACTCTCTTTACGCGGTCTCCCGCTCTGTGCCTTCTCATCTGCCGGACC
GTGTGCACTTCGCTTCACTCTGCAGTTCGATGGAACACCGTGAACGCCCCGCAAGGCTTGCCTAAGGCTTTATATAAGA
GGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTGTATTTAAAGACTGGGAGGAGTTG
GGGAGGAGATTAGTTAATGATCTTTGTACTGGGAGGCTGTAGGCATAAATGGTCTGTTTACCAGCACCATTGCAACTTTTT
CACTCTGCCTAATCATCTCATGTTTCAATGCTCCACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGCATGGACAT
TGACCCGTATAAAGAATTTGGAGCTTCTGTGGAGTTACTCTTTTTTGCCTTCTGACTTCTTTCCCTTCCATTCTGATCTCC
TCGACACCGCTCTGCTCTGTATCGGGAGGCTTAGAGTCTCCGGAACATTGTTACCTCACCATACAGCACTAAGGCAAGCC
ATTCTGTGTTGGGGTGGAGTTGATGAATTTGGCCACTGGGTGGGAGTAATTTGGAAGACCAGCATCCAGGGAATAGTAGT
AAGCTATGTCAACGATAATATGGGCATAAAATCAGACAACATTTGTGGTTTTCATTTCCAGTCTTACTTTTGAAGAGAAA
CTGTCTTGTAGTATTTGGTGTCTTTTGGAGTATGGATTGCACTCCTCCCGCTTACAGACCACCAAAAGCCCTATCTTATCA

ACACTTCCGGAACTACTGTTGTTAGACGACGAGGACAGGTCCTTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAGGTCTCA
ATCGCCGCGTCGAGAAGATCTCAATCTCGGGAATCTAAATGTTAGTATCCCTTGACTCATAAGGTGGGAACTTTACTGGG
CTTTATTCTTCTACTGTACCTGTCTTTAATCCTGACTGGCAAACCTCCCTCTTTTCCCTCACATTCATTTGAAAGAGGATATTAT
CAATAGATGTCAACAATATGTGGGCCCTCTTACAGTTAACGAAAAAGGAGGTTAAAATTGATTATGCCTGCTAGGTTTTATC
CTAACCTACCAATATTTGCCCTTAGATAAAGGCATTAAACCTTATTATCCTGAACATATAGTTAATCATTACTTCCAAACT
AGGCATTATTTACATACTCTGTGGAAGGCGGGTATTTTATATAAGAAGAACTACTCGCAGCGCCTCATTTTTGTGGGTACC
ATATCTTGGGAACAAGAGCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGACAAGGCATGGGGACGAATCTTTCTGTCCC
AATCTCTGGGATCTTTCCCGTACCAGTTGGACCCGCGATTCCGAGCCAATTCAAACAATCCAGATTGGGACTTCAACCC
CAACAAGGATCAATGGCCAGCGGCAAACAGGTAGGAGTGGGATCATTTCGGCCAGGGTTCACTCCACCACACGGCAATCTTT
TGGGGTGGAGCCCTCAGGCTCAGGGTATATTGACAACAATACCAGCCGCGCTCCTCCTGCCTCCACCAATCGGCAGTCAGGA
AGACAGCCTACTCCATCTCTCCACCTCTAAGAGACAGTCATCCTCAGGCCATGCAGTGGAA
3216

9.7.21 Patient MH

1884

CCTTGGGTGGCTTTGAGGCATGGACATTGACCCTTATAAAGGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCT
GACTTCTTTCCCTCCGTCAGAGATCTCCTAGACACCGCCTCAGCTCTGTATCGAGAAGCCTTGGAGTCTCCTGAGCATTGCTC
ACCTCACCATACTGCACTCAGGCAAGCCATTCTCTGCTGGGTGGAATTGATGACTCTAGCTACCTGGGTGGGTAATAATTTGG
AAGATCCAGCATCCAGGGATCTAGTAGTCAATTATGTTAATACTAACATGGGTTTAAAGATCAGGCAACTATTGTGGTTTCAT
ATATCTTGCCTTACTTTTGAAGAGAGACTGTACTTGAATATTTGGTCTCTTTCCGAGTGTGGATTGCACTCCTCCAGCCTA
TAGACCACCAATGCCCCATCTTATCAACACTTCCGAAACTACTGTTGTTAGACGACGGGACCGAGGCAGGTCCCTTAGAA
GAAGAATCCCTCGCCTCGCAGACGCAGATCTCAATCGCCGCGTCGAGAAGATCTCAATCTCGGGAATCTCAATGTTAGTAT
TCCTTGGACTCATAAGGTGGGAACTTTACGGGGCTTTATTCCTCTACAGTACCTATCTTTAATCCTGAATGGCAAACCTCCTT
CCTTTCCTAAGATTCATTTACAAGAGGACATTATTAATAGGTGTCAACAATTTGTGGCCCTCTCACTGTAATGAAAAGAGA
AGATTGAAATTAATATGCCTGCTAGATTCTATCTACCCACACTAAATATTTGCCCTTAGACAAAGGAATTAACCTTATTA
TCCAGATCAGGTAGTTAATCATTACTTCCAAACCAGACATTATTTACATACTCTTTGGAAGGCTGGTATTCTATATAAAGG
AAACCACACGTAGCGC
2811

9.7.22 Patient NT

1862

CTGTTTCAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATTGACCCTTATAAAGAATTTGGAGCTACTGTGGAGT
TACTCTCGTTTTTGCCTTCTGACTTCTTTCCCTCCGTCAGAGATCTCCTAGACACCGCCTCAGCTCTGTATCGAGAAGCCTTG
GAGTCTCCTGAGCATTGCTCACCTCACCATACTGCACTCAGGCAAGCCATTCTCTGCTGGGTGGAATTGATGACTCTAGCTAC
CTGGGTGGGTAATAATTTGGAAGATCCAGCATCCAGGGATCTAGTAGTCAATTATGTTAATACTAACATGGGTTTAAAGATCA
GGCAACTATTGTGGTTTCATATATCTTGCCTTACTTTTGAAGAGAGACTGTACTTGAATATTTGGTCTCTTTCCGAGTGTGG
ATTCGCACTCCTCCAGCCTATAGACCACCAATGCCCCATCTTATCAACACTTCCGAAACTACTGTTGTTAGACGACGGGA
CCGAGGCAGGTCCCTTAGAAGAAGAACTCCCTCGCCTCGCAGACGCAGATCTCAATCGCCGCGTCGAGAAGATCTCAATCTC
GGAAATCTCAATGTTAGTATTCCTTGGACTCATAAGGTGGGAACTTTACGGGGCTTTATTCCTCTACAGTACCTATCTTTAA
TCCTGAATGGCAAACCTCCTTCTTTCCTAAGATTCAATTTACAAGAGGACATTATTAATAGGTGTCAACAATTTGTGGCCCTC
TCACTGTAATGAAAAGAGAAGATTGAAAGTTAATATGCCTGCTAGATTCTATCTACCCACACTAAATATTTGCCCTTAGA
CAAAGGAATTAACCTTATTATCCAGATCAGGTAGTTAATCATTACTTCCAAACCAGACATTATTTACATACTCTTTGGAGGC
TGGTATTCTATATAAAGGGAAACCACACGTAGCGCATCATTTG
2818

9.7.23 Patient SAAO

1863

TGTTTCAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATTGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTT
ACTCTCGTTTTTGCCTTCTGACTTCTTTCCCTCCGTCAGAGATCTCCTAGACACCGCCTCAGCTCTGTATCGAGAAGCCTTGG
AGTCTCCTGAGCATTGCTCACCTCACCATACTGCACTCAGGCAAGCCATTCTCTGCTGGGTGGAATTGATGACTCTAGCTACC
TGGGTGGGTAATAATTTGGAAGATCCAGCATCCAGGGATCTAGTAGTCAATTATGTTAATACTAACATGGGTTTAAAGATCAG
GCAACTATTGTGGTTTCATATATCTTGCCTTACTTTTGAAGAGAGACTGTACTTGAATATTTGGTCTCTTTCCGAGTGTGGA
TTCGCACTCCTCCAGCCTATAGACCACCAATGCCCCATCTTATCAACACTTCCGAAACTACTGTTGTTAGACGACGGGAC
CGAGGCAGGTCCCTTAGAAGAAGAACTCCCTCGCCTCGCAGACGCAGATCTCAATCGCCGCGTCGAGAAGATCTCAATCTCG
GGAATCTCAATGTTAGTATTCCTTGGACTCATAAGGTGGGAACTTTACGGGGCTTTATTCCTCTACAGTACCTATCTTTAAT
CCTGAATGGCAAACCTCCTTCTTTCCTAAGATTCAATTTACAAGAGGACATTATTAATAGGTGTCAACAATTTGTGGGCCCTCT
CACTGTAATGAAAAGAGAAGATTGAAATTAATATGCCTGCTAGATTCTATCTACCCACACTAAATATTTGCCCTTAGACA
AAGGAATTAACCTTATTATCCAGATCAGGTAGTTAATCATTACTTCCAAACCAGACATTATTTACATACTCTTTGGAAAGCT
GGTATTCTATATAAAGGGAAACCACACGTAGCGC
2811

9.7.24 Patient WN

1

TTCCACTGCCTTCCACCAAGCTCTGCAGGATCCCAGAGTCAGGGGTCTGTATTTTCTGCTGGTGGCTCCAGTTCAGGAAACAG
TAAACCCTGCTCCGAATATTGCCTCTCACATATCGTCAATCTCCGCGAGGACTGGGGACCCTGTGACGAACATGGAGAACATC
ACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGGTTTTCTCGTTGACAAGAATCCTCACAATACCGCAGAGTCT
AGACTCGTGGTGGACTTCTCTCAGTTTTCTAGGGGATCACCCGTGTGTCTTGGCCAAAATTCGCAgTCCCCAACCTCCAATC
ACTACCAACCTCCTGTCTCCAACCTGTCTGGTTATCGCTGGATGTATCTGCGCGGTTTTATCATATTTCTCTTCTCATCTG
CTGCTATGCCATCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCGTTTGTCCCTAATTCAGGATCAACAAC
AACCAGTACGGGACCATGCAAAACCTGCACGACTCCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACCTA
CGGATGGAAATTCACCTGTATTTCCATCCCATCGTCTGGGCTTTTCGAAAATACCTATGGGAGTGGGCCTCAGTCCGTTTC
TCTTGGCTCAGTTACTAGTGCCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCACTGTTTGGCTTTCAGCTATATGGATGAT

TGTTGATTGGGGCCAAAGTCTGTACAGCATCGTGAGGCCCTTTATACCGCTGTTACCAATTTTCTTTTCGTCTCTGGGTATAACA
TTTAAATCCTAACAAAAAAGATGGGGTTATCCCTAAACTTCATGGGTTACATAAATTGGAAGTTGGGGAACTTTGCCAC
AGGATCATATAGTACAAAAGATCAAACACTGTTTTAGAAAACCTCCTGTTAACAGGCCTATTGATTGGAAAGTGTGTCAACGA
ATTTGTGGGTCA

1006

1224

CCATCAGCGCATGCGTGGAACCTTTGTGGCTCCTCTGCCGATCCATACTGCGGAACCTCTAGCCGCTTGTGTTTGGCTCGCAGCC
GGTCTGGAGCAAAGCTCATCGGAACCTGACAATTTCTGTCGTCCTCTCGCGGAAATATACATCGTTTTCCATGGCTGCTAGGCTGT
ACTGCCAACTGGATCCTTTCGCGGGACGTCTTTTGTCTACGTCCCGTTCGCGCTGAATCCCGCGGACGACCCTCTCGGGCCG
CTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCAGCCGACCACGGGGCGCACCTCTCTTTACGCGGTCTCCCGTCTG
TGCTTCTCATCTGCCGGTCCGTGTGCACCTTCGCTTCACCTCTGCACGTTGCATGGAGACCACCGTGAACGCCCATCAGATCC
TGCCCAAGGCTTACATAAGAGGACTCTTGGACTCCCAGCAATGTCAACGaccgaCCTTGAGGCCCTACTTCAAAGACTGTGTA
TTTAAAGGACTGGGAGGAGCTGGGGGAGGAGATTAGGTTAATGATCTTTGTATTAGGAGGCTGTAGGCATAAAATTTGGTCTGCGC
ACCAGCACCATTGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTACATGTCCCACTGTTCAAGCCTCAAAGCTGTGCCTGG
GTGGCTTTGGGGCATGGACATTGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCT
TTCCCTCCGTCAGAGATCTCCTAGACACCGCCTCAGCTCTGTATCGAGAAGCCTTAGAGTCTCCTGAGCATTGCTCACCTCAC
CATACTGCACTCAGGCAAGCCATTCTCTGCTGGGTGGAATTGATGACTCTAGCTACCTGGGTGGGTAATAATTTGGAAGATCC
AGCATCCAGGGATCTAGTAGTCAATTATGTTAATACTAACATGGGTTTAAAGATCAGGCAACTATTGTGGTTTTCATATATCTT
GCCTTACTTTTTGGAGAGAGACTGTACTTGAATATTTGGTCTCTTTCGGAGTGTGGATTTCGCACCTCTCCAGCCTATAGACCA
CCAAATGCCCTATCTTATCAACACTTCCGAAACTACTGTTGTTAGACGACGGGACCGAGGCAGGTCCCCTAGAAGAAGAAC
TCCCTCGCTCGCAGACGCAGATCTCAATCGCCGCTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAGTATTCCTTGG
ACTCATAAGGTGGAAACTTTACGGGCTTTTATCTCTACAGTACCTATCTTTAATCCTGAATGGCAAACTCCTTCCCTTCC
TAAGATTCAATTTACAAGAGGACATTATTAATAGGTGTCAACAATTTGTGGGCCCTCTCACTGTAATGAAAAGAGAAGATTGA
AATTAATTATGCCTGCTAGATTCTATCCTACCCACACTAAATATTTGCCCTTAGACAAAGGAATTAACCTTATTATCCAGAT
CAGGTAGTTAATCATTACTTCCAACCAGACATTAATTACATACTCTTGGAAAGGCTGGTATTCTATATAAAAGGGAACCAC
ACGTAGCGCATCATTTTGGGGTCCACCATATCTTGGGAACAAGAGCTACAGCATGGGAGGTTGGTTCATCAAAACCTCGCAAA
GGCATGGGAGCAATCTTCTGTTCCCAACCTCTGGGATCTTTCCCGATCATCAGTTGGACCTGCATTCCGAGCCAACTC
AAACAATCCAGATTGGGACTTCAACCCCATccggGAacctGGCCAACAGCAACCAGGTAGGAGTGGGAGCATTCCGGCCAGG
GCTCACCCCTCCACACGGCGGTATTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAATTGACCACAGTGTCAACAATTCCTCCTC
CTGCCTCCACCAATCGGCAGTCAGGAAGGCAGCCTACTCCCATCTCTCCACCTCTAAGAGACAGTCATCTCAGGCCATGCAG
TGGA

3221

9.7.25 Patient XA

1859

ACTGTTACGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGCATGGACATTGACCCGTATAAAGAATTTGGAGCTTCTGTGGAG
TTACTCTCTTTTTTGCTTCTGACTTCTTTCCCTCTATTTCGAGATCTCCTCGACACCGCCTCAGCTCTGTATCGGGAGGCCTT
AGAGTCTCCGGAACATTGTTACCTCACCATACAGCACTCAGGCAAGCTATCCTGTGTTGGGGTGGATTGATGAATCTGGCCA
CCTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGGGAATTAGTAGTCAGCTATGTCAATGTTAATATGGGCCATAAAATC
AGACAACACTACTGGTTCACATTTCCCTGTCTTACTTTTGGAAAGAGAACTGTTCTTGAGTATTTGGTGTCTTTTGGAGTGTG
GATTCGCACCTCTCTGCTTACAGACCACCAATGCCCTATCTTATCAACACTTCCGAAACTACTGTTGTTAGACGACGAG
GCAGGTCCCCTAGAAGAAGAACCTCCTCGCCTCGCAGACGAAGGTCTCAATCGCCGCTCGCAGAAGATCTCAATCTCGGGAA
TCTCAATGTTAGTATCCCTTGGACTCATAAGGTGGAAACTTTACTGGGCTTTATCTTCTACTGTACCTGTCTTTAATCCTG
AGTGGCAAACCTCCTCTTTCCCTCACATTCATTTACAGGAGGACATTATTAATAGATGTCAACAATATGTGGGCCCTCTTACA
GTTAATGAAAAAAGGAGATTAAAATTAATTATGCCTGCTAGGTTCTATCCTAACCTTACCAATATTTGCCCTTGGATAAAGG
CATTAAACCATATATCCTGAACAAGCAGTTAATCATTACTTCAAACACTAGGCATTATTTACATACTCTGTGGAAGGCGGGTA
TTCTATATAAGAGAGAACTACACGCAG

2800

9.7.26 Patient Y2

1185

TGCTGACGCAACCCCACTGGTTGGGGCTTGGCCATAGGCCATCAGCGCATGCGCGGAACCTTTGTGTCTCCTCTGCCGATCC
ATACTGCGGAACCTCTAGCCGCTTGTGTTNGCTCGCAGCCGGTCTGGAGCGAACTTATCGGGACTGACAATTTCTGTCGTGCTC
TCCCAGAGTATACATCGTTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCC
GTCCGGCTGAATCCCGCGGACGACCCTCCCGGGGCCGCTTGGGGCTCTACCGCCGCTTCTCCGCTGCGGATCCGACCAT
CCACGGGGCGCACCTCTCTTTACGCGGACTCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACCTTCGCTTCACTCTG
CACGTGCGATGGAGACCACCGTGAACGCCACCGGAACCTGCCAAGGTCTTACATAAGAGGACTCTTGGACTTTGCAAGATG
TCAACGACCGACTTGGAGCATACTTCAAAGACTGTGTGTTTACTGAGTGGGAGGAATCGGGGGAGGATAGGTTAATGAT
TTTTGTATTAGGAGGCTGTAGGCATAAAATTTGGTCTGTTTACCAGCACCATGCAACTTTTTTACCTCTGCCTAATCATCTCATG
TTCATGTCTACTGTTCAAGCCTCAAAGCTGTGCCTTGGGTGGCTTTAGGGCATGGACATTGACCCGTATAAAGAATTTGGAG
CTTCTGTGGAGTTACTCTCTTTTTTGCCTTCTGACTTCTTTCCCTCTATTAANAATCTTCTCGACACCGCCTCTGCTCTGTAT
CGGGAGGCCTTAGAGTCTCCGGAACATTGTTTCTCACCATACGCAATCAGGCAAGCTATTCTGTGTTGGGGTGGATTGAT
GAATCTAGCCACCTGGGTGGGAAGTAATTTGGAAGACCCAGCATCCAGGGAATTAGTAGTCGGCTATGTCAATGTTAATATAG
GTCTAAAACCTCANACAATTTGTGGTTTTCACATTTCTGTCTTACTTTTGGAAAGACAAGTTGTTCTTGAATATTTGGTGTCT
TTTGGAGTGTGGATTCTGCACCTCCTGCTGNTATAGACCACCAATGCCCTATCTTATCAACACTTCCGGAACCTACTGNTGN
TAGACAACGAGGNANGTCCCCTAGAAGAAGAANTNCCCTCGCCTCGCAGA

2395

9.7.27 Patient Y3

1185

TGCTGACGCAACCCCACTGGTTGGGGCTTGGCCATAGGCCATCAGCGCATGCGTGGAAACCTTTGTGTCTCCTCTGCCGATCC
ATACTGCGGAACCTCTAGCCGCTTGTGTTGCTCGCAGCAGGCTTGGGGCAAACCTCATCGGGACTGACAATTTCTGTCGTGCTC
TCCCAGAGTATACATCATTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCC

GTCGGCGCTGAATCCCGCGGACGACCCCTCCCGGGCCGCTTGGGGCTCTACCGCCCGCTTCTCCGCTGTTGTACCGACCGA
CCACGGGGCGCACTCTCTTTACGCGGACTCCCGCTGTGCTTCTCATCTGCCGACCGTGTGCACTTCGCTTCACTCTG
CACGTGCGATGGAGACCACCGTGAACGCCACAGGAACCTGCCAAGGTCTTGCATAAGAGGACTCTTGGACTTTCAGCAATG
TCAACGACCGACCTTGAAGCATACTTCAAAGACTGTGTGTTTACTGAGTGGGAGGANTTGGGGGAGGAGNTAGGTTAATGAT
CTTTGTACTAGGAGGCTGTAGGCATAAATTGGTGCCTTACCAGCACCATGCAACTTTTTACCTTCGCTGATCATCTCNTG
TTCATGTCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTAGGGCATGGACATCGACCCGTATAAAGAATTTGGAG
CTTCTGTGGAGTTACTCTCTTTTTGCCTTCTGACTTCTTTCCCTTCTATTTCGAGATCTCCTCGACACCGCCTCTGCCTGTAT
CGGGAGCCCTTAGAGTCTCCGGAACATTGTTACCTCACCATACGGCAATCAGGCAAGCTATTCTGTGTTGGGGTGAGTTGAT
GAATTTAGCCACCTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGGAATTAGTAGTCGGCTATGTCAACGTTAATATGG
GCATAAAAATCAGACAACCTATTGTGGTTTACATTTCTGTATTACTTTTGGGAGAGAACTGTTCTTGAATATTTGGTGTCT
TTTGGAGTGTGGATTTCGACTCCTCCTGCATATAGACCACCAATGCCCTATCTTATCNACACTTCCGAAACTACTGT
2343

9.7.28 Patient Y4

1184

TTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCCATAGGCCATCAGCGCATGCGGGAACCTTTGTGTCTCCGTTGCCGATC
CATACTGCGGAACCTTAGCCGCTTGTGTTTGTCTCGCAGCCGGTCTGGAGCGAAACTTATCGGGACTGACAATTCGTGTGCT
CTCCCGCAAGTATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCTTTGTTTACGTCC
CGTCGGCGCTGAATCCCGCGGACGACCCCTCCCGGGCCGCTTGGGGCTCTACCGCCCGCTTCTCCGCTGTCGCTACCGACCG
ACCACGGGGCGCACTCTCTTTAGCGGACTCCCGCTGTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACTCT
GCACGTGCGATGGAGACCACCGTGAACGCCACCGGAACCTGCCAAGGTCTTACATAAGAGGACTCTTGGACTTTCAGCAAT
GTCAACGACCGACCTTGAAGCATACTTCAAAGACTGTGTGTTTACTGAGTGGGAGGAGTTGGGGGAGGAGATGAGGTTAAAGG
TCTTTGTATTAGGAGGCTGTAGGCATAAATTGGTCTGTTTACCAGCACCACGCAACTTTTTACCTTCGCTAATCATCTCAT
GTTTCATGCTCTACTGTTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATTGACCCGTATAAAGAATTTGGA
GCTTCTGTGGAGTTACTCTCTTTTTGCCTTCTGACTTCTTTCCCTTCTATTTCGAGATCTTCTCGACACCGCCTCTGCCTGT
TCGGGAGGCCCTTAGAGTCTCCGGAACATTGTTTACCTCACCATACGGCACTCAGGCAAGCTATTCTGTGTTGGGGTGAGTTGA
TGAATCTAGCCACCTGGGTGGGAAGTAATTTGGAAGACCCAGCATCCAGGGAATTAGTAGTCAGCTATGTCAATGTTAATATG
GGCCTAAAAATCAGACAACCTATTGTGGTTTACATTTCTGTCTTACTTTTGGAAAGAGAACTGTTCTTGAATATTTGGTGTCT
TTTTGGAGTGTGGATTTCGACTCCTCCTGCATATAGACCACCAATGCCCTATCTTATCAACACTTCCGAAACTACTGTTG
TTAGACGACGAGGCAGGTCCCTTAGAAGAAGAACTCCCTCNCTCGCAGACGAAGNCTCANTCACCGCGTCG
2419

9.7.29 Patient Y5

1174

TTGCCAAGTGTGTTGCTGACGCAACCCCCACTGGATGGGGCTTGGCTATTGGCCATCNCCGATGCGTGAACCTTTGTGGCTC
CTCTGCCGATCCATACTGCGGAACCTTAGCAGCTTGTGTTTGTCTCGCAGCCGGTCTGGAGCGAAACTTATCGGAACCGACAAC
TCTGTGTCCTCTCTCGGAAATACACCTCCTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCTCT
TGCTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCCCTCCCGGGCCGCTTGGGGCTTACCGTCCCTTCTTTCATCTGC
CGTTCCGGCAACCCAGGGGCGAACCTCTCTTTACGCGGTCTCCCGTCTGTGCTTCTCACCTGCCGGACCGTGTGCTTTC
GCTTACCTCTGCACGTGCGATGGAGACCACCGTGAACGCCACCGAGCCCTTGCCAAGGTCTTACATAAGAGGACTCTTGGGA
CTCTCAGCAATGTCAACGACCGACCTTGAAGCGTACTTCAAAGACTGTTTGTGTTAAGGACTGGGAGGAGTTGGGGGAGGAGAT
CAGGTTAATGATTTATGTACTAGGAGGCTGTAGGCATAAATTGGTCTGTTTACCAGCACCAT
1815

1965

ACTTCTTTCCCTTCTGTTCRAGATCTCCTCKACACCTCCTCTGCTCTGTATCGGGAGGCCCTTAGAGTCTCCGGAACATTTGTTCA
CCTCATCATAACGAATCAGGCAAGCTATTCTGTGTTGGGGTGAGTTGATGAATCTGGCCACCTGGGTGGGAAGTAATTTGGA
CGATCCAGCATCCAGGGACGCAGTAGTCACCTATGTCAATGTTAATATGGGCCAAAAATTCAGACAACCTACTGTGGTTTCA
TTTCTGTCTTACTTTTGGAAAGAGATACTGTTCTTGTGATTTTGGTGTCTTTCCGAGTGTGGATTTCGACTCCTCCTGCTTAC
AGACCACCAAAATGCCCTATCTTATCAACACTCCGGAAA
2336

9.7.30 Patient Y6

1174

TTGCCAAGTGTGTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCCATAGGCCATCAGCGCATGCGTGAACCTTTGTGTCTC
CTCTGCCGATCCATACTGCGGAACCTTAGCCGCTTGTGTTTGTCTCGCAGCAGGTCTGGGGCAAANCTCATCGGGACTGACAAT
TCTGTGCTGCTCTCCCGCAAGTATACATCCTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCTCT
TGTTTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCCCTCCCGGGCCGCTTGGGGCTTACCGCCCGCTTCTCCGCTCT
TGTAACGACCGACCGGGGCGCACCTCTCTTTACGCGGACTCCCGCTGTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTC
GCTTCACTCTGCACGTGCGATGGAGACCACCGTGAACGCCACAGGAACCTGCCAAGGTCTTGCATAAGAGGACTCTTGGGA
CTTTCGGCAATGTCAACGACCGACCTTGAAGCATACTTCAAAGACTGTGTGTTAATGAGTGGGAGGAGTTGGGGGAGGAGGT
TAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTGTGTTTACCAGCACCATCTCAACTTTTTACCTCTGCCT
AATCATCTCTTGTTCATGTCTTACTTTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTAGGGCATGGACATTGACCCGTATA
AAGAATTTGGAGCTTCTGNAGGAGTACTCTCTTTTTTGCCTTCTGACTTCTTTCCCTTCCATTTCGAGATCTCCTCGACNCCGCC
TCTGCTCTGTATCGGGAGGCCCTTAGAGTCTCCGGAACATTGTTTCACTCACCATACGGCAATCAGGCAAGCTATTTTGTGTTG
GGGTGAGTTGATGAATCTAGCCACCTGGGTGGGAGGTAATTTGGAAGATCCAGCATCCAGGGAATTAGTAGTCGGCTATGTCA
ACGTTAATATGGGCCAAAAATCAGACAACCTATTGTGGTTTACATTTCTGTCTTACTTTTGGGAGAGAAAGTTGTTCTTGA
TATTTGGTGTCTTTTGGAGTGTGGATTTCGACTCCTCCTGCATATAGACCACAAAAATGCCCTATCTTATCAACACTTCCGGA
AACTACTGTTG
2345

9.8. Sequences of virus-cell DNA junctions

This appendix lists all the detected virus-cell DNA junctions detected by invPCR (as described in Sections 2.14 and 2.21) of DNA extracts from liver tissue fragments (Section 5.3.3), liver tissue sections (Section 5.3.5) and hepatocyte foci isolated by laser-microdissection (Section 7.3.4) of patients with chronic HBV infection (listed in Table 4.1).

