Histone Gene " Knock-Out " in Mouse Embryonic Stem Cells

A thesis submitted to the University of Adelaide
for the degree of Doctor of Philosophy

by

Varaporn Thonglairoam, M.Sc.

Department of Biochemistry,
University of Adelaide,
Adelaide, South Australia,
Australia

September, 1994

To

Dr. Julian Wells

Statement

This thesis contains no material which has been accepted for the award of any other degree or diploma by any university. To the best of my knowledge, it contains no material that has been previously published by any other person, except where due reference is made in the text. I consent to the thesis being made available for photocopying and loan.

Varaporn Thonglairoam

Acknowledgments

I would like to acknowledge the assistance of and express my sincere thanks to the following people:

Professor G. E. Rogers, for his permission to work in the Department of Biochemistry at the University of Adelaide.

Dr. Julian R. E. Wells for his excellent supervision, encouragement, enthusiasm and understanding. His invaluable suggestions and consideration were deeply appreciated, and will always be remembered.

Dr. Barry Egan for continued administrative supervision.

Dr. Peter Wigley, Dr. Ian Lyons and Dr. Allan Robins, for their advice in constructing and critically reading this thesis.

Jacquie Beall for performing all tissue culture work described in this thesis, her friendship and encouragement throughout the years.

Anita Peura, for blastocyst injection and production of chimaeric mice, described in chapter 5, and to staff members of the animal house for animal care.

To the members of the Department of Biochemistry, Adelaide University for providing a friendly environment. The friendship and the assistance of Briony Forbes and Rob King are important to me.

To the past and present members of lab 120, lab 109 and lab 110 for their helpful discussions, friendship and for providing a stimulating environment in which to work. Rob King, Allan Robins, Jacquie Beall, Lesley Crocker and Blair Hopwood gave technical advice during the initial stages of this work. Ian Lyons gave advice on gene targeting experiments and helped with final proof-reading of this text. Lena Rondahl provided the analysis of progeny of mice heterozygous for the disrupted H2A.Z gene. Tim Blake assisted with the electronic data storage. Renate Faast, Michael Clarkson, Lesley Crocker, Kym Duncliffe and Juliana Beltrame answered all the questions.

Finally, I would like to thank my parents, brothers and sisters for their love, continued support and encouragement throughout my academic career.

During the course of this work I was supported by a Merit Scholarship Scheme awarded by the Australian International Development Assistance Bureau.

Summary

This thesis describes studies of the mouse histone variant H2A.Z. The overall aim was to investigate the biological significance of this protein.

The first part of the thesis describes the isolation and characterisation of mouse histone H2A.Z genomic clones from different mouse genomic libraries. A clone containing a mouse H2A.Z gene was initially isolated from a mouse genomic library derived from the BALB/c strain. H2A.Z clones were also isolated from a mouse genomic library derived from the 129 strain, for gene targeting experiments. Southern analysis indicated that the histone H2A.Z gene is present as a single copy per haploid genome, with several pseudogenes. Northern analysis of total RNA isolated from embryonic stem (ES) cells demonstrated that the H2A.Z gene is expressed in ES cells.

Two mouse histone H2A.Z pseudogenes were also isolated during the course of this work. One of the pseudogenes contains sequences identical to a cDNA copy of the H2A.Z gene. A useful adjunct to this finding was that this mouse histone H2A.Z pseudogene sequence was used in an *Escherichia coli* expression system to obtain H2A.Z protein. A high level of H2A.Z protein was produced for use as an antigen for monoclonal antibody production.

The second part of the thesis describes H2A.Z gene targeting in mouse E14 ES cells, derived from the 129 strain. Several targeting vectors were constructed. The successful vector, pHZHR3, was derived from 129 strain DNA. It contains a disrupted H2A.Z gene with a fragment containing a neomycin-resistance gene and approximately 18 kb sequence homology to the chromosomal locus, in order to obtain a high targeting efficiency. Targeted ES cell clones in which one allele of the H2A.Z gene was disrupted by homologous recombination were obtained.

The third part of the thesis describes an attempt to generate ES cell lines and mice which lack the functional H2A.Z protein, in order to investigate H2A.Z function *in vitro* and *in vivo*.

To generate ES cells in which both alleles were disrupted, targeted ES cells, in which one allele had been disrupted, were subjected to a second round of gene targeting using the targeting vector pHZHR4. pHZHR4 was constructed to obtain a high targeting efficiency, and to use a hygromycin-resistance gene as a selectable marker. The targeted ES cell clone, designated ES-193 was electroporated with pHZHR4 DNA. 282 hygromycin-resistant ES cell clones were analysed and none of these clones had both alleles disrupted. To verify this result, the experiment was performed in reverse order: using first the hygromycin resistance based construct, pHZHR4, and then the neomycin resistance based construct, pHZHR3. The second round of gene targeting yielded 293 G418-resistant ES cell clones were analysed. None of these clones had both alleles disrupted. These results imply that ES cells which lack a functional H2A.Z do not survive.

To generate mice which lack a functional H2A.Z gene, two targeted ES cell clones were used. Chimaeric mice were generated from both clones. However, only one of these clones contributed to the germ line. The heterozygous mice derived from the germ line chimaeras were cross-bred in order to obtain mice homozygous for the disrupted H2A.Z gene. Preliminary results indicate that lack of a functional H2A.Z gene is lethal during embryonic development.

The results presented here demonstrate for the first time that H2A.Z has an essential role in mammals.

CONTENTS

Statement i
Acknowledgments ii
Summary iv

Chapter 1 Introduction

1.1 INTRODUCTION 1
1.2 CHROMATIN 1
1.2.1 Nucleosome structure 1
1.2.2 Higher order structure 2
1.2.3 Active chromatin 3
1.3 HISTONES 6
1.3.1 Histone proteins 6
1.3.2 Post-translational modifications of histones 7
(i) Acetylation 7
(ii) Phosphorylation 8
(iii) Ubiquitination 9
1.3.3 Histone structural variants 10
1.3.4 Histone H1 variants 11
1.3.5 Histone H2A variants 12
1.4 HISTONE H2A.Z 13
1.4.1 Post-translational modification of histone H2A.Z 14
1.4.2 Histone H2A.Z gene 14
1.4.3 Possible biological function of H2A.Z 15
1.5 GENE TARGETING BY HOMOLOGOUS RECOMBINATION 17
1.5.1 Embryonic stem (ES) cells 17
1.5.2 Gene targeting in ES cells 18
1.5.3 Targeting vector 19
1.5.4 Enrichment procedure 20
1.5.5 Screening assays for homologous recombinants 21
(i) Polymerase Chain Reaction 21
(ii) Southern analysis 22
1.5.6 Summary 22
1.6 AIMS OF THESIS 23

Chapter 2 Materials and Methods

2.1 ABBREVIATIONS 24

2.2 MATERIALS 25
2.2.1 General reagents and materials 25
2.2.2 Reagents for cell culture 26
2.2.3 Kits 26
2.2.4 Enzymes 26
2.2.5 Radiochemicals 27
2.2.6 Bacterial strains 27
2.2.7 Mouse genomic DNA libraries 27
2.2.8 Tissue culture cell Lines 28
2.2.9 Bacterial growth media 28
2.2.10 Tissue culture media 29
2.2.11 Buffers 30
2.2.12 Cloning vectors 31
2.2.13 Recombinant DNA clones 31
2.2.14 Molecular size markers 31
2.2.15 Oligonucleotides 32

2.3 DNA METHODS 32
2.3.1 Isolation and purification of plasmid DNA 32
2.3.2 Isolation of phage DNA 34
2.3.3 Isolation of mouse genomic DNA 34
(i) Isolation of genomic DNA from tissue samples 34
(ii) Isolation of genomic DNA from mouse tail biopsy 35
(iii) Isolation of genomic DNA from tissue culture cells 35
2.3.4 Restriction enzyme digestion of plasmid and phage DNA 36
2.3.5 Agarose gel electrophoresis of phage and plasmid DNA 36
2.3.6 Isolation of DNA fragment from agarose gel 36
2.3.7 Polyacrylamide gel electrophoresis 37
2.3.8 Labelling of DNA fragments 38
(i) End-labelling of synthetic oligonucleotides 38
(ii) Oligo-labelling of DNA fragments 38
2.3.9 Southern transfer and hybridisation of DNA on GeneScreen 38
2.3.10 Southern analysis of mouse genomic DNA 39
2.3.11 Screening of mouse genomic λ libraries 40
2.3.12 Subcloning of restriction fragments into plasmid vectors 41
(i) Preparation of plasmid vectors 41
(ii) Preparation of DNA restriction fragments 42

(iii) End-filling of vectors and DNA fragments for blunt end ligation 42
(iv) Ligation into plasmid vectors 42
(v) Transformation of plasmid recombinants into bacteria 42
2.3.13 Subcloning of DNA restriction fragments into M13 Vectors 43
(i) Preparation of M13 replicative form (RF) 43
(ii) Subcloning into M13 vectors 43
(iii) Transformation of recombinant M13 into *E. coli* JM101 43
(iv) Preparation of single-stranded M13 DNA 44
2.3.14 Oligonucleotide-directed *in vitro* mutagenesis 44
(i) End-labelling of the oligonucleotides 44
(ii) Mutagenesis reaction and transformation 45
(iii) Screening plaques for mutants 45
(iv) Washing filters with TMACl 45
2.3.15 Dideoxy chain-termination sequencing 46
(i) Sequencing reaction 46
(ii) Sequencing gels 46
2.3.16 Polymerase chain reaction (PCR) analysis 47
2.3.17 Preparation of plasmid DNA for electroporation into ES cells 47
2.4 RNA METHODS 48
2.4.1 Isolation of total RNA from mouse ES cells 48
2.4.2 Northern transfer of RNA 48
2.5 PROTEIN METHODS 49
2.5.1 Preparation of protein from *E. coli* 49
2.5.2 SDS/polyacrylamide gel electrophoresis of proteins 49
2.5.3 Western transfer 50
2.6 TISSUE CULTURE METHODS 51
2.6.1 Preparation of STO feeder layer for ES cell culture 51
2.6.2 Passaging and maintenance of ES cells 52
2.6.3 Electroporation of ES cells 52
2.6.4 Picking and freezing ES cells colonies 53
2.6.5 Processing of the positive ES cell clones 53
2.6.6 Blastocyst injection and production of chimaeric mice 54
2.7 AUTORADIOGRAPHY 54
2.8 UV SPECTROPHOTOMETRY 55
2.9 COMPUTER PROGRAMS 55
2.10 CONTAINMENT FACILITIES AND ANIMAL ETHICS 55

Chapter 3 Isolation and Characterisation of the Mouse Histone H2A.Z Gene

3.1 INTRODUCTION 56

3.2 RESULTS 57

3.2.1 Analysis of previous isolated clones, pBS14-2.9 and pBS17-3 57
(i) Sequencing analysis of pBS14-2.9 57
(ii) Sequencing analysis of pBS17-3 57

3.2.2 Mouse genomic library I screening 58
(i) Analysis of five positive clones with the 573 bp *Pvu* II probe 59
(ii) Subcloning and sequencing analysis of clones derived from λ17w 59

3.2.3 Isolation of an authentic mouse histone H2A.Z genomic clone 61
(i) Mouse genomic library II screening 61
(ii) PCR analysis of eight positive clones 61
(iii) Southern analysis and subclone of clone λ13 62
(iv) Sequencing of the mouse histone H2A.Z gene 62

3.2.4 Southern analysis of mouse genomic DNA 65

3.2.5 Northern analysis of total RNA from ES cells 66

3.2.6 Expression of mouse histone H2A.Z in bacteria 67
(i) Cloning of mouse histone H2A.Z cDNA sequences into the expression vector 68
(ii) Production of mouse histone H2A.Z protein 69
(iii) Western analysis 70
(iv) The use of histone H2A.Z expression product 70

3.3 SUMMARY AND DISCUSSION 71

Chapter 4 Histone H2A.Z Gene " Knock-Out " by Homologous Recombination in Mouse Embryonic Stem Cells

4.1 INTRODUCTION 74

4.2 RESULTS 75

4.2.1 Electroporation of E14 ES cells with targeting vector pHZHR1 75
(i) Construction of pHZHR1 75
(ii) Electroporation and Southern analysis 76

4.2.2 Isolation of a mouse histone H2A.Z genomic clone derived from mouse strain 129 76
(i) Screening mouse genomic library III 77

(ii) Characterisation of the three positive lambda clones (λ1, λ2, λ10) 77

4.2.3 Electroporation of E14 ES cells with targeting vector pHZHR2 78
(i) Construction of pHZHR2 78
(ii) Electroporation of pHZHR2 into ES cells and Southern analysis 79

4.2.4 Homologous recombination in ES cells using pHZHR3 79
(i) Construction of pHZHR3 79
(ii) Isolation of an overlapping clone containing sequences downstream from clone λ1 80
(iii) Electroporation of pHZHR3 and Southern analysis 81
(iv) Verification 83

