Loss of ABO Antigens in Haematological Malignancies

Tina Bianco–Miotto
B.Sc., B.Hlth.Sc. (Hons.)

A thesis submitted in total fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Medicine

University of Adelaide
Adelaide, South Australia

May, 2002

Abstract

This thesis describes the investigation of the alteration of ABH antigen expression on the surface of red blood cells in patients with haematological malignancies. ABH antigens are complex carbohydrates present on the surface of erythrocytes. These antigens are generated by the stepwise addition of monosaccharides, with each step catalysed by a specific glycosyltransferase. The *A* and *B* alleles of the *ABO* locus encode the A and B glycosyltransferases, which act on the precursor H antigen to produce the A and B antigens. The precursor H antigen is determined by a fucosyltransferase coded for by the *FUT1* locus.

A flow cytometric technique was developed to analyse the red blood cells of patients for changes in ABH antigen expression. In normal individuals flow cytometric profiles were associated with particular *ABO* genotypes. Nearly 50% of patients had some alteration of the A, B or H antigens.

RT-PCR was used to investigate *ABO* mRNA expression in bone marrow (BM) cells of patients with alterations in ABH antigens. Of eleven patients with loss of A or B antigens detected by flow cytometry, only 3 had loss at the mRNA level. The lack of correlation between loss on the flow cytometer and loss at the mRNA level was not surprising as the population with loss was often small and the predominant normal population would mask any losses.

The expression of *FUT1* mRNA was also investigated using RT-PCR. *FUT1* has three transcripts arising from three different transcription initiation start sites. The individual expression of each transcript was assessed by RT-PCR in normal individuals and in patients with loss of ABH antigens. It was found that the exon 2 transcript is the most important in determining H antigen expression.

The molecular mechanisms underlying loss of ABH antigens were investigated by assessing CpG methylation of the *ABO* and *FUT1* promoters, and by assessing loss of heterozygosity (LOH) of genes adjacent to *ABO*. Treatment of leukaemia cell lines, which did not express *ABO*, with a demethylating agent led to re-expression of *ABO*. There was no association between patients with loss of ABH antigens and *ABO* methylation. However, of 7 patients with loss of ABH antigens detected in a previous PhD study, 5 were methylated in this study. This difference may be attributable to a greater majority of cells (>50%) with loss of ABH antigens in the samples detected previously compared to the samples in this study.

LOH of genes adjacent to *ABO* on the chromosomal region 9q34 was investigated. This chromosomal region has been shown to be frequently lost in acute myeloid leukaemia. Single nucleotide polymorphisms in various 9q34 loci, were assessed for LOH by a PCR-RFLP method. Samples from 42% of patients showed loss of expression in the tumour suppressor gene *DAPK1*. Two less well-known genes, *POMT1* and *PPP2R4*, had allele dosage alterations in patient samples.

In conclusion, loss of ABH antigens in haematological malignancies is a frequent event. Loss of cell surface ABH antigen expression is a marker for alterations at other 9q loci besides *ABO*, suggesting that this region harbours potential tumour suppressor genes or that 9q undergoes regional silencing.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, to be available for loan and photocopying.

Signed ..... Date 6/11/02

Acknowledgements

I thank my supervisor Alex for allowing me to choose such an interesting project. Thanks for tirelessly reading my thesis and thanks for all the faxes and express post mail.

I thank Damian Hussey, my co-supervisor for many years, for everything. For the moral support, for someone to bounce ideas and problems with, for all the technical help and for encouraging me to do all those experiments that I was never motivated to do. Thank you! Thanks for reading drafts of my thesis.

I am indebted to Belinda Farmer for developing the flow technique and then teaching me all the flow technicalities. Thank you to Terence Gooley for looking up blood groups of patients. Thanks to David Beeslee and Margaret Colbeck for patient diagnoses. A big thankyou to Leonie and all TQEH Blood Transfusion staff for blood groupings. Thanks to Emanuel and Glen for assistance on the flow cytometer. To David Peterson, thankyou for letting me invade TQEH Haematology lab so often. Thanks to the Haematology lab TQEH/IMVS for the use of the cell counter.

I thank past lab colleagues that assisted me. Nat Albanese, I owe you heaps of thanks for showing me the finer details of tissue culture work. Thanks to Michael Norman for helping me to do lab duties. To Tanya Sanders, what can I say. Only you understand the mental, physical and emotional suffering that we went through together to do all those cell line experiments, not once, not twice but three times! I think we always thought too big, way too big.

I thank current lab colleagues for all their help. Thanks to Sally Stephenson for reading drafts of my thesis and providing such detailed scrutiny, it was a great help. To Chris Slape for going through the writing agony at the same time as myself. It was nice to know someone else was suffering. To Michael Raynor, thanks for putting up with my singing in the lab, it kept me sane even though it may have driven you crazy (The Brady Bunch). To Rebecca Gully and Evelyn Douglas, fellow PhD students, thanks for all the laughs and chats. A special thanks to Sarah Swinburne for meticulously reading drafts of my thesis. It was much appreciated and always a good reason to catch up for breakfast.

Finally I must thank the people that have suffered the most: MY FAMILY. I think they must have thought I would never finish studying. My parents have always supported me, financially and through endless meals (and there were lots of meals). I thank my brother Lui, my sister-in law Pam and my little nieces Vanessa and Adriana, who were always my excuse to come home early to play with. I am sorry 'zia' Tina has been busy lately.

Last but not least I must thank my husband, Albino, whom I met at the beginning of my PhD so he has suffered all the frustrations of Science with me. Many a weekend he came to work with me to keep me company and even after we were married was constantly helping me around the house so I could concentrate on finishing my PhD. I'll make a scientist out of him yet! Alby, I am sincerely indebted to you and fortunate to have had such great support. This would not have been possible without your help, support and encouragement. Thanks for designing a program that calculated the Tms of my primers.

This thesis is dedicated to all my family and friends who have shared the ups and downs of this PhD with me. It would not have been achievable without all their support.

Publications

Publications arising directly from this thesis:

1. Dobrovic A, **Bianco T**, Tan LW and Sanders T. Rapid screening for methylation differences using methylation sensitive single strand conformation analysis. In Methods: A companion to Methods in Enzymology (Profiling genomic DNA methylation). *Methods*, 2002, **27**, 134-138.

2. **Bianco T**, Farmer BJ, Sage RE, Dobrovic A. Loss of red cell A, B, and H antigens is frequent in myeloid malignancies. *Blood*, 2001, **97**, 3633-3639.

3. **Bianco T**, Hussey D, Dobrovic A. Methylation-sensitive, single-strand conformation analysis (MS-SSCA): A rapid method to screen for and analyze methylation. *Human Mutation*, 1999, **14**, 289-293.

Other publications:

1. Tan LW, **Bianco T**, Dobrovic A. Variable promoter region CpG island methylation of the putative tumour suppressor gene Connexin 26 in breast cancer. *Carcinogenesis*, 2002, **23**, 231-236.

2. Butler LM, Dobrovic A, **Bianco T**, Cowled PA. Promoter region methylation does not account for the frequent loss of expression of the Fas gene in colorectal carcinoma. *British Journal of Cancer*, 2000, **82,** 131-135.

3. **Bianco T**, Chenevix-Trench G, Walsh DC, Cooper JE, Dobrovic A. Tumour-specific distribution of BRCA1 promoter region methylation supports a pathogenetic role in breast and ovarian cancer. *Carcinogenesis*, 2000, **21**, 147-151.

Conference Presentations

Epigenetic silencing of *ABO* and *FUT1* in haematological malignancies and leukaemia cell lines. **Bianco T**, Sanders T, Hussey DJ, Sage E, Dobrovic A. The Joint Annual Scientific Meeting of HSANZ and ASB, Brisbane, Australia, October 2001.

Screening of candidate chromosome 16 tumour suppressor genes in breast cancer by methylation analysis. Gully R, **Bianco T**, Dobrovic A. The Australian Society for Medical Research, South Australian Division, Annual Scientific Meeting, Adelaide, Australia, June 2001.

DNA methylation in breast cancer. Sanders T, **Bianco T**, Tan LW, Dobrovic A. The Australian Society for Medical Research, South Australian Division, Annual Scientific Meeting, Adelaide, Australia, June 2001.

DNA methylation in breast cancer. Sanders T, **Bianco T**, Tan LW, Dobrovic A. Breast cancer 2001 - Emerging possibilities. The Monash Medical Centre, Melbourne, Victoria, Australia. March 2001.

Screening of candidate chromosome 16 tumour suppressor genes in breast cancer by methylation analysis. Gully R, **Bianco T**, Dobrovic A. 13th Lorne Cancer Conference, Australia, February 2001.

Loss of ABO antigens in myeloid malignancies. **Bianco T**, Farmer B, Sage E, Dobrovic A. Queen Elizabeth Hospital Research Day, Adelaide, Australia, October 2000. Winner of the best presentation prize in the category of Higher Degrees (Basic Science).

Screening of candidate chromosome 16 tumour suppressor genes in breast cancer by methylation analysis. Gully R, **Bianco T**, Tan LW, Sanders T, Dobrovic A. Queen Elizabeth Hospital Research Day, Adelaide, Australia, October 2000.

Loss of ABO antigens in myeloid malignancies. **Bianco T**, Farmer B, Sage E, Dobrovic A. Annual Meeting of the Haematology Society of Australia, Perth, Australia, July 2000.

Loss of ABO antigens in myeloid malignancies. **Bianco T**, Farmer B, Sage E, Dobrovic A. The Australian Society for Medical Research, South Australian Division, Annual Scientific Meeting, Adelaide, Australia, June 2000.

Connexin-26 is methylated in a high proportion of breast cancer patients. Tan LW, **Bianco T**, Dobrovic A. The Australian Society for Medical Research, South Australian Division, Annual Scientific Meeting, Adelaide, Australia, June 2000.

Connexin-26 is methylated in a high proportion of breast cancer patients. Tan LW, **Bianco T**, Dobrovic A. Annual Meeting of the American Association of Cancer Research, San Francisco, USA, April 2000. Proc Amer. Assoc. Cancer Research 41:346 (2000).

Connexin-26 is methylated in a high proportion of breast cancer patients. Tan LW, **Bianco T**, Dobrovic A. 12^{h} Lorne Cancer Conference, Australia, February 2000.

DNA methylation in the BRCA1 promoter region. Dobrovic A, Tan LW, **Bianco T**. Gordon Conference "Epigenetics", Holderness School, New Hampshire, August 1999.

Abbreviations

ALL	Acute Lymphocytic Leukaemia
AML	Acute Myeloid Leukaemia
5-aza	5-aza-2'-deoxycytidine
BM	Bone Marrow
bp	base pair
cDNA	complementary DNA
CLL	Chronic Lymphocytic Leukaemia
CML	Chronic Myeloid Leukaemia
DEPC	Diethylpyrocarbonate
DIG	Digoxigenin
DNMT	DNA Methyltransferase
dNTP	deoxynucleotide triphosphates
DOVAM	Detection Of Virtually All Mutations
F	Forward
FITC	Fluorescein Isothiocyanate
Hx	Helix
HSC	Haematopoietic Stem Cell
LOH	Loss Of Heterozygosity
MDS	Myelodysplastic Syndrome
min	minute
MQW	MilliQ Water
mRNA	messenger RNA
PCR	Polymerase Chain Reaction
PB MNC	Peripheral Blood Mononuclear Cell
PBSC	Peripheral Blood Stem Cell
R	Reverse
RBC	Red Blood Cell
RFLP	Restriction Fragment Length Polymorphism
RSB	Routine Staining Buffer
RT	Reverse Transcription
sec	second
SNP	Single Nucleotide Polymorphism
SSB	Spherical Staining Buffer
SSCA	Single Stranded Conformational Analysis
TQEH	The Queen Elizabeth Hospital
U	Unit
UPW	Ultra Pure Water
UTR	Untranslated Region
UV	Ultra Violet
Ux	Ulex
V	Volt
VNTR	Variable Number Tandem Repeat

CHAPTER 1

Introduction

1.1 Overview

ABH antigens are complex carbohydrate structures present on the surface of red blood cells (erythrocytes), endothelial cells and most epithelial cells. ABH antigens may play a role in immunity but their actual function is unknown. In 1957, van Loghem *et al* reported that patients with haematological malignancies lose ABH antigens from their red blood cells. This thesis is a presentation of the experiments designed and performed to further investigate van Loghem *et al*'s observation, by testing the hypothesis that the loss of ABH antigens reflects genetic and epigenetic changes in the progenitor stem cell population that gives rise to the cells of the myeloid lineage during haematopoiesis. The investigation of the changes in these cells will assist in the identification of general mechanisms for loss of gene function in leukemogenesis.

Previously, changes in ABH antigens were detected using serological techniques that were not highly sensitive as it was found that a mixture of 50:50 A:O red blood cells, that is, effectively 50% loss of A, were classified as normal by blood transfusion staff (Denise O'Keefe, PhD thesis, Adelaide University, 1995). Therefore, the first aim of this PhD project was to use a quantitative assay (developed in collaboration with Ms. Belinda Farmer) to assess the frequency and amount of ABH antigen change on erythrocytes in patients with leukaemia (Chapter 3).

The second aim was to investigate the molecular mechanisms behind ABH changes, which may include genetic changes such as mutation and loss of heterozygosity (LOH), or epigenetic changes like methylation, for which we developed a new screening technique (Chapters 4 and 5).

Although ABH changes have been shown to be important in solid tumours and related to prognosis and metastatic ability (reviewed in Le Pendu *et al*, 2001), any potential association between ABH changes and leukaemia had not yet been investigated. ABH changes in leukaemia may reflect genetic or epigenetic changes at the chromosomal region where *ABO* is mapped and may indicate the presence of a tumour suppressor gene in this region. This is supported by reports of simultaneous decreases in the ABH antigens and the adenylate kinase 1 (AK1) enzyme, which is in the same chromosomal region as *ABO* (Salmon *et al*, 1968; Kahn *et al*, 1971; Marsden *et al*, 1992). Therefore, the third aim of this research project was to study genes on 9q34, where the *ABO* locus is located, with high rates of heterozygosity and to use single nucleotide polymorphisms (SNPs) to analyse these genes for LOH and mono-allelic expression as an indicator for tumour suppressor gene function (Chapter 6).

In the following introduction, a brief description of haematopoiesis, the normal process of blood cell formation that occurs in the bone marrow, will precede an outline of the various haematological malignancies that may arise when normal haematopoiesis is disrupted. Since ABH antigens are expressed on red blood cells, which are derived from the myeloid lineage, the discussion will focus on myeloid leukaemias. Finally, there will be an outline of methylation as an epigenetic mechanism for altering gene expression.

1.2 Haematopoiesis

Haematopoiesis is the process of the formation of mature blood cells from pluripotent stem cells in the bone marrow, the sponge-like tissue that fills up the cavities of bones, and produces red blood cells (erythrocytes), white blood cells (leukocytes) and platelets for release into the blood stream (Figure 1.1). Haematopoiesis is a carefully regulated system that needs to produce the correct number of each type of cell as approximately 3 million red blood cells and 120,000 white blood cells are made every second (reviewed in Ruscetti *et al*, 1998).

The production of blood cells requires the establishment, in early embryogenesis, of pluripotent cells that not only give rise to all of the different blood cells but are also capable of self-replication. These cells are called haematopoietic stem cells (HSCs) and represent only 0.01% of the bone marrow (reviewed in Ruscetti *et al*, 1998). HCSs have a high capacity for self-renewal and are also able to respond to extracellular signals to differentiate into myeloid or lymphoid progenitors (reviewed in Punzel and Ho, 2001). Committed progenitor cells are the immediate precursor of blasts, which are the earliest morphologically recognisable precursor cells of the two lineages. Progenitors of the myeloid lineage differentiate into red blood cells, platelets or the white blood cells that include the granulocytes and monocytes. The lymphoid lineage includes lymphocytes and natural killer cells (Figure 1.1).

1.3 Leukaemia

Leukaemia is a malignant disease of the haematopoietic system, which is characterised by uncontrolled proliferation or expansion of haematopoietic cells that retain little or no capacity to differentiate normally into mature blood cells (definition from Sawyers *et al*, 1991). Normally blasts (immature white blood cells) constitute 5% of the bone marrow but in leukaemia the blasts constitute between 30-100% of the bone marrow. These blasts gradually replace normal cells in the bone marrow and are eventually released into the circulation. As the number of immature cells increases, the number of fully differentiated cells decreases, contributing to the signs and symptoms of the disease including frequent infections and flu-like symptoms. This is because the abnormal blasts cannot help the body fight off infection. Pain and swelling results because the blasts tend to collect in the lymph nodes, organs and the joints, and anaemia is observed because the production of the blasts interferes with normal production of red blood cells and platelets.

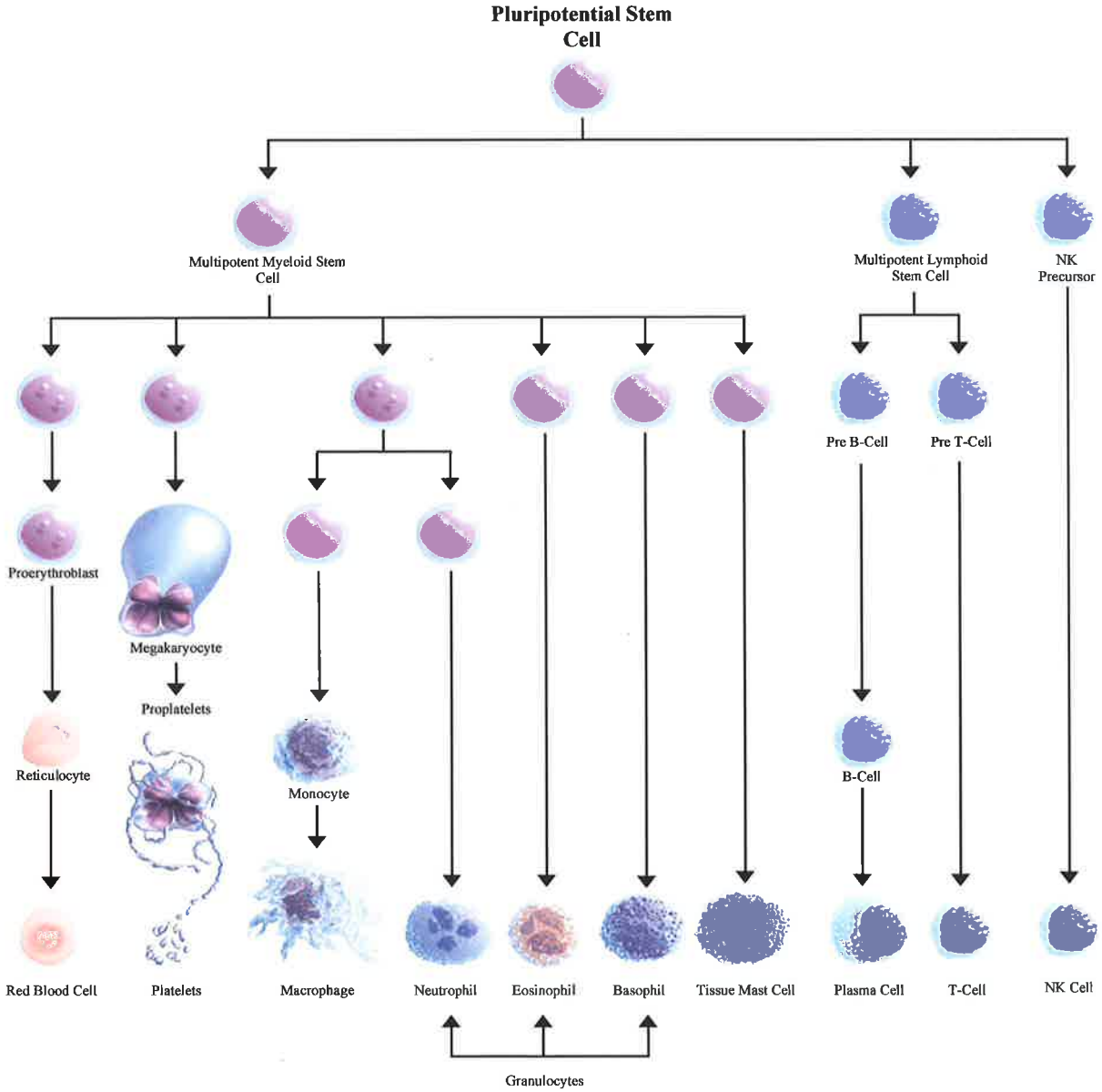


Figure 1.1. Haematopoiesis.

The diagram depicts normal blood cell formation in the bone marrow (adapted from promotional material from AMGEN). NK refers to natural killer.

1.3.1. Causes of Leukaemia

Like all cancers, leukaemia results from the accumulation of mutations in the DNA within one or more early progenitor cells (Ruscetti *et al*, 1998). Why it occurs in males more often than in females and in white people more often than in black people is not understood, but certain factors are known to increase the risk of developing the disease. The first of these is age. Approximately 60% to 70% of leukaemias occur in people who are older than 50. Environmental factors such as ionising radiation and chemical agents such as benzene, which can cause chromosomal abnormalities, are associated with an increased incidence of acute disease, especially myelogenous leukaemias.

Certain genetic conditions can also increase the risk of leukaemia. For example, children with Down's syndrome have a 15-fold increased risk of developing leukaemia when compared with other children. Patients with the genetically inherited disorders Fanconi's anaemia, Bloom's syndrome, and ataxia telangiectasia also have a greatly increased risk of leukaemia (Wetzler and Bloomfield, 1998). In addition, the incidence of leukaemia varies among racial and ethnic groups with different genetic make-ups. For example, rates of chronic lymphocytic leukaemia (CLL) are especially high in some Jewish populations, whereas Asian populations rarely develop CLL.

Certain viruses in some parts of the world have also been associated with contributing to leukaemia. The human T-cell leukaemia virus I (HTLV-I) is related to a rare acute T-cell leukaemia in parts of the Caribbean and Asia but is uncommon in the United States and Europe (Freedman and Nadler, 1998). An increase in viral-related leukaemia incidence has also been reported in butchers, slaughterhouse workers, and veterinary practitioners who are exposed to animal viruses.

Individuals who have received chemotherapy and radiation therapy for previous cancers have a greater chance of getting secondary leukaemia (leukaemia that arises after therapy). Acute myelogenous leukaemia (AML) has been reported after chemotherapy and/or radiotherapy for various solid tumours (breast cancer, ovarian cancer), blood malignancies, and non-malignant conditions. The risk of AML is increased when these chemotherapeutic drugs are combined with radiation therapy (Wetzler and Bloomfield, 1998).

1.3.2 Types of Leukaemia

Leukaemias can be classified as myeloid or lymphoid depending on which lineage of haematopoiesis is disrupted. When the cells that give rise to platelets, red blood cells, granulocytes and monocytes are affected then it is a myeloid leukaemia. When the leukaemia affects cells that give rise to lymphocytes it is a lymphoid leukaemia. The disease is then classified as acute or chronic.

Acute leukaemias are rapidly progressing diseases affecting the cells ability to differentiate and reach maturity giving rise to immature cells (blasts). These immature cells are incapable of adequately differentiating to the full spectrum of mature cells.

Chronic leukaemia progresses more slowly and occurs when more mature cells are affected. These more mature cells can carry out some of their normal functions and hence these patients do not require immediate treatment, if at all.

The four common types of leukaemia are acute lymphocytic leukaemia (ALL), chronic lymphocytic leukaemia (CLL), acute myeloid leukaemia (AML) and chronic myeloid leukaemia (CML).

1.3.2.1 Lymphoid Leukaemia

ALL is the most common childhood leukaemia responsible for approximately 75% of childhood leukaemias. However, ALL also occurs in adults, though more frequently in males. The abnormal cells in ALL are immature lymphocytes called lymphoblasts. In ALL patients there are too many lymphoblasts in the blood and bone marrow.

CLL is a slow developing disease that also affects lymphocytes and occurs commonly in elderly patients but not in children. In CLL there is an accumulation of lymphocytes that have a prolonged lifespan due to a delay in apoptosis. In CLL the lymphocytes do not mature correctly and though they appear normal they cannot fight infection as effectively as normal lymphocytes. CLL is the most common adult leukaemia in the western countries and is also more common in males (Freedman and Nadler, 1998). As this is a slowly progressing disease, some patients may never require treatment.

1.3.2.2 Myeloid Leukaemia

AML affects the myeloid lineage of cells that includes granulocytes, monocytes, red blood cells and platelets. AML is characterised by the accumulation of immature myeloid cells, which cannot carry out their normal function, and can block blood vessels. In AML there is also a halt in the production of normal marrow cells leading to a deficiency of red cells (anaemia), platelets (thrombocytopenia) and normal white cells in the blood (Wetzler and Bloomfield, 1998).

AML can be divided into seven subclasses, M1-M7, depending on the cell morphology observed in a bone marrow examination. M1 and M2 involve myeloblastic (undeveloped) cells except that in M1 the blasts are all undeveloped but in the M2 subclass some cells show a slightly more mature phenotype. M3 are promyelocytic leukaemias while M4 is

myelomonocytic. The M5 subclass involves monocytic cells while M6 involves the erythrocytes (red blood cells) and is called erythroleukaemia. The M7 subclass are megakaryocytic leukaemias (Wetzler and Bloomfield, 1998).

CML has three stages. The initial phase of CML is a chronic phase. This phase usually perseveres for a number of years and is followed by an accelerated phase in which the disease becomes more aggressive. The last phase, blast crisis, resembles acute leukaemia, and the median survival is only 3 to 6 months (Wetzler and Bloomfield, 1998).

CML is characterised by abnormal mature myeloid cells that retain their usual function in the chronic phase. CML is associated with a chromosomal abnormality in the leukaemic blasts, called a Philadelphia chromosome, which is a shortened chromosome 22, a product of a cytogenetic translocation t(9;22)(q34;q11) producing the fusion gene *BCR-ABL* (reviewed in Wetzler and Bloomfield, 1998). This cytogenetically distinct translocation is present in more than 95% of CML patients.

1.3.2.3 Myelodysplastic Syndromes

Myelodysplastic syndrome (MDS) is a disease of the bone marrow and blood. MDS is often called a 'pre-leukaemia', that is, a disorder that may progress to acute myelogenous leukaemia (AML) (Castro-Malaspina and O'Reilly, 1998). A MDS patient's bone marrow produces too few of one or more of the products of myeloid differentiation, commonly red blood cells, white blood cells and platelets. The bone marrow is characterised by cells displaying morphological abnormalities (Castro-Malaspina and O'Reilly, 1998). MDS has been shown to be a clonal disorder affecting haematopoietic stem cells and it has also been shown that MDS is a result of neoplastic transformation of HSC with involvement of myeloid cell lineages. The myeloid cells in MDS do not lose their capability to proliferate or

differentiate but they do not reach maturity, resulting in a lack of mature blood cells. The median age of onset of MDS is between 70 and 80 years, and primary MDS is more common in males (Castro-Malaspina and O'Reilly, 1998). The incidence of therapy related (secondary) MDS, especially in younger patients, is increasing as more tumours are treated with radiation therapy and chemotherapy. Many MDS patients often have a chromosomal abnormality, particularly of chromosomes 5 and 7, associated with the disease (Castro-Malaspina and O'Reilly, 1998).

1.4 The ABO Blood Group System

After the discovery of the circulation of blood in 1616 by William Harvey, transfusions of human or animal blood were often attempted but commonly fatal. In 1900, the reasons for these fatal transfusion reactions became clear when Karl Landsteiner discovered that the sera of some individuals agglutinated the red cells of others. One year later, Landsteiner classified individuals into three groups based on the two antigens, A and B, and two antibodies, anti-A and anti-B. In this classification an individual's red blood cells either had A antigens (group A), B antigens (group B) or neither A or B (group C, later renamed group O). Landsteiner found that an individual's serum does not contain the antibody for the antigen present on their own red cells but that both anti-A and anti-B are present in group O serum which has neither A or B antigens on the red cells. One year after Landsteiner's discovery, Decastello and Sturli, in 1902, discovered the fourth and rarest blood group, AB, which has both antigens on the red cells and neither antibody in the serum (reviewed in Watkins, 2001; see Table 1.1).

Blood Group Phenotype	Antigens Present on Red Blood Cells	Antibodies Present in Serum
A	A	Anti-B
B	B	Anti-A
AB	A and B	None
O	None	Anti-A and anti-B

Table 1.1. The antigens and antibodies associated with ABO blood groups.
(Adapted from Daniels, 1995).

1.4.1 The Structure and Biosynthesis of the A, B and H Antigens

The presence of ABH antigens on blood cells, endothelial cells and epithelial cells led them to be referred to as histo-blood group antigens. These histo-blood group antigens are carbohydrate structures. ABH antigens are oligosaccharide chains conjugated to polypeptides to form glycoproteins or with ceramide to form glycosphingolipids. The carbohydrate chains on glycoproteins are either highly branched *N*-glycans or simple *O*-glycans. ABH antigens are associated with both glycoproteins and glycolipids on human erythrocyte membranes. ABH antigens are determined by specific monosaccharides attached to various precursor disaccharides. There are six precursor disaccharides or peripheral core structures. Types 1-4 and 6 are found in humans while Type 5 has only been chemically synthesised (Figure 1.2).

Type 1	Galβ1→3GlcNAcβ1→R
Type 2	Galβ1→4GlcNAcβ1→R
Type 3	Galβ1→3GalNAcα1→R
Type 4	Galβ1→3GalNAcβ1→R
Type 5	Galβ1→3Galβ1→R
Type 6	Galβ1→4Glcβ1→R

Figure 1.2. Peripheral core structures of the ABH antigens.
Gal is D-galactose, GalNAc is *N*-acetyl-D-galactosamine, GlcNAc is *N*-acetyl-D-glucosamine, Glc is glucose, R is remainder of molecule (taken from Daniels, 1995). The numbers refer to which carbon is involved in the link.

The ABH antigens are generated in a stepwise fashion by addition of monosaccharides to the various peripheral core structures (Figure 1.3). Type 1 ABH structures are found in secretions, plasma and endodermally derived tissues. Type 1 structures are not made by red cells but are passively incorporated into the red cell membrane from the plasma. Type 2 chain antigens are the major ABH active oligosaccharides on red cells and are found in secretions and ectodermally and mesodermally derived tissues. Two Type 3 ABH antigens exist, one is not found on red blood cells but the other, repetitive Type 3 chains are present on red cell

glycolipids and secreted mucins from group A individuals only. Type 4 ABH antigens are located on glycolipids in small amounts on red blood cells but abundantly in kidney. Type 6 ABH antigens are found as free oligosaccharides in milk and urine (Figure 1.3).

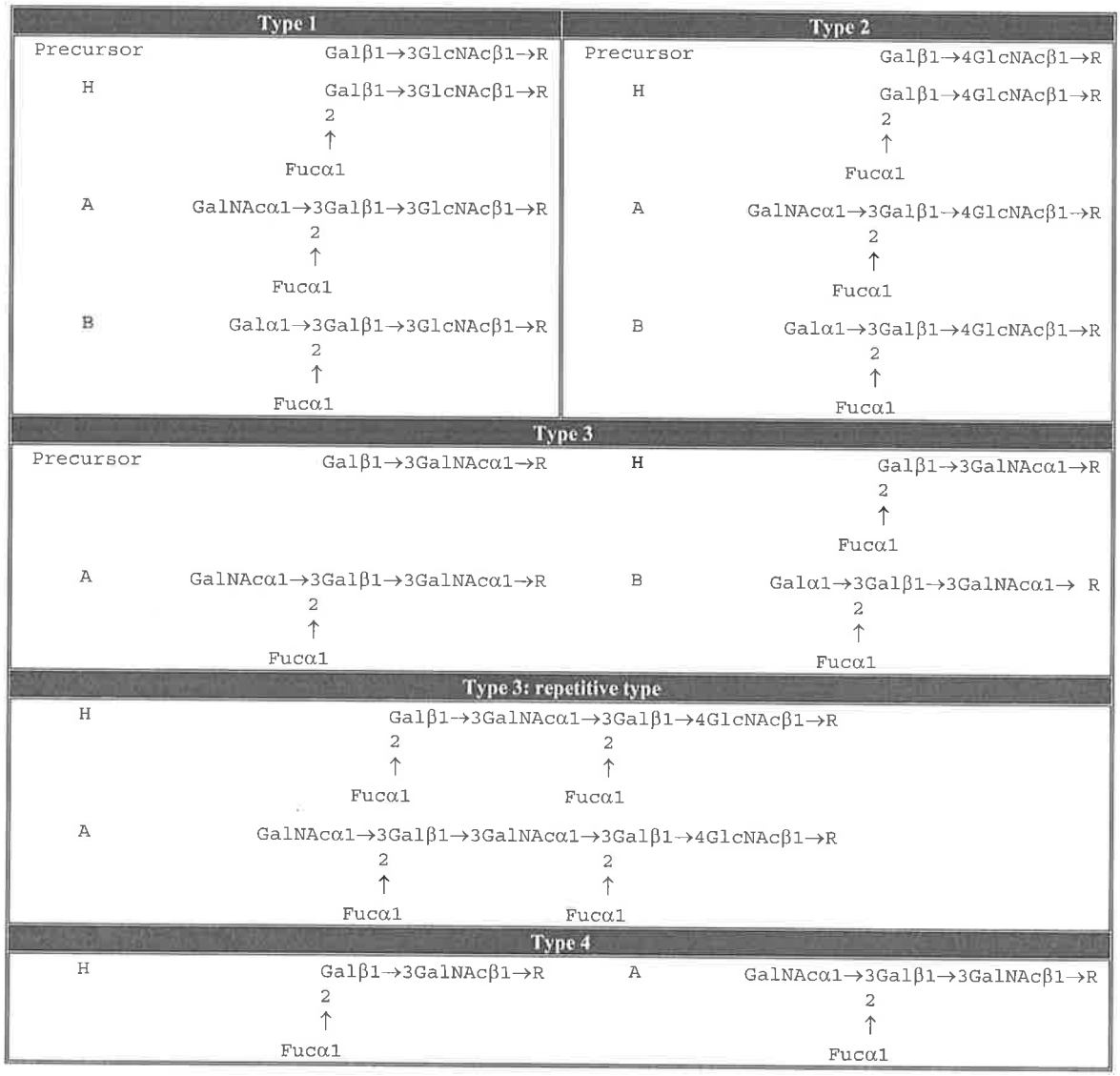


Figure 1.3. ABH antigens expressed on different precursor chains.

Abbreviations as per Figure 1.2. Fucα1 refers to the monosaccharide L-fucose.

The addition of each monosaccharide to the peripheral core structure is catalysed by a specific glycosyltransferase (Figure 1.4). The H antigens have L-fucose α-linked to the terminal galactose (Gal) while A and B antigens have *N*-acetyl-D-galactosamine and D-galactose respectively, α-linked to the fucosylated galactose (H) residue.

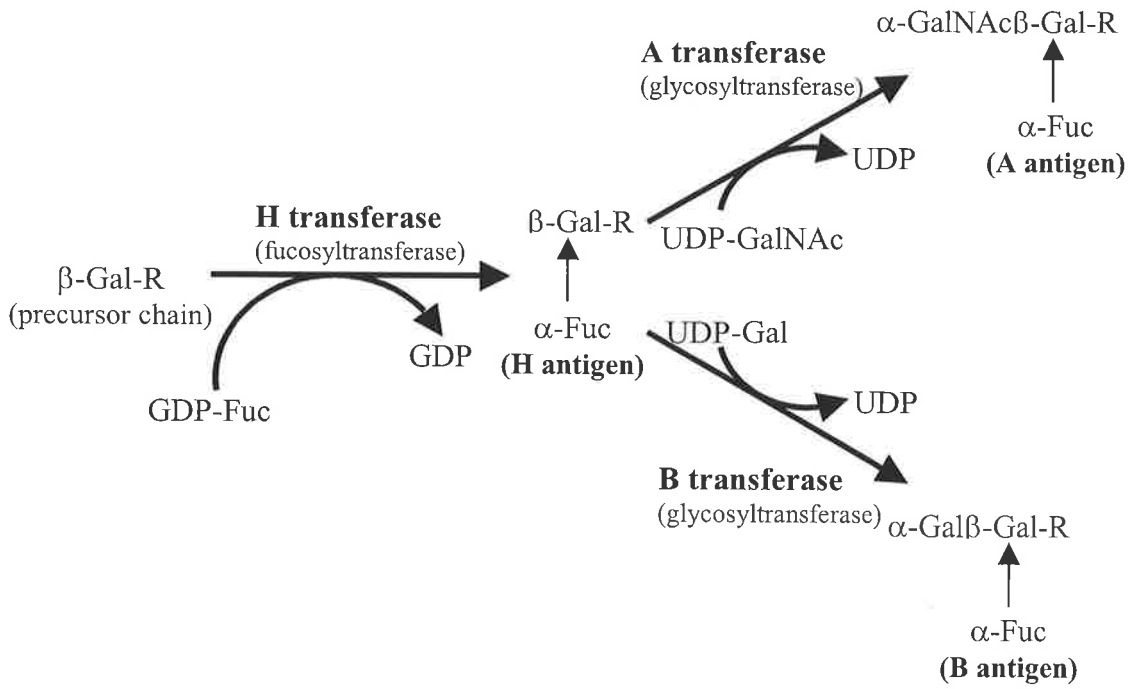


Figure 1.4. Formation of the ABH antigens by the action of glycosyltransferases.
Depiction of the conversion of the H antigen to A and B antigens. R is the core structure (taken from Olsson and Chester, 2001).

1.4.2 The Genes Encoding the Glycosyltransferases

In red blood cells, the H antigen is determined by a fucosyltransferase coded for by the *FUT1* gene. This has been cloned and mapped to the long arm of chromosome 19 at band 13.3 (Larsen *et al*, 1990; Reguigne-Arnould *et al*, 1995). The A and B glycosyltransferases that act on the precursor H antigen are encoded by separate alleles of the *ABO* gene (Yamamoto *et al*, 1990b; Yamamoto and Hakomori, 1990). The transferases encoded by the major alleles of *ABO* are *A* which adds *N*-acetylgalactosamine to the H antigen to give the A antigen, *B* which adds galactose to the H antigen to give the B antigen (see Figure 1.4) and *O* which is a null allele coding for an inactive transferase incapable of modifying the H antigen (Table 1.2). In addition to these three major alleles, several mutations within the *A* and *B* alleles result in the expression of glycosyltransferases with different activities (more detail in Section 1.4.5). For example, Schachter *et al* (1973) showed that both A^1 and A^2 alleles encode α-3-*N*-acetyl-D-galactosaminyltransferases but the two enzymes differ in their kinetic properties such that the

A^2 transferase is less effective than the A^1 transferase in converting H antigens into A antigens.

Red Cells		Serum	
Genotypes	**Phenotype**	**Antibodies**	**Glycosyltransferase(s)**
A^1A^1, A^1A^2, A^1O	A_1	Anti-B	α-3-*N*-acetylgalactosaminyl
A^2A^2, A^2O	A_2	Anti-B[a]	α-3-N-acetylgalactosaminyl
BB, BO	B	Anti-A	α-3-galactosyl
A^1B	A_1B	None	α-3-N-acetylgalactosaminyl, α-3-galactosyl
A^2B	A_2B	None[a]	α-3-N-acetylgalactosaminyl, α-3-galactosyl
OO	O	Anti-A & Anti-B	-

Table 1.2. The ABO blood group system.

[a]Anti-A_1 is sometimes present in the sera of A_2 and A_2B individuals (taken from Watkins, 2001).

1.4.3 The Bombay Phenotype and Secretor Status

The identification of individuals with no A, B or H antigens on their red blood cells or in their secretions but with anti-A, anti-B and anti-H in their serum (Bombay individuals, O_h phenotype; Bhende *et al*, 1952) supported the theory that the H antigen was an independent blood group system from *ABO*. Bombay individuals are homozygous for the inactive *h* allele of *FUT1* and are incapable of producing the H antigen that is the necessary precursor for the A and B transferases (Watkins and Morgan, 1955) (Table 1.3, Group 4).

The capability of an individual to secrete the ABH antigens is dependent on whether they are a secretor or non-secretor (Lehrs, 1930; Putkonen, 1930) and is determined by the genetic inheritance of the *Se* (secretor) or *se* (non-secretor) alleles in a Mendelian dominant fashion. Approximately eighty percent of Caucasians are secretors.

The discovery of the para-Bombay phenotype (*hh*, *Se*), individuals who lacked ABH antigens on their red cells but secreted these substances in saliva (Table 1.3, Group 3), led to a better understanding of the role of the *Se* gene. Oriol *et al* (1981) suggested that *Se* encodes a second α-2-fucosyltransferase that differs in its acceptor properties from the *FUT1* gene product in that it prefers Type 1 chains, whereas the *FUT1* gene shows no preference for any chain. The *Se* (*FUT2*) gene has been cloned (Kelly *et al*, 1995) and assigned to 19q13.3 where it is closely linked to *FUT1* (Reguigne-Arnould *et al*, 1995).

		Antigens Detectable on Red Cells	Activities Detectable in Secretions
Group	**Gene Combination**	**ABH**	**ABH**
1	*ABO, H, Se*	+++	+++
2	*ABO, H, sese*	+++	-
3	*ABO, hh, Se*	-	+++
4	*ABO, hh, sese*	-	-

Table 1.3. ABH antigens on red blood cells and in secretions.
The presence of ABH antigens in secretions is dependent on the *Se* locus. Group 3 represents the para-Bombay phenotype characterised by expression of ABH antigens in secretions but not on red blood cells while Group 4 represents the Bombay phenotype characterised by a lack of ABH antigens in secretions and on red blood cells.

1.4.4 Cloning and Regulation of the *ABO* Gene

The A and B glycosyltransferases catalyse the final steps in an enzymatic reaction where the A and B antigens are synthesised. A candidate protein for the A transferase was first isolated from the lungs of human blood group A individuals (Clausen *et al*, 1990). The protein was pure enough to allow partial amino acid sequencing, which assisted in the cloning of the cDNA of the A transferase. RNA from a blood group A gastric carcinoma cell line, MKN 45, was used to construct cDNA libraries (Yamamoto *et al*, 1990a). The amino acids from the

partial sequence obtained were reverse translated into nucleotide sequences and three degenerate oligodeoxynucleotides (oligos) were designed. DNA from individuals of groups A, B and O was amplified by PCR and underwent Southern hybridisation with a third internal oligo as a probe. Hybridisation was seen in all three blood groups and this fragment was purified and used to screen the MKN 45 and human lung cDNA libraries for full-length clones. Several positive clones were isolated and sequenced. Yamamoto *et al* then constructed two additional cDNA libraries from the human colon adenocarcinoma cell lines SW 48 (group AB) and SW 948 (group O) and were able to isolate the cDNAs corresponding to the *A*, *B* and *O* alleles and elucidate the nucleotide differences between the three alleles.

The molecular basis of the *ABO* polymorphisms was determined using the cloned transferase cDNAs. The *A* and *B* alleles encode 354 amino acid transferases with the difference in donor nucleotide-sugar specificity caused the amino acid substitutions at codons 176, 235, 266 and 268. Yamamoto and Hakomori (1990) tested the four polymorphic sites individually, or in combination, for their effect on nucleotide sugar substrate recognition. The entire coding region of the cDNAs of A and B transferases were cloned into eukaryotic expression vectors and transfected into HeLa cells (HeLa cells were derived from a human uterus adenocarcinoma that expresses H antigens on the cell surface and has the genotype *OO*). The cells were then analysed by FACS to confirm A and B antigen expression. After this, chimeric cDNAs that had all the possible combinations of the four relevant polymorphisms, at positions 176, 235, 266 and 268, were made. The amino acid at codon 176 did not affect transferase specificity while the amino acids at codons 266 or 268 were crucial for donor nucleotide-sugar specificity. When they were leucine and glycine respectively only the A antigen was expressed but when they were methionine and alanine only the B antigen was expressed. When these amino acids were methionine and glycine, enzymes with strong A and B transferase activity were generated, providing an explanation for expression of the rare *cis*-

AB phenotype (see 1.4.5.5). When they were leucine and alanine the result depended on the amino acid present at codon 235 (Yamamoto and McNeill, 1996).

Genomic cloning of the *ABO* locus and the determination of exon and intron boundaries was achieved by Bennett and colleagues (1995). This group confirmed the localisation of the *ABO* locus to 9q34. The gene is composed of seven exons with the last being the largest and codes for 225 amino acids (Figure 1.5).

1.4.5 Polymorphisms of the *ABO* Gene

1.4.5.1 *A* Alleles

In addition to the consensus *A* allele, another major A^1 allele common in Asian populations has a C467T substitution that results in the amino acid change Pro156Leu (Fukumori *et al*, 1996; Ogasawara *et al*, 1996; Kang *et al*, 1997; Yip, 2000). This additional substitution does not affect the enzyme activity or A^1 antigen synthesis, as shown by expression constructs (Yamamoto and Hakomori, 1990).

Two additional A^1 alleles have also been identified. The first has the C467T substitution plus an additional silent substitution C564T. The second A^1 allele differs from the consensus by a silent mutation A297G and though this does not affect enzyme activity, this mutation is also present in normal B alleles and thus if this site is used for genotyping, errors would occur.

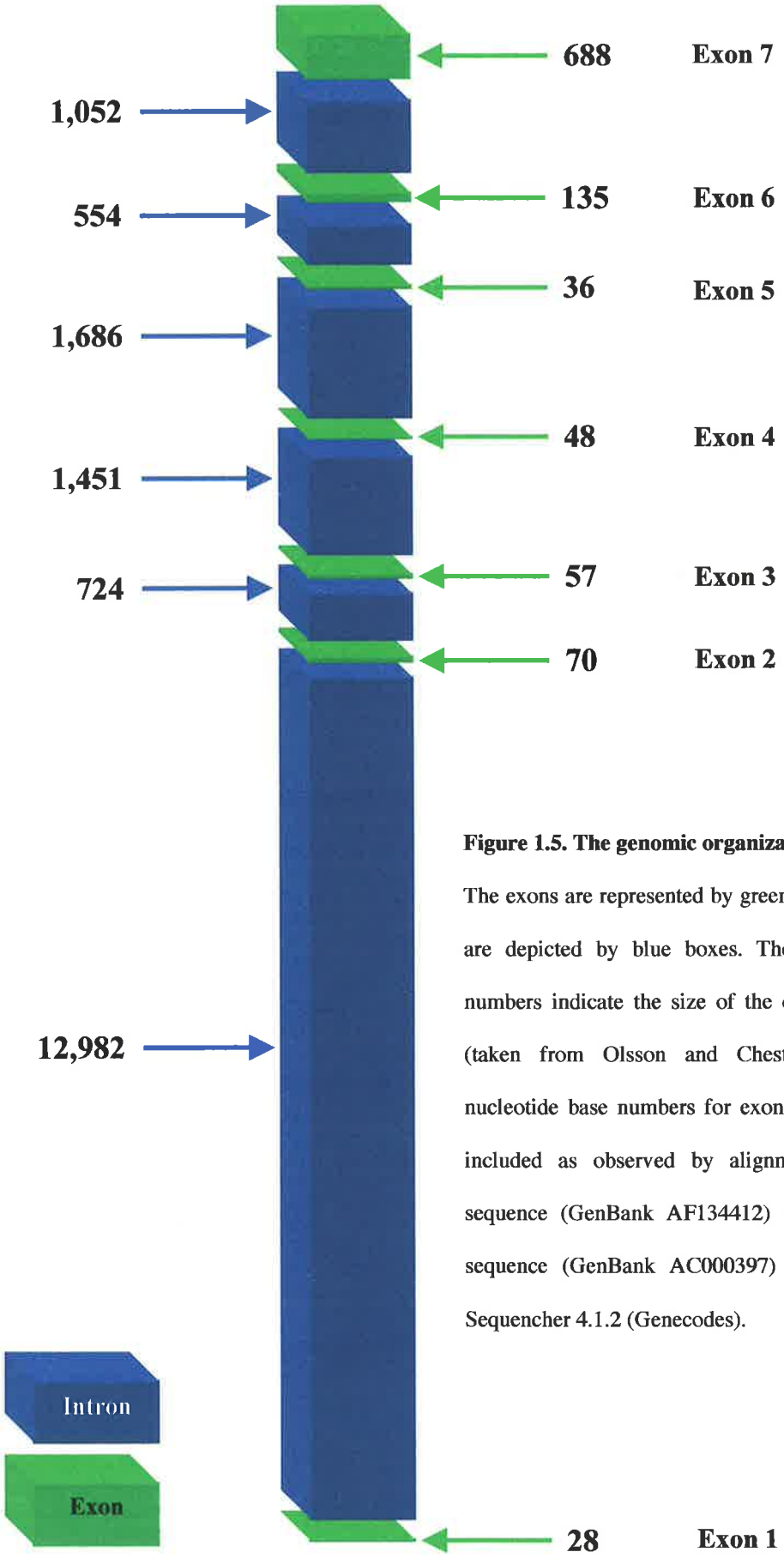


Figure 1.5. The genomic organization of the *ABO* gene. The exons are represented by green boxes while the introns are depicted by blue boxes. The arrows with adjacent numbers indicate the size of the exons and introns in bp (taken from Olsson and Chester, 2001). The exact nucleotide base numbers for exons and introns have been included as observed by alignment of the A^1 cDNA sequence (GenBank AF134412) and the genomic DNA sequence (GenBank AC000397) by using the program Sequencher 4.1.2 (Genecodes).

The A^2 phenotype is common in the Caucasian population, with 20% of individuals having this allele, but is rare in the Japanese representing less than 1% of individuals in the population (Mourant *et al*, 1976). The A^2 transferase has a weaker activity compared to the A^1 transferase. Changes in the A^1 consensus sequence that leads to the A^2 allele are the C467T (Pro156Leu) substitution and a single C deletion in a sequence of three consecutive cytosines at positions 1059-1061. The deleted cytosine causes a reading frame shift after codon 354 such that a protein with 21 extra amino acids is made (Yamamoto *et al*, 1992). This longer transferase is functional but has different activity strength and substrate specificity than the A^1 allele. Introduction of this single-base deletion into the A^1 transferase cDNA expression construct drastically decreased A transferase activity in DNA-transfected HeLa cells, whereas the C to T nucleotide substitution had no effect on A transferase activity (Yamamoto *et al*, 1992). No other mutation besides the deleted cytosine has been detected that creates an A^2 allele (see Table 1.4).

1.4.5.2 *B* Alleles

There are seven nucleotide differences between the A and B enzymes (Yamamoto *et al*, 1990b), which results in four amino acid changes. An eighth substitution (G1096A) that occurs beyond the stop codon has been utilised for *ABO* genotyping (Olsson and Chester, 1995). Three other normal *B* alleles have been identified in the Japanese population, each lacking one of the substitutions differentiating A^1 from *B* alleles (Ogasawara *et al*, 1996).

1.4.5.3 *O* Alleles

Theoretically, an *O* allele can be any allele that has a mutation resulting in the loss of the translated proteins ability to transfer *N*-acetylgalactosamine or galactose to an oligosaccharide acceptor. The original *O* allele, termed the O^1 allele and considered the consensus allele, has a single deleted G in exon 6 at position 261. This deletion causes a frame shift that introduces a

premature stop codon that alters the protein after amino acid 88 and terminates the protein chain after amino acid 117. The resulting enzyme has no activity. When the G deletion was introduced into an A^1O^1 expression construct, A transferase activity was abolished (Yamamoto, 1995). A second *O* allele was originally detected in the cell line COLO205, which has the deleted G but also has other point mutations located throughout exons 3 to 7. This allele has been called $O^{1variant}$ (O^{1v}) (Yamamoto *et al*, 1990b).

The first *O* allele (O^2) to be discovered, which did not contain the deleted G, had a new mutation G802A, which caused the amino acid change Gly268Arg (Yamamoto *et al*, 1993d). This allele had two additional mutations, which are also found in the *B* alleles (A297G and C526G). The glycine to arginine change leads to substantial loss of enzymatic activity of this allele (Yamamoto and McNeill, 1996; Amado *et al*, 2000) (see Table 1.4).

Rare *O* alleles that do not contain the deleted G include O^3, which has the common A^2 deletion (C1061-) plus the A^{el} allele insertion (an extra guanosine in a seven guanosine stretch at nucleotides 798-804) (Olsson *et al*, 1995; Olsson and Chester, 1996). The double frame shift is hypothesised to translate a normal length protein but the insertion and subsequent deletion changes the amino acid sequence between amino acids 269–353.

The allele O^4 has an inserted G at nucleotides 87-88, which results in a frame shift and a stop codon at codon 56, and the allele O^5 has a C322T mutation creating a stop codon directly (Hosseini-Maaf *et al*, unpublished observation).

Allele	Symbol in Text	Nucleotide Change	Amino Acid Change
A101	A^1	Standard Reference	Standard Reference
A102	Common Asian A^1 allele	467C>T	P156L
A201	A^2	467C>T; 1060delC	P156L;354FS +21aa
A301	A^3	871G>A	D291N
Ax01	A^x	646T>A	F216I
B101	*B*	297A>G; 526C>G; 657C>T; 703G>A; 796C>A; 803G>C; 930G>A	R176G; G235S; L266M; G268A
B301	B^3	297A>G; 526C>G; 657C>T; 703G>A; 796C>A; 803G>C; 930G>A; 1054C>T	R176G; G235S; L266M; G268A; R352W
cis-AB01	*cis-AB*	467C>T; 803G>C	P156L; G268A
B(A)01	*B(A)*	297A>G; 526C>G; 796C>A; 803G>C; 930G>A	R176G; L266M; G268A
O01	O^1	261delG	88FS + truncation
O02	$O^{1variant}$	106G>T; 188G>A; 189C>T; 220C>T; 261delG; 297A>G; 646T>A; 681G>A; 771C>T; 829G>A;	V36F; P74S; 88FS + truncation
O03	O^2	297A>G; 526C >G; 802G>A	R176G; G268R

Table 1.4. Various *ABO* alleles and their associated nucleotide differences.
Adapted from Yamamoto (2001). The abbreviations 'del' indicates deletion, and 'FS' denotes frameshift.

1.4.5.4 Weak *A* or *B* Alleles

In addition to the major A^1 and A^2 subgroups, there are many variants in which the expression of A is reduced. Weak A subgroups include A^3, A^x, A^m, and A^{el} (reviewed in Yamamoto, 1995). A^3 cells demonstrate a characteristic mixed-field pattern of agglutination (small agglutinated clumps with a majority of unagglutinated clumps) with anti-A and anti-A and B sera. The amount of A transferase in individuals with the A^3 phenotype ranges from 50% of the control to low or no detectable A transferase activity. A^x red cells are not agglutinated by

anti-A, but weak agglutination is observed with anti-A and B. A very weak A transferase activity has been detected in the serum of this subgroup (reviewed in Yamamoto, 1995). The red cells of A^x and A^m individuals are frequently first grouped as O but the presence of an A antigen is suspected because anti-A is missing from the serum. The cells can be shown to have the capacity of absorbing human anti-A even though they are not agglutinated by this reagent (reviewed in Watkins, 1980).

Weak B subgroups include B^3, B^x, B^m and B^{el} (reviewed in Yamamoto, 1995) but group B antigenic expression cannot be classified into well-defined subgroups corresponding to A^1, A^2, and A^3. A number of reports of weak B variants behaving in a similar manner to A^x and A^m have been published (reviewed in Watkins, 1980). The B^3 phenotype is characterised by a mixed-field pattern agglutination with anti-B and anti-A and B sera. B transferase activity is present in the serum of these individuals however it varies from 10 to 100 percent of the control levels (reviewed in Pitiglio, 1986).

Yamamoto and colleagues (1993) studied the nucleotide substitutions involved in some of the rare alleles of *ABO*. Primers were used to amplify genomic DNA fragments spanning the last two coding exons of *ABO,* which accounts for 90% of the coding sequence of the A and B transferases. The partial nucleotide sequence of one A^x, four A^3 and three B^3 individuals was examined (Yamamoto *et al*, 1993a; Yamamoto *et al*, 1993c). Two cases of the A^3 allele, one case of the B^3 allele and the A^x allele contained single-base substitutions that resulted in an amino acid change. Two of the A^3 alleles had a G to A substitution at cDNA position 871 (amino acid change at residue 291) but the other two A^3 alleles had no substitutions in exon 6 or 7. The fact that the substitution was not found in all the cases of A^3 demonstrated the heterogeneity among these weak subgroups (reviewed in Yamamoto, 1995). However, the coding sequence outside of the sequenced region and the promoter region remains to be

investigated. One of three B^3 alleles had a C to T substitution at 1054 (amino acid change at 352) but not the other two B^3 alleles. The A^x allele contained a T to A substitution at cDNA nucleotide position 646.

Previously, only eight A and six B subgroup alleles were defined by missense mutations in exon 7, however several had no deviations from the exon 6 and 7 consensus sequence (Yamamoto *et al*, 1993 a, b, c, d; Olsson *et al*, 1995; Ogasawara *et al*, 1996; Olsson and Chester, 1996; Olsson and Chester, 1998; Hansen *et al*, 1998; Barjas-Castro *et al*, 2000). Exons 1 to 5 in these rare ABO phenotypes remained to be investigated. Olsson and colleagues (2001) determined the *ABO* genotype of 324 blood samples that had ABO grouping discrepancies by sequencing exons 1 to 7 and two possible 5' untranslated regulatory regions.

Genotyping of fifty three samples with suspected weak A subgroups showed heterozygosity for previously identified variant alleles including A^{el}, *B(A)* and different types of A^x. Ninety-four samples from healthy donors had weak A or B phenotypes and *ABO* genotyping excluded previously identified weak alleles. Exons 1-7 and two regulatory motifs of *ABO* were sequenced in 22 of the weak A individuals and 12 of the weak B individuals. In doing so, a further 15 novel *A* and *B* alleles were discovered (Olsson *et al*, 2001) and together with the 14 previously discovered weak subgroup *ABO* alleles brings the total to 29 *ABO* alleles associated with weak subgroups.

The 324 clinical samples with ABO grouping discrepancies were divided into groups with the first being acquired variant ABO phenotypes, which included pregnancy, haematological disorders and a variety of other disorders. Pregnancy was responsible for 13 cases of variable weak A expression and retesting of five available women showed that their A antigen

expression went back to normal after the delivery of the baby (this observation was also noted by O'Keefe, PhD thesis, 1995). All 13 samples were heterozygous for the *A* allele. Weak, but variable A (and in 2 cases B) antigen expression was noted in 16 non-transfused patients with various haematological malignancies including acute and chronic myelogenous leukaemia, multiple myeloma, non-Hodgkins lymphoma and myelodysplastic syndrome. *ABO* genotyping confirmed their non-O blood group status hence transfusion with O blood was avoided. Acquired B was seen in nineteen other patients with non-haematological malignancies including adenocarcinomas of the colon, rectum, uterus or brain and others with gastrointestinal infections. *ABO* genotyping showed that they did not have *B* alleles but had either A^1 or A^2 alleles. Thirteen patients with other medical conditions had weak A expression and genotyping showed they had *A* alleles.

1.4.5.5 The *Cis-AB* and *B(A)* Alleles

The *cis*-AB phenotype was first described in 1964 by Seyfried and her colleagues who found a family with an unusual inheritance pattern of the ABO blood groups. The blood group of the father was O and the blood group of the mother was A^2B but the blood group of their two children was A^2B. The following year Yamaguchi *et al* reported a similar observation with the inheritance of the A^2B^3 phenotype in a family where it was suggested that both the A^2 and B^3 alleles were inherited simultaneously on one chromosome. To discriminate this rare phenotype from ordinary AB phenotypes, the authors named it *cis*-AB (reviewed in Yamamoto, 1995). *Cis-AB* alleles make enzymes that have A and B activity and sequencing of Japanese *cis-AB* alleles identified the presence of three out of four A specific nucleotides in exon 7 and only one B substitution. An Australian *cis-AB* allele was identified that had three B substitutions and only one A substitution (Mifsud *et al*, 2000). In nearly all instances of the *cis*-AB phenotype the B antigen on the red cell was atypical in its behaviour and the serological behaviour of the A antigen on the red cells more closely resembled that of normal A^2 cells than A^1 cells.

In individuals with the B(A) phenotype, B antigens are predominant on red blood cells, but these cells can also react with anti-A reagents. Sequencing of these alleles shows a BABB pattern at the four amino acid substitutions responsible for the differences between the *A* and *B* alleles. A *B(A)* allele has been identified that has the sequence BBBB, which appears to be a normal *B* allele except that there is a mutation responsible for an amino acid substitution adjacent to the second B which causes A glycosyltransferase activity.

1.4.5.6 Hybrid Alleles

Numerous alleles have been identified that can be explained by crossing-over between two different alleles. This phenomenon was first described in 1996 when a blood sample indicated homozygosity for the O^1 allele at position 261 and heterozygosity at position 1061 for an A^2 allele (Gassner *et al*, 1996). Numerous hybrid alleles have been identified (reviewed in Olsson and Chester, 2001) including a sequence with four characteristic *B* substitutions and two O^{1v} nucleotide substitutions. A hybrid allele caused by the crossing over of *B* and O^1 alleles was capable of coding for an A^1 transferase (for a more detailed review on hybrid alleles of *ABO* see Olsson and Chester, 2001). Many other hybrid alleles exist suggesting that the *ABO* locus is a hotspot for recombination (Olsson and Chester, 2001).

1.4.5.7 Allelic Variation in Untranslated Regions of the *ABO* Gene

Luciferase reporter constructs, which carried various portions of the 5' upstream region of the translation initiation codon sequence of *ABO*, determined that deletion of the upstream region from –3950 to –666 (+1 being the transcription initiation site) resulted in a marked decrease in luciferase activity (Kominato *et al*, 1997). This indicated the presence of an enhancer element in this region. Further deletions closer to the transcription initiation region (to either –202 or –117) led to an increase in promoter activity indicating the presence of a negative regulator in this region.

Further examination of the enhancer region identified a minisatellite sequence composed of four 43 bp repeats (in the region –3950 to –3252). The enhancer activity of this sequence was proven using further luciferase constructs that contained the repeat sequence downstream of the luciferase gene. Electrophoretic mobility shift assays (EMSA) with the repeat sequence showed that there was a near-consensus binding site for the transcription factor CBF/NF-Y on the 5' half of the 43 bp repeat unit. Binding of CBF/NF-Y to the repeat sequence was examined by a supershift assay using a CBF/NF-Y antibody and shown to be required for *ABO* transcription (Kominato *et al*, 1997).

Although initial identification of the minisatellite identified four copies of the 43 bp sequence, this enhancer was later shown to be a variable number of tandem repeat (VNTR) region with four copies associated with the A^2, *B*, O^1 and O^{1v} alleles but only a single 43 bp motif associated with the A^1 and O^2 alleles (Figure 1.6) (Irshaid *et al*, 1999; Yu *et al*, 2000). In total, three different enhancer structures existed for the *ABO* alleles. The first, found in the *B* allele and 35% of the O^1 alleles consists of 4 repeat units. The second found in O^{1v} and 65% of the O^1 alleles had 4 repeats plus a G to C substitution at the 41st nucleotide of the 43 bp repeat unit in the first of the four repeats (Yu *et al*, 2000). The third enhancer region is found in A^1 alleles and had only one 43 bp unit with a G to A substitution at the 41st nucleotide of the 43 bp repeat unit (Yu *et al*, 2000).

Yu *et al* (2000) investigated whether there were differences in the transcriptional activities of the *A* and *B* allele associated enhancers. The promoter of *ABO* (-118 to +32 bp) was inserted into a construct with no promoter and was shown to induce a 10-fold increase in transcription. The addition of the *A* allele enhancer to the *ABO* promoter sequence increased transcription 5-fold. Addition of the *B* allele enhancer to the *ABO* promoter increased transcription 300-fold more (Yu *et al*, 2000).

Another interesting set of experiments carried out by Yu and colleagues (2000) was the insertion of one *A* allele enhancer unit, and 1 to 4 units of the *B* allele enhancer. Transcription increased as the number of *B* enhancer units was increased from one to four. However, a 20% increase with the one *B* allele enhancer unit was seen when compared to only one *A* allele enhancer unit, even though they only differ by one base pair (Yu *et al*, 2000). The A^1 transferase is more efficient than the A^2 or B transferase despite having less enhancer activity, suggesting that the mutations found in the A^2 and *B* alleles are more detrimental to the transferase function than first thought (Olsson and Chester, 2001).

1.4.5.8 Allelic Variations in the *ABO* Intron Sequence

Additional allelic variations in *ABO* have also been found in intron 6 (Suzuki *et al*, 1997; Olsson and Chester, 1998). The original sequence of *ABO* (GenBank AC000397) indicated that intron 6 was 1052 bp in length. Intron 6 from A^1, *B*, O^1 and O^{1v} alleles was sequenced and described as being 1049 bp with identical A^1 and O^1 sequences but different *B* and O^{1v} sequences (Suzuki *et al*, 1997). The *B* allele differed at 12 positions, while the O^{1v} allele differed at 13 positions, of which 6 were shared between the *B* and O^{1v} alleles. Other investigations of intron 6 by Olsson and Chester (1998) (GenBank AF016622-AF016625) for B/O^2, O^{1v} and two A^x hybrids indicated a 1052 bp intron 6 sequence. This discrepancy between intron 6 sequences may be due to sequencing errors or ethnic variation (Olsson and Chester, 2001).