9.8.1. Patient C

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
C i	1788	ATTAAAGGTCCTTTGTTATTAGGAG SCTGTAGGCTCTCACCAACGTTAG CTTCCAGAGTGGCGTTT	73000	>100000	73000	7	51	33	Guanine nucleotide binding protein (G protein) alpha 12	Slide	NcoI	16(3-85)	Y
C ii	1799	TGTATATAGGAGGCTGTAGGCAT AAATTTGTTCCCGAGGCTGGAGTGT CAGTAGTCAAATCTTGGC	2100	71600	2100	5			360821 bp at 5' side: CD180 antigen precursor 668831 bp at 3' side: phosphatidylinositol 3-kinase regulatory subunit alpha isotype 1	Slide	NcoI	3(1-60)	N
C iii	1800	TGTAFTAGGAGGCTGTAGGCATA AAATTTGTTCCCGAGGCTGGAGTGT TAGACAGGGTTTCACC	3600	>100000	3600	1			102452 bp at 5' side: mitogen-activated protein kinase kinase MLK4 128681 bp at 3' side: potassium channel subfamily K member 1	Slide	NcoI	3(1-60)	N
C iv	1802	TATTAGGAGGCTGTAGGCATAAA TTGGTCTGATGGGATTTTGGT GTAACACCATGACAAAT	12700	10600	10600	2			52279 bp at 3' side: unknown protein	Slide	NcoI	8(1-68)	Y
C v	1806	AGGAGGCTGTAGGCATAAAATTGG TCTGGCACTTCAGTAGTGTCTCT TCAGGAGGTTTTAGAA	86300	15200	15200	2			2293824 bp at 5' side: neuexin-1-beta isoform beta precursor 1030569 bp at 3' side: ankyrin repeat and SOCS box protein 3 isoform b	Slide Fragment	NcoI NcoI	8(1-68) 10(1.4-73)	Y Y
C vi	1816	AGGCATAAATGGTCTGGCCACC ACCACATGAGGCTGTGCACAG GCAGGTCCTCAACCTTGG	100	>100000	100	17	217	125	24358 bp at 5' side: septin-9 isoform f 200330 bp at 3' side: hypothetical protein LOC100132174	Slide	NcoI	8(1-68)	Y
C vii	1818	SCATAAATGGTCTGGCCACCAG CACCATGGAGTTGACTCAATAGG GAAACTGAGAGCAGTTC	87300	>100000	87300	6	0	0	70890 bp at 5' side: HLA class II histocompatibility antigen, DP beta 1 chain ... 6405 bp at 3' side: collagen alpha-2(XI) chain isoform 1 preproprotein	Slide	NcoI	8(1-68)	Y
C viii	1825	TGGTCTGGCCACAGCACCATG CAACTTATGAGTTCCCTAAGT TAACTACAGACACTTT	21100	>100000	21100	4	0/0	0/0	42100 bp at 5' side: kappa-casein precursor 43375 bp at 3' side: testis development protein NYD-SP26	Slide	NcoI	48(15-157)	Y
C ix	1826	TGGTCTGGCCACAGCACCATG AACTTTTCTGTTTTTGTFTA CGACGAGTCCAGGAGGA	65200	13000	13000	1	62/0	38/0	22221 bp at 5' side: extracellular matrix protein 1 isoform 2 precursor 16514 bp at 3' side: ADAMTS-like protein 4 isoform 1	Fragment	NcoI	41(15-110)	Y
C xi	1773	GGGGAGGAGATTAGATTAAGG TCTTTGTACCACCCCTCAGGAA ATGGAGAACTAGTGTGG	47700	0	0	1	0/72	0/48	18402 bp at 5' side: hypothetical protein LOC284546 70441 bp at 3' side: RING finger protein 11	Fragment	NcoI	31(10-98)	Y
C xii	1788	ATTAAGGTCCTTTGTTATTAGGAG GCTGTAGGCAATAATAGATCAATC TACTATAAAGAACAATGC	48900	9200	9200	6	52	62	3000bp at 3' side NADP-dependent malic enzyme	Fragment	NcoI	10(1-73)	N
C xiii	1825	TGGTCTGGCCACAGCACCATG CAACTTTTATCTCCCAATGCT ATCCCTCCCGCTCCCA	81900	27800	27800	1	0	0	neuronal growth regulator 1 precursor	Fragment	NcoI	10(2-73)	N

9.8.1. Cont.

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosomes ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
C xiv	1706	ANTGCAAGCAGCCACCTTGAGG CATACTGGACCCAGTTAGAATGG CAATCATTAATAAAGTCAG	10500	>100000	10500	2			573631 bp at 5' side: A-kinase anchor protein SPHKAP isoform 2 161632 bp at 3' side: PTB-containing, cubilin and LRP1-interacting protein isof...	Fragment	Ncol	10(2-73)	N
C xv	1827	GTCTGGCAGCAGCACCATTGCA CTTTTTCAGTGCCTTAACAATTT TTTCCTTCATTTCAATTT	>100000	15700	15700	4			1179292 bp at 5' side: bifunctional heparan sulfate N-deacetylase/N-sulfotransferase... 825927 bp at 3' side: translocating chain-associated membrane protein 1-like 1	Fragment	Ncol	90(12-660)	N
C xvi	1719	CCGACCTTGAAGCATACTTCAA GACTGTTTGAACCCATAAAAACC CTAGAAGAAAACCTAGGC				Contig			Contiguous	Fragment	Ncol	670 (310-1500)	Y
C xvii	1800	TGTAATGAGAGCCTGTAGNNA NATTTGNCCTTTTGTATTNNNN TAGACACAGGGTTCCACC	3600	>100000	3600	1			102452 bp at 5' side: mitogen-activated protein kinase kinase MLK4 128681 bp at 3' side: potassium channel subfamily K member 1	Fragment	Ncol	93(13-660)	N
C xviii	1821	AGGCATAAATTTGGTCTGGGCAC AGCACCAATCACTTTTACAGT TCGGCTTGTAGAAGTCA	4900	>100000	4900	5			566781 bp at 5' side: casein kinase I isoform gamma-3 isoform 4 456504 bp at 3' side: zinc finger protein 608	Fragment	Ncol	93(13-660)	N
C xx	1826	TGTTCTGGCCACCCAGCACATGC AACTTTTCCCAACTCCCGGTTT CTGAACCTTGAGAGCCNN	91900	70400	70400	2	114/0	62/0	6005 bp at 5' side: dynactin subunit 1 isoform 1 28090 bp at 3' side: hypothetical protein LOC388963	Fragment	Ncol	93(13-660)	N

9.8.2. Patient CN

No virus-cell DNA junctions detected

9.8.3. Patient CY

CY i	1785	TAGATTAATGATCTTTTGTACTGG GAGGCTGTAAATCCTGATCCTTTG CCCTGGGGCAACATCT	92600	>100000	92600	16	20/280	14/163	42296 bp at 5' side: hypothetical protein LOC57707 26288 bp at 3' side: coactosin-like protein	Slide	Ncol	104 (6-1882)	N
CY ii	1801	GTAAGCTGTAGGCAATAATTTGGC ATTTGCTCTGAAGATATATAGGAT TTAGCCCTAATAATTTAACT	76300	>100000	76300	1			52961 bp at 3' side: olfactory receptor 4F3/4F16/4F29	Slide	Ncol	104 (6-1882)	N
CY iii	1808	GAGGCTGTAGGCAATAATTTGGC TGTTCCCAACCTTCAATAAGGG CTCGAAAGGACCCCAAAA	>100000	27500	27500	10	10	4	269628 bp at 5' side: bone morphogenetic protein 3B precursor 30022 bp at 3' side: protein tyrosine phosphatase, non-receptor type 20 isoform 8	Slide	Ncol	104 (6-1882)	N
CY iv	1817	GGCATAAATTTGGTCTGTTCACCA GCACCATGGCTTGTCTTTGTT GAGTGGGGAGAAGCTGGT	80200	5900	5900	22	10	9	hypothetical protein LOC85379	Fragment	Ncol	800 (360-1800)	Y
CY v	1818	GCATAAATTTGGTCTGTTCACCA CACCATCCAGTTCTTACACACTG AGCAGACATCCTTCCCC	81900	24000	24000	2	0	0	2414 bp at 5' side: cryptic family protein 1B 62367 bp at 3' side: cryptic protein precursor	Slide	Ncol	104 (6-1882)	N

9.8.4. Patient CYRY

Clone ID ¹	HBV site ²	Sequence ³	SIMAR ⁴			Chromosomes ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹
			Upstream	Downstream	Closest								
CYRY i	1800	TGTAATAGGAGCTAGGCATAAA ATTGGTCTCTCATAGCACTGTTT TCACCTTTGTCATTTA	4800	8200	4800	0	4		nuclear receptor coactivator 5	Fragment Ncol	40(6-290)	N	
CYRY ii	1814	GTAGGCATAAATTTGGTCTGGCCAC CAGACCAATGTCATGACTTTTCT CTGCTCTTGGGACCA	13100	10600	10600	20	9		peroxisomal carnitine O-octanoyltransferase isoform 1 peroxisomal carnitine O-octanoyltransferase isoform 2	Fragment Ncol	41(6-290)	N	
CYRY iii	1816	AGGCATAAATTTGGTCTGGCCACCA GCACCATGGCTGTTAAGCTGGGA	94600	47800	47800	0/10	0/4		35666 bp at 5' side: cell cycle control protein 50B 5275 bp at 3' side: protein kinase C eta type	Slide Fragment Ncol	295(38-2419) 41(6-290)	Y N	
CYRY iv	1825	TTGGTCTGGCCACCAAGCACCATGC AACTTTTTGAACCTATGAAA	72500	>100000	72500	11			122631 bp at 5' side: serine/threonine-protein kinase PAK 1 isoform 1 74830 bp at 3' side: aquaporin-11	Slide Fragment Ncol	104(6-1882)	N	
CYRY v	1827	GGTCTGGCCACCAAGCACCATGCAA CTTTTTCAATTCAAAACCTTTGCA AAATTTTG	7600	>100000	7600	3			77318 bp at 5' side: family with sequence similarity 19 (chemokine (C-C motif)... 111302 bp at 3' side: family with sequence similarity 19 (chemokine (C-C motif)...)...	Fragment Ncol	40(6-290)	N	
CYRY vi	Unkn.	AGCGAATAAAGTTTCAA				Contig.			Contiguous	Slide Ncol	104(6-1882)	N	

9.8.5. Patient DG

DG i	1960	GAGCTTCTGGAGTTACTCTCTT TTTTGTCTAAGGCATTAAGAAGAG AATTTTCCCCCTTGC	66900	82200	66900	11	0		C11orf16 protein	Fragment Ncol	610 (340-1100)	Y
DG ii	1949	TAAAGAATTTGGAGCTTCTGGCGGA GTACTCTATAGCCAAATAAAT TAAFTGTATATTTTAA	26000	8800	8800	8			55280 bp at 5' side: dihydropyrimidinase-related protein 2 isoform 2 33706 bp at 3' side: alpha-1A adrenergic receptor isoform 2	Fragment Ncol	40(6-290)	N
DG iii	1738	GAGTTGGGAGGAGATTAGTTA ATGATCTTGAAGAAATGCAAAATG TAAAGTCTGTGAGTT	12000	>100000	12000	1	48/0		5303 bp at 5' side: SH3-containing GRB2-like protein 3-interacting protein 1 6222 bp at 3' side: tctex1 domain-containing protein 1	Fragment Ncol	40(6-290)	N
DG iv	1802	GCTGTAGGCATAAATTTGGTCTGTT CACAGCATAAATAAAGTTGGCA AAATAGCAFTGCTGG	>100000	90500	90500	7	33/20		49156 bp at 5' side: 28S ribosomal protein S33, mitochondrial 14106 bp at 3' side: transmembrane protein 178-like	Fragment Ncol	40(6-290)	N
DG ix	1801	TGGTCTGTGGGCCAATGGCTCT AGGCTGTG	23700	27500	23700	16	52		29999 bp at 5' side: protein BANP isoform e	Slide Ncol	6450 (2419-17699)	Y
DG v	1774	AGGATTTGGGAGGAGATTAGGT TAAATGATCTTTTGTCTCTTTCATG TCCTTTTCTATTTCT	7000	44700	7000	1			520778 bp at 5' side: peptidylprolyl isomerase A (cyclophilin A)-like 4F 225242 bp at 3' side: peptidylprolyl cis-trans isomerase A-like 4A/B/C	Fragment Ncol	40(6-290)	N
DG vi	1826	TGGTCTGTACCCAGCATCATGCA ACTTTTTCTCTCTACCAAGCTG GAGTCCAGTAGCAATC	91600	>100000	91600	1	24		metal regulatory transcription factor 1	Fragment Ncol	40(6-290)	N
DG vii	1731	TCTCAGCAATGCAACGACCGACC TTGAGCAACTAGCCTCTGTTTG GATCAATGTTGCCAAA	50500	78300	50500	6	0		tumor necrosis factor receptor superfamily member 21 precursor	Fragment Ncol	240(60-100)	Y
DG viii	1906	GAGCTTCTGGAGTTACTCTCTT TTTTGTCTAAGGCATTAAGAAGAG AATTTTCCCCCTTGC	66600	82500	66600	11	0		hypothetical protein LOC56673	Slide Ncol	7645 (2998-19909)	Y

9.8.5. Cont.

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/hear ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
DG x	1823	AATTGGTCTGTTACCCAGCACCA TCCAGAAACTCTGACAAAGGA ATGAAATAAATAAAGT	25900	>100000	25900	4			1862173 bp at 5' side: polyadenylate-binding protein 4-like 1457470 bp at 3' side: protocadherin-18 precursor	Slide	NcoI	7239(2703-18793)	Y
DG xi	1815	TAGGCATAAATGGTCTGTTTAC CAGCACCAATTTCTTTG				Contig			Contiguous	Slide	NcoI	460(60-3617)	Y
DG xii	1790	TAAAGGCTTTTACTGGAGGC TGTAGCATGTAACCTGTAGGT TCAAATCCTTGGGAGAA	>100000	12200	12200	5	83	77	ras GTPase-activating-like protein IQGAP2	Slide	NcoI	281635 (135137-559679)	Y
DG xiii	1826	TGGTCTGTTCCACAGCACCATGC AACTTTTCTTTCTAGCTTGTAG GGGAGATATTCCTTTA				Contig			Contiguous	Slide	NcoI	157(8-2998)	N
DG xiv	1763	GGGAGGCTTGGGGAGGAGATT AGGTTAATGTGTAC				Contig			Contiguous	Slide	NcoI	4615 (1559-13546)	Y
DG xv	1763	GGGAGGCTTGGGGAGGAGATT AGGTTAATGTGTACTTGTACTGG AGGCTGTA	38000	16600	16600	19	41	43	U6 snRNA-associated Sm-like protein LSm4	Slide	NcoI	758(125-4615)	Y
DG xvi	1796	TCTTTGTACTGGAGGCTGTAGG CATAAATGGCAAAACATAATA AAGCTACATTTTAACT	>100000	42800	42800	8			hypothetical protein LOC116328 isoform 1	Slide	NcoI	460(60-3617)	Y

9.8.6. Patient FMC1

FMC1 i	1827	GGTGTGTTCACTGCACCAATGCA ACTTTTTCAGTCTTACCAACCAC TGCAATGGCCCTTTGGA	48900	>100000	48900	14	0	0	59417 bp at 5' side: dapper homolog 1 isoform 1 551228 bp at 3' side: disheveled-associated activator of morphogenesis 1	Fragment	NcoI	67(10-470)	N
--------	------	---	-------	---------	-------	----	---	---	---	----------	------	------------	---

9.8.7. Patient FMC2

No virus-cell DNA junctions detected

9.8.8. Patient FMC3

FMC3 i	1812	CTGTAGGCATAAATGGTCTGTT CACCAGCACATGTTCTCTTTGG ACCAATAATTCCTCAT	>100000	68400	68400	1			219767 bp at 5' side: dermatopontin precursor 152974 bp at 3' side: sodium/potassium-transporting ATPase subunit beta-1	Fragment	NcoI	510 (240-1100)	Y
FMC3 ii	1825	TCTGTTACCAGCACCATGCAAC TTTTTTGTTTCTCTTCAAC TGTTCRAACCTTGCCTG	2300	>100000	2300	5			398562 bp at 5' side: histone-lysine N-methyltransferase PRDM9 560813 bp at 3' side: cadherin-10 preproprotein preproprotein	Fragment Fragment	DpnII NcoI	830 (330-2400) 410 (100-1700)	Y Y
FMC3 iii	1774	GGGAGGAGATGAGGTTAAGGTF CTTGTACTATGGAGAAATAAT TTGCAATCTACTCATCTG	17600	96200	17600	12			876917 bp at 5' side: protein NEDD1 isoform c 681679 bp at 3' side: thymopoietin isoform beta	Fragment	NcoI	67(10-470)	N
FMC3 iv	1803	TACTAGGAGGCTGAGGCATAAA TTGGTCTGTTAAGCTCCAAGTC AGATTGCAATTTAAAT	46700	>100000	46700	2	0/20	0/14	12146 bp at 5' side: dual specificity protein phosphatase 19 isoform 2 16444 bp at 3' side: nucleoporin NUP53	Fragment	DpnII	67(10-470)	N
FMC3 v	1775	GGGAGGATTAGGTTAAGGTC TTTGTACTGAAAAGTTTGTGGA AACATGGAAATTAATAAT	64000	>100000	64000	3	41	57	vacuolar protein sorting-associated protein 8 homolog isoform 1. vacuolar protein sorting-associated protein 8 homolog isoform 2	Fragment	DpnII	67(10-470)	N

9.8.9. Patient FMC4

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
FMC4 i	1722	CGACCTTGAGTCAATCTTCNNN NANNNGTGTATACCTCCTCT CCTTCTTCCCTATTGTG	65600	54200	54200	16			483552 bp at 5' side: NEDD4-binding protein 1 180721 bp at 3' side: cerebellin-1 precursor	Fragment	NcoI	67(10-470)	N

9.8.10. Patient FMC5

FMC5 i	1819	GGTCTGGCACCAGCACCATGCG GACTGTCTCAAAAAATAAATTA AAAAAATAAATCTCTGTAT	62800	>100000	62800	13	51	28	7918 bp at 5' side: serine/threonine-protein kinase Nek3 isoform b 125188 bp at 3' side: hypothetical protein LOC100652836	Fragment	NcoI	1200 (700-2100)	Y
FMC5 ii	1815	AATTGGTCTGGCCACAGCACCC ATGCAGCTTAAAAACTTGACC ATTCGCCCTCCTCCTCCTCC	32300	54700	32300	10	10	9	25201 bp at 5' side: pterin-4-alpha-carbinolamine dehydratase precursor 298906 bp at 3' side: netrin receptor UNC5B precursor	Fragment	NcoI	67(10-470)	N
FMC5 iii	1827	GGTCTGGCACCAGCACCATGCG AACTTTTTCAAGAAGACTCTTT ACTTTTTAATTAATAACTA	9300	>100000	9300	6	103	81	inhibitor of Bruton tyrosine kinase	Slide	NcoI	9640 (1632-51599)	Y
FMC5 iv	1818	TGGTCTGCGCACAGCACCATG CAACTTTTTCTGCCCTTTTTTA ACTAGGTTACTAGCCGCCCC	>100000	99400	99400	16	20	28	NFATC2-interacting protein	Slide	NcoI	9782 (1882-51599)	Y
FMC5 v	1788	AAAGTCTTTGTACTAGGAGGC TGTFAGGCAAGTCTTCGGGG TTCTTGGATGCGAGAGGGGA	>100000	21600	21600	7	10	4	79517 bp at 5' side: putative protein cTAGE-6 16654 bp at 3' side: hypothetical protein LOC9747 isoform 1	Slide	NcoI	1632 (95-31524)	N
FMC5 vi	1800	TGTFAGGAGGCGTGTAGGCAT AAATTGGTGGCTCCAGCCCGA AGGCAGTGGAGCCCCCATTA	12900	>100000	12900	5	0	0	foliastatin-related protein 4 precursor	Slide	NcoI	8906 (1607-50059)	Y
FMC5 vii	Unkn.	CACTCTGANTACTGNCTTANA CCTCCCTGC	4200	>100000	4200	8			89813 bp at 5' side: arf-GAP with SH3 domain, ANK repeat and PH domain-contains... 288424 bp at 3' side: adenylate cyclase type 8	Slide	NcoI	1632 (95-31524)	N

9.8.11. Patient FMC6

FMC6 i	1763	GGGAGGAGTTGGGGAGGAGGT TAGGTTAATGAATATATTTCAA AATAGAAATGTTGACGGTATG	>100000	55100	55100	6	0	9	bone morphogenetic protein 5 preproprotein preproprotein	Fragment	NcoI	690 (650-1400)	Y
FMC6 ii	1824	AATTGGTGTTCACCCAGCACCA TGCAACTTTCCGACTGTCTG CTGTACCAGAGAACATTTGC	43300	29800	29800	2	134	163	11074 bp at 5' side: methylmalonic aciduria and homocystinuria type D protein,... 871693 bp at 3' side: rho-related GTP-binding protein RhoE precursor	Fragment	NcoI	67(10-470)	N
FMC6 iii	1810	GGCTGAGGCATATAATTGGTGC GTTACCAAGCTAACACTAATAT AGCCTTCATTTTATAGCCCAA	>100000	45400	45400	11	0	0	67361 bp at 5' side: calcitonin CT preproprotein CT preproprotein 35570 bp at 3' side: calcitonin gene-related peptide 2 precursor	Fragment	NcoI	1200 (670-2100)	Y

9.8.12. Patient FMC7

No virus-cell DNA junctions detected

9.8.13. Patient GSI

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosomes ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
GS1 i	1749	AAGGCTTTGTACTGGGAGGCTG TAGGCATAAGCCCTGAATAAAGTC AATTGCATGTAAAGAA	63000	5100	5100	1	0	19	nicotinamide mononucleotide adenylyltransferase 2 isoform 1	Slide	NcoI	41(6-270)	N

9.8.14. Patient GS2

No virus-cell DNA junctions detected

9.8.15. Patient GS3

No virus-cell DNA junctions detected

9.8.16. Patient GS4

No virus-cell DNA junctions detected

9.8.17. Patient GS5

No virus-cell DNA junctions detected

9.8.18. Patient HN

HN i	1762	TGGGAGGTCGGGGAGGAGAA TAGATTAAAGTTTATTTCTGC AAACAAGTTGTGCAATT	63800	>100000	63800	5			2704519 bp at 5' side: methionine adenosyltransferase 2 subunit beta isoform 1 1061692 bp at 3' side: teneurin-2	Fragment	NcoI	40(6-290)	N
HN ii	1807	GGAGGCTGTAAGCATATAATTGGT CTCGCACCTGGAGCTCCCTTCAT CTGTTACACTGGAGACGT	>100000	83200	83200	2			816649 bp at 5' side: beta-galactoside alpha-2,6-sialyltransferase 2 isoform b 166105 bp at 3' side: ranBP2-like and GRIP domain-containing protein 4	Fragment	NcoI	40(6-290)	N
HN iii	1902	TCCAAGCTGGCTTGGGTGGCT TTAGGACATGGACATTGATCAAG CGGGCTTCATCCCTTGAA TGCAA G	>100000	37700	37700	8	0	0	collagen alpha-1(XIV) chain precursor	Fragment	DpnII	41(6-290)	N
HN iv	1911	TGCCTTGGGTGGCTTTAGGACAT GGACATGATCGTCTGAAAGCCTT CTTCTCAGCTCGTCAA	3100	>100000	3100	1	0	0	32801 bp at 5' side: hypothetical protein LOC84970 isoform b 505538 bp at 3' side: gap junction beta-5 protein	Fragment	DpnII	20000 (4900-79000)	Y

9.8.19. Patient HS

No virus-cell DNA junctions detected

9.8.20. Patient L

L i	1776	TGGTCTGTCCACGACACCATGC AACTTTCCATGGCTGCTAGGCT GGA	34800	>100000	34800	19	0/0	0/0	490 bp at 5' side: CD177 antigen precursor 43680 bp at 3' side: testis-expressed protein 101 isoform 1	Slide	NcoI	95(157-1882)	N
L ii	1786	AGGCATAAAATGGTCTGTTCACC AGCACCATGCTGCTAGGCTGTA AAC	>100000	10400	10400	9			198097 bp at 5' side: tumor suppressor candidate gene 1 protein 218161 bp at 3' side: putative deoxyuridine 5-triphosphate nucleotidohydrolase...	Slide	NcoI	95(157-1882)	N
L iii	1805	TGGGAGGCTGTAGGCATAAATTG GTCGTGTCA TATA TAGAGAGA GAGAGAGAGAGAGAGA				Contig.			Contiguous	Slide	NcoI	1124 (309-4273)	Y
L iv	1825	TGGTCTGTTCACGACCAATG CAACTTTTTTTTTTCGGTGGCCT TGGGGCTTCCT	88700	>100000	88700	11	186	105	mitogen-activated protein kinase kinase kinase 2	Slide	NcoI	95(157-1882)	N