4.3 SUMMARY AND DISCUSSION 83

Chapter 5 Production of H2A.Z Null Mutants *In vitro* and *In vivo*

5.1 INTRODUCTION 85

5.2 RESULTS 86

5.2.1 Test for electroporation conditions resulting in single copy integration of introduced DNA 86

5.2.2 An attempt to create a homozygous H2A.Z gene "knock-out" ES cell line 87
(i) Construction of pHZHR4 87
(ii) Gene targeting using pHZHR4 88

5.2.3 Verification of the double allele H2A.Z gene "knock-out" experiment 91
(i) One allele gene targeting using pHZHR4 91
(ii) Second allele gene targeting using pHZHR3 92

5.2.4 Generation of H2A.Z deficient mice 93
(i) Blastocyst injection and production of chimaeric mice 93
(ii) Germ line chimaera testing 94
(iii) Detection of the disrupted H2A.Z gene in F1 mice 94

5.3 DISCUSSION 95

Chapter 6 Final Discussion

6.1 CONCLUSIONS AND DISCUSSION 98
6.1.1 Isolation and characterisation of the mouse histone H2A.Z gene 98
6.1.2 Histone H2A.Z gene targeting in mouse embryonic stem cells 98

6.2 WHAT IS THE FUNCTION OF HISTONE H2A.Z? 100

6.3 FUTURE WORK 101

Appendix: Towards Histone H1^{o} Gene "Knock-Out" in Mouse Embryonic Stem Cells

A.1 INTRODUCTION 103

A.2 RESULTS 104

A.2.1 Electroporation of E14 ES cells with targeting vector pH1^{o}HR1 104
(i) Construction of pH1^{o}HR1 104
(ii) Electroporation of pH1^{o}HR1 into ES cells and Southern analysis 106

A.3 DISCUSSION 106

A.4 ISOLATION OF A MOUSE HISTONE H1^{o} GENOMIC CLONE DERIVED FROM MOUSE STRAIN 129 107

A.4.1 Screening mouse genomic library III 107

A.4.2 Characterisation of the two positive lambda clones (λ2 and λ4) 108

A.5 STRATEGIES FOR GENE TARGETING USING H1^{o} ISOGENIC DNA .. 109

A.5.1 Construction of pH1^{o}HR2 109

A.5.2 Southern analysis of mouse genomic DNA isolated from E14 ES cells 110

A.5.3 Strategy for gene targeting using pH1^{o}HR2 DNA 110

A.6 SUMMARY 111

A.7 FUTURE WORK 111

Bibliography 113

Chapter 1

Introduction

1.1 Introduction

The work presented in this thesis describes studies of the mouse histone H2A variant, H2A.Z. The overall aim was to investigate the biological significance of this protein. Gene targeting by homologous recombination in mouse embryonic stem (ES) cells has been used as a tool for this investigation.

This chapter summarises the general background relevant to these studies. These include the structure and function of chromatin, transcriptionally active chromatin, histone proteins, histone variants, and in particular histone H2A.Z. Gene targeting is also briefly described.

1.2 Chromatin

The DNA in the eukaryotic nucleus is packaged in association with other macromolecules, primarily the histones and nonhistone chromosomal proteins, generating a complex referred to as chromatin. Chromatin is a highly dynamic complex that continuously changes its composition and conformation to accommodate different stages of genetic activity, e.g., transcription and replication. This dynamic behaviour of chromatin plays a key role in the regulation of gene expression. The structure of chromatin and transcriptionally active chromatin will be discussed below.

1.2.1 Nucleosome structure

The nucleosome is the fundamental unit of chromatin. A nucleosome core particle consists of 146 bp of DNA wrapped around a histone octamer, containing two each of the histones H2A, H2B, H3 and H4 (reviewed by McGhee and Felsenfeld, 1980; van Holde, 1988). Histone H1 is located in the region where the DNA enters and exits the core particle and is able to link two turns of DNA

around the histone octamer.

The structure of the nucleosome core particle was previously determined by using an electron microscope and X-ray diffraction of the octamer at 22 Å resolution (Finch *et al.*, 1977; Klug *et al.*, 1980) and at 7 Å resolution (Richmond *et al.*, 1984). The results demonstrated that the core particle is a flattened wedge-shaped disc approximately 11 nm in diameter and 5.7 nm in height. DNA (146 bp) is wrapped in 1.75 turns of a left-handed superhelix around the histone core. Recently, the structure of the histone octamer has been determined by X-ray crystallography to a resolution of 3.1 Å (Arents *et al.*, 1991). The histone octamer is a tripartite assembly in which a centrally located $(H3\text{-}H4)_2$ tetramer is flanked by two H2A-H2B dimers. The area of interaction between the H2A-H2B dimers and the $(H3\text{-}H4)_2$ tetramer is more extensive than the interface between the two H3-H4 dimers that make up the $(H3\text{-}H4)_2$ tetramer. The two H2A-H2B dimers and the $(H3\text{-}H4)_2$ tetramer create a protein spool consisting of a left-handed protein superhelix with an apparent pitch of 28 Å.

1.2.2 Higher order structure

The nucleosome is only the first level of DNA packaging inside the cell. At the second level of chromatin organisation is a 10 nm fibre generated by folding of the internucleosomal linker DNA into the bead-like particles. This structure is then coiled into a solenoid 30-nm diameter fibre with six nucleosomes per solenoid turn (Finch and Klug, 1976; Felsenfeld and McGhee, 1986; Butler, 1988). This higher structure is dependent on the presence of histone H1 (Thoma *et al.*, 1979) which is located inside of the 30-nm fibre (Graziano *et al.*, 1994).

At the third level of hierarchical structure, seen in both interphase and metaphase chromosomes, the 30-nm chromatin fibre appears to be folded into loops, or domains (Benyajati and Worcel, 1976; Igo-Kemenes and Zachau, 1977;

Paulson and Laemmli, 1977). These domain structures are believed to be anchored by specific nonhistone proteins to the supporting nuclear structure designated as nuclear matrix or chromosomal scaffold (Mirkovitch *et al.*, 1984). The domains are suggested to be supercoiled stretches of chromatin comprising from 5 to 100 kb of DNA. Domains appear to be essential elements of higher order chromatin structure and may be related to the units of replication and transcription (reviewed by Gasser and Laemmli, 1987; Georgiev *et al.*, 1991; Zlatanova and van Holde, 1992).

1.2.3 Active chromatin

The major functions of chromosomes are replication and transcription. Active chromatin is defined as the regions of the chromosome which are actively transcribed. Studies have revealed that active chromatin has several characteristics which are different from the chromatin structure described above. The features of active chromatin and factors that are involved and associated with transcriptionally active chromatin will be briefly summarised below. There have been several recent reviews which cover a number of aspects of active chromatin (Elgin, 1990; Grunstein, 1990a, 1990b; Svaren and Chalkley, 1990; Kornberg and Lorch, 1991; Felsenfeld, 1992; Kornberg and Lorch, 1992; Wolffe, 1992; Svaren and Hörz, 1993).

Nucleosomes are present on both transcriptionally active and inactive genes but actively transcribed genes are apparently packaged in an altered nucleosome structure and are associated with domains of chromatin that are less condensed or more open than inactive domains (Reeves, 1988). The open, decondensed structure associated with transcriptionally active chromatin is detectable by a general increase in the accessibility to DNA of a variety of DNA-modifying agents including DNase I (reviewed by Eissenberg *et al.*, 1985). Several factors are responsible for the generation of an open conformation.

In addition to the general nuclease-sensitivity of active chromatin, actively transcribed genes often have DNase I hypersensitive sites. These sites are found within the DNase I sensitive region and have been demonstrated in a cell- and tissue-specific manner near the ends of most active or potentially active genes (reviewed by Elgin *et al.*, 1988; Gross and Garrard, 1988). The structure and function of DNase I-hypersensitive sites are not fully understood. They are thought to be nucleosome-free regions available for binding of trans-acting factors to facilitate initiation of transcription (Svaren and Chalkley, 1990).

The transient displacement of the histone octamer as a result of RNA polymerase activity may contribute to the altered conformation of chromatin within actively transcribed regions (Lorch *et al.*, 1987; Pfaffle *et al.*, 1990; Clark and Felsenfeld, 1991). RNA polymerase-mediated changes in chromatin structure may be involved in displacement of histones from the path of transcription to another part of the template (Clark and Felsenfeld, 1991, 1992).

Nucleosomes sometimes adopt precise, well-defined locations with respect to specific DNA sequence (i.e. these nucleosome are said to be positioned). The precise positioning of nucleosomes around eukaryotic promoters and enhancers is important for gene transcription (reviewed by Simpson, 1991; Felsenfeld, 1992; Wolffe, 1994). Studies have demonstrated that nucleosome positioning allows *trans*-acting factors to gain access to the recognition elements which is the initial step for gene activation process (Straka and Hörz, 1991, Bresnick *et al.*, 1992; Lee and Archer, 1994). In addition, nucleosome positioning can also facilitate transcription process by bringing together distant regulatory elements (Lu *et al.*, 1993; Schild *et al.*, 1993).

The presence of linker histone and higher order structure on actively transcribed chromatin have been observed by several investigators (reviewed by Garrard, 1991). Kamakaka and Thomas (1990) have provided evidence that histone H1 is present in active chromatin but depleted in its amount and bound

in a somewhat different manner, leading to a more open chromatin conformation.

Histone post-translational modifications (see section 1.3.2), in particularly histone acetylation, and histone structural variants, may be involved in active chromatin by generating structurally and functionally heterogeneous nucleosomes and thereby providing a means for altering chromatin structure.

Specific base modifications have been proposed as a potential mechanism for altering the transcriptional template capacity of eukaryotic genes. About 2 to 4% of the cytosine residues in the DNA of higher eukaryotes are methylated. The usual site of methylation is the dinucleotide CpG. The demethylation of specific CpG dinucleotides correlates with active transcription (Bird, 1986). Studies have shown that DNA methylation can cause gene silencing either by direct interference with transcription factor binding or indirectly by promoting the formation of inactive chromatin (Watt and Molloy, 1988; Holliday and Ho, 1991; Boyes and Bird, 1991). It remains unclear whether inhibition of transcription by DNA methylation is direct or indirect effect (reviewed by Bird, 1992; Tate and Bird, 1993).

Alteration in DNA structure and conformation has also been postulated to have a role in control of transcription process. B-DNA is the most commonly form of DNA structure. Non-B-DNA structures such as Z-DNA, triplex DNA and cruciform DNA can also exist, and these structures can be formed under physiological conditions and are stabilised by unconstrained negative supercoiling. The mechanisms by which the non-B-DNA structure effects the transcription process remain under speculation (reviewed by van Holde and Zlatanova, 1994).

1.3 Histones

1.3.1 Histone proteins

Histones are the fundamental structural proteins of eukaryotic chromatin. They are small polypeptides with 100 to 200 amino acid residues which do not contain tryptophan, but are rich in lysine and/or arginine (Isenberg, 1979; Wu *et al.*, 1986). The histones can be classified into five classes. Histones H2A, H2B, H3 and H4 interact with DNA to form the nucleosome core. Histones H1 or linker histones play a key role in the compaction of nucleosomes into a higher-order structure (section 1.2.2). The primary structure of histones have been highly conserved throughout evolution (Isenberg, 1979) suggesting that their functions may be identical in all eukaryotes. Histones H3 and H4 are among the most conserved proteins known in evolution. Histones H2A and H2B are more divergent than H3 and H4 (Thatcher and Gorovsky, 1994) while H1 is the most variable of the histones.

The primary structure of histones is known. Each of the four core histones contains three domains: a long extended hydrophilic amino-terminal tail, containing a number of positively charged amino acids; a globular, hydrophobic core with a high α-helix content; and a very short hydrophilic carboxy-terminal tail. The hydrophobic cores are involved in both the histone-histone and the histone-DNA interactions that allow nucleosome assembly and stability *in vitro* (McGhee and Felsenfeld, 1980; van Holde, 1988). Enzymatic removal of the N-terminal tails has a minor effect on stability or assembly of the core particle (Whitlock and Stein, 1978; Ausio *et al.*, 1989), which suggests that their principle function is likely to be in histone and protein (Johnson *et al.*, 1990) or histone and DNA (Hill and Thomas, 1990) interactions outside of the core particle. The basic N-terminal tails are the sites of most of the post-translational histone modifications, including acetylation, methylation and phosphorylation.

Histone H1s have a tripartite structure consisting of a globular central domain flanked by lysine-rich, highly charged amino-terminal and carboxy-terminal tails (van Holde, 1988). The central hydrophobic domain appears to interact with the nucleosome core particle while the positively charged tails interact with the linker DNA and are important in chromatin condensation (Allan *et al.*, 1980; Staynov and Crane-Robinson, 1988).