Olsson and colleagues (2000) were the first to identify an intronic mutation that led to a phenotypic change in the ABO system. A mutation at nucleotide 4 of intron 6 was associated with the A_{finn} phenotype in four Finnish blood donors. The authors suggested that this mutation affected splicing so that the majority of transcripts lack exon 6. A premature stop codon at the junctions between exons 5 and 7 would abolish enzyme activity in the truncated protein product.

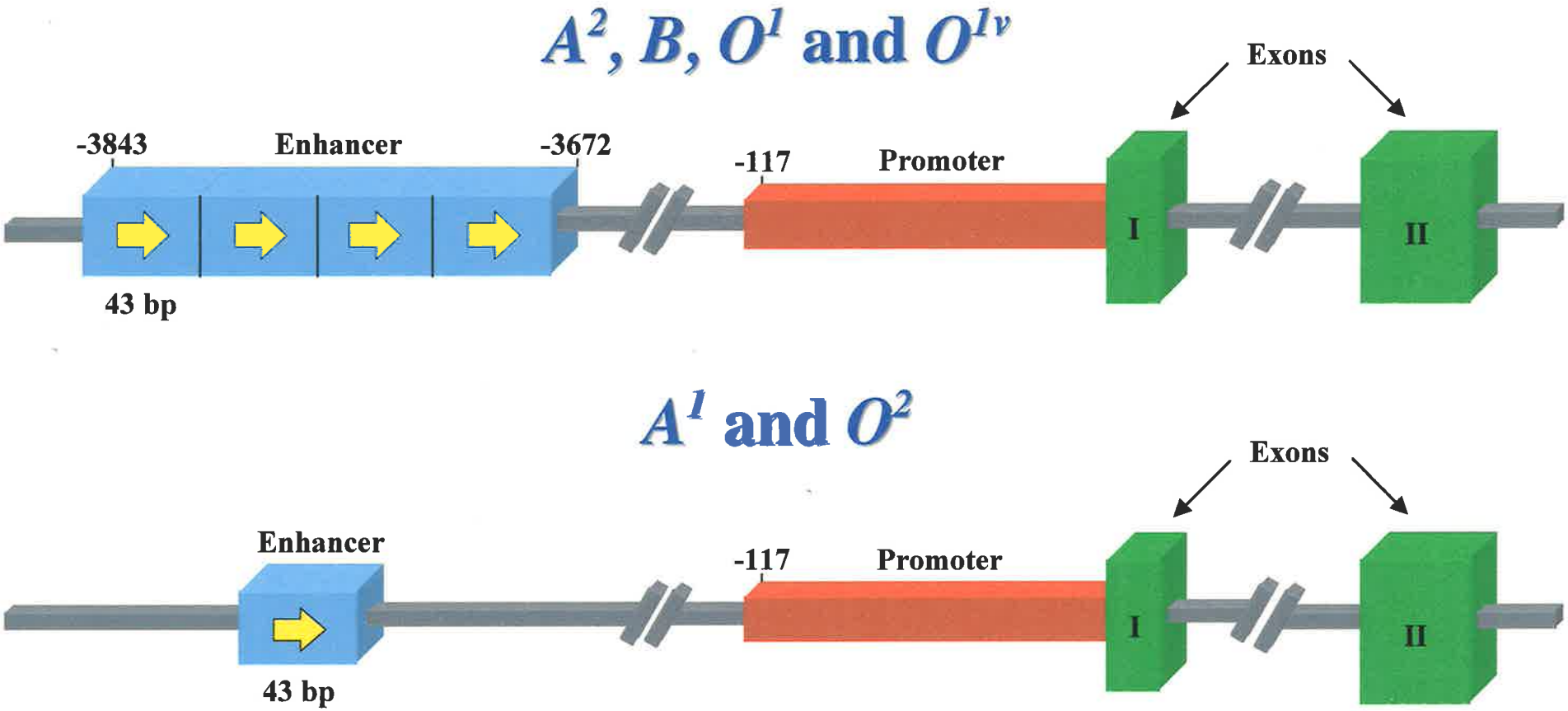


Figure 1.6. The VNTR enhancer associated with various *ABO* alleles.

The different enhancer structures associated with various *ABO* alleles are shown. One yellow arrow represents one repeat unit.

(Adapted from Yu *et al*, 2000).

1.4.5.9 Utilising the Allele Polymorphic Differences for Genotyping

The different nucleotide substitutions between the three major *ABO* alleles create allele specific restriction enzyme sites. Yamamoto and colleagues (1990b) discovered the differences between the common *ABO* alleles at the DNA level and were the first to propose a method for discriminating the *A*, *B* and O^1 alleles. The deleted G in the O^1 allele created a *Kpn*I site while the *A* and *B* alleles had a *BstE*II site at the corresponding position (Yamamoto *et al*, 1990b). Southern hybridisation after restriction enzyme digestion of DNA with *Kpn*I and *BstE*II was used to detect the O^1-specific deletion. Utilisation of a PCR-RFLP method allowed the discrimination of the *A* and *B* alleles. The PCR-RFLP method has been modified and expanded by many and is the easiest and most efficient way of genotyping for *ABO* (Yamamoto *et al*, 1990b; Yamamoto and Hakomori, 1990; Lee and Chang, 1992; O'Keefe and Dobrovic, 1993; Olsson and Chester, 1995; O'Keefe and Dobrovic, 1996). Other novel techniques for *ABO* genotyping have been developed and will be discussed in more detail in Chapter 4.

1.4.6 ABH Antigen Expression in Adult Tissues and Embryonic Development

ABH antigens are not restricted to red cells as they are found as cell surface antigens on endothelial cells of blood vessels, in glands and goblet cells, and in various organs of excretion and secretion (reviewed in Watkins, 2001). At approximately five weeks after fertilisation strong cell surface expression of A, B and H antigens can be detected on the endothelium and epithelium of most early organs but in epithelia the cell surface expression diminishes after approximately nine weeks when secretion of mucin begins. A systematic disappearance of the ABH antigens has been observed during the development of embryonic epithelial organs with the antigens reappearing after the developing organs have reached their final differentiation states (Szulman, 1964). Muscle and bone cells completely lose the ability to make ABH antigens.

1.4.7 Alterations of ABH Antigen Expression in Solid Tumours

In malignancy, the A and B antigens may reappear in tissues from which they disappeared early in development, hence the term onco-developmental antigens. An example is the distal colon in which A and B expression is not present in normal adult tissue but there is re-expression of these antigens in carcinoma of the colon (Piller *et al*, 1979; Orntoft *et al*, 1991).

Continued expression of A and B antigens in tumours derived from tissues that normally do not express ABH antigens has been shown to be a favourable prognostic factor (Lee *et al*, 1991; Matsumoto *et al*, 1993), whereas loss in tumours that normally express ABH antigens is linked with a poor prognosis and is associated with the degree of metastasis or malignancy (Davidsohn and Stejskal, 1972; Ichikawa *et al*, 1998). No essential function for tumour suppression has been attributed to the products of the *ABO* locus. The inhibitory role of the antigens in malignancy fits in with Koscielak's hypothesis that the ABH antigens are capping structures that cover up oligosaccharide chain endings that otherwise would interact with endogenous lectins or cell surfaces (Koscielak, 1986, 2001).

Loss of ABH antigens from tumour tissue is frequently seen in carcinomas (Table 1.5) of the buccal epithelium, stomach, proximal colon, pancreas, larynx, lung, endometrium, ovary, prostate, bladder, thyroid, head and neck, and breast (Lange *et al*, 1978; Hirohashi *et al*, 1984; Coon *et al*, 1985; Itzkowitz *et al*, 1987; Schoentag *et al*, 1987; Dabelsteen *et al*, 1988; Orntoft *et al*, 1989; Memon *et al*, 1990; Perlman *et al*, 1990; Inai *et al*, 1992; Stenersen *et al*, 1992; Welshinger *et al*, 1996; Nakagoe *et al*, 1998; Moldvay *et al*, 2000). Loss of ABH antigen is associated with poor prognosis, tumour grade and metastatic potential (Lange *et al*, 1978; Limas *et al*, 1979; Ernst *et al*, 1984; Lindgren *et al*, 1986; Wolf *et al*, 1990; Bryne *et al*, 1991; Lee *et al*, 1991; Matsumoto *et al*, 1993). Tumour cells with decreased expression of A/B antigens were shown to have a higher metastatic tendency (Miyake *et al*, 1992). The survival

of blood group A or AB non-small cell lung carcinoma patients with A antigen negativity was shown to be significantly shorter than for patients with A antigen positive tumours (Lee *et al*, 1991). Bryne *et al* (1990) found that 71% of oral squamous cell carcinomas showed less than 10% cell membrane staining for the H antigen. Matsumoto *et al* (1993) demonstrated a correlation of expression of ABH antigens with metastatic potential in human lung carcinomas with 54% having partial or complete loss of antigen expression. Orlow *et al* (1998) found that 59% of bladder tumours of A, B or AB individuals showed lack of expression and/or loss of an *ABO* allele using PCR and Southern analysis.

Investigation of the loss of A/B antigen expression in immortalised human urothelial cell lines, HCV 29 and Hu 609, which had no A transferase activity or any A antigens on their cell surface, showed that they retained an intact *A* allele and transcribed A mRNA (Meldgaard *et al*, 1994). To investigate this further, this same group investigated loss of heterozygosity (LOH) of *ABO* in 22 bladder tumours that were negative for A/B antigens. They found that loss of ABH antigens was not due to loss of the transferase producing allele or to down-regulation of *ABO* gene transcription, but suggested that perhaps there was an alteration in post-translational control. This is contradictory to the observation of Orlow *et al* (1998) above who found loss of an *ABO* allele in some bladder tumours with loss of ABH expression.

Another observation in tumours is 'incompatible A expression' which refers to the expression of A antigens in tumours of blood group O and B individuals, which normally do not produce A antigens (Clausen *et al*, 1986; Metoki *et al*, 1989). David *et al* (1993) investigated incompatible A antigen expression in 31 gastric tumours of O phenotype and *OO* genotype. They found eight cases that expressed A antigen immunohistologically and concluded that incompatible A antigen expression was due to A transferase expression somehow deriving from the *O* allele. It has been generally accepted that the regulation of ABO phenotypes is

based on structural DNA encoding, therefore the expression of inappropriate antigens in tumours may indicate that regulation occurs at an epigenetic level (reviewed in Clausen and Hakomori, 1989).

Type of Cancer	Change in Expression of ABH Antigens
Stomach	Loss of A and B
Proximal Colon	Loss of A and B
Distal Colon	Appearance of ABH antigens Decrease of A, B and H antigens in metastases
Pancreas	Loss of A, B and H
Lung	Loss of A and B
Bladder	Loss of A, B and H
Prostate	Loss of A and B Increase in H
Breast	Loss of A and B Increase in H
Endometrium	Loss of A and B Increase in H
Ovary	Loss of A and B
Head and Neck	Loss of A and B
Thyroid	Appearance of A, B and H Decrease of A and B in metastases

Table 1.5. Changes in ABH antigen expression in carcinomas.
Modified from Le Pendu *et al* (2001).

1.4.8 Loss of ABH Antigens in Haematological Malignancies

Alteration of ABH antigens in haematological malignancy was first reported by van Loghem *et al* (1957) who described very weak A antigen expression on the red cells of an AML patient, who had previously shown normal A antigen expression. Loss of A, B, or H antigens from the surface of red blood cells is now recognised as a recurrent observation in haematological malignancy (Race and Sanger, 1975; Daniels, 1995). In a normal individual of type A, B or AB, complete agglutination of the red blood cells is observed after incubation with antibodies reactive against their blood group. In patients with loss of ABH antigens, a

varying proportion of red blood cells do not agglutinate giving a characteristic mixed field reaction. Mixed field reactions can also occur in normal individuals where they are associated with rare alleles of the *ABO* gene such as A^3 and B^3.

In epithelial tumours, loss of ABH antigens is seen in the cancer cells which are often dedifferentiated, however, in myeloid malignancies, loss of ABH antigens is seen only on red blood cells since most haematopoietic cells normally do not express *ABO*. Malignant stem cells regularly retain the ability to differentiate along several lineages including the erythroid lineage (Fialkow *et al*, 1981; Keinanen *et al*, 1988; Suciu *et al*, 1993). Loss of ABH antigens in a subset of red blood cells derived from a malignant stem cell is indicative of genetic changes that have occurred in the malignant stem cell. Red blood cells that are deficient in A or B antigens have been shown to have decreased transferase activities, which supports the notion that the loss of antigens reflects a genetic change at the *ABO* locus and not at the cell surface or membrane precursors (Salmon *et al*, 1984). Changes in H antigen expression may result from genetic alterations at the *FUT1* locus.

Whether the mechanisms of loss of ABH antigens in epithelial tumours and haematological malignancies are truly analogous to each other remains to be determined.

1.5 Genetic and Epigenetic Changes in Cancer Development

Cancer is a clonal disorder originating from a single aberrant cell that has acquired numerous genetically inheritable changes (Nowell, 1976). Multiple progressive events are required to change a normal cell into a cancer cell. These steps have been defined as initiation, promotion and progression. Initiation involves an irreversible genetic change that is not enough to cause cancer on its own. Promotion is the clonal expansion of the initiated cells into a benign or pre-neoplastic growth involving reversible epigenetic changes. Progression involves events that are required for the change from a benign or pre-neoplastic population to a malignant population (reviewed in Boyd and Barrett, 1990).

Carcinogenesis requires numerous genetic changes in two groups of genes, oncogenes and tumour suppressor genes. Oncogenes are involved in the cellular growth and differentiation of a cell. For cancer to develop, oncogenes must be activated either by point mutation, chromosomal rearrangement or DNA amplification. The second group of genes involved in carcinogenesis are tumour suppressor genes, which are involved in the regulation of cellular proliferation. These genes are often lost or inactivated in tumours.

Genetic alterations that are involved in haematological malignancies have only been partially characterised, but inactivation of tumour suppressor genes and activation of oncogenes have been identified as essential. Both copies of a tumour suppressor gene need to be inactivated for neoplastic transformation. Thus, two sequential molecular events are required for tumour suppressor gene inactivation with one event usually leading to loss of chromosome material (loss of heterozygosity - LOH) and the second being a point mutation, deletion, chromosomal rearrangement or methylation (Boyd and Barrett, 1990). LOH is the process where the remaining wild type allele of a gene is lost either by deletion of the chromosomal region harbouring the gene or by recombination leading to the replacement of the wild type allele with a duplicated copy of the mutant allele (reviewed in Thiagalingam *et al*, 2002).

1.5.1 CpG Islands

Inactivation of tumour suppressor genes by methylation is currently under extensive investigation. Methylation predominantly occurs at CpG dinucleotides, which are rare, occurring at twenty percent of their expected frequency (Bird, 1986). CpG denotes a cytosine adjacent to a guanosine on the same DNA strand and not cytosine hydrogen bonding to guanosine on the complementary DNA strand. The rarity of CpGs results from their methylation, with 60 - 90% of CpGs methylated in the genome. There is a tendency for 5-methylcytosine to deaminate to thymine, which is not efficiently repaired by DNA repair

mechanisms (Bird, 1986). It is more difficult for the cell to correct the T:G mismatch created by 5-methylcytosine because thymine is a normally occurring DNA base. The correct or incorrect sequence depends on whether the thymine or the guanosine is repaired (Jones, 1996).

CpG islands are clusters of unmethylated CpGs located near and spanning the promoters of many genes. These CpG islands are areas of DNA 0.5 - 5 kb in length, which are GC rich (60 – 70%), and have a CpG/GpC ratio of 0.6 – 1.0 indicating little or no suppression of the frequency of CpG, that is, these CpG islands show the expected CpG frequency (Bird, 1986).

CpG islands have maintained their high level of CpG dinucleotides due to the absence of methylation (Bird, 1986). CpG islands are almost always associated with promoters or coding regions of genes. CpG islands in 'housekeeping' genes are protected from methylation but some genes with a tissue-specific expression pattern become methylated and remain so in adult tissues. They only become demethylated during differentiation of tissues in which these genes are expressed (reviewed in Versteeg, 1997).

DNA methyltransferases (DNMT) are the enzymes that catalyse the transfer of the methyl group to a cytosine in a CpG dinucleotide. *DNMT1* was the first DNMT cloned and is responsible for duplicating the DNA methylation pattern during DNA replication (Bestor and Verdine, 1994). *DNMT1* is also the most abundant of the DNA methyltransferases and is capable of directly binding histone deacetylases (HDAC) (Fuks *et al*, 2000; Robertson *et al*, 2000; Rountree *et al*, 2000). Knockout of both alleles of *DNMT1* in cultured human colon cancer cells eliminated a large amount of DNA methylating activity but it did not result in a complete loss of abnormal methylation in many of the hypermethylated genes investigated (Rhee *et al*, 2000). *DNMT3a* and *3b*, like *DNMT1*, have been shown to be essential for mammalian development (Okano *et al*, 1998; Okano *et al*, 1999).

In eukaryotes, genomic DNA is packaged with histone proteins into chromatin (Grant, 2001). Core histones are modified by acetylation of lysine residues in the amino terminal tail domains that interact with the negatively charged phosphates of the DNA backbone. Histone acetylation leads to a charge neutralisation and separation of DNA from histones, allowing transcriptions factors to access the DNA. Hence, histone hyperacetylation leads to transcriptionally active regions of the genome while hypoacetylation leads to silenced regions (reviewed in Archer and Hodin, 1999). The level of histone acetylation is based on the activity of two groups of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs). The chromatin that contains CpG islands is usually heavily acetylated, lacks histone H1 and includes a nucleosome-free region (Tazi and Bird, 1990). This open configuration allows for transcription factors to interact with gene promoters (Cross and Bird, 1995).

1.5.2 Methylation and Transcription

Many studies support an inverse relationship between CpG island methylation and gene expression. Hypomethylated DNA is associated with active chromatin whereas methylated DNA is associated with inactive chromatin (Razin, 1998). Whether methylation is a cause or consequence of events leading to transcriptional inactivation is not known. There are two examples of the relationship between CpG methylation and transcriptional inactivation in normal human cells. They are X-inactivation and imprinting.

1.5.2.1 X-Inactivation and Imprinting

X-inactivation in female mammals is associated with the methylation of CpG islands of X-linked housekeeping genes, with methylation occurring at, or soon after, gene inactivation (Jones, 1996). CpG islands associated with transcriptionally silenced genes on the inactive X chromosome are heavily methylated, while the homologous sequences on the active X chromosome are unmethylated (Mohandas *et al*, 1981).

Methylation of the CpG islands on the inactive X chromosome contributes to the repression of the inactivated genes, since the genes can be reactivated by using an inhibitor of DNA methylation, such as 5-azacytidine (Mohandas *et al*, 1981). Further evidence that methylation correlates with gene activity on the X chromosome, comes from the fact that a CpG island associated with a gene that escapes X inactivation remains unmethylated on both the active and inactive X chromosomes (Mohandas *et al*, 1981).

Methylation is also observed in parental imprinting, which is the inherited expression of a gene from either the paternal or the maternal allele. In imprinting, the active allele is often unmethylated while the silenced allele is methylated (reviewed in Tycko, 2000).

1.5.2.2 Transfection Experiments

The inverse relationship between methylation and gene expression has also been shown by transfection experiments. Unmethylated exogenous genes, when stably transfected into cultured cells, are expressed at a basal level. However, transcription of *in vitro* methylated DNA sequences is inhibited when transfected into different cell types (Yisraeli *et al*, 1988). DNA methylation can inhibit *β-globin* gene expression in mouse fibroblast and erythroleukaemia cells (Yisraeli *et al*, 1988). Studies of the human *β-globin* gene have shown that in fibroblasts the endogenous gene is heavily methylated and unexpressed, yet unmethylated copies of this gene are transcriptionally active when introduced into the same cell type (reviewed in Cedar, 1988).

Further evidence supporting the inverse relationship between CpG island methylation and gene expression comes from transfection experiments with an artificially methylated hamster gene (*APRT*). Methylation of the body of the gene did not affect transcription, but methylation of the 5' CpG island led to a decrease in the level of transcription (reviewed in

Bird, 1986). Lu and Davies (1997) found that methylation of CpG rich regions of the transglutaminase promoter correlated with a reduced level of transcriptional activity. Demethylation of the cells by treating them with 5-azacytidine led to increased enzyme expression.

1.5.2.3 Methylation Mechanisms Which Affect Gene Expression

There are two ways in which methylation can affect gene expression. Firstly, methylated CpG residues can directly interfere with the binding of specific transcription factors to DNA. Transcription factors such as AP-2, c-Myc/Myn, cyclic AMP-dependent activator CREB, E2F and NF-κB recognise sequences with CpG residues. Binding of each of these transcription factors has been shown to be abolished by methylation (reviewed in Singal and Ginder, 1999). The transcription factors Sp1 and CTF are not affected by methylation of their binding sites (Tate and Bird, 1993).

Secondly, direct binding of specific factors termed methyl-CpG binding factors to methylated DNA, like MeCP1 and MeCP2, can mediate repression. MeCP1 binds to DNA with many methylated CpGs while MeCP2 is more abundant than MeCP1 and can bind DNA with one methylated CpG (Meehan *et al*, 1989; Meehan *et al*, 1992). MeCP2 has recently been shown to associate with a co-repressor complex that consists of the transcriptional repressor mSin3A and histone deacetylase (Nan *et al*, 1998; Jones *et al*, 1998). Homozygous deletion of mouse *MeCP2* is lethal suggesting an important role for this protein in embryonic development (Zingg and Jones, 1997). It has been suggested that a high concentration of methyl-CpGs in a methylated CpG island led to binding of methyl-CpG binding factors. These proteins might then ensure transcriptional suppression by stabilisation of higher order folding of chromatin (reviewed in Antequera *et al*, 1990). Unmethylated DNA adopts a DNase I sensitive structure typical of other active genes, while fully methylated sequences are resistant to DNase I and are structurally similar to inactive genes of the cell (reviewed in Cedar, 1988).

1.5.3 Alterations of Methylation in Cancer

In recent years, there has been considerable interest in detecting altered patterns of methylation during cancer development. It is recognised that critical changes in methylation during tumorigenesis occur in some of the normally unmethylated CpG islands that encompass the promoter regions of many genes (Baylin *et al*, 1998). Particular interest has been focussed on tumour suppressor genes such as *RB*, *p16* and *BRCA1* that can have methylated promoter CpG islands in their respective tumours (Greger *et al*, 1989; Gonzalez-Zulueta *et al*, 1995; Herman *et al*, 1995; Merlo *et al*, 1995; Dobrovic and Simpfendorfer, 1997; Stirzaker *et al*, 1997).

Alterations in DNA methylation are very common in cancer cells. They include widespread loss of methylation from normally methylated sites, hypomethylation of genes (Feinberg and Vogelstein, 1983), increased levels of the enzyme DNA methyltransferase (reviewed in Jones, 1996) and *de novo* methylation of CpG islands (Baylin *et al*, 1987). DNA hypomethylation of oncogenes in human tumours has been correlated with gene expression. An example of this is *bcl-2* in B-cell CLL and *K-ras* in lung and colon carcinomas (reviewed in Singal and Ginder, 1999).

1.5.3.1 Hypomethylation and Cancer

While the CpG islands associated with tumour suppressor genes become methylated, the DNA of tumour cells shows widespread hypomethylation of CpG dinucleotides located in the coding region and introns of the gene (reviewed in Esteller and Herman, 2002). DNA hypomethylation contributes to cancer by causing chromosomal instability and through loss of imprinting. Lack of methylation favours mitotic recombination leading to LOH and chromosomal rearrangements (Esteller and Herman, 2002). This is supported by the observation that mice lacking *DNMT1* have increased rates of genetic deletions (Chen *et al*, 1998) and patients with mutations in *DNMT3b* have many chromosomal aberrations (Xu *et al*, 1999).

1.5.3.2 Levels of DNMTs

There is increased activity of DNA methyltransferases in cancers with significant over expression of *DNMT3b* in several human cancers (Robertson *et al*, 1999). Significant increases in the expression of *DNMT1*, *3a* and *3b* have also been reported in AML and the acute phase of CML (Mizuno *et al*, 2001).

The evidence for the role of DNA methylation in carcinogenesis is strengthened by work done with methyltransferase-defective mice. Experiments using these mice established that they had decreased CpG methylation and this correlated with a decrease in colon polyp formation in mice carrying a predisposing cancer-causing mutation (Laird *et al*, 1995). Also, homozygous embryos had a three-fold decrease in DNA methylation and died before day eleven (Li *et al*, 1992). This indicates an essential role for DNA methylation in normal mammalian development.

1.5.3.3 Targets for Mutation

CpG sites are seen as hot spots for inactivating mutations in tumour suppressor genes. Coulondre *et al* (1978) were the first to demonstrate that cytosine methylation sites in prokaryotes were mutational hot spots due to the hydrolytic deamination of 5-methylcytosine. In mammals, CpG sites are also mutational hot spots and contribute to the generation of polymorphisms, germline mutations and cancer causing mutations (reviewed in Jones, 1996). Cytosine methylation is responsible for the generation of a large number of disease causing point mutations in tumour suppressor genes in somatic and germline cells (Jones, 1996). This is due either to the spontaneous deamination of 5-methylcytosine, or to the side reactions involved in the enzymatic modification of cytosine in DNA (Jones, 1996). An example of this is *p53* in which 25% of point mutations in this gene in all human cancers studied occurred at CpG sites and 50% occurred at CpG sites in colon cancer (Jones, 1996).

1.5.3.4 *Calcitonin*

One of the first reports of a change in methylation in carcinogenesis came from work on the *calcitonin* (*CT*) gene. Hypermethylation of the 5' region of *CT* was found in 90% of patients with non-Hodgkin's T and B cell lymphoid neoplasms, and in 95% of tumour DNA samples from patients with acute non-lymphocytic leukaemia (ANLL) (Baylin *et al*, 1987). Hypermethylation was not found in DNA from non-neoplastic adult tissues such as sperm or lymphocytes.

1.5.3.5 *Retinoblastoma*

There is much support for the role of aberrant methylation in carcinogenesis, with *de novo* methylation an epigenetic pathway for the inactivation of tumour suppressor genes. An example of this is the retinoblastoma gene, *Rb*, in which hypermethylation of the 5' CpG island has been found in a small fraction of sporadic retinoblastomas (Greger *et al*, 1989; Sakai *et al*, 1991; Greger *et al* 1994). It has also been shown that *in vitro* methylation of the *Rb* promoter, in a reporter gene construct, reduced gene expression by 92% (Ohtani-Fujita *et al*, 1993). Reduced levels of mRNA were also found in a hypermethylated retinoblastoma cell line, suggesting that *Rb* can be inactivated by hypermethylation *in vivo* (Greger *et al*, 1994).

Recently, Stirzaker *et al* (1997) used bisulphite genomic sequencing to study in detail the methylation pattern of the *Rb* promoter region from primary retinoblastomas. In this way the methylation status of all the CpG sites in the *Rb* promoter region were characterised rather than only the sites that can be studied using restriction enzyme digestion. They wanted to determine whether the transcription factor binding sites were specifically methylated in each patient shown to have promoter methylation by restriction enzyme analysis, and whether other regions of the promoter were also methylated. Six out of seven of the methylated patients showed LOH for the *Rb* gene, therefore, in these tumours one allele is lost and the

other is methylated. They found hypermethylation was not confined to specific CpG sites. Although methylation was extensive, the overall level varied between patients from 20-30% for some sites to nearly 100% for other sites. This suggests that complete methylation at all CpG sites is not essential for silencing of the *Rb* gene.

1.5.3.6 *Von Hippel Lindau*

The von Hippel Lindau (*VHL*) gene is another tumour suppressor gene that was found to be hypermethylated and inactivated in a subset of renal cell lines and tumours that did not have a mutated *VHL* gene (Herman *et al*, 1994). Support for abnormal CpG island methylation associated with inactivation of this gene came from re-expression of the *VHL* gene in renal cell cancers after treatment with the demethylating agent 5-aza-2'-deoxycytidine (Herman *et al*, 1994). Prowse *et al* (1997) detected hypermethylation of *VHL* in 33% of tumours without LOH and were the first to report somatic methylation in a familial cancer syndrome. This would suggest that the first hit was a germline mutation and methylation of the other *VHL* allele was the second somatic inactivation event.

1.5.3.7 *p16* and *p15*

Expression of the *p16* tumour suppressor gene was found to be controlled by methylation of its 5' CpG island, with *de novo* methylation of this island the mechanism for *p16* inactivation in bladder carcinomas. A significant correlation was found between methylation of exon 1 of *p16* and transcriptional silencing of the gene (Gonzalez-Zulueta *et al*, 1995). Silencing of *p16* expression by methylation was also found in 24% of human gliomas, but was not present in normal brain (Costello *et al*, 1996) as well as in 19% of oesophageal squamous cell carcinomas (Maesawa *et al*, 1996). Hypermethylation of the CpG island surrounding the first exon of *p16* was also observed in neuroblastomas, which often lacked *p16* expression (Takita *et al*, 1997). Merlo *et al* (1995) found *de novo* methylation of the 5' CpG island of *p16* in about 20% of different primary neoplasms but not in normal tissues.

The tumour suppressor gene, *p15*, has been reported to be commonly inactivated in gliomas and leukaemias in association with promoter region hypermethylation. Treatment with the demethylating agent 5-aza-2'-deoxycytidine led to re-expression of *p15* mRNA (Herman *et al*, 1996). *p15* and *p16* are both located on chromosome 9p21 and in a study of multiple myeloma they were both investigated for methylation since they were not inactivated by deletions or mutations (Ng *et al*, 1997). It was found that inactivation of these two genes was due to hypermethylation of their 5' CpG islands with hypermethylation observed for *p16* in 75% of multiple myeloma patients and 67% for *p15* (Ng *et al*, 1997).

1.5.3.8 Methylation as a Therapeutic Target

Abnormal methylation is an attractive target for therapeutic intervention through the use of demethylating agents, especially since the growth regulating genes are not mutated or lost but simply silenced. Expression of methylated genes can be restored, in principle, through the use of inhibitors of DNA methylation like the cytidine analogue 5-azacytidine and 5-aza-2'-deoxycytidine (Jones, 1996).

5-azacytidine, can be integrated into DNA and its incorporation leads to hypomethylation of DNA, which is attributed to impaired methylation at the sites of substitution. Another DNA methylation inhibitor is the cytosine analogue 5-aza-2'-deoxycytidine. 5-aza-2'-deoxycytidine can integrate into DNA at the target site for DNA methylation, but it cannot accept the transfer of a methyl group because of the nitrogen in the fifth position of the cytosine ring. The covalent binding of DNA methyltransferase to 5-aza-2'-deoxycytidine substituted DNA mediates its cytotoxic effects. A phase II study, involving the treatment of 29 elderly patients with MDS with continuous infusion of low dose 5-aza-2'-deoxycytidine, found that 15 out of 28 evaluable patients responded to this treatment (Wijermans *et al*, 1997). Although nine reached complete remission, five patients died due to its toxicity. Preliminary findings

suggested that decreasing the dose of 5-aza-2'-deoxycytidine maintained the efficacy of this DNA methylation inhibitor with an associated decrease in toxicity.

Patients with β-thalassemia show foetal globin expression after administration of 5-aza-2'-deoxycytidine (reviewed in Jackson-Grusby *et al*, 1997). This drug has also been used *in vivo* to demonstrate the inverse relationship between DNA methylation and gene expression and to reactivate epigenetically silenced genes. Laird *et al* (1995) have shown that 5-aza-2'-deoxycytidine reduces intestinal tumour multiplicity in multiple intestinal neoplasia (*Min*) mice by reducing genomic methylation levels. *Min* mice are heterozygous for the *Min* mutation of the adenomatous polyposis coli (*APC*) gene. Since 5-aza-2'-deoxycytidine is mutagenic its utility must be questioned and highlights the need to identify novel specific inhibitors of DNA methyltransferase.

1.5.4 DNA Methylation in the Promoter Region of *ABO*

The 5' upstream *ABO* genomic sequence, which contains potential transcription initiation sites, lacks either a TATA or a CAAT box, but is rich in CpG dinucleotides (Yamamoto, 2001). There is a G + C content of 76% and a CpG/GpC ratio of 0.9 so this upstream region fulfils the criteria for a CpG island (Kominato *et al*, 1997).

Analysis of the *ABO* gene promoter region in several human carcinoma cell lines showed that DNA methylation of the *ABO* gene promoter region was inversely correlated with its gene expression (Kominato *et al*, 1999). Treatment of a non-expressing gastric cancer cell line MKN 28 with the DNA methyltransferase inhibitor 5-aza-2'-deoxycytidine resulted in demethylation of the promoter and re-expression of the *ABO* gene.

Iwamoto *et al* (1999) analysed the methylation patterns of *ABO* in A antigen positive and negative clones derived from the colon adenocarcinoma cell line SW480. The antigen positive clones were unmethylated and the antigen negative clones were methylated. They found that the methylation was observed only on the sense strand but not on the antisense strand though they did not elaborate on the significance of this strand-specific methylation.

1.6 Aims

The aim of this project was to investigate the molecular mechanisms behind the loss of ABH antigens in leukaemia and to assess whether methylation was a possible mechanism for silencing of the *ABO* gene. Since there have been reports of coordinate decreases in the A transferase and the AK1 transferase, another aim of this project is to assess whether there is a silencing event affecting a possible tumour suppressor gene in the 9q34 region that is the target in these haematological malignancies.

CHAPTER 2

Materials & Methods

2.1 Materials

Reagent	Reagent Supplier or Composition
Ammonium Persulfate	Progen Industries Limited.
5-aza-2'-deoxycytidine	Sigma. Made up in ultra pure water (UPW) and filter sterilised.
BigDye™ Terminator	Applied Biosystems.
Concert™ Rapid PCR purification System	Invitrogen.
DDS Buffer	95% formamide, 10 mM sodium hydroxide (NaOH), 0.25% xylene cyanole (ICN Biomedicals Inc), 0.25% bromophenol blue (BioRad).
Dextran Sulphate (50% w/v)	50 g of dextran sulphate sodium salt (Pharmacia) was dissolved in a final volume of 100 ml with MilliQ water (MQW) on a magnetic stirrer for several hours. The solution was autoclaved*.
Digoxigenin-11-Uridine-5'-Triphosphate (DIG-11-UTP)	Roche.
DIG Luminescent Detection Kit	Roche.
DIG Washing Buffer	Add 0.3% (v/v) Tween® 20 to Maleic acid buffer.
Dimethyl Sulphoxide (DMSO)	Sigma.
DNAzol BD Reagent	Invitrogen. For extraction of DNA from blood.
Deoxynucleotide triphosphates (dNTPs) (10 mM) each	ICN. 100 mg of each dNTP (dATP, dCTP, dGTP, dTTP) were resolvated in UPW and the pH adjusted to 7.0 - 8.0 using sterile 1 M Tris base. The final volume was made up to 18 ml with UPW.
Ethylene-diamene tetra-acetic acid (EDTA) (0.5 M, pH 8.0)	Disodium ethylene diamine tetra-acetate.$2H_2O$, 186.1 g, was dissolved in 800 ml of MQW. The pH was adjusted to 8.0 with NaOH, volume made up to 1 litre and sterilised by autoclaving.
Ethidium Bromide (1 mg/ml)	One 10 mg pellet dissolved in 10 ml of sterile MQW.
Foetal Bovine Serum (FBS)	JRH Biosciences.
Formaldehyde Solution (37%)	Sigma.
Formamide	Deionisation required 5 g of Molecular Biology Grade AG® 501-X8 (D) Resin (BioRad) per 100 ml of formamide (AnalaR) and stirring for 1 hour. Deionised formamide was stored at –20°C.
Glutamine (200X)	Dissolve 0.292 g of L-glutamine (Sigma) in 10 ml of UPW. Solution was then filtered through a 0.22 μm filter and stored at –20°C.

Helix Pomatia Lectin (detects the A antigen)	Lyophilised Helix Pomatia lectin conjugated to either fluorescein isothiocyanate (FITC) or biotin (Sigma), was resuspended in RSB to make a 0.25 mg/ml stock solution.
Hydroquinone (100 mM)	0.1 g hydroquinone (Sigma) was dissolved in 9 ml of MQW by inversion.
Loading Buffer (2X) (for Northerns)	For 1 ml: 500 µL deiononised formamide, 100 µl 10X MOPS buffer, 167 µl 37% deionised formaldehyde, 133 µl UPW, 100 µl glycerol, 0.08 g bromophenol blue, 0.08 g xylene cyanole, 3 µl of a 10 mg/ml EtBr stock
Loading Buffer (6X)	50 ml glycerol, 40 ml 0.5 M EDTA (pH 8.0), 0.25 g of bromophenol blue and 0.25 g of xylene cyanol were combined. The solution was made up to 100 ml with sterile MQW and stored at 4°C.
Lymphoprep	Pharmacia.
Maleic Acid Buffer	0.1 M maleic acid (Sigma), 0.15 M NaCl. Adjusted to pH to 7.5 with NaOH pellets and autoclaved before use.
MDE (2X) Gel Solution.	FMC BioProducts.
MMLV Reverse Transcriptase (200 U/µl)	Invitrogen.
Morpholinopropansulfonic acid (MOPS) (10X)	200 nM MOPS, 50 nM sodium acetate, 10 mM EDTA (pH 7.0). Made up in sterile MQW and autoclaved.
Novaclone (detects the B antigen)	Novaclone anti-B blood grouping reagent (Dominion Biologicals Ltd). The secondary antibody was goat F(ab')2 anti-mouse IgG (H+L) PE conjugate (Caltag). The antibody was diluted 1:10 in SSB.
Penicillin-Streptomycin (500X)	To 180 ml UPW, 6 g of penicillin G and 10 g of streptomycin sulphate were added. Final volume of 200 ml, filtered through a 0.22 µm filter and stored at –20°C.
Phosphate Buffered Saline (PBS)	The following reagents were dissolved in MQW: 8.0 g of NaCl, 0.2 g of KCl, 1.15 g of Na_2HPO_4 and 0.2 g of KH_2PO_4. The solution was made up to 1 litre with MQW before autoclaving.
Proteinase K (10 mg/ml)	Proteinase K (Merck), 100 mg, was dissolved in 10 ml of sterile MQW and stored at –20°C.
pUC19/*Hpa*II (500 ng/µl)	Geneworks. Made up to 50 ng/µl using the 2X loading buffer provided and TE.
Random Hexamers	Supplied as a sodium salt (Amersham Pharmacia) were dissolved in UPW to give a final concentration of 250 ng/µl and stored at –20°C.

Restriction Endonucleases	These were purchased from New England Biolabs, Roche or MBI Fermentas.
RNasin® RNase Inhibitor (40 U/µl)	Promega.
Routine Staining Buffer (RSB)	Isoton II (Coulter), 1% foetal calf serum, 0.1% NaN_3, pH 7.2. The solution was filtered using a 0.45 µm filter.
RPMI 1640 medium	Invitrogen. Medium was made according to the manufacturer's instructions. To 450 ml of RPMI 1640 media, 50 ml of foetal bovine serum was added (final concentration 10%). 1 ml of 500X penicillin-streptomycin sulphate and 2.5 ml of 200X glutamine were then added.
Salmon Sperm DNA (10 mg/ml)	100 mg of salmon sperm DNA (Sigma type III sodium salt) was dissolved in 10 ml of MQW and 200 µl of 10 M NaOH by incubation at 37°C overnight. The salmon sperm was sheared by several treatments with the Branson Sonifier (Selby Anax) using Constant Cycle and an Output of between 2 and 3.
Sodium Acetate (3 M, pH 5.2)	To 800 ml of MQW, 246.1 g of sodium acetate.$3H_2O$ was added. The pH was adjusted to 5.2 with glacial acetic acid, the final volume was made up to 1 litre and sterilised by autoclaving.
Sodium Bisulphite (4.8 M)	10 g of sodium bisulphite (Sigma) was dissolved in 10 ml of MQW by gentle agitation. The pH was adjusted to 5.0 using freshly made 10 M NaOH. The final volume was adjusted to 20 ml with MQW.
Sodium Dodecyl Sulphate (20% SDS)	To 180 ml of sterile MQW, 40 g of sodium dodecyl sulphate (Sigma) was added and the volume was adjusted to 200 ml with sterile MQW.
Sodium Chloride (5 M NaCl)	To 800 ml of MQW, 292.2 g of NaCl was added. The final volume was adjusted to 1 litre and sterilised by autoclaving.
Sodium Hydroxide Solution (60% w/v)	Merck.
Spherical Staining Buffer (SSB)	0.15 M NaCl, 0.0072 M Na_2HPO_4, 0.0028 M NaH_2PO_4, 5 mg/ml BSA, 0.01% Nonidet P40, pH 7.2. The solution was filtered through a 0.45 µm filter.
SPP1/*EcoR*I (500 ng/µl)	Geneworks. Made up to 50 ng/µl using the 2X loading buffer provided and TE.
SSC (20X)	175.3 g of NaCl and 88.2 g of tri-sodium citrate were dissolved in 800 ml of MQW. The volume was adjusted to 1 litre with MQW and the solution was autoclaved before use.
Streptavidin Phycoerythrin PE	Streptavidin Phycoerythrin PE (Becton Dickinson) was the secondary reagent for the biotinylated Helix.

Strip-EZ™ RNA T7 Kit	Ambion.
SYBR® Gold	Diluted 1:10,000, 40 µl in 400 ml of 0.6X TBE. The gel was stained for 30 min by gentle rocking.
TBE (5X)	Dissolve 54 g of Tris base, 27.5 g of boric acid and 20 ml of 0.5 M EDTA (pH 8.0) in a final volume of 1 litre of MQW.
TE	10 mM Tris-HCl pH 7.5, 1 mM EDTA (pH 8.0).
TEMED	Sigma.
TES	10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.1 M NaCl.
Tris Base (1 M)	Tris base, 121.1 g in 1 litre of MQW and autoclaved.
Tris HCl (1 M, pH 8.0)	Tris base, 121.1 g, was dissolved in 800 ml of MQW. The pH was adjusted to 8.0 using concentrated HCl. The final volume was made up to 1 litre and was autoclaved before use.
Trizol	For RNA extraction from mononuclear cells (Invitrogen).
Trizol LS	For RNA extraction from blood (Invitrogen).
Trypan Blue	Sigma.
Ulex Europaeus lectin (UEA-I) (detects the H antigen)	Lyophilised UEA-I conjugated to FITC (Sigma) was resuspended in RSB to make a 0.5 mg/ml stock solution.
ULTRAhyb™ Ultrasensitive Hybridization Buffer	Ambion.
Water	1) Ultra Pure Water (UPW – Biotech). 2) MilliQ water (MQW) has been deionised and filtered by passage through a MilliQ filtering system.
Wizard™ PCR Preps DNA Purification System	Promega.

* Solutions requiring sterilisation were autoclaved at 103 kPa, 121°C for 20 minutes on a Smith Pressure Steriliser.

2.1.1 List of Primers

Primer Name	Primer (5' - 3')	Length (bp)	AT %	CG %	Tm
ABL *BstN*I F	ggacttgcagtccacgggaagaca	24	42%	58%	76
ABL *BstN*I R	actgccctttccagcttcttcctt	24	50%	50%	72
ABL *Hha*I/*Mae*II F	cacgccgggagaagccaccttcGc	24	29%	71%	82
ABL *Hha*I/*Mae*II R	gcacacgccacttagaaaagagcgtc	26	46%	54%	80
AO1 R	actgctcgttgaggatgtcgatgt	24	50%	50%	72
AO1 F	ttggagtcgcatttgcctctggtt	24	50%	50%	72
ABO2 F	agacggcggagaagcacttcatgg	24	42%	58%	76
ABO2 R	ctcgatgccgttggcctggtcga	23	35%	65%	76
new ABO2 F	tgccttgcagatacgtggctttcc	24	46%	54%	74
RT-AO1 F	cttgtttggttacggggtcctaag	24	50%	50%	72
PA5'	ttgcggacgctggccggaaaac	22	36%	64%	72
RT-AO1 R	ggaaagccacgtatttcttgatggc	25	52%	48%	74
RT-AO ex 4/5 F	tcgttgccaaggatggtctaccc	23	43%	57%	72
RT-AO ex 5/6 F	caaaggtgctgacaccgtgtagga	24	46%	54%	74
RT-AO ex 7 R	acacggtggcccaccatgaagtg	23	39%	61%	74
RT-ABO ex 6/7 F	tgccatcaagaaatactgtgctttc	25	60%	40%	70
ABO x7 F	catgatggtcgaccaggccaacgg	24	38%	63%	78
ABO x7 R	tgccagcgcttgtaggcgcgcac	23	30%	70%	78
ABO CpG F	gcgccgtcccttcctagcaggggt	24	29%	71%	82
ABO CpG R	tgcgggtcccaggcacccacagcg	24	25%	75%	84
ABO CpG T7 R	ggatcctaatacgactcactatagggaggTGCGGGTCCCAGGCACCCACAGCG	53	40%	60%	170
ABO BIS F	ggaTagggTTTTaaggtaTTagggTTa	27	63%	37%	74

ABO BIS R1*	gcaacacctcgAccatAActcc	22	45%	55%	68
AGAPT2 *Nci*I F	gtccagcccacaagctgcatcCg	23	35%	65%	76
AGAPT2 *Nci*I R	catcgatggaagtagtatagggtgat	26	58%	42%	74
AGAPT2 *BstU*I F	cttctgtcactggcctcagacacag	25	44%	56%	78
AGAPT2 *BstU*I R	gccagcaggagcagcaggccGc	22	23%	77%	78
AK1 *Taq*I F	gaacccgtcatcgccttctatgag	24	46%	54%	74
AK1 *Taq*I R	caagcacatgaatcaactccaactg	25	56%	44%	72
RT-AK1 ex 6/7 F	cattgtgcgcaaggtcaacgctga	24	46%	54%	74
ALAD F	tggtgccatatgcagcttggtttc	24	50%	50%	72
ALAD R	gtagctgggactacaggcacatac	24	46%	54%	74
RT-ALAD F	agtccgttctgcacagcggctact	24	42%	58%	76
RT-ALAD R	gaccatgggaggtgtagggacaca	24	42%	58%	76
C8G *Dde*I F	aggcaacggatggacatggacttc	24	46%	54%	74
C8G *Dde*I R	cacaacttcacctgcttccctctc	24	46%	54%	74
RT-C8G ex 1 F	ctgggactgcgaccctcttgactc	24	38%	63%	78
RT-C8G ex 2/3 R	cagcagatcccatccagctttcgg	24	42%	58%	76
CDK9 *Eco*NI F	ttgtccttgctggttttctggatg	24	54%	46%	70
CDK9 *Eco*NI R	gaacaaaacacacaagtcctggag	24	54%	46%	70
DAPK1 *Tas*I F	tacaaggaaagcttcagcagcatcat	26	58%	42%	74
DAPK1 *Tas*I R	aggaaatccttgggagccccgtAa	24	46%	54%	74
DBH *Eco*NI F	ctggagtgaccagaaggggcagat	24	42%	58%	76
DBH *Eco*NI R	cgcatctagaaaccatgtcagctac	25	52%	48%	74
RT-DBH ex 3/4 R	gacgatgggctcgtacttgataatg	25	52%	48%	74
EPB72 *Rsa*I F	tctgcttactggccagaaagctgg	24	46%	54%	74
EPB72 *Rsa*I R	ggtatctggagacaggagaactca	24	50%	50%	72
FBXW2 *BstU*I F	gcgccttggctgtttttctttccc	24	46%	54%	74

FBXW2 *BstU*I R	**tttcgcagctaaggaaatggaggc**	**24**	**50%**	**50%**	**72**
FPGS *BstN*I F	**gaagctgctggagcccgcactgt**	**23**	**35%**	**65%**	**76**
FPGS *BstN*I R	**ccaggccactgcaacagtgagaga**	**24**	**42%**	**58%**	**76**
new FPGS *BstN*I F	**gaagctgctggagcccgcactgt**	**23**	**35%**	**65%**	**76**
new FPGS *BstN*I R	**caataagctgggcctgagccagga**	**24**	**42%**	**58%**	**76**
FUT1 Northern Probe F	**ATGtggctccggagccatcgtc**	**22**	**36%**	**64%**	**72**
FUT1 Northern Probe T7 R	**ggatcctaatacgactcactatagggaggTCAAGGCTTAGCCAATGTCCAGAG**	**53**	**51%**	**49%**	**158**
FUT1 ex 1 F	**tgactggagttggcagccaagcc**	**23**	**39%**	**61%**	**74**
FUT1 ex 2 F	**ctaaccggtctctgcgtcttgatg**	**24**	**46%**	**54%**	**74**
FUT1 ex 4/5 R	**tgccgcgtcctcctgcactgcac**	**23**	**30%**	**70%**	**78**
FUT1 ex 7 F	**actgtggatctgccacctgcaag**	**23**	**43%**	**57%**	**72**
FUT1 ex 8 R	**ggcaggcagaagatggccactg**	**22**	**36%**	**64%**	**72**
FUT1 ex 7 bef TS F	**gggagggcggacgctgcgaga**	**21**	**24%**	**76%**	**74**
FUT1 ex 7 aft TS F	**tgtggcctgcctttgcggagggt**	**23**	**35%**	**65%**	**76**
FUT1 ex 7/8 R	**ggctccggagccacatggccga**	**22**	**27%**	**73%**	**76**
FUT1 *Mae*II F	**ctcacacagtgcaaccacaccatt**	**24**	**50%**	**50%**	**72**
FUT1 *Mae*II R	**ctagaactgcctgccagccatca**	**23**	**43%**	**57%**	**72**
FUT1 MS-SSCA ex 1 F	**ggggtgTtgggTtTTtggaaaaaT**	**24**	**58%**	**42%**	**68**
FUT1 MS-SSCA ex 1 R	**AAtctAaAcatctAaAcaAAAacaAatA**	**28**	**82%**	**18%**	**66**
FUT1 MS-SSCA ex 2 F	**ggTtgTTtaaaagggaagggTTTT**	**24**	**63%**	**38%**	**66**
FUT1 MS-SSCA ex 2 R	**AAAtAAAtAttcttcccctAtAaAAaA**	**27**	**81%**	**19%**	**64**
FUT1 MS-SSCA ex 2 anti F	**TTtgtgaggagaggagagattgtaag**	**26**	**58%**	**42%**	**74**
FUT1 MS-SSCA ex 2 anti R	**ctAacatatActAccttctctttctAc**	**27**	**67%**	**33%**	**72**
FUT1 MS-SSCA ex 2 anti R2	**AacatatActAccttctctttctActtcAA**	**30**	**70%**	**30%**	**78**
FUT1 MS-SSCA ex 7 F	**gggTggggTtTTaTTttTTagTTatT**	**26**	**69%**	**31%**	**68**
FUT1 MS-SSCA ex 7 R	**ActActtAcaAAtAAcaAatccacaA**	**26**	**73%**	**27%**	**66**

HSPA5 *BstN*I F	**ctagacctgcccttcgcctggttc**	**24**	**38%**	**63%**	**78**
HSPA5 *BstN*I R	**catccgcaaccccacttaccagga**	**24**	**42%**	**58%**	**76**
LCN2 *Mae*II F	**ctttattctgctgtccccatctcg**	**24**	**50%**	**50%**	**72**
LCN2 *Mae*II R	**gcacatgtttatttagcagacaaCg**	**25**	**60%**	**40%**	**70**
RT-PBGD x1 F **	**ctttccaagcggagccatgtctgg**	**24**	**42%**	**58%**	**76**
RT-PBGD x6/7 R **	**catgagggttttcccgcttgcaga**	**24**	**46%**	**54%**	**74**
POMT1 F	**ccagcgtaggagtcattgacaacaa**	**25**	**52%**	**48%**	**74**
POMT1 R	**actgtaacggttttgtcgagcttag**	**25**	**56%**	**44%**	**72**
PPP2R4 *Hinf*I F	**cttggggcccattgtatctggagag**	**25**	**44%**	**56%**	**78**
PPP2R4 *Hinf*I R	**aaagcctgggcttgggaggctgAc**	**24**	**38%**	**63%**	**78**
RALGDS *BsiE*I F	**cgtcttctatgccatgaactcGac**	**24**	**50%**	**50%**	**72**
RALGDS *BsiE*I R	**aatcccagcagcgggagaggttca**	**24**	**42%**	**58%**	**76**
SURF2 *Alu*I F	**tgggagcccacatccagtgatgag**	**24**	**42%**	**58%**	**76**
SURF2 *Alu*I R	**ggtacacggtcgtctctctccagc**	**24**	**38%**	**63%**	**78**
RT-SURF2 ex 4/5 F	**gacctgtacccacctgagctattcac**	**26**	**46%**	**54%**	**80**
SURF3 *BstX*I F	**aagagaccacctgtccttcgagca**	**24**	**46%**	**54%**	**74**
SURF3 *BstX*I R	**tcaacaggcaaacgcacctcgatg**	**24**	**46%**	**54%**	**74**
SURF4 *Alu*I F	**ctgtcgtctgcaccatagtaaatgc**	**25**	**52%**	**48%**	**74**
SURF4 *Alu*I R	**gagaacccatgtcctgaagccatAg**	**25**	**48%**	**52%**	**76**
SURF4 *Dde*I F	**acgccgaggctttggagtgaacag**	**24**	**42%**	**58%**	**76**
SURF4 *Dde*I R	**ctttgggtgtccctaaatgttctcac**	**26**	**54%**	**46%**	**76**

Capital letters that are double underlined denote mismatches introduced to create restriction enzyme sites. Capital letters for bisulfite primers denote where the cytosines are converted to thymines, that is, c -> T on forward primer but on reverse primer g -> A. For the primer FUT1 Northern Probe F, ATG denotes the translation initiation codon while the capital letters in the primer FUT1 Northern Probe T7 R denotes the T7 promoter sequence. * Primer designed by Ms. Tanya Sanders. ** Primers from Hussey *et al* (2001).

2.2 General Molecular Biology Methods

2.2.1 DNA Extraction from Peripheral Blood Mononuclear Cells

Blood samples were centrifuged at 400 x g for 10 min in a bench top centrifuge (Beckman TJ6). The platelet rich plasma was removed without disturbing the buffy coat. The red cells and buffy coat were diluted to twice the original volume with PBS in a 50 ml centrifuge tube. Lymphoprep, approximately 12 ml, was carefully layered under the diluted cells. The tube was then centrifuged for 25 min at 400 x g. The lymphocytes were collected from the interface with a pasteur pipette (no more than 4 - 5 ml), placed in a 10 ml centrifuge tube and PBS added to a final volume of 10 ml. The cells were counted, using a haemocytometer, before being pelleted at 300 x g for 10 min. The supernatant was removed and the pellet was resuspended in the residual PBS before adding 1 – 1.5 ml of PBS and transferring the contents to a 1.5 ml tube. The 1.5 ml tube was centrifuged in an Eppendorf centrifuge (Eppendorf 5415) at 13000 x g to pellet the cells. The supernatant was removed and the pellet (no more than 20 x 10^6 cells) was resuspended in 500 µl of TES plus 30 µl of Proteinase K. The tube was mixed by inversion before adding 30 µl of 20% SDS and incubating at 37°C overnight or longer. Following proteinase K digestion, an equal volume of 3 M NaCl was added, the solution mixed vigorously, and then placed on ice for 10 min to allow the protein to precipitate. The solution was spun at 13000 x g for 10 min and the supernatant was collected and placed in a 1.5 ml tube (500 µl maximum). Twice the volume of absolute ethanol was added and the contents were mixed by inversion. DNA was precipitated by spinning for 30 sec at 13000 x g. The pellet was washed with 70% ethanol before drying at room temperature for 1 hour. The pellet was redissolved in an appropriate amount of TE and stored at 4°C.

2.2.2 Preparation of Genomic DNA and RNA from Bone Marrow

For bone marrow (BM) samples frozen in nitrogen, the ampoule was thawed in a beaker of warm water (37°C) with shaking. Once thawed the sample was transferred to a 1.5 ml tube and the cells pelleted by centrifugation. The supernatant was removed and for RNA extraction no more than 10^7 cells were resuspended in 1 ml of Trizol and then processed as per manufacturer's instructions. For DNA extraction 5 – 20 x 10^6 cells were resuspended in 500 µl of TES. Proteinase K, 30 µl, was added and the solution was mixed by inversion before 30 µl of 20% SDS was added. The final solution was mixed by inversion and incubated at 37°C overnight or longer. The DNA was then obtained as per 2.2.1.

2.2.3 Preparation of DNA and RNA from Peripheral Blood Stem Cells (PBSCs)

The ampoule of frozen cells was thawed rapidly by shaking in a beaker of warm water. Once thawed the cells were transferred to a 10 ml EDTA tube. 10 ml of PBS/10% FBS solution was added in a drop wise fashion. The cells were counted and 10^7 cells were immediately lysed in Trizol and frozen at –80°C until needed. The remaining cells were placed into a 50 ml centrifuge tube and the volume made up to 35 ml with PBS. 12 – 13 ml of lymphoprep was then layered underneath the cells before they were centrifuged at 400 x g for 25 min. The cells at the interface were collected into PBS, counted and aliquots were taken to make RNA or DNA as per 2.2.1 and 2.2.2.

2.2.4 Isolation of DNA from Frozen Blood

A spatula was used to scrape some frozen blood, no more than 250 µl, into a 1.5 ml tube containing DNAzol BD. The DNA was extracted using a scaled-down version based on the manufacturer's instructions. DNA was also obtained from fresh blood using DNAzol BD.

2.2.5 Quantitation of DNA and RNA

To quantitate DNA, 5 µl of the DNA sample was added to 995 µl of MQW and the optical density (OD) at 260 nm was determined using the GeneQuant DNA calculator (Pharmacia). For double stranded DNA, each OD unit is equivalent to 50 µg of DNA in a 1 ml solution. For quantitation of RNA, 1 µl of RNA was diluted in 49 µl of UPW and placed in a microcell and the reading was taken using a Beckman DU-600.

2.2.6 DNA Precipitation

DNA was precipitated by the addition of 10% 3 M sodium acetate (pH 5.2) and 2.5 volumes of absolute ethanol. The solution was placed at – 80°C for one hour before being centrifuged at 13000 x g for 30 min. The supernatant was decanted and the DNA pellet was washed with 70% ethanol. The pellet was dried at room temperature before redissolving in TE (for genomic DNA) or in UPW (for PCR products).

2.3 Cell Culture Methods

2.3.1 Thawing Cells

Human leukaemia cell lines CCRF-CEM, EM2, HEL, HL-60, K562, KCL-22, Jurkat and Raji were grown in RPMI 1640 supplemented with 10% FBS, penicillin and streptomycin. Frozen cells were thawed rapidly in a 37°C water bath and the outside of the ampoule was sterilised with 70% ethanol before the contents were transferred into a sterile 10 ml centrifuge tube. 10 ml of cold growth medium RPMI 1640 was then added in a drop wise fashion. Cells were then pelleted by spinning at 200 x g for 10 min. The supernatant was removed and the cells were resuspended in 1 ml of growth medium that was then distributed between two 25 cm^2 tissue culture flasks. After 24 hours, the growth medium was changed to remove the dead cells and any residual DMSO.

2.3.2 Maintenance and Subculturing of Cells

Cells were maintained in a humid atmosphere containing 5% CO_2 at 37°C. When cells grew to confluence, the medium was discarded and the cells washed twice with PBS. The cells were centrifuged at 200 x g for 10 min, the supernatant was discarded, and the cell pellets were resuspended in 1 ml of growth medium that was then aliquotted into new tissue culture flasks.

2.3.3 Viability Tests

Ninety microlitres of a 0.4% solution of trypan blue was mixed with 10 µl of a cell suspension. This mixture was added to a haemocytometer and allowed to settle for at least 10 seconds before counting. Viable cells, those that excluded the trypan blue dye, were counted using the haemocytometer. The total viable cell count was calculated using the following formula: the number of viable cells per ml = haemocytometer count x 10^4 x 10.

2.3.4 Freezing Cells

The cells were collected as described in 2.3.2 but the medium was reserved to freeze the cells in. The pellet of cells was resuspended in the conditioned medium (approximately 1 – 1.5 ml per ampoule to be frozen) with 10% DMSO. The cells, no more than 1.5 ml, were distributed into cryopreservation tubes (Nunc) and placed at –80°C overnight before being transferred to liquid nitrogen.

2.3.5 5-Aza-2'-Deoxycytidine Treatments

10^6 leukaemia cells were seeded in flasks and incubated at 37°C with 5% CO_2. Cells were serum starved in medium supplemented with 0.1% FBS for 48 h prior to treatment. Following this, the medium was changed to include 10% FBS and cells were treated with 5-aza-2'-deoxycytidine (1 µM, 2 µM or vehicle – UPW). Fresh 5-aza-2'-deoxycytidine was added

daily for 3 days. Twenty-four hours after the final treatment the media was removed. Cells were then washed with PBS and fresh media was added to each flask. Cells were allowed to recover for 24 h and then were harvested at 48, 72 and 96 h post treatment. RNA was isolated using Trizol and DNA was extracted as per 2.2.1. If there were less than 10^4 cells, the cells were lysed with 0.3% Nonidet P40, 20 U RNAsin, 0.01 M DTT and the supernatant was placed in Trizol for RNA extraction while the cell nuclei were directly bisulphite modified.

2.4 Making Spherical Red Blood Cells and Flow Cytometric Methods

2.4.1 Making Formalin Fixed Spherical Red Blood Cells

The method used was adapted from Langlois *et al* (1985). All solutions were filtered through 0.45 μm membranes to remove particulate matter that might interfere with flow cytometry. One hundred microlitres of whole blood, or 50 μl of packed red blood cells, was added to 1 ml of routine staining buffer (RSB) (see section 2.1). After 60 sec, 10 ml of a second solution comprising 3% formaldehyde and 0.001% SDS in RSB was added. The tube was then incubated on its side for 90 min at room temperature after which 0.7 ml of formaldehyde was added and the tube laid on its side for cell fixation overnight at room temperature. After resuspension, the spheres were pelleted and then washed at 200 x g in spherical staining buffer (SSB) (see section 2.1). Red blood cell counts were performed on a Coulter STKS with 0.3 x 10^6 red blood cells stained with the appropriate antibodies (blood groups were determined serologically by the Blood Transfusion Service of The Queen Elizabeth Hospital).

2.4.2 Labelling of Spherical Red Blood Cells

For A blood group individuals 300,000 spheres were incubated with 20 μl of lyophilised Helix Pomatia lectin conjugated to biotin, in a final volume of 100 μl SSB, for 15 min at room temperature. They were then washed with SSB before the addition of 5 μl of Streptavidin Phycoerythrin PE, the secondary reagent for the biotinylated Helix. After a 15

min incubation at room temperature and a SSB wash, 10 µl of Ulex Europaeus lectin conjugated to FITC was added. The B antigen was detected using 10 µl of a mouse monoclonal antibody (Novaclone anti-B blood grouping reagent) and 50 µl of a goat F(ab')2 anti-mouse IgG (H+L) PE conjugate. This was followed with incubation with Ulex. For AB blood group antigens, the B versus A antigen staining was detected using Novaclone for the B antigen but 15 µl of Helix Pomatia lectin conjugated to FITCd for the A antigen. Cells were washed two times with SSB before flow analysis. Lectin and antibody concentrations were optimised using serial dilutions to find optimal fluorescence with minimal agglutination for each batch.

Normal samples of relevant blood groups were included in each run of patient samples as controls for variations in the fluorescence that was seen as the lectin stock solutions, particularly the Ulex, aged. The normal samples were used as references to determine if there was a change in the patient sample from the normal expected pattern.

2.4.3 Flow Cytometric Analysis

Flow cytometric analysis was performed on a Coulter Epics Profile II. 30,000 gated events were collected from a sample volume of 100 µl with a flow rate of 30 µl/min. Log fluorescence data was gated on a linear forward scatter versus linear side scatter dot plot.

2.5 Polymerase Chain Reaction Methods

2.5.1 Primer Design

Primers were designed using Amplify v1.2 (Bill Engels) and DNA Strider™ 1.1 (Commissariat ã l'Energie Atomique). Primers for single nucleotide polymorphisms (SNPs) were designed to give a product that could be restriction enzyme digested for genotyping. If the SNP did not create a restriction enzyme site, then the primer sequence was modified to

create one. Primers for RT-PCR spanned exon – exon junctions so that genomic DNA was not amplified. If this was not possible then the RNA was DNase I treated to destroy contaminating DNA prior to reverse transcription (see section 2.5.6).

2.5.2 PCR Reaction and Conditions

PCR reactions were carried out in a final volume of 50 µl, to which 100 ng of genomic DNA, 5 µl of bisulfite modified DNA or 2 µl of cDNA was added. Each PCR reaction contained 100 ng of each oligonucleotide primer, 0.2 mM dNTPs, 0.5 U of AmpliTaq Gold® Polymerase (Applied Biosystems) or 0.5 U of HotStarTaq DNA Polymerase (Qiagen), 1X PCR buffer (10 mM Tris-HCl, 50 mM KCl) and 1.5-2.5 mM $MgCl_2$. The volume was made up to 50 µl by the addition of UPW. DMSO was used for difficult PCRs especially those amplifying CG rich templates. Another PCR additive that was used for difficult PCRs was betaine (0.9 – 1.5 M). PCRs were carried out with an initial denaturation step of 94°C for 9 min for AmpliTaq Gold® Polymerase or 95 °C for 15 min for HotStarTaq DNA Polymerase. The PCR reaction was then cycled on a PTC-100™ Programmable Thermal Cycler (MJ Research Inc) for 30-45 cycles of 96°C for 30 sec, annealing at 60 – 65°C for 1 min (or 60-50°C for PCRs involving bisulphite modified DNA) and extension at 72°C for 1 min. The annealing temperature for the PCR was determined by calculating the melting temperature of each primer, assigning 4°C to each G or C and 2°C to each A or T (see Table 2.1.1), then subtracting 5°C from this temperature.