9.8.21. Patient MH

Clone ID ¹	HBV site ²	Sequence ³	SMAR ⁴		Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream								
MH i	Unkn.	TATAATTCACATTCCTGGCCCTCC CATATAFC	8000	5900	4			2587552 bp at 5' side: polyadenylate-binding protein 4-like 732941 bp at 3' side: protocadherin-18 precursor	Fragment	NcoI	3300 (2000-5500)	Y

9.8.22. Patient NT

No virus-cell DNA junctions detected

9.8.23. Patient SAO

SAO i	1662	CTTACATAAGAGGACAGAAAAGT TTTTAAGCCCTGATTATGCTAT	>100000	28800	4	0	4	1795 bp at 5' side: kelch-like protein 8 107370 bp at 3' side: 17-beta-hydroxysteroid dehydrogenase 13 isoform A	Fragment Fragment	NcoI NcoI	510(270-950) 40(6-290)	Y Y
SAO ii	1757	AAGACTGGAGGAGTTGGGGAGG AGATTAGAGAAATGTTATATCACC TAGGT	71900	87200	3			104922 bp at 5' side: nephrocystin-3 402019 bp at 3' side: transmembrane protein 108 precursor	Fragment	DpnII	83(21-330)	Y
SAO iii	1763	GGGAGGAGTTGGGGAGGAGATTA GATTAAGTGAAAATATTTTCA TTGCATATATGTAACCA	32000	28800	6			2052963 bp at 5' side: protein eyes shut homolog isoform 3 1089827 bp at 3' side: brain-specific angiogenesis inhibitor 3 precursor	Fragment	DpnII	130(40-390)	Y
SAO iv	1788	GATTAAGGCTTTGTTATAGGAG GCTGTAGGCTTCTCTTCTTCT TTCTCTCTCTCTCT	>100000	23700	3			570955 bp at 5' side: otolin-1 precursor 2904144 bp at 3' side: sucrose-isomaltase, intestinal	Fragment	DpnII	130(43-400)	Y
SAO xvi	1810	GGCTGTAGGCATAAATGGTCTGC GCACCCACAGAGTAGGAGAGTA ATCAATCAATAGAAC	50900	>100000	6			1105570 bp at 5' side: protein quaking isoform HQK-5 568672 bp at 3' side: hypothetical protein LOC728316	Fragment	NcoI	83(21-330)	Y
SAO ix	1814	GTAGGCATAAATGGTCTGGGCAC CAGCACAGAAATATGGCATGT ATTTCCCTCCACCTT	27900	6300	3	0	0	family with sequence similarity 19 (chemokine (C-C motif)-like), member A1 (FAM19A1)	Fragment	DpnII	170(61-460)	Y
SAO v	1798	TTTGTATTAGAGGCTGTAGGCAT TGAATGTGACCTCCAAA	59400	>100000	18			2253773 bp at 5' side: cadherin-2 preproprotein preproprotein 563347 bp at 3' side: desmocollin-3 isoform Dsc3a preproprotein	Fragment	DpnII	40(6-290)	N
SAO vi	1808	GAGGCTGAAAAGCAFAAATGGTCT GCGCACAGAGAGTAGATCTC AGTGAATCCGTGCATG	73900	>100000	12			2503395 bp at 5' side: transmembrane protein 132B 108984 bp at 3' side: transmembrane protein 132C	Slide	DpnII	37(2-702)	N
SAO vii	1811	CACAGCAAGGATCTCCAAAATC TTTTCATCTTCCAAAG	19700	>100000	7	242	308	caldesmon isoform 2 caldesmon isoform 3	Fragment	DpnII	41(6-290)	N
SAO viii	1814	GTAGGCATAAATGGTCTGGGCAC TATATGATATATATA	56800	>100000	13			75515 bp at 5' side: 5-hydroxytryptamine receptor 2A isoform 1 971825 bp at 3' side: succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrion...	Fragment	NcoI	40(6-290)	N
SAO x	1821	TAAATGGTCTGGCCACCCAGCAC ATGCAACTATAGATGATATTTCC TTAATGCAAGAGGCCA	68400	100	7			74240 bp at 5' side: ETS translocation variant 1 isoform f 88680 bp at 3' side: diacylglycerol kinase beta isoform 1	Fragment	NcoI	40(6-290)	N
SAO xi	1824	ATTGGTCTGGCCACCCACCATG CAACTTTTAAAGCTACTTTGGGAG TTCCAGTAAAGTATTAG	29900	2400	3	51	33	TSC22 domain family protein 2	Fragment	DpnII	220(93-540)	Y

9.8.23. Cont.

Clone ID ¹	HBV site ²	Sequence ³	SIMAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
SAAO xii	1826	TGGTTCGCCACACAGCACCATGC AACTTTTCTTAGTCCAGCTAT TCGAGAGGCTGAGCGGG	>100000	90600	90600	13	134	77	chronic lymphocytic leukemia deletion region gene 6 protein isoform 1	Fragment	DpnII	40(6-290)	N
SAAO xiii	1826	TGGTTCGCCACACAGCACCATGC AACTTTTTCGGAATGGGACTAT AGGGAFTATTATTTTT	64200	>100000	64200	17	19/0	24/0	30601 bp at 5' side: cell division protein kinase 12 isoform 1 43449 bp at 3' side: neurogenic differentiation factor 2	Fragment	DpnII	82(20-330)	Y
SAAO xiv	1826	TGGTTCGCCACACAGCACCATGC AACTTTTTCCTTCAAACGTGTT TCCTTGAATAATGGCCTT	>100000	83000	83000	19	435/20	274/8	12717 bp at 5' side: potassium-transporting ATPase alpha chain 1 36363 bp at 3' side: HAUS augmin-like complex subunit 5	Fragment	DpnII	41(6-290)	N
SAAO xv	1832	GCGCACCGACCATCCAACTTTT TTACACTTCTTTCAGATTCCAC TTATGTACCTAAAGGAGC	57200	22300	22300	13	0	0	382830 bp at 5' side: Krueppel-like factor 12 36641 bp at 3' side: hypothetical protein LOC100288208	Fragment	DpnII	40(6-290)	N

9.8.24. Patient WN

WN i	1684	AAAGGACTCTTGGACTCCAGCA ATGCTCTACTCACCAGCAAG CTCTTTTAGGGGA	11700	500	500	X	0/0	0/0	34672 bp at 5' side: transcription factor Dp family member 3 49755 bp at 3' side: glypican-4 precursor	Fragment	NcoI	40(6-290)	N
WN ii	1768	GAGCTGGGGAGAGATTAGGTC ATGATCTTCAGGACAAATTG AGCGGAATGGCACTC	>100000	81400	81400	15			160626 bp at 5' side: gamma-aminobutyric acid receptor subunit gamma-3 precursor 61789 bp at 3' side: P protein	Fragment	NcoI	40(6-290)	N
WN iii	1805	TAGAGGCTGTAGGCATAAATTG GTCTGGCATGGCAGACCCGGT AATCGCATAAAACTTAAA				mtDNA			Mitochondrial DNA	Fragment	NcoI	40(6-290)	N
WN iv	1816	AGCATAAAATGGTCTGGCACC AGCACCATTAATCTTCTCACCT TTTTATTCGTAAGCCCTC	72100	65800	65800	6	10	28	432992 bp at 5' side: hairyenhancer-of-split related with YRPW motif protein 2 13099 bp at 3' side: HD domain-containing protein 2	Fragment	NcoI	40(6-290)	N
WN v	1818	GCATAAATGGTCTGGCACCAG CACCATGCATGCTAGTAAAGGA TCGTTCAACTCTGTGAGT				Contig.			alpha satellite DNA	Fragment	NcoI	40(6-290)	N
WN vi	1826	TGGTTCGCCACACAGCACCATGC AACTTTTTCGACAGTCCACAAA TCCTGGCACCCCTCCT	>100000	11700	11700	6	0	0	110783 bp at 5' side: G1/S-specific cyclin-D3 isoform 4 2986 bp at 3' side: transcription initiation factor TFIID subunit 8	Fragment	NcoI	40(6-290)	N
WN vii	1826	TGGTTCGCCACACAGCACCATGC AACTTTTTCGGTCTCAAACCC TATCTTCTCCTCAGGGATG	>100000	31600	31600	4			4364462 bp at 5' side: N(4)-(beta-N-acetylglucosaminy)-L-asparaginase isoform 2 517092 bp at 3' side: teneurin-3	Fragment	NcoI	40(6-290)	N
WN viii	1827	GGTTCGCCACACAGCACCATGC ACTTTTTCATAGGAGAAAAGAA TTAA GACTATCTGTTTTT	41000	3800	3800	9			1472645 bp at 5' side: receptor-type tyrosine-protein phosphatase delta isoform 2487387 bp at 3' side: 5,6-dihydroxyindole-2-carboxylic acid oxidase precursor	Fragment	NcoI	40(6-290)	N
WN ix	Unkn.	ACAAATTACTTTGAGATTTGACCT TAGATTGAT	5800	>100000	5800	8			2048991 bp at 5' side: KH domain-containing, RNA-binding, signal transduction associated 3 436377 bp at 3' side: hypothetical protein LOC51059	Fragment	NcoI	40(6-290)	N

9.8.25. Patient XA

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
XA i	1701	ACTCGAGCAATGTC AACGACCG ACCTTGAGTCACTGAATTTGA						Contiguous	Fragment	DpnII	40(5.6-290)	N	
XA ii	1714	GACTCGACCAATGTC AACGACCG GACCTTGAGTTCAGTGAAATTTGA ATTTTATAAAGTGAACA	18200	>100000	18200	1100	967	Homo sapiens fibronectin 1 (FN1)	Fragment	DpnII	2300 (590-9200)	Y	
XA iii	1800	TAAATTTGGTCTGTTCCACGACCG CATGCAACTTATAGGCATGTTCA GTTTCTATTCATCTAGCA	13800	>100000	13800	0/176	0/134	3443 bp at 5' side: meiotic recombination protein SPO11 isoform b 6928 bp at 3' side: mRNA export factor	Fragment	DpnII	6000 (2400-15000)	Y	
XA iv	1820	ATAAATTTGGTCTGTTCCACGACCG CCATGCAACACTGGATATTTGAG GAGTTTGAATAAAGGCTTT	34500	>100000	34500	4		554225 bp at 5' side: glucosamine-6-phosphate isomerase 2 764503 bp at 3' side: gamma-aminobutyric acid receptor subunit gamma-1 precursor	Fragment LMD LMD	DpnII DpnII DpnII	5800 (2400-14000) 29575 (6067-59530) 104(6-2144)	Y Y Y	
XA v	1817	GGCATAAATTTGGTCTGTTCCACCGA GCACATGGGGAAGATAGCCATT GTAGAAATTTCTCTAACTT	72500	>100000	72500	11	86	ciliary neurotrophic factor	Fragment Fragment Fragment Fragment Fragment Fragment	NcoI NcoI NcoI NcoI DpnII NcoI	980 (590-1600) 40(5.6-290) 41(5.6-290) 4500 (2800-7200) 6100 (2500-15000) 3600 (2200-5900)	Y Y Y Y Y Y	
XA vi	1749	TTGTTTAAAGGACTGGGAGGAGTT GGTGGAGGGAAGTGTATATTAG GCCCAAAAGCAACGGTGC	54800	79700	54800	4	38/0	9628 bp at 5' side: fibroblast growth factor 5 isoform 2 precursor 39558 bp at 3' side: hypothetical protein LOC255119	Fragment LMD LMD	NcoI DpnII DpnII	8479 (3617-19349)	Y	
XA vii	1720	ACCGACTTTGAGGCAACTTTCAA GCACCTGTTTAAAGTCTCAGCCTA CACGGTACTGTCGGGCTC	>100000	85700	85700	19	0/43	25030 bp at 5' side: immunoglobulin-like domain and transmembrane domain-containing protein 1 4779 bp at 3' side: U6 snRNA-associated Sm-like protein LSm7	LMD	DpnII	1882 (660-6450)	Y	
XA viii	1804	CTATGAGGCTGTAGGCTAAATTT GCCTGTTCCCTTTTAAAGGTTT CAACAAATTTTCATTG	>100000	3800	3800	12	Unknown/2303	47960 bp at 5' side: scavenger receptor class B member 1 isoform 1 6 bp at 3' side: polyubiquitin-C	Fragment	DpnII	40(5.6-290)	N	
XA x	1817	GGCATAAATTTGGTCTGCGGACCGA GCACCAATGCAACCAATGCTGCTA GGCTG	98400	9500	9500	16	0	sarcoplasmic/endoplasmic reticulum calcium ATPase 1 isoform 1 sarcoplasmic/endoplasmic reticulum calcium ATPase 1 isoform 2	Fragment	NcoI	41(5.6-290)	N	
XA xi	1823	AAATGGTCTGTTCCACGACCGA TCGCAACTTTTATATATATATATA TGAAAAGTTCAGCTCAAAA	800	65100	800	7		466579 bp at 5' side: hypothetical protein LOC100131871 1175863 bp at 3' side: POM121-like protein 12	Slide	NcoI	192(36-1081)	Y	

9.8.26. Patient Y2

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
Y2 i	1814	GTAGGCATAAATGGTCTGTTC ACCACGCCAAAATTTGCACCC ATTTCGATCTCATAGCAA	47500	4800	4800	7	155	14	52544 bp at 5' side: C-C motif chemokine 24 precursor 12892 bp at 3' side: rhomboid domain-containing protein 2 isoform a	Slide LMD	DpnII NcoI	146(8-2998) 40(2-739)	Y Y
Y2 ii	280	TTGGACTGCGAAATTTGGCCA AGACACAGGGTGTCTCCACG TGGCCCTCACATGGCTCCGC	>100000	3800	3800	7	0	0	receptor-type tyrosine-protein phosphatase N2 isoform 2 precursor receptor-type tyrosine-protein phosphatase N2 isoform 3 precursor	Slide	NcoI	146(8-2998)	N
Y2 iii	1657	CTTACATAAGATTAAAGACATA AACTGATAAAATGCCAGTACA	74600	>100000	74600	4			1392935 bp at 5' side: heparan sulfate glucosamine 3-O-sulfotransferase 1 precursor 580155 bp at 3' side: ras-related protein Rab-28 isoform 1	Slide	NcoI	157(8-2998)	N
Y2 iv	1733	CATACTTCAAAGACTGTGTGTT TACTGAGTGGACCTAGACCTGT ATCAGCTGTGGTGTAGTCCGA	>100000	77600	77600	7			136280 bp at 5' side: V-type proton ATPase subunit e 2 isoform 2 232792 bp at 3' side: actin-related protein 3C precursor	Slide	NcoI	146(8-2998)	N
Y2 ix	1873	TCTCTTTTTTGGCTTCTGACTT CTTTCCTTCTATAAAA				Contig.			Contiguous	LMD	DpnII	125(15-956)	Y
Y2 v	1793	ATGAGGTTAAAGGCTCTTTGTAT TAGGAGGCTGCACGCTGTGT TGCTGCATCTCCCATCCCC	16300	17000	16300	6			80407 bp at 5' side: FERM domain-containing protein 1 isoform 2 151131 bp at 3' side: dapper homolog 2	Slide	NcoI	146(8-2998)	N
Y2 vi	1784	TGAGGTTAAAGTCTTTGTATTT AAGAGGCTGTCAAAAATGATGG TTACTACCCAGAGAAATTTT	>100000	18000	18000	2	10	28	29146 bp at 3' side: FAM110C protein	Slide	NcoI	146(8-2998)	N
Y2 vii	1789	TTAAAGGCTTTTGTATTAGGAG GCTGTAGCAGTGTACGTCGT CAGGGCTAAGTCTGTGAG	68500	61100	61100	6	0	0	cAMP and cAMP-inhibited cGMP 3':5'-cyclic phosphodiesterase 10A isoform 1	Slide	NcoI	146(8-2998)	N
Y2 viii	1821	GCATAAATTTGGTCTGTTCACCA GCACCATGATGATAGCCCAT CCTGGTGGCAGTCTTTTTCT	>100000	18800	18800	4	0	0	16020 bp at 5' side: transmembrane protease serine 11B 52424 bp at 3' side: YTH domain-containing protein 1 isoform 1	Slide LMD LMD	DpnII DpnII DpnII	2145 (268-18243) 125(15-956) 40(2-739)	Y Y Y
Y2 x	1825	TTGGTCTGTTACCCAGCACCAT GCAACTTTTTGGGATTCGGGG	75700	5500	5500	12	10/176	14/129	proline-rich protein 13 isoform 1 proline-rich protein 13 isoform 2	LMD	DpnII	125(15-956)	Y
Y2 xi	1791	CTTTGTTTTTAGGAGGCTGTAGG AATAATTTGGAGGTTTACAA CACATCCCCACAGTCA	>100000	50500	50500	11	10/10	43/4	37683 bp at 5' side: nucleobindin-2 precursor 17699 bp at 3' side: ATP-sensitive inward rectifier potassium channel 11	Slide	NcoI	157(8-2998)	N
Y2 xii	1811	GCTGTAGGCATAAATTTGGTCTG TTCACCGCAATAAACAAGAA CTGACTTGGATATATTCAG	37800	>100000	37800	1	51	48	FGGY carbohydrate kinase domain-containing protein isoform a FGGY carbohydrate kinase domain-containing protein isoform b	Slide	DpnII	146(8-2998)	N
Y2 xiii	1825	TTGGTCTGTTACCCAGCACCAT GCAACTTTTTTAGCCTTGGCC TCCAGGCTTAAAGCAATCTT				Contig.			Contiguous	Slide	NcoI	157(8-2998)	N
Y2 xiv	1825	TTGGTCTGTTACCCAGCACCAT GCAACTTTTTTAGCCTTGGCC TCCAGGCTTAAAGCAATCTT	29800	>100000	29800	1	164	96	splicing factor 3A subunit 3	Slide	NcoI	816 (146-4615)	Y
Y2 xv	1827	GGTGTTCACAGGACACATGC AACTTTTTCATAAAGTGTCTCA CTCAGTGTCTCACACAGAA	>100000	46900	46900	1			473751 bp at 5' side: tenascin-R precursor 64338 bp at 3' side: E3 ubiquitin-protein ligase RFW2D2 isoform a	Slide	NcoI	145(15-956)	N
Y2 xvi	1831	TGTTCCACGACCATCCCAACT TTTTCAACTCACTCCACTTTC ACTCTGGCTTTTTATACAGT	5000	73000	5000	8			594660 bp at 5' side: netrin receptor UNC5D precursor 398917 bp at 3' side: potassium channel subfamily U member 1	Slide	NcoI	145(15-956)	N

9.8.27. Patient Y3

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
Y3 i	1722	CGACCGACCTTGAGCCATACCT CAAGACACGTATGCTGTTTGT GTCCTCATTTTACAAAAA	>100000	77500	77500	4	41	28	Slide	NcoI	146(8-2998)	N	
Y3 ii	1756	AATGAGTGGAGGAGTTGGGG AGGAGGTTAGCAAGAGGATT ACTGCTAGCTTTTCTAT	>100000	90700	90700	16			Slide	NcoI	146(8-2998)	N	
Y3 iii	1813	TGTAGGCATAAATGGTGTGT CACCAGACCTAAGAGATGTC CTGACAACCTTTCAGAAAT	>100000	4400	4400	4	0	0	Slide	NcoI	146(8-2998)	N	
Y3 iv	1827	GGTGTGTTCCAGCCATCC AACTTTTCAAATCTTTAGGCT TGGCGATCCCAAGTCTTAG	>100000	29500	29500	3	83	38	Slide	NcoI	157(8-2998)	N	

9.8.28. Patient Y4

Y4 i	1786	AGGTTAAAGGCTTTGTATTAG GAGGCTGTAGAAAAAGGATCA AAC				Contig.			Contiguous	LMD	DpnII	40(2-739)	N
Y4 ii	1788	GGTTAGAGTCTTTGTATTAGG AGGCTGAGGCTACTCGGAGG CTGACACAGGAAATGGCTGA				Contig.			Contiguous	LMD	DpnII	40(2-739)	N
Y4 iii	1827	TGGTCTGTTCACCCAGCCATG CAACTTTTCAAAAGTCTC ACTCAGTGTCCACACAGAGAA	>100000	47400	47400	1			473751 bp at 5' side: tenascin-R precursor 64398 bp at 3' side: E3 ubiquitin-protein ligase RFW2 isoform a	LMD	DpnII	125(15-956)	Y
Y4 iv	1718	AGCACCACCTTGAGCCATAC TCAAAGCTGTATCTAGAGAA CTTAGTGGATGGAGCCAGTA	60000	>100000	60000	20	14		probable phospholipid-transporting ATPase IIA	LMD	DpnII	40(2-739)	N
Y4 ix	1826	CAACTTTTCTGTATGTATT GTAGTGTATTTGTAGATGT	>100000	41300	41300	13	0	0	transmembrane protein FAM155A	LMD	DpnII	40(2-739)	N
Y4 v	1773	GGGGAGGAGATGAGTTAAG GCTTTTGTATGCCCTCACCTTAA CCATCAGAAATTAATTTTG	34900	>100000	34900	X	10/0	4/0	20978 bp at 5' side: hypothetical protein LOC55026 isoform 1 29720 bp at 3' side: protein ATP1B4 isoform A	LMD	NcoI	40(2-739)	N
Y4 vi	1829	TCTGTTCCACCCAGCCATGCAA ATGCTGCTAGGCTGTAAA				Contig.			Contiguous	LMD	NcoI	40(2-739)	N
Y4 viii	1767	GGAGTTGGGGAGGAGATGAGG TTAAAGGCTCATCTGTGGTA AAATAATCTAGAGATTAG	29100	79800	29100	13			536224 bp at 5' side: sAR-related lipid transfer protein 13 isoform gamma 139522 bp at 3' side: replication factor C subunit 3 isoform 2	LMD	DpnII	40(2-739)	N
Y4 x	1813	TGTAGGCATAAATGGTGTGT CACCAGACCCGATCTGAGAG AGGACATTTTCAGGAAAGC	11200	>100000	11200	1	41	24	ubiA prenyltransferase domain-containing protein 1	Slide	DpnII	146(8-2998)	Y
Y4 xi	1819	CATAAATGGCTGTCTCACCG CACCATGCAAGATTTGCAATCAT				Contig.			Contiguous	LMD	DpnII	125(15-956)	Y

9.8.28. Cont.

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
Y4 xii	1801	GTATTAGAGGCTGTAGGCATAAA TTGGTGGATCTCTGGCTAACACA GTGAAACCCCACTCTCT	>100000	69300	69300	15	41	24	LOC100128982 chromosome 15 open reading frame 29	LMD	DpnII	242(43-1301)	Y
Y4 xiii	1805	TAGAGGCTGTAGGCATAAATTGG TCTGTTCAGAAAGAGTGTGAAG CAGTTGGTGATAGGAC	24500	>100000	24500	7			322880 bp at 5' side: STAG3-like protein 4 1956519 bp at 3' side: autism susceptibility gene 2 protein isoform 1	LMD	DpnII	40(2-739)	Y
Y4 xiv	1827	GGTCTGTTCCACAGCACCATGCCAA CTTTTCAACTCACTCCACTTTTC ACTCTGGCTTTTATA	4000	74000	4000	8			594657 bp at 5' side: netrin receptor UNC5D precursor 398917 bp at 3' side: potassium channel subfamily U member 1	LMD	NcoI	40(2-739)	Y
Y4 xix	1733	CATCTTCAAAGACTGTGTGTTTA CTGAGTGGACCTAGACCCCT				Contig.			Contiguous	LMD	DpnII	40(2-739)	Y
Y4 xv	1693	ACTCTTGGACTTTCAGCAATGTCA ACGACCGAGTTGAATCACCTGAGT GATTAGTGAGAGATA				Contig.			Contiguous	LMD	DpnII	2418 (1875-7239)	Y
Y4 xvi	1791	AAAGGCTTTTGTATTAGGAGGCTG TAGGCATAGGTGACCATAGATAAC TACTTTTAAATGCCATA	24900	76200	24900	4			108433 bp at 5' side: sorbin and SH3 domain-containing protein 2 isoform 8 311382 bp at 3' side: toll-like receptor 3 precursor	LMD	DpnII	125(15-956)	Y
Y4 xvii	1817	GGCATAAATGGTCTGTTCACCAAG CACCATGCTACGTTG				Contig.			Contiguous	LMD	DpnII	125(15-956)	Y
Y4 xx	1825	AATTGGTCTGTTCCACGACCCAT GCAACTTTTATTTGCAGGAGAACT TACCAAAAGCCCAAGCAAC	29800	>100000	29800	1	164	96	splicing factor 3A subunit 3	LMD	NcoI	125(15-956)	Y
Y4 xxi	485	AGTCTGCAATGTCGGGTGCTGG TGGTTCATGATCACCCGCCACCTG CTGCTGGCTGGCAGTA	46300	>100000	46300	11	0	0	tripartite motif-containing protein 66	Slide	NcoI	146(8-2998)	N
Y4 xxiii	1791	AAAGGCTTTTGTATTAGGAGGCTG TAGGCATAGGTGACCATAGATAAC ATGGAGCTGGAGGCC	57900	>100000	57900	3	41	33	Homo sapiens ATPase, Ca++ transporting, plasma membrane 2 (ATP2B2)	Slide	DpnII	146(8-2998)	N
Y4 xxiv	1801	GTATTAGAGGCTGTAGGCATAAA TTGGTCTGAAGGAGGTGTGCCA CACTTTTAAACAACCA	46000	67400	46000	5			633898 bp at 5' side: EGF-like repeat and discoidin I-like domain-containing pr... 1599697 bp at 3' side: cytochrome c oxidase subunit 7C, mitochondrial precursor	Slide	DpnII	146(8-2998)	N
Y4 xxv	1814	GTAGGCATAAATTTGGTCTGTTCAC CAGCACAAAATFTGCACCAATTT TCGATCTCATCAGCAA	36500	15800	15800	7	155	14	52544 bp at 5' side: C-C motif chemokine 24 precursor 12891 bp at 3' side: rhomboid domain-containing protein 2 isoform a	Slide	NcoI	2998 (976-9782)	Y
Y4 xxvi	1821	GCATAAATGGTCTGTTCACCCAG ACCATGCATGATTAGCCCATCTCTG CGTGGCAGTCTTTTCT	55900	>100000	55900	4	0	0	15576 bp at 5' side: transmembrane protease serine 11B 52424 bp at 3' side: YTH domain-containing protein 1 isoform 1	Slide	NcoI	413 (250-3302)	Y