1.3.2 Post-translational modifications of histones

Histones can be post-translationally modified in several ways. These include acetylation, phosphorylation, ubiquitination, methylation and ADP-ribosylation (reviewed by Matthews, 1988; Sterner *et al.*, 1989). Histone modifications provide means for altering the chromatin structure. Three of these modifications will be briefly discussed below.

(i) Acetylation

The core histones can be modified by reversible, post-translational acetylation of specific lysine residues within the amino-terminal domains. The level of acetylation is the result of a balance between acetylating and deacetylating enzymes, and inhibition of the latter with sodium butyrate results in a shift to more highly acetylated isoforms (Boffa *et al.*, 1978; Candido *et al.*, 1978; Sealy and Chalkley, 1978). Addition of an acetyl group to the lysine residue of core histones neutralises their positive charge. This modification weakens the electrostatic interaction of the histones with DNA and may facilitate the unfolding of the chromatin fibre (McGhee and Felsenfeld, 1980).

Histone acetylation has been correlated with all aspects of DNA processing in eukaryotes: replication, transcription, spermatogenesis (reviewed by Allfrey, 1980; Csordas, 1990). Histone acetylation, in particular highly acetylated forms of H3 and H4, have been strongly correlated with elevated levels of transcription

(Allegra *et al.*, 1987; Johnson *et al.*, 1987; Chan *et al.*, 1988; Hebbes, *et al.*, 1988). Recent experiments of Lee *et al.*, (1993) have demonstrated that increasing the level of histone acetylation allows TFIIIA transcription factor free access to a 5S RNA gene in a nucleosome. Acetylation is presumed to release the N-terminal tail of the histones from the DNA, and thereby either make the nucleosomal DNA accessible, or induce a conformational change in the nucleosome which promotes TFIIIA binding. Acetylation of core histones has been found to alter the capacity of the H1 histones to form compact higher order chromatin structures (Ridsdale *et al.*, 1990). This will also facilitate trans-acting factor and polymerase access to DNA in the nucleosome.

Several other findings support the role of histone acetylation in the transcription process. In the ciliated protozoan *Tetrahymena,* studies using antibodies against acetylated synthetic peptides of histone H4 demonstrated that histone acetylation is specific to the transcriptionally active macronuclei at all stages of the life cycle except when the transcriptionally inert micronuclei undergo periods of rapid replication and chromatin assembly (Lin *et al.*, 1989). Studies in yeast have found that the silent mating type loci are specifically associated with hypoacetylated core histones (Braunstein *et al.*, 1993).

(ii) Phosphorylation

Phosphorylation of linker histones has been associated with progression through the cell cycle and specifically with the process of chromatin condensation. Phosphorylation occurs during S-phase, during mitosis, and as a response to hormonal stimulation. In each case, different and highly specific serine and threonine residues are phosphorylated (Isenberg 1979). Each molecule becomes phosphorylated at multiple sites (3 to 6 sites/H1) in both amino-terminal and carboxy-terminal tails (Gurley *et al.*, 1978). Phosphorylation of the tails will decrease their positive charges and the interaction of H1 with DNA will

be weakened. It has also been shown that the amino acid sequence SPKK in H1, the H1-growth associated kinase phosphorylation site, has a specific interaction with AT-rich DNA segments (Churchill and Suzuki, 1989). The nature of this interaction is not understood, but it is expected that phosphorylation of these sites would probably abolish their interaction with DNA. It has been suggested that H1 phosphorylation is not in itself sufficient to promote chromosome condensation (Krystal and Poccia *et al.*, 1981). Therefore, it is possible that H1 phosphorylation acts to facilitate the dissociation of the protein from chromatin, thereby allowing access to DNA of the nonhistone proteins that might directly mediate chromatin condensation (Jerzmanowski and Cole, 1990).

(iii) Ubiquitination

Histone H2A and H2B can be combined with a small, highly conserved protein, ubiquitin (Goldstein *et al.*, 1975; Hershko, 1983), at their carboxy-terminal tails to form proteins uH2A and uH2B, respectively. Ubiquitin is covalently attached to H2A and H2B by an isopeptide bond between the carboxy-terminus of ubiquitin and the ε-amino group of the target lysine side chain (Goldknopf and Busch, 1977). All mammalian H2A variants can be modified by ubiquitination (West and Bonner, 1980).

The function of ubiquitination is not known. Conjugation of ubiquitin to proteins in the cytoplasm appears to be necessary for intracellular protein degradation (Ciechanover *et al.*, 1984; Hershko *et al.*, 1984; Jentsch, 1992). Several indirect observations suggested that ubiquitinated histones might be preferentially enriched in transcriptionally active regions of chromatin. In *Drosophila*, uH2A has been found enriched in the transcribed chromosomal regions and absent from nontranscribed satellite chromatin (Levinger and Varshavsky, 1982). It has also been reported that fractionated nucleosomes containing active DNA sequences are enriched in uH2A and particularly in

uH2B (Nickel *et al.*, 1989). It has been suggested that histone ubiquitination in transcribed regions may prevent the formation of higher-order chromatin structure by modifying nucleosome-nucleosome interactions (Levinger and Varshavsky, 1982). Variations in the level of ubiquitinated histone appears to be cell cycle dependent (Matsui, *et al.*, 1979; Mueller *et al.*, 1985), since the ubiquitinated H2A is absent during mitosis but ubiquitin is bound again to H2A when the cell enters G1 phase.

1.3.3 Histone structural variants

Although the primary sequences of histones have been conserved throughout evolution, variations in primary structure do occur in all types of histones. They are referred to as subtypes or variants.

Within a single species, non-allelic histone variants exist (reviewed by Isenberg, 1979; Wu *et al.*, 1986). The structural variants of a particular histone class can be differentially expressed during development, during the cell cycle or in specific cell types.

The best studied example of developmental regulation of histone variant expression is found in the sea urchin, where different subtypes of histones are expressed during embryogenesis (reviewed by Maxson *et al.*, 1983; Romano, 1992).

Histone variants can be grouped into four classes on the basis of their relationship to DNA synthesis (Zweidler, 1984; Old and Woodland, 1984):

(i) Replication-dependent variants: These are induced at the start of DNA synthesis and are repressed at the end.

(ii) Partially replication-dependent variants: These are expressed at the start of DNA synthesis, but are not completely repressed at the end of S phase.

(iii) Replication-independent variants: These are constitutively expressed during the cell cycle or in nondividing cells, e.g., $H1^{o}$

(iv) Minor histones: These variants are replication-independent and are present in small amounts in somatic cells, e. g., H2A.X, H2A.Z.

Tissue-specific histone variants have also been described. In birds, histone H5, a variant of H1, is specific to nucleated red blood cells (Neelin *et al.*, 1964). In mammals, there are sperm-specific histone variants which are synthesised in spermatocytes, but are not expressed in other cell types (Zweidler, 1984).

The existence of histone variants, whose synthesis is differentially regulated during development and at different times, suggests the possibility that nucleosomes containing different variants may perform distinct functions.

1.3.4 Histone H1 variants

Histone H1 is the most variable of all histones. There are several distinct subtypes with tissue-specific and species-specific differences in amino acid sequence.

As noted above, histone H5 is an extreme variant of the H1 linker histone and is found in the nucleated erythrocytes of birds (Neelin *et al.*, 1964; Aviles *et al.*, 1978). Histone H5 accumulates as a major DNA-associated protein during erythrocyte maturation (Appels and Wells, 1972). It replaces the majority of H1 molecules on chromatin and is correlated with chromatin condensation and a decrease in transcription and replication (Weintraub, 1978; Appels *et al*, 1972). Studies using microinjection of H5 into rat myoblasts (Bergman *et al.*, 1988) or transient expression of H5 sequence in transfected rat sarcoma cells have shown that overexpression of H5 inhibits DNA replication (Sun *et al.*, 1989; Sun *et al.*, 1990).

Histone H1° was originally found in mammalian tissues and is more abundant in non-proliferating tissues (Panyim and Chalkley, 1969). H1° shows closer similarity to H5 than it does to H1. H1° is similar to H5 with respect to sequence homology (Pehrson and Cole, 1981), structural features (Cary *et al.*, 1981)

and immunological cross-reactivity (Mura and Stollar, 1981). These results suggested that H1° and H5 histones might have a similar function, distinct from those of the other H1 variants. However, H1° is not tissue specific as is H5. It is found in several tissues (Eisen *et al.*, 1981; Lennox and Cohen, 1983). The level of histone H1° increases when cells undergo terminal differentiation *in vivo* (Varricchio, 1977; Lennox and Cohen, 1983, 1984; Pina *et al.*, 1984). Histone H1° appears to be hormonally regulated in those tissues which depend on hormones to maintain their differentiated functions (Gjerset *et al.*, 1982). H1° has been found preferentially associated with repressed genes (Roche *et al.*, 1985). H1° expression is responsive to induction by agents that induce differentiation (Alonso *et al.*, 1988; Rousseau *et al.*, 1991). These results lead to the proposal that H1° may be involved in the differentiation process.

In an attempt to explore the function of histone H1° in mammals, gene targeting by homologous recombination in mouse embryonic stem cells was also initiated during the course of this work (described in Appendix).

1.3.5 Histone H2A variants

Histone H2A has the most variant forms of the core histones. Mammalian cells contain four histone H2A isoprotein species, called H2A.1, H2A.2, H2A.Z and H2A.X (West and Bonner, 1980; West and Bonner, 1983). The former two are the predominant forms and are synthesised in concert with DNA replication, while the latter two are minor forms and are synthesised throughout the cell cycle (Wu and Bonner, 1981).

Histone H2A.X and H2A.Z represent 5 to 20% of the total histone H2A protein. They were shown to be H2A variants by the following characteristics: (i) the arginine/lysine ratio of H2A.X and H2A.Z are typical of H2A and different from the other core histones, (ii) they contain the conserved H2A sequence AGLQFPVGR, known as the H2A box, (iii) they can be modified by the covalent

attachment of ubiquitin (Wu *et al.*, 1986).

Human histone H2A.X cDNA has been isolated and characterised by Mannironi *et al.* (1989). The deduced amino acid sequence shows that it contains 142 residues, 13 more than human H2A.1. The H2A.X sequence is almost identical to that of human H2A.1 up to amino acid residue 120. However the carboxyl-terminal region beyond residue 120 is unrelated to any known sequence in vertebrate histone H2A, but it contains a sequence homologous with those of several species of lower eukaryotes, e.g., *Saccharomyces cerevisiae* H2A.1 and H2A.2 (Choe, *et al.*, 1982), *Aspergillus nidulans* H2A (May and Morris, 1987), *Tetrahymena* H2A.1 and *Schizosaccharomyces pombe* H2A (Matsumoto and Yanagida, 1985). It has been hypothesised that this carboxyl-terminal sequence homology has been maintained during evolution because it participates in an essential cellular function (Mannironi *et al.*, 1989). The biological role of histone H2A.X is unknown, but it has been found to be enriched in active chromatin (Bhatnagar *et al.*, 1984; Huang *et al.*, 1986).

1.4 Histone H2A.Z

There is a potential for confusion due to the naming of the variant H2A.Z gene and its protein. The nomenclature of the H2A.Z gene from different species is summarised below:

mammals	H2A.Z, M1
chicken	H2A.F
sea urchin	H2A.F/Z
Drosophila	H2A.2, H2AvD
Tetrahymena	hv1

Histone H2A.Z protein was initially characterised from histone extracts of mouse nucleosomes. Subsequently, it has been found in all cell types and species

studied where it represents 5 to 10 % of the total H2A (West and Bonner, 1980). A chicken H2A.F cDNA clone which was later shown to code for the H2A.Z protein, was originally isolated and sequenced in this laboratory (Harvey *et al.*, 1983). Counterparts have been cloned from sea urchin (Ernst *et al.*, 1987), *Tetrahymena* (White *et al.*, 1988), *Drosophila* (van Daal *et al.*, 1988), rat, cow, and human (Hatch and Bonner, 1988; Hatch and Bonner, 1990).

Histone H2A.Z is only about 60% identical in amino acid sequence to major H2As within the same species. Comparison of the derived protein sequences of H2A.Z in different species shows that H2A.Z is highly conserved across species (85-100% identity) and is more conserved than major H2As (White *et al.*, 1988; van Daal *et al.*, 1990; Thatcher and Gorovsky, 1994). These results indicate that the H2A.Z variants have evolved differently than the major H2As, are under different selective pressures and therefore, histone H2A.Z variants and the major H2As must have distinct and important functions.

1.4.1 Post-translational modification of histone H2A.Z

H2A.Z is ubiquitinated as are the other H2As. It is not phosphorylated and lacks the N-terminal serine residue that is normally phosphorylated in H2A subtypes. Its acetylation pattern is different from that of H2A.1 (Pantazis and Bonner, 1981).

1.4.2 Histone H2A.Z gene

The gene coding for H2A.Z has unusual features, that are summarised below.