2.5.3 Purification of PCR Products

Wizard™ PCR Preps DNA Purification System (Promega): DNA was purified directly from PCR reactions following the manufacturer's instructions.

From low melting point agarose gel: The PCR was run out on a low melting point gel with 0.25 µg/ml of ethidium bromide. The band of interest was then cut out using a sterile scalpel

blade and placed in a 1.5 ml tube. This was then incubated at 65°C for 5 - 10 min to melt the agarose. Wizard™ prep was then performed without the use of the direct purification buffer.

Concert™ Rapid PCR Purification System (Gibco-BRL)

Used as per manufacturer's instructions for purifying PCR products directly from the PCR reaction mix.

2.5.4 Restriction Enzyme Digestion of PCR Products

PCR products (5 – 10 µl) were digested with 5 – 10 U of enzyme in a total volume of 20 µl with 1X buffer, 1X BSA and the remaining volume made up with UPW. PCR products were digested at the appropriate temperature for 3 – 16 hours.

2.5.5 Reverse Transcription

A RNA mix containing 1 µg of RNA, 2 µl of Random Hexamers (250 ng/µl) in a final volume of 20 µl by addition of UPW was denatured at 70°C for 10 min before being placed on ice for 3 min. To this RNA mix, 20 µl of the RT (reverse transcriptase) mix was added. The RT mix contained 8 µl of 5X First Strand Buffer, 4 µl of 0.1 M DTT, 2 µl of 10 mM dNTPs, 5 µl of UPW and 1 µl of MMLV Reverse Transcriptase (200 U/µl). After the addition of the RT mix the tubes were incubated at 37°C for 1 – 2 hours.

2.5.6 DNase I Treatment of RNA

Two µg of RNA, 2.5 µl of 10X DNase I buffer and 1 µl (1 unit) of DNase I (DNA-*free*™ from Ambion) were made up in a final volume of 25 µl with UPW. The sample was incubated at 37°C for 30 - 60 min followed by addition of 5 µl of DNase I Inactivation Buffer. The sample was mixed well before incubation at room temperature for 2 min. The sample was then centrifuged at 10,000 x g for 1 min, the supernatant was removed and reverse transcription performed as per 2.5.5.

2.5.7 Agarose Gel Electrophoresis

Agarose gels were prepared by melting the appropriate amount of DNA Grade Agarose powder (Progen), depending on whether a 1.5% or 3% (w/v) gel was required, in 0.5X TBE and ethidium bromide was added (final concentration 0.1 μg/ml). DNA samples were loaded onto the gel after mixing with 6X loading buffer. The gels were run at 100 V in 0.5X TBE buffer which also contained 0.1 μg/ml ethidium bromide. DNA was visualised by placing gels under UV (254 nm) light.

2.5.8 Sequencing of PCR Products

Sequencing reactions were carried out in a final volume of 20 μl with 30 – 90 ng of template DNA, 25 ng of primer and 8 μl of Big Dye™ (Applied Biosystems) before being placed on a PTC-100 for 25 cycles of 96°C for 30s, 50°C for 30s, 60°C for 4 min. The sequencing reactions were precipitated with 3 M sodium acetate (pH 5.2) and 100% EtOH and sent to the Department of Molecular Pathology, IMVS, Adelaide for analysis.

2.6 Methylation Analysis Methods

2.6.1 Bisulphite Modification of DNA [adapted from Herman *et al* (1994) & Raizis *et al* (1995)]. DNA (1 – 2 μg) was dissolved in a final volume of 18 μl with MQW. Two microlitres of 3 M NaOH, freshly prepared before each bisulphite modification reaction, was added to give a final volume of 20 μl and a final concentration of 0.3 M NaOH. The solution was incubated at 37°C for 15 min. During this time, a 4.8 M sodium bisulphite solution and a 100 mM hydroquinone solution were freshly prepared as described in 2.1. After the 15 min incubation, 278 μl of 4.8 M sodium bisulphite and 2 μl of 100 mM hydroquinone were added to give a final concentration of 4.45 M sodium bisulphite and 0.67 mM hydroquinone. The samples were mixed, centrifuged and then incubated on a PTC-100 thermal cycler for 20 cycles of 95°C for 30s and 55°C for 15 min. The bisulfite modified DNA was purified using the

Wizard™ PCR Prep DNA Purification System (Promega) or the Concert™ Rapid PCR Purification System (Invitrogen) as described in 2.5.3. The DNA was eluted using 50 µl of preheated TE (at 65 – 70°C) before the addition of 5.5 µl of 3 M NaOH to the eluted DNA. Incubation was performed at 37°C for 15 min before precipitating the DNA as described in section 2.2.6. The precipitated pellet was redissolved in 50 µl of TE and used immediately, or stored at –20°C.

2.6.2 Primer Design

Bisulphite modified primers were designed to avoid CpGs and CpNpG as the primers were designed with the intention of amplifying bisulphite modified DNA only, but they did not differentiate from methylated or unmethylated DNA. If a CpG could not be avoided it was placed as far as possible to the 5' end of the primer were it would cause the least instability. The antisense primer was designed first and was designed from the bisulphite modified sense strand. The sense primer binds to the strand synthesised by the antisense primer. The primer was designed to have a T (that is, a cytosine that became a thymine) at the 3' end of the primer to ensure that only the modified DNA would be amplified.

2.6.3 PCR of Bisulphite Modified DNA and MDE Gel Electrophoresis

5 µl of bisulphite modified DNA was used in the PCR reaction as described in section 2.5.2. The PCR product was denatured by adding 6 µl of the sample to 14 µl of DDS (see section 2.1) and heated at 95°C for 5 min on a PTC-100™ Programmable Thermal Cycler. The samples were then quickly chilled on ice and then loaded onto a 0.5X or 0.75X MDE gel. The gel was run in 0.6X TBE with 0.5X MDE gels run at 280 V for 4-5 hours and 0.75X MDE gels run at 130 V for 20 hours. The gel was stained with SYBR® Gold, which was diluted 10,000 fold in 0.6X TBE, covered with the solution and placed on a rocking platform in the dark for 30 – 40 min. The gel was then visualised on UV transilluminator 254 nm.

2.6.4 Band Stabs and Sequencing

Variant or extra bands were band stabbed with a yellow tip and placed in 10 µl of UPW and left at room temperature for at least 1 hour. Bandstabs were reamplified by using 1 µl of the bandstab DNA solution with the original PCR conditions. The PCR product was then purified as per 2.5.3 and sequenced as per 2.5.8.

2.7 Northern Blotting Methods

2.7.1 Generation of Probes

Primers were designed so that they had T7 promoter sequence (lower case letters in Table 2.1.1) on the reverse primer in an orientation that would create antisense RNA transcripts from the PCR DNA template. For the *FUT1* gene, the antisense RNA transcripts generated consisted of the entire protein coding region of the gene, 1127 bp. After the DNA template was generated using the primers FUT1 Northern Probe F and FUT1 Northern Probe T7 R, 500 ng of DNA was used in a T7 transcription reaction using the Strip-Ez™ RNA T7 kit (Ambion). The manufacturer's instructions were followed except that DIG-11-UTP (Roche) was used to generate antisense RNA transcripts that were DIG labelled. Five microlitres of the reaction was used directly and added to the Northern membrane, which had prehybridised at 68°C in ULTRAhyb™, the rest of the reaction was stored at –20°C until required.

2.7.2 RNA Gel Electrophoresis and Blotting

One to 4 µg of RNA was electrophoresed through a denaturing 1% agarose/1.2 M formaldehyde gel. Prior to electrophoresis, all RNA samples were brought to the same volume using UPW and an equal volume of 2X LB was added. RNA was denatured at 70°C for 10 min, quickly chilled on ice and then loaded onto the gel. Gels were run in 1X MOPS at 80 V until the bromophenol blue had reached the bottom of the gel. Following electrophoresis, the gel was downward blotted in 5X SSC/10 mM NaOH onto Brightstar plus

membrane (Ambion) according to the manufacturer's protocol. The RNA was crosslinked to the membrane with 120,000 microjoules/cm^2 of UV irradiation using UV Stratalinker 1800 (BioRad).

2.7.3 Hybridisation

The membrane was initially pre-hybridised at 68°C in ULTRAhyb™ and after the addition of the probe, the membrane was hybridised overnight at 68°C. The membrane was then washed in 2X SSC, 0.1% SDS twice for 5 min at 68°C and then in 0.1X SSC, 0.1% SDS twice for 15 min at 68°C. The membrane then underwent detection as per the maufacturer's instructions and was exposed to autoradiograph film (Hyperfilm, Amersham) using Dupont Cronex® Hi-Plus intensifying screens until the appropriate exposure time was achieved (15 – 60 min).

2.8 Southern Blotting Methods

2.8.1 Restriction Digests

10 µg of DNA was digested with 100 U of *Taq*I (New England Biolabs) in a total volume of 200 µl which consisted of 1X buffer (supplied by the manufacturer), 1X BSA (supplied by the manufacturer) and MQW. The final mixture was placed at 65°C overnight. The *Taq*I digested DNA was then precipitated as outlined in 2.2.6 with the pellet redissolved in MQW. For the second digest, 100 U of the methylation sensitive enzymes were used: *Cfo*I (Roche) and *Hpa*II (New England Biolabs). These incubations were carried out at 37°C overnight.

2.8.2 Gel Electrophoresis

The restriction digests were precipitated as outlined in 2.2.6 and before loading the restricted fragments were heated to 56°C for 5 min to disrupt base pairing between cohesive termini of restriction fragments. DNA was electrophoresed through a 1.5%, 300 ml, 20 x 20 cm agarose gel containing ethidium bromide (0.25 µg/ml). The gel was run at 100 V until the

bromophenol blue dye had migrated approximately two-thirds of the way down the gel (approximately 5 hours). The DNA was visualised under UV (254 nm) and photographed using Polaroid 667 film.

2.8.3 DNA Transfer to Nylon Membrane

DNA was transferred to a Hybond-N^+ nylon membrane (Amersham) overnight in a transfer solution consisting of 0.4 M NaOH. After transfer the membrane was washed with 2X SSC for a few minutes to neutralise the NaOH and remove any gel pieces. The DNA was crosslinked to the membrane as per 2.7.2.

2.8.4 Hybridisation

Membranes were prehybridised for at least two hours at 42 – 52°C in 20 ml of a solution containing 50% deionised formamide, 1 M NaCl, 10% dextran sulphate, 1% SDS, and 0.2 mg/ml of sonicated and denatured salmon sperm DNA. The final volume was made up to 20 ml by the addition of MQW. The radioactive probe was prepared using the GIGAprime DNA labelling kit (Geneworks), according to the manufacturers instructions. Template DNA (25 ng) was used for the probe and 5 µl of 10 mCi/ml (50 µCi) of radioactive α-^{32}P dCTP (Geneworks) was used in a total volume of 25 µl. The hybridisation mixture was heated at 100°C for 5 min to denature the labelled DNA before being added to the prehybridisation mix. Hybridisation was carried out at 42 – 52°C overnight.

2.8.5 Washing the Membranes

Membranes were washed at room temperature in 2X SSC. This was followed by a 60 min wash at 65°C in 1X SSC, 0.1% SDS and finally a 60 min wash at 65°C in 0.1X SSC, 0.1% SDS. The membrane was wrapped in plastic cling film, exposed to autoradiograph film (Hyperfilm, Amersham) using Dupont Cronex® Hi-Plus intensifying screens and placed at -80°C for the appropriate time.

CHAPTER 3

Flow Cytometric Analysis of ABH Antigens

3.1 Introduction

The relatively infrequent reporting of the loss of ABH antigens under represents the actual frequency of this loss and the probable significance of this loss as a marker of the recurrent mechanisms involved in the development of haematological malignancy. Serological techniques detect loss of ABH antigens by a mixed field reaction (a mixture of agglutinated and unagglutinated cells) but are neither reliable nor quantitative. It has been found that blood with as many as 50% of cells showing loss of A antigens can be scored as normal in transfusion laboratories (O'Keefe, PhD thesis, 1995).

Therefore, a sensitive flow cytometric method was necessary to detect alterations in ABH antigens. This flow cytometric technique can quantitatively and sensitively assess alterations of ABH antigen expression on the surface of red blood cells (RBCs) using antibodies specific to the A, B and H antigens. As such, alterations in ABH antigen expression of cells from patients with haematological malignancies can be assessed.

Flow cytometers are laser-based instruments used to characterise the physical and biochemical properties of cells (reviewed in Jaroszeski and Radcliff, 1999). Flow cytometers scan cells as they flow in a liquid stream of droplets past a laser beam. The cells scatter light and emit fluorescence as the laser strikes the cells. Light scattering and fluorescence is measured for each particle that passes the laser beam. Fluorescence is introduced by addition of an antibody conjugated to a fluorescent dye which binds to cell surface proteins (Wedemeyer and Pötter, 2001).

Various cell properties can be deduced from the scattered light and emitted fluorescence of a cell. Light scattered in the same direction as the laser beam is forward scatter (FSC), while light scattered at 90° relative to the laser beam is termed side scatter (SSC). FSC is due to

diffraction and provides physical information about the cell, such as cell size. SSC results from refraction and reflection of light and provides information about the cell surface. The FSC and SSC correlation allows for the discrimination of cell subpopulations in a heterogeneous sample and exclusion of cellular debris from the analysis (reviewed in Jaroszeski and Radcliff, 1999).

Immunohaematologists have been analysing RBCs using blood group antibodies ever since Landsteiner discovered ABO blood groups in 1901. Flow cytometric analysis was first applied to RBCs by van der Meulen *et al* (1980) and since then numerous flow cytometric techniques have been adapted to study various RBC group antigens. Agglutination of RBCs presents difficulties for flow cytometry as this technique relies on single cell analysis. ABO antibodies are especially problematic, due to their strong agglutination properties. Agglutination problems have been overcome in recent times by disrupting the agglutinates by vortexing the RBC sample before flow cytometry (Sharon and Fibach, 1991; Fibach and Sharon, 1994). Other techniques use chemicals like formaldehyde to fix the cells, allowing antibody binding but decreasing agglutination (Berneman *et al*, 1991). Problems with chemical fixation include haemolysis, decrease in antigen strength, non-specific antibody uptake and increased autofluorescence (reviewed in Garratty and Arndt, 1999).

The flow cytometry methodology presented in this chapter was developed by Belinda Farmer and Alex Dobrovic. At the time I began my PhD, two-thirds of the patients and normal controls had been analysed. This methodology was used to analyse more patients and normal individuals. In addition, all of the patients were genotyped, which was critical to the analysis, by a methodology developed during this PhD study (Chapter 4). The results and analysis of the flow cytometric profiles are the basis of the work presented in this chapter.

3.2 Results

3.2.1 Normal Individuals

The first aim was to ascertain the patterns expected in normal individuals for the various blood groups. This would allow for easy identification of aberrant patterns when analysing the patient flow profiles.

3.2.1.1 Flow Patterns in O and Bombay Individuals

For normal blood group O individuals (51 examined) only one bright Ulex (H antigen expression) peak was seen (Figure 3.1). No staining with the Helix Pomatia lectin (A antigen) or the anti-B antibody was seen. The O blood group has the strongest Ulex (H antigen) staining of all the blood groups, as expected, since all the H sites are unconverted because the O allele encodes an inactive transferase that is unable to convert the H antigen. Bombay individuals have no A, B or H antigens, irrespective of their blood group genotype, since they are homozygous for the inactive *h* allele and thus have no H antigen besides some minimal background staining (see Figure 3.1). As a result, a Bombay individual's spherical RBCs is similar to unstained spherical RBCs, as expected.

3.2.1.2 Flow Patterns in B Individuals

For blood group B individuals (35 examined), there was heterogeneous staining with the B antibody, with most cells expressing high amounts of B antigens. However, a small number of cells had almost no B antigen (Figure 3.1). There was a population of RBCs, 3 – 10%, that resembled O RBCs by virtue of this lack of labelling with anti-B, yet they had similar labelling with Ulex (H) as O RBCs. Individuals of blood group B show considerable variation for Ulex staining, indicating that not all the H sites have been converted to B antigens by the B transferase. In comparison to the A blood groups, the B blood group individuals have a higher proportion of cells that label with Ulex (H). This suggests that the B transferase is less efficient than the A transferase in converting H antigens into A or B antigens.

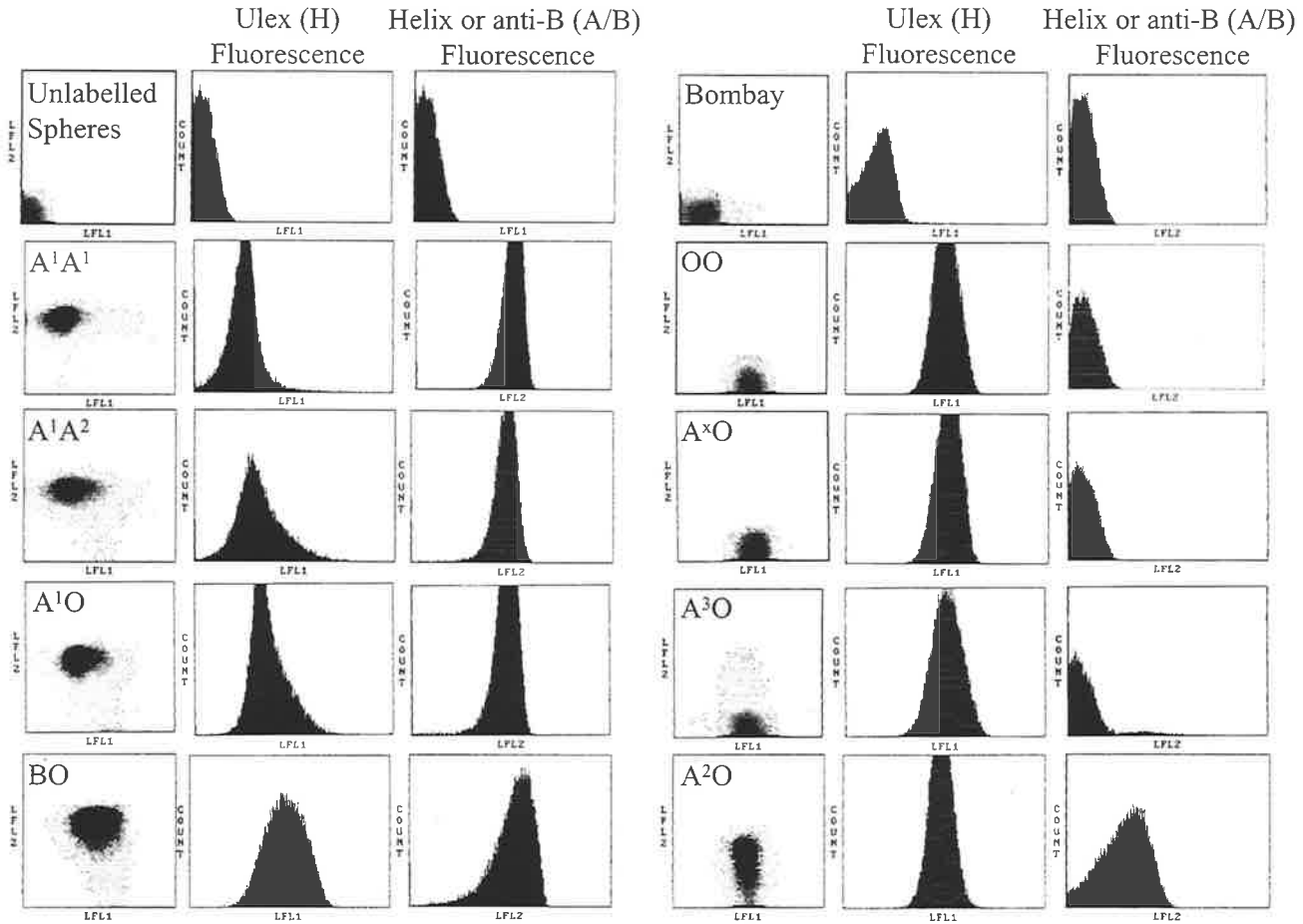


Figure 3.1. Dot plots and ABH antigen expression on red blood cells from healthy individuals as related to their genotypes. Typical results are presented for red blood cell spheres from 8 ABO genotypes as well as a Bombay individual and a control of unstained spheres. The first panel of each pair shows the 2-colour dot plot; the second panel shows the FITC (Ulex) fluorescence only. The Bombay and OO samples were stained with Ulex only whereas the A samples were stained with Helix and Ulex, and the B samples with anti-B and Ulex. LFL indicates log of fluorescence. The LFL1 axis shows Ulex (H) fluorescence and the LFL2 axis shows Helix (A) or anti-B fluorescence as appropriate.

3.2.1.3 Flow Patterns in A Individuals

For A blood group individuals several patterns of Helix-Ulex staining were observed with the pattern dependent on the individual's genotype. The A^2 pattern (8 examined, all presumed to be A^2O since no material was available for genotyping) was different from the A^1 pattern and showed more heterogenous staining with Helix (A antigen). This was expected as it is known that the A transferase encoded by the A^2 allele is less efficient at converting the H antigen than the A^1 transferase (Mollison *et al*, 1993). The A^2 pattern showed less heterogenous staining with Ulex (H) than the B pattern. Nearly 50% of cells had moderate to no staining with Helix, though all RBCs had a similar level of H (Ulex) expression (Figure 3.1). The Ulex peak closely resembled that for an O individual indicating many unconverted H sites. This is in accordance with reports that the A^2 transferase is less efficient than the B transferase, which is less efficient than the A^1 transferase, at converting the H antigen (Mollison *et al*, 1993).

When the flow cytometric analysis was first performed, differences in the flow profiles within the A^1 blood group normal individuals (34 examined) were not apparent. Using the *ABO* genotyping method (see Chapter 4), we were able to ascertain that different genotypes presented with different flow profiles. The profile of an A^1A^1 individual was different from an A^1O or A^1A^2 individual (Figure 3.1). The A^1A^1 flow profile showed the least amount of Ulex staining and hence the least amount of unconverted H antigen when compared to the other A^1 genotypes, that is A^1O and A^1A^2. The Ulex (H antigen) population of the A^1A^1 individual is to the left of the A^1O Ulex fluorescence peak indicating that the A^1A^1 population has less H antigen (Figure 3.2). The A^1A^2 Ulex fluorescence lies in between the A^1A^1 and A^1O fluorescence peaks and is more heterogeneous. This indicates that the amount of H antigen on the RBCs of the A^1 blood group is in the order $A^1O > A^1A^2 > A^1A^1$ (Figure 3.2). This is not surprising since an A^1A^1 individual would be expected to produce more A^1 transferase as a

result of the two A^1 alleles producing more A^1 transferase which results in the conversion of more H antigens. In the attempts to find the various A^1 genotypes, 48 A^1 blood group samples were genotyped. Forty-one were A^1O genotype, 6 were A^1A^1 for which the flow profiles were all the same, and only one A^1A^2 sample was found.

3.2.1.4 Flow Patterns in Weak A Individuals

Two weak A blood groups show different flow profiles when compared to the other A blood groups. These are the rare A^3 and A^x blood groups. The A^3 pattern (two examined, see Figure 3.1) resembled that of an A^2 pattern in terms of the very small population of cells with the heterogenous Helix staining. Some cells (less than 5%) had a considerable amount of A antigen while the majority of cells had no A staining. The majority of the A^3 cells resembled that of an O blood group, with almost the same amount of Ulex (H) as the *OO* sample, and like the A^2O pattern, the A^3 cells all have a similar amount of H regardless of the tiny population that has a considerable amount of A fluorescence. The A^x pattern (four examined, see Figure 3.1) was indistinguishable from an O pattern and a 50:50 mixture of RBC spheres from an O individual and an A^x individual showed only a single Ulex peak, indicating that the vast majority of H residues were unconverted (see Figure 3.3).

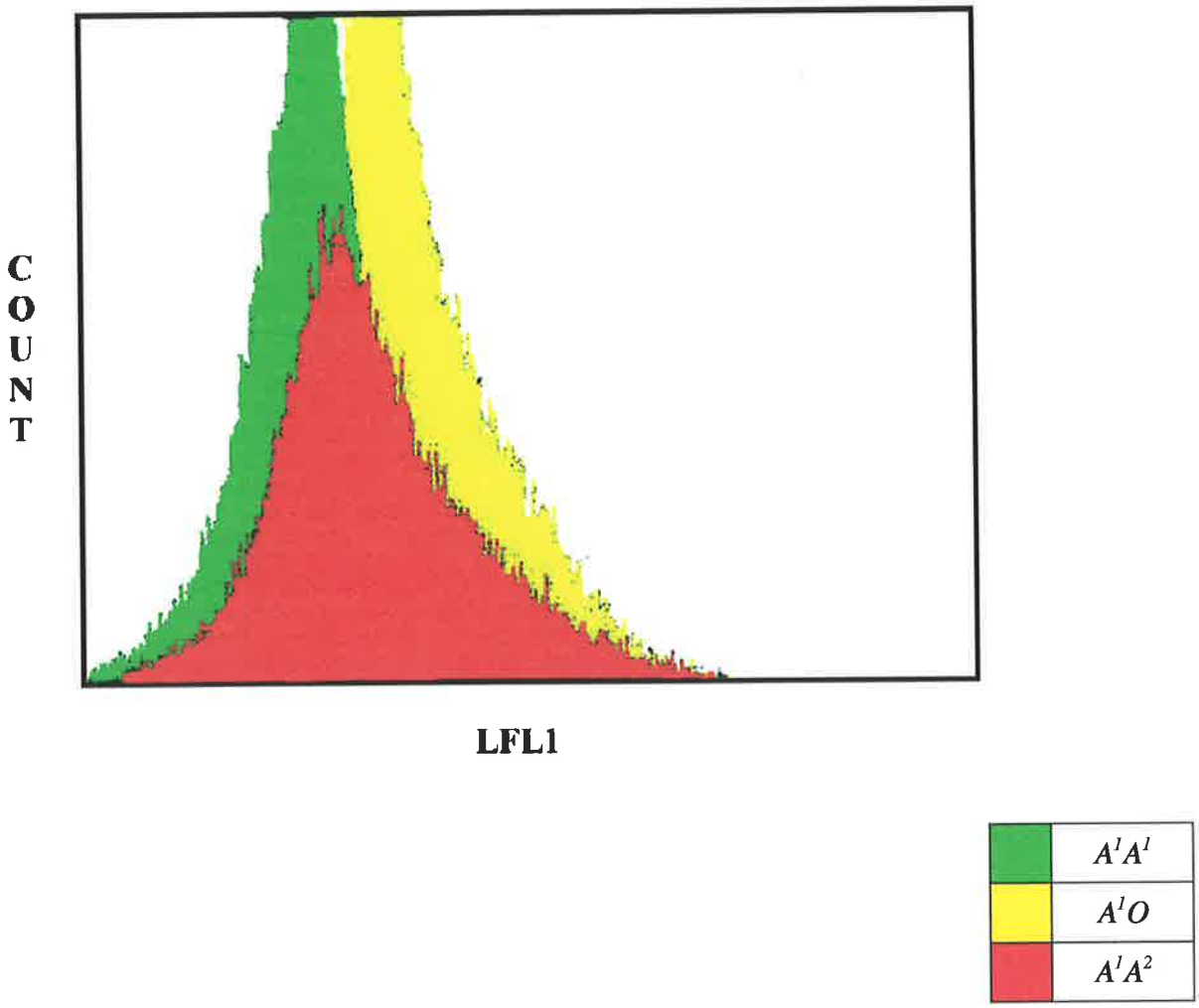


Figure 3.2. Ulex overlays from A^1 blood group genotypes.

The Ulex fluorescence peaks of the three A^1 genotypes presented were all generated on the same day with the same amount and batch of Ulex. This diagram shows the variation in the amount of unconverted H in three different A^1 blood group genotypes.

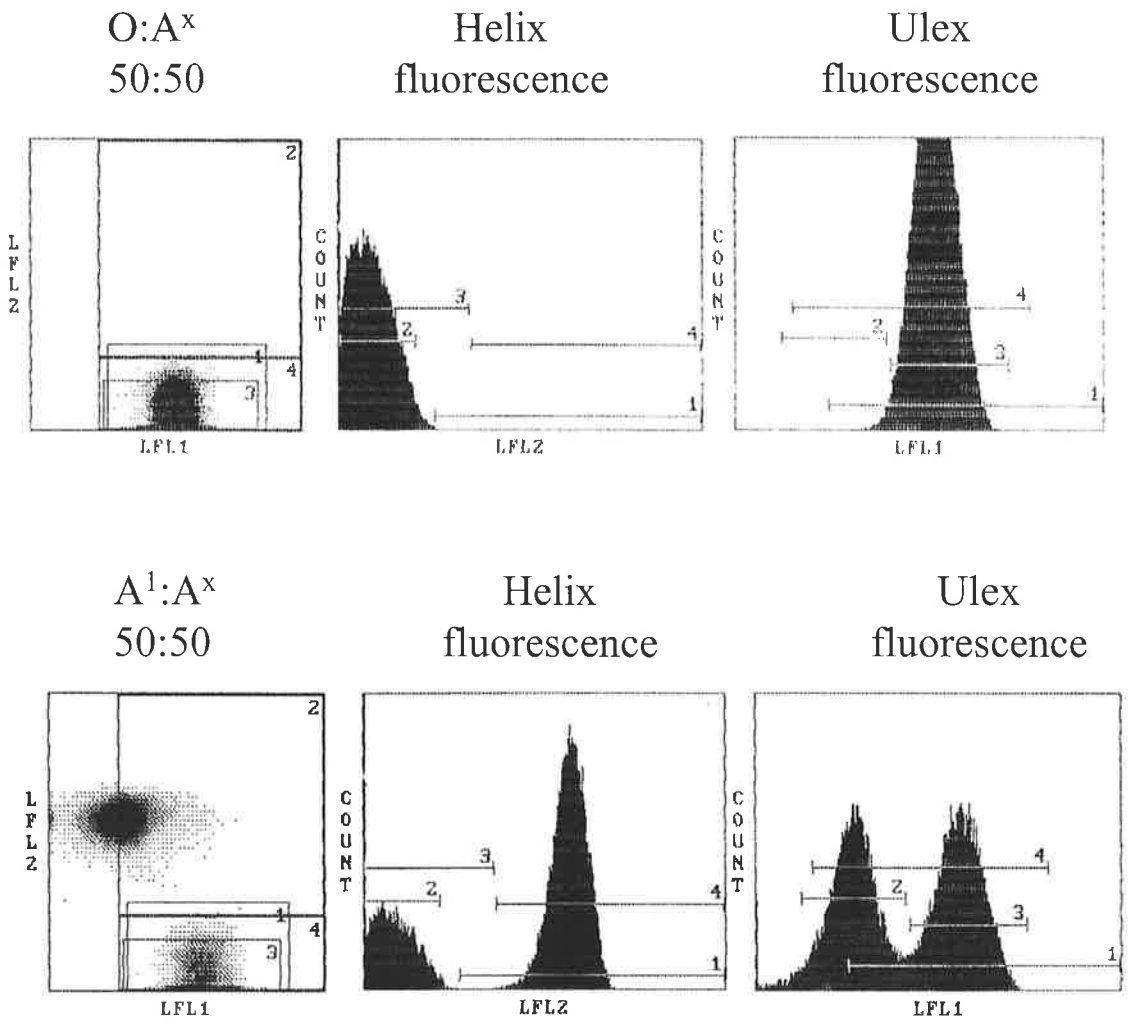


Figure 3.3. Mixtures of 50:50 O:A^x red blood cells and 50:50 A^1:A^x red blood cells. For the O:A^x mixture the two red blood cell populations are indistinguishable as only one population is visible. The A^1:A^x mixture clearly indicates the normal A^1 pattern and that the A^x population resembles an O population.

3.2.1.5 Flow Patterns in AB Individuals

The AB blood groups (29 A^1B and 15 A^2B examined) also produced distinctive flow profiles. The two colour flow cytometer used required the analysis of three combinations of antibody staining: Helix (A) vs Ulex (H), B vs Ulex and B vs Helix. The staining patterns for the AB blood groups were quite complex. For an A^1B individual the Helix – Ulex staining illustrates that there is a considerable amount of A antigen and little H antigen, with the RBC population very close to the LFL2 axis (Figure 3.4). This is not surprising as A^1B individuals produce A^1 and B transferases that convert the H antigen to A^1 or B antigens. The A^1B Helix – Ulex profile is similar to the A^1A^1 profile as these genotypes have two alleles that encode for transferases, hence converting more H antigens, unlike the A^1O genotype which has only one allele that encodes an active transferase. The B – Ulex pattern confirms that there is little H, as per the Helix – Ulex staining. However, unlike the Helix – Ulex staining which shows that all cells have the same amount of A, the B – Ulex staining has cells with variations in the level of B staining ranging from moderate to none. The B – Helix staining confirms that all the cells have approximately the same amount of A antigens but that the amount of B antigens varies from none to high.

The Helix – Ulex staining of an A^2B individual resembles that of an A^3 individual and not an A^2 individual (see Figure 3.1). There is a major O-like population with Ulex staining (H antigen positive) but no Helix staining (A antigen negative) and then a smaller A^2-like population that contains cells with variable A antigen expression. The B – Ulex staining for an A^2B individual resembles that of a *BO* individual. The majority of cells have the same amount of B antigen expression but a tiny population of cells have a heterogenous pattern of B antigen expression ranging from cells with no B antigen but similar amounts of H as for an *OO* population to cells with high B antigen expression. For an A^2B individual, the cells have heterogenous amounts of B antigen but when compared to an A^1B pattern there are more cells

with high B antigen expression (compare BUx for A^1B and A^2B Figure 3.4). Also, for an A^1B individual, there is minimal Ulex (H) staining when compared to an A^2B individual which has a major population of cells with similar amounts of H as for an *OO* population (compare HxUx for A^1B and A^2B Figure 3.4). In the A^2B pattern, the A antigen is very much diminished compared to the A^1B pattern.

It has previously been reported that the A^1 transferase is more efficient than the B transferase, which is more efficient again than the A^2 transferase (Mollison *et al*, 1993). Hence, the amount of H antigen on an O RBC is highest but decreases for a B RBC and finally the A^1 blood group RBCs have the lowest amount of H antigens. This relationship could be demonstrated using the flow cytometry method described here. Ulex alone staining of RBCs from individuals of blood groups O, B and A^1 demonstrated that the amount of H was greatest on O RBCs, then B RBCs and finally A^1 RBCs had the least amount of H antigen expression (Figure 3.5). This clearly demonstrates the changes in efficiency of the transferases in converting H antigens into A or B antigens.

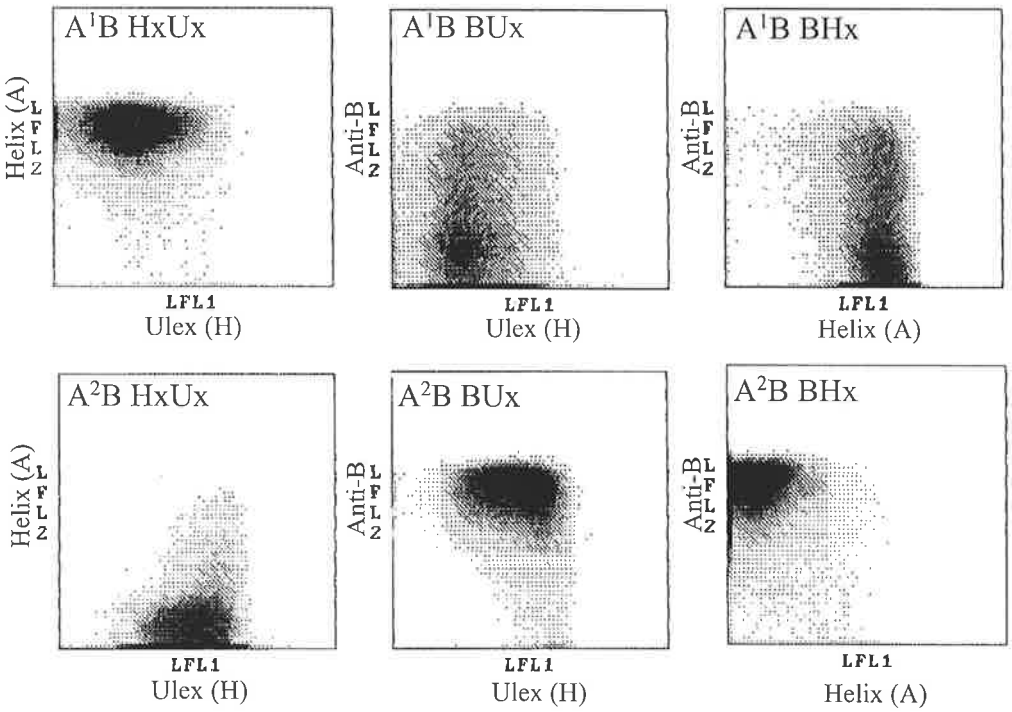
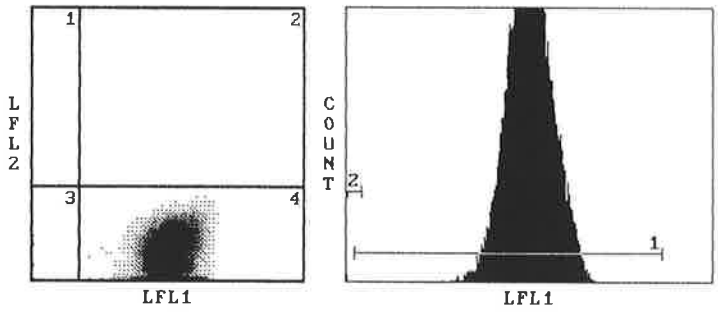


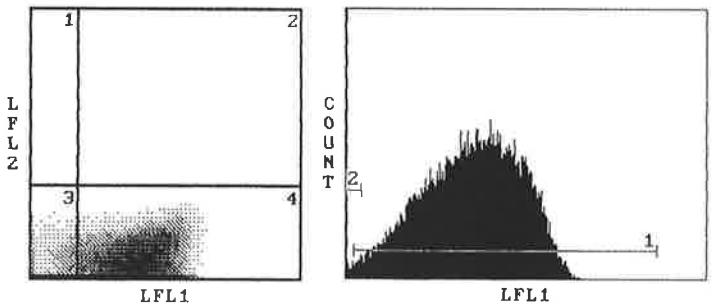
Figure 3.4. Dot plots of healthy A^1B and A^2B individuals.

The 2-colour system used required 3 dot plots for the full analysis of AB individuals. HxUx refers to the dual staining with Helix (A) on the LFL2 axis and Ulex (H) on the LFL1 axis. BUx refers to the dual staining with anti-B on the LFL2 axis and Ulex (H) on the LFL1 axis. BHx refers to the dual staining with anti-B on the LFL2 axis and Helix (A) on the LFL1 axis.

O with Ulex alone



B with Ulex alone



A^1 with Ulex alone

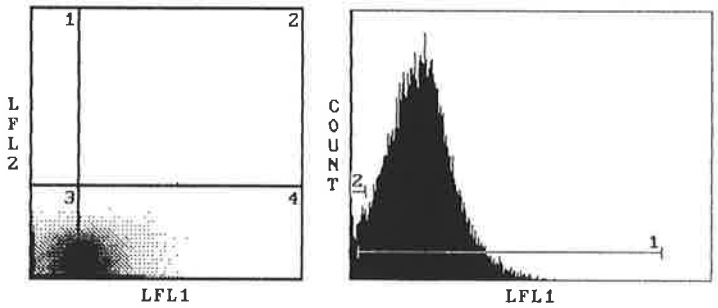


Figure 3.5. Ulex alone staining on red blood cells from blood groups O, B and A^1.

The red blood spheres from O, B and A^1 blood groups were stained with the same batch and amount of Ulex on the same day to observe variations in free H among the blood groups. The amount of H increased in the order O>B>A^1 as reported (Mollison *et al*, 1993).

3.2.2 AML and MDS Patients

Sixty one patients with myeloid disorders were examined by flow cytometry for changes to the ABH antigens (see Table 3.1). There was a variant RBC population in a considerable number of the patients, as specified below. This varied from a minor to a major proportion of the cells. In a few cases, there was no normal population of cells. The descriptions that follow refer to the variant RBC population. All patients with a variant population were genotyped.

Diagnosis	Blood Group	Loss of A or B	Loss of H	Loss of A or B and H	No loss	Total
AML	A, B	3	3	2	8	**16**
	O	N/A	6	N/A	16	**22**
MPD	A, B, AB	1	2	0	3	**6**
	O	N/A	0	N/A	1	**1**
MDS	A, B, AB	4	0	1	2	**7**
	O	N/A	0	N/A	5	**5**
		8	**11**	**3**	**35**	**57**

Table 3.1. Loss of ABH antigens according to ABO blood group and patient disease.
Patients were classified according to disease. Acute myeloid leukaemia (AML), myeloproliferative disorder (MPD), which includes chronic myeloid leukaemia (CML), and myelodysplastic disorder (MDS) including chronic myelomonocytic leukaemia (CMML). NA (not applicable) refers to analysis of the O blood group where loss of A or B could not be assessed.

Four patients with altered cell populations (F3, F5, F21 and F40; see Table 3.2) were omitted from the analysis as they had received blood transfusions. Blood transfusion can give rise to two distinct cell populations as patients, especially of group A, can be transfused with blood from a different blood group. Figure 3.6 illustrates the patterns observed for patients before and after transfusion where an A^1 patient received A^2 blood and vice versa. Transfusions have the potential to obscure some blood group changes, making it difficult to follow changes in patient RBC ABH antigens.

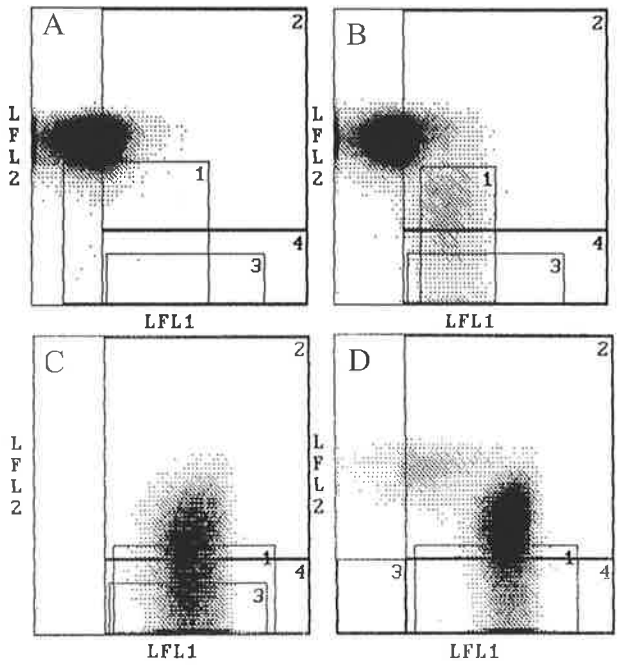


Figure 3.6. The effect of transfusion on patient flow cytometry results.

Dot plots for the dual staining of red blood cell spheres with Helix and Ulex are shown for an A^1 patient before transfusion (A) and after an A^2 blood transfusion (B). The A^2 population is located within quadrant 1 and comprises 18% of the red blood cell population. An A^2 blood group patient before (C) and after (D) transfusion with A^1 blood is also shown with the A^1 population comprising 16% of the red blood cell population after transfusion. The LFL1 axis shows Ulex fluorescence and the LFL2 axis shows Helix (A antigen) fluorescence.

AML AND MDS PATIENTS

ID	AGE	SURV	DIAGNOSIS	PRED	GRP	GENO	FLOW	CYTOGENETICS
F1*	48	23	AML M1	No	O	O1O1	**NO LOSS**	46XY
F2*	NA	NA	MDS		A1	A1A2	**LOSS of A**. A1 to A2 shift (22%)	ND
F3*	48	55+	AML M4	MDS	A1	A1O1	A1 & A2 pattern. **Transfusions**	XX
F4*	51	9	AML M1	No	O	O1O1	**LOSS of H**	XX
F5*	NA	NA	MDS		A1	ND	A1 & A2 pattern. **Transfusions**	
F6*	88	1	AML NR	No	O	ND	**NO LOSS**	
F7*	41	22	AML M2	No	A2	A2O1	**NO LOSS**	46XX, i(7)(p10)/47XX +11
F8*	86	NA	MDS		O	ND	**NO LOSS**	46XX , del(5)(q14,q33)
F9	72	NA	AML		A	A1O1	**LOSS of A & loss of H**	46XY, der(10),t(1;10)(q25;p15)/47XY+8
F10*	NA	NA	MDS		A1	A1O1	**LOSS of A**	ND
F11	70	0	AML M1	MDS	A	A1O1	**NO LOSS**	
F12*	37	NA	RAEB-T/AML M4		B	BO1	**LOSS of B & loss of H**	46XY,del22(p10)
F13	67	9	AML M6	MDS	O	O1O1	**LOSS of H**	del(5)(q11.2;q34)der(14;17)(q10;q10),der16,t(1;16)(p13,p13)
F14*	79	3	AML M2	No	A1	A1O1	**NO LOSS**	XX
F15*	60	64	CML blast crisis		A	A1A2	**NO LOSS**	insufficient metaphases
F16*	84	NA	MDS		A1	ND	**NO LOSS**	XY
F17*	57	NA	MDS		A1	A1O1	**NO LOSS**	44XY,dic(3;21),-4,del5,-7,t(14;?),+r(6)
F18	42	NA	CML blast crisis		A	ND	**NO LOSS**	46XY,der4,ins(4;22)(q etc +Ph
F19*	65	NA	AML	MDS	O	O1O1	**NO LOSS**	44-45XY,der5,-8,add8,der9,-10,der(11;14)(q14;q21)
F20*	69	45+	CML		A2	A1O1	**LOSS of H**	XY
F21*	36	85	CML blast crisis		A1	A1O1	A1 & A2 pattern. **Transfusions**	t(9;22)
F22*	70	0	AML	No	O	O1O1	**NO LOSS**	67-70XY,+X,+1,+1,+2,+6,+del6,+7,+8,etc
F23*	56	45+	CML chronic		A2B	A2B	**NO LOSS**	t(9;22) p210
F24*	72	7	AML	MDS	B	BO1	**LOSS of B** approx. 20%	XY
F25*	87	NA	MDS		A1	A1A1	**LOSS of A** approx 15%	45X-Y
F26*	22	46+	AML M3	No	A1	A1O1	**NO LOSS**	t(15;17)
F27*	24	33	AML M4	No	A1	A1O1	**LOSS of A & loss of H**	46XY, inv1 etc t(1;12)
F28	73	2	AML	MDS	O	ND	**NO LOSS**	44 XX,add3p13,-5,-7,-10,del12p,der17 etc
F29*	71	NA	AML	MDS	O	O1O1	**NO LOSS**	46XX
F30	54	28+	CML chronic		B	BO1	**LOSS of H**	46Y, t(X;13)(p11;q14)-no t(9;22)

F31*	40	9	AML M5	No	O	O1O1	**NO LOSS**	XY
F32*	44	NA	AML M3	No	O	O1O1	**NO LOSS**	t(15;17)
F33	29	NA	CML chronic		O	ND	**NO LOSS**	t(9;22)
F34*	36	NA	AML?MDS		O	ND	**NO LOSS**	46XX, del 5(q13,q31),dic(3;?)
F35*	81	1	AML M4Eo	No	O	ND	**NO LOSS**	inv 16
F36	84	NA	MDS		O	ND	**NO LOSS**	XY
F37	60	12	AML	MDS	O	O1O1	**NO LOSS**	XY
F38*	54	11	AML M1	No	A1	A1A1	**NO LOSS**	XY
F39*	57	0	AML M4	No	A2	A2O1	**LOSS of H**	46XY t(10;11)(q21;q21)/47XY,t(10;11)+8
F40	67	25	AML M2	No	A2	ND	A2 & A1 pattern. **Transfusions**	XX
F41*	78	6	AML M2	MDS	O	O1O1	**LOSS of H**	47XX+14
F42*	68	61+	AML M3	No	B	BO1	**NO LOSS**	t(15,17)
F43*	85	1	AML M1	No	O	O1O1	**LOSS of H**	47XX+8
F44*	78	NA	MDS		O	O1O1	**NO LOSS**	XY
F45*	76	53+	CMML		O	O1O1	**NO LOSS**	XY
F46*	33	1	AML M3	No	A1	A1O1	**LOSS of A** approx 15%	t(15;17)
F47	74	NA	MDS		O	ND	**NO LOSS**	45X, -Y
F48	23	4	AML M5	No	O	O1O1	**LOSS of H**	47XY t(2;11)(q37;q23) +8
F49*	61	45+	AML M2	TC	A2	ND	**NO LOSS**	t(9;11)(p22,p15?)
F50	75	1	AML M2	MDS	O	O1O1	**LOSS of H**	XY
F51	75	4	AML M4	PRV	A	A1O1	**LOSS of A** 100%	45X, -Y. karotyping on 30/1/97 -ve for p210, p190
F52*	61	NA	MDS		AB	A1B	**LOSS of A** with increase in H	XY
F53	81	1	AML M1	No	A	A1O1	**LOSS of H** 10-20%	XY
F54*	65	19	AML M2	No	O	O1O1	**NO LOSS**	4/2/97 41-45XY, -3, -5,del7,-9,der11,t(11;15),del17p etc
F55*	56	2	AML M4	MDS	O	ND	**NO LOSS**	XX
F56*	28	46+	AML M2	No	O	ND	**NO LOSS**	t(8;21)
F57*	26	3	AML M7	No	B	BO1	**LOSS of H** approx. 20%	43-45XY, add2,-5,-6,del7 etc +markers
F58*	61	15	AML NR	MDS	O	O1O1	**NO LOSS**	XX
F59*	61	22	AML NR	No	O	O1O1	**NO LOSS**	XY
F60*	54	34	CML blast crisis		A1	A1A2	**LOSS of A.** A1->A2->O.	t(9;22)
F61*	77	3	AML NR	MDS	A1	ND	**NO LOSS**	47XY+4

OTHER PATIENTS

ID	DIAGNOSIS	GRP	GENO	FLOW
OT1*	ALL	O1	O1O1	**NO LOSS**
OT2*	ALL	A1	A1O1	**NO LOSS**
OT3*	ALL	A1	A1A1	**NO LOSS**
OT4*	ALL	O	O1O1	**NO LOSS**
OT5*	ALL	A1	A1O1	**NO LOSS**
OT6	ALL	O	ND	**NO LOSS**
OT7*	CLL	A1	ND	**NO LOSS**
OT8*	CLL	O	ND	**NO LOSS**
OT9	CLL	A1	ND	**NO LOSS**
OT10*	CLL	O	ND	**NO LOSS**
OT11*	CLL	O	ND	**NO LOSS**
OT12*	CLL	A2	ND	**NO LOSS**
OT13*	CLL	O	ND	**NO LOSS**
OT14*	CLL	A2	ND	**LOSS of H**
OT15*	CLL	O	ND	**LOSS of H**
OT16*	CLL	A1	A1O1	**LOSS of A**
OT17*	Heart Condition	A2	A2A2O1	A2 with **INCREASE in A**
OT18*	AML	A1	ND	A1 with **INCREASE in H**

Table 3.2. Summary of patient flow results.

* denotes patients analysed by B. Farmer. AGE is at presentation of disease.

SURV denotes survival in months with + indicating patient still alive. NA is not available and NR is no record.

PRED denotes predispostion, with PRV indicating polycythemia rubra vera, and TC denotes thrombocytopenia.

GRP denotes blood group determined by serological methods while GENO indicates *ABO* genotype as determined by the method developed in Chapter 4.

3.2.2.1 Loss of the H Antigen

In one group of patients, the H antigens were lost or diminished creating a variant population of red cells. This was most readily seen for the O blood group patients. Six of the 28 O patients (F4, F13, F41, F43, F48, F50; Table 3.2) studied showed loss of H antigen expression. This was seen as a population of cells (approximately 25%) with reduced Ulex fluorescence, in addition to the normal O population, in an O^1O^1 AML patient (Figure 3.7A). Alternatively, a total reduction in Ulex fluorescence of the entire population was seen in another O^1O^1 AML patient (Figure 3.7B). Loss of H was not seen in any of the 51 normal O individuals.

Alterations in H antigen expression was responsible for the alterations in A or B antigens in 5/29 A, B or AB patients (F20, F30, F39, F53, F57; Table 3.2). Figure 3.7C shows an AML patient (genotype A^1O) where partial loss of A due to loss of H can readily be seen as the red cell population has less Ulex fluorescence causing it to resemble an A^1A^1 pattern. There is also a continuum of cells that have decreased Helix fluorescence while showing no increase in Ulex fluorescence.

Another example of loss of A due to loss of H can be seen in a CML A^1O patient in which the entire population shows diminished H relative to a normal A^1O individual (Figure 3.7D). The pattern resembles an A^2O with diminished H indicating that the loss of H in this instance is causing a decrease in A antigens. This patient was grouped as an A^2 by the blood transfusion service, however the patient genotyped as an A^1O^1 (Table 3.2).

Figure 3.7E shows a CML BO patient with loss of H. The fluorescent population is shifted to the left compared to a normal B pattern, indicating decrease of H. In this case, total B fluorescence was diminished due to the loss of H sites.

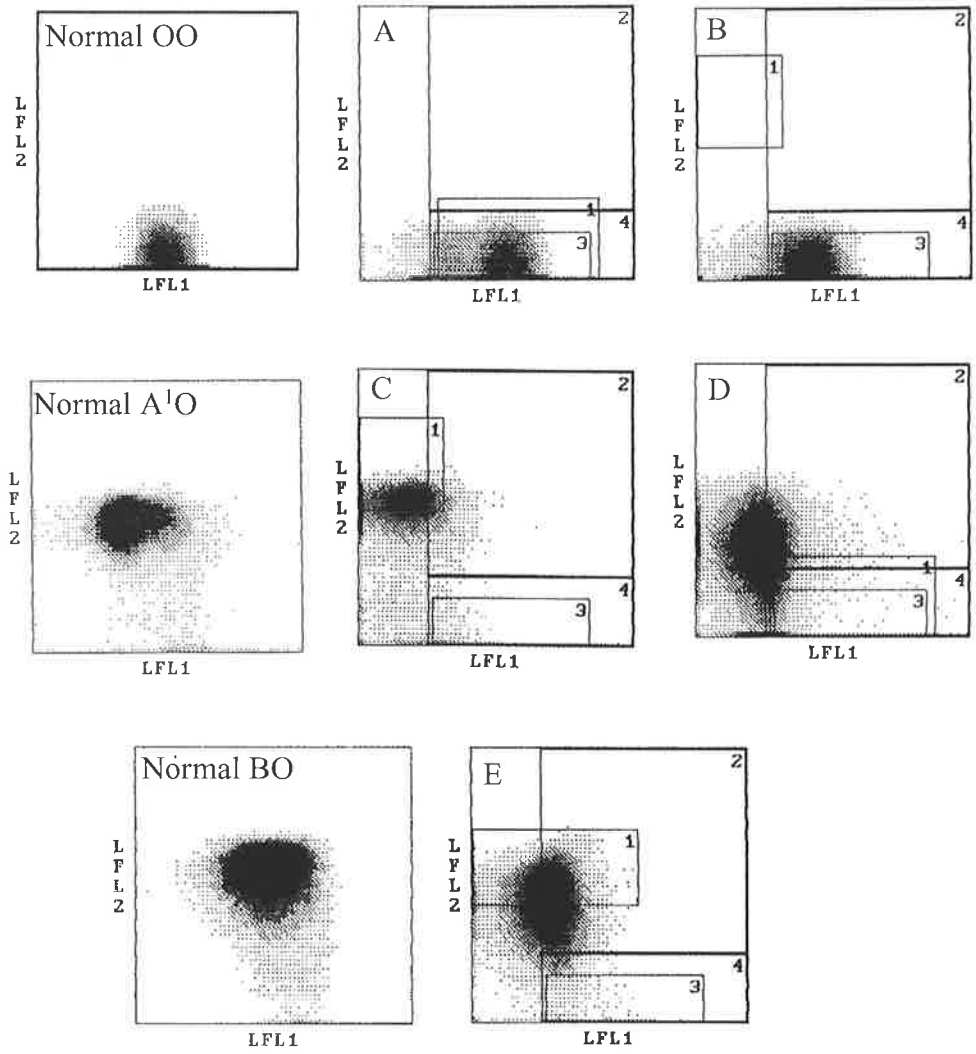


Figure 3.7. Dot plots of patients with loss of H.

Red blood cell spheres stained with Ulex (LFL1) from two O patients with AML (A) and (B) both with a normal population of RBCs and a population of RBCs having diminished H fluorescence. The RBC spheres from an A^1O patient with AML (C) and an A^1O patient with CML (D) stained with Helix (LFL2) and Ulex (LFL1). The loss of H is more extensive in the patient with CML and is affecting the levels of the A antigen as well. Loss of H was also seen in the RBC spheres from a *BO* patient with CML (E) stained with anti-B (LFL2) and Ulex (LFL1).

3.2.2.2 Loss of A or B Antigen

Loss of A or B without loss of H occurred in 8/29 A, B and AB patients (F2, F10, F24, F25, F46, F51, F52, F60; Table 3.2). These patients showed some compensatory increase in H on the cells where A or B antigens were decreased. Figure 3.8A shows an AML patient, who had presented with a mixed field but had previously grouped as an A and genotyped as an A^1O, but showed complete loss of A. The patient showed the same aberrant pattern at the time of his death, a year later.

Figure 3.8B shows the RBCs of a MDS patient (genotype A^1A^2) with two populations, a typical A^1A^2 population and an O-like population. The O like population is derived from the normal A^1A^2 population losing A antigens which would free up H antigens and hence create a population of cells with only H expression, like an O pattern. There is also a small population of cells with decreasing A fluorescence and increasing H fluorescence.

Figure 3.8C shows an A^1A^2 individual with end stage CML with a second A^2 - like population of cells (about 15%) in addition to the normal A^1A^2 cells. This A^2 population had diminished A and a corresponding increase in H when compared to the normal A^1A^2 population of cells.

Figure 3.8D shows the RBCs from an AML patient (genotype *BO*) with 20% of cells showing loss of B and an increase in H. This population of cells with loss of B resembles an O-like population with its lack of B antigen expression but high H antigen expression.

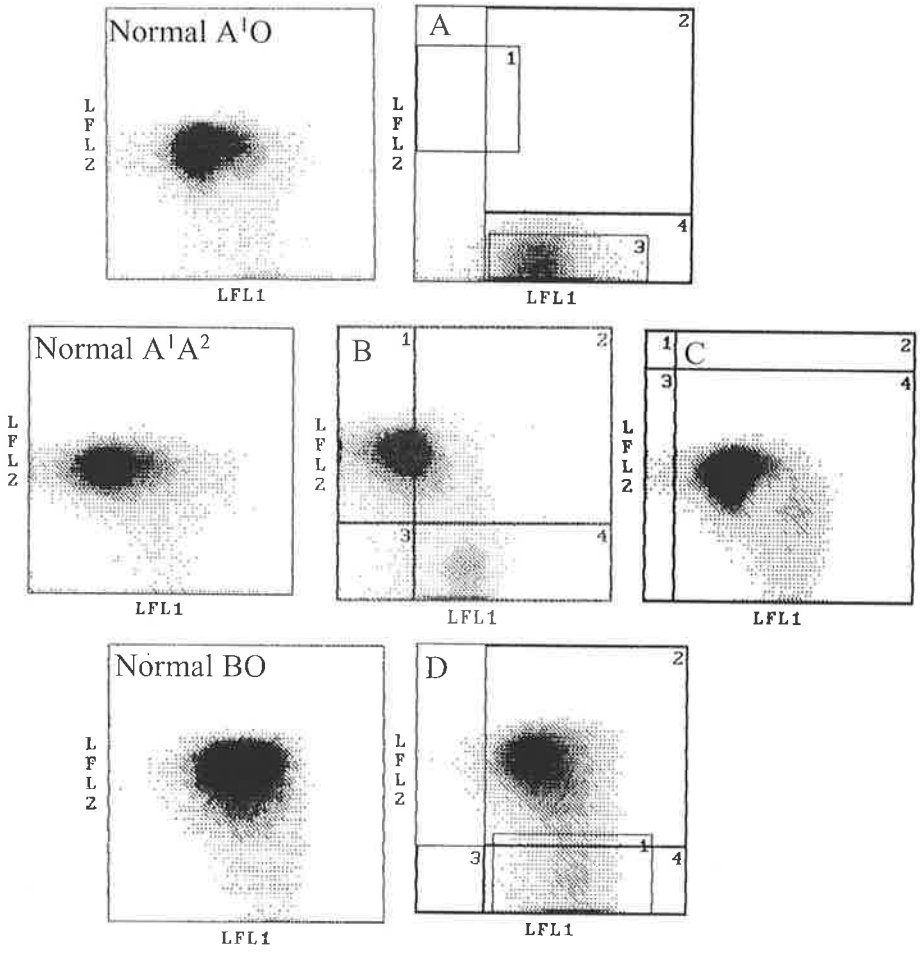


Figure 3.8. Dot plots of patients with loss of A or B.

RBC spheres from an A^1O patient with AML with 100% of cells showing no Helix staining (A), an A^1A^2 MDS patient with 20% of RBCs, located in quadrant 4, having no Helix staining (B), and an A^1A^2 CML patient with 15% of cells having reduced Helix staining (C). (A), (B), and (C) were stained with Helix (LFL2) and Ulex (LFL1). (D) A *BO* AML patient with RBC spheres stained with anti-B (LFL2) and Ulex (LFL1) and with 20% of cells, located in quadrant 1, having minimal B staining.

3.2.2.3 Loss of H and A or B Antigens

A third group of patients (3/29; F9, F12, F27; Table 3.2) appeared to have both loss of A or B and H antigens. One A^1O AML patient had two populations of cells, a minor population resembling normal A^1O cells and a major population with no A fluorescence. The abnormal population showed a continuum of H fluorescence extending from no evident expression to weak expression characteristic of most patients with loss of H (Figure 3.9A). This patient had several samples taken during the terminal stages of the disease and the normal A^1O cells decreased, presumably as normal haematopoiesis shut down.

Figure 3.9B shows an A^1O AML patient with two populations; one that had loss of H giving an A^1A^1 pattern (quadrant 1). The loss of H results in the normal A^1O population being shifted to the left due to the loss of H, but the A antigens are unaffected, with the pattern resembling that of an A^1A^1 population. The second population has varying degrees of ABH antigen changes. The cells directly under the A^1A^1 like population have no H but some A hence loss of A due to H. However, there is a group of cells beyond the A^1A^1 like population which have low levels of A but varying amounts of H indicating loss of A with a corresponding increase in H.

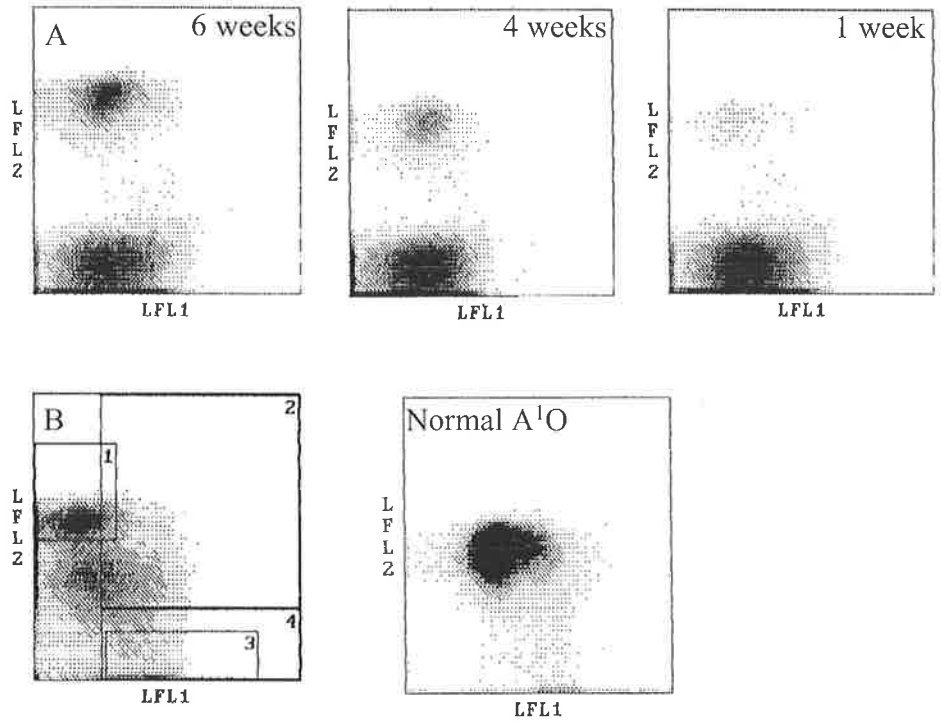


Figure 3.9. Dot plots of patients with loss of H and loss of A.