9.8.28. Cont.

Clone ID ¹	HBV site ²	Sequence ³	SMAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/hear ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
Y4 xxvii	1825	TTGGTCTGTTCCACGACCCAT GCAACTTTTTCACAACTCAT TCTCTCTCTCTCTCTGAAT	66800	>100000	66800	5			294487 bp at 5' side: hypothetical protein LOC100128508 429637 bp at 3' side: F-box/LRR-repeat protein 7	Slide	NcoI	444 (52-3617)	Y
Y4 xxviii	1826	TGGTCTGTTCCACGACCCATG CAACTTTTCTTCGTAGAGCG GGTTTACCAATGTTGGTCA				Contig.			Contiguous	Slide	DpnII	146(8-2998)	N
Y4 xxx	1782	GATGAGGTTAAAGGTCCTTTGTA TTAGAGGCTCCCATAACTCTA AAATGCACCGTAAACTCAG	20400	>100000	20400	17	20	19	breast carcinoma-amplified sequence 3 isoform 1 breast carcinoma-amplified sequence 3 isoform 2	LMD	DpnII	40(2-739)	N
Y4 xxxi	1694	CTCTTGGACTTTCAGCAATGTC AACGACCGACCTTAATCCCTG CCCCCTACCCACTGGGTTCC	>100000	56000	56000	8	0	0	speritolin isoform 1 (SPATC1) speritolin isoform 2 (SPATC2)	LMD	DpnII	40(2-739)	N

9.8.29. Patient Y5

Y5 i	1796	TGGCTGTTCCACGACCCATG CAACTTTTTCGAAAGGGGATGT TGGTTGAAAACACTACCCATT	62900	32200	32200	4			510308 bp at 5' side: cystine/glutamate transporter 263338 bp at 3' side: nocturnin	Slide LMD	NcoI NcoI	413 (250-3302) 114(6-2145)	Y Y
Y5 ii	1787	GGTTAAATGATCTTTGACTAGG AGGCTGTAGGACCG				Contig.			Contiguous	LMD	NcoI	114(6-2145)	N
Y5 iii	1796	AGGAGTTGGGGAGGAGATCAG GTTAATGATCCCCACACACAG AAAGTCAGAAACACACAGC	80400	14700	14700	1	238	245	36075 bp at 5' side: ERBB receptor feedback inhibitor 1 272345 bp at 3' side: proton-associated sugar transporter A	LMD	NcoI	666 (125-3617)	Y
Y5 ix	1827	GGTCTGTTCCACGACCCATGC AACTTTTTCAGGTACTCGAGAA GGTGTACTGTTTTCAGAAAT	43400	75300	43400	1	20	14	VPS13D gene for vacuolar protein sorting 13D	Slide LMD	NcoI NcoI	815 (146-4615) 666 (125-3617)	Y Y
Y5 v	1704	GTACTTCAAAGACTGTTTGT AAGGACTGGGTCCAGCTCC ATCGAAGCACTTCCCAAGGG	93200	>100000	93200	14	10	4	ras and Rab interactor 3	Slide	NcoI	146(8-2998)	N
Y5 vi	1729	GAGCGTACTTCAAAGACTGTT TGTTTAAGGAAAATTCCTCAGC CATTGTCCTTTTAAATATTG	51700	>100000	51700	21	29/0	33/4	21436 bp at 5' side: human dsRNA adenosine deaminase DRADA2b, EC 3.5. 23707 bp at 3' side: human mRNA for KIAA0958 protein	Slide LMD LMD LMD	NcoI NcoI NcoI NcoI	1301 (295-5692) 114(6-2145) 114(6-2145) 338 (42-2703)	Y Y Y Y
Y5 vii	1800	GTTCCACGACCCAFGTAATTT TTTTTCACTGGTATTTGAGA GCCTATGAGTCACATAAGCA	51700	>100000	51700	10			205779 bp at 5' side: rab11 family-interacting protein 2 58746 bp at 3' side: hypothetical protein LOC63877	Slide	NcoI	157(8-2998)	N
Y5 viii	1816	AAGCATAAATGGCTGTTCAC CAGCACCATGGCTGCTTTGGTG T				Contig.			Contiguous	Slide	NcoI	157(8-2998)	N

9.8.30. Patient Y6

Clone ID ¹	HBV site ²	Sequence ³	S/MAR ⁴			Chromosome ⁵	mRNA in tumour tissue ⁶	mRNA in normal liver ⁶	Gene integrated in/near ⁷	Extract ⁸	Design ⁹	Calculated clone size ¹⁰	Repeat ¹¹
			Upstream	Downstream	Closest								
Y6 i	1727	TTGAGGCA TACTTCAAAGACTG TCGTGTTAATCA TGCCTCATAAA TATTCCTATCA TTATTATCA	60400	71400	60400	4	0	9	ellis-van Creveld syndrome protein	LMD	DpnII	114(6-2145)	Y
Y6 iii	1794	GATCTTTG TACTAGGAGGCTGT AGGCATAAAT ACATCATGAGAT CAGTAGATG TCAAAATTCATA	81800	97000	81800	7	10	24	juxtaposed with another zinc finger protein 1	Slide	NcoI	13054 (9340-28934)	Y
Y6 iv	1821	TAAATGGT GTGTTCCACCAGCA CCATGCCAACT AAATAAATTAAT AATCATGGCA ACTCTAGAAG	>100000	41800	41800	13	0/20	24/24	5658 bp at 5' side: DNA ligase 4 12261 bp at 3' side: abhydrolase domain-containing protein 13	Slide	NcoI	21047 (7239-39330)	Y
Y6 v	1735	TACTTCAA AAGACTGTGTGTTTA ATGAGTGGGA TTAAAGAGTCTC ATTATTACAAA AGGCCAAAAG	44200	>100000	44200	13			736582 bp at 5' side: SLIT and NTRK-like protein 5 precursor 2983978 bp at 3' side: glypican-5 precursor	LMD	DpnII	338 (42-2703)	Y
Y6 vi	1747	TCGTGT TTTAAATCAGTGGGAGG AGTTGGGGA ATGGCTTCCTTC TCCCA GTGATGTGACT	21300	59000	21300	17	0	0	254544 bp at 5' side: monocyte to macrophage differentiation protein precursor 44400 bp at 3' side: transmembrane protein 100	LMD	DpnII	1632 (460-6450)	Y

¹ Each unique virus-cell DNA junction detected by invPCR as described in Sections 2.14 and 2.21 was given a unique identifier.

² The HBV DNA nucleotide at which of the integration occurred was measured with respect to distance from the single *EcoRI* site of the HBV sequence. This was determined by aligning HBV DNA sequences of the virus-cell DNA junction using BLAST (NCBI).

³ For each virus-cell DNA junction, the HBV sequence is shown in blue and the cellular DNA sequence is shown in black.

⁴ Expected cellular sequences 100 kb either side of each virus-cell DNA junction was examined for S/MARs using freely-available, online program MAR-Wiz (<http://www.futuresoft.org/MarFinder/>). The upstream, downstream and minimum distance from a predicted S/MAR was recorded. No virus-cell DNA junction was >100 kb from a predicted S/MAR site.

⁵ The chromosome in which the HBV DNA is integrated and the flanking open reading frames were determined by BLAST (NCBI).

⁶ For those integrants within 50 kb of a open reading frame, the mRNA expression in transcripts per million of the cellular gene in liver tumour tissue and normal liver tissues were determined using online databases Unigene and EST (NCBI).

⁷ Virus-cell DNA junctions were detected in DNA extracts of ~5mg fragments of liver tissue (Fragment), entire 5 µm liver tissue sections (Slide) or foci of hepatocytes isolated by laser-microdissection (LMD)

⁸ The invPCR design used to detect each virus-cell DNA junction is shown. *DpnII* invPCR design was only compatible with the HBV DNA sequences found in the tissues of the following patients: CN, FMC3, GS4, HN, HS, NT, SAAO, XA, Y2, Y4 and Y6.

⁹ The number of unique virus-cell DNA junctions, therefore hepatocyte clone sizes, in the original extract was determined using the most probable number method, calculated by MPN calculator (Curiale, 2000). For virus-cell junctions found in slide or LMD DNA extracts, the three dimensional clone sizes were calculated as previously described in Section 2.22. using the calculated nuclear concentration from Section 3.2.1.1. Bracketed numbers represent the lower and upper 95% confidence intervals, respectively.

¹⁰ “Y” indicates that the particular virus-cell DNA junction was detected multiple times in the liver tissue DNA extracts of a particular patient. “N” indicates that the virus-cell DNA junction was detected only once and therefore caution should be taken in accepting the calculated clone size.

9.9 Computer Programming

In Chapter 6, a 3-dimensional simulation of a liver undergoing random cell turnover was written to determine the size of hepatocyte clones that could have arisen stochastically. This program is outlined in Figure 6.1.

9.9.1 Random number generation

A XORshift random number generator was used instead of the in-built Java random number generator based on a modified linear congruential formula. XORshift RNGs provides a full period cycle of $2^{64}-1$ and passes Marsaglia's “Diehard battery” statistical tests for randomness (Marsaglia 2003).

9.9.2 Cell selection

A random walk model with steps ranging from zero to as large as the entire length of LiverArray was used to nominate random cells to die. This circumvents the “falling on planes” bias that occurs when random number generators are used to select dimensional coordinates (Coffey 2010).

9.9.3 Source code for the simulation of stochastic clonal proliferation

```
import java.util.*;
import java.io.*;

public class ClonesLiverNoGrowthPTD{
    /*XOR shift Random number generator for integers*/
    protected static int next() {
        long seed = System.nanoTime();
        seed ^= (seed << 21);
        seed ^= (seed >>> 35);
        seed ^= (seed << 4);
        return Math.abs((int) seed);
    }
    /*XOR shift Random number generator for floating point numbers*/
    protected static float nextFloat() {
        float seed = (next() % 1000)/1000;
        return (Math.abs(seed));
    }

    /*Calculates the 3D distance from the nearest Portal Tract*/
    protected static double PTDist(double x, double y, double z, double LOBULE_RADIUS)
    {
        double CubeSide = 2 * Math.sqrt(Math.pow(LOBULE_RADIUS,2)/3);
        x = x % CubeSide;
        y = y % CubeSide;
        z = z % CubeSide;
        double xDist, yDist, zDist;

        if (x > (CubeSide/2)){
```

```

    xDist = CubeSide - x;
}else{
    xDist = x;
}
if (y > (CubeSide/2)){
    yDist = CubeSide - y;
}else{
    yDist = y;
}

if (z > (CubeSide/2)){
    zDist = CubeSide - z;
}else{
    zDist = z;
}
return Math.sqrt((xDist*xDist)+(yDist*yDist)+(zDist*zDist));
}

public static void main(String[] args) {
    long h, LiveXYZ, FreqIndex, CellDeaths, PrinterA, PrinterB, CloneCellCount,
MaxSize;
    float Complexity, Extinct;
    int CloneID, e, v, CheckerA, CheckerB, CheckerC, DieX, DieY, DieZ, LiveX, LiveY,
LiveZ, LiveIDX, LiveIDY, LiveIDZ, DieIDX, DieIDY, DieIDZ, Above10000;
    short FillerA, FillerB, FillerC;

    final short LIVER_ARRAY_SIZE = 200;
    /*length of each side of LiverArray*/
    final int LOBULE_RADIUS = 20;
    /*distance between PT and CV*/
    final float PT_CLUSTER = 0;
    /*Focal death around portal tract, 0 = No effect from PT, 1 = proportional to
distance, 2 = proportional to distance^2*/
    final int PT_RAND_RATIO = 0;
    /*% of death due to Portal tract distance as opposed to random cell death*/
    final long NUM_CELLS = LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE;
    /*No. of cells in LiverArray*/
    final int ITERATIONS = 20;
    /*number of iterations of main program*/
    final int CONNECT = 1;
    /*connectivity*/
    final int TO_PER_INT = 500;
    /*liver turnovers per iteration*/
    final int UPDATE_INT = 1;
    /*update interval measured in liver fold changes*/
    final int GROWTH_FACTOR = 10;
    /*factor that liver grows by*/
    final int NUM_UPDATES = GROWTH_FACTOR/UPDATE_INT;
    /*number of updates per iteration*/
    final long ORIG_SIZE = NUM_CELLS/GROWTH_FACTOR;
    /*original size of the liver*/

    final int DUPDATE_INT = 1;
    /*updates per liver turnover for death phase*/
    final int DNUM_UPDATES = (TO_PER_INT*DUPDATE_INT);
    /*number of updates per iteration for death phase*/
    final int SUAD = 10;
    /*Survival advantage (in %) of cells with a survival advantage*/
    final int SACELLNUM = 80;
    /*Number of cells with survival advantage*/

    short[][][][] LiverArray = new
short[LIVER_ARRAY_SIZE][LIVER_ARRAY_SIZE][LIVER_ARRAY_SIZE][5];
    /*analogous to a liver starting w/ unique clones, size of final liver if
growing*/
    int[][][] CloneSizeArray = new
int[LIVER_ARRAY_SIZE][LIVER_ARRAY_SIZE][LIVER_ARRAY_SIZE];
    /*tracks clone size of each unique ID*/
    float [][] ComplexityArray = new float[ITERATIONS][DNUM_UPDATES+1];
    /*tracks complexity of liver at each update*/
    long [][] MaxCloneArray = new long[ITERATIONS][NUM_UPDATES];
    /*tracks largest clone size at each update in growth phase*/
    int [][] FreqArray = new int[ITERATIONS][15*TO_PER_INT];
    /*Looks at clonal frequency of the final liver in growth phase*/

```



```

long[][] DeathMaxCloneArray = new long[ITERATIONS][DNUM_UPDATES];
/*tracks largest clone size at each update in death phase*/
int[][] DeathFreqArray = new int[ITERATIONS][30*TO_PER_INT];
/*Looks at clonal frequency of the final liver in death phase*/
int[][] Above10000Array = new int[ITERATIONS][DNUM_UPDATES];
/*tracks clones above 10000 cells at each update*/

long[] SatArray = new long[ITERATIONS];

String FilePrefix = LIVER_ARRAY_SIZE + "Size"+ TO_PER_INT +"TO"+ PT_CLUSTER +
"PTD" + PT_RAND_RATIO + "SUAD" + SUAD;

System.out.println("Output files are prefixed with " + FilePrefix);

for(int i = 0; i < ITERATIONS; i++) {
    System.out.println("Started iteration " + (i+1) + " of " + ITERATIONS + "...");
    long CurrentSize = ORIG_SIZE;
    h = 0; v = 0; FillerA = 0; FillerB = 0; FillerC = 0; CloneID = 0; Above10000 =
0;

    /*Fills liver array with ID (corresponds to place in CloneSize array)*/
    for(FillerA=0; FillerA<LIVER_ARRAY_SIZE;FillerA++){
        for(FillerB=0; FillerB<LIVER_ARRAY_SIZE;FillerB++){
            for(FillerC=0; FillerC<LIVER_ARRAY_SIZE;FillerC++){
                LiverArray[FillerA][FillerB][FillerC][0] = FillerA;
                LiverArray[FillerA][FillerB][FillerC][1] = FillerB;
                LiverArray[FillerA][FillerB][FillerC][2] = FillerC;
                LiverArray[FillerA][FillerB][FillerC][3] = 0;
                /*Gives survival advantage due to distance from portal tracts, in
necessary*/
                if (PT_CLUSTER>0 && PTDist(FillerA,FillerB,FillerC,LOBULE_RADIUS)>0 &&
PT_RAND_RATIO > 0) {
                    double Survival = PT_RAND_RATIO *
(Math.pow((PTDist(FillerA,FillerB,FillerC,LOBULE_RADIUS)/LOBULE_RADIUS),PT_CLUSTER)
);
                    LiverArray[FillerA][FillerB][FillerC][4] = (short) Survival;
                }else{
                    LiverArray[FillerA][FillerB][FillerC][4] = 0;
                }
            }
        }
    }
    /*Output LiverArrayDistances to file*/
    try {
        PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"LiverArrayDistance.txt"));
        for(int m=0;m<(LIVER_ARRAY_SIZE - 1);m++){
            for(int l=0;l<(LIVER_ARRAY_SIZE - 1);l++){
                for(int n=0;n<(LIVER_ARRAY_SIZE - 1);n++){
                    out.print(LiverArray[l][m][n][4] + "\t");
                }
                out.println();
            }
            out.println();
        }
        out.close();
    } catch (FileNotFoundException f) {
        f.printStackTrace();
    }

    /*Sets CloneSize array to read clone size of 1 for every clone*/
    FillerA = 0; FillerB = 0; FillerC = 0;
    for(FillerA=0; FillerA<LIVER_ARRAY_SIZE;FillerA++){
        for(FillerB=0; FillerB<LIVER_ARRAY_SIZE;FillerB++){
            for(FillerC=0; FillerC<LIVER_ARRAY_SIZE;FillerC++){
                CloneSizeArray[FillerA][FillerB][FillerC] = 1;
            }
        }
    }

    /*While the number of cells is less than the fully grown liver*/
    while(CurrentSize < NUM_CELLS){
        DieX = (int) (CurrentSize / (LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE));

```

```

        DieY = (int) ((CurrentSize % (LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE)) /
LIVER_ARRAY_SIZE);
        DieZ = (int) (CurrentSize % LIVER_ARRAY_SIZE);

        /*chooses cell to divide to expand into spot*/
        LiveXYZ = nextFloat() * (CurrentSize - 1);
        LiveX = (int) (LiveXYZ / (LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE));
        LiveY = (int) ((LiveXYZ % (LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE)) /
LIVER_ARRAY_SIZE);
        LiveZ = (int) (LiveXYZ % LIVER_ARRAY_SIZE);

        /*Replace the ID of hole with replacing cell*/
        LiverArray[DieX][DieY][DieZ][0] = LiverArray[LiveX][LiveY][LiveZ][0];
        LiverArray[DieX][DieY][DieZ][1] = LiverArray[LiveX][LiveY][LiveZ][1];
        LiverArray[DieX][DieY][DieZ][2] = LiverArray[LiveX][LiveY][LiveZ][2];

        /*Get ID of replacing cell*/
        LiveIDX = LiverArray[LiveX][LiveY][LiveZ][0];
        LiveIDY = LiverArray[LiveX][LiveY][LiveZ][1];
        LiveIDZ = LiverArray[LiveX][LiveY][LiveZ][2];

        /*Change number of clone in CloneSize array*/
        CloneSizeArray[LiveIDX][LiveIDY][LiveIDZ] =
CloneSizeArray[LiveIDX][LiveIDY][LiveIDZ] + 1;

        if(h == (ORIG_SIZE*UPDATE_INT) - 1){
        /*Find maximum clone size by scanning through CloneSizeArray*/
        MaxSize = 0; CheckerA= 0; CheckerB= 0; CheckerC = 0;

        for(CheckerA=0; CheckerA<LIVER_ARRAY_SIZE;CheckerA++){
        for(CheckerB=0; CheckerB<LIVER_ARRAY_SIZE;CheckerB++){
        for(CheckerC=0; CheckerC<LIVER_ARRAY_SIZE;CheckerC++){
        if (CloneSizeArray[CheckerA][CheckerB][CheckerC] > MaxSize){
        MaxSize = CloneSizeArray[CheckerA][CheckerB][CheckerC];
        }
        }
        }
        }

        /*Put maximum clone size in MaxCloneArray*/
        MaxCloneArray[i][v] = MaxSize;
        v++; h = 0;
        }

        CurrentSize++; h++;
    }

    //initialise row of FreqArray
    for(FillerA=0;FillerA < MaxCloneArray[i].length; FillerA++){
        FreqArray[i][FillerA] = 0;
    }

    //Get frequency array filled in with clone frequency of the final liver
    for(CheckerA=0; CheckerA<LIVER_ARRAY_SIZE;CheckerA++){
        for(CheckerB=0; CheckerB<LIVER_ARRAY_SIZE;CheckerB++){
            for(CheckerC=0; CheckerC<LIVER_ARRAY_SIZE;CheckerC++){
                CloneID = CloneSizeArray[CheckerA][CheckerB][CheckerC];
                FreqArray[i][CloneID] = FreqArray[i][CloneID] + 1;
            }
        }
    }
}

DeathFreqArray [i][1] = LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE*LIVER_ARRAY_SIZE;

//Random Death phase
h = 0; v = 0; DieX = 0; DieY = 0; DieZ = 0;
FillerA = 0; FillerB = 0; FillerC = 0; Extinct = 0;
//Gives a given number of random cells a given growth advantage
for(int SuAdCount = 0; SuAdCount < SACELLNUM; SuAdCount++){
    int x, y, z;
    do{
        //Pick random cell that doesn't already have survival advantage
        x = next() % LIVER_ARRAY_SIZE;
        y = next() % LIVER_ARRAY_SIZE;

```

```

    z = next() % LIVER_ARRAY_SIZE;
  }while (LiverArray [x][y][z][3] == SUAD);
  //give cell survival advantage
  LiverArray [x][y][z][3] = SUAD;
}

/*Until the number of cells to die has been reached*/
for(long j=0; j<(NUM_CELLS*TO_PER_INT); j++){
  /*Pick cell at random to die*/
  do{
    if(next() % 2 == 0){
      DieX = DieX - (next() % LIVER_ARRAY_SIZE);
    }
    else{
      DieX = DieX + (next() % LIVER_ARRAY_SIZE);
    }
  }
  if(next() % 2 == 0){
    DieY = DieY - (next() % LIVER_ARRAY_SIZE);
  }
  else{
    DieY = DieY + (next() % LIVER_ARRAY_SIZE);
  }
  if(next() % 2 == 0){
    DieZ = DieZ - (next() % LIVER_ARRAY_SIZE);
  }
  else{
    DieZ = DieZ + (next() % LIVER_ARRAY_SIZE);
  }

  /*keeps cell in LiverArray*/
  while (DieX >= LIVER_ARRAY_SIZE){
    DieX = DieX - LIVER_ARRAY_SIZE;
  }
  while (DieX < 0){
    DieX = DieX + LIVER_ARRAY_SIZE;
  }
  while (DieY >= LIVER_ARRAY_SIZE){
    DieY = DieY - LIVER_ARRAY_SIZE;
  }
  while (DieY < 0){
    DieY = DieY + LIVER_ARRAY_SIZE;
  }
  while (DieZ >= LIVER_ARRAY_SIZE){
    DieZ = DieZ - LIVER_ARRAY_SIZE;
  }
  while (DieZ < 0){
    DieZ = DieZ + LIVER_ARRAY_SIZE;
  }
  } while
(Math.Max(LiverArray[DieX][DieY][DieZ][3],LiverArray[DieX][DieY][DieZ][4]) >
(next()%100));

  /*chooses neighbouring cell to replace lost cell*/
  e = CONNECT + 1;
  do{
    if(next() % 2 == 0){
      LiveX = DieX + (next() % e);
    }else{
      LiveX = DieX - (next() % e);
    }
    if(next() % 2 == 0){
      LiveY = DieY + (next() % e);
    }else{
      LiveY = DieY - (next() % e);
    }
    if(next() % 2 == 0){
      LiveZ = DieZ + (next() % e);
    }else{
      LiveZ = DieZ - (next() % e);
    }
  }while (DieX == LiveX && DieY == LiveY && DieZ == LiveZ);

  /*keeps the neighbouring cell in the array*/
  while (LiveX >= LIVER_ARRAY_SIZE){

```

```

    LiveX = LiveX - LIVER_ARRAY_SIZE;
}
while (LiveX < 0){
    LiveX = LiveX + LIVER_ARRAY_SIZE;
}
while (LiveY >= LIVER_ARRAY_SIZE){
    LiveY = LiveY - LIVER_ARRAY_SIZE;
}
while (LiveY < 0){
    LiveY = LiveY + LIVER_ARRAY_SIZE;
}
while (LiveZ >= LIVER_ARRAY_SIZE){
    LiveZ = LiveZ - LIVER_ARRAY_SIZE;
}
while (LiveZ < 0){
    LiveZ = LiveZ + LIVER_ARRAY_SIZE;
}
/*Get ID of the dying cell*/
DieIDX = LiverArray[DieX][DieY][DieZ][0];
DieIDY = LiverArray[DieX][DieY][DieZ][1];
DieIDZ = LiverArray[DieX][DieY][DieZ][2];

/*Get ID of replacing cell*/
LiveIDX = LiverArray[LiveX][LiveY][LiveZ][0];
LiveIDY = LiverArray[LiveX][LiveY][LiveZ][1];
LiveIDZ = LiverArray[LiveX][LiveY][LiveZ][2];

/*Change number of clones in CloneSize array*/
CloneSizeArray [DieIDX][DieIDY][DieIDZ] =
CloneSizeArray[DieIDX][DieIDY][DieIDZ] - 1;
if (CloneSizeArray [DieIDX][DieIDY][DieIDZ] == 0) {
    Extinct ++;
} else if (CloneSizeArray [DieIDX][DieIDY][DieIDZ] == 9999){
    Above10000 --;
}
CloneSizeArray [LiveIDX][LiveIDY][LiveIDZ] =
CloneSizeArray[LiveIDX][LiveIDY][LiveIDZ] + 1;
if (CloneSizeArray [LiveIDX][LiveIDY][LiveIDZ] == 10000){
    Above10000 ++;
}
/*Replace the ID of dead cell with replacing cell*/
LiverArray[DieX][DieY][DieZ][0] = LiverArray[LiveX][LiveY][LiveZ][0];
LiverArray[DieX][DieY][DieZ][1] = LiverArray[LiveX][LiveY][LiveZ][1];
LiverArray[DieX][DieY][DieZ][2] = LiverArray[LiveX][LiveY][LiveZ][2];
LiverArray[DieX][DieY][DieZ][3] = LiverArray[LiveX][LiveY][LiveZ][3];
h++;

/*if it's at the update interval*/
if(j %(NUM_CELLS/DUPDATE_INT) == 0){

/*compute complexity and add to Complexity array*/
    Extinct = DeathFreqArray[i][0];
    Complexity = (NUM_CELLS - Extinct)/NUM_CELLS;
    ComplexityArray[i][v] = Complexity;

/*Find maximum clone size by scanning through CloneSizeArray*/
    MaxSize = 0; CheckerA= 0; CheckerB= 0; CheckerC = 0;
    for(CheckerA=0; CheckerA<LIVER_ARRAY_SIZE;CheckerA++){
        for(CheckerB=0; CheckerB<LIVER_ARRAY_SIZE;CheckerB++){
            for(CheckerC=0; CheckerC<LIVER_ARRAY_SIZE;CheckerC++){
                if (CloneSizeArray[CheckerA][CheckerB][CheckerC]>MaxSize){
                    MaxSize = CloneSizeArray[CheckerA][CheckerB][CheckerC];
                }
            }
        }
    }
/*Put maximum clone size in MaxCloneArray*/
    DeathMaxCloneArray[i][v] = MaxSize;
    Above10000Array[i][v] = Above10000;
    v++;
}
}
/*Output FreqArray to file*/
try {