H2A.Z gene organisation is different from that of the other histone genes. It exists as a single copy per haploid genome. The genes encoding the major core and histone H1 proteins are typically present in multiple copies (reviewed by Hentschel and Birnstiel, 1981; Maxson *et al.*, 1983; Stein *et al.*, 1984; Hnilica *et al.*,

1989). The genes coding for H2A.Z isolated from chicken (Harvey *et al.*, 1983), human (Hatch and Bonner, 1990), *Tetrahymena* (White *et al.*, 1988) and *Drosophila* (van Daal *et al.*, 1988) have been shown to appear as a single copy. The exception is sea urchin which has four copies of the gene (Ernst *et al.*, 1987). The presence of several pseudogenes in the human genome has been reported by Hatch and Bonner (1990).

The H2A.Z gene contains introns. Histone genes typically lack introns. This feature is not unique to the variant H2A.Z gene since the gene coding for chicken histone H3.3 also contains introns (Engel *et al.*, 1982). The gene for hv1 in *Tetrahymena* and H2AvD in *Drosophila* have two and three introns respectively (van Daal *et al.*, 1990). The gene for H2A.F in chicken (Dalton *et al.*, 1989) has four introns as does the gene for H2A.Z in human (Hatch and Bonner, 1990).

H2A.Z genes are expressed throughout the cell cycle and their mRNAs are polyadenylated. The genes encoding the major core and histone H1 proteins are expressed primarily in concert with DNA replication, producing mRNAs with a stem loop structure (reviewed by Schümperli, 1988). The conserved polyadenylation signal in the H2AvD gene is AATAA. For the H2A.F gene, the polyadenylation signal is GATAAA and a sequence capable of forming a stable hairpin loop is found in the 3' untranslated region (Dalton *et al.*, 1989). In mammals, cDNAs from rat, cow and human H2A.Z have been characterised (Hatch and Bonner, 1988). All three have a polyadenylation signal, AATAAA, and a potential stem loop-forming sequence.

1.4.3 Possible biological function of H2A.Z

The high degree of conservation of H2A.Z throughout evolution suggests that this protein must have an essential function. The biological significance of histone H2A.Z is still unclear. There are many observations that indicate a role

for this protein in association with transcriptionally active chromatin.

Vegetative cells of the ciliated protozoan *Tetrahymena thermophila* contain two types of nuclei, the transcriptionally active macronucleus and the transcriptionally inert micronucleus. These nuclei are products of a single postzygotic division during conjugation. In this organism, hv1 is found exclusively in the transcriptionally active macronucleus and is absent from the transcriptionally inert micronucleus (Allis *et al.*, 1980). Recent studies have suggested that hv1 is required for the formation of active chromatin rather than appearing secondarily as a consequence of transcriptional activity (Stargell *et al.*, 1993). Antibodies specific to hv1 stain nucleoli in mammalian cells, presumably in the highly transcriptionally active rDNA chromatin (Allis *et al.*, 1982). In addition, hv1 has been found to be an essential gene in *Tetrahymena* (Liu and Gorovsky, unpublished observation, cited in Thatcher and Gorovsky, 1994).

In the sea urchin, H2A.F/Z mRNA is found in all embryonic stages and adult tissues tested (Ernst *et al.*, 1987), in contrast to the developmental class switching shown for the major H2As (Newrock *et al.*, 1978). *In situ* hybridization studies demonstrate that maternal H2A.F/Z mRNA is stored in the egg cytoplasm and is present at equal levels in all cells of the mesenchyme blastula-stage embryo, suggesting that H2A.F/Z is not coordinately regulated with DNA synthesis (McIsaac *et al.*, 1992).

The distribution of chicken histone variant H2A.F has been investigated in this laboratory using Western blot analysis and immunofluorescence technique (Whiting, 1988). Western blots were performed on histone extracts from various adult organs, and from embryonic tissue from days 4 to 10. The results showed that H2A.F is present in all tissues examined and the levels of H2A.F do not vary in any of the stages or tissues studied. The level of H2A.F in histone extracts from transcriptionally-inert mature erythrocytes and transcriptionally active early chicken erythroid cells (cultured AEV ts34 cells)

were also analysed by Western blots. The data showed that, in these cells of the same lineage, the H2A.F levels are the same, suggesting that no correlation exists between the H2A.F content and either the state of cell differentiation or transcriptional activity of the cell. However, immunofluorescence staining of sectioned chicken embryos showed that H2A.F is present in all cell types and is localised to the nucleolus and nuclear periphery, known sites of transcriptional activity.

In *Drosophila*, H2AvD transcript is abundantly expressed in early embryogenesis, but also remains present in all later developmental stages. Antibodies to H2AvD bind specifically to the interband region of polytene chromosomes (van Daal *et al.*, 1988). Significantly, deletion of the H2AvD gene is lethal, demonstrating an essential role for this protein in development. Null mutant flies only survive until the third larval instar stage, presumably utilising maternal supplies of H2AvD up to this stage (van Daal and Elgin, 1992).

* THIS MUTATION [l(8)9] IS A 311 bp EMS-INDUCED DELETION THAT REMOVES THE ENTIRE SECOND EXON AND SOME FLANKING SEQUENCE.

1.5 Gene targeting by homologous recombination

Gene targeting or gene "knock-out" has been made possible by joining two technologies: embryonic stem (ES) cells and homologous recombination. This technology allows the generation of cells or mice deficient for a defined gene (reviewed by Capecchi, 1989; Koller and Smithies, 1992; Capecchi, 1994). This technique has been used in this study to investigate the function of H2A.Z in mice. The background to this technique is outlined in this section, and strategies used to disrupt genes are described.

1.5.1 Embryonic stem (ES) cells

Murine embryonic stem (ES) cells are pluripotent stem cells derived from the inner cell mass of the mouse blastocyst 3.5 days post coitum (Evans and

Kaufman, 1981; Martin, 1981). ES cells can be cultured for long periods, subjected to manipulation *in vitro* and, upon return to the interior of another blastocyst, can contribute to the development of all cell lineages. Significantly, the ES cells can contribute to the development of germ cells, allowing the transmission of the (manipulated) genome of the ES cells. Therefore, ES cells provide a method for the introduction of specific mutations into mice. Several ES cell lines that are competent for introduction of mutations into the mouse germ line have been isolated. Most of them have been derived from the 129 mouse strain.

In order for ES cells to maintain germ line competence, they must be grown in specialised conditions to avoid differentiation and loss of totipotency. ES cells can be grown in an undifferentiated state *in vitro* either on monolayers of mitotically inactivated fibroblast cells (also referred to as feeder cells) (Evans and Kaufman, 1981; Martin, 1981), or on gelatinised tissue culture dishes in Buffalo rat liver (BRL) cell conditioned medium (Smith and Hooper, 1987), or in the presence of leukemia inhibitory factor (LIF, also termed differentiation inhibitory activity, DIA). LIF is a cytokine required for maintaining ES cells in vitro (Smith *et al.*, 1988; Williams *et al.*, 1988).

1.5.2 Gene targeting in ES cells

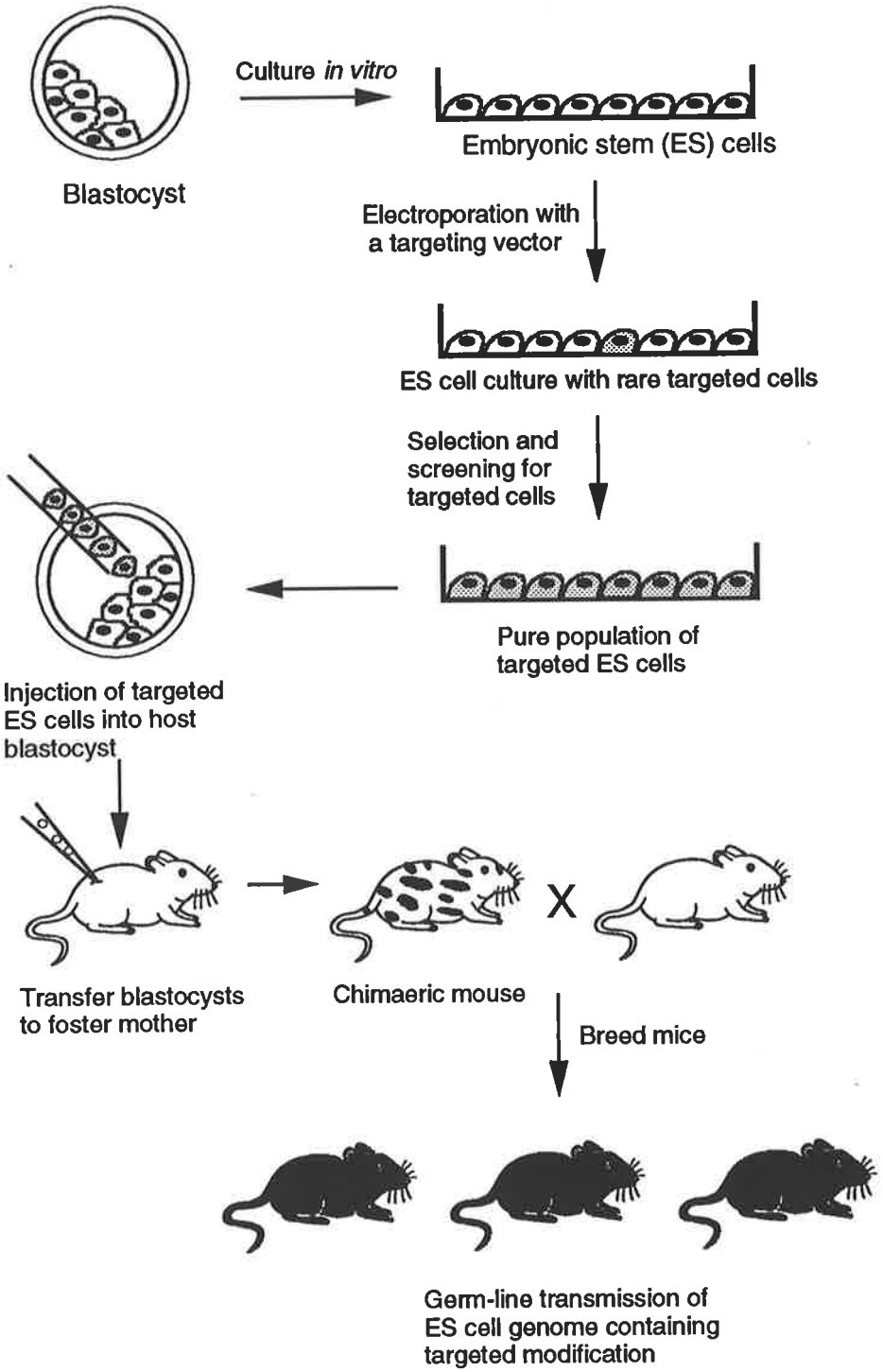
Gene targeting, the homologous recombination between chromosomal DNA sequences and introduced DNA sequences, provides the means for systematically altering the mammalian genome (Lin *et al.*, 1985; Smithies *et al.*, 1985; Thomas and Capecchi, 1986). Gene targeting has been widely used, particularly in mouse embryonic stem (ES) cells, to mutate many different loci. The phenotypic consequences of these genetic modifications can be assessed in ES cells or in mice.

The general procedure for gene targeting to generate mice with specific mutations is summarised in Fig. 1-1. A targeting vector containing the desired

Figure 1-1

General strategy for gene targeting experiment

ES cells are derived from inner cell mass of the mouse blastocyst 3.5 days post coitum. The targeting vector is introduced into ES cells by electroporation and targeted ES cell clones are selected. Targeted ES cells are introduced into a host blastocyst which has different genetic markers from the ES-cell derived blastocyst. In this example, the ES cells are derived from a mouse homozygous for the black coat colour allele and the host blastocyst is derived from an albino mouse. The blastocysts are transferred into a foster mother. The resulting chimaeric animal is born with mixed coat colours. Breeding of the chimaeric mouse to an albino mouse yields some black mice, indicating that the ES cells contributed to the formation of the germ line. These progeny are screened for the presence of the modified gene. The heterozygotes are crossed to generate homozygous animals carrying the modified gene.


Blastocyst
Culture *in vitro*
Embryonic stem (ES) cells
Electroporation with a targeting vector
ES cell culture with rare targeted cells
Selection and screening for targeted cells
Pure population of targeted ES cells
Injection of targeted ES cells into host blastocyst
Transfer blastocysts to foster mother
Chimaeric mouse
X
Breed mice
Germ-line transmission of ES cell genome containing targeted modification

alteration is introduced into ES cells by standard transfection methods. In some cells, the targeting vector pairs with the cognate chromosomal DNA sequence and the mutation is transferred to the genome by homologous recombination. Screening and enrichment procedures are used to identify ES cells in which the targeted event has occurred. A targeted ES cell will then have the newly mutated allele in the genome. Chimaeric mice are formed by microinjecting a small number (5-15) of targeted ES cells into the blastocoel of a blastocyst. The injected blastocyst is transferred into the uterus of a pseudopregnant mother. The resulting animal is a chimaera in that it is composed of cells derived from both the donor stem cells and the recipient blastocyst. If ES cells contribute to the germ line, interbreeding of heterozygous siblings yields animals homozygous for the desired mutation. However, this can be achieved only if modification of both alleles of the particular gene do not lead to embryonic lethality.