(A) RBC spheres from an A^1O patient with AML at 6 weeks, 4 weeks, and 1 week before death. The normal A^1O population decreased from 25% at 6 weeks before death to 10% at 4 weeks before death before falling to 1% just before death. (B) The dot plot of an A^1O AML patient showing a population of cells (28%) with loss of H (quadrant 1) and a second population of cells (quadrant 2, 3 and 4) with loss of A. The LFL1 axis shows Ulex fluorescence and the LFL2 axis shows Helix fluorescence.

3.2.3 Other Patients Analysed by Flow Cytometry

Flow cytometric analysis was also performed on other patients to assess loss of ABH antigens. These included both ALL and CLL patients, in which loss of ABH antigens would not be expected as RBCs are derived from the myeloid lineage, while ALL and CLL cells are of the lymphoid lineage. The 6 ALL patients examined (3 A^1 blood group and 3 O blood group) showed no loss (data not shown). Of the ten CLL patients analysed (3 A^1 blood group, 2 A^2 blood group and 5 blood group O) two showed loss of H (1 A^2 and 1 blood group O patient) while only one showed some RBCs with loss of A (approximately 7%) (Table 3.2).

Two other patients showed interesting results although they did not show loss of ABH antigens. One patient, OT17, was blood group A^2 and did not have leukaemia but rather a heart condition and on blood group typing presented with a mixed field. This patient was genotyped as an A^2A^2O showing the presence of a gene duplication. This explained the increase in A fluorescence by flow cytometry (see Figure 3.10A). Another patient, OT18, was a blood group A^1 AML patient (unable to be genotyped) whose RBCs showed an increase in H without loss of A (see Figure 3.10B). We were unable to genotype the A^1 patient since there was no sample but it is possible that they may have a duplication at the *FUT1* locus.

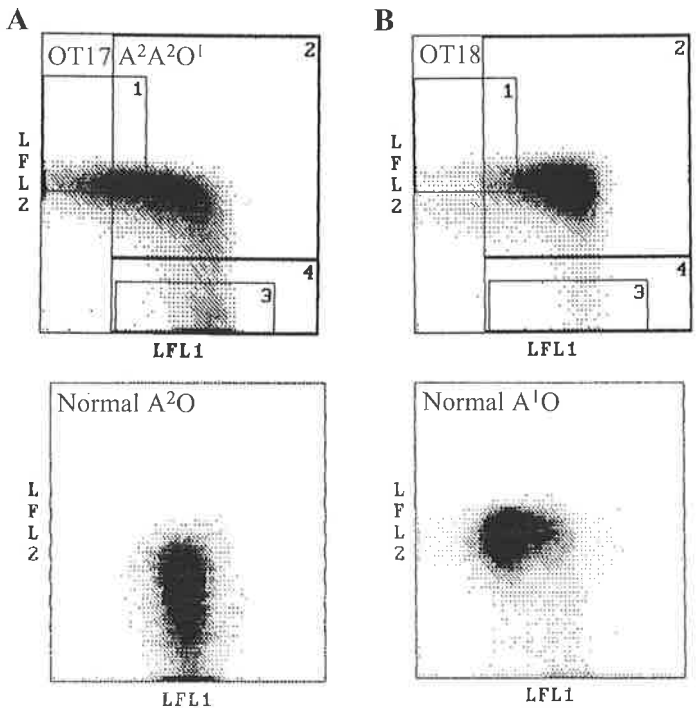


Figure 3.10. Increase in ABH antigens.

A. OT17 is a patient with a heart condition that presented with a mixed field reaction but flow cytometric analysis indicated an increase in A antigen staining when compared to the normal A^2O pattern. Genotyping indicated that the individual had an extra copy of the A^2 allele.

B. OT18 is an AML patient of blood group A^1 with a flow pattern similar, but not identical to an A^1O pattern. This patient seems to have an increase in H antigen, without loss of A, possibly due to a duplication at the *FUT1* locus.

3.3 Discussion

Alterations in the expression of the ABH antigens in haematological malignancies have been considered rare, however, few methodical approaches have been utilised to try to establish how frequently this phenomenon occurs. Some reports have indicated quite high rates of alterations (Majsky *et al*, 1967; Starling and Fernbach, 1970). However, the less subtle changes are likely to have been missed, since samples with 50% loss can be grouped as normal in transfusion laboratories (O'Keefe, PhD thesis, 1995). The aim of this work was to determine the frequency of loss of ABH antigens in the RBCs of patients with myeloid malignancies in a methodical fashion using two-colour flow cytometry. Flow cytometry is a sensitive method for determining loss of antigenic expression as it can resolve multiple populations, and allow partial or full loss of expression to be determined as well as quantification of the number of cells involved.

The normal blood group O individuals had high amounts of H antigen, as expected, due to the inability of the O transferase to convert H antigens. Bombay individuals had no staining for any antigens using this flow cytometric method apart from background staining (Figure 3.1).

Normal B blood group individuals showed heterogeneous staining for H and B antigens. Most cells had high levels of B however there was a population, 3-10%, with no B fluorescence. The H staining was also heterogeneous in the normal blood group B, indicating large amounts of unconverted H. Sharon and Fibach (1991) found a reciprocal relationship between cells expressing high amounts of B and low expressors of H, and vice versa. This was seen in the flow cytometry pattern for the normal B individuals with a small population of cells with no B antigen the ones with the most H antigen expression (Figure 3.1). This in accordance with other reports that B blood group RBCs have more free H (Ulex staining) than A^1 blood groups (Mollison *et al*, 1993).

The normal A^2 blood group also had a very heterogeneous A staining but unlike the B pattern, normal A^2 individuals had less variation in the H staining with almost all cells having the same amount of H regardless of the amount of A. This may imply that the total amount of A is fairly low so that even in the strongest staining there is only a marginal decrease in the total amount of free H antigen. For A^2O individuals, A antigen staining ranged from high to little or no A staining. This observation is in accordance with the study of Reyes *et al* (1976) who used immunoelectron microscopy and observed that A^2 cells showed marked variation in their cell surface staining for A antigens, with minimal to high level cell staining. They remarked that the A^2 blood group is a mixture of several antigenically different populations of cells. A similar heterogeneity was observed for A^1 cells but to a lesser extent (Reyes *et al*, 1976).

For the normal A^1 individuals, it was noted that slight, but obvious, variations in the pattern of antigen expression existed. On further investigation it was possible to demonstrate that the A^1 genotypes (A^1A^1, A^1A^2 and A^1O) had different patterns of A and H antigen fluorescence. The A^1A^1 had the most A staining and the least H staining, while the A^1A^2 had less A and more H, and the A^1O had even less A and the most H staining of the three A^1 genotypes investigated (Figure 3.2). There was a noticeable dosage effect in that the presence of 2 A alleles led to more conversion of H sites than one alone. This observation is in accordance with Sharon and Fibach (1991) who used A/H ratios and found that A antigen expression was greatest on A^1A^1 RBCs that express almost no H antigen, followed by A^1A^2, A^1O and A^2O RBCs, respectively. Sharon and Fibach (1991) found that A^1A^1 RBCs express almost no H antigen and that A^1A^1 RBCs showed the highest level of A and lowest level of H. This is accordance with the observations shown here. Similarly, the results here are in accordance with Sharon and Fibach (1991) that the A^2O RBCs have more free H antigen than the A^1O RBCs.

Two other weak A blood groups were analysed by flow cytometry. These were the A^3 and A^x blood groups. The A^3 blood group had a small population of cells with a very heterogeneous pattern of A antigen expression, much like the A^2 flow profile, except that the proportion of cells with high or intermediate expression of A was very low. For the A^3 individuals the majority of cells had no A expression but had H expression similar to an O individual, hence the O-like population for the A^3 RBCs was bigger than the O-like population in the A^2 RBCs. The A^3 blood group gives a mixed field by serological techniques and this is not surprising since most of the cells look like O cells upon flow cytometric analysis. The A^x blood group could not be distinguished from an O blood group using flow cytometry (Figure 3.3).

For the A^2 blood group a considerable number of RBCs had negligible A expression yet this blood group does not give a mixed field. We were unable to find an A^2A^2 individual but it would have been interesting to see whether the flow cytometer would pick up any differences in A or H antigen expression between an A^2O individual and an A^2A^2 individual. The expectation would be that an A^2A^2 would have less H antigen and more A antigen expression than an A^2O, like and A^1A^1 had less H and more A antigen expression than an A^1O^1. The flow cytometric pattern seen for patient OT17 (Figure 3.10A) may represent the flow profile of an A^2A^2 individual.

A^1B red cells have been reported to often give an A^2 pattern serologically with A^2B red cells reported to give an A^3 pattern (reviewed in Race and Sanger, 1975; Daniels, 1995). However, a considerable decrease in A was not evident by flow cytometry in any of the 29 A^1B individuals examined (Figure 3.4). The Helix versus Ulex plot of A^1B was similar to that observed in an A^1A^1 control and not that of an A^2O individual. An A^1B has 4-8x10^5 A antigen sites per RBC while an A^1O RBC has 8-12x10^5 and an A^2O has 1.5-4x10^5 A antigen sites (Economidou *et al*, 1967). Based on these numbers one would expect the Helix-Ulex (A vs H)

pattern of an A^1B to be intermediate between an A^1O and A^2O however, it more closely resembles an A^1A^1 individual and not an A^2O. This may be due to any free H being converted by the B transferase, hence the A vs H pattern has more A than H antigen staining. The discrepancy may also be due to technical differences between the flow cytometric method and the technique used for determining the A antigen site numbers, which relied on agglutination of the antigens with the appropriate antibodies.

The B fluorescence in A^1B individuals was markedly weaker than seen in a normal BO and this can be explained by the fact that the A^1 transferase is a more efficient transferase than the B, and once all of the H sites have been converted, there are no more H sites for the B transferase to convert.

By contrast, the A^2B samples gave results consistent with previous reports (Figure 3.4). The B fluorescence was practically identical to that seen in a BO individual. The A fluorescence was substantially weaker than that seen in an A^2O and gave an A^3-like pattern with the majority of cells appearing negative for A. This is surprising since there is a considerable amount of H that could be converted into A antigen (data not shown) however, the B antigen is more efficient at converting the H antigen. It has been suggested that when both the A and B alleles are inherited, phenotypic expression of the A antigen is depressed due to competition between the A and B glycosyltransferases for the H antigen (Hakomori, 1981). Though this is may explain why the A vs H pattern in A^2B resembles an A^3, it does not explain why an A^1B A vs H pattern does not look like an A^2 flow profile. The activities of the 2 alleles in AB patients do not act additively. The reduction of B in A^1B individuals may be due to direct competition of the A and B transferases as there is near saturation of the H sites. However, in an A^2B individual, there are a substantial amount of free H sites that could be modified so there is no clear explanation why the A fluorescence is diminished. It has been suggested that UDP-

galactose, the sugar that the B transferase adds to the H antigen to make the B antigen, inhibits the A transferase (Salmon *et al*, 1984). This inhibitory effect may be seen in an A^2B individual and not an A^1B individual because the A^1 transferase is more efficient than the A^2, hence only a small amount of A^1 transferase is required to convert H into A antigens.

Sharon and Fibach (1991) found that A^1B RBCs differed from A^2B cells in A/H ratios without significant differences in their B/H ratios. The data presented here have shown that for the A^1B blood group, the level of B expression was significantly less than the level of B antigens seen on the RBCs of a normal B individual, while for the A^2B blood group the B antigen levels were equivalent to those seen for a B individual. The A^1B individual had over 50% of cells with no B antigen expression while an A^2B only had 10% of cells with no B antigen expression. This observation is contradictory to Sharon and Fibach (1991) which found that the level of B fluorescence was nearly 50% higher in A^1B individuals when compared to A^2B individuals. At the moment, this discrepancy in results cannot be explained.

The heterogeneity seen for the weaker blood groups confirms previous agglutination and electron microscopy studies (Reyes *et al*, 1976; Cartron *et al*, 1977). The mechanism for this observation is unclear but may be a consequence of the association of the weaker blood groups with a different sequence in the upstream region of the *ABO* gene. It has recently been shown that an upstream repeat element of the *ABO* gene, which varies in copy number according to the allele, has enhancing ability (Kominato *et al*, 1997; Yu *et al*, 2000). Surprisingly, the enhancing ability of the single repeat associated with the A^1 allele was reported to be weaker than that of the multiple repeat associated with the A^2 and B alleles in *in vitro* studies (Yu *et al*, 2000). Recent studies show that enhancer mediated gene expression operates in a stochastic manner (reviewed in Fiering *et al*, 2000). We would then expect a weak enhancer to give a higher degree of null cells whereas A^1 is associated with lower

amounts of null cells than A^2 and B. It will be necessary to consider the whole upstream region rather than isolated segments of it to fully establish the relationship of the enhancer sequence with ABH antigen expression. The number of repeats associated with the various weak blood groups remains to be determined.

Sharon and Fibach (1991) also reported a flow cytometry based approach for studying RBCs of normal individuals. They were able to differentiate the genotypes A^1A^1, A^1A^2, A^1O and A^2O by analysing the ratio of A/H and B/H fluorescence. The results shown here confirm that flow cytometry can be used to deduce these as well as other genotypes. The methodology presented here differs from that of Sharon and Fibach in important technical details such as the use of Helix Pomatia lectin rather than a monoclonal antibody to detect A, as well as the use of RBC spheres to decrease agglutination problems. The use of RBC spheres has several other advantages, the principal one being the decreased scatter of the cells, which allows better visualisation of the entire population, and lends itself to accurate gating to exclude agglutinated cells and debris. In addition, formalin fixed spheres are much more stable than unfixed RBCs and may be used for several months after fixation.

Moreover, visual analysis of the two-parameter flow cytometry patterns is used. Pattern analysis allows identification of differences not only in normal individuals where the different patterns reflect the genotype, but more importantly in leukaemia patients where variant and normal red cell populations may co-exist. Pattern variations in a panel of 180 normal individuals (34 A^1, 8 A^2, 2 A^3, 4 A^x, 35 B, 29 A^1B, 15 A^2B, 51 O and 2 Bombay individuals) was related only to genotype, rather than changes caused by disease. However, in a considerable number of leukaemic patients, 22/57 (39%) a variant RBC population was present. In almost all cases, the changes were not detected by routine serological typing. Leukaemia patients with unusual patterns were genotyped as this was necessary to allow

comparison to the normal pattern for that genotype. It was not always possible to analyse the patient RBCs as blood transfusion with different genotypes gave rise to aberrant patterns (see Figure 3.6).

The changes in ABH antigens in leukaemia patients were of three different types. Firstly, a group of patients with decreased H antigen staining to which the loss of A or B antigens was secondary (Figure 3.7). Presumably, there is a loss of expression of the gene coding for the H antigen (*FUT1* in erythroid precursors) since loss of A or B would lead to an increase in H, although it is possible the loss may occur in a precursor of H or may even be due to masking of the H epitope. Investigation of possible *FUT1* alterations in this group of patients is the focus of Chapters 4 and 5. A second group of patients had a decrease of activity at the *ABO* locus. This can be seen by the increase of H antigen expression, which is less obvious in some cases than might be expected (Figure 3.8). A third group had a decrease or loss of H as well as of A or B, sometimes resulting in two different abnormal populations of cells (Figure 3.9). This indicates that the ABH changes may occur in subpopulations after the malignant clone is established.

One common pattern observed for A^1O patients was a shift to an A^1A^1 like pattern due to a marked decrease in the amount of H (Figure 3.7). This can most readily be explained by a reduction in H activity perhaps due to inactivation of one of the two *FUT1* alleles. Most of the losses of H observed are compatible with inactivation of one *FUT1* allele as Ulex fluorescence is diminished and not lost. No patients showed a total loss of H for the whole cell population, however, some had a sub-population with total loss of H (Figure 3.7 and 3.9).

Genotyping of all patients was an important aspect of this study. In particular, A^1O patients who showed a shift to an A^1A^1 like pattern demonstrated the necessity of genotyping. It was

also necessary to verify the genotype for other patients with complete or near complete loss of antigens. For example, Figure 3.7D shows a pattern which looks like an A^2O pattern but is actually from an A^1O patient with almost total loss of H.

In some patients, we were able to track changes in ABH antigen expression during disease progression. Figure 3.9 shows an AML patient (genotype A^1O) in which the population of normal A cells gradually decreased during the final months of disease, presumably due to continuing loss of normal haematopoiesis. Several other groups have also noted changes over the course of disease. Gold *et al* reported an AML patient where 98% of the cells had loss of A at presentation. This fell to 35% in remission before rising again at relapse (Gold *et al*, 1959). This indicates that loss of ABH antigens can be used to monitor disease as they are known differentiation markers (Ravn and Dabelsteen, 2000). In remission, the shift of the normal and leukaemic cell populations reverts to the normal population being predominant but the cells which show loss are still leukaemic and hence relapse occurs at a later date.

There were similar frequencies of loss of H in the O patients (6/28) and the A/B patients (8/29). This data combined gives an overall loss of 25%. Loss of A or B was seen in 11/29 (38%) of the A, B or AB patients. Fifty-five percent (16/29) of A/B patients showed alterations at one or both of the loci. The rate of loss of H that we observed is similar to that of Majsky *et al* (1967) who reported that 27% of acute leukaemia patients had loss of H.

The high frequency of changes seen is not age related as alterations in blood group are also seen in younger patients. Omitting the transfusion patients from the analysis, there were losses in 5 out of 14 patients (36%) younger than 50 and in 15 out of 41 patients (36%) older than 50 (Table 3.2). No alterations in ABH antigens were observed in the 178 healthy

individuals examined who ranged in age from 21 to 74. This supports the hypothesis that the alterations in ABH antigen expression on RBCs is related to the disease process.

Kaplan-Meier analysis showed no significant difference in survival between patients with and without ABH antigen loss (data not shown). Two analyses were done; considering all patients and considering patients with AML alone. There was, nevertheless, a tendency for patients with loss to have a shorter survival time. Patients with AML with loss showed a median survival time after presentation of 5 months, whereas patients with AML without loss showed a median survival of 11 months (Table 3.2).

Examination of the association of antigenic loss with patient diagnosis, age at presentation, and survival after diagnosis showed that there was no significant difference in the proportion of patients with loss among patients with AML and the other haematological patients (Fisher exact test). Subclass information was available for 26 patients with AML. In each subgroup there were 4/7 M1 patients, 3/8 M2 patients, 1/4 M3 patients, 2/4 M4 patients, 1/2 M5 patients, and 1/1 M6 patient that showed variant cell populations. As such, loss does not appear to be restricted to particular subclasses of AML although further analysis is required as the patient sample size is low. Thirty-six patients with AML had information regarding prior predisposing disease such as MDS or MPD. The percentage of patients with loss of ABH antigens and a predisposing disease (5/14) or those with no predisposing disease (8/22) was 36% in both cases, suggesting that the predisposing disease was not related to the loss. This may also suggest that alterations in ABH antigens are a later change in the disease progression.

Loss of ABH antigens may arise in some cases from loss of heterozygosity. The *ABO* gene is located on chromosome band 9q34, whereas the *H* (*FUT1*) gene is on chromosome band 19q13. There is a recurrent deletion of the 9q23-31 region in myeloid malignancy and this

may extend to 9q34 in some cases (Evdokiou *et al*, 1993). None of the patients with ABH antigen loss and that had cytogenetic data available, had visible 9q or 19q (or -9 or -19) deletions (Table 3.2) indicating that the inactivation of *ABO* and *FUT1* may be due to epigenetic mechanisms (see Chapter 5). The 9q34 region is studied in more detail in Chapter 6 as there may be a tumour suppressor gene in the same chromosomal region as *ABO* that may be the target for inactivation in haematological malignancies.

The results in the patients can best be explained by considering one or more silencing mechanisms acting at both the *ABO* and *FUT1* loci. In some cases, just one locus appears to be silenced and in other cases, both are silenced. In all the cases of loss of H, it seemed that only one allele was silenced. Methylation of the *ABO* and *FUT1* gene promoter regions is one such mechanism for silencing and has been reported for the *ABO* gene in human cell lines (Kominato *et al*, 1999). In line with an epigenetic explanation, it is notable that in some cases, loss of expression of A or B seems incomplete (for example see Figure 3.8). In this instance a decrease in gene activity may be occurring, rather than a complete loss. Epigenetic mechanisms are also supported by parent of origin effects on loss of A and B alleles (Dobrovic *et al*, 1993). Epigenetic mechanisms of loss of *ABO* and *FUT1* are the focus of Chapter 5.

In conclusion, flow cytometry was used to demonstrate that red cells from normal individuals showed patterns of antigenic expression consistent with their genotype. In contrast, red cells from leukaemic individuals frequently showed substantial disorder of antigenic expression at the two major loci affecting expression of the ABO blood group. The results in this chapter suggest that alterations in ABH antigen expression in patients with haematological malignancies may be an important signpost to fundamental alterations leading to leukemogenesis. The remainder of this thesis will attempt to clarify the importance of this alteration.

CHAPTER 4

ABO & *FUT1* Genotyping & Expression Analysis

4.1 Introduction

Utilising a previously isolated A transferase probe, Yamamoto and colleagues (1990a) cloned and sequenced the cDNA from cell lines derived from different ABO blood groups. By comparison of the cloned cDNA to the A transferase cDNA sequence they discovered that the O^1 allele had a single base deletion at cDNA position 261 (Yamamoto *et al* 1990b). This group demonstrated that deletion of the G nucleotide at this position caused a shift in the reading frame of the O^1 allele sequence, introducing a premature stop codon. The O transferase is truncated at amino acid residue 117, explaining the lack of O transferase activity.

Seven single base substitutions were identified in the cDNA sequence of the B transferase when compared to the A transferase cDNA sequence. These changes occurred at nucleotide positions 297, 526, 657, 703, 796, 803 and 903 (Yamamoto *et al* 1990b). Four nucleotide substitutions led to amino acid changes to the B transferase at residues 176, 235, 266 and 268 (corresponding to nucleotides 526, 703, 796 and 803 respectively).

The nucleotide differences between the three major *ABO* alleles create allele specific restriction enzyme sites. The deleted G in the O^1 allele creates a *Kpn*I site while the *A* and *B* alleles have a *BstE*II site at the corresponding position (Yamamoto *et al*, 1990b). The C to G substitution at nucleotide 526 in the *A* and O^1 alleles creates a *BssH*II site, however, a *Nar*I site is present in the *B* allele. The G to A substitution at nucleotide 703 creates a *Hpa*II site in the *A* and O^1 alleles and an *Alu*I site in the *B* allele. The C to A substitution at nucleotide 796 creates a *BstN*I site in the *A* and O^1 alleles and a *Nla*III site in the *B* allele.

Yamamoto and colleagues were the first to utilise some of these restriction site polymorphisms to determine the *ABO* genotype of individuals using PCR and restriction enzyme digestion (Yamamoto *et al*, 1990b; Yamamoto and Hakomori, 1990). Southern hybridisation after restriction enzyme digestion of DNA with *Kpn*I or *BstE*II was used to detect the O^1-specific deletion. A PCR-RFLP method was utilised to discriminate the *A* and *B* alleles. A region which spanned the C to G substitution at position 526 and the G to A substitution at 703, was PCR amplified and restriction enzyme digested with *BssH*II, *Nar*I, *Hpa*II or *Alu*I. This PCR-RFLP method has undergone numerous modifications and refinement by many groups and is the easiest and most efficient way of genotyping the *ABO* locus (Yamamoto *et al*, 1990b; Yamamoto and Hakomori, 1990; Lee and Chang, 1992; O'Keefe and Dobrovic, 1993; Olsson and Chester, 1995; O'Keefe and Dobrovic, 1996b).

Grunnet and colleagues (1994) were the first group to differentiate the O^1 and O^2 alleles by genotyping. They amplified the region spanning the deleted G found in O^1 alleles and digested with *BstE*II to discriminate the *A*, *B* and O^2 alleles from the O^1 allele. A second PCR amplified the region containing the nucleotide changes that alter amino acids at 526 and 703. Digestion with *BssH*II discriminated the *A* and O^1 alleles from the *B* and O^2 alleles. Digestion with *Hpa*II elucidated the *A*, O^1 and O^2 alleles from the *B* allele. Using this method, they were able to determine the following genotypes: AA, AO^1, BB, BO^1, AB, O^1O^1 and O^1O^2.

Olsson and Chester (1995) developed a genotyping method that allowed the additional identification of the A^2 allele. This method used the *Hpa*II site, which is present in the 3'-untranslated region of A^1 and O^1 alleles but not of *B* and O^2 alleles. The substitution responsible for the presence or absence of this *Hpa*II site (G1096A) is situated between the stop codons for the consensus and A^2 allele reading frames. The presence of mutations associated with alleles A^2 (C467T), *B* (G703A and G1096A) and O^2 (G1096A) removes

*Hpa*II sites present in the *ABO* consensus sequence. A multiplex PCR was used to generate two products, the first of which spanned exon 6 and thus position 261, for which *Kpn*I would be used to elucidate O^1 alleles. The second PCR product spanned all of exon 7 and did not contain any *Kpn*I sites. The two products were amplified simultaneously and restriction enzyme digested with *Kpn*I and *Hpa*II jointly, and separated by agarose gel electrophoresis. The O^1 allele was easily distinguished from all other alleles by its *Kpn*I site. The other five common *ABO* alleles (A^1, A^2, B, O^1 and O^2) were resolved by their *Hpa*II patterns and all 15 different genotype combinations could be elucidated by this method.

In addition to PCR-RFLP techniques, other genotyping methods use allele-specific PCR (ASP) (Gassner *et al*, 1996; Procter *et al*, 1997; Pearson and Hessner, 1998) or PCR amplification followed by single-stranded conformational polymorphism analysis (SSCP) (Ogasawara *et al*, 1996a; Ogasawara *et al*, 1996b; Yip, 2000). One *ABO* genotyping technique uses ^{32}P-labelled PCR products, amplified with allele-specific primers (ASP) against polymorphisms in exons 6 and 7, that are detected by polyacrylamide gel electrophoresis and autoradiography (Ugozzoli and Wallace, 1992). The Olsson and Chester (1995) PCR-RFLP method described above uses a single PCR while the ASP methods use two to eight PCRs to generate the correct DNA fragments. PCR-RFLP methods have the disadvantage of incorrect genotyping of rare *ABO* alleles for which some have no known RFLPs associated with them. PCR-SSCP methods are the best for detection of unknown mutations in the *ABO* gene, however, these can be very complex to perform (Ogasawara *et al*, 1996a; Ogasawara *et al*, 1996b; Yip, 2000). Another novel method utilises denaturing gradient gel electrophoresis (DGGE) of exon 6 PCR amplified products (Johnson and Hopkinson, 1992). Like SSCP, DGGE can detect mutations anywhere in the PCR amplified products but the type of mutation cannot be determined.

Genotyping of the *ABO* locus was essential to this project as analysis of allelic expression in patients with haematological malignancies required the knowledge of their genotype. *ABO* genotyping information also became important for full understanding of the results obtained from the flow cytometric analysis. For example, a patient typed as an A^1A^1 by flow cytometry may actually be an A^1A^1 or an A^1O^1 with loss of H (Chapter 3). Genotyping this patient is essential for determining which of these two possibilities is true. Some of the above *ABO* genotyping methods described are quite complex but as we had flow cytometry information, which would identify patients with weak *ABO* blood groups, a PCR-RFLP technique for genotyping *ABO* could be used easily. The description of the PCR-RFLP technique and the allelic expression of *ABO* is the focus of the first part of this chapter.

4.2 Results

4.2.1 Genotyping of *ABO*

4.2.1.1 Genotyping for the O^1 Allele

The PCR-RFLP methodology employed for *ABO* genotyping was similar in principle to that previously reported (Yamamoto *et al*, 1990b; Yamamoto and Hakomori, 1990; Lee and Chang, 1992; O'Keefe and Dobrovic, 1993; Olsson and Chester, 1995; O'Keefe and Dobrovic, 1996b). Restriction enzyme digestion of two PCR products was used to genotype the *ABO* locus. The first PCR product, the AO1 PCR, spanned exon 6 and was used to identify the O^1 allele. The AO1 product was amplified using primers AO1 F and AO1 R to give a 263 bp PCR product (nucleotides 1400-1662 of Genbank accession AC000397; see Table 2.1.1 for primer sequences). The deleted G in the O^1 allele creates a *Kpn*I restriction enzyme site at position 196 of the AO1 PCR product. All of the other alleles cut with the restriction enzyme *BstE*II at this site. Hence, complete cutting with *Kpn*I, but no digestion with *BstE*II will be seen for an O^1O^1 homozygote. Partial cutting with *Kpn*I and *BstE*II will be seen for O^1 heterozygotes. Genotypes that contain no O^1 alleles will not cut with *Kpn*I but will be completely digested by *BstE*II (Table 4.1, Figure 4.1).

		Restriction Enzyme	
		*Kpn*I	*BstE*II
Genotype	O^1O^1	196, 67	263
	$?O^1$	196, 67	196, 67
	$??$	263	196, 67

Table 4.1. Expected sizes from *Kpn*I and *BstE*II digestion of the AO1 PCR product.
? refers to the unknown alleles, that is, A^1, A^2, B or O^2.

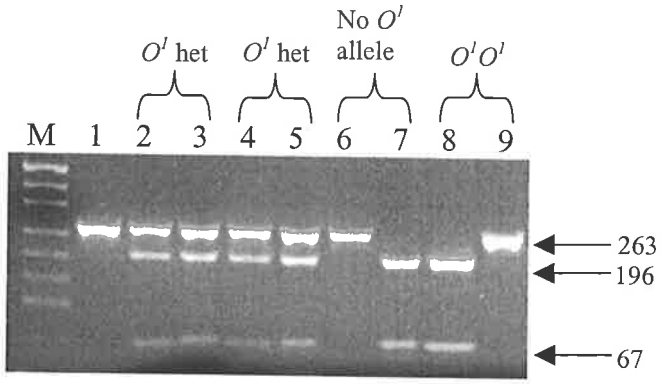


Figure 4.1. Genotyping for the O^1 allele.
M is pUC19/*Hpa*II marker and lane 1 is uncut AO1 PCR product. Lanes 2, 4, 6 and 8 are AO1 PCR products digested with *Kpn*I while lanes 3, 5, 7 and 9 are digestions with *BstE*II. Lanes 2 and 3, and lanes 4 and 5 are paired samples and are examples of O^1 heterozygotes (het). Lanes 6 and 7 are a paired sample of an individual with no O^1 alleles. Lanes 8 and 9 are paired and is an O^1O^1 genotype.

4.2.1.2 Genotyping for the A^2 Allele

A second PCR product, called ABO2, was designed within exon 7 to allow differentiation of the other major *ABO* alleles; A^1, A^2, B and O^2. The initial primers used were ABO2 F and ABO2 R to give a PCR product of 488 bp (nucleotides 80-567 of Genbank accession X84752; see Table 2.1.1 for primer sequences). The A^2 allele has a C467T compared to the A^1 consensus cDNA that creates a *Pvu*II site at position 60 of the ABO2 PCR product (a *Pvu*II site already exists at position 102 of the ABO2 PCR product). Restriction enzyme digestion with *Pvu*II allowed for the differentiation of the A^2 allele (Table 4.2).

		Restriction Enzyme
		*Pvu*II
Genotype	$A^1/B/O$	386, 102
	$?A^2$	386, 102, 60, 42
	A^2A^2	386, 60, 42

Table 4.2. Expected sizes from digestion of the ABO2 PCR product with *Pvu*II.
? refers to unknown alleles, that is, A^1, B, O^1 or O^2.

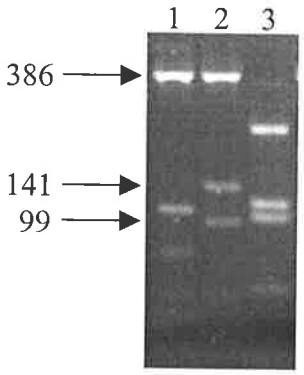
The smaller sized fragments from the *Pvu*II digest were difficult to distinguish in the initial experiment (results not shown) so a new ABO sense primer was designed to optimise the genotyping procedure. This primer, new ABO2 F (see Table 2.1.1), in combination with ABO2 R gave a PCR product of 527 bp (nucleotides 41-567 of Genbank accession X84752). The C to T nucleotide substitution that is found in the A^2 allele creates a *Pvu*II site at position 103 of the new ABO2 PCR product (a *Pvu*II site already exists at position 141). Digestion of this PCR product with *Pvu*II for A^2 genotyping gave larger product sizes and hence clearer genotyping results (see Table 4.3, Figure 4.2).

		Restriction Enzyme
		*Pvu*II
Genotype	$A^1/B/O$	386, 141
	$?A^2$	386, 141, 99, 42
	A^2A^2	386, 99, 42

Table 4.3. Expected sizes from digestion of the new ABO2 PCR product with *Pvu*II.
? refers to unknown allele, that is, A^1, B, O^1 or O^2.

Figure 4.2. Genotyping for the A^2 allele.

Lane 1 is the new ABO2 PCR product digested with *Alu*I, lane 2 and lane 3 show the same product cut with *Pvu*II or *BstU*I, respectively. The arrows indicate the sizes of the *Pvu*II bands in lane 2 with the 99 bp band characteristic of the A^2 allele. The 141 bp band indicates that the individual is a heterozygote for the A^2 allele. The reasons for the other digests are explained in sections 4.2.1.3 and 4.2.1.4.



4.2.1.3 Genotyping for the *B* Allele

Restriction enzyme digestion of the 527 bp new ABO2 PCR product with *Alu*I, allowed differentiation of the *B* allele from the *A* alleles. An *Alu*I site is created in the *B* allele due to a G to A substitution at position 703 of the cDNA sequence, which corresponds to position 339 of the new ABO2 PCR product. Two other *Alu*I sites already exist in the PCR product at positions 28 and 142, however the *B* allele is readily distinguishable as *Alu*I digestion gives a 188 and 197 bp bands, which are present on the gel as a characteristic doublet. (Table 4.4, Figure 4.3).

		Restriction Enzyme
		*Alu*I
Genotype	*A/O*	385, 114, 28
	?A²	385, 114, 72, 42, 28
	A²A²	385, 72, 42, 28
	A²B	385, 197, 188, 114, 72, 42, 28
	?B	385, 197, 188, 114, 28
	BB	197, 188, 114, 28

Table 4.4. Digestion of the new ABO2 PCR product with *Alu*I.

? refers to unknown allele, that is, A^1, O^1 or O^2.

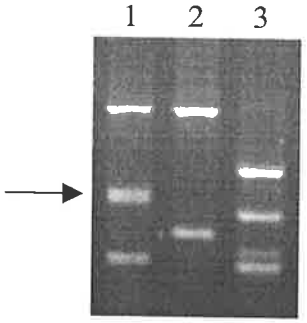


Figure 4.3. Genotyping for the *B* allele.
Lanes 1, 2 and 3 are digests of the new ABO2 PCR product from a BO^1 individual with *Alu*I (lane 1), *Pvu*II (lane 2) and *BstU*I (lane 3). The arrow indicates the 188 and 197 bp bands, which form a doublet that is characteristic for the *B* allele.

The *Alu*I digestion was originally used for the *B* allele but it was noted that the A^2 allele had a distinctively different pattern from the A^1, O^1 and O^2 alleles when cut with *Alu*I. Further investigation revealed that the C to T substitution that was assessed for the A^2 allele using *Pvu*II also created an *Alu*I site. *Alu*I was then used to differentiate the A^2 and *B* alleles such that the *Pvu*II digest was no longer required (Table 4.4, Figure 4.4).

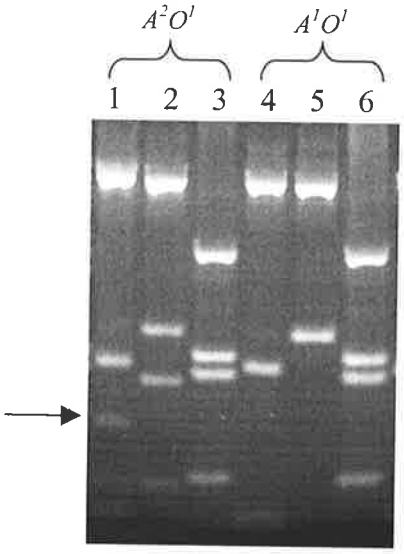


Figure 4.4. Genotyping of the A^2 allele with *Alu*I.
Lanes 1, 2 and 3 show digests of the new ABO2 PCR product for an A^2O^1 individual with *Alu*I (lane 1), *Pvu*II (lane 2) and *BstU*I (lane 3). Lanes 4, 5 and 6 are digests of the new ABO2 PCR product for an A^1O^1 individual with *Alu*I, *Pvu*II and *BstU*I respectively. The arrow points to the unique 72 bp band present in the *Alu*I digest of an A^2 allele.

4.2.1.4 Genotyping for the O^2 Allele

The restriction enzyme *Nar*I was used to detect the C526G substitution present in *B* and O^2 alleles (position 162 of new ABO2 PCR product). However, as reported previously, and observed in this study, the site in the new ABO2 PCR product was resistant to digestion with this restriction enzyme (O'Keefe and Dobrovic, 1996). *Bss*HII was used instead to distinguish the O^2 allele (O'Keefe and Dobrovic, 1996). The A^1, A^2 and O^1 alleles have a *Bss*HII site at position 161 of the new ABO2 PCR product but the C526G substitution present in the *B* and O^2 alleles abolishes this site (Table 4.5).

		Restriction Enzyme
		*BssH*II
Genotype	A/O^1	366, 161
	$BB/O^2O^2/BO^2$	527
	$?B$ or $?O^2$	527, 366, 161

Table 4.5. Restriction enzyme digestion of new ABO2 PCR product with *BssH*II.
? refers to unknown alleles, that is, A^1, A^2 or O^1.

*BssH*II was used initially for the genotyping of the O^2 allele but a cheaper and simple alternative, *BstU*I, which has not been reported elsewhere, was later utilised. The A^1, A^2 and O^1 alleles have four *BstU*I sites in the new ABO2 PCR product found at positions 107, 117, 162 and 285. The C526G substitution, which is at 162 in the PCR product, abolishes this site such that the B and O^2 alleles only have *BstU*I sites at 107, 117 and 285 (Table 4.6, Figure 4.3 lane 3 compared with Figure 4.4 lanes 3 and 6).

		Restriction Enzyme
		*BstU*I
Genotype	A/O^1	242, 123, 107, 45, 10
	$BB/O^2O^2/BO^2$	242, 168, 107, 10
	$?B$ or $?O^2$	242, 168, 123, 107, 45, 10

Table 4.6. Restriction enzyme digestion of the new ABO2 PCR product with *BstU*I.
? refers to unknown alleles, that is, A^1, A^2 or O^1.

With the genotyping of *ABO* optimised, all of the patients analysed by flow cytometry (Chapter 3) and the remaining patients were genotyped.

4.2.1.5 Patient Genotyping Results

Normal controls, leukaemic cell lines and patient samples were all genotyped for the *ABO* locus. Genotyping of the *ABO* locus identified shifts in allele dosage, visualised as a shift in band intensities when compared to normal individuals of the same genotype (Figure 4.5). In samples with equivalent allele dosage, the top band, after restriction enzyme digestion, is heavier due to the inability of restriction enzymes to cut the heteroduplexes created from complementary strands of DNA which differ at the RFLP sites.

The *ABO* genotypes for the patients analysed by flow cytometry are presented in Chapter 3 (Table 3.2). One patient (OT17) had an allele dosage shift (lanes 8 and 9; Figure 4.5). No differences in allele dosage were seen for 41 O^1 heterozygous patients, however, there were alterations seen for the leukaemic cell lines Jurkat and KCL22 (Figure 4.5).

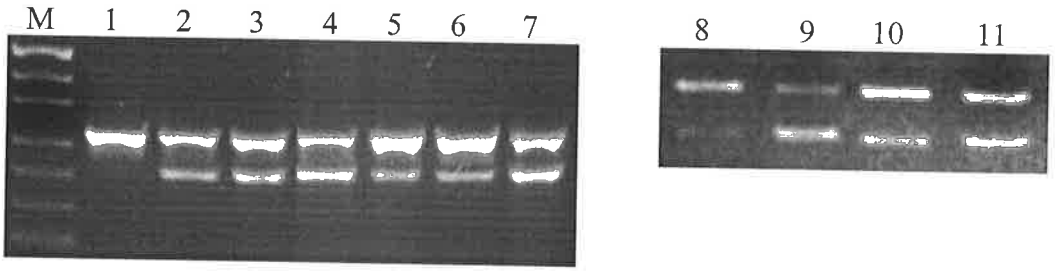


Figure 4.5. *ABO* allele dosage changes.

M is the pUC19/*Hpa*II marker and lane 1 is uncut AO1 PCR product. Lanes 2, 4, 6, 8 and 10 are AO1 PCR products digested with *Kpn*I while lanes 3, 5, 7, 9 and 11 are AO1 PCR products digested with *BstE*II. Lanes 2 and 3 are a paired sample showing equal allele dosage, while lanes 4 and 5 are from the leukaemic cell line Jurkat, which has more cutting with *Kpn*I but less cutting with *BstE*II. Lanes 6 and 7 are from KCL22 which has a shift in *ABO* allele dosage seen as less cutting with *Kpn*I but more cutting with *BstE*II. Lanes 8 and 9 are from OT17 which by flow cytometry has an increase in A resulting from a duplication of the A^2 allele (Figure 3.10A), which is consistent with digestion results of more cutting with *BstE*II (lane 9) and minimal cutting with *Kpn*I (lane 8). Lanes 10 and 11 are digests of a patient sample with no allele dosage shifts.

4.2.2 Expression of *ABO*

ABO genotyping not only allowed for the analysis of LOH at this locus, but the genotype meant that *ABO* allelic expression could be analysed as the majority of patients were heterozygous. *ABO* allelic expression was analysed using a RT-PCR RFLP technique. A limited number of groups have assessed *ABO* mRNA expression by RT-PCR. Orntoft *et al* (1996) used RT-PCR to detect *ABO* mRNA and found that it was present in normal urothelium and low-grade tumours but absent in high-grade human bladder tumours. They did not assess the relative levels of expression of the various *ABO* alleles.

Hosoi *et al* (1998) used RT-PCR to detect mRNA in two O blood group cell lines, both of which were CML cell lines at immature stages of haematopoietic differentiation: K-562 and KOPM-28. They digested the products with *Kpn*I, which confirmed that only the *O* allele was expressed.

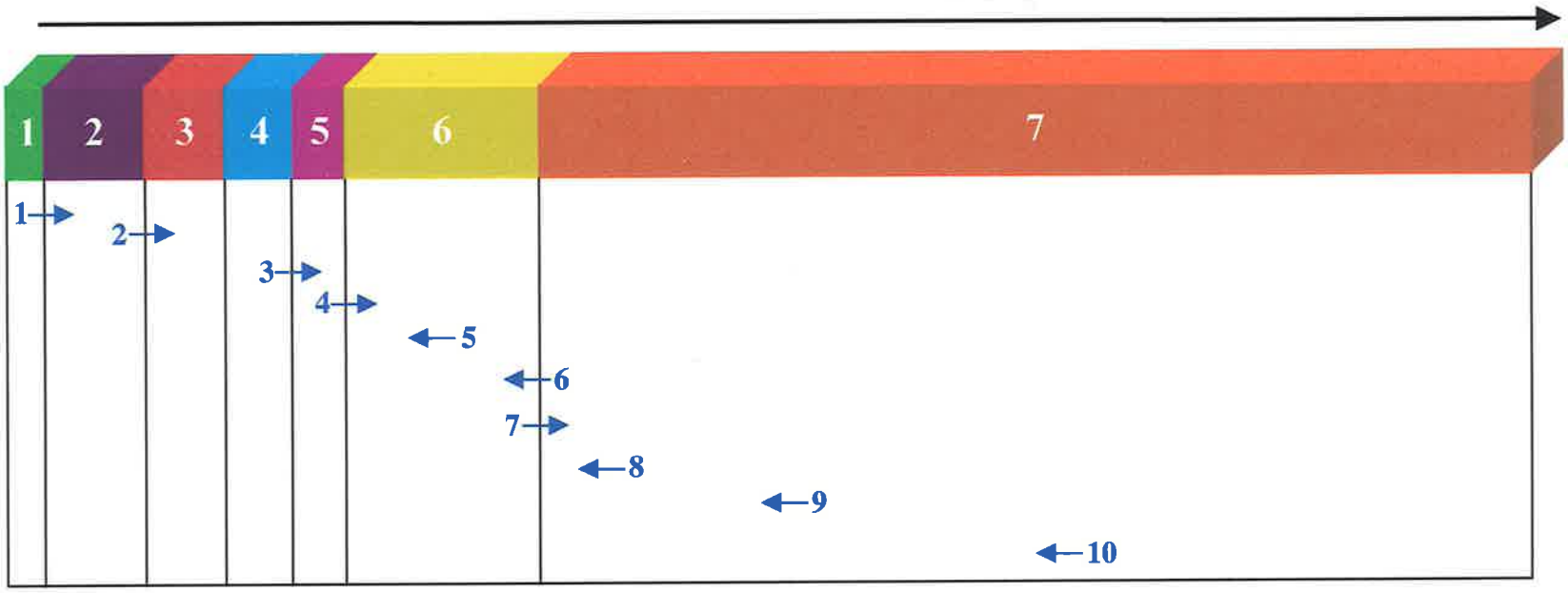
O'Keefe and Dobrovic (1996) used primers that amplified nucleotides 12-329 of *ABO* cDNA spanning the G deletion present in the O^1 allele. The 316 bp PCR product, when digested with *Kpn*I, gave a 249 bp band for the O^1 allele. For assessment of *ABO* expression it was decided that a RT-PCR RFLP method was the best way to assess the allelic expression of *ABO*.

4.2.2.1 Alternative Splicing of the *ABO* mRNA

The first aim was to assess the relative allelic expression in normal individuals so that any variations in the patients could be accurately observed.

Allelic expression of the *ABO* gene was analysed using RT-PCR followed by restriction enzyme digestion of the PCR products. The primers RT-AO1 F with AO1 R (see Table 2.1.1) were used to amplify a PCR product from cDNA of 236 bp. Figure 4.6 shows the location of all the primers used in the analysis of *ABO* expression. The primers RT-AO1 F and AO1 R together did not amplify genomic DNA due to the presence of the introns, that is, the primers were located in different exons. The PCR product spanned the deleted G/*Kpn*I restriction enzyme site in O^1 alleles, which was at position 169 of the PCR product. Digestion of the 236 bp product with *Kpn*I or *BstE*II allowed the discrimination of O^1 alleles and non-O^1 alleles. However, gel electrophoresis of the PCR products identified the presence of extra bands. Modification of the PCR conditions by the addition of the PCR specificity enhancers DMSO or betaine did not prevent the amplification of the extra bands (data not shown). The extra bands suggested that alternative splicing of the mRNA was taking place. As these additional products were also digested, allelic expression analysis was very complex.

The first primer combination used consisted of RT-AO1 F, which spans the exon 2 – exon 3 junction but the reverse primer, AO1 R, lies within exon 6. Hence, one new primer (RT-AO1 R) was designed in an attempt to diminish alternative splicing as this new primer spans the exon 6 – exon 7 junction. Another primer, PA5', which is a forward primer used previously for expression analysis, was tried as no problems with alternative splicing were reported (Denise O'Keefe, PhD thesis, 1995).



	Primer Name
1	PA5'
2	RT-AO1 F
3	RT-AO ex 4/5 F
4	RT-AO ex 5/6 F
5	AO1 R
6	RT-AO1 R
7	RT-ABO ex 6/7 F
8	RT-AO ex 7 R
9	ABO x 7 R
10	ABO2 R

Figure 4.6. Primers used in *ABO* expression analysis.

The location of all the primers utilised in the analysis of *ABO* expression by RT-PCR is shown. The numbers in the coloured boxes indicate the exon number. The arrows depict the primer location. The direction of the arrow denotes whether the primer is in the forward or reverse orientation.

Using the four primers, RT-AO1 F, AO1 R, PA5' and RT-AO1 R in all possible combinations, that is, PA5' and RT-AO1 R, PA5' and AO1 R, RT-AO1 F and RT-AO1 R, and RT-AO1 F and AO1 R yielded PCR products of sizes 376 bp, 313 bp, 299 bp and 236 bp respectively. The various primer combinations contained at least one primer that spanned an exon-exon boundary and thus amplification of genomic DNA did not occur as verified by the lack of amplification of the genomic DNA controls. All primer combinations spanned the deleted G region to allow for the analysis of allelic expression by digestion with *Kpn*I and/or *BstE*II. All four combinations of primers utilised the same PCR conditions with extra bands present for all the combinations when analysed by gel electrophoresis (Figure 4.7).

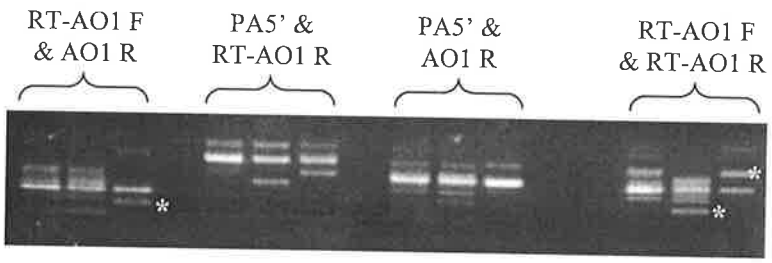


Figure 4.7. Alternative *ABO* transcripts.
The various primer combinations used are shown above the lanes and the three samples for each primer combination consist of a peripheral blood stem cell sample followed by two peripheral blood mononuclear samples. The same samples were used for each primer combination. The alternative transcripts that were sequenced are indicated with a white star and described in Figure 4.9 1A, 2A and 2B.

ABO is expressed in normoblasts of the erythroid lineage (Wada *et al*, 1990) (Figure 4.8). As bone marrow (BM) controls were scarce, it was hoped that the peripheral blood stem cell (PBSC) samples would be equivalent to BM because they are enriched for haematopoietic precursor cells normally only found in the BM.

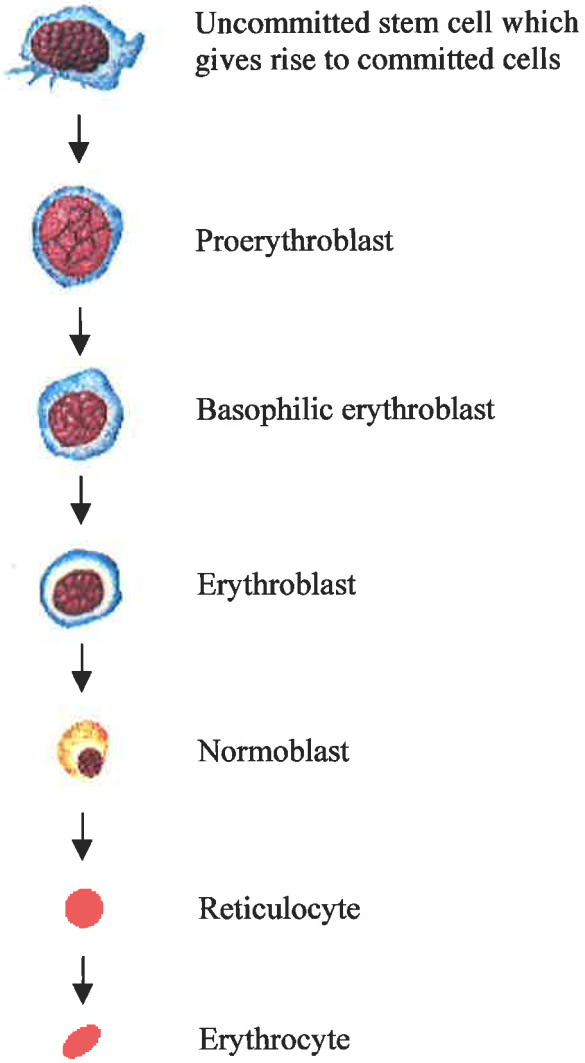


Figure 4.8. The erythroid lineage.

ABO is expressed in the normoblasts

Products not corresponding to the expected size for a given primer combination were isolated by band stabbing, re-amplified by PCR and sequenced. The smaller than expected band amplified using the RT-AO1 F and AO1 R combination was identified as having exon 4 completely spliced out (Figure 4.9 1A).

Sequencing of the lower molecular weight band from PCR amplification with RT-AO1 F and RT-AO1 R identified a transcript with exons 4 and 5 spliced out (Figure 4.9 2A). In addition a higher molecular weight band than the expected product from RT-AO1 F and RT-AO1 R was observed, which had exon 3, exon 4 and exon 5 sequence (exon 5 is nucleotides 2040-2075 of Genbank sequence AC000397) but then had an insert which corresponded to retention of part of intron 5 (1641-1700 of AC000397) followed by exon 6 sequence (1351-1485 of AC000397) (Figure 4.9 2B).

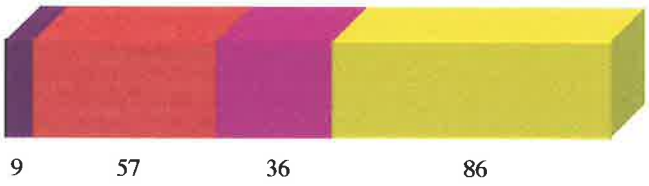
It appeared that the combination of PA5' and RT-AO1 R resulted in the least amount of alternative transcript amplification. An attempt was made to optimise this PCR by performing the PCR more stringently. However, many extra bands were still present. The presence of the alternatively spliced transcripts varied between samples, with some peripheral blood mononuclear (PB MNC) cDNAs giving extra bands while others gave only the one distinct and correct band (data not shown). Amplification of cDNA from the colon cancer cell lines LIM1215 and SW48 did not show extra bands, however, the leukaemic cell line K562 showed many extra bands, while KCL22 gave none (data not shown). In addition, PCR amplification of some patient cDNAs gave extra bands while others gave only a single product (data not shown). Additional products from the leukaemic cell line K562 and an O^lO^l patient were sequenced to determine the composition of the transcripts.

One of the K562 splice variants amplified by PA5' and RT-AO1 R had exons 4 and 5 spliced out (Figure 4.9 3A). The splice variant from the O^1O^1 patient amplified using PA5'and RT-AO1 R had exon 5 spliced out (Figure 4.9 3B). This same alternative transcript was also found for K562.

Based on the information regarding the composition of the alternatively spliced transcripts more primers were designed. As the alternative transcripts involved exons 4 and 5 an exon 4 – 5 junction primer (RT-AO ex 4/5 F) and an exon 5 – 6 junction primer (RT-AO ex 5/6 F) were designed (Figure 4.6). These exon junction primers were used with a new exon 7 primer (RT-AO ex 7 R) to ensure that the product was of sufficient size as the site of interest for distinguishing $A/B/O^2$ alleles from the O^1 allele lies in the beginning of exon 6 and the product had to be large enough to digest.

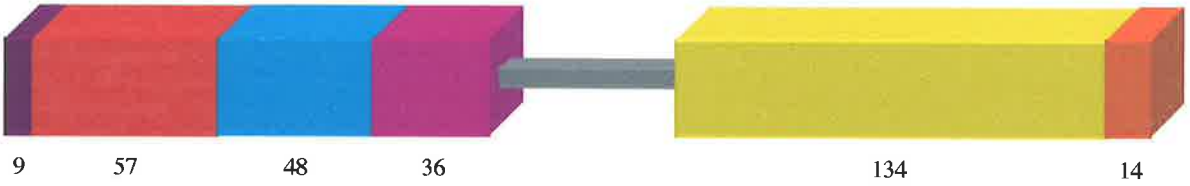
Amplification with RT-AO ex 4/5 F and RT-AO ex 7 R gave a PCR product of 248 bp while RT-AO ex 5/6 F and RT-AO ex 7 R gave a PCR product of 220 bp. Extra products were still amplified with the RT-AO ex 4/5 F primer but these were not as apparent with the RT-AO ex 5/6 F primer. The extra bands from the ex 4/5 combination are likely to be due to the absence of exon 4 as the 3' end of the primer will still bind to exon 5 and amplify, which would explain lower bands. The higher bands observed could be due to the retention of part of intron 5. The ex 5/6 primer would not be expected to pick up the intron insert, as the 3' end of the primer is complementary to the exon 6 sequence.

1A. RT - AO1 F & AO1 R



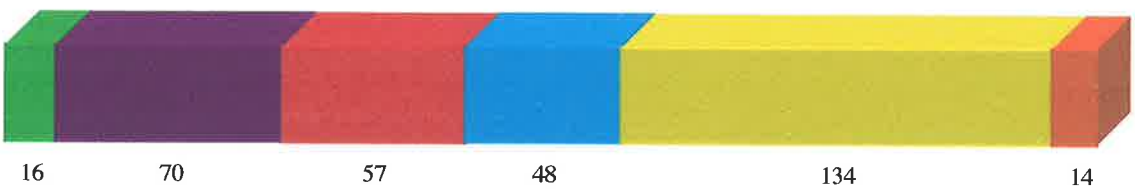
2A. RT - AO1 F & RT - AO1 R

2B.

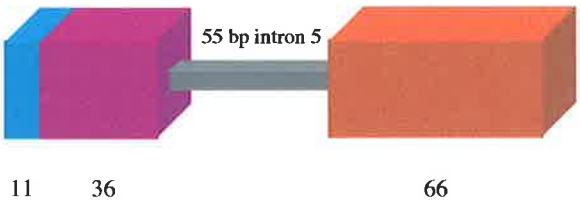


3A. PA5' & RT - AO1 R

3B.



4A. RT – AO ex 4/5 F & RT – AO ex 7 R



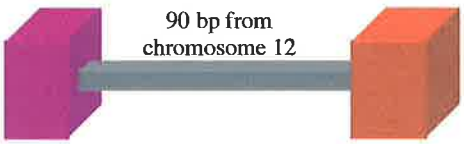
11 36 66

4B.

11 36 66

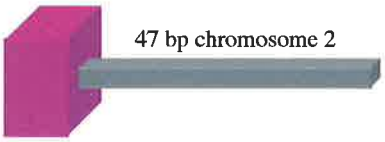
ABO Exon 1
ABO Exon 2
ABO Exon 3
ABO Exon 4
ABO Exon 5
ABO Exon 6
ABO Exon 7

5A. RT – AO ex 5/6 F & RT – AO ex 7 R



19 23

5B.



19

Figure 4.9. Alternative splice variants of *ABO*.

The alternative spliced variants generated from various combinations of primers in the RT-PCR analysis of *ABO* expression. The numbers underneath the exons (coloured boxes) indicate the size in bp.

When the PCR products from the two primer combinations, RT-AO ex4/5 F and RT-AO ex 7 R, and RT-AO ex 5/6 F and RT-AO ex 7 R, were resolved on a 3% gel, numerous extra bands were observed. Some of these extra bands were band stabbed, re-amplified and sequenced to determine whether these were splice variants or non-specific products (Figure 4.10).

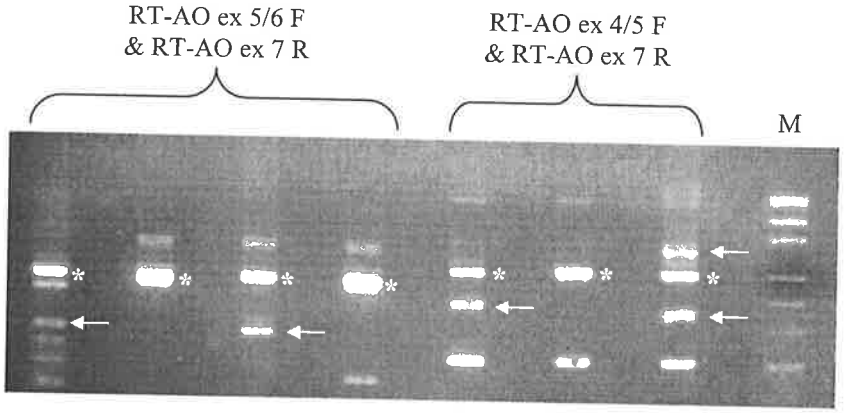


Figure 4.10. *ABO* alternative transcripts.
M is pUC19/*Hpa*II marker with the white stars indicating the correct size of the PCR products. The white arrows indicate the bands that were band stabbed, re-amplified and sequenced.

The sequence obtained from the higher molecular weight bands amplified with RT-AO ex4/5 F and RT-AO ex 7 R was the expected sequence, suggesting that these were heteroduplexes. Some of the extra bands sequenced from the RT-AO ex4/5 F and RT-AO ex7 R combination of primers included one variant with exon 6 spliced out and some of intron 5 (55 bp) immediately adjacent to exon 5 retained (Figure 4.9 4A). Another variant amplified with this combination of primers had only exon 6 spliced out (Figure 4.9 4B).

The RT-AO ex 5/6 and RT-AO ex 7 combination appeared to be the optimal primer combination, however, some extra bands were present in some samples. To determine what these were, some of them were sequenced (Figure 4.10). Sequencing of some of the variant bands amplified by the RT-AO ex 5/6 F and RT-AO ex7 R primer combination had unusual composition. One variant had some exon 5 sequence and exon 7 sequence due to the primers,

however, in between there was 90 bp with 97% homology to a BAC from 12q13-15 (nucleotides 128222-128135 from Genbank accession number AC024941) which is a different chromosomal region to *ABO* (Figure 4.9 5A). The other variant contained exon 5 sequence (primer) followed by 47 bp with 95% homology to a BAC clone from chromosome 2 (nucleotides 85039-85085 Genbank Accession number AC007040; Figure 4.9 5B) which is also a different chromosomal region to *ABO*. These products are therefore due to non-specific amplification.

4.2.2.2 *ABO* Expression in Normal PB MNCs and PBSCs

ABO cDNA from PB MNCs and PBSCs was PCR amplified with the primers RT-AO ex 5/6 F and RT-AO ex 7 R. As normal BM specimens were scarce, PBSC were used at it was thought that they would resemble normal BM controls because they represent mobilised stem cells. Previous attempts to amplify an *ABO* transcript from PB were unsuccessful (Denise O'Keefe, PhD thesis, 1995) but it was re-attempted in this study as RT-PCR has become much more sensitive. In O^1O^1 PB MNC samples only O^1 allele expression was detected. In BO^1 PB MNC samples the majority of *ABO* transcript expression came from the *B* allele and very little, if any, from the O^1 allele. *BB* PB MNC samples only cut with *BstE*II indicating that the *B* transcript alone was expressed. Unexpectedly, PB MNC samples from *A* genotypes showed no *A* allele expression.

In A^1O^1 and A^2O^1 PB MNC samples, the digestion pattern with *Kpn*I and *BstE*II was identical to that of an O^1O^1 individual, that is, only O^1 expression was seen. This was not expected since it had previously been shown that in A^1O^1 BM specimens that the A^1 transcript was predominant and very little, if any, O^1 transcript was present (Denise O'Keefe, PhD thesis, 1995). The cell type from which RNA was isolated could explain this discrepancy since PB MNCs do not express *ABO*. However, it remains unclear why the BO^1 PB MNC samples

show the expected allelic expression when the A^1O^1 PB MNC samples do not. Perhaps the different enhancers associated with the various alleles of *ABO* are responsible for the difference in expression of the *A* and *B* alleles in PB MNC samples. The *B* and *O* allele enhancer has recently been shown to increase transcription 300-fold more than the *A* allele associated enhancer (Yu *et al*, 2000). This may also explain why the *O* allele is expressed in *AO* samples even though the *A* allele is not.

As the RNA from the PB samples was isolated from MNCs, RNA isolation from whole blood using Trizol LS was performed. This was marketed as able to isolate RNA from the entire blood cell population and reticulocytes are known to express *ABO*. Investigation of A^1O^1 RNA taken from a Trizol LS extraction showed the same allelic expression as an A^1O^1 RNA isolated from PB MNCs. That is, there was only expression of the O^1 allele (data not shown). However, patient samples investigated at the same time expressed *A* demonstrating that PB MNCs were not suitable controls.

The lack of suitability of PB MNC samples as a normal control led to the investigation of PBSC as a normal equivalent to BM since they are enriched for precursor haematopoietic cells. RNA was extracted from PBSC harvested from seven breast cancer patients, as it was expected that these patients would not have any alteration in their haematopoietic cells. The PBSC samples were genotyped as described in 4.2.1. Five samples were genotype O^1O^1, one was A^1O^1 and another was A^2O^2. Again PBSC from the A^1O^1 individual did not express the A^1 allele and the A^2O^2 PBSC sample did not express A^2 (data not shown). It was considered that perhaps the primers were not sensitive enough so 11 different combinations were tried using RT-AO1 F, AO1 R, PA5', RT-AO1 R, RT-AO ex4/5 F, RT-AO ex5/6 F, ABO x7 R (see 2.1.1 and Figure 4.6) all under the same PCR conditions.