```

```

        PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"FreqArray.txt"));
        for(int m=0;m<FreqArray[0].length;m++){
            for(int l=0;l<FreqArray.length;l++){
                out.print(FreqArray[l][m] + "\t");
            }
            out.println();
        }
        out.close();
    } catch (FileNotFoundException f) {
        f.printStackTrace();
    }

    /*Output DeathFreqArray to file*/
    try {
        PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"DeathFreqArray.txt"));
        for(int m=0;m<DeathFreqArray[0].length;m++){
            for(int l=0;l<DeathFreqArray.length;l++){
                out.print(DeathFreqArray[l][m] + "\t");
            }
            out.println();
        }
        out.close();
    } catch (FileNotFoundException f) {
        f.printStackTrace();
    }

    /*Output ComplexityArray to file*/
    try {
        PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"ComplexityArray.txt"));
        for(int m=0;m<ComplexityArray[0].length;m++){
            for(int l=0;l<ComplexityArray.length;l++){
                out.print(ComplexityArray[l][m] + "\t");
            }
            out.println();
        }
        out.close();
    } catch (FileNotFoundException f) {
        f.printStackTrace();
    }

    /*Output Above10000Array to file*/
    try {
        PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"Above10000Array.txt"));
        for(int m=0;m<Above10000Array[0].length;m++){
            for(int l=0;l<Above10000Array.length;l++){
                out.print(Above10000Array[l][m] + "\t");
            }
            out.println();
        }
        out.close();
    } catch (FileNotFoundException f) {
        f.printStackTrace();
    }

    /*Output MaxCloneArray to file*/
    try {
        PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"MaxCloneArray.txt"));
        for(int m=0;m<MaxCloneArray[0].length;m++){
            for(int l=0;l<MaxCloneArray.length;l++){
                out.print(MaxCloneArray[l][m] + "\t");
            }
            out.println();
        }
        out.close();
    } catch (FileNotFoundException f) {
        f.printStackTrace();
    }

    /*Output DeathMaxCloneArray to file*/
    try {

```

```

    PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"DeathMaxCloneArray.txt"));
    for(int m=0;m<DeathMaxCloneArray[0].length;m++){
        for(int l=0;l<DeathMaxCloneArray.length;l++){
            out.print(DeathMaxCloneArray[l][m] + "\t");
        }
        out.println();
    }
    out.close();
} catch (FileNotFoundException f) {
    f.printStackTrace();
}

/*Output SaturationArray to file*/
try {
    PrintStream out = new PrintStream(new FileOutputStream(FilePrefix +
"DeathSaturationArray.txt"));
    for(int l=0;l<SatArray.length;l++) {
        out.print(SatArray[l]);
        out.println();
    }
    out.close();
} catch (FileNotFoundException f) {
    f.printStackTrace();
}
}
}
}
}

```

9.9.4 Source code for the 2D visualisation of stochastic clonal proliferation

```

Sub kill()

Dim u, e, p, h, v, f, g, k, l, SanityCounter, InsaneCounter, xVis, yVis, RVal,
GVal, BVal, FreqIndex, LiveID, DieID, FillerA, FillerB, GrAdCount, CellDeaths,
CheckerA, CheckerB, Extinct, PrinterA, PrinterB, CloneCellCount As Integer
Dim GrAd, Complexity, PCREff, xZero, yZero, xMod, yMod As Single
Dim LiverArray(0 To 100, 0 To 100, 0 To 100, 0 To 5) As Single
'Analogous to a liver starting with unique clones
Dim CloneSizeArray(0 To 1, 0 To 10000) As Integer
'Tracks the Clone size of each unique ID
Dim ComplexityArray(0 To 2, 0 To 800) As Single
'Tracks Complexity at each update
Dim FreqArray(0 To 1, 0 To 10000) As Integer
'Tracks frequency of each clone
Dim MaxCloneArray(0 To 1, 0 To 800) As Long
'Tracks maximum clone size at each update
Dim InsaneArray(0 To 1, 0 To 1000) As Long
'If LiverArray ends up composed of unkillable cells, tracks endpoint
Dim Many As Integer
'Iteration While loop
Dim Lob As Integer
'radius of lobule
Dim PVDist As Single
'Distance from PV

u = 0 'Iteration counter
InsaneCounter = 0 'Allows filling of InsaneArray

'Stops simulation if arrays aren't big enough
If Sheets("sheet1").Range("ex130") / Sheets("sheet1").Range("ex129") < 12.5 Then
    MsgBox "Please increase Update Interval or decrease TO/Iterations"
    Exit Sub
End If

For Many = 1 To Sheets("sheet1").Range("ex122").Value '# of iterations
    Sheets("sheet1").Range("bx114").Value = "Iteration " & u + 1 & " of " &
Sheets("sheet1").Range("ex122").Value
    e = Sheets("sheet1").Range("ex123").Value + 1 'Connectivity
    p = 0 'Counter for update
    h = 0 'Cell death counter

```

```

v = 0 'makes inserting into array easier
f = 0
g = 0

xMod = 0
yMod = 0
Randomize

'Initiates liver array w/ unique IDs, no growth advantage and random colour
values
FillerA = 0
FillerB = 0
GrAd = 0
Randomize
For FillerA = 0 To 99
  For FillerB = 0 To 99
    LiverArray(FillerA, FillerB, 0) = (FillerA) + (FillerB * 100) 'Unique id
    Lob = 20
    xMod = FillerA Mod (2 * Sqr(Lob ^ 2 / 2))
    yMod = FillerB Mod (2 * Sqr(Lob ^ 2 / 2))

    If xMod > Lob ^ 2 / 2 Then
      xZero = 2 * Sqr(Lob ^ 2 / 2)
    Else
      xZero = 0
    End If

    If yMod > Lob ^ 2 / 2 Then
      yZero = 2 * Sqr(Lob ^ 2 / 2)
    Else
      yZero = 0
    End If

    PVDist = (Sqr((xMod - xZero) ^ 2 + (yMod - yZero) ^ 2)) / Lob
    LiverArray(FillerA, FillerB, 1) = 0.7 * (PVDist) 'Distance
    LiverArray(FillerA, FillerB, 2) = Int((255 - 0 + 1) * Rnd + 0) 'Red
    LiverArray(FillerA, FillerB, 3) = Int((255 - 0 + 1) * Rnd + 0) 'Green
    LiverArray(FillerA, FillerB, 4) = Int((255 - 0 + 1) * Rnd + 0) 'Blue
  Next FillerB
Next FillerA

'Colours visual array with liver array values, if visualisation active
If Sheets("sheet1").Range("ex131").Value = 1 Then
  xVis = 0
  yVis = 0
  For yVis = 0 To 99
    For xVis = 0 To 99
      RVal = LiverArray(yVis, xVis, 2)
      GVal = LiverArray(yVis, xVis, 3)
      BVal = LiverArray(yVis, xVis, 4)
      Sheets("sheet1").Cells(10 + yVis, 10 + xVis).Interior.Color = RGB(RVal, GVal,
BVal)
      Sheets("sheet1").Cells(10 + yVis, 10 + xVis).Value = LiverArray(yVis, xVis,
0)
    Next xVis
  Next yVis
End If

'Resets Clone size array to 1
FillerA = 0
For FillerA = 0 To 9999
  CloneSizeArray(1, FillerA) = 1 'Clone count
Next FillerA

'Gives a given number of random cells a given survival advantage
GrAdCount = 0
GrAd = Sheets("sheet1").Range("ex128").Value
Do Until GrAdCount = Sheets("sheet1").Range("ex127").Value
  Do
    FillerA = Int(Rnd * 100)
    FillerB = Int(Rnd * 100)
    Loop While LiverArray(FillerA, FillerB, 5) = GrAd
    LiverArray(FillerA, FillerB, 0) = (FillerA) + (FillerB * 100)
    LiverArray(FillerA, FillerB, 1) = GrAd
  Do

```

```

    GrAdCount = GrAdCount + 1
Loop

For CellDeaths = 0 To ((10000 * Sheets("sheet1").Range("ex129").Value) - 1)
'pick a random cell to die taking growth advantage into account
SanityCounter = 0
Do
    Do
        f = (f + Int(Rnd * 100)) Mod 100
        g = (g + Int(Rnd * 100)) Mod 100
        SanityCounter = SanityCounter + 1
        If SanityCounter > 100000 Then
            Exit Do
        End If
    Loop While Rnd < LiverArray(f, g, 1)
Loop While Rnd < (f, g, 5)

If SanityCounter > 100000 Then
    InsaneArray(0, InsaneCounter) = u
    InsaneArray(1, InsaneCounter) = CellDeaths
    InsaneCounter = InsaneCounter + 1
    Exit For
End If

'randomly pick a nearby cell to replace dying cell
Do
    If Rnd > 0.5 Then
        k = f + Int(Rnd * e)
    Else
        k = f - Int(Rnd * e)
    End If
    If Rnd > 0.5 Then
        l = g + Int(Rnd * e)
    Else
        l = g - Int(Rnd * e)
    End If
Loop While (l = g And k = f)

'keep it in the array
If k > 99 Then
    k = k - 100
ElseIf k < 0 Then
    k = k + 100
End If
If l > 99 Then
    l = l - 100
ElseIf l < 0 Then
    l = l + 100
End If

'Update CloneSizeArray
DieID = LiverArray(f, g, 0)
LiveID = LiverArray(k, l, 0)
CloneSizeArray(1, DieID) = CloneSizeArray(1, DieID) - 1
CloneSizeArray(1, LiveID) = CloneSizeArray(1, LiveID) + 1

'Replace with progeny of replicating cell in visual array, if visualisation
active
If Sheets("sheet1").Range("ex131").Value = 1 Then
    RVal = LiverArray(k, l, 2)
    GVal = LiverArray(k, l, 3)
    BVal = LiverArray(k, l, 4)
    Sheets("sheet1").Cells(10 + f, 10 + g).Interior.Color = RGB(RVal, GVal, BVal)

'Makes unique ID form of "xID#x" if cell has growth advantage
If LiverArray(k, l, 5) > 0 Then
    Sheets("sheet1").Cells(10 + f, 10 + g).Value = "x" & LiverArray(k, l, 0) &
"x"
Else
    Sheets("sheet1").Cells(10 + f, 10 + g).Value = LiverArray(k, l, 0)
End If
End If

'Replace with progeny of replicating cell in LiverArray

```



```

LiverArray(f, g, 0) = LiverArray(k, l, 0)
LiverArray(f, g, 2) = LiverArray(k, l, 2)
LiverArray(f, g, 3) = LiverArray(k, l, 3)
LiverArray(f, g, 4) = LiverArray(k, l, 4)
LiverArray(f, g, 5) = LiverArray(k, l, 5)

p = p + 1
h = h + 1

If p = Sheets("sheet1").Range("ex130").Value Then
Randomize
'Checks CloneSizeArray for extinct clones
Extinct = 0
CheckerA = 0
For CheckerA = 0 To 9999
If CloneSizeArray(1, CheckerA) = 0 Then
Extinct = Extinct + 1
End If
Next CheckerA

'Calculate complexity
Complexity = (10000 - Extinct) / 10000

'Puts min and max complexity in ComplexityArray
ComplexityArray(0, v) = h
If u = 0 Then 'Sets ComplexityArray values to start with
ComplexityArray(1, v) = Complexity
ComplexityArray(2, v) = Complexity
Else If Complexity < ComplexityArray(1, v) Then
ComplexityArray(1, v) = Complexity
Else If Complexity > ComplexityArray(2, v) Then
ComplexityArray(2, v) = Complexity
End If
End If

'Print Complexity Array
Sheets("sheet3").Cells(2 + v, 2).Value = ComplexityArray(0, v)
Sheets("sheet3").Cells(2 + v, 3).Value = ComplexityArray(1, v)
Sheets("sheet3").Cells(2 + v, 4).Value = ComplexityArray(2, v)

'Shows frequency of clones
'Resets FreqArray
FillerA = 0
For FillerA = 0 To 9999
FreqArray(0, FillerA) = FillerA
FreqArray(1, FillerA) = 0
Next FillerA

'Fills in FreqArray by going through CloneSizeArray and tallying clone sizes
CheckerA = 0
For CheckerA = 0 To 9999
FreqIndex = CloneSizeArray(1, CheckerA)
FreqArray(1, FreqIndex) = FreqArray(1, FreqIndex) + 1
Next CheckerA

>Delete previous FreqArray
Sheets("sheet2").Range("B2:C10002").ClearContents

'Print FreqArray
PrinterA = 0
PrinterB = 0
CloneCellCount = 0
Do
If FreqArray(1, PrinterA) > 0 Then
Sheets("sheet2").Cells(2 + PrinterB, 2).Value = FreqArray(0, PrinterA)
Sheets("sheet2").Cells(2 + PrinterB, 3).Value = FreqArray(1, PrinterA)
PrinterB = PrinterB + 1
End If
CloneCellCount = CloneCellCount + FreqArray(0, PrinterA) * FreqArray(1,
PrinterA)
PrinterA = PrinterA + 1
Loop While CloneCellCount < 9999

'Updates MaxCloneArray

```

```

MaxCloneArray(0, v) = h
If u = 0 Then 'Sets MaxCloneArray values to start with
    MaxCloneArray(1, v) = PrinterA - 1
    Else
    If (PrinterA - 1) > MaxCloneArray(1, v) Then
        MaxCloneArray(1, v) = PrinterA - 1
    End If
End If

'Prints MaxCloneArray
Sheets("sheet2").Cells(2 + v, 19).Value = MaxCloneArray(0, v)
Sheets("sheet2").Cells(2 + v, 20).Value = MaxCloneArray(1, v)

Sheets("sheet1").Range("ex125").Value = h
p = 0
v = v + 1
End If
Next CellDeaths
u = u + 1
Sheets("sheet1").Range("ex125").Value = h
Next Many

'Prints out InsaneArray if program has been crazy
PrinterA = 0
Do While InsaneArray(1, PrinterA) > 0
    Sheets("sheet3").Cells(2 + PrinterA, 5).Value = InsaneArray(0, PrinterA)
    Sheets("sheet3").Cells(2 + PrinterA, 6).Value = InsaneArray(1, PrinterA)
    PrinterA = PrinterA + 1
Loop

End Sub

```

Appendix 9.10. Ion maps of mass peaks corresponding to proteins identified by IMS and HPLC-nESI-LTQ Orbitrap MS in multiple liver tissue sections of HBV-negative and -positive patients. Empty boxes represent mass peaks which were not automatically selected by software analysis of IMS sum spectra as described in Section 2.29. Patient data is listed in Table 4.1.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6
ACOX_HUMAN	Peroxisomal acyl-coenzyme A oxidase OS=Homo sapiens GN=ACOX1 PE=1 SV=3	1716.9	SKEVAWNLTSVDLVR								
		1556.8	LTNILDGGAQNTALR								
ACSL1_HUMAN	Long-chain-fatty-acid--CoA ligase 1 OS=Homo sapiens GN=ACSL1 PE=1 SV=1	1351.7	VLQPTVFPVPR								
		1709.9	IIVVMDAYGSELVER								
ACTN4_HUMAN	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	1421.7	GYEEWLLNEIR								
		1429.7	TINEVENQLTR								
		1537.8	FAIQDISVEETSAK								

Appendix 9.10. Cont.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6	
ALBU_HUMAN	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	1149.6	LVNEVTEFAK									
		1511.8	VPQVSTPTLVEVSR									
		1467.7	RHPDYSWLLLR									
ALDH2_HUMAN	Aldehyde dehydrogenase, mitochondrial OS=Homo sapiens GN=ALDH2 PE=1 SV=2	1639.9	DVFLGMFLYEYAR									
		1527.8	ANNSTYGLAAAVFTK									
ALDOB_HUMAN	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2	1385.7	LGPALATGNVVMK									
		1243.7	ALQASALAAWGGK									
		1490.8	GILAADESVTMGNR									

Appendix 9.10. Cont.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6
ALDOB_HUMAN	Fructose-bisphosphate aldolase B OS=Homo sapiens GN=ALDOB PE=1 SV=2	1458.7	ELSEIAQSIVANGK								
		1916	KYTPQVAMATVTALHR								
ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1	1171.6	VVDALGNAIDGK								
		1316.7	TSAIDTIINQK								
		2104.2	GMSLNLEPDNVGVVFGNDK								
ATPB_HUMAN	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3	1457.7	TVLIMELINNVAK								
		1650.8	LVLEVAQHLGESTYR								

Appendix 9.10. Cont.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6	
CPSM_HUMAN	Carbamoyl-phosphate synthase [ammonial, mitochondrial] OS=Homo sapiens GN=CPST PE=1 SV=2	1306.7	AADTGYPMIR									
		1607.8	VLGTSVESIMATEDR									
		1419.7	AVNTLNEALEFAK									
DOPD_HUMAN	D-dopachrome decarboxylase OS=Homo sapiens GN=DDT PE=1 SV=3	1499.8	PFLELDTNLPANR									
		2194.2	PFLELDTNLPANRVPAGLEK									
ECHA_HUMAN	Trifunctional enzyme subunit alpha, mitochondrial OS=Homo sapiens GN=HADHA PE=1 SV=2	1318.7	MQLLEITTEK									
		1521.8	DLNSDMDSILASLK									
		1626.8	TIEYLEEVAITFAK									

Appendix 9.10. Cont.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6
ENOA_HUMAN	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2	1406.7	GNPTVEVDLFTSK								
		1804.9	AAVPSGASTGIYEALER								
		2033.1	FTASAGIQVGGDDLTVTNPK								
HBA_HUMAN	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=3	1071.5	MFLSFPTTK								
		1087.5	MFLSFPTTK								
		1529.7	VGAHAGEYGAEALER								
HBB_HUMAN	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2	1274.7	LLVVPWTQR								
		1314.7	VNVDEYVGEALGR								

Appendix 9.10. Cont.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6
HS90B_HUMAN	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4	1349.7	TLTLVDTGIGMTK								
		1544.7	ELISNASDALKIR								
PDIA3_HUMAN	Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4	1619.9	DLIIYYDVDYEK								
		1370.7	ELSDFISYLQR								
PYC_HUMAN	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2	1547.8	AEAEQAQAEILSFPR								
		1654.9	DFTATFGPLDSLINTR								
RPN_HUMAN	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase OS=Homo sapiens GN=RPN1 PE=1 SV=1	1301.7	ALTSEIALLQSR								
		1372.7	SIVEEIEDLVAR								

Appendix 9.10. Cont.

Protein accession	Protein description	m/z	Sequence	NHL1	NHL2	NHL3	Y2	Y3	Y4	Y5	Y6
SBP1_HUMAN	Selenium-binding protein 1 OS=Homo sapiens GN=SELENBP1 PE=1 SV=2	1313.7	AVFWIEFVMR								
		1906.1	NTGTEAPDYLATVDVDPK								
SPYA_HUMAN	Serine-pyruvate aminotransferase OS=Homo sapiens GN=AGXT PE=1 SV=1	1494.7	IMAAGGLQMGMSK								
		2505.4	ALLKPLSIPNQLLLGFGPSNLP PR								
TBB2C_HUMAN	Tubulin beta-2C chain OS=Homo sapiens GN=TUBB2C PE=1 SV=1	1143.6	LAVNMPFPR								
		1696.9	NSSFYVEWIPNNVK								
THTR_HUMAN	Thiosulfate sulfurtransferase OS=Homo sapiens GN=TST PE=1 SV=4	1452.7	TYEQVLENLESK								
		2010	FLGTEPEPDAVGLDSGHIR								

Chapter 10 - References

- Aihara, T., S. Noguchi, et al. (1994). "Clonal analysis of regenerative nodules in hepatitis C virus-induced liver cirrhosis." *Gastroenterology* **107**(6): 1805-1811.
- Alison, M. R., S. Islam, et al. (2009). "Stem cells in liver regeneration, fibrosis and cancer: the good, the bad and the ugly." *J Pathol* **217**(2): 282-298.
- Allice, T., F. Cerutti, et al. (2007). "COBAS AmpliPrep-COBAS TaqMan hepatitis B virus (HBV) test: a novel automated real-time PCR assay for quantification of HBV DNA in plasma." *J Clin Microbiol* **45**(3): 828-834.
- Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." *Journal of molecular biology* **215**(3): 403-410.
- Andreani, T., L. Serfaty, et al. (2007). "Chronic hepatitis B virus carriers in the immunotolerant phase of infection: histologic findings and outcome." *Clin Gastroenterol Hepatol* **5**(5): 636-641.
- Anthony, P. P., K. G. Ishak, et al. (1977). "The morphology of cirrhosis: definition, nomenclature, and classification." *Bull World Health Organ* **55**(4): 521-540.
- Anthony, P. P., K. G. Ishak, et al. (1978). "The morphology of cirrhosis. Recommendations on definition, nomenclature, and classification by a working group sponsored by the World Health Organization." *J Clin Pathol* **31**(5): 395-414.
- Aragona, E., R. D. Burk, et al. (1996). "Cell type-specific mechanisms regulate hepatitis B virus transgene expression in liver and other organs." *J Pathol* **180**(4): 441-449.
- Arthur, J. L., G. D. Higgins, et al. (2009). "A novel bocavirus associated with acute gastroenteritis in Australian children." *PLoS Pathog* **5**(4): e1000391.
- Arzumanyan, A., T. Friedman, et al. (2011). "Does the hepatitis B antigen HBx promote the appearance of liver cancer stem cells?" *Cancer Res* **71**(10): 3701-3708.
- Asabe, S., S. F. Wieland, et al. (2009). "The size of the viral inoculum contributes to the outcome of hepatitis B virus infection." *J Virol* **83**(19): 9652-9662.
- Asada, K., Y. Kotake, et al. (2006). "LINE-1 hypomethylation in a choline-deficiency-induced liver cancer in rats: dependence on feeding period." *J Biomed Biotechnol* **2006**(1): 17142.
- Attolini, C. S. and F. Michor (2009). "Evolutionary theory of cancer." *Ann N Y Acad Sci* **1168**: 23-51.
- Axelrod, R., D. E. Axelrod, et al. (2006). "Evolution of cooperation among tumor cells." *Proc Natl Acad Sci U S A* **103**(36): 13474-13479.
- Azechi, H., N. Nishida, et al. (2001). "Disruption of the p16/cyclin D1/retinoblastoma protein pathway in the majority of human hepatocellular carcinomas." *Oncology* **60**(4): 346-354.
- Balkwill, F., K. A. Charles, et al. (2005). "Smoldering and polarized inflammation in the initiation and promotion of malignant disease." *Cancer Cell* **7**(3): 211-217.
- Bannasch, P. (1986). "Preneoplastic lesions as end points in carcinogenicity testing. I. Hepatic preneoplasia." *Carcinogenesis* **7**(5): 689-695.
- Barash, H., R. G. E, et al. (2010). "Accelerated carcinogenesis following liver regeneration is associated with chronic inflammation-induced double-strand DNA breaks." *Proc Natl Acad Sci U S A* **107**(5): 2207-2212.
- Barboza, L., S. Salmen, et al. (2007). "Antigen-induced regulatory T cells in HBV chronically infected patients." *Virology* **368**(1): 41-49.
- Bartenschlager, R., M. Junker-Niepmann, et al. (1990). "The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation." *J Virol* **64**(11): 5324-5332.
- Bastie-Sigeac, F. and G. Lucotte (1983). "Optimal use of restriction enzymes in the analysis of human DNA polymorphism." *Hum Genet* **63**(2): 162-165.
- Beasley, R. P. (1988). "Hepatitis B virus. The major etiology of hepatocellular carcinoma." *Cancer* **61**(10): 1942-1956.
- Bedossa, P. and T. Poynard (1996). "An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group." *Hepatology* **24**(2): 289-293.

- Benali-Furet, N. L., M. Chami, et al. (2005). "Hepatitis C virus core triggers apoptosis in liver cells by inducing ER stress and ER calcium depletion." *Oncogene* **24**(31): 4921-4933.
- Bernuau, J., A. Goudeau, et al. (1986). "Multivariate analysis of prognostic factors in fulminant hepatitis B." *Hepatology* **6**(4): 648-651.
- Berquist, K. R., J. M. Peterson, et al. (1975). "Hepatitis B antigens in serum and liver of chimpanzees acutely infected with hepatitis B virus." *Infect Immun* **12**(3): 602-605.
- Bertoletti, A. and A. J. Gehring (2006). "The immune response during hepatitis B virus infection." *J Gen Virol* **87**(Pt 6): 1439-1449.
- Bichko, V., P. Pushko, et al. (1985). "Subtype ayw variant of hepatitis B virus. DNA primary structure analysis." *FEBS Lett* **185**(1): 208-212.
- Bill, C. A. and J. Summers (2004). "Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration." *Proc Natl Acad Sci U S A* **101**(30): 11135-11140.
- Bioulac-Sage, P. and C. Balabaud (2010). "Human cirrhosis: Monoclonal regenerative nodules derived from hepatic progenitor cells abutting ductular reaction." *Gastroenterol Clin Biol*.
- Bock, C. T., S. Schwinn, et al. (2001). "Structural organization of the hepatitis B virus minichromosome." *J Mol Biol* **307**(1): 183-196.
- Bode, J., M. Stengert-Iber, et al. (1996). "Scaffold/matrix-attached regions: topological switches with multiple regulatory functions." *Crit Rev Eukaryot Gene Expr* **6**(2-3): 115-138.
- Boguski, M. S. and G. D. Schuler (1995). "ESTablishing a human transcript map." *Nature genetics* **10**(4): 369-371.
- Bortolotti, F., M. Guido, et al. (2006). "Chronic hepatitis B in children after e antigen seroclearance: final report of a 29-year longitudinal study." *Hepatology* **43**(3): 556-562.
- Boulikas, T. (1993). "Nature of DNA sequences at the attachment regions of genes to the nuclear matrix." *J Cell Biochem* **52**(1): 14-22.
- Bourdi, M., T. P. Reilly, et al. (2002). "Macrophage migration inhibitory factor in drug-induced liver injury: a role in susceptibility and stress responsiveness." *Biochem Biophys Res Commun* **294**(2): 225-230.
- Bowden, S. (2002). Laboratory Diagnosis of Hepatitis B Infection. *Hepatitis B virus*. C. L. Lai and S. Locarnini. London, International Medical Press. **1**: 145-159.
- Boxall, E. H., J. Sira, et al. (2004). "Natural history of hepatitis B in perinatally infected carriers." *Arch Dis Child Fetal Neonatal Ed* **89**(5): F456-460.
- Boyd, C. D., K. Weliky, et al. (1986). "The single copy gene coding for human alpha 1 (IV) procollagen is located at the terminal end of the long arm of chromosome 13." *Hum Genet* **74**(2): 121-125.
- Bralet, M. P., S. Branchereau, et al. (1994). "Cell lineage study in the liver using retroviral mediated gene transfer. Evidence against the streaming of hepatocytes in normal liver." *Am J Pathol* **144**(5): 896-905.
- Bravo, R. and H. Macdonald-Bravo (1987). "Existence of two populations of cyclin/proliferating cell nuclear antigen during the cell cycle: association with DNA replication sites." *J Cell Biol* **105**(4): 1549-1554.
- Brechot, C. (2004). "Pathogenesis of hepatitis B virus-related hepatocellular carcinoma: old and new paradigms." *Gastroenterology* **127**(5 Suppl 1): S56-61.
- Brechot, C., D. Gozuacik, et al. (2000). "Molecular bases for the development of hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC)." *Semin Cancer Biol* **10**(3): 211-231.
- Brechot, C., D. Kremsdorf, et al. (2010). "Hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC): Molecular mechanisms and novel paradigms." *Pathol Biol (Paris)*.
- Bruni, R., E. D'Ugo, et al. (2003). "Scaffold attachment region located in a locus targeted by hepadnavirus integration in hepatocellular carcinomas." *Cancer Detect Prev* **27**(3): 175-181.