It is also possible to generate null mutant ES cells in which both alleles have been disrupted *in vitro* (te Riele *et al.*, 1990). The methods involve sequential targeting by two separate constructs using two different selectable markers. These mutant cells allow the investigation of questions to be addressed easier in cell lines than in the whole organism.

1.5.3 Targeting vector

A summary of gene targeting vector design is presented in Fig. 1-2. For more detailed descriptions see review by Hasty and Bradley (1993), Ramírez-Solis *et al.*, (1993).

In general, targeting vectors contain DNA homologous to the target locus, a positive selectable marker, and sometimes a negative selectable marker in a plasmid. The most commonly used selectable makers are the bacterial genes neomycin phosphotransferase (*neo*^{r}), hygromycin B phosphotransferase (*hyg*^{r}) and the mammalian gene hypoxanthine phosphoribosyl transferase (*hprt*). The

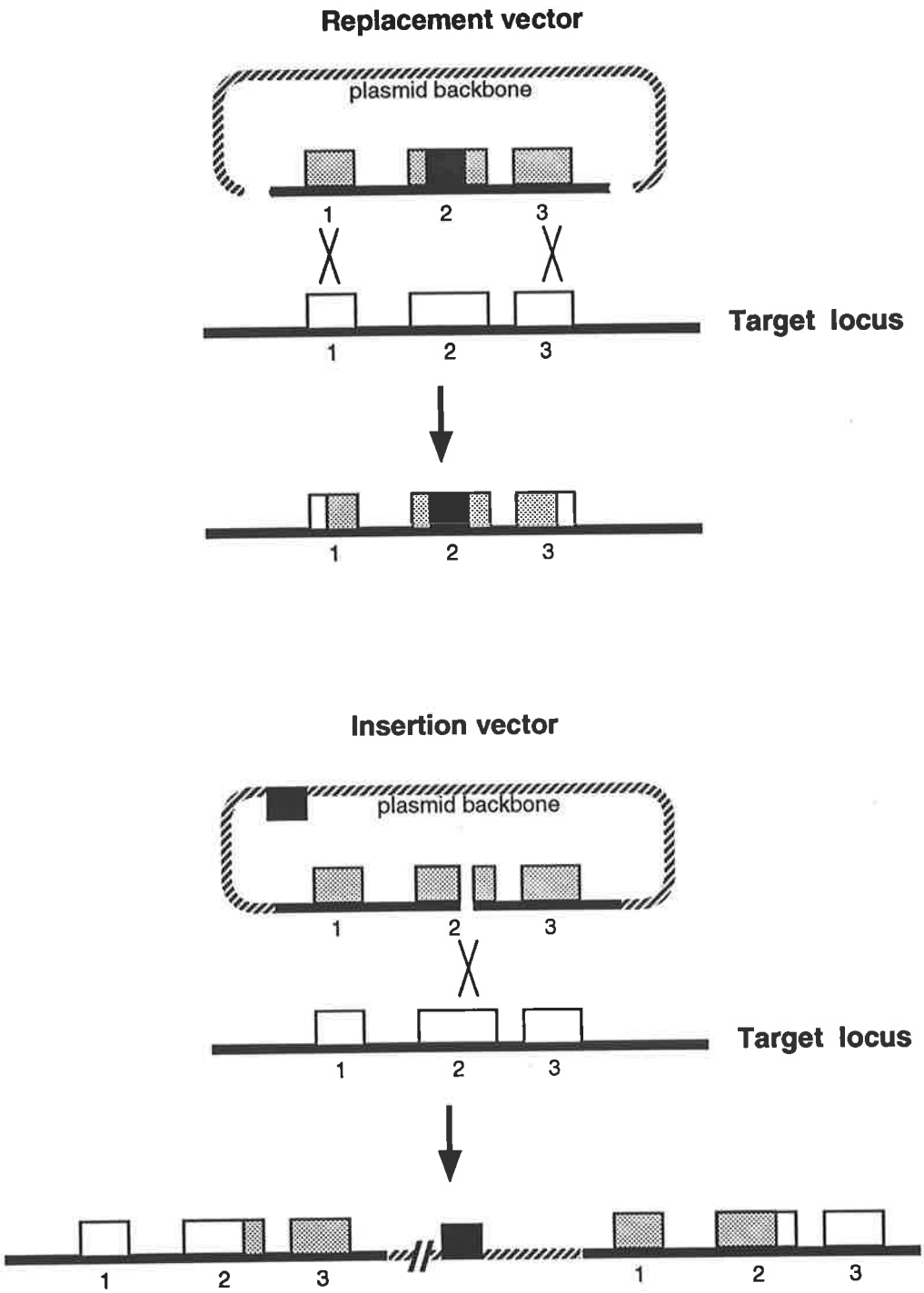
Figure 1-2

Two types of targeting vector used for homologous recombination

Two types of vector, replacement and insertion vectors, which are commonly used in gene targeting experiments are illustrated. Exons are boxed and numbered. The black box represents a positive selectable marker.

In a replacement vector, the positive selectable marker is inserted into a construct that contains DNA that is homologous to the target locus. The insertion shown disrupts exon 2 in the coding sequence. The replacement vector is linearised outside the target homology. Two crossover events result in the replacement of the endogenous sequences, with the corresponding sequence derived from the targeting vector.

In a insertion vector, the positive selectable marker is inserted outside the region of homology. The insertion vector is linearised within the target homology. A single recombination event results in an integration of the entire targeting vector.


Replacement vector
plasmid backbone
1
2
3
Target locus
1
2
3
1
2
3
Insertion vector
plasmid backbone
1
2
3
Target locus
1
2
3
1
2
3
1
2
3

positive selectable marker is necessary because it allows isolation of cells that have incorporated the targeting vector into the genome.

The targeting vector commonly used in gene targeting can be classified into two types, replacement type and insertion type (Fig. 1-2). They differ in configuration. After homologous recombination has occurred, they yield different integration products.

The targeting efficiency is greatly influenced by several factors including the length of homology. The targeting efficiency is defined as a porportion of transformed cells which have integrated the targeting construct by homologous recombination. The homology between the targeting vector and the chromosomal locus is critical to high targeting efficiency. Therefore it is important to construct the targeting vector from isogenic DNA derived either from ES cells or the mouse strain from which they were derived. van Deursen and Weiringa (1992) have demonstrated that targeting efficiency of the creatine kinase M gene using isogenic DNA vector is 25-fold more efficient than that using non-isogenic DNA.

1.5.4 Enrichment procedure

The majority of integration events are random. Several approaches have been used to enrich for homologous integration events. Mansour *et al.* (1988) used positive/negative selection (PNS) to target the *hprt* and *int*-2 genes. Their construct included homologous DNA and a *neo*^{r} expression cassette to allow positive selection of cells that had integrated the vector. External to the targeting homology, an herpes simplex virus-thymidine kinase (HSV-*tk*) cassette was added which was lost during homologous recombination. Cells which had integrated the HSV-*tk* cassette were killed using purine analogues such as acyclovir or gancyclovir. Mansour *et al.*, (1988) and Thomas and Capecchi (1990) have shown that they can obtain as much as 1000-fold or greater enrichment for

homologous recombination with PNS. Although similar levels of enrichment have not been observed by other investigators (Mann *et al.*, 1993), it is clear that PNS significantly enhances the selection process and the chances of finding targeting events at some loci (Johnson *et al.*, 1989; DeChiara *et al.*, 1990).

Other approaches that lower the background of non-homologous integration use targeting vectors containing a positive selectable marker that lack either a promoter or a polyadenylation signal. These selectable markers are expressed only after homologous recombination or if randomly integrated into another gene in an appropriate fashion. Although transcription of the targeted locus in ES cells is required for this approach to succeed, very low levels of transcription seem to be sufficient (Jeannotte *et al.*, 1991). Polyadenylation signal-less markers (Joyner *et al.*, 1989) require integration close to a genomic polyadenylation signal to stabilise the transcript sufficiently to confer antibiotic resistance. These approaches have been used successfully to enrich of about 5 to 100-fold for targeted clones (Schwartzberg *et al.*, 1990; Donehower *et al.*, 1992).

1.5.5 Screening assays for homologous recombinants

(i) Polymerase Chain Reaction

Unique DNA regions generated by homologous recombination can be detected by polymerase chain reaction (PCR) (Saiki *et al.*, 1985). PCR has been used to screen a population or a pool of ES cell clones and detect the ES cells in which the targeting events occurred (Zimmer and Gruss, 1989; Joyner *et al.*, 1989). The advantage of PCR is its sensitivity. Kim and Smithies (1988) reported identification of homologous recombination events in as few as 1 in a background of 10,000 nontargeted ES cells. The result obtained by PCR is not conclusive. Homologous recombinants identified by PCR are require confirmation by Southern analysis for correct integration.

(ii) Southern analysis

Southern blotting (Southern, 1975) can be used to detect a restriction enzyme digestion pattern characteristic of the modified genotype. This technique is suitable for the detection of single copy sequences and can therefore be used to screen clones. An example of Southern analysis is showed in Fig. 1-3. The probe used for the primary screening should be to a region external to the targeting vector to avoid detecting random integration events. Homologous recombinants identified in the primary screen can be further characterised using different enzyme digestions and probes. Complete characterisation of the 5' end and 3' end of the targeted locus will eliminate the possibility of incomplete homologous recombination or "pick-up" events. This involves a strand exchange of the vector with the target in the vector picking up some sequences which flank the target sequence. Subsequently, the vector dissociates from the target and may integrated at a random location to generate a partial third allele (Adair *et al.*, 1989). This will be discussed more fully in chapter 4.

1.5.6 Summary

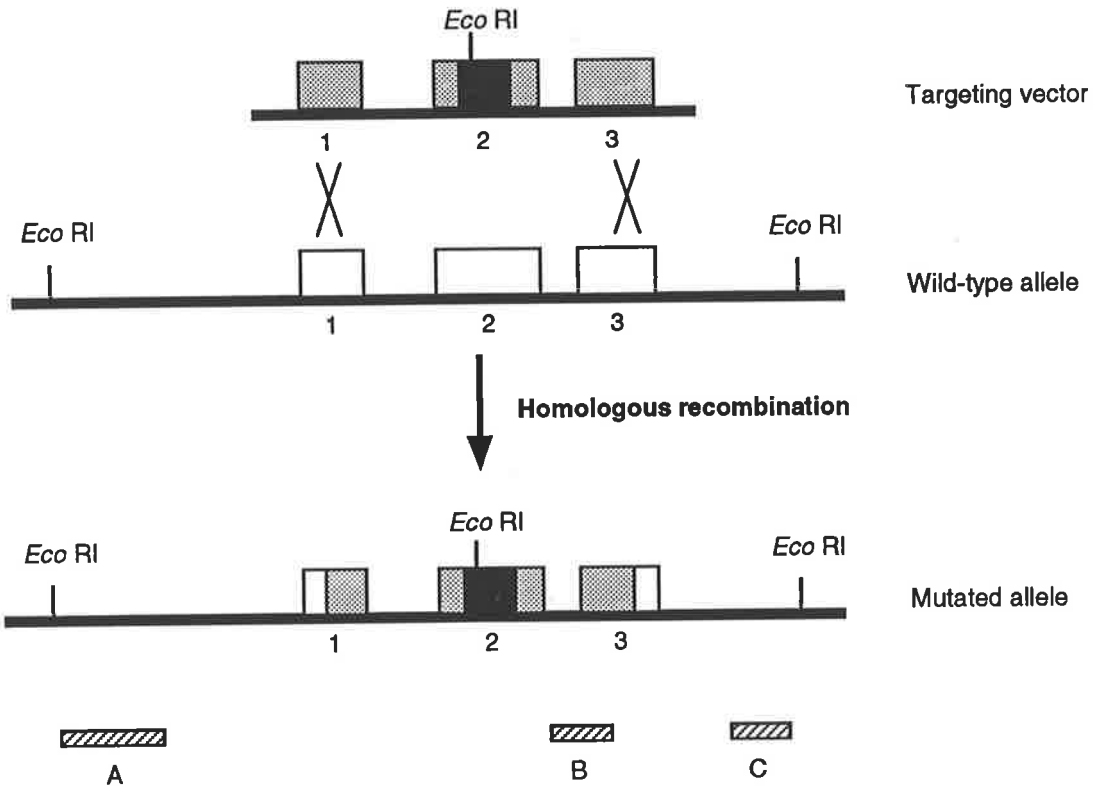
Gene targeting can be achieved in ES cells, via a number of strategies. The frequency of homologous recombination reported has been as high as 33% (te Riele *et al.*, 1992). Many kinds of change, to any gene in a genome can be introduced, and mouse with specific, predetermined genetic alterations can be produced. At present, this technology has an important application in basic research. Several animal models have been generated for investigating the biological functions of a number of genes involved in cell growth, differentiation, development and human genetic disorders (reviewed by Zimmer, 1992; Capecchi, 1994).