Of the eleven combinations, the best combination, which did not amplify alternative transcripts and gave a very good yield (at 40 cycles), was RT-AO ex 5/6 F and RT-AO1 R with a product size of 168 bp and the *KpnI*/*BstE*II at position 38 of the PCR product. This combination was then used instead of RT-AO ex 5/6 F and RT-AO1 ex 7 R. Some of the extra bands amplified from A^1O^1 PB MNC and the A^1O^1 PBSC samples with the various primer combinations were digested but expression of the *A* allele was never detected. Despite the many attempts to use PB MNC or PBSC samples as normal controls, it was clear that they were not adequate, hence the necessity for analysing BM specimens.

4.2.2.3 *ABO* Expression in BM

The only normal BM specimens available were those taken from breast cancer patients. These were deemed adequate as it was unlikely that the haematopoietic system in these patients was affected by the disease. Analysis of blood group A breast cancer BM specimens showed that the *A* allele was expressed in BMs but not in PB MNCs or PBSC specimens of the same genotype (Figure 4.11).

4.2.2.4 The RT-ABO PCR Product

To analyse allelic expression in patients that were not heterozygous for the O^1 allele required the development of a new combination of primers that spanned the nucleotide substitutions responsible for the A^2 and *B* alleles. The primers, RT-ABO ex 6/7 F and ABO2 R (Figure 4.6), generate a product that is 529 bp in length and digestion of this product with *Pvu*II and *Alu*I can differentiate the A^2 and *B* alleles respectively.

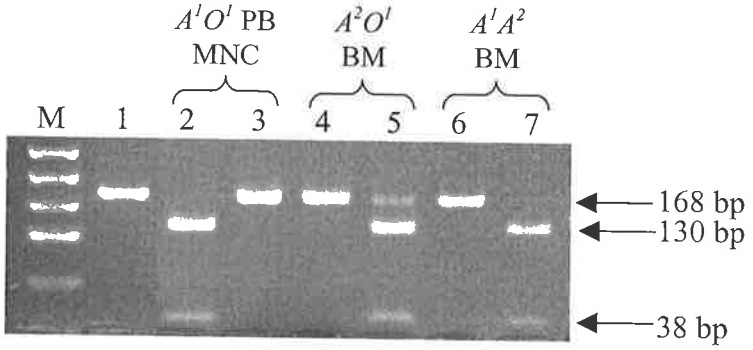


Figure 4.11. *A* expression in PB MNCs vs BM.

Lane M is the pUC19/*Hpa*II marker while lane 1 is the uncut RT-AO ex5/6 F and RT-AO ex 6/7 R PCR product. Lanes 2 and 3 are an A^1O^1 PB MNC sample while lanes 4 and 5 are an A^2O^1 BM and lanes 6 and 7 are an A^1A^2 BM. Lanes 2, 4 and 6 are *Kpn*I digests while lanes 3, 5 and 7 are *BstE*II digests. The PB A^1O^1 is completely digested with *Kpn*I and not *BstE*II hence only the O^1 allele is expressed. The BM specimens are digested with *BstE*II and not *Kpn*I indicating that the *A* alleles are expressed.

4.2.2.5 *ABO* Allelic Expression in Patients

Analysis of *ABO* expression in patients with haematological malignancies is summarised in Table 4.7. The flow cytometry studies described in Chapter 3 showed that 16/29 A, B or AB patients showed alterations of A or B and/or H. Since loss of *ABO* expression is expected to affect loss of A or B and not H, which is encoded by a different locus, the number of patients with alterations at the *ABO* locus is 11/29. Of these eleven with loss of A or B, only 3 had loss at the mRNA level. One patient with loss of A, F10 (A^1O^1 genotype), had no expression of either the A^1 or O^1 alleles while another patient, F27, which had loss of A and H, was A^1O^1 genotype but only expressed the O^1 allele, as expected (Figure 4.12A). It was not surprising that some of the blood group A or B patients that had loss on the flow did not have complete loss of the A or B allele as the flow patterns indicated that the variant population was small and that the majority of the population was normal.

However, one A^1O^1 patient with complete loss of A, F51, still had A^1 allele expression although the flow analysis indicated that no A positive cells were present. This was surprising, however other mechanisms may result in *A* transferase mRNA being made but it may contain a mutation that renders the transferase inactive or some other post transcriptional control may be affected. Sequencing of exon 7 of this patient's PCR product did not detect any mutations, however, if a gene mutation is present it may lie elsewhere in the gene.

Another interesting patient, F60, was of genotype A^1A^2 with loss of A. There were two samples of different dates available for analysis of this patient. The flow analysis was performed in 1996 and RNA from a specimen of this date failed to show any *ABO* expression, however, a sample taken 3 years prior was positive for *ABO* expression. Though there was no flow analysis performed for that sample it seems likely that loss of *ABO* occurred as the malignancy progressed from chronic to blast crisis CML (Figure 4.12B; Table 4.7).

Blood group A or B patients with loss of H would not be expected to have loss of *ABO* at the mRNA level but rather be affected at the *FUT1* locus. However, there were five blood group A and B patients with loss of H only and two of these, F39 and F53, had loss of *A* allele expression (Table 4.7; Figure 4.12C).

There were 6 blood group O patients with loss of H and 5 were analysed. All had O^1 expression as expected since loss of the O^1 allele would not affect the levels of the H antigen which is determined by the *FUT1* gene.

Patient samples previously studied (Denise O'Keefe, PhD thesis, 1995; Table 4.7) were analysed using the above RT-PCR RFLP technique. The results from this analysis were in accord with those found previously. Loss of A at the allelic level was seen for 5 samples that had shown loss of A by serology (Figure 4.12C). The percentage of loss is higher in the group of patients previously studied as 50% or more of the cells showed loss while patient samples analysed by flow cytometry had some cells with only 20% loss. This demonstrates that detecting loss at the mRNA level is difficult since the majority of cells are not affected and will have normal bi-allelic expression.

There was also loss of *A* allelic expression in three other patients, two of which, F7 and F11, showed no loss in ABH antigens by flow cytometric analysis. The other, OT24, was a patient that had not been analysed by flow cytometry. The three samples, 2 A^1O^1 and 1 A^2O^1, all had loss of A, that is, a PCR product generated from their cDNA did not digest with *BstE*II but did for *Kpn*I. Their patterns resembled that of an O^1O^1 individual (data not shown). This result may be due to the population of cells sampled for flow cytometric analysis not being representative of the population of cells with loss of *A* allelic expression.

Myeloid Patients Analysed By Flow Cytometry

ID	DIAGNOSIS	GENO	FLOW	EXPRESSION
F1	AML M1	O1O1	**NO LOSS**	nd
F2	MDS	A1A2	**LOSS of A**	+
F3	AML M4	A1O1	**Transfusions**	mostly A1
F4	AML M1	O1O1	**LOSS of H**	O1
F5	MDS	nd	**Transfusions**	nd
F6	AML NR	nd	**NO LOSS**	nd
F7	AML M2	A2O1	**NO LOSS**	O1
F8	MDS	nd	**NO LOSS**	nd
F9	AML	A1O1	**LOSS of A & H**	little A1
F10	MDS	A1O1	**LOSS of A**	-
F11	AML M1	A1O1	**NO LOSS**	O1
F12	RAEB-T/AML M4	BO1	**LOSS of B & H**	nd
F13	AML M6	O1O1	**LOSS of H**	O1
F14	AML M2	A1O1	**NO LOSS**	A1 & O1
F15	CML blast crisis	A1A2	**NO LOSS**	A1 & A2
F16	MDS	nd	**NO LOSS**	nd
F17	MDS	A1O1	**NO LOSS**	A1
F18	CML ?blast crisis	nd	**NO LOSS**	nd
F19	AML	O1O1	**NO LOSS**	O1
F20	CML	A1O1	**LOSS of H**	A1
F21	CML blast crisis	A1O1	**Transfusions**	A1 & O1
F22	AML	O1O1	**NO LOSS**	O1
F23	CML chronic	A2B	**NO LOSS**	A2 & B
F24	AML	BO1	**LOSS of B**	B & O1
F25	MDS	A1A1	**LOSS of A**	A1
F26	AML M3	A1O1	**NO LOSS**	A1 & O1
F27	AML M4	A1O1	**LOSS of A & H**	O1
F28	AML	nd	**NO LOSS**	nd
F29	AML	O1O1	**NO LOSS**	nd
F30	CML chronic	BO1	**LOSS of H**	B & O1
F31	AML M5	O1O1	**NO LOSS**	O1
F32	AML M3	O1O1	**NO LOSS**	O1
F33	CML chronic	nd	**NO LOSS**	nd
F34	AML?MDS	nd	**NO LOSS**	nd
F35	AML M4Eo	nd	**NO LOSS**	nd
F36	MDS	nd	**NO LOSS**	nd
F37	AML	O1O1	**NO LOSS**	O1
F38	AML M1	A1A1	**NO LOSS**	A1
F39	AML M4	A2O1	**LOSS of H**	O1
F40	AML M2	nd	**Transfusions**	nd
F41	AML M2	O1O1	**LOSS of H**	nd
F42	AML M3	BO1	**NO LOSS**	B & O1
F43	AML M1	O1O1	**LOSS of H**	O1
F44	MDS	O1O1	**NO LOSS**	O1
F45	CMML	O1O1	**NO LOSS**	nd
F46	AML M3	A1O1	**LOSS of A**	A1 & O1
F47	MDS	nd	**NO LOSS**	nd
F48	AML M5	O1O1	**LOSS of H**	O1
F49	AML M2	nd	**NO LOSS**	nd
F50	AML M2	O1O1	**LOSS of H**	O1
F51	AML M4	A1O1	**LOSS of A**	A1
F52	MDS	A1B	**LOSS of A**	+
F53	AML M1	A1O1	**LOSS of H**	O1
F54	AML M2	O1O1	**NO LOSS**	nd
F55	AML M4	nd	**NO LOSS**	nd
F56	AML M2	nd	**NO LOSS**	nd
F57	AML M7	BO1	**LOSS of H**	B & O1
F58	AML NR	O1O1	**NO LOSS**	nd
F59	AML NR	O1O1	**NO LOSS**	O1
F60	CML blast crisis	A1A2	**LOSS of A**	'93 +, '96 -
F61	AML NR	nd	**NO LOSS**	ND

Other Patients Analysed By Flow Cytometry

ID	DIAGNOSIS	GENO	FLOW	EXPRESSION
OT1	ALL	O1O1	**NO LOSS**	nd
OT2	ALL	A1O1	**NO LOSS**	A1 & O1
OT3	ALL	A1A1	**NO LOSS**	A1
OT4	ALL	O1O1	**NO LOSS**	nd
OT5	ALL	A1O1	**NO LOSS**	A1 & O1
OT6	ALL	nd	**NO LOSS**	nd
OT7	CLL	nd	**NO LOSS**	nd
OT8	CLL	nd	**NO LOSS**	nd
OT9	CLL	nd	**NO LOSS**	nd
OT10	CLL	nd	**NO LOSS**	nd
OT11	CLL	nd	**NO LOSS**	nd
OT12	CLL	nd	**NO LOSS**	nd
OT13	CLL	nd	**NO LOSS**	nd
OT14	CLL	nd	**LOSS of H**	nd
OT15	CLL	nd	**LOSS of H**	nd
OT16	CLL	A1O1	**LOSS of A**	A1
OT17	Heart Condition	A2A2O1	**INCREASE in A**	nd
OT18	AML	nd	**INCREASE in H**	nd

Other Patients Not Analysed By Flow

ID	DIAGNOSIS	GENO	EXPRESSION
OT19	unknown	O1O1	nd
OT20	MDS	O1O1	O1
OT21	CML	O1O1	O1
OT22	CML	A1B	A1 & B
OT23	AML	A1O1	A1 & O1
OT24	unknown	A1O1	O1
OT25	AML	BO1	nd
OT26	unknown	A2O1	A2
OT27	unknown	O1O1	O1
OT28	unknown	A2O1	nd
OT29	unknown	A1O1	A1

Previously Studied Patients

ID	DIAGNOSIS	GENO	SEROLOGY BY DOK	EXPRESSION
DOK1	RAEB	A2O1	weak A2 expression	O1
DOK2	MDS	A1O1	loss of A	O1
DOK3	CML	A2O1	loss of A	A2
DOK4	AML M1	A1B	mfr with anti-A,B, & AB	nd
DOK5	AML M2	A1O1	mfr->only H antigen	O1
DOK6	AML M3	A2B	mfr with anti-A	B
DOK7	AML	BO1	mfr with anti-B	B & O1
DOK8	AML M0	A1O1	loss of A	O1

Normals

ID	DIAGNOSIS	GENO	EXPRESSION
N1	Normal PB MNC	BO1	B & O1
N2	Normal PB MNC	O1O1	nd
N3	Normal PB MNC	O1O1	O1
N4	Normal PB MNC	BO1	nd
N5	Normal PB MNC	A2O1	nd
N6	Normal PB MNC	A2O1	nd
N7	Normal PB MNC	A1O1	O1
N8	Normal PB MNC	A1O1	O1
N9	Normal PB MNC	O1O1	O1
N10	Normal PB MNC	A2O2	O2
N11	Normal PB MNC	A2O1	O1
N12	Normal PB MNC	A1O1	nd
N13	Normal PB MNC	A1B	nd
N17	Normal PB MNC	BO1	nd
N18	Normal PB MNC	BO1	B & O1
N19	Normal PB MNC	O1O1	O1
N20	Normal PB MNC	A2O1	O1
N21	Normal PB MNC	O1O1	O1
N22	Normal PB MNC	BB	B
N23	Normal PB MNC	A1B	nd
N24	Normal PB MNC	O1O1	O1
N25	Normal PB MNC	O1O1	O1
N26	breast cancer PBSC	O1O1	O1
N27	breast cancer PBSC	O1O1	O1
N28	breast cancer PBSC	O1O1	O1
N29	breast cancer PBSC	A2O2	O2
N30	breast cancer PBSC	A1O1	O1
N31	breast cancer PBSC	O1O1	O1
N32	breast cancer PBSC	O1O1	O1
K562	leukaemic cell line	O1O1	O1
KCL22	leukaemic cell line	A2O1	-
SW48	colon cance cell line	A1B	A1 & B
LIM1215	colon cance cell line	BO1	B & O1
HEL	leukaemic cell line	O1O1	O1
HL60	leukaemic cell line	A1O1	nd
CCRF-CEM	leukaemic cell line	O1O1	-
EM2	leukaemic cell line	A1A1	-
RAJI	leukaemic cell line	O1O1	-
Jurkat	leukaemic cell line	O1O2	-
BM1	bone marrow	O1O1	O1
BM2	bone marrow	A1A2	+
BM3	bone marrow	A2O1	A2 & O1

Patients That Were Genotyped Only

ID	DIAGNOSIS	GENO
OT30	AML	O1O1
OT31	AML	A1A2
OT32	AML	A2O2
OT33	T4 CLL	O1O1
OT34	ALL	A1O1
OT35	AML	O1O1
OT36	CLL	BO1
OT37	ALL	O1O1
OT38	CLL	A1A1
OT39	CLL	O1O1
OT40	AML	BO1
OT41	CLL	A2O1
OT42	AML	A1O1

Table 4.7. *ABO* genotyping and expression analysis of patients and normals.
'nd' is not done. '+' refers to postivie expression of *ABO* but allelic expression not able to be determined. '-' refers to no expression of *ABO*. Expression column shows which allele(s) are expressed.

A

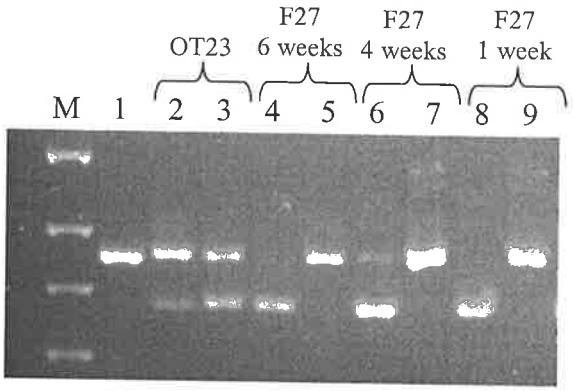
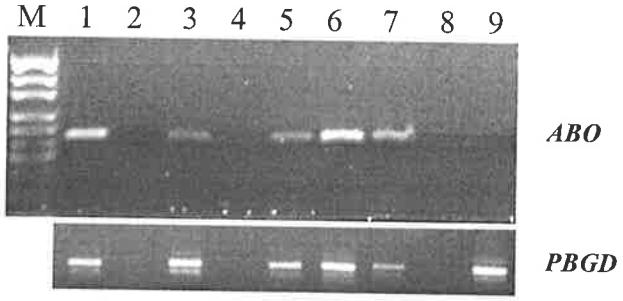


Figure 4.12. A. Loss of *A* expression by RT-PCR.

M is pUC19/*Hpa*II marker and lane 1 is uncut PCR product. Lanes 2, 4, 6 and 8 are digests with *Kpn*I while lanes 3, 5, 7 and 9 are digests with *BstE*II. Lanes 2 and 3 is sample OT23 which shows expression of the A^1 (digestion with *BstE*II) and O^1 (minimal digestion with *Kpn*I) transferase mRNA. Lanes 4 and 5 are sample F27 6 weeks before death, lanes 6 and 7 are sample F27 4 weeks before death and lanes 8 and 9 are F27 1 week before death. All three samples only show O^1 allelic expression, that is digestion with *Kpn*I only, despite genotyping as A^1O^1.

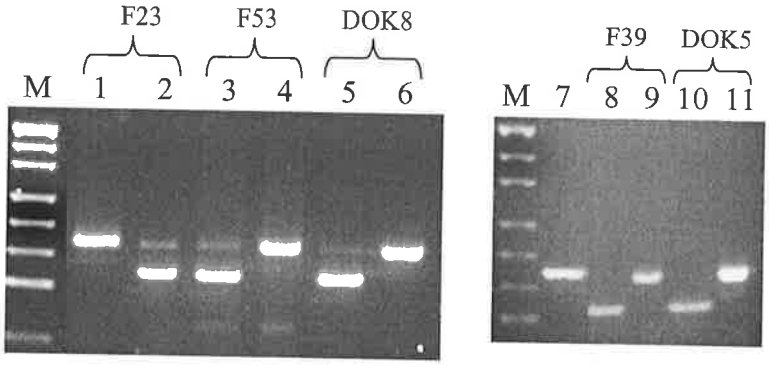
B.



B. *ABO* expression in patient F60 at two different time points.

M is pUC19/*Hpa*II marker, lane 1 is patient OT29 while lane 2 is a negative control. Lane 3 is F51 and lane 4 is a negative control. Lane 5 is OT16 and lane 6 is F59. Lane 7 is a 1993 sample of F60, lane 8 is a negative control and lane 9 is a 1996 sample of F60. The bottom gel is the corresponding *PBGD* (housekeeping gene) for each of the cDNA samples to control for differences in cDNA amounts.

C.



C. Loss of *A* in loss of H patients.

M is pUC19/*Hpa*II marker while lane 7 is uncut PCR product. Lanes 2, 4, 6, 8 and 10 are PCR products digested with *Kpn*I while lanes 1, 3, 5, 9 and 11 are PCR products digested with *BstE*II. Lanes 1 and 2 are from cDNA of F23, an A^2B genotype, hence no cutting with *Kpn*I is expected. Lanes 3 and 4 are F53, an A^1O^1 patient, and lanes 8 and 9 are F39, an A^2O^1 patient, and both had loss of H on the flow but have loss of *A* at the mRNA level. Lanes 5 and 6 are DOK8, and lanes 10 and 11 are DOK 5. Both are A^1O^1 genotype and have loss of *A*, which is in agreement with previous results (O'Keefe PhD thesis, 1995).

4.2.3 Expression of *FUT1*

4.2.3.1 Introduction

Cytometric analysis of the ABH antigens on the red blood cells of patients with haematological malignancies indicated that loss of H was responsible for the loss of A and/or B in 8/29 patients (Chapter 3). Loss of H was also seen in 6/28 patients of blood group O, hence there were 14 patients with alterations of H. Therefore, it was decided that an investigation of the *FUT1* locus, which encodes a fucosyltransferase that acts on a precursor molecule to form the H antigen, was necessary since this cohort of patients accounted for 25% of changes seen in patients with haematological malignancies.

The importance of looking at the *FUT1* gene is strengthened by the report that anti-sense transfection of an *FTA* cDNA fragment, the rat *FUT1* homologue, into a rat colon carcinoma cell line resulted in reduced cell surface H antigen. This cell line was significantly more tumorigenic in syngeneic animals (Hallouin *et al*, 1999). Gorelik *et al* (1997) transfected the cDNA of the mouse homologue of *FUT1* into a melanoma mouse cell line and reported H antigen expression and a significant reduction in the metastatic ability of the transfected cells in immunocompetent and immunosuppressed (X-irradiated) mice. Aubert *et al* (2000) demonstrated that over-expressing *FUT1* in a human pancreatic tumour cell line, which originally expressed low amounts of the fucosyltransferase, resulted in an increase in H antigen expression and a significant decrease in the metastatic potential.

Loss of the H antigen has been reported in numerous carcinomas, including those of the urinary bladder, uterine cervix, lung, ovary and oral squamous cells, and is associated with stage and grade of the tumour, prognosis and metastatic ability (Juhl *et al*, 1986; Sakamoto *et al*, 1986; Ogawa *et al*, 1987; Metoki *et al*, 1989 Bryne *et al*, 1991). Yang (1990) found that deletion of the H antigen in transitional cell carcinomas predicted recurrence and mortality.

The H antigen expression of the blasts of 123 AML patients was investigated and it was found that reactivity of the H antigen antibody had prognostic value, predicted clinical remission and probability of survival (Schwarzinger *et al*, 1990). Blasts from patients that negatively stained for the H antigen correlated with an increased likelihood of clinical remission and a higher survival rate. These results are contradictory to those of solid tumours where expression of the H antigen is a favourable prognostic marker. This indicates that the mode of function of the H antigen differs between solid tumours and haematological malignancies. Ivanov *et al* (1990) investigated 94 leukaemia patients and found that the level of H antigen was decreased in all types of leukaemia, but the levels of A and B antigen were unchanged when compared to 1014 normal controls.

All the above studies indicate the importance and significance of the alterations of the H antigen in solid tumours and leukaemia, hence investigation of *FUT1* was deemed important, as it was associated with patient survival and disease progression. Though there were difficulties with analysing the *ABO* gene at the mRNA level due to the alternative splicing, investigation of *FUT1* mRNA expression was originally investigated by non-radioactive Northern analysis based on the results of Koda *et al* (1997).

However, the results of Koda and colleagues (1997, 1998) indicated that the *FUT1* gene had three transcription start sites with the use of each dependent on the stage of differentiation of the human erythroid lineage. Since loss of H occurs on red blood cells, which are derived from the erythroid lineage, another aim was to test whether there were changes in the use of the transcription start sites in patients with haematological malignancies.

4.2.3.2 Northern Analysis of *FUT1*

Koda *et al* (1997) analysed *FUT1* mRNA expression by Northern blot analysis using a digoxigenin-labelled *FUT1* antisense RNA probe of the complete coding region. Two primers were designed, FUT1 Northern Probe F and FUT1 Northern Probe T7 R so that antisense *FUT1* RNA transcripts would be generated using the T7 polymerase. The Koda *et al* (1997) Northern hybridisation experiment indicated that a single transcript was detected in HEL and this cell line was used as a control in the studies described here (Figure 4.13A).

Our initial experiments indicated that more than one product was found in HEL cells, however, only two transcripts were found in normal PB MNC samples (6 tested; Figure 4.13B). The Northern analysis of patient RNA was less clear as multiple transcripts could be seen and further investigation of the literature found reports from Koda *et al* (1998) that displayed a Northern of HEL identical to ours, that is, more that one transcript was present

(Figure 4.13C). This group also reported that there were three transcription start sites for *FUT1*, with the site used determined by the stage of differentiation of the human erythroid lineage. The transcription sites utilised in patients with haematological malignancies, and whether there were differences in the loss of H group, was subsequently investigated by RT-PCR which would enable the analysis of the three different transcripts which used alternative promoters.

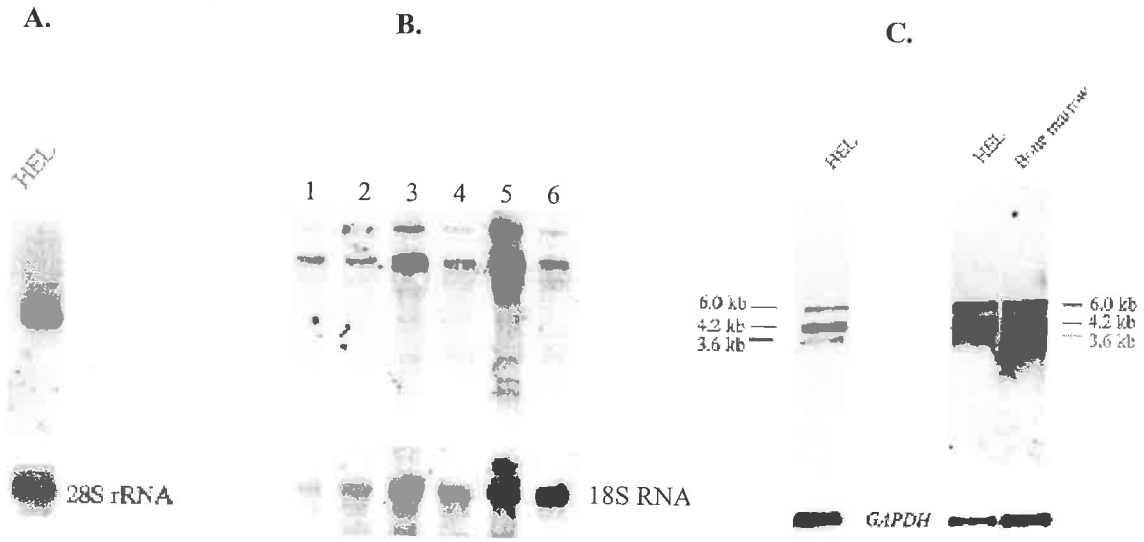


Figure 4.13. *FUT1* Northerns.

A. *FUT1* Northern taken from Koda *et al* (1997). A single band was seen for the cell line HEL, even though the same probe was used for the Northern in Figure 4.13B. 28S used as a control for loading.

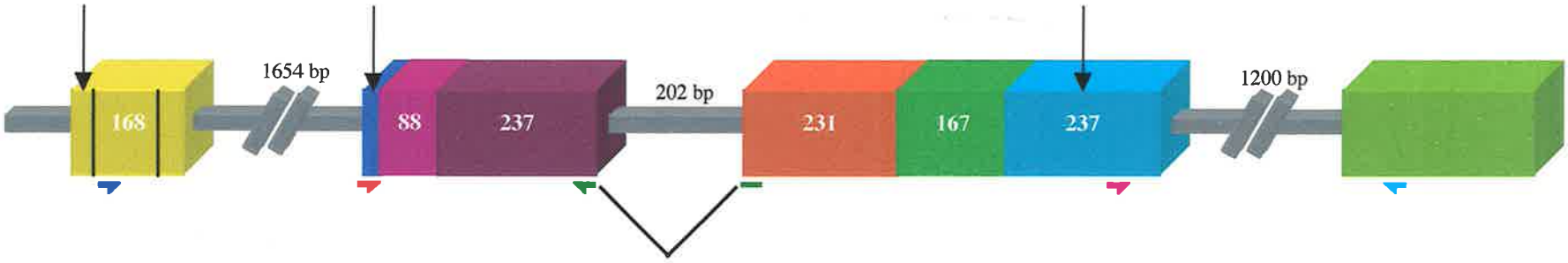
B. Lane 1 is 2 μg of PB MNC RNA, lane 2 is 4 μg of PB MNC RNA, lanes 3 and 4 are the same PB MNC sample, lane 3 is 5 μg and lane 4 is 1 μg. Lanes 5 and 6 are the leukaemic cell line HEL but lane 5 is 5 μg and lane 6 is 2 μg. 18S used as a control for loading.

C. *FUT1* Northern taken from Koda *et al* (1998) with poly(A)-rich RNA. The HEL *FUT1* hybridisation resembles the one in Figure 4.13B with the pattern seen for the BM specimen similar to that for patient RNA samples (data not shown). GAPDH used as a control for loading.

4.2.3.3 RT-PCR Analysis of the Alternative Transcripts of *FUT1*

There were only two reports on the nature of the alternative transcripts of *FUT1* (Koda *et al*, 1997 and 1998) and this made the design of the RT-PCR quite complex. Using the two published reports, the primers designed for the exon 1 containing transcript and its alternative spliced forms, were FUT1 ex 1 F and FUT1 ex 4/5 R which produce PCR products of 305 and 341 bp (Figure 4.14). An additional primer, FUT1 ex 2 F, was used in conjunction with the primer FUT1 ex4/5 R, for the exon 2 containing transcripts and its alternatively spliced forms. The primers for the *FUT1* exon 2 transcripts produce two PCR products of 273 and 361 bp. Two other primers, FUT1 ex 7 F and FUT1 ex 8 R were designed for the exon 7 transcript, however, this primer combination, which yielded a PCR product of 206 bp, also generated PCR products from the exon 1 and exon 2-containing transcripts. Using the primers in a multiplex PCR and using a subtractive approach, the signal from the exon 1 and exon 2 transcripts were subtracted from the signal from all three transcripts to give the exon 7 transcript signal.

A preliminary study, to test the reliability and reproducibilty of this technique, was carried out using a multiplex PCR, in triplicate. CDNA was amplified from the cell lines HEL, which was published to have exon 2 and exon 7 transcripts, K562 which was published to have only exon 1 transcripts and Jurkat, which in this experiment was found to only produce a product with the exon 7 primers and hence had only the exon 7 transcript (Figure 4.15) (Koda *et al*, 1998).



	FUT1 ex 1 F
	FUT1 ex 2 F
	FUT1 ex 4/5 R
	FUT1 ex 7 F
	FUT1 ex 8 R

	Transcription start sites

	Exon 1
	Exon 2
	Exon 3
	Exon 4
	Exon 5
	Exon 6
	Exon 7
	Exon 8

Figure 4.14. The genomic structure of *FUT1*.

The different alternative splice variants and transcription start sites are based on Koda *et al* (1997, 1998). The positions of the primers used in the RT-PCR analysis of *FUT1* transcripts are shown. Numbers inside the exons are the exon sizes. Exon 2 is 25 bp and not all of exon 8 is shown.

For the HEL cell line, which expresses only the exon 2 and exon 7 transcripts, the ratio of the three transcripts was reproducible in the same experiment (Figure 4.15). The limitations of the subtractive method became apparent when the two alternative exon 2 transcripts (273 bp and 361 bp bands; Figure 4.15) had a greater combined intensity than the exon 7 transcript band. Since we did not know whether one exon 2 alternative transcript was expressed more efficiently than the other, the subtractive method was no longer applicable.

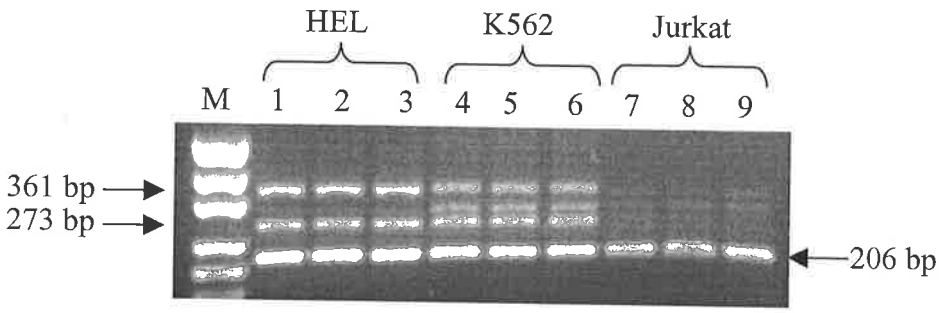


Figure 4.15. *FUT1* RT-PCR of the three alternative transcripts. M is the pUC19/*Hpa*II marker, while lanes 1-3 are the cell line HEL, lanes 4-6 are K562 and lanes 7-9 are Jurkat. Each PCR has been performed in triplicate. The 273 and 361 bp sizes are alternatively spliced forms of *FUT1* exon 2 containing transcripts. K562 has the two products from exon 2 and an exon 1 containing transcript (305 bp). The 206 bp band is from exon 7 containing transcripts.

For K562, only exon 1 containing transcripts were expected, as published, however both exon 1 and exon 2 transcripts were detected. The Jurkat cell line only expressed the exon 7 transcript. Due to these difficulties, the subtractive method was abandoned in favour of a different PCR method. Expression of the exon 1 transcript was subsequently determined in a single PCR using the primers FUT1 ex1 F and FUT1 ex4/5 R (Figure 4.16). Expression of the exon 2 transcript was determined in a single PCR using the FUT1 ex2 F and FUT1 ex4/5R primers (Figure 4.17).

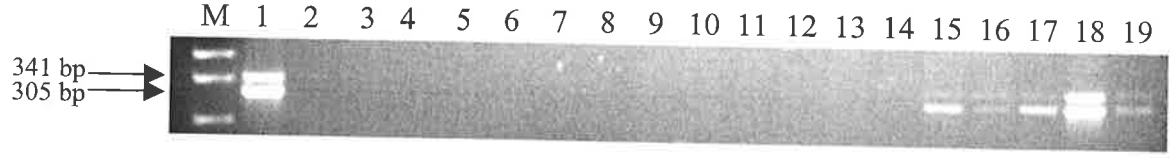


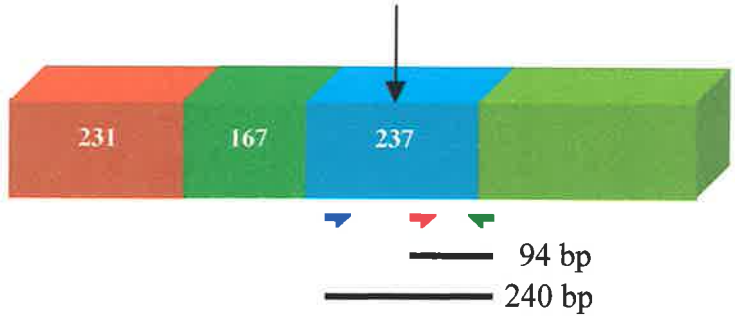
Figure 4.16. *FUT1* exon 1 transcript expression.

M is pUC19/*Hpa*II marker and lane 1 is a positive control which is PMC42, a breast cancer cell line. Lanes 2-6 are PBSC samples and lanes 7 – 12 are normal PB MNC samples. Lanes 13 – 17 and 19 are the leukaemic cell lines CCRF-CEM, EM2, HEL, Jurkat, KCL22 and Raji respectively. Lane 18 is the colon cancer cell line LIM1215. The arrows indicate the two exon 1 alternative transcripts of *FUT1*. Patient samples showed no expression as seen for the PBSC and normal PB MNC.

Figure 4.17. *FUT1* exon 2 transcript expression.

M is pUC19/*Hpa*II marker and lane 1 is a PBSC sample. Lanes 2-7 are PB MNC samples. Lanes 8 and 10 are BM specimens while lanes 9 and 11 are negative controls. Lanes 12 – 19 are patient samples. Lane 12 is OT2, lane 13 is F2, lane 14 is F3, lane 15 is DOK1, lane 16 is OT3, lane 17 is F4, lane 18 is DOK2 and lane 19 is F7. The arrows indicate the two exon 2 alternative transcripts of *FUT1* which are 273 and 361 bp.

For determining the expression of the *FUT1* exon 7 transcript, three new primers were designed: FUT1 ex 7 bef TS F, FUT1 ex 7 aft TS F and FUT1 ex 7/8 R (Table 2.1.1; Figure 4.18). FUT1 ex 7 bef TS is a primer placed before the exon 7 transcription start site and when used with FUT1 ex 7/8 R will yield a PCR product of 240 bp that will only be amplified from exon 1 and exon 2 containing transcripts. FUT1 ex 7 aft TS is a primer that lies after the exon 7 transcription start site and when used with FUT1 ex 7/8 R will yield a PCR product of 94 bp that will be amplified from exon 1, 2 and 7 containing transcripts. The two PCRs are performed on the same sample and if there is no PCR product generated with the primer situated before the exon 7 transcription start site, but a PCR product is seen with the primers after the exon 7 transcription start site then this sample expresses only the exon 7 transcript (Figure 4.19A).



	FUT1 ex 7 bef TS F
	FUT1 ex 7 aft TS F
	FUT1 ex 7/8 R

	Exon 5
	Exon 6
	Exon 7
	Exon 8

↓	Transcription start sites

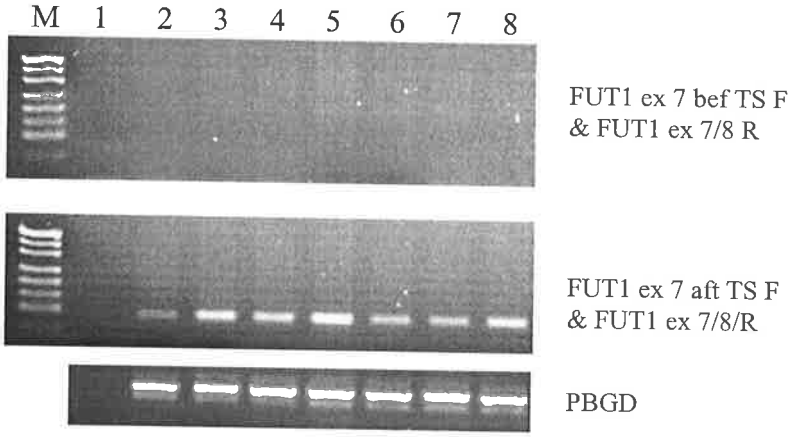
Figure 4.18. Position of RT-PCR primers for the exon 7 transcript of *FUT1*.

The primers used in the RT-PCR analysis of *FUT1* exon 7 transcripts are shown. Numbers inside the exons are the exon sizes. Black bars underneath the primers represent the PCR product and their sizes. Not all of exon 8 is shown.

Analysis of the various *FUT1* transcripts for all samples is summarised in Table 4.8. One of the surprising features was that the normal BM specimens expressed *FUT1* exon 2 and exon 7 transcripts while the patient samples predominantly expressed *FUT1* exon 7 transcripts (Figure 4.19B). *FUT1* expression was analysed for 43 patients for which only 1, F17, was positive for exon 1 transcripts (Table 4.8). This patient was a MDS patient with no loss of ABH antigens by flow and no loss of *ABO* at the mRNA level. K562 expressed transcripts containing exon 1 of *FUT1* and is thought to be a multipotential haematopoietic cell line able to differentiate into the progenitor for erythrocytes, monocytes and granulocytes (Andersson *et al*, 1979). MDS is characterised by an increase in cells of the myeloid lineage that do not reach maturity (see section 1.3.2.3) hence it is not surprising that the patient with MDS expresses exon 1 *FUT1* transcripts. Other MDS patients were analysed, but the sample size was to small to determine if *FUT1* exon 1 containing transcripts are common to most MDS's.

For exon 2 *FUT1* transcripts, greater variation in expression was observed, with 11/43 patients expressing this transcript (Table 4.8), which is thought to be expressed from cells at an early stage of differentiation of the erythroid lineage (Koda *et al*, 1998). Of the 11 patients that expressed the exon 2 containing transcripts, 3 had MDS, 3 had CML, 1 was RAEB and 4 had AML. There did not appear to be an association between disease and *FUT1* exon 2 expression, however, 4/11 patients showed some alteration of *ABO*. F24 had loss of B antigen by flow cytometry and F30 and F48 had loss of H by flow cytometry analysis. DOK1 had loss of A by serology and mRNA analysis.

A.



B.

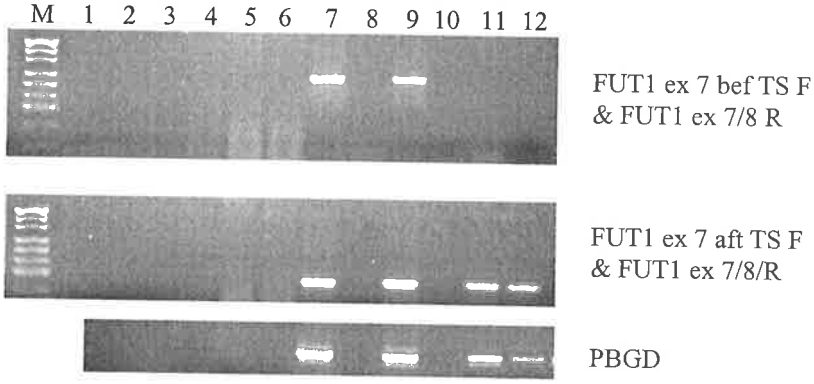


Figure 4.19. *FUT1* exon 7 transcripts.

A. M is pUC19/*Hpa*II marker, lane 1 is a negative control, lane 2 is a PB MNC sample and lanes 3 – 8 are PBSC samples.

B. M is pUC19/*Hpa*II marker, lanes 1 – 3 are no template controls, lanes 4 – 6 are genomic DNA samples and lanes 8 and 10 are negative controls. Lanes 7 and 9 are BM specimens and lanes 11 and 12 are two patients, OT2 and F2 respectively, which have the expression pattern characteristic of the majority of patient samples analysed for exon 7 containing transcripts. PBGD is the control for ensuring the RT has worked and for loading.

Determining the significance of loss of H with the expression of a certain transcript was not clear, however, as the BM reflects the normal situation and has expression of exon 2 *FUT1* transcripts, it appeared that loss of exon 2 transcript expression was significant. There were 32/43 patients that failed to express the exon 2 containing *FUT1* transcripts. The 32 patients without exon 2 *FUT1* transcript expression consisted of 19 AML patients, 2 MDS patients, 4 CML patients, 3 ALL, 1 CLL, and 4 unknown. For the AML patients where subclass information was available, lack of *FUT1* exon 2 expression was seen in 4/4 M1, 3/4 M2, 2/2 M3, 3/4 M4, 1/2 M5, and 1/1 M6. The numbers for each subclass were too small to draw any conclusions about loss of exon 2 transcript expression and disease type or AML subclass. Using the information in Table 3.2 and Table 4.8 no association between *FUT1* exon 2 transcript expression with age or survival could be made. Samples from 8 patients exhibiting loss of H by flow cytometry were tested for exon 2 containing transcripts and 6 of these lacked exon 2 containing transcripts.

All patients expressed the exon 7 *FUT1* transcript, like normal bone marrow except for F39, OT3, OT27 and DOK2. F39 is a loss of H by flow cytometry patient with loss of *A* allelic expression while OT3 and OT27 show no *ABO* alterations. DOK2 had loss of *A* allele expression. The common observation for these 4 patients is that they lack expression of any *FUT1* transcripts.

ID	DIAGNOSIS	FUT1 ex1	FUT1 ex2	FUT1 ex7
F2	MDS	-	-	+
F3	AML M4	-	+	+
F4	AML M1	-	-	+
F7	AML M2	-	-	+
F11	AML M1	-	-	+
F13	AML M6	-	-	+
F14	AML M2	-	+	+
F15	CML blast crisis	-	+	+
F17	MDS	+	+	+
F19	AML	-	-	+
F21	CML blast crisis	-	-	+
F22	AML	-	-	+
F24	AML	-	+	+
F27	AML M4	-	-	+
F30	CML chronic	-	+	+
F31	AML M5	-	-	+
F32	AML M3	-	-	+
F37	AML	-	-	+
F38	AML M1	-	-	+
F39	AML M4	-	-	-
F43	AML M1	-	-	+
F44	MDS	-	+	+
F48	AML M5	-	+	+
F50	AML M2	-	-	+
F51	AML M4	-	-	+
F59	AML NR	-	-	+
F60 1993	CML blast crisis	-	-	+
F60 1996	CML blast crisis	-	-	+
OT2	ALL	-	-	+
OT3	ALL	-	-	-
OT5	ALL	-	-	+
OT16	CLL	-	-	+
OT20	MDS	-	+	+
OT21	CML	-	-	+
OT22	CML	-	+	+
OT23	AML	-	-	+
OT24	Unknown	-	-	+

OT26	Unknown	-	-	+
OT27	Unknown	-	-	-
OT29	Unknown	-	-	+
DOK1	RAEB	-	+	+
DOK2	MDS	-	-	-
DOK5	AML M2	-	-	+
DOK6	AML M3	-	-	+
N7	Normal	-	-	-
N10	Normal	-	-	-
N19	Normal	-	-	-
N23	Normal	-	-	-
N24	Normal	-	-	-
N25	Normal	-	-	-
N26	Normal	-	-	+
N27	Normal	-	-	+
N28	Normal	-	-	+
N29	Normal	-	-	+
N30	Normal	-	-	+
N31	Normal	-	-	+
BM1	Normal	-	+	+
BM3	Normal	-	+	+
CCRF-CEM	Leukaemic	-	-	-
EM2	Leukaemic	-	-	-
HEL	Leukaemic	+	+	+
Jurkat	Leukaemic	-	-	+
K562	Leukaemic	+	+	unknown
KCL22	Leukaemic	+	-	+
LIM1215	Colon caner	+	+	+
RAJI	Leukaemic	-	-	-

Table 4.8. *FUT1* alternative transcript expression.

+' is positive expression and '-' is negative expression.

4.3 Discussion

ABO genotyping is essential for investigating *ABO* allelic expression in patients with haematological malignancies. Furthermore, it can also give information about LOH. Allele dosage changes were only seen for one patient sample that did not have a haematological malignancy. This patient, OT17, presented with a mixed field and had an increase in A antigen staining when analysed by flow cytometry. Two other allele dosage changes were seen but both were in leukaemic cell lines. In this study no differences in genotype frequencies between the normal and AML populations were seen.

In this study, it has been demonstrated that *ABO* cDNA can be amplified from PB MNC but that the allelic expression is not a true reflection of the individual's genotype. Amplification of the *ABO* transcript had previously not been possible (O'Keefe PhD thesis, 1995) but RT and PCR technology has become more sensitive. In PB MNC, mRNA expression in O^1O^1 and BO^1 individuals was identical to expression in BM whereas A^1O^1 PB MNCs did not express the A^1 allele, unlike A^1O BM samples. This was not limited to the A^1 allele as the A^2 allele was not expressed in A^2O PB MNC or A^2O PBSC samples but was expressed in an A^2O^1 BM sample.

Alignment of *ABO* ests and use of the G1096A nucleotide to determine which est came from which allele, showed that the 3' UTR of the *A/O* alleles are longer than the *B* allele 3' UTR. It has been reported that transcripts from the same gene which have shorter 3' UTRs are expressed in higher amounts since transcripts with longer 3' UTRs are less stable and degraded more rapidly (Canete-Soler *et al*, 1998; Borrmann *et al*, 2001). This may explain why the A^1 and A^2 alleles are not expressed in PB MNC samples but the *B* allele is. A small population of cells may express *ABO* in the PB MNCs but the A^1 and A^2 transcripts would be degraded rapidly while the *B* transcript accumulated and was hence detected by RT-PCR.

However, O^1 allelic expression is seen in A^1O^1 PB MNCs and PBSCs that do not express the A^1 or A^2 allele. This may be explained by the recent work on the promoter of *ABO*, which indicates that various alleles have different enhancer structures. The enhancer associated with the *B* and O^1 alleles is 300 fold more active than the A^1 allele enhancer (Yu *et al*, 2000). This may explain why the O^1 allele is expressed when the A^1 allele is not. However, this does not explain why the A^2 allele is not expressed when it has the same enhancer as the O^1 allele (Irshaid *et al*, 1999). The sequence of the A^2 allele enhancer has not been elucidated and it may have nucleotide substitutions that render this enhancer less efficient than the *B* and O^1 enhancers, like the one repeat unit of the *B* allele which increases transcription 20% more than one repeat unit of the A^1 allele when they only differ by one nucleotide substitution (Yu *et al*, 2000). Although the A^2 allele has four repeats and is expected to have an enhancer as efficient as the *B* allele, nucleotide differences in the A^2 allele enhancer may alter the enhancer's ability to increase transcription (Irshaid *et al*, 1999; Yu *et al*, 2000). Other regulatory regions may also contribute to the expression of the various *ABO* alleles.

Investigation of PBSCs as an equivalent for BM samples showed that they could not replace BM as a normal control. This was shown by the fact that *AO* PBSC samples did not express the *A* allele as seen for *AO* PB MNCs. The *ABO* RT-PCR results suggested that PBSC samples resemble PB MNC more than BM samples. However, identification of the various *FUT1* alternative transcripts indicated that PB MNC samples did not express *FUT1*. This contrasts with PBSC samples, which expressed exon 7 transcripts while BM samples expressed exon 2 and exon 7 transcripts. These results on their own indicate that these three cell types have their own unique mRNA expression of *FUT1* consistent with them being distinct populations of cells.

Analysis of samples for *ABO* expression at the mRNA level indicated that this gene undergoes extensive alternative splicing. Previous reports of *ABO* alternative splicing include a variant not found in this study, an alternative transcript that retains 220 bp of intron 2 (O'Keefe PhD thesis, 1995). Hakomori (1999) reported alternatively spliced transcripts that had exon 6, or exons 2, 3 and 6, or exons 2-6 spliced out. In this study novel alternative transcripts were identified that have not been previously reported, such as exon 4 spliced out, exon 4 and 5 spliced out, and a splice variant with a portion of intron 5 retained. The alternative splicing was not unique to a certain type of sample although further investigation would be required. Since our study, the alternative splicing out of exons 4–6 has been reported yet the intron 5 retained splice variant has not been reported elsewhere (Kominato *et al*, 1999). The extensive alternative splicing may be due to improper RNA splicing and may not yield functional proteins, though this would require further investigation.

Analysis of most patient samples with alterations of ABH antigens by flow cytometric analysis did not show a correlation with changes in *ABO* mRNA expression. This was not surprising as the majority of patients had ABH antigen alterations on a minor population of RBCs, however the majority of the RBC population was normal. This normal population would express *ABO* mRNA, masking any changes in *ABO* mRNA expression in the cell population with loss of ABH antigens. Flow cytometry is performed on RBCs which live for 120 days whereas allelic expression of *ABO* is investigated in blasts, hence changes at the mRNA level may not have been translated as changes on the RBCs which have not yet entered the circulation. If a RBC sample of these patients could be analysed 3-4 months down the track, then a corresponding alteration may be seen.

There were five blood group A and B patients with loss of H and two of these, F39 and F53, had loss of *A* allele expression (Table 4.7; Figure 4.12C). The corresponding loss of A antigen was not seen on the flow because the loss of H may mask the loss of A since there is no H to begin with, and no A antigen could be made regardless of whether the A transferase was made or not. One of these patients, F39, did not express *FUT1* which supports the finding that the loss of H antigen seen by flow cytometry may mask the loss of A allele expression seen by RT-PCR.

There may also be other alterations responsible for loss of ABH antigens besides changes in mRNA expression. These may include mutations, such that the mRNA is still made but a mutation may mean that the transferase made is inactive or less efficient. This is supported by the extensive work that demonstrates that various weak *ABO* alleles are generated by a single nucleotide substitution (see section 1.4.5.4). For example, the A^3 phenotype which has minimal A antigen (Figure 3.1) results from a single nucleotide substitution such that the A transferase made is much weaker than the A^1 transferase (Table 1.4).

Analysis of *FUT1* alternative transcripts did not correlate with loss of H but this is not surprising as there was never complete loss of H. This suggests that perhaps only one allele of the *FUT1* gene may have been lost and not both. Mono-allelic or bi-allelic expression of *FUT1* could be determined by using a restriction fragment length polymorphism (RFLP) such that allelic expression could be determined by restriction enzyme digestion of PCR products, much like *ABO* allelic expression is determined.

Normal BM samples expressed exon 2 and exon 7 *FUT1* transcripts while the majority of patients did not express exon 2 *FUT1* transcripts. This would indicate that the majority of patient samples have cells at a well differentiated state whereas normal BM have cells at different stages of differentiation, though the majority are at a well differentiated state. This result is surprising since the majority of patients have AML, which is characterised by an accumulation of immature myeloid cells.

We found that K562 cells express exon 1 and exon 2 containing *FUT1* transcripts although it has been reported that K562 cells use the exon 1 transcription start site exclusively (Koda *et al*, 1998). Koda and colleagues (1998) suggested that the transcription start site changes during differentiation of the erythroid lineage and that K562 is a multipotential haematopoietic cell line that can differentiate into progenitors of erythrocytes, monocytes and granulocytes. The only other samples expressing exon 1 transcripts were the cell lines HEL, KCL22 and one MDS patient which expressed all three *FUT1* alternative transcripts.

HEL is thought to be a cell line at an early stage of differentiation of the erythroid lineage and has been reported to predominantly use the exon 2 transcription start site (Koda *et al*, 1998). In this study we have found that HEL expresses all the alternative transcripts of *FUT1* indicating that there is heterogeneity among the cell population in terms of the stage of differentiation. Koda *et al* (1998) found that bone marrow cells predominantly use the exon 7 transcription start site though in this study they express exon 2 transcripts as well. Koda *et al* (1998) used Northern analysis but our RT-PCR technique may have been more sensitive at detecting the various alternative transcripts of *FUT1* which would explain the differences in results.

The expression of ABH antigens occurs in a stage-specific and tissue specific manner (Szulman, 1980). It has been suggested that the use of multiple promoters in a single gene is a way for controlling expression in a stage-specific and tissue-specific manner (Kozak, 1991). The use of the alternative promoters in the *FUT1* gene has been suggested to be a way of controlling the ABH antigen expression in various tissues and stages of differentiation of the erythroid lineage (Koda *et al*, 1998). If this is so, then the leukaemic patients can be divided into groups depending on their expression. Of the 44 patient analysed only one (2%) expressed the exon 1 transcript and 11 (25%) expressed the exon 2 transcript. There were 40 patients (91%) that expressed the exon 7 transcript, hence, it seems that expression of the exon 2 transcript may differentiate the patients into 2 groups. Normal BM expresses exon 2 and exon 7 transcripts, hence, the group of patients not expressing exon 2 transcripts may have cells at a higher stage of differentiation.

In the patients with loss of H, 6/8 of those tested did not express the *FUT1* exon 2 containing transcripts (Table 4.8). Since the exon 2 containing transcripts are expressed in normal BM, lack of this transcript may be responsible for the loss of H seen on the flow cytometer. To determine if there is an association between loss of the H antigen and lack of the *FUT1* exon 2 containing transcript would require further studies. These would include characterising the various proteins encoded from the different *FUT1* transcripts.

Studies have shown that promoters within Alu sequences, such as the exon 2 promoter, may be suppressed by methylation (Yoder *et al*, 1997) and the focus of the following chapter is to identify whether *ABO* and *FUT1* expression is regulated by methylation.

CHAPTER 5

ABO & FUT1 Epigenetic Analysis

5.1 Introduction

Loss of ABH antigens is a common occurrence in haematological malignancies (Chapter 3). In some of these patients, there is also a corresponding loss of *ABO* and *FUT1* mRNA expression (Chapter 4). The frequency of loss of *ABO* and *FUT1* mRNA expression may have been underestimated because the normal red cell precursors may mask the loss of expression. In this chapter, methylation is investigated as an epigenetic mechanism for silencing *ABO* and *FUT1* in haematological malignancies. Methylation has been the subject of much investigation as it is a way to silence genes without changing the DNA sequence, making methylation a potential therapeutic target.

Methylation predominantly occurs at CpG dinucleotides, that is, a cytosine adjacent to a guanosine on the same DNA strand (Bird, 1986). CpG dinucleotides are rare and do not occur at the expected frequency due to their methylation, as 5-methylcytosine can deaminate to thymine (Bird, 1986). CpG islands are clusters of unmethylated CpGs showing the expected CpG frequency due to the absence of methylation (Bird, 1986).

Many studies show an inverse relationship between CpG island methylation and gene expression. Hypomethylated (lack of CpG methylation) DNA is associated with active chromatin whereas methylated DNA is associated with inactive chromatin (Razin, 1998). Examples of the relationship between CpG methylation and transcriptional inactivation in normal human cells are X-inactivation and imprinting (see 1.5.2.1).

Critical changes in methylation during tumorigenesis occur in some of the normally unmethylated CpG islands that encompass the promoter regions of many genes such as *RB*, *p16, calcitonin* and *BRCA1* (Baylin *et al*, 1987; Greger *et al*, 1989; Gonzalez-Zulueta *et al*, 1995; Herman *et al*, 1995; Merlo *et al*, 1995; Dobrovic and Simpfendorfer, 1997; Stirzaker *et al*, 1997).

Analysis of the *ABO* gene promoter region in several human carcinoma cell lines showed that DNA methylation of the *ABO* gene promoter region was inversely correlated with its gene expression (Kominato *et al*, 1999). Treatment of a non-expressing gastric cancer cell line, MKN 28, with the demethylating agent, 5-aza-2'-deoxycytidine, resulted in re-expression of the *ABO* gene.

Iwamoto *et al* (1999) analysed the methylation patterns of *ABO* in A antigen positive and negative clones derived from the colon adenocarcinoma cell line SW480. The antigen positive clones were unmethylated and the antigen negative clones were methylated. These results suggest that lack of A antigen expression on these cells was associated with methylation.

The aim of this chapter was to investigate whether methylation may be responsible for loss of *ABO* and *FUT1* expression in haematological malignancies. Another aim was to develop a new screening technique to sensitively assess methylation without the requirement for large amounts of high molecular weight DNA.

5.2 Results

Assessment of the *ABO* and *FUT1* genes for promoter methylation commenced with Southern analysis, which had previously proven to be successful for analysing the methylation status of the *BRCA1* gene in breast, ovarian, colon cancers and leukaemia (Bianco *et al*, 2000).

5.2.1 Methylation Status of the *ABO* Promoter by Southern Analysis

The *ABO* promoter region CpG island has been defined by Kominato *et al* (1999). This *ABO* sequence (Genbank sequence AC000397) was analysed using a CpG island prediction program called GrailEXP v3.3 (http://compbio.ornl.gov/grailexp). This program predicted 2 CpG islands, both fulfilling the criteria for a CpG island, that is, G + C content greater than

50% and CpG/GpC ratio of at least 0.6 (Gardiner-Garden and Frommer, 1987). The first predicted CpG island was 1450 bp, commencing 1000 bp before exon 1 and finishing 350 bp into intron 1, with a G + C content of 73% and a CpG/GpC ratio of 0.82. The second predicted CpG island was 380 bp in length, commencing 130 bp before exon 7 and finishing 250 bp into exon 7, with a G + C content of 67% and a CpG/GpC ratio of 0.73.

The methylation status of the CpG island preceding exon 1 was investigated as methylation of promoter CpG islands has been associated with the silencing of many genes. For Southern hybridisation the primers ABO CpG F and ABO CpG R (Table 2.1.1) were used to generate a 419 bp probe, which was α-P^{32} labelled and used to hybridise DNA samples from individuals whose DNA was digested with the bordering enzyme *Taq*I, and the methylation sensitive enzymes *Cfo*I or *Hpa*II (Figure 5.1).

For methylation analysis using Southern blotting, an enzyme that is not methylation sensitive and does not cut within the target (CpG island) sequence is used to cut the DNA. For *ABO, Taq*I was chosen as the bordering enzyme. A bordering enzyme allows for a better interpretation of the bands expected to be seen if the sample is unmethylated and allows determination of the origin of any bands that are aberrant, when compared to the unmethylated pattern. A bordering enzyme is required to define the upper size limit of restriction fragments contained in the target sequence so that adequate resolution on an agarose gel can be achieved.

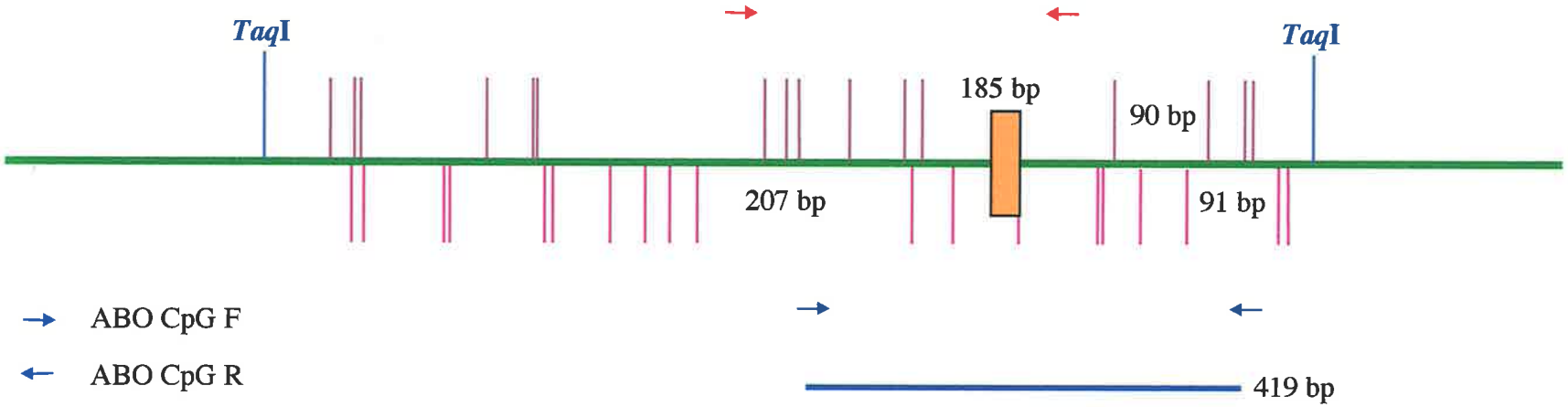


Figure 5.1. The *ABO* CpG island.

The diagram shows the position of the 419 bp probe used for Southern blot analysis and the regions that will hybridise to unmethylated samples digested with *Taq*I and *Cfo*I, 185 and 90 bp and the band that the probe will hybridise to when unmethylated samples are digested with *Taq*I and *Hpa*II, 207 and 91 bp. The primers used for MS-SSCA analysis are also shown.

Once the DNA sample to be analysed was restricted with *Taq*I, it was digested with the methylation sensitive enzymes, *Cfo*I (recognition sequence GCGC) and *Hpa*II (CCGG). These two enzymes contain a CpG site within their recognition sequence but they do not cut if the cytosine is methylated. *Taq*I also has a CpG in its recognition sequence, TCGA, but this enzyme is not methylation sensitive, that is, *Taq*I will digest DNA independent of the methylation status of the CpG in its recognition site.

If there is no methylation within the recognition sequences of the methylation sensitive enzymes, then the *Taq*I fragment homologous to the probe sequence will be cut with *Cfo*I or *Hpa*II to give the expected band sizes (5.2.1). If there is hypermethylation, extensive CpG methylation, then neither enzyme would be able to cut the DNA, producing the same bands that would be expected for a control sample digested with *Taq*I only. Varying amounts of methylation will produce patterns differing from the expected normal unmethylated pattern.

The expected fragment sizes when assessing *ABO* for methylation by Southern analysis included a 1012 bp band from the *Taq*I digest to which the probe shares 100% homology. Other fragments expected from unmethylated samples include 185 bp and 90 bp bands from the *Taq*I and *Cfo*I digest, to which the probe shares complete homology (Figure 5.1). All the other fragments were too small to be detected, however methylation of any of the *Cfo*I sites spanning the probe would create larger fragments that could easily be detected. From the *Taq*I and *Hpa*II digest a 207 and 91 bp band would be detected if the sample was unmethylated (Figure 5.1). The probe had 109 bp of homology with the 207 bp band and 48 bp of homology with the 91 bp band.

There were extreme difficulties with the hybridisation due to the GC richness of the probe (80.7%), which caused the probe to hybridise non-specifically. Previous Southern analysis

with a *BRCA1* probe had not caused these problems but this probe had a GC content of this probe was only 59.3% (Bianco *et al*, 2000).

The hybridisation difficulties were overcome, to some extent, by increasing the hybridisation temperature to 51°C. Figure 5.2 shows a Southern blot hybridised with the *ABO* probe with all of the expected bands present. There was an extra band in one sample (Figure 5.2; lane 1 arrow), which most likely resulted from the first *Cfo*I site after exon 1 being methylated. This would create a band of 185 + 90 bp, that is, 275 bp. The blot also demonstrates extreme background hybridisation, a result of numerous hybridisations at various increasing temperatures to reduce non-specific binding. The use of a digoxigenin labelled probe was tried on the Southern blots but no hybridisation could be achieved due to the condition of the blots, which had been damaged by repeated hybridisations with the α-P^{32} probe. Since 20 µg of DNA is required for a Southern blot it was not feasible to repeat them, as this quantity of DNA was not available for all patients. Instead, a new technique to investigate for methylation changes was developed.

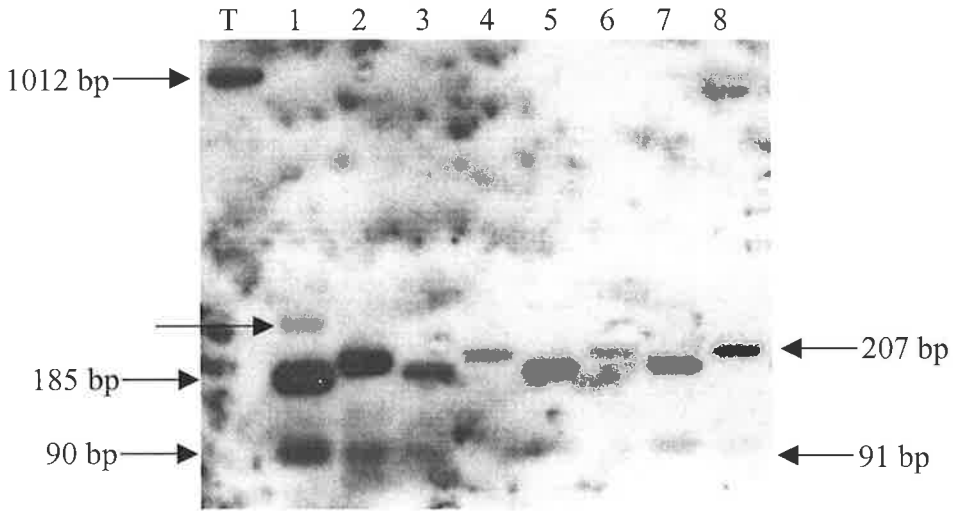


Figure 5.2. An *ABO* Southern blot.