- Bruni, R., E. D'Ugo, et al. (1999). "Activation of the N-myc2 oncogene by woodchuck hepatitis virus integration in the linked downstream b3n locus in woodchuck hepatocellular carcinoma." Virology **257**(2): 483-490.
- Bucher, N. L. and M. N. Swaffield (1964). "The Rate of Incorporation of Labeled Thymidine into the Deoxyribonucleic Acid of Regenerating Rat Liver in Relation to the Amount of Liver Excised." Cancer Res **24**: 1611-1625.
- Budhu, A. and X. W. Wang (2006). "The role of cytokines in hepatocellular carcinoma." J Leukoc Biol **80**(6): 1197-1213.
- Bui-Nguyen, T. M., S. B. Pakala, et al. (2010). "NF-kappaB signaling mediates the induction of MTA1 by hepatitis B virus transactivator protein HBx." Oncogene **29**(8): 1179-1189.
- Bulmer, M. (1987). "A statistical analysis of nucleotide sequences of introns and exons in human genes." Mol Biol Evol **4**(4): 395-405.
- Burrell, C. J., E. J. Gowans, et al. (1984). "Correlation between liver histology and markers of hepatitis B virus replication in infected patients: a study by in situ hybridization." Hepatology **4**(1): 20-24.
- Buscher, M., W. Reiser, et al. (1985). "Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription." Cell **40**(3): 717-724.
- Calvisi, D. F., S. Ladu, et al. (2007). "Mechanistic and prognostic significance of aberrant methylation in the molecular pathogenesis of human hepatocellular carcinoma." J Clin Invest **117**(9): 2713-2722.
- Canbay, A., S. Friedman, et al. (2004). "Apoptosis: the nexus of liver injury and fibrosis." Hepatology **39**(2): 273-278.
- Cao, F. and J. E. Tavis (2004). "Detection and characterization of cytoplasmic hepatitis B virus reverse transcriptase." The Journal of general virology **85**(Pt 11): 3353-3360.
- Castello, G., S. Scala, et al. (2010). "HCV-related hepatocellular carcinoma: From chronic inflammation to cancer." Clin Immunol **134**(3): 237-250.
- Cha, C. and R. P. Dematteo (2005). "Molecular mechanisms in hepatocellular carcinoma development." Best Pract Res Clin Gastroenterol **19**(1): 25-37.
- Chaurand, P., D. S. Cornett, et al. (2006). "Molecular imaging of thin mammalian tissue sections by mass spectrometry." Curr Opin Biotechnol **17**(4): 431-436.
- Chaurand, P., J. C. Latham, et al. (2008). "Imaging mass spectrometry of intact proteins from alcohol-preserved tissue specimens: bypassing formalin fixation." J Proteome Res **7**(8): 3543-3555.
- Chemin, I. and F. Zoulim (2009). "Hepatitis B virus induced hepatocellular carcinoma." Cancer Lett **286**(1): 52-59.
- Chen, C. J., L. Y. Wang, et al. (2000). "Epidemiology of hepatitis B virus infection in the Asia-Pacific region." J Gastroenterol Hepatol **15** Suppl: E3-6.
- Chen, C. J., H. I. Yang, et al. (2006). "Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level." JAMA **295**(1): 65-73.
- Chen, J. Y., T. J. Harrison, et al. (1986). "Detection of hepatitis B virus DNA in hepatocellular carcinoma." Br J Exp Pathol **67**(2): 279-288.
- Chen, J. Y., T. J. Harrison, et al. (1988). "Detection of hepatitis B virus DNA in hepatocellular carcinoma: analysis by hybridization with subgenomic DNA fragments." Hepatology **8**(3): 518-523.
- Chen, L., D. W. Ho, et al. (2010). "Enhanced Detection of Early Hepatocellular Carcinoma by Serum SELDI-TOF Proteomic Signature Combined with Alpha-Fetoprotein Marker." Ann Surg Oncol.
- Chen, L., L. Hu, et al. "Dysregulation of beta-catenin by hepatitis B virus X protein in HBV-infected human hepatocellular carcinomas." Front Med China **4**(4): 399-411.
- Chen, R., P. S. Rabinovitch, et al. (2005). "The initiation of colon cancer in a chronic inflammatory setting." Carcinogenesis **26**(9): 1513-1519.

- Chen, X. P., S. Q. He, et al. (2003). "Expression of TNF-related apoptosis-inducing Ligand receptors and antitumor tumor effects of TNF-related apoptosis-inducing Ligand in human hepatocellular carcinoma." World J Gastroenterol **9**(11): 2433-2440.
- Chen, Y. W., Y. M. Jeng, et al. (2002). "P53 gene and Wnt signaling in benign neoplasms: beta-catenin mutations in hepatic adenoma but not in focal nodular hyperplasia." Hepatology **36**(4 Pt 1): 927-935.
- Chiba, T., K. Kita, et al. (2006). "Side population purified from hepatocellular carcinoma cells harbors cancer stem cell-like properties." Hepatology **44**(1): 240-251.
- Chisari, F. V., P. Filippi, et al. (1987). "Structural and pathological effects of synthesis of hepatitis B virus large envelope polypeptide in transgenic mice." Proc Natl Acad Sci U S A **84**(19): 6909-6913.
- Chisari, F. V., P. Filippi, et al. (1987). "Structural and pathological effects of synthesis of hepatitis B virus large envelope polypeptide in transgenic mice." Proceedings of the National Academy of Sciences of the United States of America **84**(19): 6909-6913.
- Chiu, C. C., G. T. Huang, et al. (2007). "Characterization of cytokeratin 19-positive hepatocyte foci in the regenerating rat liver after 2-AAF/CCl4 injury." Histochem Cell Biol **128**(3): 217-226.
- Chu, C. M., S. J. Hung, et al. (2004). "Natural history of hepatitis B e antigen to antibody seroconversion in patients with normal serum aminotransferase levels." Am J Med **116**(12): 829-834.
- Chu, C. M., P. Karayiannis, et al. (1985). "Natural history of chronic hepatitis B virus infection in Taiwan: studies of hepatitis B virus DNA in serum." Hepatology **5**(3): 431-434.
- Chu, M. L., W. de Wet, et al. (1985). "Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter structure, AluI repeats, and polymorphic transcripts." J Biol Chem **260**(4): 2315-2320.
- Churukian, C. J. (2008). Pigments and Minerals. Theory and Practice of Histological Techniques. J. D. Bancroft and M. Gamble. Philadelphia, PA, Churchill Livingstone Elsevier. **1**: 233-260.
- Ciupe, S. M., R. M. Ribeiro, et al. (2007). "The role of cells refractory to productive infection in acute hepatitis B viral dynamics." Proc Natl Acad Sci U S A **104**(12): 5050-5055.
- Coffey, N. (2010). "Randomness of bits with LCGs." Retrieved 10 January 2010, 2010, from http://www.javamex.com/tutorials/random_numbers/lcg_bit_positions.shtml.
- Cohen, N., T. Dagan, et al. (2005). "GC composition of the human genome: in search of isochores." Mol Biol Evol **22**(5): 1260-1272.
- Colnot, S., T. Decaens, et al. (2004). "Liver-targeted disruption of Apc in mice activates beta-catenin signaling and leads to hepatocellular carcinomas." Proc Natl Acad Sci U S A **101**(49): 17216-17221.
- Crawford, J. M. (2007). Basic mechanisms in hepatopathology. MacSween's Pathology of the Liver. A. D. Burt, B. C. Portmann and L. D. Ferrell. Philadelphia, PA, USA, Churchill Livingstone Elsevier. **1**: 75-117.
- Csepregi, A., C. Rocken, et al. (2008). "APC promoter methylation and protein expression in hepatocellular carcinoma." J Cancer Res Clin Oncol **134**(5): 579-589.
- Curiale, M. (2004). MPN Calculator: Freely available calculator of most probable number statistics - <http://www.i2workout.com/mcuriale/mpn/index.html>.
- D'Ugo, E., R. Bruni, et al. (1998). "Identification of scaffold/matrix attachment region in recurrent site of woodchuck hepatitis virus integration." DNA Cell Biol **17**(6): 519-527.
- Dadarkar, S. S., L. C. Fonseca, et al. (2010). "Phenotypic and genotypic assessment of concomitant drug-induced toxic effects in liver, kidney and blood." J Appl Toxicol.
- Dandri, M., M. R. Burda, et al. (2002). "Increase in de novo HBV DNA integrations in response to oxidative DNA damage or inhibition of poly(ADP-ribosylation)." Hepatology **35**(1): 217-223.

- Dane, D. S., C. H. Cameron, et al. (1970). "Virus-like particles in serum of patients with Australia-antigen-associated hepatitis." *Lancet* **1**(7649): 695-698.
- Dara, L., C. Ji, et al. (2010). "The contribution of endoplasmic reticulum stress to liver diseases." *Hepatology* **53**(5): 1752-1763.
- Das, S., J. C. Noe, et al. (2005). "An improved arbitrary primed PCR method for rapid characterization of transposon insertion sites." *J Microbiol Methods* **63**(1): 89-94.
- de Franchis, R., G. Meucci, et al. (1993). "The natural history of asymptomatic hepatitis B surface antigen carriers." *Ann Intern Med* **118**(3): 191-194.
- de La Coste, A., B. Romagnolo, et al. (1998). "Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas." *Proc Natl Acad Sci U S A* **95**(15): 8847-8851.
- De Meyer, S., Z. J. Gong, et al. (1997). "Organ and species specificity of hepatitis B virus (HBV) infection: a review of literature with a special reference to preferential attachment of HBV to human hepatocytes." *J Viral Hepat* **4**(3): 145-153.
- Delpuech, O., J. B. Trabut, et al. (2002). "Identification, using cDNA macroarray analysis, of distinct gene expression profiles associated with pathological and virological features of hepatocellular carcinoma." *Oncogene* **21**(18): 2926-2937.
- Dragan, Y., J. Teeguarden, et al. (1995). "The quantitation of altered hepatic foci during multistage hepatocarcinogenesis in the rat: transforming growth factor alpha expression as a marker for the stage of progression." *Cancer Lett* **93**(1): 73-83.
- Dragan, Y. P., H. A. Campbell, et al. (1994). "Focal and non-focal hepatic expression of placental glutathione S-transferase in carcinogen-treated rats." *Carcinogenesis* **15**(11): 2587-2591.
- Dragan, Y. P., J. Hully, et al. (1994). "Incorporation of bromodeoxyuridine in glutathione S-transferase-positive hepatocytes during rat multistage hepatocarcinogenesis." *Carcinogenesis* **15**(9): 1939-1947.
- Dragan, Y. P., C. Laufer, et al. (1993). "Quantitative comparison of initiation and mutation phenotypes in hepatocytes of the analbuminemic rat." *Jpn J Cancer Res* **84**(2): 175-183.
- Dugaiczuk, A., H. W. Boyer, et al. (1975). "Ligation of EcoRI endonuclease-generated DNA fragments into linear and circular structures." *J Mol Biol* **96**(1): 171-184.
- Edamoto, Y., A. Hara, et al. (2003). "Alterations of RB1, p53 and Wnt pathways in hepatocellular carcinomas associated with hepatitis C, hepatitis B and alcoholic liver cirrhosis." *Int J Cancer* **106**(3): 334-341.
- Eddleston, A. L. and M. Mondelli (1986). "Immunopathological mechanisms of liver cell injury in chronic hepatitis B virus infection." *Journal of hepatology* **3 Suppl 2**: S17-23.
- Eilers, F., S. Modaressi, et al. (1995). "Predominant periportal expression of the fructose 1,6-bisphosphatase gene in rat liver: dynamics during the daily feeding rhythm and starvation-refeeding cycle." *Histochemistry and cell biology* **103**(4): 293-300.
- Elias, H. and J. C. Sherrick (1969). *Liver Lobules. Morphology of the Liver*. New York, Academic Press, Inc. **1**: 125-137.
- Elias, H. and A. Sokol (1953). "Dependence of the lobular architecture of the liver on the porto-hepatic blood pressure gradient." *Anat Rec* **115**(1): 71-85.
- Enzmann, H. and P. Bannasch (1987). "Potential significance of phenotypic heterogeneity of focal lesions at different stages in hepatocarcinogenesis." *Carcinogenesis* **8**(11): 1607-1612.
- Esumi, M., Y. Tanaka, et al. (1989). "Clonal state of human hepatocellular carcinoma and non-tumorous hepatocytes." *Cancer Chemother Pharmacol* **23 Suppl**: S1-3.
- Evans, A., A. Riva, et al. (2008). "Programmed death 1 expression during antiviral treatment of chronic hepatitis B: Impact of hepatitis B e-antigen seroconversion." *Hepatology* **48**(3): 759-769.

- Evarts, R. P., P. Nagy, et al. (1987). "A precursor-product relationship exists between oval cells and hepatocytes in rat liver." *Carcinogenesis* **8**(11): 1737-1740.
- Factor, V. M., S. A. Radaeva, et al. (1994). "Origin and fate of oval cells in dipin-induced hepatocarcinogenesis in the mouse." *Am J Pathol* **145**(2): 409-422.
- Fan, Y. F., C. C. Lu, et al. (2001). "Prevalence and significance of hepatitis B virus (HBV) pre-S mutants in serum and liver at different replicative stages of chronic HBV infection." *Hepatology* **33**(1): 277-286.
- Farinati, F., R. Cardin, et al. (1996). "Hepatocyte proliferative activity in chronic liver damage as assessed by the monoclonal antibody MIB1 Ki67 in archival material: the role of etiology, disease activity, iron, and lipid peroxidation." *Hepatology* **23**(6): 1468-1475.
- Farinati, F., M. Piciocchi, et al. (2010). "Oxidative stress and inducible nitric oxide synthase induction in carcinogenesis." *Dig Dis* **28**(4-5): 579-584.
- Farrell, J. and J. Dienstag (2002). Epidemiology and prevention of Hepatitis B virus infection. *Hepatitis B virus*. C. L. Lai and S. Locarnini. London, International Medical Press. **1**: 115-130.
- Fattovich, G., F. Bortolotti, et al. (2008). "Natural history of chronic hepatitis B: special emphasis on disease progression and prognostic factors." *J Hepatol* **48**(2): 335-352.
- Fattovich, G., T. Stroffolini, et al. (2004). "Hepatocellular carcinoma in cirrhosis: incidence and risk factors." *Gastroenterology* **127**(5 Suppl 1): S35-50.
- Fausto, N. (2000). "Liver regeneration." *J Hepatol* **32**(1 Suppl): 19-31.
- Fausto, N., J. S. Campbell, et al. (2006). "Liver regeneration." *Hepatology* **43**(2 Suppl 1): S45-53.
- Fausto, N. and K. J. Riehle (2005). "Mechanisms of liver regeneration and their clinical implications." *J Hepatobiliary Pancreat Surg* **12**(3): 181-189.
- Fawcett, D. W. (1994). The liver and gall-bladder. *Bloom and Fawcett, a Textbook of Histology*. D. W. Fawcett, Chapman and Hall. **1**: 655.
- Fedorova, L. V., I. Dizadex, et al. (2001). "[In silico analysis of the restriction fragments length distribution in the human genome]." *Genetika* **37**(4): 456-466.
- Feitelson, M. A. and J. Lee (2007). "Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis." *Cancer Lett* **252**(2): 157-170.
- Feitelson, M. A., H. M. Reis, et al. (2009). "Putative roles of hepatitis B x antigen in the pathogenesis of chronic liver disease." *Cancer Lett* **286**(1): 69-79.
- Fellous, T. G., S. Islam, et al. (2009). "Locating the stem cell niche and tracing hepatocyte lineages in human liver." *Hepatology* **49**(5): 1655-1663.
- Feo, F., M. Frau, et al. (2009). "Genetic and epigenetic control of molecular alterations in hepatocellular carcinoma." *Exp Biol Med (Maywood)* **234**(7): 726-736.
- Foo, J., K. Leder, et al. (2011). "Stochastic dynamics of cancer initiation." *Phys Biol* **8**(1): 015002.
- Fowler, M. J., C. Greenfield, et al. (1986). "Integration of HBV-DNA may not be a prerequisite for the maintenance of the state of malignant transformation. An analysis of 110 liver biopsies." *J Hepatol* **2**(2): 218-229.
- Fredriksson, S., M. Gullberg, et al. (2002). "Protein detection using proximity-dependent DNA ligation assays." *Nat Biotechnol* **20**(5): 473-477.
- Frelin, L., T. Wahlstrom, et al. (2009). "A mechanism to explain the selection of the hepatitis e antigen-negative mutant during chronic hepatitis B virus infection." *Journal of virology* **83**(3): 1379-1392.
- Frisch, M., K. Frech, et al. (2002). "In silico prediction of scaffold/matrix attachment regions in large genomic sequences." *Genome Res* **12**(2): 349-354.
- Fujiwara, K., Y. Tanaka, et al. (2005). "Novel type of hepatitis B virus mutation: replacement mutation involving a hepatocyte nuclear factor 1 binding site tandem repeat in chronic hepatitis B virus genotype E." *J Virol* **79**(22): 14404-14410.
- Galibert, F., T. N. Chen, et al. (1982). "Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence." *J Virol* **41**(1): 51-65.

- Galipeau, P. C., X. Li, et al. (2007). "NSAIDs modulate CDKN2A, TP53, and DNA content risk for progression to esophageal adenocarcinoma." *PLoS Med* **4**(2): e67.
- Gamble, M. (2008). The Hematoxylin and Eosin. *Theory and Practice of Histological Techniques*. J. D. Bancroft and M. Gamble. Philadelphia, PA, Churchill Livingstone Elsevier. **1**: 121-134.
- Ganem, D., J. R. Pollack, et al. (1994). "Hepatitis B virus reverse transcriptase and its many roles in hepadnaviral genomic replication." *Infect Agents Dis* **3**(2-3): 85-93.
- Gao, W., Y. Kondo, et al. (2008). "Variable DNA methylation patterns associated with progression of disease in hepatocellular carcinomas." *Carcinogenesis* **29**(10): 1901-1910.
- Gerdes, J., H. Lemke, et al. (1984). "Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67." *J Immunol* **133**(4): 1710-1715.
- Giannelli, G., V. Quaranta, et al. (2003). "Tissue remodelling in liver diseases." *Histol Histopathol* **18**(4): 1267-1274.
- Giles, R. H., J. H. van Es, et al. (2003). "Caught up in a Wnt storm: Wnt signaling in cancer." *Biochim Biophys Acta* **1653**(1): 1-24.
- Golden-Mason, L., B. Palmer, et al. (2007). "Upregulation of PD-1 expression on circulating and intrahepatic hepatitis C virus-specific CD8+ T cells associated with reversible immune dysfunction." *J Virol* **81**(17): 9249-9258.
- Gonda, T. A., S. Tu, et al. (2009). "Chronic inflammation, the tumor microenvironment and carcinogenesis." *Cell Cycle* **8**(13): 2005-2013.
- Gong, L., Y. H. Li, et al. (2010). "Clonality of nodular lesions in liver cirrhosis and chromosomal abnormalities in monoclonal nodules of altered hepatocytes." *Histopathology* **56**(5): 589-599.
- Goodman, Z. D. and L. Terracciano (2007). Tumours and tumour-like lesions of the liver. *MacSween's Pathology of the Liver*. A. D. Burt, B. C. Portmann and L. D. Ferrell. Philadelphia, PA, USA, Churchill Livingstone Elsevier. **1**: 761-814.
- Gowans, E. J. and C. J. Burrell (1985). "Widespread presence of cytoplasmic HBcAg in hepatitis B infected liver detected by improved immunochemical methods." *J Clin Pathol* **38**(4): 393-398.
- Gowans, E. J., C. J. Burrell, et al. (1985). "Cytoplasmic (but not nuclear) hepatitis B virus (HBV) core antigen reflects HBV DNA synthesis at the level of the infected hepatocyte." *Intervirology* **24**(4): 220-225.
- Grisham, J. W. (1994). "Migration of hepatocytes along hepatic plates and stem cell-fed hepatocyte lineages." *Am J Pathol* **144**(5): 849-854.
- Grover, D., P. P. Majumder, et al. (2003). "Nonrandom distribution of alu elements in genes of various functional categories: insight from analysis of human chromosomes 21 and 22." *Mol Biol Evol* **20**(9): 1420-1424.
- Grunau, C., S. J. Clark, et al. (2001). "Bisulfite genomic sequencing: systematic investigation of critical experimental parameters." *Nucleic Acids Res* **29**(13): E65-65.
- Guedj, N., D. Dargere, et al. (2006). "Global proteomic analysis of microdissected cirrhotic nodules reveals significant biomarkers associated with clonal expansion." *Lab Invest* **86**(9): 951-958.
- Gust, I. D. (1986). "Towards the control of hepatitis B: an historical review." *Aust Paediatr J* **22**(4): 273-276.
- Gustafsson, J. O. (2012). *Molecular characterization of metastatic ovarian cancer by MALDI imaging mass spectrometry*, University of Adelaide.
- Gustafsson, J. O., M. K. Oehler, et al. (2011). "MALDI Imaging Mass Spectrometry (MALDI-IMS)-Application of Spatial Proteomics for Ovarian Cancer Classification and Diagnosis." *Int J Mol Sci* **12**(1): 773-794.

- Ha, S. J., E. E. West, et al. (2008). "Manipulating both the inhibitory and stimulatory immune system towards the success of therapeutic vaccination against chronic viral infections." *Immunol Rev* **223**: 317-333.
- Hall, P. (2007). Alcoholic liver disease. *MacSween's Pathology of the Liver*. A. D. Burt, B. C. Portmann and L. D. Ferrell. Philadelphia, PA, USA, Churchill Livingstone Elsevier. **1**: 327-365.
- Han, E. C., Y. S. Lee, et al. (2010). "Direct tissue analysis by MALDI-TOF mass spectrometry in human hepatocellular carcinoma." *Clin Chim Acta*.
- Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." *Cell* **144**(5): 646-674.
- Hanai, R. and A. Wada (1988). "The effects of guanine and cytosine variation on dinucleotide frequency and amino acid composition in the human genome." *J Mol Evol* **27**(4): 321-325.
- He, L. and G. J. Hannon (2004). "MicroRNAs: small RNAs with a big role in gene regulation." *Nat Rev Genet* **5**(7): 522-531.
- Herath, N. I., B. A. Leggett, et al. (2006). "Review of genetic and epigenetic alterations in hepatocarcinogenesis." *J Gastroenterol Hepatol* **21**(1 Pt 1): 15-21.
- Heymsfield, S. B., D. Gallagher, et al. (2007). "Scaling of human body composition to stature: new insights into body mass index." *Am J Clin Nutr* **86**(1): 82-91.
- Hino, O., T. Kitagawa, et al. (1984). "Detection of hepatitis B virus DNA in hepatocellular carcinomas in Japan." *Hepatology* **4**(1): 90-95.
- Hino, O., S. Tabata, et al. (1991). "Evidence for increased in vitro recombination with insertion of human hepatitis B virus DNA." *Proc Natl Acad Sci U S A* **88**(20): 9248-9252.
- Hirsch, R. C., J. E. Lavine, et al. (1990). "Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription." *Nature* **344**(6266): 552-555.
- Hoofnagle, J. H. (2009). "Reactivation of hepatitis B." *Hepatology* **49**(5 Suppl): S156-165.
- Horney, J. T. and J. T. Galambos (1977). "The liver during and after fulminant hepatitis." *Gastroenterology* **73**(4 Pt 1): 639-645.
- Hsia, C. C., R. P. Evarts, et al. (1992). "Occurrence of oval-type cells in hepatitis B virus-associated human hepatocarcinogenesis." *Hepatology* **16**(6): 1327-1333.
- Hsia, C. C., S. S. Thorgeirsson, et al. (1994). "Expression of hepatitis B surface and core antigens and transforming growth factor-alpha in "oval cells" of the liver in patients with hepatocellular carcinoma." *J Med Virol* **43**(3): 216-221.
- Hsu, H. C., Y. M. Jeng, et al. (2000). "Beta-catenin mutations are associated with a subset of low-stage hepatocellular carcinoma negative for hepatitis B virus and with favorable prognosis." *Am J Pathol* **157**(3): 763-770.
- Hsu, H. C., I. J. Su, et al. (1987). "Biologic and prognostic significance of hepatocyte hepatitis B core antigen expressions in the natural course of chronic hepatitis B virus infection." *Journal of hepatology* **5**(1): 45-50.
- Hsu, Y. S., R. N. Chien, et al. (2002). "Long-term outcome after spontaneous HBeAg seroconversion in patients with chronic hepatitis B." *Hepatology* **35**(6): 1522-1527.
- Huang, Y. H., H. H. Hung, et al. (2010). "Core Antigen Expression Is Associated with Hepatic Necroinflammation in e Antigen-negative Chronic Hepatitis B Patients with Low DNA Loads." *Clin Vaccine Immunol*.
- Hui, C. K., N. Leung, et al. (2007). "Natural history and disease progression in Chinese chronic hepatitis B patients in immune-tolerant phase." *Hepatology* **46**(2): 395-401.
- Huovila, A. P., A. M. Eder, et al. (1992). "Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment." *J Cell Biol* **118**(6): 1305-1320.
- Hyams, K. C. (1995). "Risks of chronicity following acute hepatitis B virus infection: a review." *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **20**(4): 992-1000.