Figure 1-3

An example of Southern analysis for detection of homologous recombination events in ES cells

The replacement type targeting vector is shown at the top. Exons are boxed and numbered. The black box represents a positive selectable marker. The probes used in Southern analysis are shown in striped boxes. Probes A and C are external probes that do not contain any sequence within the targeting vector and can be used for the primary screening. Probe B is an internal probe containing the sequence present in the vector.

After homologous recombination, the endogenous gene is replaced with the corresponding sequence derived from the targeting vector. In this example *Eco* RI can be used to detect the mutated allele, which will have a smaller hybridising fragment when using an external probe for the primary screening. The targeted ES cell clones isolated by the initial screening require verification by several restriction enzyme digests and probing with the internal probe to determine the number of integration sites of the vector within the genome.


Eco RI
Targeting vector
1
2
3
Eco RI
Eco RI
Wild-type allele
1
2
3
Homologous recombination
Eco RI
Eco RI
Eco RI
Mutated allele
1
2
3
A
B
C

1.6 Aims of thesis

Work in this laboratory has focused on studies of histone genes, in particular histone variants H2A.Z and H5. The cDNA of chicken histone H2A.F was initially isolated and characterised in this laboratory. The functional significance of this histone variant is still unclear. Recent studies indicate an essential role of this variant during *Drosophila* development. However, the biological role of this protein has not been established.

Histone H2A.Z gene "knock-out" in mouse embryonic stem cells was used in these studies with an aim to create ES cell lines, and animals, lacking H2A.Z, allowing a functional investigation of this gene *in vitro* and *in vivo*.

The gene coding for mouse histone H2A.Z had not been cloned. Therefore the initial aim was to isolate and characterise the mouse histone H2A.Z gene.

Once the H2A.Z genomic clone was obtained, the second aim was to create homozygous null mutant ES cell lines and animals by gene targeting.

Chapter 2

Materials and Methods

2.1 Abbreviations

Abbreviations are as described in "Instructions to authors" (1994) *Biochem. J.* **297,** 1-15. In addition:

A_{xxx}	absorbance at xxx nm
APS	ammonium persulphate
BCIG	5-bromo-4-chloro-3-indolyl-β-D-galactoside
bisacrylamide	N, N'-methylene-bisacrylamide
BRL	Buffalo rat liver cell
BRL-CM	Buffalo rat liver cell conditioned medium
CAPS	3-(cyclohexylamino)-1-propanesulphonic acid
DEPC	diethylpyrocarbonate
DMEM	Dulbecco's modified Eagle's medium
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DTE	dithioerythritol
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetate
EGTA	ethyleneglycol bis(amino-ethyl ether)tetra-acetate
ES cells	embryonic stem cells
FBS	foetal bovine serum
FIAU	2'-fluoro-5'-iodo-1-β-D-arabinofuranosyluracil
G418	geneticin™, Gibco
HPLC	High performance liquid chromatography
IPTG	isopropyl-β-D-thio-galactopyranoside
LIF	leukaemia inhibitory factor
MT PBS	mouse tonicity phosphate-buffered saline
NBT	nitro blue tetrazolium
PBS	phosphate-buffered saline

PEG	polyethylene glycol
pfu	plaque forming unit
PSB	phage storage buffer
RSP	reverse sequencing primer
TEMED	N, N, N', N'-tetramethylethylenediamine
TMACl	tetra methyl ammonium chloride
USP	universal sequencing primer

2.2 Materials

2.2.1 General reagents and materials

Reagents used were generally of analytical reagent grade, or the highest available purity. Chemicals were obtained from a number of suppliers, the sources of the more important materials are listed below:

acrylamide, agarose (type I), ampicillin, BCIG, dNTPs, DTT, DTE, EDTA, IPTG, PEG 6000, salmon sperm DNA, SDS and spermidine: Sigma Chemical Co.

bisacrylamide, low melting point agarose, mix bed ion exchange resin AG 501-X8 (D), TEMED and Zeta-Probe blotting membrane : Bio-Rad Laboratories

CsCl: Boehringer Mannheim

dextran sulfate, Sepharose CL-6B: Pharmacia

GeneScreen™, PlaqueScreen™ : New England Nuclear

Immobilon-P: Millipore

methanol (HPLC grade) : Waters Associates Pty. Ltd

TMACl: Aldrich

urea (ultra-pure): Merck

2.2.2 Reagents for cell culture

The sources of the chemicals and reagents are as follows:

DMEM (cat. No 430-2100), Dulbecco's PBS (without Ca^{2+} and Mg^{2+}), FBS, G418 Sulfate, Trypsin-EDTA: Gibco

EGTA, gelatin, nonessential amino acid, mitomycin C: Sigma Chemical Co.

FIAU; Bristol-Myers Squibb

gentamycin: Delta West

hygromycin B: Calbiochem

LIF (ESGRO^{TM}) : Amrad (Australia)

2.2.3 Kits

Erase-a-base kit: Promega

Geneclean kit: Bio 101

Megaprime DNA labelling system kit: Amersham

Sequencing kit: Bresatec, United States Biochemical Corporation

2.2.4 Enzymes

Enzymes were obtained from the following sources:

Calf intestinal phosphatase: Sigma, Boehringer Mannheim

E. coli DNase I: Worthington Biochemical Corporation

E. coli DNA-polymerase I, Klenow fragment: Bresatec

Proteinase K: Boehringer Mannheim

Restriction Endonucleases: Boehringer Mannheim, New England Biolabs, Pharmacia and Bresatec

Ribonuclease A: Sigma

Taq DNA polymerase: Cetus Corporation

T4 DNA ligase: Boehringer Mannheim, Bresatec

T4 DNA polynucleotide kinase: Bresatec

2.2.5 Radiochemicals

α-^{32}P-dATP (specific activity, 3000 Ci/mmole), γ-^{32}P-dATP (specific activity, 4000 Ci/mmole) and α-^{35}S-dATP (specific activity, 1500 Ci/mmole) were purchased from Bresatec.

2.2.6 Bacterial strains

Escherichia coli :

JM101: *supE, thi, Δ (lac-pro AB),* [F' *lacI*^{q}ZΔM15, *traD36, proAB*] (Messing, 1979)

ED8799: *hsdS, metB7, supE,* (*gln*V)44, *supF,* (*tyrT*)58, Δ(*lacZ*)M15, r_K^- , m_K^- (gift from Dr. S. Clarke, Biotechnology Australia)

DH5α: *supE*44, *ΔlacU*169 (ø80 *lacZ*ΔM15), *hsdR*17, *recA*1, *endA*1, *gyrA*96 *thi*-1, *relA*1 (Sambrook *et al.*, 1989)

LE392: F$^-$ *supE*44, *supF*58, *hsdR*514 (r_K^-, m_K^-), *lacY*1 or Δ(*lacIZY*)6, *galK*2, *galT*22, *metB*1, *trpR*55, λ$^-$ (Sambrook *et al.*, 1989)

BL21: (DE3)pLysS, F$^-$, *ompT*, $hsdS_B$ (r_B^- , m_B^-) (Studier *et al.*, 1990)

2.2.7 Mouse genomic DNA libraries

(i) Mouse genomic library I (lambda Charon 4A):

a gift from E. Webb (The Walter and Eliza Hall Institute of Medical Research (WEHI), Melbourne, Australia), was prepared from DNA isolated from BALB/c mouse embryos. Partially *Eco* RI-digested DNA fragments were cloned into lambda Charon 4A vector.

(ii) Mouse genomic library II (lambda EMBL3 SP6/T7):

a gift from A.Dunn (Ludwig Institute for Cancer Research, Melbourne, Australia), was prepared from BALB/c adult mouse liver. Partially *Sau* 3A-

digested DNA fragments were cloned into *Bam* HI-digested lambda EMBL3 SP6/T7 vector.

(iii) Mouse genomic library III (lambda Fix II™, from Stratagene): a gift from R. Harvey (WEHI, Australia), was prepared from DNA isolated from D3 embryonic stem cells (derived from 129/Sv mouse).

2.2.8 Tissue culture cell Lines

(i) Mouse embryonic stem cell line E14 (derived from 129/Ola mouse) was kindly provided by A. Michalska with permission from M. Hooper (Department of Pathology, University of Medical School, Edinburgh, UK).

(ii) G418-resistant STO cells (Transformed mouse embryonic fibroblast cell line) (E. Robertson, Columbia University, New York) were kindly provided by R. Harvey, WEHI, Australia.

(iii) Hygromycin-resistant STO cells (L. Robb, WEHI, Australia) were kindly provided by R. Harvey, WEHI, Australia.

(iv) Buffalo rat liver cell line (BRL-3A, ATCC No. CRL 1442)

2.2.9 Bacterial growth media

Bacterial growth media were prepared with distilled and deionised water and sterilised by autoclaving, except heat labile reagents, which were filter sterilised.

All bacteria, except *E. coli* JM101, were grown in L-broth or L-agar plates. *E. coli* JM101 was grown in minimal medium, 2xYT broth and on minimal plus glucose plates. Ampicillin (50-100 µg/ml) was added where required for selection.

L-broth (Luria broth):	1% (w/v) amine A 0.5% (w/v) yeast extract 1% (w/v) NaCl, pH 7.0.
L-agar plates:	L-broth

1.5% (w/v) bacto-agar.

LMN-broth: L-broth
0.2% (w/v) maltose
10 mM $MgSO_4$.

LMN-agarose: LMN-broth
0.7% (w/v) agarose.

Minimal medium: 2.1% (w/v) K_2HPO_4
0.9% (w/v) KH_2PO_4
0.2% (w/v) $(NH_4)_2SO_4$
0.1% (w/v) tri-sodium citrate.

Minimal plate: Minimal medium
0.4% (w/v) glucose
0.0001% (w/v) thiamine
1.5% (w/v) bacto-agar.

2xYT broth: 1.6% (w/v) bacto-tryptone
1% (w/v) yeast extract
0.5% (w/v) NaCl, pH 7.0.

2.2.10 Tissue culture media

Incomplete ES medium: DMEM
3.7% sodium bicarbonate
50 µg/ml gentamycin
100 µM β–mercaptoethanol

Complete ES medium: 85% Incomplete ES medium
15% FBS
1000 U/ml LIF

2xFreezing medium: 80% FBS
20% DMSO

BRL-CM: BRL-3A was cultured in incomplete ES media until confluence. The conditioned medium was collected every 3 days and cell debris was removed by filtration. Fresh complete ES medium was added to the conditioned medium at the dilution of 2:3 before use with ES cells.

2.2.11 Buffers

Commonly used buffers were:

PBS:	7.5 mM Na_2HPO_4 2.5 mM $NaH_2PO_4.2H_2O$ 145 mM NaCl
PSB:	10 mM Tris-HCl pH 7.4 10 mM NaCl 10 mM $MgCl_2$
SSC:	150 mM NaCl 15 mM sodium citrate
TAE:	40 mM Tris-acetate pH 8.2 1 mM EDTA
TBE:	50 mM Tris-borate pH 8.3 1 mM EDTA
TE:	10 mM Tris-HCl pH 8.0 1 mM EDTA
Acrylamide load buffer:	50% (v/v) glycerol 50 mM Tris-HCl pH 7.4 5 mM EDTA 0.1% (w/v) Bromophenol blue 0.1% (w/v) Xylene cyanol
Agarose gel-loading buffer:	30% (v/v) glycerol 0.25% (w/v) bromphenol blue 0.25% (w/v) Xylene cyanol FF
Formamide load buffer:	80% (v/v) formamide 250 mM EDTA 0.1% (w/v) Bromophenol blue 0.1% (w/v) Xylene cyanol
RNA load buffer:	50% (v/v) glycerol 1 mM EDTA 0.1% (w/v) Bromophenol blue
2 x SDS load buffer:	10% (v/v) glycerol 375 mM Tris-HCl pH 8.8 5% (w/v) SDS

0.1% (w/v) Bromophenol blue
5% (v/v) β-mercaptoethanol
(added just prior to use)

2.2.12 Cloning vectors

pBluescript KS+, pBluescript SK+ (Stratagene), pUC19 (Yanisch-Perron *et al*, 1985) and M13mp18, M13mp19 (Yanisch-Perron *et al*, 1985): gift from A. Robins (Department of Biochemistry, University of Adelaide, Adelaide, Australia).

pTZ19R : Pharmacia

2.2.13 Recombinant DNA clones

(i) pBS14-2.9 (Wait, Hons. thesis 1990)

(ii) pBS17-3 (Wait, Hons. thesis 1990)

(iii) pgkHYG (te Riele *et al.*, 1990): a gift from H. te Riele (Division of Molecular Genetics of the Netherlands Cancer Institute, Amsterdams, The Netherlands).