T refers to the *Taq*I alone digest and clearly shows the 1012 bp band. Lanes 1, 3, 5 and 7 are patient DNA samples digested with *Taq*I and *Cfo*I while lanes 2, 4, 6 and 8 are digests with *Taq*I and *Hpa*II. An arrow points to a methylated band in lane 1.

5.2.2 MS-SSCA: the Development of a New Methylation Analysis Technique

Using samples and a gene that we had already shown the methylation status of by Southern hybridisation (Bianco *et al*, 2000), a new technique for methylation detection, called methylation sensitive – single stranded conformational analysis (MS-SSCA), was developed. This technique required that 1 μg of genomic DNA was bisulphite treated (see 2.6.1) so that all unmethylated cytosines are converted into thymines but methylated cytosines are protected and remain cytosines. Primers were then designed within the CpG island and spanning exon 1. They were designed to amplify bisulphite modified DNA only by placing thymines resulting from modified cytosines on the 3' end of the primers. CpG sites were not placed in the primers, but if they were unavoidable, they were placed as close to the 5' end of the primer as possible. The primers amplified modified methylated and modified unmethylated DNA, with the PCR product 250 – 300 bp in length since it has been reported that SSCA is most sensitive when used to detect one sequence variation in PCR amplified DNA molecules of 200 bp or less (Sheffield *et al*, 1993). However, larger fragments can be used when detecting every sequence variation is not critical, that is, for methylation one would expect more than one CpG to be methylated. This would mean more than one base difference, therefore, a greater difference in sequences between the methylated sample and unmethylated control, which is seen as a band shift from the normal unmethylated pattern.

The PCR products were then denatured with sodium hydroxide and run on a Mutation Detection Enhancement (MDE) gel under optimal conditions for detecting nucleotide shifts when compared to normal unmethylated samples. This new technique, MS-SSCA, was validated by using the gene *BRCA1* and breast cancer samples which had previously been shown to be methylated for this gene by Southern blot analysis (Bianco *et al*, 1999).

5.2.2.1 MS-SSCA of the *ABO* Promoter

For MS-SSCA of the *ABO* promoter the primers used were ABO BIS F and ABO BIS R1 (Figure 5.1), which amplified a PCR product from bisulphite modified DNA of 269 bp. The *ABO* MS-SSCA product was analysed for methylation by electrophoresis through a 0.5X MDE gel or a 0.75X MDE gel (Figure 5.3). Some band shifts were better resolved on a 0.5X MDE gel while others were more obvious and clearer on a 0.75X MDE gel. All of the normal unmethylated samples and K562 showed the same pattern on the 0.5X and the 0.75X MDE gel. For the methylated cell lines, the 0.75X MDE gel tended to give better resolution of the bands. A 0.75X MDE gel with 5% glycerol was also tried but this showed less resolution than the other gel running conditions.

Restriction enzyme digestion of the *ABO* MS-SSCA PCR product could also be used to detect methylation. As depicted in Figure 5.4, various restriction enzyme sites are created or destroyed due to methylation. An unmethylated CpG will be converted to TpG, while a methylated CpG will remain CpG. For example, the *Taq*I recognition sequence is TCGA and methylation of this CpG will prevent this site from being bisulphite modified. Unmethylated sequences will not cut with *Taq*I as the recognition site will be abolished. The restriction enzymes used for methylation analysis are *Taq*I, *Hinf*I (if the final C in the recognition sequence is part of a CpG) and *BstU*I, which will only digest the MS-SSCA product if the samples are methylated (Figure 5.4). Examples of these digests are shown in Figure 5.5.

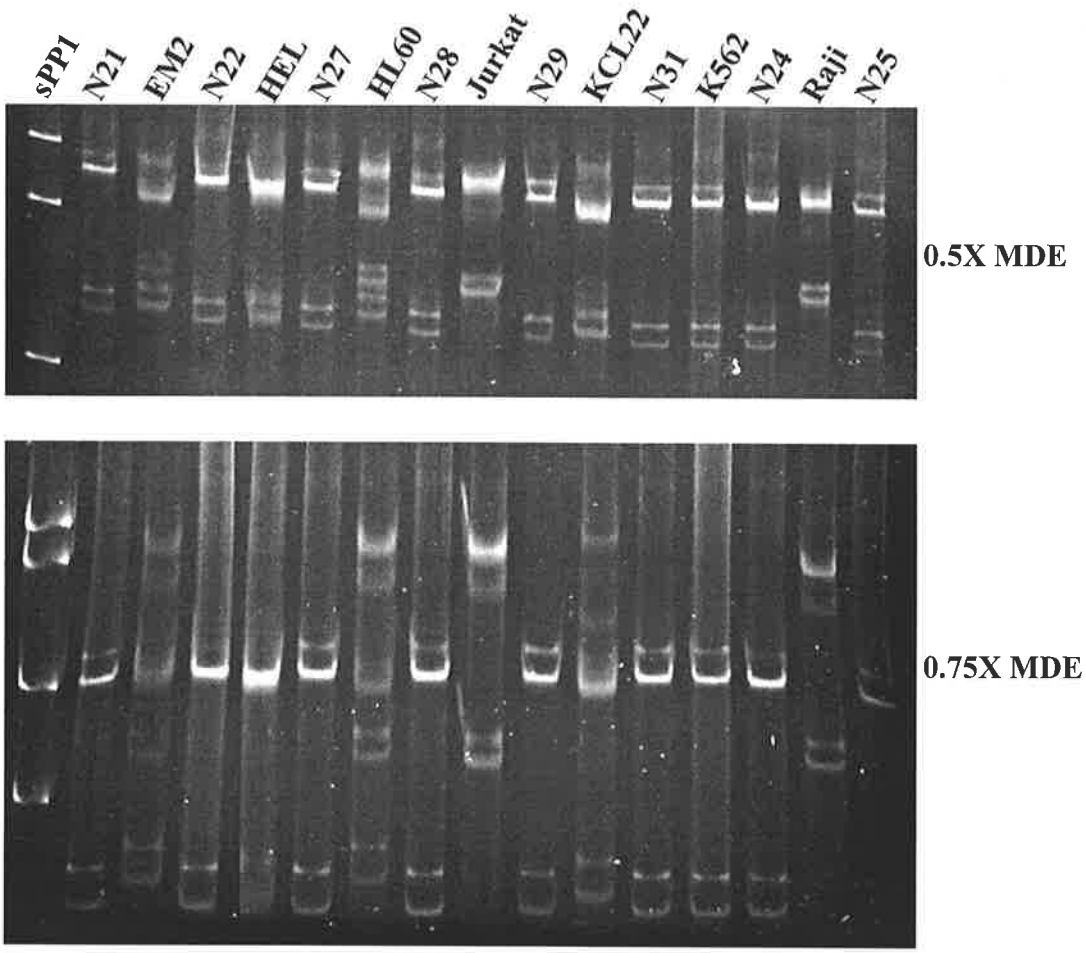


Figure 5.3. *ABO* MS-SSCA analysis on different MDE conditions. The same samples were run on two different gel conditions to determine which was optimal. N21, N22, N24, N25, N27, N28, N29 and N31 are *ABO* MS-SSCA products derived from the DNA of normal peripheral blood mononuclear cells and are unmethylated controls for *ABO*. The remaining samples are from leukaemic cell lines.

Unmethylated and bisulfite modifed *ABO* MS-SSCA sequence.

1 ggaTagggTTTTaaggtaTTagggTTa**Tg**agggg**TgTgTg**ggtTTTttgg 50

51 ggatg**TgTgTg**aggagg**Tg**T**Tg**tTTTttTTtagTaggggtTTTtggggaT 100

101 T**TgTg**gT**Tg**TTtTT**TgTg**TTTTtTtgtTTTTtTT**Tg**tgtt**Tg**gTTt**Tg**gg 150

151 aagt**Tg**ggg**Tg**g**Tg**gg**Tg**g**TgTg**ggT**Tg**ggaggggg**Tg**TTt**Tg**ggTtTaT 200

201 TT**Tg**TTTTagggT**Tg**T**Tg**gg**Tg**gaagg**Tg**gaggT**Tg**agaTTaga**TgTg**ga 250

251 gTTatggT**Tg**aggtgttgT 269

Methylated and bisulfite modifed *ABO* MS-SSCA sequence.

1 ggaTagggTTTTaaggtaTTagggTTa**cg**agggg**cgcgcg**ggtTTTttgg 50

51 ggatg**cgcgcg**aggagg**cg**T**cg**tTTTttTTtagTaggggtTTTtggg**gaT** 100

101 **Tcgcg**gT**cg**TTtTT**cgcg**TTTTtTtgtTTTTtTT**cg**tgtt**cg**gTTt**cg**gg 150

151 aagt**cg**ggg**cg**g**cg**gg**cg**g**cgcg**ggT**cg**ggaggggg**cg**TTt**cg**ggTt**TaT** 200

201 **TTcgTTTTag**ggT**cg**T**cg**gg**cg**gaagg**cg**gagg**Tcga**gaTTaga**cgcg**ga 250

251 gTTatgg**Tcga**ggtgttg**c** 269

Figure 5.4. Unmethylated and methylated *ABO* MS-SSCA products.

The capital Ts identify thymines that are a result of bisulfite modification of cytosine and the unmethylated CpGs (TpG) and methylated CpGs are underlined. The different restriction enzymes used for assessing methylation by digestion are as follows: eight *BstUI* sites, two *TaqI* sites (however one is found in the primer and hence will cut regardless of methylation status), one *HinfI* site. Yellow highlight denotes Sp1 sites (Hata et al, 2002).

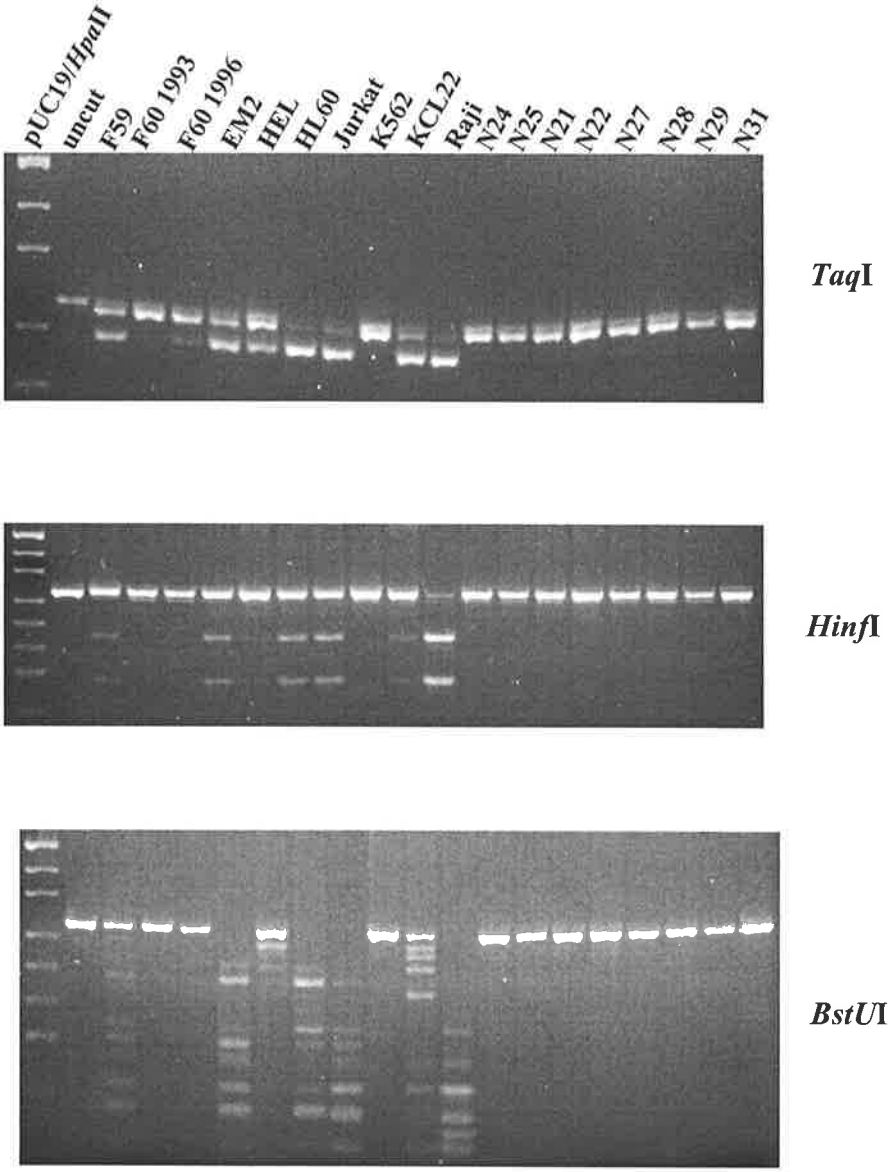


Figure 5.5. Restriction enzyme digests of the *ABO* MS-SSCA PCR product. Digestion with any of the restriction enzymes is indicative of methylation at that CpG site within the restriction enzyme recognition sequence. All the products will cut with *Taq*I since there is a *Taq*I site in the reverse primer.

A summary of the restriction enzyme digestion results and the SSCA results of the *ABO* MS-SSCA PCR product is shown in Table 5.1. For the cell lines and the normal controls, the digests and the two MDE gel conditions, gave comparable results. Analysis of the methylation status for the AML and CML patients analysed by flow cytometry yielded 3/28 patients with methylation. None of these patients, F15, F21 and F59, showed alterations in *ABO* by flow cytometric analysis or by RT-PCR. An interesting observation is that two of the patients that had *ABO* methylation were in CML blast crisis in which *ABL*, a gene linked to *ABO*, undergoes methylation (Asimakopoulos *et al*, 1999). Also, in this case, the blasts would be unlikely to show any differentiation to red cells as in blast crisis, the blasts show little maturation, that is, there will be no normal or abnormal RBCs.

In another group of patients 6/11 showed methylation, these were OT2, OT3, OT16, OT19, OT26 and OT29. OT2 and OT3 had no alterations in *ABO* by flow cytometric analysis or RT-PCR but OT16 did have a small population of cells with loss of A detected by flow cytometry. The remaining 3 patients had no alterations in *ABO* mRNA expression.

The group of patients analysed previously that had ABH antigen changes detected by serology (O'Keefe PhD thesis, 1995) showed the greatest amount of methylation, with 5/7 methylated. This was not surprising as changes detected by serology require a larger proportion of the cells to have alterations in *ABO* antigens. The patients analysed by flow cytometry often had very small populations of cells with changes in ABH antigen expression and methylation patterns in this abnormal population would be hard to detect against the large background of normal cells.

ID	*Taq* I	*Hinf* I	*BstU* I	0.5X MDE	0.75X MDE
F4	U	U	U	ND	U
F9	U	U	U	U	U
F10	U	U	U	U	U
F11	U	U	U	ND	U
F12	U	U	U	U	U
F15	M	M	M	M	M
F17	U	U	U	ND	U
F19	U	U	U	U	U
F20	U	U	U	U	U
F21	M	U	M	ND	M
F23	U	U	U	ND	U
F24	U	U	U	U	U
F25	U	U	U	U	U
F27	U	U	U	ND	U
F30	U	U	U	U	U
F37	ND	ND	ND	U	ND
F39	U	U	U	U	U
F43	U	U	U	U	U
F44	U	U	U	U	U
F45	U	U	U	U	U
F48	U	U	U	ND	U
F51	U	U	U	U	U
F52	U	U	U	ND	U
F53	U	U	U	U	U
F57	U	U	U	U	U
F58	U	U	U	U	U
F59	M	M	M	M	M
F60	U	U	U	U	U

OT2	M	M	M	M	M
OT3	M	U	U	ND	M
OT16	M	U	M	U	U
OT19	M	M	M	ND	M
OT21	ND	ND	ND	U	U
OT22	U	U	U	U	U
OT24	U	U	U	U	U
OT25	U	U	U	U	U
OT26	M	U	U	U	M
OT27	U	U	U	U	U
OT29	M	U	M	M	M

DOK1	U	U	U	U	U
DOK2	M	U	U	M	M
DOK3	U	U	U	U	ND
DOK4	M	U	U	U	U
DOK5	M	U	M	M	M
DOK6	M	M	M	U	M
DOK7	M	U	M	M	M

ID	*Taq* I	*Hinf* I	*BstU* I	0.5X MDE	0.75X MDE
N21	U	U	U	U	U
N22	U	U	U	U	U
N24	U	U	U	U	U
N25	U	U	U	U	U
N27	U	U	U	U	U
N28	U	U	U	U	U
N29	U	U	U	U	U
N31	U	U	U	U	U
K562	U	U	U	U	U
KCL22	M	M	M	M	M
HEL	M	U	M	M	M
HL60	M	M	M	M	M
EM2	M	M	M	M	M
RAJI	M	M	M	M	M
Jurkat	M	M	M	M	M

Table 5.1. ABO MS-SSCA results.
'M' refers to methylation found either by restriction enzyme digestion or by band shifts on a SSCA. 'U' refers to an unmethylated sample determined by lack of restriction enzyme digestion or by no shifts on a SSCA when compared to an unmethylated control. 'ND' is not done.

Another interesting observation in Table 5.1 is that two patients, OT26 and DOK6, appeared unmethylated on a 0.5X MDE gel but were found to be methylated on a 0.75X MDE gel. This is not surprising in light of the report by Liu *et al*, (1999) where they describe the detection of virtually all mutations (DOVAM) by SSCP. They tried seven different running conditions, in which the polyacrylamide, additives, running buffers and running temperatures were altered and found that each changed condition alone was unable to detect mutations 100% of the time. However, using five different running conditions enabled 100% detection of mutations. Buzin *et al* (2000) used the DOVAM approach to analyse mutations in three genes using five different conditions and found that for each gene, a mutation would have been missed if only four conditions had been used.

Restriction enzyme digestion was used to circumvent this problem to some extent, however, there were 34 CpGs in the *ABO* MS-SSCA PCR product (excluding one in the primer) and restriction enzyme digestion with *Taq*I, *Hinf*I and *BstU*I allowed analysis of only 16 of these CpGs. SSCA is able to analyse all these CpGs however, numerous running conditions are required to fully rule out lack of methylation.

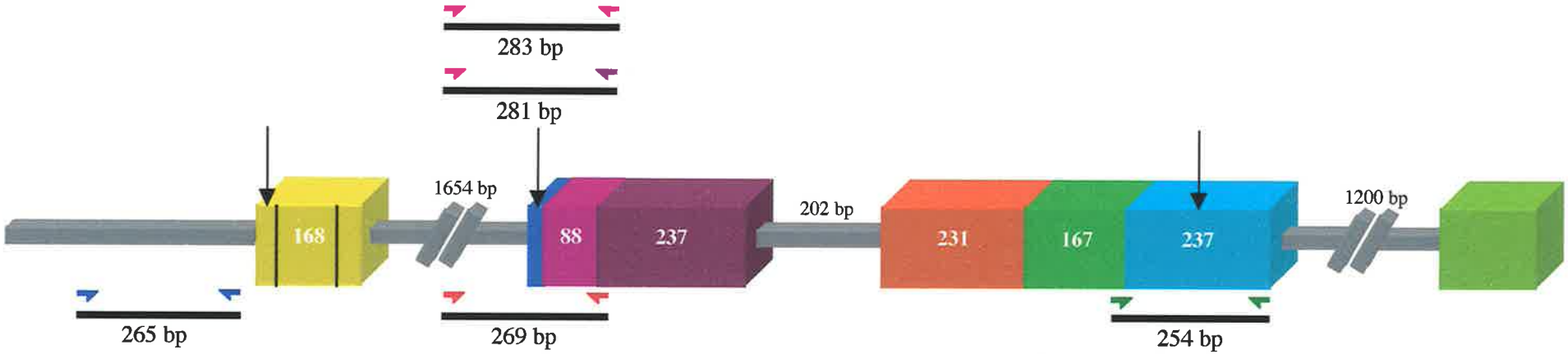
5.2.2.2 *FUT1* MS-SSCA

Methylation of the *FUT1* gene was also investigated by MS-SSCA as alterations in *FUT1* mRNA expression and alterations in the H antigen were observed in a large proportion of patients (Chapter 3 and 4). Analysis of the Genbank sequence AC009002, which contains the *FUT1* sequence, with EMBOSS (http://www.ebi.ac.uk) revealed a CpG island approximately 200 bp before the start of exon 1. There was also a very large CpG island spanning the transcription start sites of exons 2 and 7. The CpG island preceding exon 2 is formed partly by an Alu sequence which are known to be CpG rich and tend to be methylated (Liu *et al*, 1994).

Primers were designed for the CpG islands preceding all three transcription start sites in *FUT1* (Figure 5.6). The bisulphite modified primers designed to analyse the CpG island preceding exon 1 generated a product of 265 bp (Figure 5.6) and contained a *Mae*II site and an *Aci*I site which could be used to screen for methylation by restriction digestion of the MS-SSCA product (Figure 5.7).

Preliminary analysis of 5 normal PB MNC samples (N1, N4, N7, N9, N10) and 5 patient samples (F9, F27, F30, F51 and F60) of the CpG island preceding *FUT1* exon 1 by 0.5X MDE gel analysis found no altered patterns, suggesting that these samples were all unmethylated (Table 5.2). However, analysis of 8 normal samples (PB MNC and PBSC) and 7 leukaemic cell lines by restriction enzyme digestion with *Aci*I and *Mae*II indicated that all of the samples, except for the cell line Jurkat, were digested with these two enzymes (Figure 5.8). Hence, all of the samples were methylated for at least 2 of the CpGs assessed by the restriction enzymes. Analysis of these samples on a 0.75X MDE gel indicated that all the samples had the same pattern, since they were all methylated, except for the Jurkat sample which had a subtle shift upward indicating that it was less methylated (see white star in Figure 5.8).

It was not surprising that the normal samples were methylated as none of these samples expressed the exon 1 transcript of *FUT1* and it is likely that methylation may be the mechanism for this silencing or for maintaining the inactivation (Chapter 4). The methylation of the leukaemic cell lines was surprising since some of these actually expressed the *FUT1* exon 1 transcript, while the unmethylated sample Jurkat did not. The MS-SSCA band shifts were very subtle and the MDE gel running conditions needed optimisation but since none of the patients expressed the *FUT1* exon 1 transcript (except F17, Table 4.8) it was concluded that analysis of this CpG island was not essential.



	FUT1 MS-SSCA ex 1 F
	FUT1 MS-SSCA ex 1 R
	FUT1 MS-SSCA ex 2 F
	FUT1 MS-SSCA ex 2 R
	FUT1 MS-SSCA ex 7 F
	FUT1 MS-SSCA ex 7 R
	FUT1 MS-SSCA ex2 anti F
	FUT1 MS-SSCA ex2 anti R
	FUT1 MS-SSCA ex2 anti R2

	Exon 1
	Exon 2
	Exon 3
	Exon 4
	Exon 5
	Exon 6
	Exon 7
	Exon 8

↓	Transcription start sites

Figure 5.6. The primers used for *FUT1* methylation analysis.

The Genbank sequence AC009002 was used for determining the location of the primers used for MS-SSCA of *FUT1*. Numbers inside the exons are the exon sizes. Exon 2 is 25 bp. Not all of exon 8 is shown. Black solid bars represent the PCR product and its size.

Unmethylated and bisulphite modified *FUT1* exon 1 MS-SSCA sequence.

```
1     ggggtgTtgggTtTTtggaaaaaTTaTTtTttggaTTatTtatgTagatT     50

51    aTgTagaaTaagagaaatttTtgTgTTTTatTtgaatttTtaagtttggg     100

101   gggagggTgtgatTtgaTaTtgaggttTTttgatTTtTagTaaggTggTa     150

151   attgTtgtatgaaagaagTgaTTgTatTtgagaTaTaagtatTTtgTTtt     200

201   ggaagTTtTtTaTTtggTTgtgggTTaaTTtTaaTTtTatTtgtTTTtgT     250

251   tTagatgTtTagaTT  265
```

Methylated and bisulphite modified *FUT1* exon 1 MS-SSCA sequence.

```
1     ggggtgTtgggTtTTtggaaaaaTTaTTtTttggaTTatTtatgTagatT     50

51    acgTagaaTaagagaaatttTtgcgTTTTatTtgaatttTtaagtttggg     100

101   gggagggcgtgatTtgaTaTtgaggttTTttgatTTtTagTaaggcggTa     150

151   attgTtgtatgaaagaagcgaTcgTatTtgagaTaTaagtatTTtgTTtt     200

201   ggaagTTtTtTaTTtggTcgtgggTTaaTTtTaaTTtTatTtgtTTTtgT     250

251   tTagatgTtTagaTT  265
```

Figure 5.7. Unmethylated and methylated *FUT1* MS-SSCA products.

The capital T define thymines that are a result of bisulphite modification of cytosine and the unmethylated CpGs (TpG) and methylated CpGs (red) are underlined. The different restriction enzymes used for assessing methylation by digestion are as follows: *Aci*I (GGCG) sites, *Mae*II (ACGT) sites, *BstU*I (CGCG) sites, *Taq*I (TCGA) sites, *Hinf*I (GANTCg) sites. Yellow shading denote Sp1 sites and double underline indicates AP-2 sites, both found using AliBaba2.1 program (http://wwwiti.cs.uni-magdeburg.de/~grabe/alibaba2/webbaba2.cgi).

Unmethylated and bisulphite modified *FUT1* exon 2 MS-SSCA sequence.

1 ggTtgTTtaaaagggaagggTTTTTtgtTTTatgatTa**Tg**tgaTttgTtt 50

51 gaTTttatTagtTatttgga**Tg**aTtTaTTTtTTttatTTtgTTTTTTTtt 100

101 gtTttgtataTaataaatatTag**TgTg**TTTagTTatt**Tg**gggTTaTtaT**T** 150

151 **g**gtTtTtg**Tg**tTttgatggtagtggtTTTT**Tg**ggTTTagTtgttttTtTt 200

201 ttatTtTtttgtTttgtgtTttttatttTttaTaatTtTtTTtTtTTtTaT 250

251 aggggaagaaTaTTTaTTT 269

Methylated and bisulphite modified *FUT1* exon 2 MS-SSCA sequence.

1 ggTtgTTtaaaagggaagggTTTTTtgtTTTatgatT**acgt**gaTttgTtt 50

51 gaTTttatTagtTatttgga**cg**aTtTaTTTtTTttatTTtgTTTTTTTtt 100

101 gtTttgtataTaataaatatTag**cgcg**TTTagTTatt**cg**gggTTaTtaT**c** 150

151 **g**gtTtTtg**cg**tTttgatggtagtggtTTTT**cg**ggTTTagTtgttttTtTt 200

201 ttatTtTtttgtTttgtgtTttttatttTttaTaatTtTtTTtTtTTtTaT 250

251 aggggaagaaTaTTTaTTT 269

Unmethylated and bisulphite modified *FUT1* exon 2 antisense MS-SSCA sequence.

```
1     TTtgtgaggagaggagagattgtaagaaataaagaTaTaagaTaaagaga     50

51    taaagagaaaaTagTtgggTTTggggaTTaTtaTTatTaagaTgTagag     100

101   aTTggtagtggTTTTgaatggTtgggTgTgTtgatatttattgtataTaa     150

151   gaTaagggggggTaggataaggagggtgagtTgtTTaaatgaTtgataag     200

201   gtTaagTaagtTaTgtgatTatgggaTagggggTTTttTTTttttaggTa     250

251   gTTgaagTagaaagagaaggTagTatatgtT   281
```

Methylated and bisulphite modified *FUT1* exon 2 antisense MS-SSCA sequence.

```
1     TTtgtgaggagaggagagattgtaagaaataaagaTaTaagaTaaagaga     50

51    taaagagaaaaTagTtgggTTcgggggaTTaTtaTTatTaagacgTagag     100

101   aTcggtagtggTTTcgaatggTtgggcgcgTtgatatttattgtataTaa     150

151   gaTaagggggggTaggataaggagggtgagtcgtTTaaatgaTtgataag     200

201   gtTaagTaagtTacgtgatTatgggaTagggggTTTttTTTttttaggTa     250

251   gTcgaagTagaaagagaaggTagTatatgtT   281
```

Unmethylated and bisulphite modified *FUT1* exon 7 MS-SSCA sequence.

1 gggTggggTtTTaTTttTTagTTatTTgTTTggTTTggaggggTggaTgT 50

51 tgTgagaTtTTTggTTgTgTTTTtTtTTttTTtTtTTtTTTTaagTTTtTg 100

101 TtgTTagtTTggaTaggTtgTgTggaggggagggTTtgTTgggTTggata 150

151 gTTggaTgTTtggTgttTTagggggTggTTggatgtggTTtgTTtttgTgg 200

201 agggtgTgTttTggTTaTgaaaagTggaTtgtggatTtgTTaTTtgTaag 250

251 TagT 254

Methylated and bisulphite modified *FUT1* exon 7 MS-SSCA sequence.

1 ggGcgggTtTTaTTttTTagTTatTcgTTcggTTcggaggGcggacgT 50

51 tgcgagaTtTTcggTcgcgTTTtTtTTttTTtTtTTtTTTTaagTTTtcg 100

101 TtgTTagtTTcggaTaggTtgcgcggaggggagggTTtgTcgggTcggata 150

151 gTcggacgTTtggcgttTTaggggcggTcggatgtggTTtgTTttgcgg 200

201 agggtgcgTtTcggTTacgaaaagcggaTtgtggatTtgTTaTTtgTaag 250

251 TagT 254

ID	*FUT1* exon 1 0.5X MDE	*FUT1* exon 1 0.75X MDE	*FUT1* exon 1 *Aci* I	*FUT1* exon 1 *Mae* II	*FUT1* exon 1 exp
F9	U	ND	ND	ND	ND
F27	U	ND	ND	ND	-
F30	U	ND	ND	ND	-
F51	U	ND	ND	ND	-
F60	U	ND	ND	ND	-
N1	U	ND	ND	ND	ND
N4	U	ND	ND	ND	ND
N7	U	ND	ND	ND	-
N9	U	ND	ND	ND	ND
N10	U	ND	ND	ND	-
N21	ND	M	M	M	ND
N22	ND	M	M	M	ND
N24	ND	M	M	M	-
N25	ND	M	M	M	-
N27	ND	M	M	M	-
N28	ND	M	M	M	-
N29	ND	M	M	M	-
N31	ND	M	M	M	-
K562	ND	M	M	M	+
KCL22	ND	M	M	M	+
HEL	ND	M	M	M	+
HL60	ND	M	M	M	ND
EM2	ND	M	M	M	-
RAJI	ND	M	M	M	-
Jurkat	ND	U	U	U	-

ID	*FUT1* exon 2 0.5X MDE	*FUT1* exon 2 *BstU* I	*FUT1* exon 2 *Mae* II	*FUT1* exon 2 anti 0.75X MDE	*FUT1* exon 2 anti *BstU* I	*FUT1* exon 2 anti *Hinf* I	*FUT1* exon 2 anti *Mae* II	FUT1 exon 2 anti *Taq* I	*FUT1* exon 2 exp
F4	ND	ND	ND	ND	M	U	M	M	-
F9	U	M	M	ND	ND	ND	ND	ND	ND
F11	ND	ND	ND	ND	M	U	M	M	-
F15	ND	ND	ND	ND	M	ND	ND	M	+
F17	ND	ND	ND	ND	M	ND	ND	M	+
F19	ND	ND	ND	ND	U	U	M	M	-
F21	ND	ND	ND	ND	M	U	M	M	-
F24	ND	ND	ND	ND	M	U	M	M	+
F27	U	M	M	ND	M	U	M	M	-
F30	U	M	M	ND	M	U	M	M	+
F43	ND	ND	ND	ND	M	U	M	M	-
F44	ND	ND	ND	ND	M	U	M	M	+
F48	ND	ND	ND	ND	U	U	M	M	+
F51	U	M	M	ND	ND	ND	ND	ND	-
F60	U	M	M	ND	ND	ND	ND	ND	-
OT2	ND	ND	ND	ND	M	U	M	M	-
OT3	ND	ND	ND	ND	M	U	M	M	-
OT21	ND	ND	ND	ND	M	U	M	M	-
OT22	ND	ND	ND	ND	M	U	M	M	+
OT27	ND	ND	ND	ND	M	U	M	M	-
DOK1	ND	ND	ND	ND	M	U	M	M	+
DOK5	ND	ND	ND	ND	M	U	M	M	-

ID	*FUT1* exon 2 0.5X MDE	*FUT1* exon 2 *BstU* I	*FUT1* exon 2 *Mae* II	*FUT1* exon 2 anti 0.75X MDE	*FUT1* exon 2 anti *BstU* I	*FUT1* exon 2 anti *Hinf* I	*FUT1* exon 2 anti *Mae* II	FUT1 exon 2 anti *Taq* I	*FUT1* exon 2 exp
N7	U	M	M	ND	ND	ND	ND	ND	-
N10	U	M	M	ND	ND	ND	ND	ND	-
N21	ND	ND	ND	M	M	M	M	M	ND
N22	ND	ND	ND	M	M	M	M	M	ND
N24	ND	ND	ND	M	M	M	M	M	-
N25	ND	ND	ND	M	M	M	M	M	-
N27	ND	ND	ND	M	M	M	M	M	-
N28	ND	ND	ND	M	M	M	M	M	-
N29	ND	ND	ND	M	M	M	M	M	-
N31	ND	ND	ND	M	M	M	M	M	-
K562	ND	ND	ND	U	U	U	U	U	+
KCL22	ND	ND	ND	M	M	M	M	M	-
HEL	ND	ND	ND	U	U	U	U	U	+
HL60	ND	ND	ND	M	M	M	M	M	ND
EM2	ND	ND	ND	M	M	M	M	M	-
RAJI	ND	ND	ND	M	M	M	M	M	-
Jurkat	ND	ND	ND	M	M	M	M	M	-

ID	*FUT1* exon 7 0.5X MDE	*FUT1* exon 7 0.75 XMDE	*FUT1* exon 7 *Aci* I	FUT1 exon 7 *BstU* I	*FUT1* exon 7 *Mae* II	*FUT1* exon 7 exp
F9	U	ND	ND	U	ND	ND
F27	U	ND	ND	U	ND	+
F30	U	ND	ND	U	ND	+
F51	U	ND	ND	U	ND	+
F60	U	ND	ND	U	ND	+
N1	U	ND	ND	U	ND	ND
N4	U	ND	ND	U	ND	ND
N7	U	ND	ND	U	ND	-
N9	U	ND	ND	U	ND	ND
N10	U	ND	ND	U	ND	-
N21	ND	U	U	U	U	ND
N22	ND	U	U	U	U	ND
N24	ND	U	U	U	U	-
N25	ND	U	U	U	U	-
N27	ND	U	U	U	U	+
N28	ND	U	U	U	U	+
N29	ND	U	U	U	U	+
N31	ND	U	U	U	U	+
K562	ND	M	U	M	M	ND
KCL22	ND	U	U	U	U	+
HEL	ND	U	U	U	U	+
HL60	ND	U	U	U	U	ND
EM2	ND	M	M	U	M	-
RAJI	ND	M	M	M	M	-
Jurkat	ND	M	U	M	M	+

Table 5.2. *FUT1* MS-SSCA results.

'U' is unmethylated, 'M' is methylated and 'ND' is not done. '+' refers to positive expression and '-' refers to negative expression. *FUT1* exon 1, 2 and 7 MS-SSCA restriction enzyme digestion and SSCA results are shown. 'M' for the restriction enzyme digestion refers to cutting (methylation at that restriction enzyme recognition sequence). 'M' for the SSCA results indicates band shifts when compared to the banding pattern of a known unmethylated 'U' control.

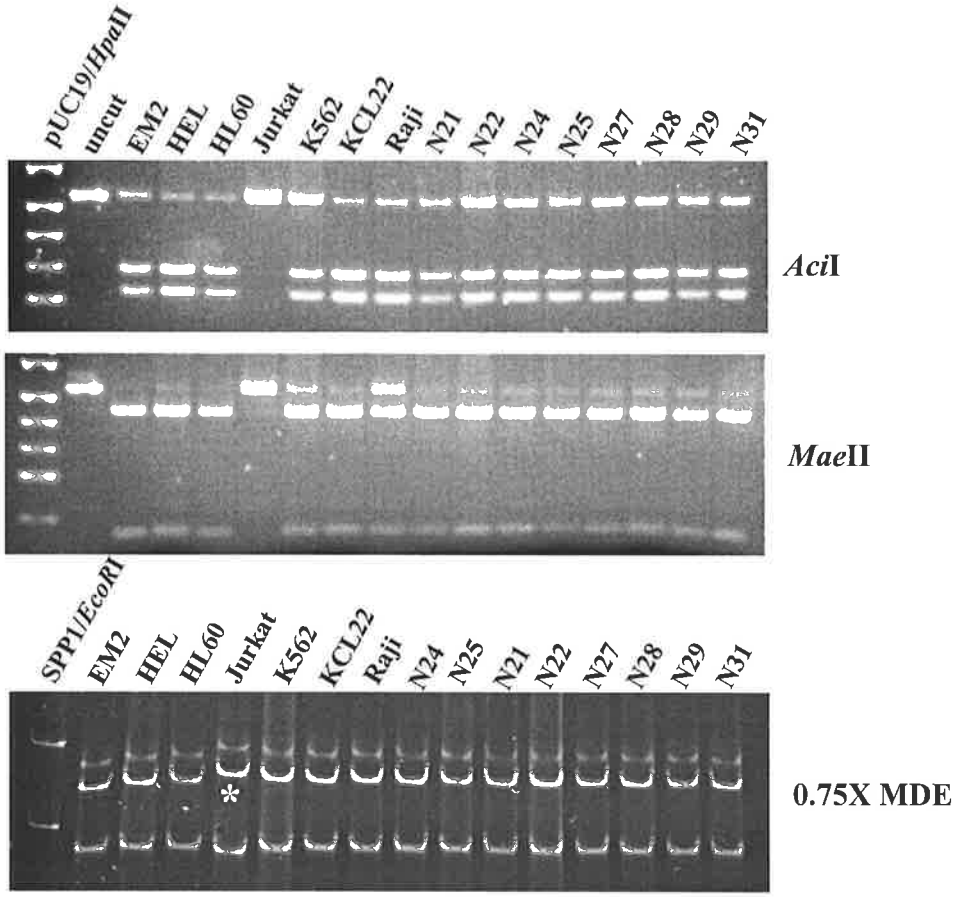


Figure 5.8. *FUT1* exon 1 MS-SSCA.
Digestion with *Aci*I and *Mae*II indicated that all normal samples and leukaemia cell lines were methylated, except for Jurkat, which had a band shift (marked with a white star) on a 0.75X MDE gel.

The primers for the CpG island preceding the exon 7 transcription start site were designed to generate a PCR product of 254 bp (Figure 5.6). The MS-SSCA product for the *FUT1* exon 7 CpG island contained 5 *Aci*I sites (excluding the one in the forward primer), 2 *BstU*I sites and two *Mae*II sites which could be used to screen for methylation of 10 of the total 26 CpG sites in the PCR product (Figure 5.7). Analysis of the CpG island associated with exon 7 found no methylation for 5 normal PB MNC samples (N1, N4, N7, N9, N10) and 5 patient samples (F9, F27, F30, F51 and F60) (Table 5.2) when assessed by a 0.5X MDE gel. Restriction enzyme digestion detected methylation in four leukaemic cell lines (Figure 5.9), 2 of which expressed the *FUT1* exon 7 transcript. The PB MNC samples which did not express the *FUT1* exon 7 transcript were unmethylated as were the PBSC samples except that they expressed this transcript (Table 5.2). The patient expression results did not correlate with the methylation status (Table 5.2) though this could be due to the DOVAM theory, which requires multiple gel running conditions to detect a single change. The cell line digestion results correlated with the band shifts seen on a 0.75X MDE gel (Figure 5.9).

Analysis of the CpG island associated with the exon 2 transcription start site of *FUT1* was quite complex. The original primers, FUT1 MS-SSCA ex 2 F and FUT1 MS-SSCA ex 2 R, yielded a product of 269 bp (Figure 5.6) and preliminary results found no band shifts for 5 normal PB MNC and 5 patient samples using a 0.5X MDE gel. However, digestion with *BstU*I and *Mae*II (Figure 5.7) revealed that there was slight digestion indicating methylation in all samples. Two of the patient samples tested, F51 and F60, had nearly 50% digestion. Running the PCR product on a 0.5X MDE gel with 5% glycerol at room temperature revealed no obvious band shifts but running the same gel at 4°C revealed some shifts but the bands were not sufficiently resolved.

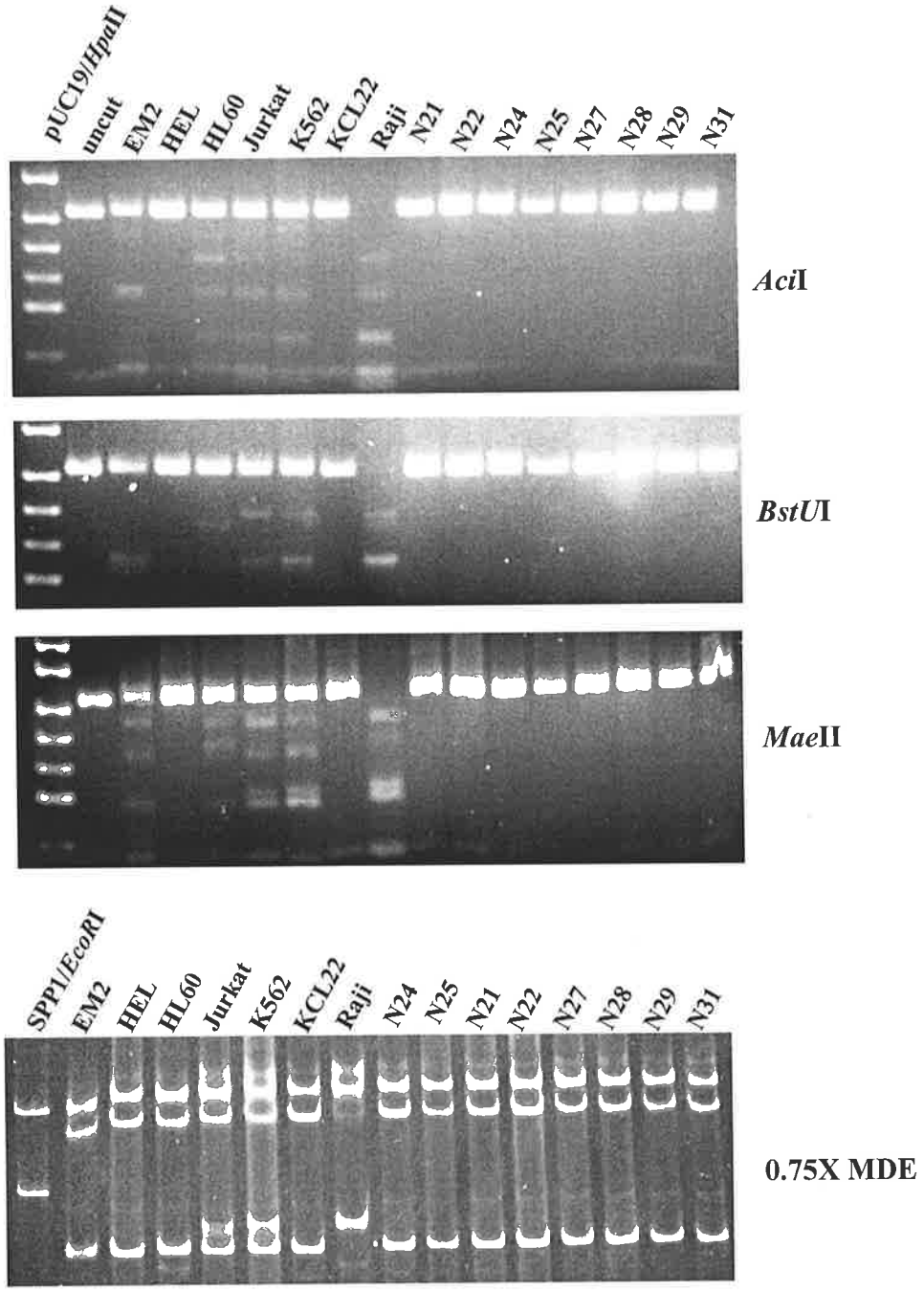


Figure 5.9. *FUT1* exon 7 MS-SSCA.

Restriction enzyme digestions with *Aci*I, *BstU*I and *Mae*II indicated that normal samples were not methylated as they all had the same pattern on a 0.75X MDE gel. The Raji cell line was fully methylated (by digestion) and this appeared as a dramatic shift on the SSCA. The other cell lines had various degrees of methylation with Jurkat and K562 having a combination of a normal unmethylated pattern and a methylated Raji-like pattern.

Two new primers were designed at approximately the same positions to the above primers but they were designed to amplify the antisense strand, as these would give different conformers and hence might resolve better on a MDE gel. The initial primers FUT1 MS-SSCA ex 2 anti F and FUT1 MS-SSCA ex 2 anti R yielded a product of 283 bp but amplified unmodified DNA. A new reverse primer, FUT1 MS-SSCA ex 2 anti R2 was designed to eliminate this problem, and yielded a PCR product of 281 bp (Figure 5.6) which contained 2 *Mae*II sites, 1 *BstU*I site, a *Taq*I site and a *Hinf*I site (Figure 5.7). Restriction enzyme digestion with *Mae*II, *Taq*I, *BstU*I and *Hinf*I allowed for the analysis of 7/9 CpG sites in the MS-SSCA PCR product.

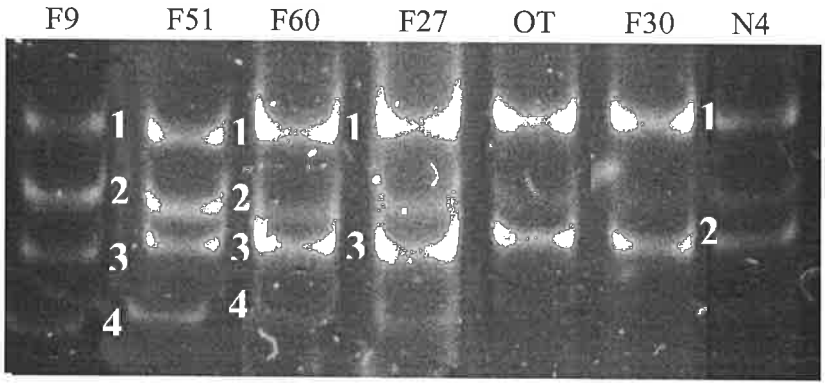
When the new FUT1 exon 2 (antisense) MS-SSCA product was resolved on a 0.75X MDE gel, band shifts were observed, so bands were stabbed, re-amplified and sequenced (Figure 5.10). There was great variation in the methylation status of each CpG site among the patient samples. For samples F30 and F60, which seemed to have 2 predominant bands rather than 4, the two bands had the same methylation pattern, that is, were methylated at all the CpG sites except the ones at 182 and 214 (Figure 5.10). F30 expresses the *FUT1* exon 2 transcript whereas F60 does not (Table 5.2) and the difference between the two is that F30 has the CpG at position 72 of the PCR product unmethylated, whereas F60 has this CpG site methylated. F9 and F51 both had four bands with the methylation status of each band different between the two samples. The expression status of F9 is unknown but F51 does not express the *FUT1* exon 2 transcript and this may be due to methylation of the CpG site at position 72.

Analysis of various normal samples indicated that they were all methylated but still had some undigested product (Figure 5.11), suggesting that perhaps one allele was methylated, though none of the normal samples express *FUT1* exon 2 transcript. The leukaemic cell lines were methylated to a greater extent than the normal samples as visualised by the presence of less undigested product except for HEL and K562 (Figure 5.11). The expression data and the

methylation status of the exon 2 *FUT1* transcript correlate for HEL and K562 which are both unmethylated at the exon 2 CpG island and express the *FUT1* exon 2 transcript.

There seemed to be an association between methylation status and expression of the *FUT1* exon 2 transcript. This transcript showed the greatest variation in expression in the patient samples (Table 4.8) suggesting that this CpG island may be important to assess. For 21 patients for which expression data was known, there was no association between methylation status and expression (Table 5.2). The methylation status was investigated by restriction enzyme digestion which assessed 7/9 CpG sites within the MS-SSCA PCR product, however, the sequencing results (Figure 5.10) would suggest that the CpG site at position 72 of the PCR product, which is part of the recognition sequence for the Sp1 transcription factor, may determine expression. This site could not be assessed by restriction digestion and SSCA analysis would be difficult since the samples were all methylated at multiple sites. The elucidation of the methylation status of this CpG site would require more sequencing which time did not permit.

MS-SSCA Gel



Sequencing Results

		CpG Site								
ID	**Band**	**72**	**94**	**103**	**115**	**127**	**129**	**182**	**214**	**253**
F9	**1**	T	C	C	T	C	C	T	T	NA
	2	T	C	T	C/T	C	C	T	NA	NA
	3	T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	NA
	4	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	NA
F30	**1**	T	C	C	C	C	C	T	T	T
	2	T	C	C	C	C	C	T	T	NA
F51	**1**	C	C	C	T	C	C	T	C	NA
	2	T	C	C	C	C	T	T	C	NA
	3	C/T	C	C	C/T	C	C	T	C	NA
	4	T	C	C	C	C	C/T	T	C	NA
F60	**1**	C	C	C	C	C	C	NA	NA	NA
	3	C	C	C	C	C	C	T	T	NA

Figure 5.10. Sequencing results of the *FUT1* exon 2 antisense MS-SSCA PCR product.
The numbers on the SSCA gel refer to the bands that were sequenced with the same numbers used in the table to indicate the methylation pattern of each band. 'NA' refers to not assessable. C denotes a cytosine unconverted after bisulphite modification due to protection by CpG methylation. T denotes a cytosine that was unmethylated and hence converted after bisulphite modification.

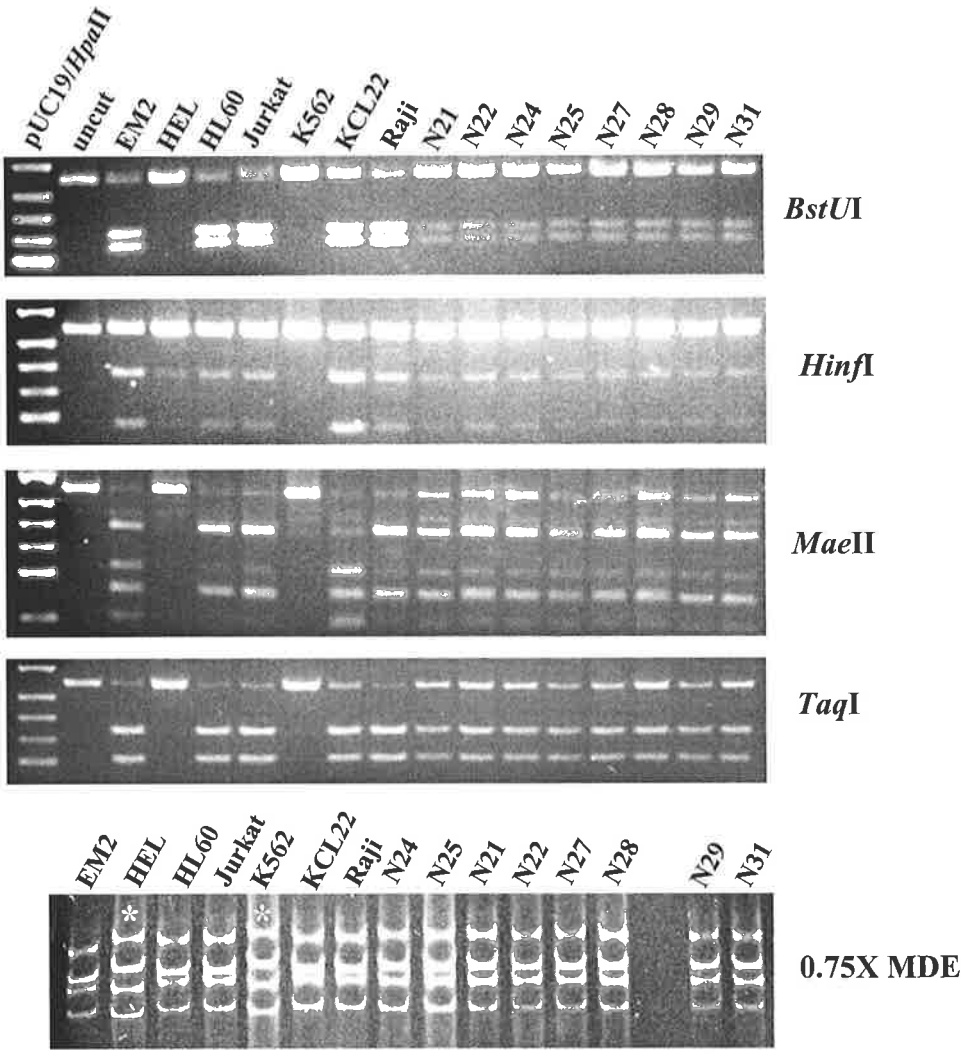


Figure 5.11. *FUT1* exon 2 antisense MS-SSCA.

The restriction enzyme digests clearly show that all normal samples are methylated to some degree but that the leukaemic cell lines are methylated to a greater extent (less undigested product) except for HEL and K562, which are unmethylated and have a different SSCA pattern (see white stars).

5.2.3 5-Aza-2'-Deoxycytidine Treatment of Leukaemic Cell Lines

Since many of the leukaemic cell lines were methylated and failed to express *ABO* and *FUT1*, an attempt to re-express these genes was performed by treating the cells with a demethylating agent, 5-aza-2'-deoxycytidine (5-aza).

K562 and HEL already expressed *ABO* and hence 5-aza treatment made no change. The cell lines Jurkat and Raji were methylated for *ABO* and did not express *ABO* but after treatment with 5-aza they re-expressed *ABO* (Figure 5.12). KCL22 and EM2 did not re-express *ABO* although they were demethylated as seen by undigested PCR product after incubation with the restriction enzymes *BstU*I, *Mae*II and *Taq*I (data not shown). Although the cell lines were demethylated but did not re-express *ABO*, this may be due to the restriction enzymes not assessing the methylation status of the CpG sites which are essential for expression.

For *FUT1*, Jurkat, KCL22 and EM2 showed no change in expression, that is, they had the same expression of *FUT1* transcripts before and after 5-aza treatment. The Raji cell line initially expressed none of the *FUT1* transcripts but after 5-aza treatment expressed all of the *FUT1* transcripts.

The re-expression of *ABO* and *FUT1* after treatment with 5-aza, in some cell lines, indicates that these genes are silenced by methylation. However, some cell lines do not re-express after 5-aza treatment indicating that there are other mechanisms responsible for silencing these genes in leukaemic cell lines. These other mechanisms may include mutation, deletion, chromosomal rearrangement or acetylation.

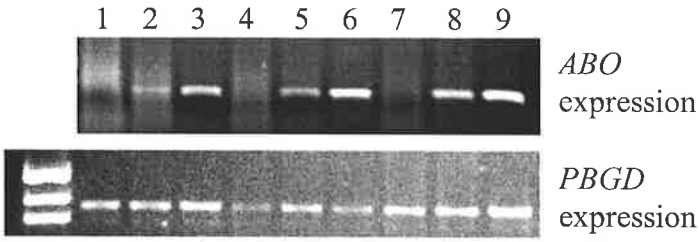


Figure 5.12. *ABO* expression in the Jurkat cell line after 5-aza-2'-deoxycytidine treatment. Lanes 1-3 is the Jurkat cell line harvested 48 h post-treatment (or after a 24 h recovery period), lanes 4-6 are Jurkat cells harvested 72 h post-treatment and lanes 7-9 are Jurkat cells harvested 96 h post-treatment. Lanes 1, 4 and 7 are Jurkat cells treated with vehicle, that is, UPW. Lanes 2, 5 and 8 are Jurkat cells treated with 1 μM of 5-aza-2'-deoxycytidine, while lanes 3, 6 and 9 are Jurkat cells treated with 2 μM of 5-aza-2'-deoxycytidine. *PBGD* is the control for cDNA synthesis and loading.

5.3 Discussion

CpG methylation of the promoters of *ABO* and *FUT1* was investigated as a mechanism responsible for the loss of ABH antigens in haematological malignancies. Assessing *ABO* for CpG island methylation was extremely difficult by Southern analysis since the probe was so GC rich that it bound non-specifically. Another disadvantage of Southern blotting was the large amount of high molecular weight DNA required, that is, 20 μg was needed for analysis with two methylation sensitive restriction enzymes. For many of the patients, this amount of DNA was not available. The other disadvantage is that Southern analysis provides information only about CpG sites found within sequences that are recognised by methylation sensitive restriction enzymes.

To overcome the limitations of Southern blotting, a new method was developed, which was based on bisulphite modification of DNA. This creates differences between the methylated and unmethylated DNA sequences since unmethylated CpGs are converted into TpGs while methylated CpGs remain as CpGs. Coupling this with PCR, which requires minimal amounts of DNA, means that numerous samples can be studied using 1 μg of DNA or less, for the analysis of more than 10 CpG islands.

Many other techniques have been developed that utilise bisulphite modification of DNA followed by PCR amplification. Bisulphite modified DNA can be amplified and sequenced, enabling the determination of the methylation status of all CpG sites (Frommer *et al*, 1992; Clark *et al*, 1994; Raizis *et al*, 1995). Genomic sequencing of bisulphite modified DNA is considered the best technique for CpG methylation analysis as each CpG site of interest can be investigated, but this is technically difficult, expensive and laborious. We needed to analyse numerous samples and genes for methylation and bisulphite sequencing was not a convenient and efficient method.

Other methods using bisulphite modification and PCR are methylation-specific PCR (MSP), which is a method that uses primers designed to distinguish methylated bisulphite modified DNA from unmethylated bisulphite modified DNA (Herman *et al*, 1996). MSP can be applied to minimal amounts of DNA and is not as laborious as sequencing. A disadvantage of MSP is that it is a qualitative measure of methylation rather than a quantitative one, and for each area of CpGs to be analysed, two sets of primers need to be designed, a methylation specific and an unmethylated specific pair.

Other groups have taken advantage of the fact that bisulphite modification of DNA leads to changes in restriction enzyme sites resulting from the action of bisulphite on methylated and non methylated CpG sites (Sadri and Hornsby, 1996). If methylation at only a few critical CpG sites is necessary to prevent binding of important transcription factors then this method is not sensitive enough (reviewed in Siegfried and Cedar, 1997) since only 25% of CpGs are amenable to analysis by restriction enzyme digestion (Sadri and Hornsby, 1996). We have used changes in restriction enzyme sites due to sequence differences between methylated and unmethylated CpGs as a supplement to SSCA.

Gonzalgo and Jones (1997) developed a quantitative method using methylation sensitive single nucleotide primer extension (Ms-SNuPE). This method is quantitative and does not use restriction enzymes. The DNA sample is bisulphite treated so unmethylated cytosines become uracil but are represented by thymine after PCR amplification, while methylated cytosines remain as cytosines after PCR amplification. Quantitation of the ratio of methylated versus unmethylated cytosines (C vs T) at original CpG sites is determined by incubating gel isolated PCR products with the appropriate Ms-SNuPE primers, *Taq* polymerase and ^{32}P-dCTP or ^{32}P-dTTP followed by denaturing polyacrylamide gel electrophoresis and phophorimage analysis. Ms-SNuPE primers are internal primers that anneal to the PCR template and terminate immediately 5' of the single nucleotide to be assayed. An advantage of this method is that many CpG sites can be analysed in a single reaction by using a combination of Ms-SNuPE primers.

All the methods published before the work on methylation carried out in this study commenced, were unable to analyse a whole area for CpG methylation, except for the Ms-SNuPE method. As such, a method was developed that utilised the bisulphite reaction, which provided the sequence differences between methylated and modified DNA, and unmethylated and modified DNA. After PCR amplification with primers specific for bisulphite converted DNA, the PCR product was denatured and run out on a MDE gel, that is, using the PCR products for single-stranded conformation analysis (SSCA). This method, MS-SSCA, utilises SSCA, which is based on single-stranded molecules taking on certain sequence-based secondary structures (conformers) under non-denaturing conditions. Strands that differ by as little as a single base substitution (that is have a thymine, unmethylated cytosine conversion, instead of a cytosine) may form different conformers and hence will not migrate through a non-denaturing polyacrylamide gel in the same way (Sheffield *et al*, 1993).

The advantage of this method is that once the normal unmethylated but modified pattern is established, any variation from the unmethylated pattern would indicate some degree of methylation. The variant bands can then be isolated and sequenced. All of the CpG sites within an island can be assessed by developing pairs of primers that will enable the whole island to be amplified in 200-250 bp overlapping fragments. The PCR product has to be less than 250 bp for it to be effective for SSCA (Sheffield *et al*, 1993). MS-SSCA PCR products between 250-300 bp have been successfully used as methylation usually occurs at more than one CpG, creating multiple sequence variations. MS-SSCA primers based on the PCR amplification of the antisense strand have also been used to create different conformers of the same CpG island which resolve better on a MDE gel.

A similar methodology was independently developed by another laboratory (Burri and Chaubert, 1999). In subsequent reports, similar methodologies have variously been called bisulphite-PCR-SSCP (BiPS) (Maekawa *et al*, 1999; Maekawa *et al*, 2001), bisulphite-SSCP (Bian *et al*, 2001) and MS-SSCP (Kinoshita *et al*, 2000) while others have used DGGE instead of SSCA (Aggerholm *et al*, 1999).

MS-SSCA analysis of the CpG island associated with *ABO* found 9/39 samples methylated in patients of which only one had alterations of *ABO* detectable by flow cytometry. The patients previously analysed for loss of ABH antigens by serology, (O'Keefe PhD thesis, 1995), were methylated at a much higher frequency (5/7). For the normal samples, 0/8 were methylated and though they all expressed *ABO* following RT-PCR, the A^2O^2 individual (N29) did not express A^2. The finding that the DOK patients had a higher rate of methylation than those identified with loss of ABH antigens by flow cytometry (Chapter 3), is not surprising since changes in ABH antigens in these patients was detected by serology, which required 50% of the cells to be abnormal. The flow cytometer could detect alterations in populations as little as 10% and the methylation status of this population would be disguised by the methylation

pattern of the larger normal population. One patient with loss by flow cytometry, but not by RT-PCR was methylated (OT16) even though they only had a small population with loss of A antigens (less than 10%). However, the methylation results of the patient may have been indicative of changes not yet translated into RBCs since the methylation analysis and the flow cytometric analysis are not assessing the same population of cells.

Recently, Hata *et al* (2002) have shown that mutation of a Sp1 site which spans the CpG site at position 203 in the ABO MS-SSCA product (Figure 5.3) causes a reduction in the promoter activity of *ABO*. While this data was not known during the course of these experiments, it clearly indicated that methylation in the *ABO* promoter may affect *ABO* expression. Analysis of this CpG site in patients with loss of *ABO* would have required analysis by sequencing. This site was not amenable to analysis by restriction enzyme digestion. However, in the future, the antisense strand could be analysed to determine whether the CpG site within the Sp1 recognition sequence coincided with a restriction enzyme recognition sequence.

Examining the *FUT1* gene for methylation was complicated since there was a CpG island associated with each of the three transcription start sites of this gene. For the exon 1 CpG island, the initial analysis of some patients and some normal samples on a 0.5X MDE did not reveal any band shifts, so the samples were all presumed to be unmethylated (Table 5.2). On further investigation with restriction enzymes, all samples were methylated, except for the Jurkat cell line (Table 5.2). When these samples were run on a 0.75X MDE gel the Jurkat cell line was the one with subtle band shifts when compared to all the other samples on the same gel (Figure 5.8). The cell line Jurkat was unmethylated for the exon 1 CpG island, yet it did not express the *FUT1* exon 1 transcript. K562, HEL and KCL22 are methylated for the exon 1 CpG island and express the *FUT1* exon 1 transcript. Restriction enzyme digestion was only able to assess 2 CpG sites with 5 not assessed. Of these 5 CpGs, three of them are within Sp1 sites and it is known that these sites are important for transcription. Sequencing or running

the samples on a MDE gel under different conditions to reveal more patterns could have been performed but as none of the patients or normal BM specimens express the exon 1 transcript, assessment of the exon 1 CpG island did not appear to be necessary.