- Ieta, K., E. Ojima, et al. (2007). "Identification of overexpressed genes in hepatocellular carcinoma, with special reference to ubiquitin-conjugating enzyme E2C gene expression." Int J Cancer **121**(1): 33-38.
- Iizuka, N., M. Oka, et al. (2004). "Molecular signature in three types of hepatocellular carcinoma with different viral origin by oligonucleotide microarray." Int J Oncol **24**(3): 565-574.
- Ikeda, H., M. Sasaki, et al. (2009). "Large cell change of hepatocytes in chronic viral hepatitis represents a senescent-related lesion." Hum Pathol **40**(12): 1774-1782.
- Ishak, K., A. Baptista, et al. (1995). "Histological grading and staging of chronic hepatitis." J Hepatol **22**(6): 696-699.
- Ito, T., N. Nishida, et al. (2004). "Alteration of the p14(ARF) gene and p53 status in human hepatocellular carcinomas." J Gastroenterol **39**(4): 355-361.
- Jardi, R., F. Rodriguez, et al. (2001). "Quantitative detection of hepatitis B virus DNA in serum by a new rapid real-time fluorescence PCR assay." J Viral Hepat **8**(6): 465-471.
- Jilbert, A. R. (2000). "In situ hybridization protocols for detection of viral DNA using radioactive and nonradioactive DNA probes." Methods Mol Biol **123**: 177-193.
- Jilbert, A. R., C. J. Burrell, et al. (2002). Hepatitis B virus replication. Hepatitis B virus. C. L. Lai and S. Locarnini. London, International Medical Press. **1**: 43-53.
- John, B., A. J. Enright, et al. (2004). "Human MicroRNA targets." PLoS biology **2**(11): e363.
- John, B., C. Sander, et al. (2006). "Prediction of human microRNA targets." Methods in molecular biology **342**: 101-113.
- Jones, M. L., J. D. Bancroft, et al. (2008). Connective Tissues and Stains. Theory and Practice of Histological Techniques. J. D. Bancroft and M. Gamble. Philadelphia, PA, Churchill Livingstone Elsevier. **1**: 135-160.
- Judo, M. S., A. B. Wedel, et al. (1998). "Stimulation and suppression of PCR-mediated recombination." Nucleic Acids Res **26**(7): 1819-1825.
- Jungermann, K. (1988). "Metabolic zonation of liver parenchyma." Semin Liver Dis **8**(4): 329-341.
- Jungermann, K. and N. Katz (1989). "Functional specialization of different hepatocyte populations." Physiol Rev **69**(3): 708-764.
- Jungermann, K. and T. Kietzmann (1996). "Zonation of parenchymal and nonparenchymal metabolism in liver." Annual review of nutrition **16**: 179-203.
- Jungst, C., B. Cheng, et al. (2004). "Oxidative damage is increased in human liver tissue adjacent to hepatocellular carcinoma." Hepatology **39**(6): 1663-1672.
- Junker-Niepmann, M., R. Bartenschlager, et al. (1990). "A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA." Embo J **9**(10): 3389-3396.
- Kawakita, N., S. Seki, et al. (1992). "Analysis of proliferating hepatocytes using a monoclonal antibody against proliferating cell nuclear antigen/cyclin in embedded tissues from various liver diseases fixed in formaldehyde." Am J Pathol **140**(2): 513-520.
- Kawakita, N., S. Seki, et al. (1992). "Immunocytochemical identification of proliferative hepatocytes using monoclonal antibody to proliferating cell nuclear antigen (PCNA/cyclin). Comparison with immunocytochemical staining for DNA polymerase-alpha." Am J Clin Pathol **97**(5 Suppl 1): S14-20.
- Kawanishi, S. and Y. Hiraku (2006). "Oxidative and nitrative DNA damage as biomarker for carcinogenesis with special reference to inflammation." Antioxid Redox Signal **8**(5-6): 1047-1058.
- Keir, M. E., L. M. Francisco, et al. (2007). "PD-1 and its ligands in T-cell immunity." Curr Opin Immunol **19**(3): 309-314.
- Khan, M. A. and B. M. Angus (1980). "Histochemical activity of acid phosphatase in rat liver after perfusion fixation." Acta Anat (Basel) **106**(3): 327-329.
- Kibbe, W. A. (2007). "OligoCalc: an online oligonucleotide properties calculator." Nucleic Acids Res **35**(Web Server issue): W43-46.

- Kim, H., B. K. Oh, et al. (2009). "Large liver cell change in hepatitis B virus-related liver cirrhosis." Hepatology **50**(3): 752-762.
- Kim, J. W., Q. Ye, et al. (2004). "Cancer-associated molecular signature in the tissue samples of patients with cirrhosis." Hepatology **39**(2): 518-527.
- Kim, J. Y., E. H. Song, et al. (2010). "HBx-induced hepatic steatosis and apoptosis are regulated by TNFR1- and NF-kappaB-dependent pathways." J Mol Biol **397**(4): 917-931.
- Kimbi, G. C., A. Kramvis, et al. (2005). "Integration of hepatitis B virus DNA into chromosomal DNA during acute hepatitis B." World J Gastroenterol **11**(41): 6416-6421.
- Knaus, T. and M. Nassal (1993). "The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop structure that is critical for its function." Nucleic Acids Res **21**(17): 3967-3975.
- Kobayashi, M., K. Ikeda, et al. (2006). "Dysplastic nodules frequently develop into hepatocellular carcinoma in patients with chronic viral hepatitis and cirrhosis." Cancer **106**(3): 636-647.
- Komarova, N. L. (2004). "Genomic instability in cancer: biological and mathematical approaches." Cell Cycle **3**(8): 1081-1085.
- Koniaris, L. G., I. H. McKillop, et al. (2003). "Liver regeneration." J Am Coll Surg **197**(4): 634-659.
- Kosaka, Y., K. Takase, et al. (1991). "Fulminant hepatitis B: induction by hepatitis B virus mutants defective in the precore region and incapable of encoding e antigen." Gastroenterology **100**(4): 1087-1094.
- Lai, C. L. (2002). Discovery and classification. Hepatitis B virus. C. L. Lai and S. Locarnini. London, International Medical Press. **1**: 1-8.
- Lai, M., B. J. Hyatt, et al. (2007). "The clinical significance of persistently normal ALT in chronic hepatitis B infection." J Hepatol **47**(6): 760-767.
- Lambert, M. P., A. Paliwal, et al. (2010). "Aberrant DNA methylation distinguishes hepatocellular carcinoma associated with HBV and HCV infection and alcohol intake." J Hepatol.
- Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.
- Lavanchy, D. (2004). "Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures." J Viral Hepat **11**(2): 97-107.
- Lawn, R. M., E. F. Fritsch, et al. (1978). "The isolation and characterization of linked delta- and beta-globin genes from a cloned library of human DNA." Cell **15**(4): 1157-1174.
- Le Bail, B., P. H. Bernard, et al. (1997). "Prevalence of liver cell dysplasia and association with HCC in a series of 100 cirrhotic liver explants." J Hepatol **27**(5): 835-842.
- Le Faouder, J., S. Laouirem, et al. (2011). "Imaging Mass Spectrometry Provides Fingerprints for Distinguishing Hepatocellular Carcinoma from Cirrhosis." J Proteome Res **10**(8): 3755-3765.
- Lee, H. S., M. H. Park, et al. (2007). "Novel candidate targets of Wnt/beta-catenin signaling in hepatoma cells." Life Sci **80**(7): 690-698.
- Lee, I. N., C. H. Chen, et al. (2005). "Identification of human hepatocellular carcinoma-related biomarkers by two-dimensional difference gel electrophoresis and mass spectrometry." Journal of proteome research **4**(6): 2062-2069.
- Lee, R. G., A. C. Tsamandas, et al. (1997). "Large cell change (liver cell dysplasia) and hepatocellular carcinoma in cirrhosis: matched case-control study, pathological analysis, and pathogenetic hypothesis." Hepatology **26**(6): 1415-1422.
- Lee, S., H. J. Lee, et al. (2003). "Aberrant CpG island hypermethylation along multistep hepatocarcinogenesis." Am J Pathol **163**(4): 1371-1378.
- Lee, Y. M., M. Kang, et al. (2010). "Sulfasalazine induces apoptosis of HBx-expressing cells in an NF-kappaB-independent manner." Virus Genes **40**(1): 37-43.

- Leenman, E. E., R. E. Panzer-Grumayer, et al. (2004). "Rapid determination of Epstein-Barr virus latent or lytic infection in single human cells using in situ hybridization." Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc **17**(12): 1564-1572.
- Legault, J., A. Tremblay, et al. (1997). "Preferential localization of DNA damage induced by depurination and bleomycin in a plasmid containing a scaffold-associated region." Biochem Cell Biol **75**(4): 369-375.
- Levy, L., C. A. Renard, et al. (2002). "Genetic alterations and oncogenic pathways in hepatocellular carcinoma." Ann N Y Acad Sci **963**: 21-36.
- Li, L., S. H. Chen, et al. (2008). "Identification of hepatocellular-carcinoma-associated antigens and autoantibodies by serological proteome analysis combined with protein microarray." Journal of proteome research **7**(2): 611-620.
- Liang, T. J., K. Hasegawa, et al. (1991). "A hepatitis B virus mutant associated with an epidemic of fulminant hepatitis." The New England journal of medicine **324**(24): 1705-1709.
- Liaw, Y. F., C. M. Chu, et al. (1982). "Chronic lobular hepatitis: a clinicopathological and prognostic study." Hepatology **2**(2): 258-262.
- Liaw, Y. F., G. K. Lau, et al. (2010). "Hepatitis B e Antigen Seroconversion: A Critical Event in Chronic Hepatitis B Virus Infection." Dig Dis Sci.
- Liaw, Y. F., D. Y. Lin, et al. (1989). "Natural course after the development of cirrhosis in patients with chronic type B hepatitis: a prospective study." Liver **9**(4): 235-241.
- Liaw, Y. F., D. I. Tai, et al. (1987). "Acute exacerbation in chronic type B hepatitis: comparison between HBeAg and antibody-positive patients." Hepatology **7**(1): 20-23.
- Liaw, Y. F., S. S. Yang, et al. (1985). "Acute exacerbation in hepatitis B e antigen positive chronic type B hepatitis. A clinicopathological study." Journal of hepatology **1**(3): 227-233.
- Liaw, Y. F., S. S. Yang, et al. (1985). "Acute exacerbation in hepatitis B e antigen positive chronic type B hepatitis. A clinicopathological study." J Hepatol **1**(3): 227-233.
- Lim, C. Y. and K. V. Kowdley (2007). "Optimal duration of therapy in HBV-related cirrhosis." J Antimicrob Chemother **60**(1): 2-6.
- Lim, S. O., S. J. Park, et al. (2002). "Proteome analysis of hepatocellular carcinoma." Biochemical and biophysical research communications **291**(4): 1031-1037.
- Lin, C. H., S. Y. Hsieh, et al. (2001). "Genome-wide hypomethylation in hepatocellular carcinogenesis." Cancer Res **61**(10): 4238-4243.
- Lin, N., H. Y. Chen, et al. (2005). "Apoptosis and its pathway in X gene-transfected HepG2 cells." World J Gastroenterol **11**(28): 4326-4331.
- Lin, W. J., J. Li, et al. (2003). "Suppression of hepatitis B virus core promoter by the nuclear orphan receptor TR4." The Journal of biological chemistry **278**(11): 9353-9360.
- Lin, W. R., S. N. Lim, et al. (2010). "The histogenesis of regenerative nodules in human liver cirrhosis." Hepatology **51**(3): 1017-1026.
- Lin, X., N. J. Robinson, et al. (2005). "Chronic hepatitis B virus infection in the Asia-Pacific region and Africa: review of disease progression." Journal of gastroenterology and hepatology **20**(6): 833-843.
- Lindberg, J. and A. Lindholm (1988). "HBsAg-positive Swedish blood donors: natural history and origin of infection." Scand J Infect Dis **20**(4): 377-382.
- Lindros, K. O. (1997). "Zonation of cytochrome P450 expression, drug metabolism and toxicity in liver." Gen Pharmacol **28**(2): 191-196.
- Liu, Q. Y., M. Ribocco-Lutkiewicz, et al. (2003). "Mapping the initial DNA breaks in apoptotic Jurkat cells using ligation-mediated PCR." Cell Death Differ **10**(3): 278-289.
- Llovet, J. M., Y. Chen, et al. (2006). "A molecular signature to discriminate dysplastic nodules from early hepatocellular carcinoma in HCV cirrhosis." Gastroenterology **131**(6): 1758-1767.

- Loeb, D. D. and R. Tian (1995). "Transfer of the minus strand of DNA during hepadnavirus replication is not invariable but prefers a specific location." J Virol **69**(11): 6886-6891.
- Loeb, D. D., R. Tian, et al. (1996). "Mutations within DR2 independently reduce the amount of both minus- and plus-strand DNA synthesized during duck hepatitis B virus replication." Journal of virology **70**(12): 8684-8690.
- Lok, A. S. (2009). "Hepatitis B: liver fibrosis and hepatocellular carcinoma." Gastroenterol Clin Biol **33**(10-11): 911-915.
- Lok, A. S., L. B. Seeff, et al. (2009). "Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease." Gastroenterology **136**(1): 138-148.
- London, W. T. and B. S. Blumberg (1983). Hepatitis B and related viruses in chronic hepatitis, cirrhosis and hepatocellular carcinoma in man and animals. Chronic active liver disease. S. Cohen and R. D. Soloway. New York, Churchill Livingstone. **2**: 147-170.
- Long, A. A. and P. Komminoth (1997). "In situ PCR. An overview." Methods Mol Biol **71**: 141-161.
- Lopez, J. B. (2005). "Recent developments in the first detection of hepatocellular carcinoma." Clin Biochem Rev **26**(3): 65-79.
- Lowes, K. N., B. A. Brennan, et al. (1999). "Oval cell numbers in human chronic liver diseases are directly related to disease severity." Am J Pathol **154**(2): 537-541.
- Ma, H., L. Wei, et al. (2008). "Clinical features and survival in Chinese patients with hepatitis B e antigen-negative hepatitis B virus-related cirrhosis." J Gastroenterol Hepatol **23**(8 Pt 1): 1250-1258.
- Macdonald, R. A., A. E. Rogers, et al. (1963). "Growth and Regeneration of the Liver." Ann N Y Acad Sci **111**: 70-86.
- Malhi, H. and R. J. Kaufman (2011). "Endoplasmic reticulum stress in liver disease." J Hepatol **54**(4): 795-809.
- Mancini, R., L. Marucci, et al. (1994). "Immunohistochemical analysis of S-phase cells in normal human and rat liver by PC10 monoclonal antibody." Liver **14**(2): 57-64.
- Manivasakam, P. and R. H. Schiestl (1998). "Nonhomologous end joining during restriction enzyme-mediated DNA integration in *Saccharomyces cerevisiae*." Mol Cell Biol **18**(3): 1736-1745.
- Marrero, J. A., M. Kudo, et al. (2010). "The challenge of prognosis and staging for hepatocellular carcinoma." Oncologist **15 Suppl 4**: 23-33.
- Marsaglia, G. (2003). "Xorshift RNGs." J Stat Soft **8**(14).
- Martinez-Chantar, M. L., F. J. Corrales, et al. (2002). "Spontaneous oxidative stress and liver tumors in mice lacking methionine adenosyltransferase 1A." FASEB J **16**(10): 1292-1294.
- Maruyama, T., S. Iino, et al. (1993). "Serology of acute exacerbation in chronic hepatitis B virus infection." Gastroenterology **105**(4): 1141-1151.
- Mason, W. S., C. Aldrich, et al. (1982). "Asymmetric replication of duck hepatitis B virus DNA in liver cells: Free minus-strand DNA." Proc Natl Acad Sci U S A **79**(13): 3997-4001.
- Mason, W. S., J. Cullen, et al. (1998). "Lamivudine therapy of WHV-infected woodchucks." Virology **245**(1): 18-32.
- Mason, W. S., A. R. Jilbert, et al. (2005). "Clonal expansion of hepatocytes during chronic woodchuck hepatitis virus infection." Proc Natl Acad Sci U S A **102**(4): 1139-1144.
- Mason, W. S., C. Liu, et al. (2010). "Clonal expansion of normal-appearing human hepatocytes during chronic hepatitis B virus infection." J Virol **84**(16): 8308-8315.
- Mason, W. S., H. C. Low, et al. (2009). "Detection of clonally expanded hepatocytes in chimpanzees with chronic hepatitis B virus infection." J Virol **83**(17): 8396-8408.
- Matsubara, K. and T. Tokino (1990). "Integration of hepatitis B virus DNA and its implications for hepatocarcinogenesis." Mol Biol Med **7**(3): 243-260.

- Matsumoto, T. and M. Kawakami (1982). "The unit-concept of hepatic parenchyma--a re-examination based on angioarchitectural studies." Acta Pathol Jpn **32 Suppl 2**: 285-314.
- Mauad, T. H., C. M. van Nieuwkerk, et al. (1994). "Mice with homozygous disruption of the *mdr2* P-glycoprotein gene. A novel animal model for studies of nonsuppurative inflammatory cholangitis and hepatocarcinogenesis." Am J Pathol **145**(5): 1237-1245.
- McConaughy, B. L., C. D. Laird, et al. (1969). "Nucleic acid reassociation in formamide." Biochemistry **8**(8): 3289-3295.
- McMahon, B. J. (2009). "The natural history of chronic hepatitis B virus infection." Hepatology **49**(5 Suppl): S45-55.
- Mendy, M. E., T. Welzel, et al. (2010). "Hepatitis B viral load and risk for liver cirrhosis and hepatocellular carcinoma in The Gambia, West Africa." J Viral Hepat **17**(2): 115-122.
- Merlo, L. M., J. W. Pepper, et al. (2006). "Cancer as an evolutionary and ecological process." Nat Rev Cancer **6**(12): 924-935.
- Merlo, L. M., N. A. Shah, et al. (2010). "A comprehensive survey of clonal diversity measures in Barrett's esophagus as biomarkers of progression to esophageal adenocarcinoma." Cancer prevention research **3**(11): 1388-1397.
- Meyers, R. B., J. L. Fredenburgh, et al. (2008). Carbohydrates. Theory and Practice of Histological Techniques. J. D. Bancroft and M. Gamble. Philadelphia, PA, Churchill Livingstone Elsevier. **1**: 161-186.
- Michalopoulos, G. K. (2010). "Liver regeneration after partial hepatectomy: critical analysis of mechanistic dilemmas." Am J Pathol **176**(1): 2-13.
- Michitaka, K., N. Horiike, et al. (2007). "Heterogeneity of hepatitis B virus genotype D in Japan." Intervirology **50**(2): 150-155.
- Milich, D. and T. J. Liang (2003). "Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection." Hepatology **38**(5): 1075-1086.
- Minami, M., K. Poussin, et al. (1995). "A novel PCR technique using Alu-specific primers to identify unknown flanking sequences from the human genome." Genomics **29**(2): 403-408.
- Mishra, L., T. Banker, et al. (2009). "Liver stem cells and hepatocellular carcinoma." Hepatology **49**(1): 318-329.
- Mizukoshi, E., J. Sidney, et al. (2004). "Cellular immune responses to the hepatitis B virus polymerase." J Immunol **173**(9): 5863-5871.
- Murakami, Y., K. Saigo, et al. (2005). "Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas." Gut **54**(8): 1162-1168.
- Murray, J. M., S. F. Wieland, et al. (2005). "Dynamics of hepatitis B virus clearance in chimpanzees." Proc Natl Acad Sci U S A **102**(49): 17780-17785.
- Murugavel, K. G., S. Mathews, et al. (2008). "Alpha-fetoprotein as a tumor marker in hepatocellular carcinoma: investigations in south Indian subjects with hepatotropic virus and aflatoxin etiologies." Int J Infect Dis **12**(6): e71-76.
- Nagaya, T., T. Nakamura, et al. (1987). "The mode of hepatitis B virus DNA integration in chromosomes of human hepatocellular carcinoma." Genes Dev **1**(8): 773-782.
- Nagy, P., A. Kiss, et al. (1998). "Dexamethasone inhibits the proliferation of hepatocytes and oval cells but not bile duct cells in rat liver." Hepatology **28**(2): 423-429.
- Nakashima, H., K. Nishikawa, et al. (1997). "Differences in dinucleotide frequencies of human, yeast, and *Escherichia coli* genes." DNA Res **4**(3): 185-192.
- Nassal, M. and A. Rieger (1996). "A bulged region of the hepatitis B virus RNA encapsidation signal contains the replication origin for discontinuous first-strand DNA synthesis." Journal of virology **70**(5): 2764-2773.
- NCBI. (2011, November 2010). "NCBI *Homo sapiens* genome map viewer." Build 37.2. Retrieved 10th March, 2011, from http://www.ncbi.nlm.nih.gov/projects/mapview/map_search.cgi?taxid=9606.

- Newbold, J. E., H. Xin, et al. (1995). "The covalently closed duplex form of the hepadnavirus genome exists in situ as a heterogeneous population of viral minichromosomes." *J Virol* **69**(6): 3350-3357.
- Nguyen, T., A. J. Thompson, et al. (2010). "Hepatitis B surface antigen levels during the natural history of chronic hepatitis B: a perspective on Asia." *J Hepatol* **52**(4): 508-513.
- Nhieu, J. T., C. A. Renard, et al. (1999). "Nuclear accumulation of mutated beta-catenin in hepatocellular carcinoma is associated with increased cell proliferation." *Am J Pathol* **155**(3): 703-710.
- Nishida, N. and A. Goel (2011). "Genetic and epigenetic signatures in human hepatocellular carcinoma: a systematic review." *Curr Genomics* **12**(2): 130-137.
- Nomoto, S., T. Kinoshita, et al. (2007). "Hypermethylation of multiple genes as clonal markers in multicentric hepatocellular carcinoma." *Br J Cancer* **97**(9): 1260-1265.
- Nussinov, R. (1980). "Some rules in the ordering of nucleotides in the DNA." *Nucleic Acids Res* **8**(19): 4545-4562.
- Nussinov, R. (1981). "Eukaryotic dinucleotide preference rules and their implications for degenerate codon usage." *J Mol Biol* **149**(1): 125-131.
- Odze, R. D. and C. C. Maley (2010). "Neoplasia without dysplasia: lessons from Barrett esophagus and other tubal gut neoplasms." *Arch Pathol Lab Med* **134**(6): 896-906.
- Oh, B. K., Y. J. Kim, et al. (2005). "Up-regulation of telomere-binding proteins, TRF1, TRF2, and TIN2 is related to telomere shortening during human multistep hepatocarcinogenesis." *Am J Pathol* **166**(1): 73-80.
- Ohkubo, K., Y. Kato, et al. (2002). "Viral load is a significant prognostic factor for hepatitis B virus-associated hepatocellular carcinoma." *Cancer* **94**(10): 2663-2668.
- Ohno, S. (1988). "Codon preference is but an illusion created by the construction principle of coding sequences." *Proc Natl Acad Sci U S A* **85**(12): 4378-4382.
- Ohshima, K., A. Ohgami, et al. (1998). "Random integration of HTLV-1 provirus: increasing chromosomal instability." *Cancer Lett* **132**(1-2): 203-212.
- Oltvai, Z. N., C. L. Milliman, et al. (1993). "Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death." *Cell* **74**(4): 609-619.
- Omata, M., T. Ehata, et al. (1991). "Mutations in the precore region of hepatitis B virus DNA in patients with fulminant and severe hepatitis." *The New England journal of medicine* **324**(24): 1699-1704.
- Ott, M., Q. Ma, et al. (1999). "Regulation of hepatitis B virus expression in progenitor and differentiated cell types: evidence for negative transcriptional control in nonpermissive cells." *Gene Expr* **8**(3): 175-186.
- Ou, Y. C., R. B. Conolly, et al. (2001). "A clonal growth model: time-course simulations of liver foci growth following penta- or hexachlorobenzene treatment in a medium-term bioassay." *Cancer Res* **61**(5): 1879-1889.
- Paku, S., J. Schnur, et al. (2001). "Origin and structural evolution of the early proliferating oval cells in rat liver." *Am J Pathol* **158**(4): 1313-1323.
- Paran, N., B. Geiger, et al. (2001). "HBV infection of cell culture: evidence for multivalent and cooperative attachment." *Embo J* **20**(16): 4443-4453.
- Park, S. G., Y. Kim, et al. (2003). "Fidelity of hepatitis B virus polymerase." *Eur J Biochem* **270**(14): 2929-2936.
- Park, S. H., C. H. Kim, et al. (2010). "Development and Validation of a Model to Predict Advanced Fibrosis in Chronic Hepatitis B Virus-Infected Patients with High Viral Load and Normal or Minimally Raised ALT." *Dig Dis Sci*.
- Park, W. W. (1957). "The incidence of sex chromatin in trophoblast." *Ned Tijdschr Geneeskd* **101**(32): 1523.
- Park, W. W. (1957). "The occurrence of sex chromatin in early human and macaque embryos." *J Anat* **91**(3): 369-373.