(iv) pgkNEOpA: a gift from R. Harvey (WEHI, Australia).

(v) pgkTKpA: a gift from R. Harvey (WEHI, Australia).

(vi) pET8.H3 (T7 expression plasmid): a gift from R. Sturm.

(vii) pH1° (Mouse histone H1° genomic clone, Breuer *et al.*, 1989): a gift from A. Alonso (Deutsches Krebsforschungszentrum, Institute of Experimental Pathology, Heidelberg, Germany).

2.2.14 Molecular size markers

(i) <u>DNA markers</u>

DMW-S1: *Eco* RI-digested SPP1 phage DNA (0.38-7.84 kb), Bresatec

DMW-P1: *Hpa* II-digested pUC19 DNA (26-502 bp), Bresatec

DMW-L1: *Hin*dIII-digested λ DNA (0.56-23 kb), Bresatec

DMW-M1: Hi-Lo mixed DNA markers (0.13-23 kb), Bresatec

(ii) Protein markers

Prestained protein molecular weight standards, low range (3000-43000 daltons): BRL

2.2.15 Oligonucleotides

All oligonucleotides were synthesised by Bresatec using an Applied Biosystems Model 380B DNA synthesiser. The sequence of the oligonucleotides used in this work are as follows:

RSP:	5'	dAACAGCTATGACCATG 3'
USP:	5'	dGTAAAACGACGGCCAGT 3'
PHZ1:	5'	dTGGATTCTCTGATCAAAGCTACC 3'
PHZ2:	5'	dTGTGGATGTGTGGGATGACAC 3'
PHZseq1:	5'	dCGGAGAAGATAGAATTAC 3'
PHZseq2:	5'	dCGCACGCTCCCACTCGCCG 3'
PHZseq3:	5'	dGTTAAGATAAGATGTGCG 3'
PHZ.Nco:	5'	dCCCCATGGCTGGCGGTAAGGCT 3'
PHZ.Hd:	5'	dAAAAGCTTAAACAGTCTTCTGTTG 3'
H1°/RV:	5'	dTCTGCCTGGATATCAGCCACGATCA 3'
H1°/Xb:	5'	dCGTAGGGTTCTAGAGGTATTGCCTA 3'
H1°/9bp:	5'	dGCGCCACCATCTGTAGTCGACGAATTCCG ACCTGGGGCTGTTGGGAG 3'

2.3 DNA Methods

2.3.1 Isolation and purification of plasmid DNA

Plasmid DNA was isolated by the alkaline lysis method as described by Sambrook *et al.* (1989). Plasmid DNA was purified by CsCl gradient

ultracentrifugation.

A single bacterial colony containing the plasmid of interest was picked in 50 ml of L-broth plus appropriate antibiotic and grown overnight at 37°C with vigorous shaking. The cells were collected by centrifugation at 5,000 rpm (Sorvall, HB-4 rotor) for 10 minutes at 4°C. The cell pellet was resuspended in 3 ml of 50 mM Glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA pH8.0 and kept on ice for 5 minutes. The suspension was mixed with 6 ml of freshly prepared 0.2 N NaOH/1% (w/v) SDS and kept on ice for 10 minutes then 4.5 ml of 3 M potassium acetate was added and the mixture was kept on ice for a further 20 minutes. The chromosomal DNA and cellular debris were removed by centrifugation at 12,000 rpm (Sorvall, HB-4 rotor) for 15 minutes at 4°C. The supernatant was collected and DNA was precipitated by adding 8 ml of ice cold isopropanol, and spinning at 12,000 rpm (Sorvall, HB-4 rotor) for 15 minutes. The DNA pellet was resuspended in 1.38 ml TE buffer, and added to 1.5 g CsCl in a 10 ml tube. After the CsCl was dissolved, 120 ml of 10 mg/ml ethidium bromide was added to give a total volume of 2.2 ml which was placed in a small Beckman TL-100 heat sealable tube. The sample was then centrifuged at 80000 rpm for 16 hours at 20°C in a Beckman TL-100 centrifuge.

The lower band in the CsCl gradient containing supercoiled plasmid DNA was removed from the tube using a needle and syringe. To remove ethidium bromide, one volume of water-saturated butanol was added, mixed, and the top layer discarded. This step was repeated three times. One volume of water, and three volumes of nuclease-free ethanol were then added, and the DNA precipitated. The DNA was pelleted by centrifugation for 5 minutes in a microfuge. The DNA pellet was washed with 70% (v/v) nuclease-free ethanol, the purified DNA was then resuspended in 0.5 ml TE buffer, and stored at -20°C.

2.3.2 Isolation of phage DNA

Approximately 10^5 λ phage were absorbed to 500 µl of an overnight culture of *E. coli* strain LE392 at 37°C for 20 minutes. This was added to 50 ml LMM-broth and incubated overnight at 37°C with shaking. Cells were removed by centrifugation (Sorvall, SS-34 rotor; 5,000 rpm for 5 minutes), and the supernatant treated with 50 µg crude DNase I, and 100 µg Ribonuclease A at room temperature for 30 minutes. The suspension was then spun (SS-34 rotor; 18,000 rpm for 15 minutes), and the supernatant collected. The phage were pelleted by centrifugation (SS-34 rotor; 20,000 rpm for 3 hours), and resuspended in 0.1 M Tris-HCl pH 8.0, 0.3 M NaCl and 0.2% (w/v) SDS. Proteinase K (200 µg) was added, and the sample incubated at 37°C for 30 minutes. This was then extracted with phenol/chloroform three times before the DNA was precipitated with ethanol and resuspended in 100 µl of TE. The yield of DNA from this method was approximately 50-100 µg, and the DNA was sufficiently clean for restriction enzyme digestion.

2.3.3 Isolation of mouse genomic DNA

(i) Isolation of genomic DNA from tissue samples

A tissue sample was cut into a small pieces and placed in liquid nitrogen before grinding to a powder. The powdered tissue was gradually added into a tube containing ten volumes of extraction buffer (10 mM Tris-HCl pH 8.0, 100 mM EDTA, 0.5% (w/v) SDS and 20 µg/ml pancreatic RNase) and the solution was mixed thoroughly. Proteinase K was added to a final concentration of 100 µg/ml and the viscous solution was incubated at 37°C for 12-16 hours. The solution was gently extracted twice with an equal volume of phenol/chloroform, once with an equal volume of chloroform before ethanol precipitation using 0.2 volume of 10 M ammonium acetate and 2 volumes of nuclease-free ethanol. The DNA pellet was collected after centrifugation at 10,000 rpm (Sorvall, SS-34

rotor) for 10 minutes. The pellet was washed twice with 70% ethanol and resuspended in TE buffer. The DNA concentration was determined by UV spectrophotometry at a wavelength of 260 nm. The DNA was kept at 4°C or at -20°C for long term storage.

(ii) Isolation of genomic DNA from mouse tail biopsy

Tail fragments were placed in Eppendorf tubes and 500 µl of tail buffer (50 mM Tris-HCl pH 7.5, 5 mM EDTA, 150 mM NaCl, 1% SDS, 200 µg/ml Proteinase K) was added. After incubation at 55°C for 12-16 hours, the solution was extracted once with 500 µl of phenol/chloroform and once with 500 µl of chloroform. The DNA was then precipitated by the addition of 0.1 volume of 4 M NaCl and 2 volumes of cold nuclease-free ethanol. The DNA pellet was obtained by centrifugation and washed with 70% ethanol and dissolved in an appropriate volume of TE buffer depending on the pellet size. Genomic DNA isolated using this protocol was pure enough to be digested with restriction enzymes and the amount of DNA obtained varied from 40-50 µg.

(iii) Isolation of genomic DNA from tissue culture cells

Cells attached to tissue culture dishes were washed twice with PBS and 200 µl of extraction buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl, 1% SDS, 500 µg/ml Proteinase K) was added. After incubation at 37°C for 12-16 hours, the solution was transferred to an Eppendorf tube containing 200 µl of 10 mM Tris-HCl pH 7.5, 1 mM EDTA and 100 mM NaCl, and then extracted with 400 µl of phenol/chloroform. The DNA was precipitated by the addition of 0.1 volume of 4 M NaCl and 2 volumes of cold nuclease-free ethanol. The DNA pellet was obtained by centrifugation and washed with 70% ethanol before being dissolved in an appropriate volume of TE buffer depending on the pellet size. Genomic DNA isolated using this protocol was pure enough to be digested with

restriction enzymes and the amount of DNA obtained from each clone varied from 20-30 µg per 1.5 cm. diameter well.

2.3.4 Restriction enzyme digestion of plasmid and phage DNA

Restriction endonuclease digestions were performed using the conditions recommended by the supplier for each enzyme.

Analytical digest was perform in 20 µl containing 0.5-1.0 µg of DNA and 2-5 units of enzyme for 1 hour. The reaction was stopped by the addition of 4 µl of agarose gel loading buffer and electrophoresed through an agarose gel.

Preparative digests were performed in 100 µl reactions containing 2-10 µg DNA and 8-15 units of enzyme. After at least an hour of incubation the reaction was stopped by phenol/chloroform extraction and DNA was recovered by ethanol precipitation. The DNA was then resuspended in appropriate volume of TE buffer.

2.3.5 Agarose gel electrophoresis of phage and plasmid DNA

Agarose was dissolved in TAE to the final concentration required (0.7-3.0%), and poured in appropriate gel mould for horizontal gels. For restriction enzyme analysis, 75 mm x 55 mm gels were used. For large scale preparation of DNA fragment, 75 mm x 150 mm gels were used. TAE was used as a running buffer and a current of 70-120 mA was applied to the gels. DNA in TE buffer was mixed with 1/6 volume of agarose gel loading buffer before loading onto the gel. After electrophoresis, the DNA was visualised under UV light after staining with 0.5 µg/ml ethidium bromide and photographed.

2.3.6 Isolation of DNA fragment from agarose gel

DNA samples were localised on agarose by illumination under long wave length UV light (365 nm) after staining with 0.5 µg/ml ethidium bromide. The

gel slice containing the required fragment was excised. DNA was recovered by either of the following methods:

(i) DNA was recovered from an agarose gel slice using the protocol and solutions provided by Bio 101 and based on the method of Vogelstein and Gillespie (1979). The agarose gel slice containing the required fragment was excised from the gel, placed into an Eppendorf tube and 3 volumes of 6 M NaI solution was added. The agarose was dissolved by incubating at 50°C for 5 minutes. A 5 µl aliquot of Glassmilk solution (containing a silica matrix suspension) was added and incubated on ice for 10 minutes to allow the DNA to bind. The Glassmilk/DNA complex was pelleted by centrifugation for 5 seconds and the supernatant discarded. The pellet was washed three times with 500 ml of a wash buffer containing NaCl and ethanol to remove all traces of NaI. The final pellet was resuspended in a small volume of water and incubated at 50°C for 5 minutes. After spinning for 2 minutes the supernatant containing the DNA was transferred to another tube. This method was used to isolated DNA fragments larger than 300 bp since the recovery of DNA smaller than this was very low.

(ii) DNA was recovered from a low melting point agarose gel. The gel slice containing the required fragment was excised from the gel, placed into an Eppendorf tube and an equal volume of TE buffer was added. The agarose was melted at 65°C for 10 minutes. DNA was recovered by one phenol extraction, one phenol/chloroform extraction, followed by ethanol precipitation of the aqueous layer.

2.3.7 Polyacrylamide gel electrophoresis

Electrophoresis of oligonucleotides was carried out on 10%-20% (w/v) polyacrylamide/bis-acrylamide (30:1), which had been deionised with mixed bed ion exchange resin. The acrylamide was polymerised in 1 x TBE buffer by the

addition of 0.1% (w/v) APS, and 0.1% (v/v) TEMED. The gel (14 cm x 14 cm x 0.5 mm) was pre-electrophoresed for 30 minutes at 400 V prior to loading. A quarter volume of acrylamide load buffer was added to the samples which were directly loaded into gel slots approximately 1 cm wide. All acrylamide gels were electrophoresed at 400 Volts. DNA was visualised under UV light (254 nm) after ethidium bromide staining, or by autoradiography if the DNA was radioactively labelled.

2.3.8 Labelling of DNA fragments

(i) End-labelling of synthetic oligonucleotides

Oligonucleotides were 5' end-labelled with T4 polynucleotide kinase and γ-^{32}P-ATP. Usually 50-100 ng of synthetic oligonucleotide was labelled in a 10 µl reaction mixture containing 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 50 µCi of γ-^{32}P-ATP and 1 unit of enzyme. The reaction mixture was incubated at 37°C for 30 minutes, then 5 µl of acrylamide loading buffer was added and the mixture loaded onto a 10%-20% (w/v) polyacrylamide gel for purification. The band corresponding to the labelled oligonucleotide was located by autoradiography for 1 minute, excised using a scalpel blade, and eluted by incubation at 37°C overnight in 400 µl of TE buffer.

ii) Oligo-labelling of DNA fragments

Oligo-labelling of DNA fragments was performed using an Amersham Megaprime DNA labelling kit according to the manufacturer' s instructions. This protocol is based on a method described by Feinberg and Vogelstein (1983).