The exon 2 CpG island of *FUT1* was difficult to assess and initially a group of samples run on a 0.5X MDE gel showed no differences in band patterns between samples, so it was assumed that they were all unmethylated (Table 5.2). However, restriction enzyme digestion indicated that the samples were all methylated to varying degrees. Various MDE gel running conditions were attempted to see if differences between moderately and heavily methylated samples could be determined. Various gel conditions, like increasing the MDE gel concentration, adding glycerol and running the gel at various temperatures, showed hints of diverse patterns but were unable to fully resolve the bands. New primers were designed to target the antisense strand, as this would yield different sequences and perhaps different conformers, which could be better resolved. Analysis of the antisense MS-SSCA products on a 0.75X MDE yielded two patterns, one with four distinct bands and the other with two predominant bands (Figure 5.10). Sequencing seemed to indicate that the CpG at position 72 of the MS-SSCA PCR product might be important in expression of this transcript since F30 expresses the *FUT1* exon 2 transcript and is unmethylated at this site while F60 does not express this transcript and is methylated at this site (Table 4.8; Figure 5.10). Analysis of the sequence reveals that an AP-2 and Sp1 site span this CpG and it is known that methylation inhibits binding of AP-2 (Chapter 1). Thus, methylation of the CpG site within an AP-2 and Sp1 recognition sequence may control the expression of the FUT1 exon 2 transcript, though more samples would need to be analysed and time did not permit this.

Since these results showed that all of the samples had some undefined level of methylation for the exon 2 CpG island, the MS-SSCA product was restriction enzyme digested to give a better indication of the amount of methylation. However, none of the restriction enzyme sites

were able to assess the CpG at position 72. Further examination of this site, by other methodologies, may have yielded a clearer picture. An association between abolishing the AP-2 site and abolishing *FUT1* exon 2 transcription would also require further investigation.

For *FUT1* exon 7 CpG island, the digestion results and the SSCA results correlated (Table 5.2). Normal samples and patients samples were all unmethylated but some of the normal samples, the PBSC, are known to express this transcript while the PB MNCs do not. The fact that the PB MNC sample did not express this transcript but were unmethylated, indicated that methylation is not responsible for silencing of this transcript. Other mechanisms, like acetylation, may be responsible for silencing of the exon 7 *FUT1* transcript.

The methylation analysis of the *ABO* and *FUT1* genes was further complicated by the heterogeneity of the samples, as the normal population was larger than the abnormal population, which had loss of ABH antigen expression. The methylation status of normal BM could not be determined due to the limited availability and quantity of these samples. PB MNC samples may not always be an adequate control for the normal pattern of methylation since their *ABO* and *FUT1* expression pattern is very different to the BM situation.

When assessed by restriction enzyme digestion most samples had some undigested product, representing unmethylated sequence, yet this was not obvious by SSCA. The DOVAM paper (Liu *et al*, 1999) shows that numerous gel running conditions are required to detect all possible mutations. The MS-SSCA methodologies adopted in these experiments may have required further optimisation, though more than one condition was tried for all MS-SSCA PCR products.

The 5-azacytidine experiments were able to establish an association between expression and methylation with some cells re-expressing *ABO* or *FUT1* after treatment with a demethylating agent. Some cell lines failed to express after 5-azacytidine treatment suggesting that other mechanisms are at work, like acetylation or mutation. These mechanisms would need to be analysed in the patients with loss of ABH antigens to determine whether they are responsible for silencing of *ABO* or *FUT1*.

The lack of correlation between methylation and expression may be due to the fact that when assessing methylation, the whole population of cells are assessed, yet the expression or lack of expression may be from a sub-population of cells. This was apparent when assessing methylation of the patients with ABH alteration by flow cytometry and those with ABH alterations by serology. The patients with ABH alterations by serology have at least 50% of cells with loss and this group of patients had a higher frequency of methylation (5/7). The patients with ABH alteration by flow cytometry have 10-20% of cells and the larger normal population may obscure methylation of the population with ABH antigen alterations.

Though the methylation results were unclear, the use of the MS-SSCA technique was quite advantageous over previous techniques. MS-SSCA is a rapid technique with a very low false negative rate and in many cases will be the method of choice for methylation analysis. In particular, MS-SSCA is an efficient approach for genomic sequencing studies of methylation as methylated samples can be identified by the presence of variant conformers, which can then be directly sequenced.

CHAPTER 6

Analysis of 9q34 Genes

6.1 Introduction

Tumour suppressor genes serve as negative regulators of cell growth and their inactivation leads to tumorigenesis (reviewed in Stanbridge, 1990). Knudson's two-hit hypothesis (1971) proposes that inactivation of both alleles of a tumour suppressor gene are required for tumorigenesis. Certain characteristics of the mechanisms resulting in the inactivation of tumour suppressor genes may be used for identifying and locating these gene; one approach is to assay for tumour-specific loss of genetic material. Detection of allele loss utilises polymorphic markers like RFLPs, VNTRs and microsatellites. The detection of only one allelic variant in the tumour (homozygous) versus two allelic variants in corresponding normal cells (heterozygous) indicates loss of genetic material in the tumour (Goddard and Solomon, 1993). Studies of chromosome-specific LOH in tumour cells affecting limited genomic regions have yielded gene localisations of tumour suppressor genes. LOH has been used to suggest the presence of tumour suppressor genes in many cancers.

Studies with inherited familial cancers led to the identification of specific tumour suppressor genes, the first of which was the tumour suppressor gene responsible for retinoblastoma (*Rb*). *Rb* was localised by the observation that normal cells from some individuals with familial retinoblastoma bore a deletion at 13q14 (reviewed in Goddard and Solomon, 1993). Sporadic cases had the same region deleted in the tumour cells but not the normal cells. For colorectal cancers, cytogenetic abnormalities are observed at 5q21 and 18q21. Positional cloning at these loci has led to the discovery of two genes involved in colorectal cancer, DCC at 18q21 and FAP at 5q21 (reviewed in Boyd and Barrett, 1990).

Investigation of LOH of the chromosomal region where *ABO* is localised, 9q34, is one focus of this chapter. The possibility of regional silencing of the region harbouring *ABO* will also be examined. LOH and regional silencing would suggest the presence of a tumour suppressor

gene in this region. Though alterations of ABH antigens are a frequent occurrence in haematological malignancies (Chapter 3), the role of this alteration in leukemogenesis is unclear. Bombay individuals have no H antigen and hence are not able to make A or B antigens yet they do not have an increased predisposition to leukaemia. In solid tumours, loss of ABH antigens is associated with increased metastatic potential of the tumour and is a marker for poor prognosis (see section 1.4.7). Such associations for the alterations of ABH antigens in haematological malignancies have not been elucidated.

It is known that the 9q22-qter region is frequently deleted in myeloid malignancies (Evdokiou *et al*, 1993) with many cases of AML having aberrations at 9q34 (Mitelman Database of Chromosome Aberrations in Cancer, 2002). Since changes at *ABO* have been associated with changes at other 9q34 loci, it is possible that alterations of *ABO* are not the leukaemia causing event but rather a marker to other events occurring at this chromosomal region. This is supported by the reported observation of individuals with decreases in both *ABO* and adenylate kinase (*AK1*) expression (Salmon *et al*, 1968; Kahn *et al*, 1971; Marsden *et al*, 1992). *AK1* is localised at 9q34.11, not too distant from *ABO* at 9q34.2. Salmon *et al* (1968) found a patient with loss of A and when they separated this sample into the A and O-like populations, it was noted that a decrease in *AK1* occurred in the O-like population. Kahn *et al* (1971) found significant decreases in *AK1* activity in an A negative population from a patient whose cells demonstrated 50% loss of A. The normal A positive population had normal levels of *AK1*.

In this study, a PCR RFLP approach was used to analyse 9q34 loci for LOH and mono-allelic expression arising from regional silencing. For some 9q34 loci, previously published RFLPs or SNPs in coding regions were utilised but when this study first commenced the SNP database was not as comprehensive as now. SNPs were identified by aligning ests from the

Unigene database using the alignment program Sequencher. If there were variations in numerous ests at a single base position, then this was identified as a potential RFLP marker.

Primers were designed to span the potential SNP and if it did not create or abolish a restriction enzyme recognition site, a primer was engineered to create one. For example, the T->A SNP used for *PPP2R4*, when there was a T a *Mae*III site was present but this enzyme is 80X more expensive than *Hinf*I. By creating a mismatch in the primer such that it is incorporated into the PCR, enabled the creation of a *Hinf*I site when the *PPP2R4* SNP is an A and not a T (Figure 6.1).

Figure 6.1. Creating a restriction enzyme site for SNP analysis.
A heterozygous sequence for the *PPP2R4* SNP is shown. The bold capital letters are the SNP and the bold lower case letter is the base pair engineered into the PCR product to form a *Hinf*I site (recognition sequence underlined) when the SNP is an A. No *Hinf*I site is created when the SNP is a T. The grey shading is the *Mae*III site.

The potential SNPs were then assessed in a group of normal individuals and if heterozygotes were common then the patient samples were genotyped at the SNPs by PCR followed by restriction enzyme digestion. Patients that were informative were also investigated for mono-allelic expression.

6.2 Results

6.2.1 Genotyping SNPs at 9q34 Loci

A total of 17 SNPs for 13 genes were analysed and the data is summarised in Table 6.1 and Table 6.2. The loci were genotyped at the SNP and the sequence at the SNP is shown in Table 6.1 and Table 6.2. If the sample was informative, that is heterozygous, then the genotype would consist of two different base pairs, one for each allele. For example, for *AK1*, a heterozygous sample is 'AC' while homozygous samples are either 'AA' or 'CC'. All primers for the various loci are shown in Table 2.1.1 and the restriction enzyme adjacent to the loci name in Table 6.1 and Table 6.2 indicates the restriction enzyme used for the PCR RFLP analysis. More than one restriction enzyme indicates that the loci had more than one SNP genotyped.

Some SNPs were not useful and these included the published RFLP for the *DBH* locus. When the expression of *DBH* was analysed, it was found that it was not expressed in PB MNC or BM. After this discovery, all subsequent loci that were investigated had to fulfil the criteria of being expressed in PB MNC or BM. Other loci, *SURF3*, *LCN2*, *HSPA5* and *FBXW2*, appeared to have SNPs when the Unigene ests were aligned but after PCR RFLP analysis of numerous samples, no SNPs were found. Further to this, heteroduplex analysis was attempted but did not identify any SNPs at these loci. Other loci had PCR RFLP systems designed and operational but due to lack of time they were not able to be fully assessed. These include *RALGDS*, *AGAPT2*, *ASS* and *SURF4*.

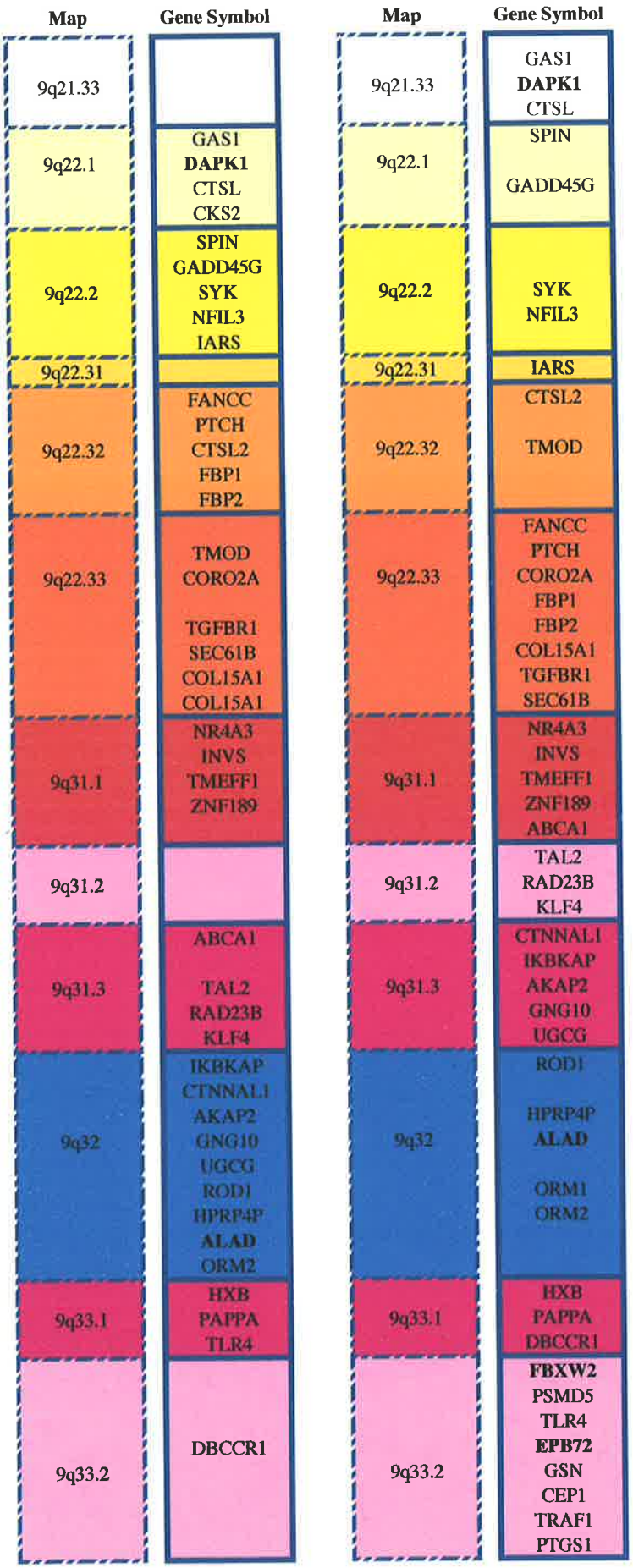
Other difficulties in gene loci included their location. Figure 6.2 shows the position of various loci on 9q based on the human genome sequence in October 2000 and December 2001. For this study, the October 2000 sequence was used for the identification of loci to be analysed. In December 2001 this map was revised and the position of a number of loci were altered.

However, fortunately all of the markers chosen retained their same approximate distance and were still valid choices for the study. For example, *ALAD* was chosen because it is localised in 9q32, is far away from *ABO*, and in the December 2001 map it remains in 9q32. *DAPK1* was chosen as it is very close to *GAS1*, the gene that localises to a chromosome band frequently deleted in AML (Evdokiou *et al*, 1993) and at the time of the analysis a SNP was not found in *GAS1*.

The region frequently deleted in AML extends from 9q21 to the telomere of 9q and it is possible that there are two regions harbouring potential tumour suppressor genes on 9q, like bladder cancer. In bladder cancer, evidence for tumour suppressor genes have been localised to the regions 9p12-q21 and telomeric of 9q33 (Simoneau *et al*, 1996). In the October 2000 sequence *DAPK1* was in 9q22.1 but in the October 2000 sequence it was in 9q21.33 (Figure 6.2). *ABO* was mapped closer to the telomere in the October 2000 map but in the December 2001 sequence was mapped closer to the centromere but still in 9q34.2 (Figure 6.2).

6.2.1.1 9q34 SNP Analysis in Normal Individuals

The genotype of the SNPs in normal individuals is shown in Table 6.1. The rate of heterozygosity for each SNP varied from 15-60%, though the sample size was small for some loci. The presence of two or more SNPs for one locus meant that more informative cases for each locus could be found. There were no allele dosage changes seen for any normal samples analysed as assessed by a PCR-RFLP technique (Table 6.1).


Map
Gene Symbol
9q21.33
9q22.1
GAS1
DAPK1
CTSL
CKS2
9q22.2
SPIN
GADD45G
SYK
NFIL3
IARS
9q22.31
9q22.32
FANCC
PTCH
CTSL2
FBP1
FBP2
9q22.33
TMOD
CORO2A
TGFBR1
SEC61B
COL15A1
COL15A1
9q31.1
NR4A3
INVS
TMEFF1
ZNF189
9q31.2
9q31.3
ABCA1
TAL2
RAD23B
KLF4
9q32
IKBKAP
CTNNAL1
AKAP2
GNG10
UGCG
ROD1
HPRP4P
ALAD
ORM2
9q33.1
HXB
PAPPA
TLR4
9q33.2
DBCCR1
Map
Gene Symbol
9q21.33
GAS1
DAPK1
CTSL
9q22.1
SPIN
GADD45G
9q22.2
SYK
NFIL3
9q22.31
IARS
9q22.32
CTSL2
TMOD
9q22.33
FANCC
PTCH
CORO2A
FBP1
FBP2
COL15A1
TGFBR1
SEC61B
9q31.1
NR4A3
INVS
TMEFF1
ZNF189
ABCA1
9q31.2
TAL2
RAD23B
KLF4
9q31.3
CTNNAL1
IKBKAP
AKAP2
GNG10
UGCG
9q32
ROD1
HPRP4P
ALAD
ORM1
ORM2
9q33.1
HXB
PAPPA
DBCCR1
9q33.2
FBXW2
PSMD5
TLR4
EPB72
GSN
CEP1
TRAF1
PTGS1

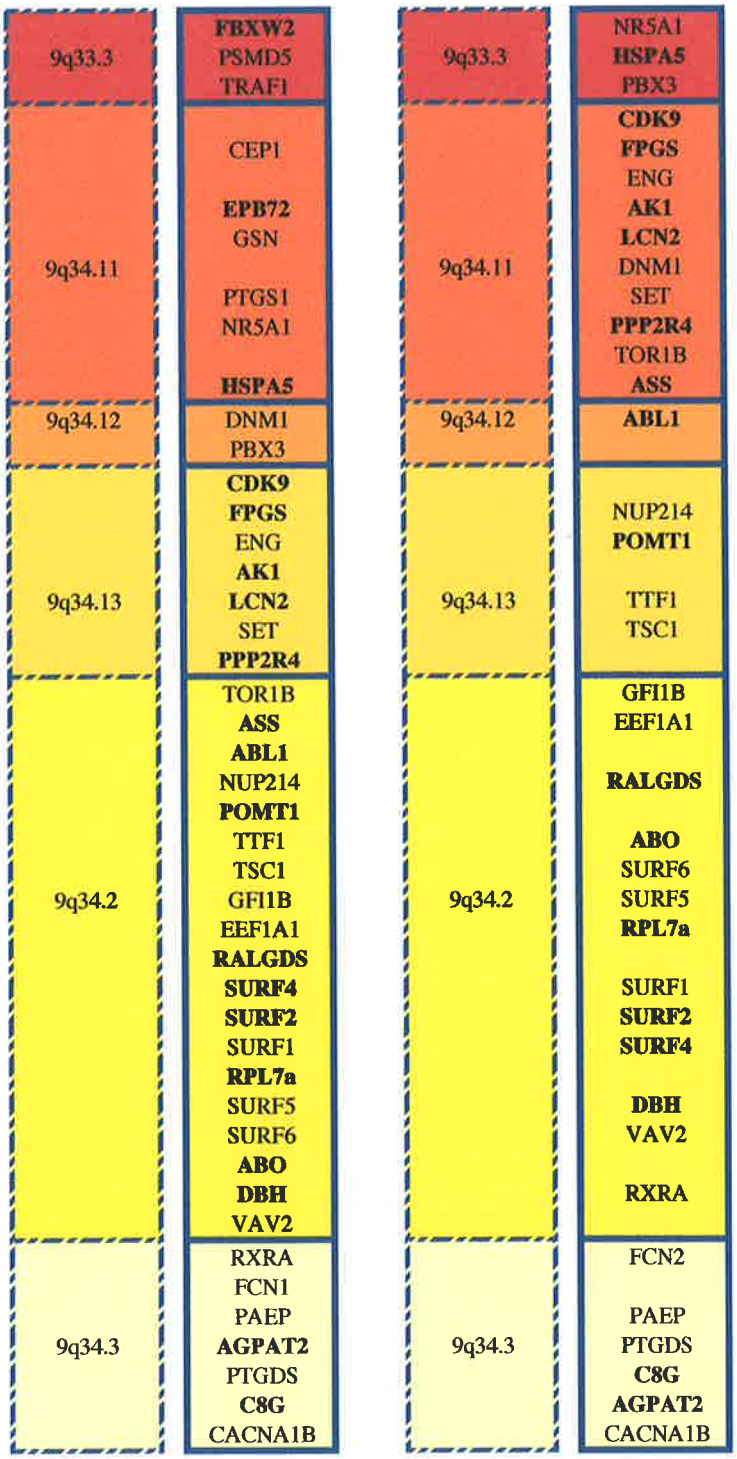


Figure 6.2. Comparison of the 9q map based on the human genome sequence from October 2000 and December 2001.
The map on the left was based on the human genome sequence as of October 2000 and was the map used for selecting loci for the 9q34 SNP analysis. The map on the right was based on the December 2001 human genome sequence and shows how the loci are now in different positions on 9q and in relation to ABO. The genes in bold were the ones analysed for LOH and for analysing allelic expression. The position of the loci was based on the USCS Genome Browser (http://genome.cse.uscs.edu).

ID	SAMPLE	*DAPK1* *Tas*I	*ALAD* *Msp*I	*ALAD* *Rsa*I	*EPB72* *Rsa*I	*CDK9* *EcoN*1	*FPGS* *BstN*I	*AK1* *Taq*I	*PPP2R4* *Hinf*1	*ABL1* *BstN*I	*ABL1* *Cfo*I	*ABL1* *Mae*II	*POMT 1 BstN*1	*POMT1* *Hpa*II	*RALGDS* *BsiE*I	*SURF2* *Alu*I	*C8G* *Dde*I
N1	PB MNC	nd	GG	CT	AA	AG	CT	AC	nd	CC	AA	CC	nd	nd	nd	GG	GT
N3	PB MNC	AG	GG	CT	AG	nd	CT	AA	nd	CC	AA	CC	nd	nd	CC	AA	GG
N4	PB MNC	nd	GG	CC	AA	nd	CT	CC	nd	CG	AG	CT	nd	nd	TT	AA	GG
N6	PB MNC	nd	GG	TT	AA	nd	nd	nd	nd	CC	AA	CC	nd	nd	nd	GG	nd
N7	PB MNC	AA	GG	CT	GG	GG	TT	AA	nd	CC	AA	CC	CC	GG	CT	GG	GT
N8	PB MNC	AA	GG	CT	AG	GG	TT	AA	nd	CC	AA	CC	CC	GG	TT	AG	GT
N9	PB MNC	nd	GG	TT	AA	GG	TT	AA	nd	CC	AA	CC	nd	nd	TT	AA	GG
N10	PB MNC	AA	CG	TT	AA	AA	CT	CC	nd	CC	AA	CC	nd	nd	CT	AA	TT
N14	PB MNC	nd	GG	CT	GG	nd	TT	AC	nd	CC	AA	CC	nd	nd	nd	GG	GG
N15	PB MNC	nd	GG	CC	GG	nd	TT	CC	nd	CC	AG	CC	nd	nd	nd	GG	GG
N16	PB MNC	nd	CG	CT	AG	nd	CT	AC	nd	CC	AA	CC	nd	nd	nd	GG	GG
N18	PB MNC	AA	GG	TT	AA	AG	CT	AA	TT	CC	AA	CC	CT	GG	CT	AG	GG
N19	PB MNC	nd	GG	CT	AA	GG	TT	AA	AA	CG	GG	CT	CC	GG	TT	AG	GT
N20	PB MNC	nd	GG	CT	AG	nd	TT	AA	TT	CC	AG	CC	nd	nd	CC	GG	TT
N21	PB MNC	nd	GG	CT	AA	nd	TT	AC	nd	CC	AA	CC	nd	nd	CC	GG	GT
N22	PB MNC	AA	GG	CT	AA	AG	CT	AA	AT	CC	AG	CC	CC	GG	CT	AA	GG
N23	PB MNC	AG	GG	TT	AA	nd	CT	AA	AT	CC	AA	CC	nd	nd	CC	AA	GT
N24	PB MNC	nd	GG	CT	AA	AG	CT	AC	nd	CC	AA	nd	nd	nd	nd	GG	GT
N25	PB MNC	AA	GG	TT	AG	AA	CT	AA	nd	CC	AA	nd	nd	nd	nd	GG	nd
N26	PBSC	nd	CG	TT	AG	nd	TT	AA	nd	nd	nd	nd	nd	nd	nd	nd	GG
N27	PBSC	AG	GG	CC	AA	AG	CT	AA	AA	CG	nd	nd	nd	nd	nd	AG	TT
N28	PBSC	GG	GG	TT	GG	GG	TT	AA	TT	CC	AA	nd	CC	GG	nd	nd	nd
N29	PBSC	AG	GG	CT	AA	GG	TT	AC	TT	CG	nd	nd	CC	GG	nd	AG	GT
N30	PBSC	AA	CG	CT	AA	AG	CT	AC	nd	CC	AG	nd	nd	nd	nd	GG	GT
N31	PBSC	AG	GG	CT	AG	AG	CT	AA	AT	CC	AA	nd	CC	GG	nd	GG	TT
N32	PBSC	GG	GG	CT	AA	GG	TT	AA	TT	CC	AA	nd	nd	nd	nd	AG	GG
K562	cell line	AA	GG	TT	AA	AA	CT	CC	TT	CC	**AG***	**CT***	TT	GG	TT	GG	GG
KCL22	cell line	AA	GG	CT	AG	AG	CT	AA	AA	CC	AA	CC	CC	GG	TT	AA	GG
SW48	cell line	nd	GG	CT	AG	nd	CT	AA	nd	CC	AA	CC	nd	nd	TT	AG	GT
LIM1215	cell line	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
HEL	cell line	AA	GG	TT	AA	AA	CT	AA	TT	CC	AA	nd	CC	GG	CC	AA	GG
HL60	cell line	nd	GG	TT	AA	GG	TT	AA	AT	CC	AA	nd	nd	nd	nd	AA	GG
CCRF-CEM	cell line	-	nd	nd	-	AA	CC	-	TT	-	nd	nd	nd	nd	nd	-	nd
EM2	cell line	GG	nd	nd	AA	GG	TT	AA	TT	CC	AA	nd	CC	GG	nd	GG	TT
RAJI	cell line	GG	GG	CT	AG	GG	CT	CC	AT	CC	GG	nd	TT	GG	nd	GG	GG
Jurkat	cell line	AA	CG	CT	AA	GG	TT	AA	TT	CC	AA	nd	**CT***	AG	nd	AG	GG
BM1	BM	AA	CG	TT	AA	AA	CT	AA	AT	CC	AG	nd	nd	nd	nd	AG	nd
BM2	BM	nd	GG	CT	AG	AA	CT	AC	nd	CC	nd	nd	nd	nd	nd	GG	GG
BM3	BM	AG	GG	CT	AA	AG	CT	AA	AT	CC	AG	nd	nd	nd	nd	AG	GG

Table 6.1. SNP genotypes for 9q loci in normal samples and leukaemic cell lines.

The loci are shown in order from centromere to telomere using the December 2001 human genome mapping data. The locus is shown above each column with the restriction enzyme used for genotyping the SNP. The genotype at both alleles is shown for each SNP. 'nd' is not done. '-' refers to the lack of a PCR product even after a secondary PCR. Samples in bold with a star had changes in allele dosage.

6.2.1.2 9q34 SNP Analysis in Leukaemic Cell Lines

The genotypes for the SNPs in the leukaemic cell lines are shown in Table 6.1. There were no allele dosage changes seen in any of the normal individuals but this was not the case for the leukaemic cell lines. The Jurkat cell line had an allele dosage shift for *POMT1* (Figure 6.3A) which is located at 9q34.13 and is adjacent to the chromosomal region harbouring *ABO*. This cell line was also shown to have an allele shift for *ABO* (Chapter 4). The leukaemic cell line K562 showed an allele dosage change for the *ABL1* locus, which is closer to *POMT1* than *ABO* (Figure 6.3B). The cell line CCRF-CEM is known to have extensive chromosome 9 deletions and rearrangements as seen by its karyotype: 90(88-101)<4n>XX, -X, -X, +20, +20, t(8;9)(p11;p24)x2, der(9)del(9)(p21-22)del(9)(q11q13-21)x2; sideline with +5, +21, add(13)(q3?3), del(16)(q12) (http://www.dsmz.de/). Therefore, it was not surprising that for some 9q loci, that is, *ABL1*, *AK1*, *DAPK*, *EPB72* and *SURF2* it was not possible to PCR amplify from CCRF-CEM, even after a secondary PCR amplification. It is interesting that CCRF-CEM has a pattern, from centromere to telomere, of negative for *EPB72*, positive for *CDK9* and *FPGS*, then negative for *AK1*, positive for *PPP2R4* and negative for *ABL1*.

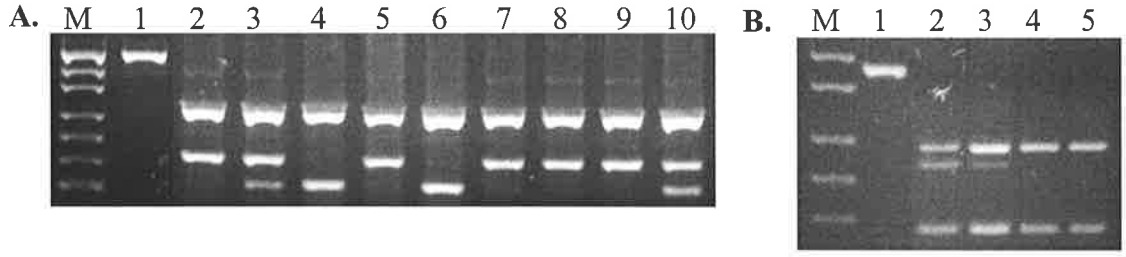


Figure 6.3. Allele dosage changes in leukaemic cell lines.

A. M refers to the marker pUC19/*Hpa*II. Lane 1 is uncut *POMT1* PCR products while lanes 2-10 are *POMT1* PCR product digested with *Bst*NI. Lanes 2-6 are leukaemic cell lines HEL, Jurkat, K562, KCL22 and Raji respectively. Lanes 7-10 are samples from normal individuals N28, N29, N31 and N18 respectively. Lanes 2, 5, 7, 8 and 9 genotype as 'CC' at the SNP while lanes 4 and 6 genotype as 'TT'. Lanes 3 and 10 are heterozygous, 'CT', but Jurkat has an allele dosage shift when the intensity of its bands are compared to the heterozygous normal sample in lane 10.

B. M refers to the marker pUC19/*Hpa*II. Lane 1 is uncut *ABL* PCR products while lanes 2-5 are *ABL* PCR product digested with *Cfo*I. Lane 2 is a heterozygous normal sample N20, lane 3 is the heterozygous sample K562 while lanes 4 and 5 are homozygous samples from KCL22 and SW48 respectively. Comparing lanes 2 and 3 shows the allele dosage shift in K562 at this locus.

6.2.1.3 9q34 SNP Analysis in Leukaemic Patients

The SNP data for 80 patients is summarised in Table 6.2. The heterozygous frequencies for the various SNPs are comparable to the normal individuals. Only a small number of shifts in allele dosage were evident. An allele dosage shift was seen for patient F3 at the *POMT1* locus. This patient had no *ABO* or *FUT1* alterations at the mRNA level when compared to normal samples but the flow results were unclear due to transfusions. *POMT1* is found at 9q34.13 while *ABO* is located at 9q34.2, in between *RALGDS* and *SURF2*.

Patient F11 had an alteration at *PPP2R4* and is an A^1O^1 patient that does not express the A^1 allele. Patient F21 had a shift at *POMT1* and had no *ABO* alterations.

F22 was the only patient that had an allele dosage shift for more than one locus. This patient had an alteration at *AK1* and *DAPK1*. *DAPK1* is localised to 9q21.33, while *AK1* is found at 9q34.11. This patient had numerous cytogenetic abnormalities (Table 3.2) but none involved chromosome 9. This patient had no ABH alterations.

Patient F54 had no ABH alterations but had an allele dosage shift at *POMT1*. F60 is a loss of A patient with a dosage alteration at *POMT1* and OT17, the non-leukaemic patient with an *ABO* duplication, had a dosage alteration at *PPP2R4*.

Another interesting observation was that all patients were informative for more that one SNP except for F24 which was informative for 1/11 SNPs, F25 which was informative for 1/15 SNPs and OT27 which was informative for 1/11 SNPs. F24 is a loss of B patient while F25 is a loss of A patient. Since the samples can be quite pure and contain high amounts of leukaemic blasts, it is possible that the reason for these patients not being informative for more that one loci is that they may have LOH but it is not seen due to the lack of normal cells

which would give the PCR signal for both alleles. A normal sample would be required, possibly from isolating the T-cells from the specimens or by a buccal cell sample.

An interesting observation is that 6/8 alterations occur at *POMT1* and *PPP2R4* (Table 6.3) which are relatively close to each other on chromosome 9q34. It may be that one of these loci is a potential tumour suppressor gene or that one lies within this area. The other interesting observation is that 50% of the alterations occur at *POMT1*, which is in 9q34.13, not too distant from *ABO* at 9q34.2.

Of the 7 patients with allele dosage alterations, 3 had alterations in *A* mRNA expression (F11, F60 and OT17; Table 6.3 and Table 4.7). Two of the patients were O^1O^1 homozygotes so loss of one *O* allele would not be seen since there are two copies. Loss of the *O* allele can not be assessed by flow cytometry since the *O* allele encodes for an inactive transferase. Therefore, 3/5 patients have allele dosage alterations and alterations in *A* mRNA expression. This may indicate that the change in *ABO* expression in these patients may be due to other 9q loci alterations or regional silencing of the chromosomal region harbouring *ABO*.

One of the loci that had allele dosage shifts was *POMT1*. There is not much data on *POMT1* except that it encodes a protein with high homology to the yeast mannosyl-transferases. It is expressed at the highest levels in testis and foetal brain. *POMT1* may be involved in O-mannosylation, an essential protein modification, and is the first of its kind to be discovered in humans (Jurado *et al*, 1999). The conservation of *POMT1* among other species supports the fact that O-mannosylation is an essential physiological process. O-mannosylation starts at the endoplasmic reticulum with the transfer of mannose to seryl or threonyl residues of secretory proteins (Strahl-Bolsinger *et al*, 1999). This reaction is catalysed by a family of protein O-mannosyltransferases, of which *POMT1* belongs. O-mannosylation is rare in mammals

occurring in a limited number of glycoproteins of brain, nerve and skeletal muscle (Endo, 1999). *POMT1* is also a homologue of the *Drosophila melanogaster* rotated abdomen gene, which is involved in embryonic muscle development (Jurado *et al*, 1999).

PPP2R4 regulates the protein phosphatase 2A (PP2A) protein. *PPP2R4* is also known as *PTPA* and is ubiquitously expressed. *PPP2R4* lacks a TATA or CAAT-box but contains a CpG island surrounding the transcription initiation site (Janssens *et al*, 1999). Studies of yeast homologues of *PPP2R4* indicate that it is essential for G1 progression in the cell cycle (Mitchell and Sprague, 2001). The loci that are mapped in between *PPP2R4* and *POMT1* include *ABL1* and *NUP214* (also known as *CAN* or *CAIN*), which are involved in translocations in myeloid leukaemias. This may indicate that this region of 9q is a potential region harbouring a tumour suppressor gene(s) implicated in myeloid malignancies.

6.2.2 Expression of 9q34 Loci SNPs

Once all the samples had been genotyped, it was very simple to use the SNPs to look at allelic expression, similar to how the restriction enzymes created by the polymorphic differences between the various *ABO* alleles were used to assess the allelic expression of *ABO* alleles. For some loci, primers could be designed that did not amplify genomic DNA, resulting in a larger pool of patients that could be analysed. For some of the SNPs, it was impossible to design new primers to avoid genomic DNA contamination so the same primers used for genotyping the SNP were used for the expression analysis, except that the RNA was DNase I treated before being reverse transcribed. There was not sufficient material to do this for all of the patients but as many as possible were analysed.

ID	*DAPK1* *Tas* I	*ALAD* *Msp* I	*ALAD* *Rsa* I	*EPB72* *Rsa* I	*CDK9* *EcoN* I	*FPGS* *BstN* I	*AK1* *Taq* I	*PPP2R4* *Hinf* I	*ASS* *Nla* III	*ABL1* *BstN* I	*ABL1* *Cfo* I	*ABL1* *Mae* II	*POMT1* *BstN* I	*POMT1* *Hpa* II	*RALGDS* *BsiE* I	*SURF2* *Alu* I	*C8G* *Dde* I
F1	nd	GG	CT	AG	AG	nd	AC	nd	nd	CC	nd	nd	CT	AG	nd	nd	nd
F2	AG	GG	TT	AG	AG	CT	AA	AT	nd	CC	AA	nd	nd	nd	nd	AG	GT
F3	AA	GG	TT	AA	AG	CT	AA	AT	CC	CC	AA	nd	**CT***	GG	nd	GG	GT
F4	AA	GG	CT	AG	AA	CT	nd	AA	nd	CC	AA	nd	nd	nd	nd	AG	GT
F7	AA	GG	CC	AG	AG	CT	AA	TT	nd	CC	AA	CC	CC	GG	nd	GG	GG
F9	AG	GG	TT	AG	GG	TT	AA	AT	nd	CG	AG	CT	CT	GG	nd	GG	GG
F10	AA	GG	CT	AG	AG	CT	AA	TT	nd	CC	AA	CC	CC	GG	nd	AG	GG
F11	AG	CG	CT	AA	AG	TT	AA	**AT***	nd	CC	AA	nd	nd	nd	nd	GG	GG
F12	AA	GG	TT	AA	AA	CT	AA	AT	nd	CC	AA	nd	nd	nd	nd	nd	GT
F13	AA	CG	TT	AG	GG	TT	AC	TT	nd	CG	nd	nd	nd	nd	nd	GG	GT
F14	AG	GG	CT	AG	GG	TT	AA	nd	nd	CC	AA	nd	nd	nd	nd	GG	GG
F15	AG	GG	CT	AA	AG	CT	AC	AT	nd	CC	AA	nd	nd	nd	nd	AG	TT
F17	AG	GG	CC	AA	AG	CT	AC	AT	nd	CC	AA	nd	nd	nd	nd	AG	TT
F19	GG	nd	nd	AG	GG	CT	nd	AT	nd	CG	nd	nd	nd	nd	nd	nd	nd
F20	AG	GG	CT	AA	AG	CT	CC	TT	CC	CC	AG	CC	CC	GG	nd	GG	GG
F21	AA	GG	TT	AA	AG	CT	AC	TT	CC	CC	AA	nd	**CT***	GG	nd	AG	GG
F22	**AG***	GG	CC	AA	nd	CT	**AC***	nd	nd	CC	AA	CC	nd	nd	nd	GG	GG
F23	AA	GG	CT	AG	AG	CT	AC	AT	nd	CG	nd	nd	nd	nd	nd	AG	TT
F24	AA	GG	TT	AA	GG	TT	AA	TT	nd	CC	AA	nd	nd	nd	nd	nd	GT
F25	AG	GG	CC	AA	GG	TT	AA	TT	nd	CC	AA	CC	CC	GG	nd	GG	GG
F27	AA	GG	CT	AA	AG	CT	AA	AA	nd	CG	GG	CT	CC	GG	CT	AA	GT
F30	AG	GG	TT	AG	GG	TT	AA	AA	nd	CC	AA	nd	CC	GG	nd	AG	GT
F31	nd	GG	TT	GG	AG	CT	AA	AT	nd	CC	AG	nd	CT	AG	nd	AG	nd
F32	AA	CG	TT	nd	nd	CT	AC	nd	CC	CC	nd	nd	nd	nd	CT	nd	GT
F37	AA	GG	CT	AG	nd	nd	AA	TT	nd	CC	AA	CC	nd	nd	CT	nd	GG
F38	GG	GG	CC	AG	AG	CT	AA	TT	nd	CC	AA	nd	CC	GG	CT	AG	GT
F39	GG	GG	CT	GG	AA	CT	AA	TT	CT	CG	AG	CT	CC	GG	nd	GG	GT
F41	nd	nd	nd	AG	GG	TT	AA	nd	nd	nd	nd	nd	nd	nd	nd	nd	GG
F42	AA	GG	TT	AA	AA	CT	AA	TT	nd	CC	AA	CC	nd	nd	nd	AG	TT
F43	AA	GG	CC	AG	GG	TT	AC	AA	nd	CC	AA	nd	nd	nd	nd	nd	nd
F44	AA	CG	TT	AG	AG	CT	AA	TT	nd	CC	nd	nd	nd	nd	nd	AG	GT
F45	AG	CG	CT	AG	GG	TT	AA	AA	nd	CC	AA	nd	nd	nd	nd	AA	TT
F46	AG	GG	CT	AA	AG	CT	AA	TT	nd	CC	AA	CC	CC	GG	nd	GG	GG
F48	GG	GG	CT	AA	GG	CT	AC	TT	nd	CC	AA	nd	nd	nd	nd	AG	GG
F50	AG	GG	CT	AA	AG	CT	AA	AT	nd	CC	AA	CC	CT	AG	nd	AA	GT
F51	AG	GG	CT	AG	AG	CT	AA	TT	nd	CC	AA	CC	TT	AA	nd	GG	TT
F52	AA	GG	CT	AG	AA	CT	AC	TT	nd	CC	AA	CC	nd	nd	nd	AG	GG
F53	AA	GG	CT	AA	GG	TT	AA	TT	nd	CG	nd	nd	CC	GG	nd	nd	GT
F54	nd	GG	CT	AG	AA	nd	AA	nd	nd	CC	nd	nd	TT	**AG***	nd	nd	nd
F57	AA	GG	CT	AG	GG	TT	AA	AA	nd	CC	AA	CC	CT	AG	nd	AA	TT
F58	AA	GG	CC	AA	GG	TT	CC	TT	nd	CC	AA	nd	CT	AG	nd	AA	GT
F59	AA	GG	CT	AG	AG	TT	AC	nd	nd	CC	nd	nd	nd	nd	nd	nd	GG
F60	AG	GG	TT	GG	GG	TT	AA	TT	nd	CC	AA	nd	CC	**AG***	TT	GG	GG

ID	*DAPK1* *Tas*I	*ALAD* *Msp*I	*ALAD* *Rsa*I	*EPB72* *Rsa*I	*CDK9* *EcoN*I	*FPGS* *BstN*I	*AK1* *Taq*I	*PPP2R4* *Hinf*1	*ASS* *Nla*III	*ABL1* *BstN*I	*ABL1* *Cfo*I	*ABL1* *Mae*II	*POMT1* *BstN*I	*POMT1* *Hpa*II	*RALGDS* *BsiE*I	*SURF2* *Alu*I	*C8G* *Dde*I
OT2	AA	GG	CC	AA	AA	CT	AA	TT	CC	CC	AA	nd	CC	GG	nd	AG	GG
OT3	AG	GG	CC	AG	AG	CT	AA	AT	CC	CC	AA	CC	CC	GG	CT	GG	GG
OT5	AA	GG	CT	AG	AA	CT	AA	TT	CC	CC	AA	nd	nd	nd	nd	AG	GT
OT9	AA	GG	CC	AG	AA	CT	AA	TT	nd	CC	AG	nd	nd	nd	nd	nd	TT
OT16	AG	GG	CT	AA	AG	TT	AC	TT	nd	CC	AA	nd	CC	GG	nd	nd	TT
OT17	GG	GG	CT	AG	AA	CT	AA	**AT***	CC	CC	AA	CC	nd	nd	nd	AG	TT
OT19	AA	GG	CC	AG	nd	CT	AC	AA	CT	CC	AG	CC	nd	nd	nd	AG	GT
OT20	AG	GG	TT	AA	GG	TT	AA	TT	CC	CG	AG	CT	nd	nd	nd	GG	GT
OT21	AG	GG	CC	AA	AA	CT	AA	AT	nd	CC	AA	CC	nd	nd	nd	GG	GT
OT22	AA	GG	CC	AA	GG	TT	AA	AT	nd	CC	AG	nd	nd	nd	nd	AG	nd
OT23	AA	GG	CT	AG	GG	TT	AA	AT	nd	CC	AA	nd	nd	nd	nd	AA	GT
OT25	nd	GG	CT	AG	GG	TT	nd	TT	nd	CG	AG	CT	nd	nd	nd	AG	TT
OT26	AG	GG	CT	AG	AG	CT	AA	TT	nd	CC	AA	nd	nd	nd	nd	GG	GG
OT27	AA	CC	TT	AA	AA	CT	AA	AA	nd	CC	AA	nd	nd	nd	nd	nd	GG
OT28	AG	GG	CT	AA	AG	TT	AA	TT	nd	CC	AA	nd	nd	nd	nd	nd	nd
OT29	AG	CG	TT	AG	AA	CT	AA	AA	nd	CC	AG	nd	nd	nd	nd	AG	GT
OT30	nd	CG	nd	AA	nd	nd	nd	nd	CT	CC	AA	CC	nd	nd	nd	AG	nd
OT31	nd	CG	nd	AA	nd	nd	nd	nd	TT	CC	AA	CC	nd	nd	nd	GG	nd
OT32	nd	GG	nd	AG	nd	nd	nd	nd	CT	CC	AA	CC	nd	nd	nd	GG	nd
OT33	nd	GG	nd	AA	nd	nd	nd	nd	CC	CC	AA	CC	nd	nd	nd	AG	nd
OT34	nd	GG	nd	AA	nd	nd	nd	nd	CC	CC	AA	CC	nd	nd	nd	AA	nd
OT35	nd	GG	nd	AA	nd	nd	nd	nd	CT	CC	AA	CC	nd	nd	nd	AG	nd
OT36	nd	GG	nd	AA	nd	nd	nd	nd	CC	CC	AA	CC	nd	nd	nd	AG	nd
OT37	nd	GG	nd	AA	nd	nd	nd	nd	TT	CC	AA	CC	nd	nd	nd	AG	nd
OT38	nd	GG	nd	AA	nd	nd	nd	nd	CT	CC	AA	CC	nd	nd	nd	GG	nd
OT39	nd	GG	nd	AA	nd	nd	nd	nd	TT	CC	AA	CC	nd	nd	nd	GG	nd
OT40	nd	GG	nd	AG	nd	nd	nd	nd	TT	CC	AG	CC	nd	nd	nd	AG	nd
OT41	nd	GG	nd	AA	nd	nd	nd	nd	CC	nd	AG	CC	nd	nd	nd	AG	nd
OT42	nd	GG	nd	AA	nd	nd	nd	nd	CC	CC	GG	CC	nd	nd	nd	AA	nd
DOK1	AG	GG	CT	AA	AG	CT	AA	TT	nd	CC	AA	nd	nd	nd	nd	AG	GT
DOK2	AG	GG	CC	AG	GG	TT	AA	AT	nd	CC	AG	nd	nd	nd	CT	AG	GG
DOK3	AA	GG	CT	AA	AG	CT	AA	TT	CC	CG	nd	nd	nd	nd	nd	nd	nd
DOK4	AA	GG	CC	nd	GG	TT	AA	AT	nd	CC	nd	nd	nd	nd	nd	nd	GT
DOK5	AA	GG	CC	AA	GG	TT	AA	TT	nd	CG	AG	CT	nd	nd	nd	GG	GT
DOK6	AG	CG	CT	AA	AA	CT	AA	TT	nd	CC	AG	CC	nd	nd	nd	AG	GG
DOK7	AG	CG	TT	AG	AG	CT	AA	TT	nd	CC	AG	nd	nd	nd	nd	nd	GG
DOK8	AG	GG	TT	AG	AG	CT	AC	AA	nd	CG	AG	nd	nd	nd	nd	nd	GT

Table 6.2. SNP genotypes for the 9q loci in patient samples.

The loci are shown above each column with the restriction enzyme used for PCR-RFLP genotyping. The genotype at the SNP for both alleles is shown. The loci are shown in order from centromere to telomere based on the December 2001 mapping information. Samples in bold with a star had changes in allele dosage. 'nd' is not done.

ID	*DAPK1 Tas* I	*ALAD Msp* I	*ALAD Rsa* I	*EPB72 Rsa* I	*CDK9 EcoN* I	*FPGS BstN* I	*AK1 Taq* I	*PPP2R4 Hinf* I	*ASS Nla* III	*ABL1 BstN* I	*ABL1 Cfo* I	*ABL1 Mae* II	*POMT1 BstN* I	*POMT1 Hpa* II	*RALGDS BsiE* I	*SURF2 Alu* I	*C8G Dde* I
F3	AA	GG	TT	AA	AG	CT	AA	AT	CC	CC	AA	nd	**CT***	GG	nd	GG	GT
F11	AG	CG	CT	AA	AG	TT	AA	**AT***	nd	CC	AA	nd	nd	nd	nd	GG	GG
F21	AA	GG	TT	AA	AG	CT	AC	TT	CC	CC	AA	nd	**CT***	GG	nd	AG	GG
F22	**AG***	GG	CC	AA	nd	CT	**AC***	nd	nd	CC	AA	CC	nd	nd	nd	GG	GG
F54	nd	GG	CT	AG	AA	nd	AA	nd	nd	CC	nd	nd	TT	**AG***	nd	nd	nd
F60	AG	GG	TT	GG	GG	TT	AA	TT	nd	CC	AA	nd	CC	**AG***	TT	GG	GG
OT17	GG	GG	CT	AG	AA	CT	AA	**AT***	CC	CC	AA	CC	nd	nd	nd	AG	TT

Table 6.3. A summary of the SNP genotypes for the 9q loci in patient samples with allele dosage alterations.
The locus is shown above each column with the restriction enzyme used for PCR-RFLP genotyping. The genotype at the SNP for both alleles is shown. The loci are shown in order from centromere to telomere based on the December 2001 mapping information. Samples in bold with a star had changes in allele dosage. 'nd' is not done.

6.2.2.1 9q34 Loci Expression in Normal Individuals and Leukaemic Cell Lines

All normal samples that were heterozygous showed bi-allelic expression (Table 6.4). Due to time constraints, not all the SNPs could be investigated in all the normal samples. However, a sufficient number of SNPs were chosen so that they were representative of the entire panel in terms of the regions of 9q that were assessed.

6.2.2.2 9q34 Loci Expression in Leukaemic Patients

The allelic expression of the various 9q loci using SNPs is shown in Table 6.5.

F11 is an AML patient that had loss of *A* expression (Table 4.7) and had mono-allelic expression of the loci *DAPK1* and *PPP2R4* (Figure 6.4). This patient had an allele dosage shift for *PPP2R4* but not for *DAPK1* (Table 6.3).

F15 was a CML patient with no ABH alterations but displays mono-allelic expression of *DAPK1*, *ALAD*, *CDK9*, *AK1*, *PPP2R4* and *SURF2* (Figure 6.4). This patient had mono-allelic expression for every locus tested though they showed no allele dosage shifts for any of these loci. Mono-allelic expression was not seen for *ABO* which lies in the midst of all these loci. There is no cytogenetic information for this patient but it is unlikely that they have a chromosome 9 deletion since this sample does not have any allele dosage shifts.

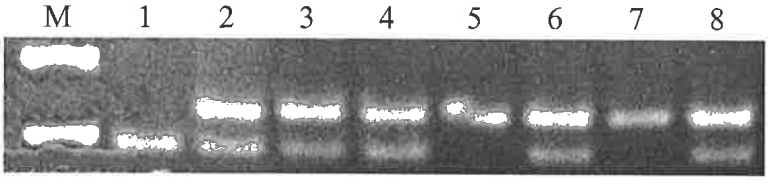


Figure 6.4. Loss of allelic expression in *PPP2R4*.

M refers to the marker pUC19/*Hpa*II and lane 1 is N32, a homozygote for *PPP2R4*. Lanes 2-4 are heterozygous *PPP2R4* samples Raji, OT3 and DOK2, respectively with bi-allelic expression seen by two bands after digestion with *Hinf*I. Lanes 5-8 are F11, OT22, F15 and F17, respectively, and they are all heterozygous for *PPP2R4*. F11 had an allele dosage shift for this locus and after digestion with *Hinf*I, F11 and F15 only had one band indicative of mono-allelic expression (lanes 5 and 7).

F23 is a CML patient with no ABH alterations but displays mono-allelic expression of *AK1*. F24 is an AML patient with loss of B and mono-allelic expression of *C8G*, which is telomeric to *ABO*. This patient was not informative for any of the other 10 loci tested.

F27 is an AML patient with loss of A and H and has mono-allelic expression of *CDK9* which is centromeric to *ABO*. This patient had bi-allelic expression for *ALAD* and *C8G*, which lie on either side of *CDK9* and *ABO* (Figure 6.2). This may suggest that a tumour suppressor gene lies in between this region, 9q34.11-q34.1 and *POMT1* and *PPP2R4,* which showed the most allele dosage shifts, lie in this region and may be potential tumour suppressor genes. This patient was not informative for these loci.

F30 is a CML patient with loss of H and does not express *DAPK1*. This patient does show bi-allelic expression for 3 other loci and positive expression for another 5 loci however, *DAPK1* is extremely distant from all the other loci. F42 is an AML patient and F44 is a MDS patient with no ABH alterations but they do not express *DAPK1*. They do express all the other 9q loci. F46 is an AML patient with loss of A and negative expression for *DAPK1*. Again, this patient expressed all other 9q loci. It seems that *DAPK1* may be targeted for silencing in these patients or that it may reside adjacent to the gene that is the target, but is itself silenced in a global chromosomal tide.

F50 is an AML patient with loss of H and mono-allelic expression of *C8G* while patient F51 is an AML patient with loss of A and mono-allelic expression of *DAPK1*. F50 and F51 express other 9q loci, though F50 was only tested for one other locus. F60 is a CML patient with loss of A and the one locus it was informative for, *DAPK1*, shows mono-allelic expression.

OT5 is an ALL patient with mono-allelic expression of C8G while OT16 is a CLL patient with loss of A and negative expression for *DAPK1*, *EPB72* and *PPP2R4*. There was no cytogenetic information available for this patient, however, they had positive 9q expression for all the other 9q loci. OT20 is a MDS patient with mono-allelic expression of *ABL1* and no expression for *DAPK1* and *PPP2R4* but positive expression for the remaining 9q loci. OT22 is a CML patient with mono-allelic expression of *ABL1* but not for other loci while OT26 is negative for *DAPK1* but positive for all other loci. OT29 had mono-allelic expression of *ABL1* and negative expression for *DAPK1* but expresses all other loci accordingly. DOK1 and DOK8 are both loss of A patients, with mono-allelic expression of *C8G* and *ABL1* respectively.

All loci had at least one sample with an alteration and *AK1* and *CDK9* had 2 samples with alterations. *ABL1* had 4/32 and *PPP2R4* 4/30 samples with alterations while *C8G* had 5/49 samples with expression changes (Figure 6.4). The most frequently affected locus was *DAPK1* with 13/31 (42%) samples showing expression alterations (summarised in Table 6.6).

Of all the patients assessed, 21/54 had expression alterations. If O^1O^1 patients are excluded, since loss of *O* can not be assessed by flow cytometry and loss at the mRNA level would not be seen due to the presence of two O^1alleles, 8/18 patients with 9q loci expression alterations also had ABH antigen or *ABO* mRNA alterations. Patients with loss of the H antigen by flow cytometry were included as loss of H can mask loss of A or B antigen expression. This means that nearly 50% of patients with *ABO* alterations have other 9q loci expression alterations supporting the notion that this chromosomal region may undergo regional silencing.

ID	SAMPLE	*DAPK*	*ALAD*	*EPB72*	*CDK9*	*AK1*	*PPP2R4*	*ABL1*	*SURF2*	*C8G*
N7	PB MNC	nd	+	nd	nd	nd	nd	nd	nd	bi
N10	PB MNC	nd	+	nd	nd	nd	nd	nd	nd	+
N18	PB MNC	nd	+	nd	nd	nd	nd	nd	nd	nd
N19	PB MNC	nd	bi	nd	nd	nd	nd	nd	nd	bi
N20	PB MNC	nd	bi	bi	nd	nd	nd	nd	nd	nd
N22	PB MNC	nd	bi	nd	nd	nd	nd	nd	nd	nd
N23	PB MNC	nd	+	nd	nd	nd	nd	nd	nd	bi
N24	PB MNC	nd	bi	nd	nd	nd	nd	nd	nd	bi
N25	PB MNC	nd	+	nd	nd	nd	nd	nd	nd	+
N26	PBSC	nd	bi	bi	nd	nd	nd	nd	nd	+
N27	PBSC	nd	+	nd	nd	nd	nd	nd	nd	+
N28	PBSC	nd	+	nd	nd	nd	nd	nd	nd	+
N29	PBSC	nd	nd	nd	nd	nd	nd	nd	nd	bi
N30	PBSC	nd	bi	nd	nd	nd	nd	nd	nd	bi
N31	PBSC	+	bi	bi	bi	+	bi	+	nd	+
N32	PBSC	+	nd	+	+	+	+	+	bi	nd
KCL22	cell line	nd	bi	-	nd	nd	nd	nd	nd	nd
SW48	cell line	nd	bi	bi	nd	nd	nd	nd	nd	bi
LIM1215	cell line	nd	+	nd	nd	nd	nd	nd	nd	+
HEL	cell line	+	+	-	+	+	nd	+	+	+
CCRF-CEM	cell line	nd	+	nd	nd	nd	nd	nd	nd	-
EM2	cell line	-	+	+	+	+	nd	+	+	+
RAJI	cell line	-	bi	bi	+	+	bi	+	+	-
Jurkat	cell line	nd	bi	nd	nd	nd	nd	nd	nd	+
BM1	BM	+	bi	+	+	+	bi	bi	bi	+
BM3	BM	bi	bi	+	bi	+	bi	bi	bi	+

Table 6.4. Allelic expression using SNPs for various 9q loci in normal samples and leukaemic cell lines.
The loci are shown in order from centromere to telomere. 'bi' refers to the presence of both alleles after restriction enzyme digestion of the PCR product amplified from cDNA, that is, bi-allelic expression. '+' refers to the locus being expressed in that sample but allelic expression could not be determined since it was not informative at that SNP. '-' means that sample did not express that gene as no PCR product was generated from the cDNA. 'nd' means not done.

ID	*DAPK*	*ALAD*	*EPB72*	*CDK9*	*AK1*	*PPP2R4*	*ABL1*	*SURF2*	*C8G*
F2	nd	+	nd	nd	nd	nd	nd	nd	bi
F3	+	+	+	bi	+	bi	+	+	bi
F4	nd	bi	nd	nd	nd	nd	nd	nd	+
F7	+	+	bi	bi	+	+	+	+	+
F11	mono	bi	+	bi	+	mono	+	+	+
F12	nd	+	nd	nd	nd	nd	nd	nd	nd
F13	nd	bi	nd	nd	nd	nd	nd	nd	bi
F14	nd	bi	nd	nd	nd	nd	nd	nd	+
F15	mono	mono	+	mono	mono	mono	+	mono	+
F17	bi	+	nd	bi	bi	bi	+	bi	+
F19	nd	+	nd	nd	nd	nd	nd	nd	+
F20	+	nd	+	bi	+	+	bi	+	+
F21	+	+	+	bi	bi	+	+	bi	+
F22	nd	+	nd	nd	nd	nd	nd	nd	+
F23	+	nd	bi	bi	mono	nd	+	bi	+
F24	+	+	+	+	+	+	+	+	mono
F27	+	bi	+	mono	+	+	+	+	bi
F30	-	+	bi	+	+	+	+	bi	bi
F31	nd	+	nd	nd	nd	nd	nd	nd	+
F32	nd	bi	nd	nd	nd	nd	nd	nd	bi
F37	nd	bi	nd	nd	nd	nd	nd	nd	+
F38	+	+	bi	bi	+	+	+	bi	bi
F39	+	bi	+	+	+	+	bi	+	bi
F42	-	nd	+	+	+	+	+	bi	+
F43	nd	+	nd	nd	nd	nd	nd	nd	+
F44	-	bi	bi	bi	+	+	+	bi	bi
F46	-	nd	+	bi	+	+	+	+	nd
F48	nd	bi	nd	nd	nd	nd	nd	nd	+
F50	nd	bi	nd	nd	nd	nd	nd	nd	mono
F51	mono	bi	bi	bi	+	+	+	+	+
F52	-	nd	bi	+	bi	+	+	bi	-
F53	nd	nd	nd	nd	+	nd	+	nd	bi
F59	+	bi	bi	bi	bi	+	+	+	+
F60	mono	+	+	+	+	+	+	+	+
OT2	+	+	+	+	+	+	+	bi	+
OT3	bi	+	bi	bi	+	bi	+	+	+
OT5	nd	bi	nd	nd	nd	nd	nd	nd	mono
OT9	nd	+	nd	nd	nd	nd	nd	nd	+
OT16	-	bi	-	bi	bi	-	+	+	+
OT20	-	+	+	+	+	-	mono	+	+
OT21	nd	+	nd	nd	nd	nd	nd	nd	bi
OT22	+	+	+	+	+	bi	mono	bi	+
OT23	nd	bi	nd	nd	nd	nd	nd	nd	+
OT26	-	bi	bi	bi	+	+	+	+	+
OT27	nd	+	nd	nd	nd	nd	nd	nd	+
OT28	nd	+	nd	nd	nd	nd	nd	nd	nd
OT29	-	bi	+	+	+	+	mono	bi	bi
DOK1	bi	bi	+	bi	+	+	+	bi	mono
DOK2	bi	+	bi	+	+	bi	bi	bi	+
DOK3	nd	+	nd	nd	nd	nd	nd	nd	nd
DOK5	nd	+	nd	nd	nd	nd	nd	nd	+
DOK6	bi	bi	+	+	+	+	bi	bi	+
DOK7	nd	bi	nd	nd	nd	nd	nd	nd	nd
DOK8	bi	nd	bi	bi	bi	+	mono	+	bi

Table 6.5. Allelic expression using SNPs for various 9q loci.
The loci are shown in order from centromere to telomere. 'bi' refers to the presence of both alleles after restriction enzyme digestion of the PCR product amplified from cDNA, that is, bi-allelic expression. 'mono' refers to mono-allelic expression at that locus. '+' refers to that locus being expressed in that sample but allelic expression could not be determined as it was not informative at that SNP. '-' refers to that sample not expressing the locus in question since no PCR product was generated from the cDNA of that sample. 'nd' is not done.

ID	*DAPK*	*ALAD*	*EPB72*	*CDK9*	*AK1*	*PPP2R4*	*ABL1*	*SURF2*	*C8G*
F11	mono	bi	+	bi	+	mono	+	+	+
F15	mono	mono	+	mono	mono	mono	+	mono	+
F23	+	nd	bi	bi	mono	nd	+	bi	+
F24	+	+	+	+	+	+	+	+	mono
F27	+	bi	+	mono	+	+	+	+	bi
F30	-	+	bi	+	+	+	+	bi	bi
F42	-	nd	+	+	+	+	+	bi	+
F44	-	bi	bi	bi	+	+	+	bi	bi
F46	-	nd	+	bi	+	+	+	+	nd
F50	nd	bi	nd	nd	nd	nd	nd	nd	mono
F51	mono	bi	bi	bi	+	+	+	+	+
F52	-	nd	bi	+	bi	+	+	bi	-
F60	mono	+	+	+	+	+	+	+	+
OT5	nd	bi	nd	nd	nd	nd	nd	nd	mono
OT16	-	bi	-	bi	bi	-	+	+	+
OT20	-	+	+	+	+	-	mono	+	+
OT22	+	+	+	+	+	bi	mono	bi	+
OT26	-	bi	bi	bi	+	+	+	+	+
OT29	-	bi	+	+	+	+	mono	bi	bi
DOK1	bi	bi	+	bi	+	+	+	bi	mono
DOK8	bi	nd	bi	bi	bi	+	mono	+	bi

Table 6.6. A summary of allelic expression changes at 9q loci in patients with haematological malignancies.

The loci are shown in order from centromere to telomere. 'bi' refers to the presence of both alleles after restriction enzyme digestion of the PCR product amplified from cDNA, that is, bi-allelic expression. 'mono' refers to mono-allelic expression at that locus. '+' indicates that locus is expressed in that sample but allelic expression could not be determined due to homozygosity. '-' refers to that sample not expressing the locus in question since no PCR product was generated from the cDNA of that sample. 'nd' is not done.

6.3 Discussion

Allele dosage shifts at the 9q loci were not common but predominantly occurred at *POMT1* and *PPP2R4*, which are both centromeric to *ABO* (Figure 6.2). For patients which had available cytogenetic information (Table 3.2), chromosome 9 was not affected except for CML patients with the BCR-ABL translocation.

POMT1 is involved in an essential protein modification process and is the first of its kind to be found in mammals (Jurado *et al*, 1999). Disruptions in the *Drosophila melanogaster* homologue of *POMT1* causes defects in embryonic muscle development, hence this gene may be a candidate for genetic disorders of the muscular system, though its function has not yet been characterised (Jurado *et al*, 1999).

PPP2R4 is highly homologous to genes in *Drosophila melanogaster*, yeast and *Xenopus laevis* and shares 96.6% identify with rabbit *PPP2R4* (Janssens *et al*, 2000). All the homologues of *PPP2R4* suggest it has a well conserved function and experiments with yeast homologues of *PPP2R4* have shown that its function is required for proper G1 progression (Mitchell and Sprague, 2001). *PPP2R4* is definitely a candidate for a potential tumour suppressor gene involved in haematological malignancies, and whose inactivation may also disrupt *ABO*. Further characterisation of *PPP2R4* is required and the expression analysis of *PPP2R4* would need to be performed in the remaining patients and other patients with haematological malignancies.

Though changes in allele dosage were not common, there were many alterations when allelic expression of these loci were examined. For patients with altered expression at more than one 9q locus, they spanned the whole chromosome. Since there were alterations spanning chromosome 9, yet no deletion of this chromosome arm was evident, it is possible that regions of 9q are targeted for global silencing by some epigenetic mechanism, like methylation or acetylation, which silences chromosomal regions in X-inactivation and imprinting.