- Pasquinelli, C., K. Bhavani, et al. (1992). "Multiple oncogenes and tumor suppressor genes are structurally and functionally intact during hepatocarcinogenesis in hepatitis B virus transgenic mice." *Cancer Res* **52**(10): 2823-2829.
- Paterlini-Brechot, P., K. Saigo, et al. (2003). "Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human telomerase gene." *Oncogene* **22**(25): 3911-3916.
- Patzer, E. J., G. R. Nakamura, et al. (1986). "Intracellular assembly and packaging of hepatitis B surface antigen particles occur in the endoplasmic reticulum." *J Virol* **58**(3): 884-892.
- Peng, G., S. Li, et al. (2008). "PD-1 upregulation is associated with HBV-specific T cell dysfunction in chronic hepatitis B patients." *Mol Immunol* **45**(4): 963-970.
- Penna, A., M. Artini, et al. (1996). "Long-lasting memory T cell responses following self-limited acute hepatitis B." *J Clin Invest* **98**(5): 1185-1194.
- Persson, H., M. Perricaudet, et al. (1979). "Purification of RNA-DNA hybrids by exclusion chromatography." *J Biol Chem* **254**(16): 7999-8003.
- Pett, M. R., W. O. Alazawi, et al. (2004). "Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes." *Cancer Res* **64**(4): 1359-1368.
- Pienta, K. J., R. H. Getzenberg, et al. (1991). "Cell structure and DNA organization." *Crit Rev Eukaryot Gene Expr* **1**(4): 355-385.
- Pineau, P., A. Marchio, et al. (1998). "Extensive analysis of duplicated-inverted hepatitis B virus integrations in human hepatocellular carcinoma." *J Gen Virol* **79** (Pt 3): 591-600.
- Pitot, H. C., Y. P. Dragan, et al. (1996). "Quantitation of multistage carcinogenesis in rat liver." *Toxicologic pathology* **24**(1): 119-128.
- Popper, H., F. Paronetto, et al. (1965). "Immune processes in the pathogenesis of liver disease." *Ann N Y Acad Sci* **124**(2): 781-799.
- Portmann, B. C., R. J. Thompson, et al. (2007). Genetic and metabolic disease. *MacSween's Pathology of the Liver*. A. D. Burt, B. C. Portmann and L. D. Ferrell. Philadelphia, PA, USA, Churchill Livingstone Elsevier. **1**: 199-326.
- Powell, C. L., O. Kosyk, et al. (2006). "Phenotypic anchoring of acetaminophen-induced oxidative stress with gene expression profiles in rat liver." *Toxicol Sci* **93**(1): 213-222.
- Puschel, G. P. and K. Jungermann (1994). "Integration of function in the hepatic acinus: intercellular communication in neural and humoral control of liver metabolism." *Progress in liver diseases* **12**: 19-46.
- Qu, Z. L., S. Q. Zou, et al. (2005). "Upregulation of human telomerase reverse transcriptase mRNA expression by in vitro transfection of hepatitis B virus X gene into human hepatocarcinoma and cholangiocarcinoma cells." *World J Gastroenterol* **11**(36): 5627-5632.
- Radziwill, G., W. Tucker, et al. (1990). "Mutational analysis of the hepatitis B virus P gene product: domain structure and RNase H activity." *Journal of virology* **64**(2): 613-620.
- Ramire, C., C. Scholtes, et al. (2008). "Transactivation of the hepatitis B virus core promoter by the nuclear receptor FXRalpha." *Journal of virology* **82**(21): 10832-10840.
- Raney, A. K., D. R. Milich, et al. (1990). "Differentiation-specific transcriptional regulation of the hepatitis B virus large surface antigen gene in human hepatoma cell lines." *Journal of virology* **64**(5): 2360-2368.
- Reuter, S., S. C. Gupta, et al. (2010). "Oxidative stress, inflammation, and cancer: how are they linked?" *Free Radic Biol Med* **49**(11): 1603-1616.
- Rhim, J. A., E. P. Sandgren, et al. (1994). "Replacement of diseased mouse liver by hepatic cell transplantation." *Science* **263**(5150): 1149-1152.
- Rieger, A. and M. Nassal (1996). "Specific hepatitis B virus minus-strand DNA synthesis requires only the 5' encapsidation signal and the 3'-proximal direct repeat DR1." *Journal of virology* **70**(1): 585-589.

- Robertson, K. D. and P. A. Jones (2000). "DNA methylation: past, present and future directions." Carcinogenesis **21**(3): 461-467.
- Rodrigues, R. R. and C. T. Barry (2011). "Gene pathway analysis of hepatocellular carcinoma genomic expression datasets." J Surg Res **170**(1): e85-92.
- Rodriguez-Diaz, J. L., V. Rosas-Camargo, et al. (2007). "Clinical and pathological factors associated with the development of hepatocellular carcinoma in patients with hepatitis virus-related cirrhosis: a long-term follow-up study." Clin Oncol (R Coll Radiol) **19**(3): 197-203.
- Rogers, A. E., J. A. Shaka, et al. (1961). "Regeneration of the liver. Absence of a "humoral factor" affecting hepatic regeneration in parabiotic rats." Am J Pathol **39**: 561-578.
- Rogler, C. E., M. Sherman, et al. (1985). "Deletion in chromosome 11p associated with a hepatitis B integration site in hepatocellular carcinoma." Science **230**(4723): 319-322.
- Roskams, T., V. J. Desmet, et al. (2007). Development, structure and function of the liver. MacSween's Pathology of the Liver. A. D. Burt, B. C. Portmann and L. D. Ferrell. Philadelphia, PA, USA, Churchill Livingstone Elsevier. **1**: 1-74.
- Saigo, K., K. Yoshida, et al. (2008). "Integration of hepatitis B virus DNA into the myeloid/lymphoid or mixed-lineage leukemia (MLL4) gene and rearrangements of MLL4 in human hepatocellular carcinoma." Hum Mutat **29**(5): 703-708.
- Sakamoto, M. (2009). "Early HCC: diagnosis and molecular markers." J Gastroenterol **44 Suppl 19**: 108-111.
- Sakamoto, T., Y. Tanaka, et al. (2006). "Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippines." J Gen Virol **87**(Pt 7): 1873-1882.
- Santoni, V., M. Molloy, et al. (2000). "Membrane proteins and proteomics: un amour impossible?" Electrophoresis **21**(6): 1054-1070.
- Sari, A., Y. Dere, et al. (2011). "Relation of hepatitis B core antigen expression with histological activity, serum HBeAg, and HBV DNA levels." Indian journal of pathology & microbiology **54**(2): 355-358.
- Sato, S., K. Suzuki, et al. (1995). "Hepatitis B virus strains with mutations in the core promoter in patients with fulminant hepatitis." Annals of internal medicine **122**(4): 241-248.
- Schulze, A., P. Gripon, et al. (2007). "Hepatitis B virus infection initiates with a large surface protein-dependent binding to heparan sulfate proteoglycans." Hepatology **46**(6): 1759-1768.
- Schutte, K., J. Bornschein, et al. (2009). "Hepatocellular carcinoma--epidemiological trends and risk factors." Dig Dis **27**(2): 80-92.
- Sebastiani, G., D. Tempesta, et al. (2012). "Hepatic iron overload is common in chronic hepatitis B and is more severe in patients coinfecting with hepatitis D virus." Journal of viral hepatitis **19**(2): e170-176.
- Seeger, C. and W. S. Mason (2000). "Hepatitis B virus biology." Microbiol Mol Biol Rev **64**(1): 51-68.
- Seeger, C., J. Summers, et al. (1991). "Viral DNA synthesis." Curr Top Microbiol Immunol **168**: 41-60.
- Seeley, E. H. and R. M. Caprioli (2008). "Molecular imaging of proteins in tissues by mass spectrometry." Proc Natl Acad Sci U S A **105**(47): 18126-18131.
- Seki, S., H. Sakaguchi, et al. (2000). "Outcomes of dysplastic nodules in human cirrhotic liver: a clinicopathological study." Clin Cancer Res **6**(9): 3469-3473.
- Shafritz, D. A., D. Shouval, et al. (1981). "Integration of hepatitis B virus DNA into the genome of liver cells in chronic liver disease and hepatocellular carcinoma. Studies in percutaneous liver biopsies and post-mortem tissue specimens." N Engl J Med **305**(18): 1067-1073.
- Sharma, S. K., N. Saini, et al. (2005). "Hepatitis B virus: inactive carriers." Virol J **2**: 82.

- Shepard, T. H., M. Shi, et al. (1988). "Organ weight standards for human fetuses." *Pediatr Pathol* **8**(5): 513-524.
- Shera, K. A., C. A. Shera, et al. (2001). "Small tumor virus genomes are integrated near nuclear matrix attachment regions in transformed cells." *J Virol* **75**(24): 12339-12346.
- Shimoda, R., M. Nagashima, et al. (1994). "Increased formation of oxidative DNA damage, 8-hydroxydeoxyguanosine, in human livers with chronic hepatitis." *Cancer Res* **54**(12): 3171-3172.
- Shimoda, T., T. Uchida, et al. (1980). "A 6-year-old boy having hepatocellular carcinoma associated with hepatitis B surface antigenemia." *Am J Clin Pathol* **74**(6): 827-831.
- Shioiri, C. and N. Takahata (2001). "Skew of mononucleotide frequencies, relative abundance of dinucleotides, and DNA strand asymmetry." *J Mol Evol* **53**(4-5): 364-376.
- Singh, G. B., J. A. Kramer, et al. (1997). "Mathematical model to predict regions of chromatin attachment to the nuclear matrix." *Nucleic Acids Res* **25**(7): 1419-1425.
- Snowberger, N., S. Chinnakotla, et al. (2007). "Alpha fetoprotein, ultrasound, computerized tomography and magnetic resonance imaging for detection of hepatocellular carcinoma in patients with advanced cirrhosis." *Aliment Pharmacol Ther* **26**(9): 1187-1194.
- Solt, D. B., A. Medline, et al. (1977). "Rapid emergence of carcinogen-induced hyperplastic lesions in a new model for the sequential analysis of liver carcinogenesis." *Am J Pathol* **88**(3): 595-618.
- Soresi, M., C. Magliarisi, et al. (2003). "Usefulness of alpha-fetoprotein in the diagnosis of hepatocellular carcinoma." *Anticancer Res* **23**(2C): 1747-1753.
- Speel, E. J., F. C. Ramaekers, et al. (1995). "Cytochemical detection systems for in situ hybridization, and the combination with immunocytochemistry, 'who is still afraid of red, green and blue?'" *The Histochemical journal* **27**(11): 833-858.
- Sperry, A. O., V. C. Blasquez, et al. (1989). "Dysfunction of chromosomal loop attachment sites: illegitimate recombination linked to matrix association regions and topoisomerase II." *Proc Natl Acad Sci U S A* **86**(14): 5497-5501.
- St Clair, E. G., S. J. Anderson, et al. (1997). "Bcl-2 counters apoptosis by Bax heterodimerization-dependent and -independent mechanisms in the T-cell lineage." *J Biol Chem* **272**(46): 29347-29355.
- Staprans, S., D. D. Loeb, et al. (1991). "Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA." *J Virol* **65**(3): 1255-1262.
- Stefaniuk, P., J. Cianciara, et al. (2010). "Present and future possibilities for early diagnosis of hepatocellular carcinoma." *World J Gastroenterol* **16**(4): 418-424.
- Stelzl, E., Z. Muller, et al. (2004). "Rapid quantification of hepatitis B virus DNA by automated sample preparation and real-time PCR." *J Clin Microbiol* **42**(6): 2445-2449.
- Su, H., J. Zhao, et al. (2008). "Large-scale analysis of the genetic and epigenetic alterations in hepatocellular carcinoma from Southeast China." *Mutat Res* **641**(1-2): 27-35.
- Su, Q. and P. Bannasch (2003). "Relevance of hepatic preneoplasia for human hepatocarcinogenesis." *Toxicol Pathol* **31**(1): 126-133.
- Su, Q., A. Benner, et al. (1997). "Human hepatic preneoplasia: phenotypes and proliferation kinetics of foci and nodules of altered hepatocytes and their relationship to liver cell dysplasia." *Virchows Arch* **431**(6): 391-406.
- Sudmant, P. H., J. O. Kitzman, et al. (2010). "Diversity of human copy number variation and multicopy genes." *Science* **330**(6004): 641-646.
- Summers, J., A. R. Jilbert, et al. (2003). "Hepatocyte turnover during resolution of a transient hepadnaviral infection." *Proc Natl Acad Sci U S A* **100**(20): 11652-11659.
- Summers, J. and W. S. Mason (1982). "Replication of the genome of a hepatitis B--like virus by reverse transcription of an RNA intermediate." *Cell* **29**(2): 403-415.
- Summers, J. and W. S. Mason (2004). "Residual integrated viral DNA after hepadnavirus clearance by nucleoside analog therapy." *Proc Natl Acad Sci U S A* **101**(2): 638-640.

- Sun, L., P. K. Chow, et al. (1999). "Liver regeneration after partial hepatectomy is non-uniform: flow cytometric bromodeoxyuridine incorporation and cell cycle studies in a porcine model." *Res Exp Med (Berl)* **198**(5): 229-236.
- Sun, S., R. T. Poon, et al. (2010). "Proteomics of hepatocellular carcinoma: serum vimentin as a surrogate marker for small tumors (<or=2 cm)." *Journal of proteome research* **9**(4): 1923-1930.
- Suzuki, F., N. Akuta, et al. (2007). "Selection of a virus strain resistant to entecavir in a nucleoside-naïve patient with hepatitis B of genotype H." *J Clin Virol* **39**(2): 149-152.
- Takada, S., Y. Gotoh, et al. (1990). "Integrated structures of HBV DNA in chronic hepatitis and hepatoma tissues." *Gastroenterol Jpn* **25 Suppl 2**: 31-37.
- Takahashi, K., Y. Akahane, et al. (1998). "Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates." *Arch Virol* **143**(12): 2313-2326.
- Tan, A., S. H. Yeh, et al. (2008). "Viral hepatocarcinogenesis: from infection to cancer." *Liver Int* **28**(2): 175-188.
- Tanaka, T., H. Miyamoto, et al. (1986). "Primary hepatocellular carcinoma with hepatitis B virus-DNA integration in a 4-year-old boy." *Hum Pathol* **17**(2): 202-204.
- Tanaka, Y., M. Esumi, et al. (1988). "Frequent integration of hepatitis B virus DNA in noncancerous liver tissue from hepatocellular carcinoma patients." *J Med Virol* **26**(1): 7-14.
- Tandon, B. N., S. K. Acharya, et al. (1996). "Epidemiology of hepatitis B virus infection in India." *Gut* **38 Suppl 2**: S56-59.
- Tavis, J. E., S. Perri, et al. (1994). "Hepadnavirus reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer." *J Virol* **68**(6): 3536-3543.
- Tennant, B. C., I. A. Toshkov, et al. (2004). "Hepatocellular carcinoma in the woodchuck model of hepatitis B virus infection." *Gastroenterology* **127**(5 Suppl 1): S283-293.
- Teutsch, H. F. (2005). "The modular microarchitecture of human liver." *Hepatology* **42**(2): 317-325.
- Teutsch, H. F., D. Schuerfeld, et al. (1999). "Three-dimensional reconstruction of parenchymal units in the liver of the rat." *Hepatology* **29**(2): 494-505.
- Tham, K. M., V. T. Chow, et al. (1991). "Diagnostic sensitivity of polymerase chain reaction and Southern blot hybridization for the detection of human papillomavirus DNA in biopsy specimens from cervical lesions." *Am J Clin Pathol* **95**(5): 638-646.
- Theise, N. D., H. C. Bodenheimer Jr., et al. (2007). Acute and chronic viral hepatitis. *MacSween's Pathology of the Liver*. A. D. Burt, B. C. Portmann and L. D. Ferrell. Philadelphia, PA, USA, Churchill Livingstone Elsevier. **1**: 399-442.
- Thompson, M. D. and S. P. Monga (2007). "WNT/beta-catenin signaling in liver health and disease." *Hepatology* **45**(5): 1298-1305.
- Tiollais, P., P. Charnay, et al. (1981). "Biology of hepatitis B virus." *Science* **213**(4506): 406-411.
- Tischoff, I. and A. Tannapfe (2008). "DNA methylation in hepatocellular carcinoma." *World J Gastroenterol* **14**(11): 1741-1748.
- Treichel, U., K. H. Meyer zum Buschenfelde, et al. (1997). "Receptor-mediated entry of hepatitis B virus particles into liver cells." *Arch Virol* **142**(3): 493-498.
- Truant, R., J. Antunovic, et al. (1995). "Direct interaction of the hepatitis B virus HBx protein with p53 leads to inhibition by HBx of p53 response element-directed transactivation." *J Virol* **69**(3): 1851-1859.
- Tsai, S. L., P. J. Chen, et al. (1992). "Acute exacerbations of chronic type B hepatitis are accompanied by increased T cell responses to hepatitis B core and e antigens. Implications for hepatitis B e antigen seroconversion." *The Journal of clinical investigation* **89**(1): 87-96.

- Tsang, P. S., H. Trinh, et al. (2008). "Significant prevalence of histologic disease in patients with chronic hepatitis B and mildly elevated serum alanine aminotransferase levels." Clin Gastroenterol Hepatol **6**(5): 569-574.
- Tsuge, M., N. Hiraga, et al. (2010). "HBx protein is indispensable for development of viremia in human hepatocyte chimeric mice." J Gen Virol.
- Tuttleman, J. S., C. Pourcel, et al. (1986). "Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells." Cell **47**(3): 451-460.
- Um, T. H., H. Kim, et al. (2010). "Aberrant CpG island hypermethylation in dysplastic nodules and early HCC of hepatitis B virus-related human multistep hepatocarcinogenesis." J Hepatol.
- Van Thiel, D. H., J. S. Gavaler, et al. (1987). "Rapid growth of an intact human liver transplanted into a recipient larger than the donor." Gastroenterology **93**(6): 1414-1419.
- Vandegraaff, N., R. Kumar, et al. (2001). "Kinetics of human immunodeficiency virus type 1 (HIV) DNA integration in acutely infected cells as determined using a novel assay for detection of integrated HIV DNA." J Virol **75**(22): 11253-11260.
- Vassilopoulos, D., I. Rapti, et al. (2008). "Cellular immune responses in hepatitis B virus e antigen negative chronic hepatitis B." J Viral Hepat **15**(11): 817-826.
- Velazquez, R. F., M. Rodriguez, et al. (2003). "Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis." Hepatology **37**(3): 520-527.
- Venter, J. C., M. D. Adams, et al. (2001). "The sequence of the human genome." Science **291**(5507): 1304-1351.
- Villa, E. and G. Fattovich (2010). "No inflammation? No cancer! Clear HBV early and live happily." J Hepatol **52**(5): 768-770.
- Villeneuve, J. P. (2005). "The natural history of chronic hepatitis B virus infection." J Clin Virol **34 Suppl 1**: S139-142.
- Vincent, T. L. and R. A. Gatenby (2008). "An evolutionary model for initiation, promotion, and progression in carcinogenesis." Int J Oncol **32**(4): 729-737.
- Vivekanandan, P., H. D. Daniel, et al. (2010). "Hepatitis B virus replication induces methylation of both host and viral DNA." J Virol **84**(9): 4321-4329.
- Vivekanandan, P., D. Thomas, et al. (2008). "Hepatitis B viral DNA is methylated in liver tissues." J Viral Hepat **15**(2): 103-107.
- Vivekanandan, P., D. Thomas, et al. (2009). "Methylation regulates hepatitis B viral protein expression." J Infect Dis **199**(9): 1286-1291.
- Walboomers, J. M., W. J. Melchers, et al. (1988). "Sensitivity of in situ detection with biotinylated probes of human papilloma virus type 16 DNA in frozen tissue sections of squamous cell carcinomas of the cervix." The American journal of pathology **131**(3): 587-594.
- Wang, C. C., L. Y. Lim, et al. (2008). "Factors predictive of significant hepatic fibrosis in adults with chronic hepatitis B and normal serum ALT." J Clin Gastroenterol **42**(7): 820-826.
- Wang, G. H. and C. Seeger (1992). "The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis." Cell **71**(4): 663-670.
- Wang, H., Y. Ye, et al. (2010). "Proteomic and functional analyses reveal the potential involvement of endoplasmic reticulum stress and alpha-CP1 in the anticancer activities of oridonin in HepG2 cells." Integr Cancer Ther **10**(2): 160-167.
- Wang, H. C., W. Huang, et al. (2006). "Hepatitis B virus pre-S mutants, endoplasmic reticulum stress and hepatocarcinogenesis." Cancer Sci **97**(8): 683-688.
- Wang, H. C., H. C. Wu, et al. (2003). "Different types of ground glass hepatocytes in chronic hepatitis B virus infection contain specific pre-S mutants that may induce endoplasmic reticulum stress." Am J Pathol **163**(6): 2441-2449.
- Wang, J., F. Zindy, et al. (1992). "Modification of cyclin A expression by hepatitis B virus DNA integration in a hepatocellular carcinoma." Oncogene **7**(8): 1653-1656.

- Wang, J. C. and N. Davidson (1966). "On the probability of ring closure of lambda DNA." J Mol Biol **19**(2): 469-482.
- Wang, J. C. and N. Davidson (1966). "Thermodynamic and kinetic studies on the interconversion between the linear and circular forms of phage lambda DNA." J Mol Biol **15**(1): 111-123.
- Wang, L., L. Sun, et al. (2011). "Cyclin-dependent kinase inhibitor 3 (CDKN3) novel cell cycle computational network between human non-malignancy associated hepatitis/cirrhosis and hepatocellular carcinoma (HCC) transformation." Cell proliferation **44**(3): 291-299.
- Wang, Y., S. H. Lau, et al. (2004). "Characterization of HBV integrants in 14 hepatocellular carcinomas: association of truncated X gene and hepatocellular carcinogenesis." Oncogene **23**(1): 142-148.
- Watanabe, S., K. Okita, et al. (1983). "Morphologic studies of the liver cell dysplasia." Cancer **51**(12): 2197-2205.
- Wei, Y., G. Fourel, et al. (1992). "Hepadnavirus integration: mechanisms of activation of the N-myc2 retrotransposon in woodchuck liver tumors." J Virol **66**(9): 5265-5276.
- Wei, Y., A. Ponzetto, et al. (1992). "Multiple rearrangements and activated expression of c-myc induced by woodchuck hepatitis virus integration in a primary liver tumour." Res Virol **143**(2): 89-96.
- WHO. (2000). "Hepatitis B: World Health Organization Fact Sheet 204." Fact sheets Retrieved 3/3/07, 2007, from <http://www.who.int/mediacentre/factsheets/fs204/en/>.
- Wieland, S. F., H. C. Spangenberg, et al. (2004). "Expansion and contraction of the hepatitis B virus transcriptional template in infected chimpanzees." Proc Natl Acad Sci U S A **101**(7): 2129-2134.
- Will, H., W. Reiser, et al. (1987). "Replication strategy of human hepatitis B virus." J Virol **61**(3): 904-911.
- Wolf, H. K. and G. K. Michalopoulos (1992). "Hepatocyte regeneration in acute fulminant and nonfulminant hepatitis: a study of proliferating cell nuclear antigen expression." Hepatology **15**(4): 707-713.
- Wong, C. M., S. T. Fan, et al. (2001). "beta-Catenin mutation and overexpression in hepatocellular carcinoma: clinicopathologic and prognostic significance." Cancer **92**(1): 136-145.
- Wu, P. C., V. K. Lau, et al. (1999). "Imbalance between cell proliferation and cellular DNA fragmentation in hepatocellular carcinoma." Liver **19**(5): 444-451.
- Xie, H., J. Song, et al. (2008). "The expression of hypoxia-inducible factor-1alpha in hepatitis B virus-related hepatocellular carcinoma: correlation with patients' prognosis and hepatitis B virus X protein." Dig Dis Sci **53**(12): 3225-3233.
- Xu, B., D. C. Hu, et al. (2003). "Chronic hepatitis B: a long-term retrospective cohort study of disease progression in Shanghai, China." Journal of gastroenterology and hepatology **18**(12): 1345-1352.
- Xu, C., T. Yamamoto, et al. (2007). "The liver of woodchucks chronically infected with the woodchuck hepatitis virus contains foci of virus core antigen-negative hepatocytes with both altered and normal morphology." Virology.
- Xuan, S. Y., Y. N. Xin, et al. (2007). "Significance of hepatitis B virus surface antigen, hepatitis C virus expression in hepatocellular carcinoma and pericarcinomatous tissues." World J Gastroenterol **13**(12): 1870-1874.
- Yachida, S., K. Imaida, et al. (2010). "Jun activation domain binding protein 1 is overexpressed from the very early stages of hepatocarcinogenesis." Ann Surg Oncol **17**(12): 3386-3393.
- Yaginuma, K., M. Kobayashi, et al. (1985). "Hepatitis B virus integration in hepatocellular carcinoma DNA: duplication of cellular flanking sequences at the integration site." Proc Natl Acad Sci U S A **82**(13): 4458-4462.

- Yamamoto, T., H. Nagano, et al. (2004). "Partial contribution of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)/TRAIL receptor pathway to antitumor effects of interferon-alpha/5-fluorouracil against Hepatocellular Carcinoma." Clin Cancer Res **10**(23): 7884-7895.
- Yang, A. S., M. R. Estecio, et al. (2004). "A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements." Nucleic Acids Res **32**(3): e38.
- Yang, J. C., C. F. Teng, et al. (2009). "Enhanced expression of vascular endothelial growth factor-A in ground glass hepatocytes and its implication in hepatitis B virus hepatocarcinogenesis." Hepatology **49**(6): 1962-1971.
- Yang, W. and J. Summers (1999). "Integration of hepadnavirus DNA in infected liver: evidence for a linear precursor." J Virol **73**(12): 9710-9717.
- Yao, D. F., Z. Z. Dong, et al. (2007). "Specific molecular markers in hepatocellular carcinoma." Hepatobiliary Pancreat Dis Int **6**(3): 241-247.
- Yasui, H., O. Hino, et al. (1992). "Clonal growth of hepatitis B virus-integrated hepatocytes in cirrhotic liver nodules." Cancer Res **52**(24): 6810-6814.
- Yeh, C. T., M. So, et al. (2010). "Hepatitis B virus-DNA level and basal core promoter A1762T/G1764A mutation in liver tissue independently predict postoperative survival in hepatocellular carcinoma." Hepatology **52**(6): 1922-1933.
- Yim, H. J. and A. S. Lok (2006). "Natural history of chronic hepatitis B virus infection: what we knew in 1981 and what we know in 2005." Hepatology **43**(2 Suppl 1): S173-181.
- Yoo, Y. G., T. Y. Na, et al. (2008). "Hepatitis B virus X protein induces the expression of MTA1 and HDAC1, which enhances hypoxia signaling in hepatocellular carcinoma cells." Oncogene **27**(24): 3405-3413.
- Yoshizumi, T., G. E. Gondolesi, et al. (2003). "A simple new formula to assess liver weight." Transplant Proc **35**(4): 1415-1420.
- Yuen, M. F. and C. L. Lai (2008). The Natural History of Chronic Hepatitis B. Hepatitis B virus. C. L. Lai and S. Locarnini. London, International Medical Press. **1**: 1-8.
- Yuen, M. F., H. J. Yuan, et al. (2005). "Prognostic determinants for chronic hepatitis B in Asians: therapeutic implications." Gut **54**(11): 1610-1614.
- Zajicek, G., N. Arber, et al. (1991). "Streaming liver. VIII: Cell production rates following partial hepatectomy." Liver **11**(6): 347-351.
- Zajicek, G., R. Oren, et al. (1985). "The streaming liver." Liver **5**(6): 293-300.
- Zhang, X., N. Dong, et al. (2005). "Effects of hepatitis B virus X protein on human telomerase reverse transcriptase expression and activity in hepatoma cells." J Lab Clin Med **145**(2): 98-104.
- Zhang, Y., R. Xue, et al. (2012). "Palmitic and linoleic acids induce ER stress and apoptosis in hepatoma cells." Lipids in health and disease **11**: 1.
- Zhou, Y., S. Wang, et al. (2010). "Hepatitis B virus protein X-induced expression of the CXC chemokine IP-10 is mediated through activation of NF-kappaB and increases migration of leukocytes." J Biol Chem **285**(16): 12159-12168.
- Zulehner, G., M. Mikula, et al. (2010). "Nuclear beta-catenin induces an early liver progenitor phenotype in hepatocellular carcinoma and promotes tumor recurrence." Am J Pathol **176**(1): 472-481.