The other possibility is that the situation in myeloid leukaemias is similar to bladder cancer, which has 2 regions on 9q harbouring potential tumour suppressor genes, one at the proximal end of 9q and one at the distal end. *DAPK1* is at the proximal end of 9q and is affected in 42% of samples tested hence it may be a potential tumour suppressor gene. This gene only showed an allele shift in one sample but had expression changes in 42% of patients tested. It is known that *DAPK1* is silenced by methylation in myeloid leukaemias (Aggerholm and Hokland, 2000) and thus it is possible that this gene is silenced by this mechanism or other mechanisms in myeloid leukaemias.

Another observation is that 21 patients had changes in expression of one or more 9q loci and 12 of these patients had ABH alterations by flow cytometry or RT-PCR, 10/21 (48%) had loss of A or B or if looking only at myeloid malignancies, 9/17 (53%) had loss of A and B and alterations at other 9q loci. Given that loss of the *O* allele can not be assessed by flow cytometry or RT-PCR, and that approximately 40% of patients with haematological malignancies will be $O^{l}O^{l}$, at least 50% of patients with alterations at *ABO*, have changes at other 9q loci. This suggests that these *ABO* alterations may be a marker for changes occurring at potential tumour suppressor gene(s) on 9q or that *ABO* is inactivated by silencing of the 9q chromosomal region by epigenetic mechanisms.

LOH of 9q has also recently been identified in childhood de novo AML implying that there may be the same gene targeted in adult AML cases (Sweetser *et al*, 2001). *DAPK1* is a likely candidate as it has shown the most frequent alterations in this cohort of patients and because it is a well-characterised tumour suppressor gene involved in apoptosis (Guzman *et al*, 2001; Velentza *et al*, 2001).

Though rates of LOH were low in patient samples, the common finding of silencing of 9q loci may indicate that chromosome 9q loci undergo regional silencing, such as in X-inactivation. This could readily be investigated in *DAPK1*, which is known to undergo methylation, and in *PPP2R4* which has a CpG island associated with its promoter.

CHAPTER 7

Conclusion & Future Directions

7.1 Concluding Discussion

The first aim of this PhD thesis was to develop a quantitative assay, using flow cytometry, to assess the frequency and amount of ABH antigen change on red blood cells from patients with leukaemia (Chapter 3). In the development of this technique, it was determined that quantitative changes in flow profiles reflected different *ABO* genotypes. In total, 22/57 patients had alterations in ABH antigen expression. Loss of H in the O patients was seen for 6/28 patients and loss of H in the A/B patients was seen in 8/29 samples. This data combined gave an overall frequency of loss of the H antigen of 25%. Loss of A or B was seen in 11/29 (38%) of the A/B patients. Fifty-five percent (16/29) of A/B patients showed alterations at one or both of the loci. Changes in ABH antigens did not correlate with age, survival, disease predisposition, AML subclass or cytogenetics. In conclusion, the systematic investigation of alterations in ABH antigens elucidated that it occurred with relatively high frequency in patients with haematological malignancies.

The second aim of this PhD was to determine the molecular mechanisms behind the frequent observation of alterations of ABH antigens in patients with haematological malignancies (Chapters 4 and 5). LOH and promoter methylation at the *ABO* locus were investigated, but did not account for ABH antigen loss in a high percentage of cases. *ABO* was silenced by methylation in some leukaemic cell lines that re-expressed *ABO* after treatment with the demethylating agent 5-aza-2'-deoxycytidine. Some cell lines did not re-express even after treatment with a demethylating agent so other mechanisms like acetylation or mutation may be responsible for the loss of ABH antigens in patients with haematological malignancies.

Some patients with alterations in ABH antigens detected by flow cytometric analysis did not have a corresponding alteration at the mRNA level. A better method to assess for ABH changes, instead of assessing mRNA levels, may have been to assess the levels of the various transferases in the individuals. This may have correlated better with the flow profiles.

Loss of *ABO* expression at the mRNA level correlated with loss of A or B antigens when assessed by serology (Denise O'Keefe, PhD thesis, 1995). Loss of A or B antigens by serology requires alterations in over 50% of the cells and this explains why loss of *A* at the mRNA level is seen in these patients. The abnormal population of cells detected by flow cytometry often comprised between 10-20% of the whole population. Loss of *ABO* at the mRNA level in these cells would be masked by the *ABO* expression of the normal cells.

Another problem encountered when assessing for *ABO* expression at the mRNA level was that PB MNC and PBSC were not adequate controls. It was found that *AO* PB MNC and PBSC samples failed to express the *A* allele while BM *AO* samples expressed the *A* allele. *BO* PB MNC samples had an identical expression pattern to *BO* BM specimens. The difference in expression patterns of the *AO* and *BO* PB MNCs may reflect the different enhancers associated with the two promoters of these alleles. It is known that the *B* allele enhancer is 300-fold more efficient at promoting transcription than the *A* allele enhancer.

Alignment of *ABO* ests for the *A* and *B* alleles, which can be determined by using the previously reported sequence variation between the alleles in the 3' UTR (Olsson and Chester, 1995), suggested that the *A* 3' UTR is longer than the *B* 3' UTR. It has been reported that longer 3' UTRs are degraded more rapidly (Canete-Soler *et al*, 1998; Borrmann *et al*, 2001). This may explain the lack of *A* expression in PB MNC and PBSC samples, as the *A* allele mRNA is degraded more rapidly than the *B* allele mRNA. The increased number of cells expressing *ABO* in BM compared to PB MNC and PBSC may account for the absence of detectable *A* allele expression in PB MNC and PBSC.

However, this does not explain why the O^1 allele is expressed since it also has a longer 3' UTR than the *B* allele. Expression of the O^1 allele may be due to its enhancer which is the same as the *B* allele, and 300-fold more active than the A^1 enhancer (Yu *et al*, 2000). The lack of expression of the A^2 allele in PB MNC and PBSC but not in BM is also seen. The A^2 allele enhancer has four repeat units like the *B* and O^1 allele but it has not been fully characterised like the A^1, *B* and O^1 enhancers, hence it may contain nucleotide substitutions that render it less efficient than the *B* and O^1 enhancers (Irshaid *et al*, 1999; Yu *et al*, 2000).

The unsuitability of PB MNC and PBSC as controls for normal mRNA expression was more apparent when expression of the various alternative transcripts of *FUT1* was investigated. PB MNC samples do not express any of the *FUT1* alternative transcripts while PBSC samples express exon 7 *FUT1* transcripts and BM samples express exon 2 and exon 7 *FUT1* transcripts. This demonstrates that the population of cells within a BM sample have different expression patterns than those of PB MNC and PBSC samples. These results indicate that PB MNC and PBSC samples are not suitable controls for investigating alterations of *FUT1* or *ABO* mRNA expression in patient BM samples.

Expression of *FUT1* in patients with haematological malignancies indicates that the exon 1 transcript does not play a role in the loss of the H antigen as it is not expressed in any of the patients, except one, but is expressed in some leukaemic cell lines. The CpG island for this exon was also methylated in normal samples and hence methylation may be responsible for the silencing of this transcript in the patient samples.

The exon 7 *FUT1* transcript was expressed in all patients except four. One of these, F39, had loss of H antigen by flow cytometry and another, DOK2, had loss of A by serology. The other two patients, OT3 and OT27 showed no alterations in ABH antigens. The CpG island associated with this exon was unmethylated in normal samples but methylated in leukaemic cell lines which did not express this transcript. Methylation may be responsible for silencing of this transcript but the patients that did not express require further investigation.

The exon 2 *FUT1* transcript showed the greatest variation in expression amongst patients and seemed therefore, to be the most relevant to investigate. Seven loss of H patients, as determined by flow cytometry, were assessed for this transcript and 5 were negative. This strengthens the hypothesis that this transcript may be responsible for the loss of H antigen seen on the flow cytometer. Analysis of the methylation status of the CpG island associated with exon 2 of *FUT1* was complicated by the fact that normal PB MNC and PBSC samples showed extensive methylation at this island, though BM samples could not be assessed due to lack of material. However, sequencing may suggest that methylation of a CpG site, which resides in the recognition sequence for the transcription factors Sp1 and AP-2, may be associated with expression. More sequencing would be required. It is known that methylation of CpGs within AP-2 sites abolishes the ability for AP-2 to bind to its recognition sequence (reviewed in Singal and Ginder, 1999).

The third aim of this research project was to study genes on 9q, where the *ABO* locus is located, and to use single nucleotide polymorphisms (SNPs) to analyse these genes for LOH and mono-allelic expression as an indicator for tumour suppressor gene function (Chapter 6). Markers were chosen on either side of *ABO* and though the localisation of most of the loci changed after they were chosen (Figure 6.2) this was advantageous as more loci were telomeric to *ABO*. Six loci were analysed centromeric to *ABO* and two telomeric to *ABO*.

POMT1 and *PPP2R4* are both potential 9q tumour suppressor loci that have not been extensively characterised. Due to time constraints, investigation of the expression of *POMT1* was not possible and *PPP2R4* and *DAPK1* expression analysis needs to be performed on the remaining patients, though this is limited by availability of patient material. *POMT1* and *PPP2R4* are highly conserved genes with important roles in protein modification and the cell cycle, respectively, so their importance requires further investigation.

There were 6/8 allele dosage shifts at *POMT1* and *PPP2R4* (Table 6.3), which are relatively close to each other on chromosome 9q34. These loci are possible tumour suppressor genes that are not too distant from *ABO* and are separated by *ABL1* and *NUP214* which are involved in translocations in myeloid leukaemias. This region may be targeted for inactivation in myeloid malignancies.

DAPK1 had altered expression in 42% of patients, clearly making this gene of great interest. This locus did not show LOH as most tumour suppressor genes do but was not expressed or only mono-allelically expressed in many samples, suggesting its inactivation by other mechanisms. This gene has been shown to be methylated in haematological malignancies thus making epigenetic silencing a possible explanation for the silencing of this gene (Katzenellenbogen *et al*, 1999; Aggerholm and Hokland, 2000). This gene is a well characterised tumour suppressor gene involved in apoptosis (Guzman *et al*, 2001; Velentza *et al*, 2001).

In conclusion, alterations in ABH antigens are a common occurrence in haematological malignancies but alterations at the *ABO* locus are not solely responsible. Changes at the *FUT1* locus may account for some loss of H observations. Changes in ABH antigens appear to be a marker for other silencing events targeting potential tumour suppressor gene(s) at 9q. The lack of numerous allele dosage alterations and the frequent mono-allelic expression indicates that 9q may be targeted for regional silencing by epigenetic mechanisms.

7.2 Future Directions

The continuing assessment of loss of ABH alterations in haematological malignancies seems to separate a group of malignancies of which 50% have alterations at other 9q loci. Further assessment of other 9q loci is required, especially genes known to be altered in leukaemia, like *ABL1* and *NUP214*. The situation in haematological malignancies may parallel bladder cancer which has two 9q regions harbouring potential tumour suppressor genes (Simoneau *et al*, 1999).

LOH did not seem to be the mechanism for the inactivation of the potential tumour suppressor gene, though the appropriate controls were not used in this study. It may be more appropriate to use patient T-cells or buccal cells as normal controls. Further analysis of the other 9q loci, for which assays were developed but time constraints prevented performing them, will also help to clarify the role of LOH.

DAPK1 which in this study was found to be altered in 42% of patient samples, is known to be methylated. *ABL1* on 9q34.12 is also methylated in CMLs. There may be global methylation of this region, 9q21-9q34, spanning *DAPK1* as the minimal region, and this may silence potential tumour suppressor genes that are important in the pathogenesis of leukaemia, especially those of the myeloid lineage. Further assessment of 9q loci for methylation is necessary to determine if this is true.

The subset of patients with loss of H is a group of patients that have *FUT1* targeted for silencing. Again, methylation appears important but bears further investigation. A larger sample size of normal BM specimens needs to be analysed for methylation, possibly by sequencing which may help determine which CpG sites are essential for expression or silencing by methylation.

In conclusion, it has been shown that the loss of ABH antigens in patients with haematological malignancies is a signpost to the disruption of gene expression on the long arm of chromosome 9. *DAPK1*, *POMT1* and *PPP2R4*, have been identified as potential tumour suppressor genes. Further studies will elucidate the importance of these genes in haematological malignancy though the lack of allele dosage changes and the frequent loss of allelic expression indicates that the 9q chromosomal region may be silenced by epigenetic mechanisms.

References

Aggerholm, A., Guldberg, P., Hokland, M. & Hokland, P. (1999). Extensive intra- and interindividual heterogeneity of p15INK4B methylation in acute myeloid leukemia. *Cancer Res*, **59,** 436-41.

Aggerholm, A. & Hokland, P. (2000). DAP-kinase CpG island methylation in acute myeloid leukemia: methodology versus biology? *Blood*, **95,** 2997-9.

Amado, M., Bennett, E.P., Carneiro, F. & Clausen, H. (2000). Characterization of the histo-blood group O(2) gene and its protein product. *Vox Sang*, **79,** 219-26.

Antequera, F., Boyes, J. & Bird, A. (1990). High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell*, **62,** 503-14.

Archer, S.Y. & Hodin, R.A. (1999). Histone acetylation and cancer. *Curr Opin Genet Dev*, **9,** 171-4.

Asimakopoulos, F.A., Shteper, P.J., Krichevsky, S., Fibach, E., Polliack, A., Rachmilewitz, E., Ben-Neriah, Y. & Ben-Yehuda, D. (1999). ABL1 methylation is a distinct molecular event associated with clonal evolution of chronic myeloid leukemia. *Blood*, **94,** 2452-60.

Aubert, M., Panicot, L., Crotte, C., Gibier, P., Lombardo, D., Sadoulet, M.O. & Mas, E. (2000). Restoration of alpha(1,2) fucosyltransferase activity decreases adhesive and metastatic properties of human pancreatic cancer cells. *Cancer Res*, **60,** 1449-56.

Barjas-Castro, M.L., Carvalho, M.H., Locatelli, M.F., Bordin, S. & Saad, S.T. (2000). Molecular heterogeneity of the A3 subgroup. *Clin Lab Haematol*, **22,** 73-8.

Baumer, A., Wiedemann, U., Hergersberg, M. & Schinzel, A. (2001). A novel MSP/DHPLC method for the investigation of the methylation status of imprinted genes enables the molecular detection of low cell mosaicisms. *Hum Mutat*, **17,** 423-30.

Baylin, S.B., Fearon, E.R., Vogelstein, B., de Bustros, A., Sharkis, S.J., Burke, P.J., Staal, S.P. & Nelkin, B.D. (1987). Hypermethylation of the 5' region of the calcitonin gene is a property of human lymphoid and acute myeloid malignancies. *Blood*, **70,** 412-7.

Baylin, S.B., Herman, J.G., Graff, J.R., Vertino, P.M. & Issa, J.P. (1998). Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res*, **72,** 141-96.

Bennett, E.P., Steffensen, R., Clausen, H., Weghuis, D.O. & Geurts van Kessel, A. (1995). Genomic cloning of the human histo-blood group ABO locus. *Biochem Biophys Res Commun*, **211,** 347.

Berneman, Z.N., van Bockstaele, D.R., Uyttenbroeck, W.M., Van Zaelen, C., Cole-Dergent, J., Muylle, L. & Peetermans, M.E. (1991). Flow-cytometric analysis of erythrocytic blood group A antigen density profile. *Vox Sang*, **61,** 265-74.

Bestor, T.H. & Verdine, G.L. (1994). DNA methyltransferases. *Curr Opin Cell Biol*, **6,** 380-9.

Bhende, Y.M., Deshpande, C.K., Bhatia, H.M., Sanger, R., Race, R.R., Morgan, W.T.J. & Watkins, W.M. (1952). A 'new' blood group character related to the ABO system. *Lancet*, **1,** 903-4.

Bianco, T., Chenevix-Trench, G., Walsh, D.C., Cooper, J.E. & Dobrovic, A. (2000). Tumour-specific distribution of BRCA1 promoter region methylation supports a pathogenetic role in breast and ovarian cancer. *Carcinogenesis*, **21,** 147-51.

Bianco, T., Hussey, D. & Dobrovic, A. (1999). Methylation-sensitive, single-strand conformation analysis (MS-SSCA): A rapid method to screen for and analyze methylation. *Hum Mutat*, **14,** 289-93.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, **321,** 209-13.

Borrmann, L., Wilkening, S. & Bullerdiek, J. (2001). The expression of HMGA genes is regulated by their 3'UTR. *Oncogene*, **20,** 4537-41.

Boyd, J.A. & Barrett, J.C. (1990a). Genetic and cellular basis of multistep carcinogenesis. *Pharmacol Ther*, **46,** 469-86.

Boyd, J.A. & Barrett, J.C. (1990b). Tumor suppressor genes: possible functions in the negative regulation of cell proliferation. *Mol Carcinog*, **3,** 325-9.

Bryne, M., Thrane, P.S. & Dabelsteen, E. (1990). Loss of expression of blood group antigen H is associated with cellular invasion and spread of oral squamous cell carcinomas. *Cancer*, **66,** 2118-23.

Bryne, M., Thrane, P.S. & Dabelsteen, E. (1991). Loss of expression of blood group antigen H is associated with cellular invasion and spread of oral squamous cell carcinomas. *Cancer*, **67,** 613-8.

Burri, N. & Chaubert, P. (1999). Complex methylation patterns analyzed by single-strand conformation polymorphism. *Biotechniques*, **26,** 232-4.

Buzin, C.H., Wen, C.Y., Nguyen, V.Q., Nozari, G., Mengos, A., Li, X., Chen, J.S., Liu, Q., Gatti, R.A., Fujimura, F.K. & Sommer, S.S. (2000). Scanning by DOVAM-S detects all unique sequence changes in blinded analyses: evidence that the scanning conditions are generic. *Biotechniques*, **28,** 746-50, 752-3.

Canete-Soler, R., Schwartz, M.L., Hua, Y. & Schlaepfer, W.W. (1998). Stability determinants are localized to the 3'-untranslated region and 3'-coding region of the neurofilament light subunit mRNA using a tetracycline-inducible promoter. *J Biol Chem*, **273,** 12650-4.

Cartron, J.P., Reyes, F., Gourdin, M.F., Garretta, M. & Salmon, C.H. (1977). Antigen site distribution among 'weak A' red cell populations. A study of A3, Ax and Aend variants. *Immunology*, **32,** 233-44.

Castro-Malaspina, H. & O'Reilly, R.J. (1998). Aplastic anemia and myelodysplastic syndromes. In *Harrison's principles of internal medicine*, Fauci, A.S. (ed), Vol. 1. pp. 672-9. McGraw-Hill: New York.

Cedar, H. (1988). DNA methylation and gene activity. *Cell*, **53,** 3-4.

Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L. & Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. *Nature*, **395,** 89-93.

Clark, S.J., Harrison, J., Paul, C.L. & Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*, **22,** 2990-7.

Clausen, H., Hakomori, S., Graem, N. & Dabelsteen, E. (1986). Incompatible A antigen expressed in tumors of blood group O individuals: immunochemical, immunohistologic, and enzymatic characterization. *J Immunol*, **136,** 326-30.

Clausen, H., White, T., Takio, K., Titani, K., Stroud, M., Holmes, E., Karkov, J., Thim, L. & Hakomori, S. (1990). Isolation to homogeneity and partial characterization of a histo-blood group A defined Fuc alpha 1----2Gal alpha 1----3-N- acetylgalactosaminyltransferase from human lung tissue. *J Biol Chem*, **265,** 1139-45.

Coon, J.S., Watkins, J.R., Pauli, B.U. & Weinstein, R.S. (1985). Flow cytometric analysis of heterogeneity in blood group-related antigen expression in a human urinary bladder carcinoma cell line, 647V. *Cancer Res*, **45,** 3014-21.

Costello, J.F., Berger, M.S., Huang, H.S. & Cavenee, W.K. (1996). Silencing of p16/CDKN2 expression in human gliomas by methylation and chromatin condensation. *Cancer Res*, **56,** 2405-10.

Coulondre, C., Miller, J.H., Farabaugh, P.J. & Gilbert, W. (1978). Molecular basis of base substitution hotspots in Escherichia coli. *Nature*, **274,** 775-80.

Cross, S.H. & Bird, A.P. (1995). CpG islands and genes. *Curr Opin Genet Dev*, **5,** 309-14.

Dabelsteen, E., Clausen, H., Holmstrup, P. & Reibel, J. (1988). Premalignant and malignant oral lesions are associated with changes in the glycosylation pattern of carbohydrates related to ABH blood group antigens. *Apmis*, **96,** 813-9.

Daniels, G. (1995). *Human Blood Groups*. Blackwell Scientific Publications: Oxford.

David, L., Leitao, D., Sobrinho-Simoes, M., Bennett, E.P., White, T., Mandel, U., Dabelsteen, E. & Clausen, H. (1993). Biosynthetic basis of incompatible histo-blood group A antigen expression: anti-A transferase antibodies reactive with gastric cancer tissue of type O individuals. *Cancer Res*, **53,** 5494-500.

Davidsohn, I. & Stejskal, R. (1972). Tissue antigens A, B and H in health and disease. *Haematologia*, **6,** 177-84.

Dobrovic, A., O'Keefe, D., Sage, R.E. & Batchelder, E. (1993). Imprinting and loss of ABO antigens in leukemia. *Blood*, **82,** 1684-5.

Dobrovic, A. & Simpfendorfer, D. (1997). Methylation of the BRCA1 gene in sporadic breast cancer. *Cancer Res*, **57,** 3347-50.

Economidou, J., Hughes-Jones, N.C. & Gardner, B. (1967). Quantitative measurements concerning A and B antigen sites. *Vox Sang*, **12,** 321-8.

Endo, T. (1999). O-mannosyl glycans in mammals. *Biochim Biophys Acta*, **1473,** 237-46.

Ernst, C., Thurin, J., Atkinson, B., Wurzel, H., Herlyn, M., Stromberg, N., Civin, C. & Koprowski, H. (1984). Monoclonal antibody localization of A and B isoantigens in normal and malignant fixed human tissues. *Am J Pathol*, **117,** 451-61.

Esteller, M. & Herman, J.G. (2002). Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J Pathol*, **196,** 1-7.

Evdokiou, A., Webb, G.C., Peters, G.B., Dobrovic, A., O'Keefe, D.S., Forbes, I.J. & Cowled, P.A. (1993). Localization of the human growth arrest-specific gene (GAS1) to chromosome bands 9q21.3-q22, a region frequently deleted in myeloid malignancies. *Genomics*, **18,** 731-3.

Feinberg, A.P. & Vogelstein, B. (1983). Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun*, **111,** 47-54.

Fialkow, P.J., Singer, J.W., Adamson, J.W., Vaidya, K., Dow, L.W., Ochs, J. & Moohr, J.W. (1981). Acute nonlymphocytic leukemia: heterogeneity of stem cell origin. *Blood*, **57,** 1068-73.

Fibach, E. & Sharon, R. (1994). Changes in ABH antigen expression on red cells during in vivo aging: a flow cytometric analysis. *Transfusion*, **34,** 328-32.

Fiering, S., Whitelaw, E. & Martin, D.I. (2000). To be or not to be active: the stochastic nature of enhancer action. *Bioessays*, **22,** 381-7.

Freedman, A.S. & Nadler, L.M. (1998). Malignancies of lymphoid cells. In *Harrison's principles of internal medicine*, Fauci, A.S. (ed), Vol. 1. pp. 695-710. McGraw-Hill: New York.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. & Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5- methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A*, **89,** 1827-31.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet*, **24,** 88-91.

Fukumori, Y., Ohnoki, S., Shibata, H. & Nishimukai, H. (1996). Suballeles of the ABO blood group system in a Japanese population. *Hum Hered*, **46,** 85-91.

Garratty, G. & Arndt, P.A. (1999). Applications of flow cytofluorometry to red blood cell immunology. *Cytometry*, **38,** 259-67.

Gassner, C., Schmarda, A., Nussbaumer, W. & Schonitzer, D. (1996). ABO glycosyltransferase genotyping by polymerase chain reaction using sequence-specific primers. *Blood*, **88,** 1852-6.

Goddard, A.D. & Solomon, E. (1993). Genetic aspects of cancer. *Adv Hum Genet*, **21,** 321-76.

Gold, E.R., Tovey, G.H., Benney, W.E. & Lewis, F.J.W. (1959). Changes in the group A antigen in a case of leukaemia. *Nature*, **183,** 892-3.

Gonzalez-Zulueta, M., Bender, C.M., Yang, A.S., Nguyen, T., Beart, R.W., Van Tornout, J.M. & Jones, P.A. (1995). Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing. *Cancer Res*, **55,** 4531-5.

Gonzalgo, M.L. & Jones, P.A. (1997). Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res*, **25,** 2529-31.

Gorelick, J.I., Oyasu, R., Golmon, M.E. & Grayhack, J.T. (1984). Correlation of ABH antigenicity and urine cytology results in transitional cell carcinoma of bladder. *Urology*, **24,** 287-90.

Grant, P.A. (2001). A tale of histone modifications. *Genome Biol*, **2.**

Greger, V., Debus, N., Lohmann, D., Hopping, W., Passarge, E. & Horsthemke, B. (1994). Frequency and parental origin of hypermethylated RB1 alleles in retinoblastoma. *Hum Genet*, **94,** 491-6.

Greger, V., Passarge, E., Hopping, W., Messmer, E. & Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet*, **83,** 155-8.

Grunnet, N., Steffensen, R., Bennett, E.P. & Clausen, H. (1994). Evaluation of histo-blood group ABO genotyping in a Danish population: frequency of a novel O allele defined as O2. *Vox Sang*, **67,** 210-5.

Guzman, M.L., Upchurch, D., Grimes, B., Howard, D.S., Rizzieri, D.A., Luger, S.M., Phillips, G.L. & Jordan, C.T. (2001). Expression of tumor-suppressor genes interferon regulatory factor 1 and death-associated protein kinase in primitive acute myelogenous leukemia cells. *Blood*, **97,** 2177-9.

Hakomori, S. (1981). Blood group ABH and Ii antigens of human erythrocytes: chemistry, polymorphism, and their developmental change. *Semin Hematol*, **18,** 39-62.

Hakomori, S. (1999). Antigen structure and genetic basis of histo-blood groups A, B and O: their changes associated with human cancer(1). *Biochim Biophys Acta*, **1473,** 247-266.

Hallouin, F., Goupille, C., Bureau, V., Meflah, K. & Le Pendu, J. (1999). Increased tumorigenicity of rat colon carcinoma cells after alpha1,2- fucosyltransferase FTA anti-sense cDNA transfection. *Int J Cancer*, **80,** 606-11.

Hansen, T., Namork, E., Olsson, M.L., Chester, M.A. & Heier, H.E. (1998). Different genotypes causing indiscernible patterns of A expression on A(el) red blood cells as visualized by scanning immunogold electron microscopy. *Vox Sang*, **75,** 47-51.

Hata, Y., Kominato, Y., Yamamoto, F.I. & Takizawa, H. (2002). Characterization of the human ABO gene promoter in erythroid cell lineage. *Vox Sang*, **82,** 39-46.

Herman, J.G., Graff, J.R., Myohanen, S., Nelkin, B.D. & Baylin, S.B. (1996a). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A*, **93,** 9821-6.

Herman, J.G., Jen, J., Merlo, A. & Baylin, S.B. (1996b). Hypermethylation-associated inactivation indicates a tumor suppressor role for p15INK4B. *Cancer Res*, **56,** 722-7.

Herman, J.G., Latif, F., Weng, Y., Lerman, M.I., Zbar, B., Liu, S., Samid, D., Duan, D.S., Gnarra, J.R., Linehan, W.M. & et al. (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A*, **91,** 9700-4.

Herman, J.G., Merlo, A., Mao, L., Lapidus, R.G., Issa, J.P., Davidson, N.E., Sidransky, D. & Baylin, S.B. (1995). Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res*, **55,** 4525-30.

Hirohashi, S., Ino, Y., Kodama, T. & Shimosato, Y. (1984). Distribution of blood group antigens A, B, H, and I(Ma) in mucus- producing adenocarcinoma of human lung. *J Natl Cancer Inst*, **72,** 1299-305.

Hosoi, E., Hirose, M., Hamano, S. & Kuroda, Y. (1998). Detection of histo-blood group ABO mRNA in human chronic myeloid leukemia cell lines using reverse transcription-polymerase chain reaction (RT-PCR). *Cancer Lett*, **133,** 191-6.

Hussey, D.J., Moore, S., Nicola, M. & Dobrovic, A. (2001). Fusion of the NUP98 gene with the LEDGF/p52 gene defines a recurrent acute myeloid leukemia translocation. *BMC Genet*, **2,** 20.

Ichikawa, D., Handa, K. & Hakomori, S. (1998). Histo-blood group A/B antigen deletion/reduction vs. continuous expression in human tumor cells as correlated with their malignancy. *Int J Cancer*, **76,** 284-9.

Inai, K., Yamamoto, A., Arihiro, K., Khatun, N., Kurihara, K. & Takeda, S. (1992). Epithelial hyperplasia of the breast. Comparison of incidence between cases of breast carcinoma and control autopsy specimens with immunohistochemical observation of blood group antigens. *Acta Pathol Jpn*, **42,** 193-200.

Irshaid, N.M., Chester, M.A. & Olsson, M.L. (1999). Allele-related variation in minisatellite repeats involved in the transcription of the blood group ABO gene. *Transfus Med*, **9,** 219-26.

Itzkowitz, S.H., Yuan, M., Ferrell, L.D., Ratcliffe, R.M., Chung, Y.S., Satake, K., Umeyama, K., Jones, R.T. & Kim, Y.S. (1987). Cancer-associated alterations of blood group antigen expression in the human pancreas. *J Natl Cancer Inst*, **79,** 425-34.

Ivanov, E.P., Levin, V.I., Smirnova, L.A. & Molochko, V.A. (1990). [Expression of A, B, and H antigens on erythrocytes of patients with leukemia]. *Gematol Transfuziol*, **35,** 23-4.

Iwamoto, S., Withers, D.A., Handa, K. & Hakomori, S. (1999). Deletion of A-antigen in a human cancer cell line is associated with reduced promoter activity of CBF/NF-Y binding region, and possibly with enhanced DNA methylation of A transferase promoter. *Glycoconj J*, **16,** 659-66.

Jackson-Grusby, L., Laird, P.W., Magge, S.N., Moeller, B.J. & Jaenisch, R. (1997). Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. *Proc Natl Acad Sci U S A*, **94,** 4681-5.

Janssens, V., van Hoof, C., Martens, E., de Baere, I., Merlevede, W. & Goris, J. (2000). Identification and characterization of alternative splice products encoded by the human phosphotyrosyl phosphatase activator gene. *Eur J Biochem*, **267,** 4406-13.

Jaroszeski, M.J. & Radcliff, G. (1999). Fundamentals of flow cytometry. *Mol Biotechnol*, **11,** 37-53.

Johnson, P.H. & Hopkinson, D.A. (1992). Detection of ABO blood group polymorphism by denaturing gradient gel electrophoresis. *Hum Mol Genet*, **1,** 341-4.

Jones, P.A. (1996). DNA methylation errors and cancer. *Cancer Res*, **56,** 2463-7.

Jones, P.A., Gonzalgo, M.L., Tsutsumi, M. & Bender, C.M. (1998). DNA methylation in bladder cancer. *Eur Urol*, **33,** 7-8.

Juhl, B.R., Hartzen, S.H. & Hainau, B. (1986). A, B, H antigen expression in transitional cell carcinomas of the urinary bladder. *Cancer*, **57,** 1768-75.

Jurado, L.A., Coloma, A. & Cruces, J. (1999). Identification of a human homolog of the Drosophila rotated abdomen gene (POMT1) encoding a putative protein O-mannosyl-transferase, and assignment to human chromosome 9q34.1. *Genomics*, **58,** 171-80.

Kahn, A., Vroclans, M., Hakim, J. & Boivin, P. (1971). Differences in the two red-cell populations in erythroleukaemia. *Lancet*, **2,** 933.

Kang, S.H., Fukumori, Y., Ohnoki, S., Shibata, H., Han, K.S., Nishimukai, H. & Okubo, Y. (1997). Distribution of ABO genotypes and allele frequencies in a Korean population. *Jpn J Hum Genet*, **42,** 331-5.

Katzenellenbogen, R.A., Baylin, S.B. & Herman, J.G. (1999). Hypermethylation of the DAP-kinase CpG island is a common alteration in B-cell malignancies. *Blood*, **93,** 4347-53.

Keinanen, M., Griffin, J.D., Bloomfield, C.D., Machnicki, J. & de la Chapelle, A. (1988). Clonal chromosomal abnormalities showing multiple-cell-lineage involvement in acute myeloid leukemia. *N Engl J Med*, **318,** 1153-8.

Kelly, R.J., Rouquier, S., Giorgi, D., Lennon, G.G. & Lowe, J.B. (1995). Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J Biol Chem*, **270,** 4640-9.

Kinoshita, H., Shi, Y., Sandefur, C. & Jarrard, D.F. (2000). Screening hypermethylated regions by methylation-sensitive single-strand conformational polymorphism. *Anal Biochem*, **278,** 165-9.

Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A*, **68,** 820-3.

Koda, Y., Soejima, M., Johnson, P.H., Smart, E. & Kimura, H. (1997). Missense mutation of FUT1 and deletion of FUT2 are responsible for Indian Bombay phenotype of ABO blood group system. *Biochem Biophys Res Commun*, **238,** 21-5.

Koda, Y., Soejima, M. & Kimura, H. (1998). Changing transcription start sites in H-type alpha(1,2)fucosyltransferase gene (FUT1) during differentiation of the human erythroid lineage. *Eur J Biochem*, **256,** 379-87.

Kominato, Y., Hata, Y., Takizawa, H., Tsuchiya, T., Tsukada, J. & Yamamoto, F. (1999). Expression of human histo-blood group ABO genes is dependent upon DNA methylation of the promoter region. *J Biol Chem*, **274,** 37240-50.

Kominato, Y., Tsuchiya, T., Hata, N., Takizawa, H. & Yamamoto, F. (1997). Transcription of human ABO histo-blood group genes is dependent upon binding of transcription factor CBF/NF-Y to minisatellite sequence. *J Biol Chem*, **272,** 25890-8.

Koscielak, J. (2001). ABH blood group active glycoconjugates from human red cells. *Transfus Med*, **11,** 267-79.

Koscielak, J., Lenkiewicz, B., Zielenski, J. & Seyfried, H. (1986). Weak A phenotypes possibly caused by mutation. *Vox Sang*, **50,** 187-90.

Laird, P.W., Jackson-Grusby, L., Fazeli, A., Dickinson, S.L., Jung, W.E., Li, E., Weinberg, R.A. & Jaenisch, R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell*, **81,** 197-205.

Lange, P.H., Limas, C. & Fraley, E.E. (1978). Tissue blood-group antigens and prognosis in low stage transitional cell carcinoma of the bladder. *J Urol*, **119,** 52-5.

Langlois, R.G., Bigbee, W.L. & Jensen, R.H. (1985). Flow cytometric characterization of normal and variant cells with monoclonal antibodies specific for glycophorin A. *J Immunol*, **134,** 4009-17.

Larsen, R.D., Ernst, L.K., Nair, R.P. & Lowe, J.B. (1990). Molecular cloning, sequence, and expression of a human GDP-L- fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can form the H blood group antigen. *Proc Natl Acad Sci U S A*, **87,** 6674-8.

Le Pendu, J., Marionneau, S., Cailleau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B. & Clement, M. (2001). ABH and Lewis histo-blood group antigens in cancer. *Apmis*, **109,** 9-31.

Lee, J.C. & Chang, J.G. (1992). ABO genotyping by polymerase chain reaction. *J Forensic Sci*, **37,** 1269-75.

Lee, J.S., Ro, J.Y., Sahin, A.A., Hong, W.K., Brown, B.W., Mountain, C.F. & Hittelman, W.N. (1991). Expression of blood-group antigen A--a favorable prognostic factor in non-small-cell lung cancer. *N Engl J Med*, **324,** 1084-90.

Lehrs, H. (1930). Uber gruppenspezifischen eigenschaften des menslichen Speichels. *Zeitscrift fur Immunforsch*, **66,** 175-92.

Li, E., Bestor, T.H. & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69,** 915-26.

Limas, C., Lange, P., Fraley, E.E. & Vessella, R.L. (1979). A, B, H antigens in transitional cell tumors of the urinary bladder: correlation with the clinical course. *Cancer*, **44,** 2099-107.

Lindgren, A., Stendahl, U., Brodin, T., Lundblad, A., Sallstrom, J. & Busch, C. (1986). Blood group antigen expression and prognosis in squamous cell carcinoma of the uterine cervix. *Anticancer Res*, **6,** 255-8.

Liu, Q., Feng, J., Buzin, C., Wen, C., Nozari, G., Mengos, A., Nguyen, V., Liu, J., Crawford, L., Fujimura, F.K. & Sommer, S.S. (1999). Detection of virtually all mutations-SSCP (DOVAM-S): a rapid method for mutation scanning with virtually 100% sensitivity. *Biotechniques*, **26,** 932, 936-8, 940-2.

Liu, W.M., Maraia, R.J., Rubin, C.M. & Schmid, C.W. (1994). Alu transcripts: cytoplasmic localisation and regulation by DNA methylation. *Nucleic Acids Res*, **22,** 1087-95.

Lu, S. & Davies, P.J. (1997). Regulation of the expression of the tissue transglutaminase gene by DNA methylation. *Proc Natl Acad Sci U S A*, **94,** 4692-7.

Maekawa, M., Sugano, K., Kashiwabara, H., Ushiama, M., Fujita, S., Yoshimori, M. & Kakizoe, T. (1999). DNA methylation analysis using bisulfite treatment and PCR-single-strand conformation polymorphism in colorectal cancer showing microsatellite instability. *Biochem Biophys Res Commun*, **262,** 671-6.

Maekawa, M., Sugano, K., Ushiama, M., Fukayama, N., Nomoto, K., Kashiwabara, H., Fujita, S. & Kakizoe, T. (2001). Heterogeneity of DNA methylation status analyzed by bisulfite-PCR-SSCP and correlation with clinico-pathological characteristics in colorectal cancer. *Clin Chem Lab Med*, **39,** 121-8.

Maesawa, C., Tamura, G., Nishizuka, S., Ogasawara, S., Ishida, K., Terashima, M., Sakata, K., Sato, N., Saito, K. & Satodate, R. (1996). Inactivation of the CDKN2 gene by homozygous deletion and de novo methylation is associated with advanced stage esophageal squamous cell carcinoma. *Cancer Res*, **56,** 3875-8.

Majsky, A., Jakoubkova, J. & Machackova, M. (1967). Frequency and serological character of changes of group ABH-antigens in blood diseases and malignant tumours. *Neoplasma*, **14,** 41-8.

Marsden, K.A., Pearse, A.M., Collins, G.G., Ford, D.S., Heard, S. & Kimber, R.I. (1992). Acute leukemia with t(1;3)(p36;q21), evolution to t(1;3)(p36;q21), t(14;17)(q32;q21), and loss of red cell A and Le(b) antigens. *Cancer Genet Cytogenet*, **64,** 80-5.

Matsumoto, H., Muramatsu, H., Shimotakahara, T., Yanagi, M., Nishijima, H., Mitani, N., Baba, K., Muramatsu, T. & Shimazu, H. (1993). Correlation of expression of ABH blood group carbohydrate antigens with metastatic potential in human lung carcinomas. *Cancer*, **72,** 75-81.

Meehan, R.R., Lewis, J.D. & Bird, A.P. (1992). Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res*, **20,** 5085-92.

Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L. & Bird, A.P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, **58,** 499-507.

Meldgaard, P., Holmes, E.H., Bennett, E.P., Clausen, H., Zeuthen, J., Wolf, H. & Orntoft, T.F. (1994). Blood group ABO-related glycosylation of urothelial cell lines: immunocytological, enzymatic, and genetic characterization. *Cancer Res*, **54,** 2440-7.

Memon, R.S., Yonezawa, S., Hasui, K. & Sato, E. (1990). Expression of blood group antigens and lectin binding in adenocarcinomas of the endometrium. *Acta Pathol Jpn*, **40,** 509-16.

Merlo, A., Herman, J.G., Mao, L., Lee, D.J., Gabrielson, E., Burger, P.C., Baylin, S.B. & Sidransky, D. (1995). 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat Med*, **1,** 686-92.

Metoki, R., Kakudo, K., Tsuji, Y., Teng, N., Clausen, H. & Hakomori, S. (1989). Deletion of histo-blood group A and B antigens and expression of incompatible A antigen in ovarian cancer. *J Natl Cancer Inst*, **81,** 1151-7.

Mifsud, N.A., Watt, J.M., Condon, J.A., Haddad, A.P. & Sparrow, R.L. (2000). A novel cis-AB variant allele arising from a nucleotide substitution A796C in the B transferase gene. *Transfusion*, **40,** 1276-7.

Mitchell, D.A. & Sprague, G.F., Jr. (2001). The phosphotyrosyl phosphatase activator, Ncs1p (Rrd1p), functions with Cla4p to regulate the G(2)/M transition in Saccharomyces cerevisiae. *Mol Cell Biol*, **21,** 488-500.

Miyake, M., Taki, T., Hitomi, S. & Hakomori, S. (1992). Correlation of expression of H/Le(y)/Le(b) antigens with survival in patients with carcinoma of the lung. *N Engl J Med*, **327,** 14-8.

Mizuno, S., Chijiwa, T., Okamura, T., Akashi, K., Fukumaki, Y., Niho, Y. & Sasaki, H. (2001). Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. *Blood*, **97,** 1172-9.

Mohandas, T., Sparkes, R.S. & Shapiro, L.J. (1981). Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science*, **211,** 393-6.

Moldvay, J., Scheid, P., Wild, P., Nabil, K., Siat, J., Borrelly, J., Marie, B., Farre, G., Labib, T., Pottier, G., Sesboue, R., Bronner, C., Vignaud, J.M., Martinet, Y. & Martinet, N. (2000). Predictive survival markers in patients with surgically resected non- small cell lung carcinoma. *Clin Cancer Res*, **6,** 1125-34.

Mollison, P.L., Engelfriet, C.P. & Contreras, M. (1993). *Blood transfusion in clinical medicine.* Blackwell Scientific Publications: Oxford.

Mourant, A.E., Kopec, A.C. & Domaniewska, K. (1976). *The Distribution of Human Blood Groups and Other Polymorphisms*. Oxford University Press: Oxford.

Nakagoe, T., Fukushima, K., Tuji, T., Sawai, T., Nanashima, A., Yamaguchi, H., Yasutake, T., Hara, S., Ayabe, H., Matuo, T. & Kamihira, S. (1998). Immunohistochemical expression of ABH/Lewis-related antigens in primary breast carcinomas and metastatic lymph node lesions. *Cancer Detect Prev*, **22,** 499-505.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. & Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393,** 386-9.

Ng, M.H., Chung, Y.F., Lo, K.W., Wickham, N.W., Lee, J.C. & Huang, D.P. (1997). Frequent hypermethylation of p16 and p15 genes in multiple myeloma. *Blood*, **89,** 2500-6.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science*, **194,** 23-8.

O'Keefe, D.S. & Dobrovic, A. (1993a). A rapid and reliable PCR method for genotyping the ABO blood group. *Hum Mutat*, **2,** 67-70.

O'Keefe, D.S. & Dobrovic, A. (1993b). A rapid and reliable PCR method for genotyping the ABO blood group [published erratum appears in Hum Mutat 1994;3(3):337]. *Hum Mutat*, **2,** 67-70.

O'Keefe, D.S. & Dobrovic, A. (1996a). Decreased stability of the O allele mRNA transcript of the ABO gene. *Blood*, **87,** 3061-2.

O'Keefe, D.S. & Dobrovic, A. (1996b). A rapid and reliable PCR method for genotyping the ABO blood group. II: A2 and O2 alleles. *Hum Mutat*, **8,** 358-61.

Ogasawara, K., Bannai, M., Saitou, N., Yabe, R., Nakata, K., Takenaka, M., Fujisawa, K., Uchikawa, M., Ishikawa, Y., Juji, T. & Tokunaga, K. (1996a). Extensive polymorphism of ABO blood group gene: three major lineages of the alleles for the common ABO phenotypes. *Hum Genet*, **97,** 777-83.

Ogasawara, K., Yabe, R., Uchikawa, M., Saitou, N., Bannai, M., Nakata, K., Takenaka, M., Fujisawa, K., Ishikawa, Y., Juji, T. & Tokunaga, K. (1996b). Molecular genetic analysis of variant phenotypes of the ABO blood group system. *Blood*, **88,** 2732-7.

Ogawa, J., Iwazaki, M., Inoue, H., Koide, S., Kawada, S. & Shohtsu, A. (1987). Relation of ABH blood group antigen expression to prognosis in lung carcinoma. *Tokai J Exp Clin Med*, **12,** 49-54.

Ohtani-Fujita, N., Fujita, T., Aoike, A., Osifchin, N.E., Robbins, P.D. & Sakai, T. (1993). CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene. *Oncogene*, **8,** 1063-7.

Okano, M., Bell, D.W., Haber, D.A. & Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99,** 247-57.

Okano, M., Xie, S. & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet*, **19,** 219-20.

Olsson, M.L. & Chester, M.A. (1995). A rapid and simple ABO genotype screening method using a novel B/O2 versus A/O2 discriminating nucleotide substitution at the ABO locus. *Vox Sang*, **69,** 242-7.

Olsson, M.L. & Chester, M.A. (1996a). Evidence for a new type of O allele at the ABO locus, due to a combination of the A2 nucleotide deletion and the Ael nucleotide insertion. *Vox Sang*, **71,** 113-7.

Olsson, M.L. & Chester, M.A. (1996b). Polymorphisms at the ABO locus in subgroup A individuals. *Transfusion*, **36,** 309-13.

Olsson, M.L. & Chester, M.A. (1998). Heterogeneity of the blood group Ax allele: genetic recombination of common alleles can result in the Ax phenotype. *Transfus Med*, **8,** 231-8.

Olsson, M.L. & Chester, M.A. (2001). Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies. *Transfus Med*, **11,** 295-313.

Olsson, M.L., Irshaid, N.M., Hosseini-Maaf, B., Hellberg, A., Moulds, M.K., Sareneva, H. & Chester, M.A. (2001). Genomic analysis of clinical samples with serologic ABO blood grouping discrepancies: identification of 15 novel A and B subgroup alleles. *Blood*, **98,** 1585-93.

Olsson, M.L., Santos, S.E., Guerreiro, J.F., Zago, M.A. & Chester, M.A. (1998). Heterogeneity of the O alleles at the blood group ABO locus in Amerindians. *Vox Sang*, **74,** 46-50.

Olsson, M.L., Thuresson, B. & Chester, M.A. (1995). An Ael allele-specific nucleotide insertion at the blood group ABO locus and its detection using a sequence-specific polymerase chain reaction. *Biochem Biophys Res Commun*, **216,** 642-7.

Oriol, R., Danilovs, J. & Hawkins, B.R. (1981). A new genetic model proposing that the Se gene is a structural gene closely linked to the H gene. *Am J Hum Genet*, **33,** 421-31.

Orlow, I., Lacombe, L., Pellicer, I., Rabbani, F., Delgado, R., Zhang, Z.F., Szijan, I. & Cordon-Cardo, C. (1998). Genotypic and phenotypic characterization of the histoblood group ABO(H) in primary bladder tumors. *Int J Cancer*, **75,** 819-24.

Orntoft, T.F., Greenwell, P., Clausen, H. & Watkins, W.M. (1991). Regulation of the oncodevelopmental expression of type 1 chain ABH and Lewis(b) blood group antigens in human colon by alpha-2-L-fucosylation. *Gut*, **32,** 287-93.

Orntoft, T.F., Meldgaard, P., Pedersen, B. & Wolf, H. (1996). The blood group ABO gene transcript is down-regulated in human bladder tumors and growth-stimulated urothelial cell lines. *Cancer Res*, **56,** 1031-36.

Orntoft, T.F., Wolf, H., Clausen, H., Dabelsteen, E. & Hakomori, S. (1989). Blood group ABH-related antigens in normal and malignant bladder urothelium: possible structural basis for the deletion of type-2 chain ABH antigens in invasive carcinomas. *Int J Cancer*, **43,** 774-80.

Pearson, S.L. & Hessner, M.J. (1998). A(1,2)BO(1,2) genotyping by multiplexed allele-specific PCR. *Br J Haematol*, **100,** 229-34.

Perlman, E.J. & Epstein, J.I. (1990). Blood group antigen expression in dysplasia and adenocarcinoma of the prostate. *Am J Surg Pathol*, **14,** 810-8.

Piller, F., Cartron, J.P. & Tuppy, H. (1980). Increase of blood group A and loss of blood group Sda activity in the mucus from human neoplastic colon. *Rev Fr Transfus Immunohematol*, **23,** 599-611.

Procter, J., Crawford, J., Bunce, M. & Welsh, K.I. (1997). A rapid molecular method (polymerase chain reaction with sequence- specific primers) to genotype for ABO blood group and secretor status and its potential for organ transplants. *Tissue Antigens*, **50,** 475-83.

Prowse, A.H., Webster, A.R., Richards, F.M., Richard, S., Olschwang, S., Resche, F., Affara, N.A. & Maher, E.R. (1997). Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am J Hum Genet*, **60,** 765-71.

Punzel, M. & Ho, A.D. (2001). Divisional history and pluripotency of human hematopoietic stem cells. *Ann N Y Acad Sci*, **938,** 72-81; discussion 81-2.

Putkonen, T. (1930). Uber gruppenspezifischen eigenschaften verschiedener korper flussigkeiten. *Acta Society Med Fenn Duodecim A*, **14**.

Race, R.R. & Sanger, R. (1975). *Blood Groups in Man*. Blackwell Scientific Publications: Oxford.

Raizis, A.M., Schmitt, F. & Jost, J.P. (1995). A bisulfite method of 5-methylcytosine mapping that minimizes template degradation. *Anal Biochem*, **226,** 161-6.

Ravn, V. & Dabelsteen, E. (2000). Tissue distribution of histo-blood group antigens. *Apmis*, **108,** 1-28.

Razin, A. (1998). CpG methylation, chromatin structure and gene silencing-a three-way connection. *Embo J*, **17,** 4905-8.

Reguigne-Arnould, I., Couillin, P., Mollicone, R., Faure, S., Fletcher, A., Kelly, R.J., Lowe, J.B. & Oriol, R. (1995). Relative positions of two clusters of human alpha-L-fucosyltransferases in 19q (FUT1-FUT2) and 19p (FUT6-FUT3-FUT5) within the microsatellite genetic map of chromosome 19. *Cytogenet Cell Genet*, **71,** 158-62.

Reyes, F., Gourdin, M.F., Lejonc, J.L., Cartron, J.P., Gorius, J.B. & Dreyfus, B. (1976). The heterogeneity of erythrocyte antigen distribution in human normal phenotypes: an immunoelectron microscopy study. *Br J Haematol*, **34,** 613-21.

Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B. & Schuebel, K.E. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature*, **404,** 1003-7.

Robertson, K.D., Keyomarsi, K., Gonzales, F.A., Velicescu, M. & Jones, P.A. (2000). Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. *Nucleic Acids Res*, **28,** 2108-13.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A. & Jones, P.A. (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res*, **27,** 2291-8.

Rountree, M.R., Bachman, K.E. & Baylin, S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet*, **25,** 269-77.

Ruscetti, F.W., Keller, J.R. & Longo, D.L. (1998). Hematopoiesis. In *Harrison's principles of internal medicine*, Fauci, A.S. (ed), Vol. 1. pp. 634-8. McGraw-Hill: New York.

Sadri, R. & Hornsby, P.J. (1996). Rapid analysis of DNA methylation using new restriction enzyme sites created by bisulfite modification. *Nucleic Acids Res*, **24,** 5058-9.

Sakai, T., Toguchida, J., Ohtani, N., Yandell, D.W., Rapaport, J.M. & Dryja, T.P. (1991). Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet*, **48,** 880-8.

Sakamoto, Y., Kamada, M., Irahara, M., Hasebe, H., Daitoh, T., Furumoto, H., Kishi, Y. & Mori, T. (1986). Differences in clinical significance of blood group antigens A, B, and H in carcinoma tissue in the uterine cervix. *Gynecol Oncol*, **25,** 195-203.

Salmon, C., Cartron, J.P., Lopez, M., Rahuel, C., Badet, J. & Janot, C. (1984). Level of the A, B and H blood group glycosyltransferases in red cell membranes from patients with malignant hemopathies. *Rev Fr Transfus Immunohematol*, **27,** 625-37.

Salmon, C., Seger, J., Mannoni, P., Bahno-Duchery, J. & Liberge, G. (1968). [A population of erythrocytes with simulated anomaly of phenotypes induced by genes of the ABO locus and adenylate kinase]. *Rev Fr Etud Clin Biol*, **13,** 296-8.

Sawyers, C.L., Denny, C.T. & Witte, O.N. (1991). Leukemia and the disruption of normal hematopoiesis. *Cell*, **64,** 337-50.

Saxon, P.J., Srivatsan, E.S. & Stanbridge, E.J. (1986). Introduction of human chromosome 11 via microcell transfer controls tumorigenic expression of HeLa cells. *Embo J*, **5,** 3461-6.

Schachter, H., Michaels, M.A., Tilley, C.A., Crookston, M.C. & Crookston, J.H. (1973). Qualitative differences in the N-acetyl-D-galactosaminyltransferases produced by human A1 and A2 genes. *Proc Natl Acad Sci U S A*, **70,** 220-4.

Schoentag, R., Primus, F.J. & Kuhns, W. (1987). ABH and Lewis blood group expression in colorectal carcinoma. *Cancer Res*, **47,** 1695-700.

Schwarzinger, I., Valent, P., Koller, U., Marosi, C., Schneider, B., Haas, O., Knapp, W., Lechner, K. & Bettelheim, P. (1990). Prognostic significance of surface marker expression on blasts of patients with de novo acute myeloblastic leukemia. *J Clin Oncol*, **8,** 423-30.

Seyfried, H., Walewska, I. & Werblinska, B. (1964). Unusual inheritance of ABO group in a family with weak B antigens. *Vox Sang*, **9,** 268-77.

Sharon, R. & Fibach, E. (1991). Quantitative flow cytometric analysis of ABO red cell antigens. *Cytometry*, **12,** 545-9.

Sheffield, V.C., Beck, J.S., Kwitek, A.E., Sandstrom, D.W. & Stone, E.M. (1993). The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics*, **16,** 325-32.

Siegfried, Z. & Cedar, H. (1997). DNA methylation: a molecular lock. *Curr Biol*, **7,** R305-7.

Simoneau, A.R., Spruck, C.H., 3rd, Gonzalez-Zulueta, M., Gonzalgo, M.L., Chan, M.F., Tsai, Y.C., Dean, M., Steven, K., Horn, T. & Jones, P.A. (1996). Evidence for two tumor suppressor loci associated with proximal chromosome 9p to q and distal chromosome 9q in bladder cancer and the initial screening for GAS1 and PTC mutations. *Cancer Res*, **56,** 5039-43.

Singal, R. & Ginder, G.D. (1999). DNA methylation. *Blood*, **93,** 4059-70.

Stanbridge, E.J. (1990). Human tumor suppressor genes. *Annu Rev Genet*, **24,** 615-57.

Starling, K.A. & Fernbach, D.J. (1970). Changes in strength of a antigen in children with acute leukemia. *Transfusion*, **10,** 3-5.

Stenersen, T.C. & Dabelsteen, E. (1992). Changes in the glycosylation pattern of histo-blood group antigens in benign, premalignant and malignant laryngeal epithelial lesions. *APMIS Suppl*, **27,** 139-48.

Stirzaker, C., Millar, D.S., Paul, C.L., Warnecke, P.M., Harrison, J., Vincent, P.C., Frommer, M. & Clark, S.J. (1997). Extensive DNA methylation spanning the Rb promoter in retinoblastoma tumors. *Cancer Res*, **57,** 2229-37.

Strahl-Bolsinger, S., Gentzsch, M. & Tanner, W. (1999). Protein O-mannosylation. *Biochim Biophys Acta*, **1426,** 297-307.

Suciu, S., Zeller, W., Weh, H.J. & Hossfeld, D.K. (1993). Immunophenotype of mitotic cells with clonal chromosome abnormalities demonstrating multilineage involvement in acute myeloid leukemia. *Cancer Genet Cytogenet*, **70,** 1-5.

Suzuki, K., Iwata, M., Tsuji, H., Takagi, T., Tamura, A., Ishimoto, G., Ito, S., Matsui, K. & Miyazaki, T. (1997). A de novo recombination in the ABO blood group gene and evidence for the occurence of recombination products. *Hum Genet*, **99,** 454-61.

Sweetser, D.A., Chen, C.S., Blomberg, A.A., Flowers, D.A., Galipeau, P.C., Barrett, M.T., Heerema, N.A., Buckley, J., Woods, W.G., Bernstein, I.D. & Reid, B.J. (2001). Loss of heterozygosity in childhood de novo acute myelogenous leukemia. *Blood*, **98,** 1188-94.

Szulman, A.E. (1964). The histological distribution of blood group substances in man as disclosed by immunofluorescence. III The A, B and H antigens in embryos and fetuses from 18mm in length. *J Exp Med*, **119,** 503-16.

Szulman, A.E. (1966). Chemistry, distribution, and function of blood group substances. *Annu Rev Med*, **17,** 307-22.

Takita, J., Hayashi, Y., Kohno, T., Yamaguchi, N., Hanada, R., Yamamoto, K. & Yokota, J. (1997). Deletion map of chromosome 9 and p16 (CDKN2A) gene alterations in neuroblastoma. *Cancer Res*, **57,** 907-12.

Tate, P.H. & Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr Opin Genet Dev*, **3,** 226-31.

Tazi, J. & Bird, A. (1990). Alternative chromatin structure at CpG islands. *Cell*, **60,** 909-20.

Thiagalingam, S., Foy, R.L., Cheng, K.H., Lee, H.J., Thiagalingam, A. & Ponte, J.F. (2002). Loss of heterozygosity as a predictor to map tumor suppressor genes in cancer: molecular basis of its occurrence. *Curr Opin Oncol*, **14,** 65-72.

Tycko, B. (2000). Epigenetic gene silencing in cancer. *J Clin Invest*, **105,** 401-7.

Ugozzoli, L. & Wallace, R.B. (1992). Application of an allele-specific polymerase chain reaction to the direct determination of ABO blood group genotypes. *Genomics*, **12,** 670-4.

van der Meulen, F.W., de Bruin, H.G., Goosen, P.C., Bruynes, E.C., Joustra-Maas, C.J., Telkamp, H.G., von dem Borne, A.E. & Engelfriet, C.P. (1980). Quantitative aspects of the destruction of red cells sensitized with IgG1 autoantibodies: an application of flow cytofluorometry. *Br J Haematol*, **46,** 47-56.

van Loghem, J.J., Dorfmeier, H. & van der Hart, M. (1957). Two A antigens with abnormal serologic properties. *Vox Sang*, **2,** 16-24.

Velentza, A.V., Schumacher, A.M., Weiss, C., Egli, M. & Watterson, D.M. (2001). A protein kinase associated with apoptosis and tumor suppression: structure, activity, and discovery of peptide substrates. *J Biol Chem*, **276,** 38956-65.

Versteeg, R. (1997). Aberrant methylation in cancer. *Am J Hum Genet*, **60,** 751-4.

Watkins, W.M. (1980). Biochemistry and Genetics of the ABO, Lewis, and P blood group systems. *Adv Hum Genet*, **10,** 1-136.

Watkins, W.M. (2001). The ABO blood group system: historical background. *Transfus Med*, **11,** 243-65.

Watkins, W.M. & Morgan, W.T.J. (1955). Some observations on the O and H characters of human blood and secretions. *Vox Sang*, **5,** 1-14.

Wedemeyer, N. & Potter, T. (2001). Flow cytometry: an 'old' tool for novel applications in medical genetics. *Clin Genet*, **60,** 1-8.

Weissman, B.E., Saxon, P.J., Pasquale, S.R., Jones, G.R., Geiser, A.G. & Stanbridge, E.J. (1987). Introduction of a normal human chromosome 11 into a Wilms' tumor cell line controls its tumorigenic expression. *Science*, **236,** 175-80.

Welshinger, M., Finstad, C.L., Venkatraman, E., Federici, M.G., Rubin, S.C., Lewis, J.L., Jr. & Lloyd, K.O. (1996). Expression of A, B, and H blood group antigens in epithelial ovarian cancer: relationship to tumor grade and patient survival. *Gynecol Oncol*, **62,** 106-12.

Wetzler, M. & Bloomfield, C.D. (1998). Acute and chronic leukemia. In *Harrison's principles of internal medicine*, Fauci, A.S. (ed), Vol. 1. pp. 684-95. McGraw-Hill: New York.

Wijermans, P.W., Krulder, J.W., Huijgens, P.C. & Neve, P. (1997). Continuous infusion of low-dose 5-Aza-2'-deoxycytidine in elderly patients with high-risk myelodysplastic syndrome. *Leukemia*, **11,** 1-5.

Wolf, G.T., Carey, T.E., Schmaltz, S.P., McClatchey, K.D., Poore, J., Glaser, L., Hayashida, D.J. & Hsu, S. (1990). Altered antigen expression predicts outcome in squamous cell carcinoma of the head and neck. *J Natl Cancer Inst*, **82,** 1566-72.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J. & Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, **402,** 187-91.

Yamaguchi, H., Okubo, Y. & Hazama, F. (1965). An A2 phenotype blood showing atypical mode of inheritance. *Proc Jap Acad*, **41,** 316-20.

Yamamoto, F. (1990). Cloning the ABH genes. *Transfusion*, **30,** 671-2.

Yamamoto, F. (1995). Molecular genetics of the ABO histo-blood group system. *Vox Sang*, **69,** 1-7.

Yamamoto, F. (2001). Cloning and regulation of the ABO genes. *Transfus Med*, **11,** 281-94.

Yamamoto, F., Clausen, H., White, T., Marken, J. & Hakomori, S. (1990a). Molecular genetic basis of the histo-blood group ABO system. *Nature*, **345,** 229-33.

Yamamoto, F. & Hakomori, S. (1990). Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. *J Biol Chem*, **265,** 19257-62.

Yamamoto, F., Marken, J., Tsuji, T., White, T., Clausen, H. & Hakomori, S. (1990b). Cloning and characterization of DNA complementary to human UDP-GalNAc: Fuc alpha 1----2Gal alpha 1----3GalNAc transferase (histo-blood group A transferase) mRNA. *J Biol Chem*, **265,** 1146-51.

Yamamoto, F. & McNeill, P.D. (1996). Amino acid residue at codon 268 determines both activity and nucleotide- sugar donor substrate specificity of human histo-blood group A and B transferases. In vitro mutagenesis study. *J Biol Chem*, **271,** 10515-20.

Yamamoto, F., McNeill, P.D. & Hakomori, S. (1992). Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem Biophys Res Commun*, **187,** 366-74.

Yamamoto, F., McNeill, P.D., Kominato, Y., Yamamoto, M., Hakomori, S., Ishimoto, S., Nishida, S., Shima, M. & Fujimura, Y. (1993a). Molecular genetic analysis of the ABO blood group system: 2. cis-AB alleles. *Vox Sang*, **64,** 120-3.

Yamamoto, F., McNeill, P.D., Yamamoto, M., Hakomori, S., Bromilow, I.M. & Duguid, J.K. (1993b). Molecular genetic analysis of the ABO blood group system: 4. Another type of O allele. *Vox Sang*, **64,** 175-8.

Yamamoto, F., McNeill, P.D., Yamamoto, M., Hakomori, S. & Harris, T. (1993c). Molecular genetic analysis of the ABO blood group system: 3. A(X) and B(A) alleles. *Vox Sang*, **64,** 171-4.

Yamamoto, F., McNeill, P.D., Yamamoto, M., Hakomori, S., Harris, T., Judd, W.J. & Davenport, R.D. (1993d). Molecular genetic analysis of the ABO blood group system: 1. Weak subgroups: A3 and B3 alleles. *Vox Sang*, **64,** 116-9.

Yang, S.C. (1990). [Correlation of the expression of A, B isoantigens with its precursor H antigen in transitional cell carcinoma and the clinical signification of the expression]. *Zhonghua Wai Ke Za Zhi*, **28,** 631-3, 639.

Yip, S.P. (2000). Single-tube multiplex PCR-SSCP analysis distinguishes 7 common ABO alleles and readily identifies new alleles. *Blood*, **95,** 1487-92.

Yisraeli, J., Frank, D., Razin, A. & Cedar, H. (1988). Effect of in vitro DNA methylation on beta-globin gene expression. *Proc Natl Acad Sci U S A*, **85,** 4638-42.

Yoder, J.A., Walsh, C.P. & Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet*, **13,** 335-40.

Yu, L.C., Chang, C.Y., Twu, Y.C. & Lin, M. (2000). Human histo-blood group ABO glycosyltransferase genes: different enhancer structures with different transcriptional activities. *Biochem Biophys Res Commun*, **273,** 459-66.

Zingg, J.M. & Jones, P.A. (1997). Genetic and epigenetic aspects of DNA methylation on genome expression, evolution, mutation and carcinogenesis. *Carcinogenesis*, **18,** 869-82.