2.3.9 Southern transfer and hybridisation of DNA on GeneScreen

Plasmid and phage DNA fragments fractionated by agarose gel electrophoresis were transferred to GeneScreen membranes by a modified

method of Southern (1975). Gels were presoaked in 0.6 M NaCl, 0.2 M NaOH and then washed twice in 25 mM sodium phosphate solution pH 6.5. Transfer of the DNA from the gel to the membrane occurred in 25 mM sodium phosphate solution pH 6.5 solution for a minimum of 4 hours. After transfer, the filter was baked in a vacuum oven at 80°C for at least 2 hours.

Membranes were prehybridised at 42°C for at least 4 hours in a solution containing 40% (v/v) formamide, 10% Dextran sulfate, 1% (w/v) SDS, 50 Tris-HCl pH 7.4, 1 M NaCl, 5 x Denhardt's, 100 ml/ml denature sonicated salmon sperm DNA. Heat denatured probe was added and allowed to hybridise for 16-24 hours at 42°C. Membranes were washed at a stringency appropriate for the probe in use. Usually filters were washed in 2 x SSC/0.1% (w/v) SDS at room temperature for 15 minutes and then in 0.2 x SSC/0.1% (w/v) SDS at 65°C for an additional 30-60 minutes depending on the signal. The filters were then autoradiographed at -80°C with a tungsten intensifying screen. If filters were to be reprobed, they were stripped by boiling in 0.5% (w/v) SDS for 10 minutes before being probed again.

2.3.10 Southern analysis of mouse genomic DNA

Generally mouse genomic DNA (7-10 µg) was restriction enzyme digested in a 200 µl reaction mixture containing 33 mM Tris-acetate pH 7.8, 62.5 mM potassium acetate, 10 mM magnesium acetate, 4 mM spermidine and 0.5 mM DTE. The amount of enzyme added in the reaction was usually in five to ten fold excess to allow complete digestion. The reaction was incubated for 12-16 hours at the temperature recommended by the supplier. After the digestion was completed, the DNA was ethanol precipitated, washed with 70% ethanol and then redissolved in 17 µl TE buffer.

For electrophoresis, the digested-DNA in TE buffer (17 µl) was mixed with 3 µl agarose gel loading buffer and loaded onto a 0.7%-1% agarose gel (190 mm x

140 mm x 5 mm). TAE was used as a running buffer and the gel was run at 25-45 volts for 16-20 hours at room temperature. After electrophoresis, the DNA was visualised under UV light after staining with 0.5 μg/ml ethidium bromide and the photograph was taken.

Genomic DNA fragments in agarose gel were transferred to Zeta-Probe membranes by the alkaline blotting procedure described by Chomczynski and Qasba (1984), as modified by Reed and Mann (1985), using an LKB 2016 VacuGene vacuum blotting apparatus. Gels were placed on the apparatus, covered with 0.25 M HCl for 8 minutes under 40 cm/H_2O pressure. The acid solution was removed, and the gel washed with distilled water. The gel was then covered with 0.4 M NaOH solution, and the transfer continued for 90 minutes at 40 cm/H_2O pressure. Filters were washed in 2 x SSC for 5 minutes and left air dry at room temperature.

Filters were prehybridised at 65°C for 5 minutes in a solution containing, 0.25 M Na_2HPO_4 pH 7.2, 1 mM EDTA, 10% (w/v) PEG 6000, 7% (w/v) SDS, and 100 μg/ml denatured sonicated salmon sperm DNA. Heat-denatured probes were added, and allowed to hybridise for 12-16 hours at 65°C. Filters were washed in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 5% (w/v) SDS, 65°C for 30-60 minutes and then in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 1% (w/v) SDS, 65°C for 30-45 minutes depending on the signal. The filters were then autoradiographed. If filters were to be reprobed, they were first stripped by washing twice in a large volume of 0.1 x SSC, 0.5% (w/v) SDS at 95°C for 40 minutes. They were checked by overnight exposure before being reprobed.

2.3.11 Screening of mouse genomic λ libraries

A culture of *E. coli* LE392 was grown overnight in LMM-broth. An aliquot representing 5 x 10^4 pfu from the λ library was mixed with 300 μl of the overnight culture for 20 minutes at 37°C before 10 ml of LMM-agarose was added and the

sample plated onto 15 cm L-agar plates. Twenty plates were screened to cover the total genome. The plates were incubated at 37°C for 8-10 hours or until almost a confluent plate of lysis was obtained, and then cooled at 4°C before the phage were lifted in duplicate onto PlaqueScreen filters. The phage were lysed by a 2 minute treatment in an autoclave set at 105°C and the DNA baked onto the filters for 2 hours at 80°C. Filters were prehybridised at 42°C for at least 4 hours in a solution containing 50% (v/v) deionised formamide, 10% (w/v) dextran sulfate, 1.0% (w/v) SDS, 50 mM Tris-HCl pH 7.4, 1 M NaCl, 5 x Denhardt's and 100 µg/ml denatured sonicated salmon sperm DNA. The prehybridisation solution was removed and was replaced by the same solution plus the heat-denatured probes. The filters were hybridised for 12-16 hours at 42°C. Filters were washed at a stringency appropriate for the probe in use. Usually filters were washed in 2 x SSC/0.1% (w/v) SDS at room temperature for 30 minutes and then in 0.2 x SSC/0.1% (w/v) SDS at 65°C for an additional 30-60 minutes depending on the signal. The filters were then autoradiographed at -80°C with a tungsten intensifying screen.

2.3.12 Subcloning of restriction fragments into plasmid vectors

(i) Preparation of plasmid vectors

Plasmid DNA was digested with the appropriate restriction enzyme(s), then dephosphorylated with calf intestinal phosphatase in a 100 µl reaction mixture containing 50 mM Tris-HCl pH 8.5, 0.1 mM EDTA and 1.0 unit of enzyme. After incubation at 37°C for one hour, the linearised, dephosphorylated vector was purified from uncut vector by running the DNA on a 0.8%-1.2% low melting point agarose gel in TAE buffer. The vector DNA was recovered from the gel slice as previously described (2.3.6).

(ii) Preparation of DNA restriction fragments

The cloned DNA was digested with the appropriate restriction enzyme(s) and electrophoresed on agarose gel or low melting point agarose gel. Restriction fragments were isolated as previously described (2.3.6).

(iii) End-filling of vectors and DNA fragments for blunt end ligation

Digested-DNA with protruding 5' termini was end-filled with 1 unit of DNA polymerase I, Klenow fragment in 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 100 mM of each dNTP at 37°C for 30 minutes. After this time, DNA was loaded onto an agarose gel and isolated as described (2.3.6).

(iv) Ligation into plasmid vectors

Generally the DNA fragment and appropriate vector were ligated in a molar ratio of 3:1 respectively, in a 20 µl reaction containing 20 ng of vector, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1mM DTT, 0.5 mM ATP, 5% PEG 6000 and 1 unit of T4 DNA ligase. The reactions were incubated for 4-16 hours at 4°C.

(v) Transformation of plasmid recombinants into bacteria

E. coli ED8799 and *E. coli* DH5α were made competent and transformed using a modification of the method described by Dagert and Ehrlich (1979).

E. coli ED8799, or *E. coli* DH5α, was grown overnight in L-broth at 37°C with constant shaking. The cells were subcultured 1/100 (v/v) into fresh L-broth and grown with shaking at 37°C until an A_{600} of 0.5-0.6 was obtained. The cells were chilled on ice for 15 minutes and pelleted by centrifugation at 5,000 rpm (Sorvall, HB-4 rotor) for 5 minutes at 4°C. The cells were gently resuspended in half the original volume of ice-cold 100 mM $MgCl_2$. The cells were pelleted again by centrifugation and gently resuspended in 1/25 volume of ice-cold 100 mM $CaCl_2$ and left on ice for at least 3 hours prior to use.

Usually a quarter of a ligation reaction was mixed with 200 µl of the

competent cells, and kept on ice for 15 minutes. After heat-shock for 2 minutes at 42°C, the cells were kept on ice for a further 15 minutes and then warmed to room temperature. L-broth (0.5 ml) was added and the cells incubated at 37°C for 20 minutes. After this time, the cells were gently pelleted, resuspended in 100 µl of L-broth, and spread directly onto an L-agar plate (with the appropriate antibiotic). When colour selection was required with *E. coli* ED8799, or *E. coli* DH5α cells, L-agar plates (with the appropriate antibiotic) were used with 60 µl of BCIG (20 mg/ml in DMF) and 40 µl of IPTG (20 mg/ml) spread directly onto the surface of each plate. The plate was incubated overnight at 37°C.

2.3.13 Subcloning of DNA restriction fragments into M13 Vectors

(i) Preparation of M13 replicative form (RF)

A single plaque was picked into 1.5 ml of 2xYT broth containing 40 µl of a fresh *E. coli* JM101 overnight culture grown in minimal media. This culture was incubated at 37°C with vigorous aeration for 6 hours and then dispensed into 50 ml of 2xYT broth and grown overnight. The method of isolation of M13 RF from this culture was the same as that used for the isolation of plasmid DNA (2.3.1).

(ii) Subcloning into M13 vectors

Vector M13 DNA and insert was prepared as described (2.3.12). Ligations generally contained 20 ng vector and sufficient insert to give a 3:1 molar ratio of vector to insert.

(iii) Transformation of recombinant M13 into *E. coli* JM101

Preparation of competent *E. coli* JM101 cells and transfection of recombinant M13 was performed as described by Messing (1983).

Competent *E. coli* JM101 cells were prepared as described (2.3.12 (v)). One quarter of the ligation reaction was mixed with 200 µl of the competent cells and the cells were kept on ice for 20 minutes. After heat-shock at 42°C for 2 minutes,

the cells were mixed with 3 ml of L-agar (warmed at 42°C) containing 30 µl of BCIG (20 mg/ml in DMF), 10 µl of IPTG (20 mg/ml in water) and 0.2 ml of a *E. coli* JM101 overnight culture (diluted 1:40 in 2xYT broth), and then plated directly onto a minimal plus glucose plate. The plate was incubated overnight at 37°C.

(iv) Preparation of single-stranded M13 DNA

M 13 phage plaques were picked into 2.0 ml of a 1:40 dilution of a fresh *E. coli* JM101 overnight culture (grown in minimal medium) in 2xYT broth. After incubation at 37°C with vigorous shaking for 5 hours, 1.5 ml of the culture was transferred into an Eppendorf tube and centrifuged for 10 minutes in a microfuge. The supernatant was transferred into an Eppendorf tube containing 200 µl of 2.5 M NaCl, 20% (w/v) PEG 6000 and left at room temperature for 15 minutes. The single-stranded M13 phage particles were obtained as a pellet after centrifugation for 5 minutes. After removal of the supernatant, the pellet was resuspended in 200 µl of TE buffer and extracted with phenol/chloroform. The DNA was recovered by ethanol precipitation, resuspended in 25 µl of TE buffer and stored frozen at -20°C.

2.3.14 Oligonucleotide-directed *in vitro* mutagenesis (Zoller and Smith, 1983)

(i) End-labelling of the oligonucleotides

The oligonucleotide primers used in mutagenesis reactions were all end-labelled before use. The end-labelling reaction (in a total volume of 10 µl) containing 100 ng of oligonucleotide, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1mM DTT, 1 mM ATP, 1 unit of T4 polynucleotide kinase was incubated at 37°C for 1 hour, then heated to 65°C for 10 minutes, diluted to 4 ng/µl and stored at -20°C.

(ii) Mutagenesis reaction and transformation

The DNA fragment to be mutated was cloned into a M13 vector and single-stranded DNA was isolated (2.3.13 (iv)). The mutagenesis reaction (in a total volume of 15 µl) containing 100 ng of single-stranded DNA, 4 ng of kinased mutagenic primer, 2.5 ng kinased USP, 200 mM NaCl, 10 mM Tris-HCl pH 8.0 and 10 mM $MgCl_2$ was heated at 65°C for 5 minutes and was cooled down at room temperature for 10 minutes to allow annealing to occur. Then 5 µl of 500 mM each dNTP, 5 µl 10 mM ATP, 5 µl 10 mM DTT, 2 units of Klenow enzyme and 1 unit of T4 DNA ligase was added. The reaction was incubated at room temperature for 4 hours and aliquots of 0.5, 1 and 2 µl transformed into *E. coli* JM101.

(iii) Screening plaques for mutants

A plate containing approximately 100 well separated plaques was chosen for screening. Duplicate nitrocellulose lifts were made by the method of Benton and Davis (1977). The filters were allowed to dry and baked at 80°C for 2 hours.

The filters were prehybridised at 42°C for 2 hours in a solution containing 0.9 M NaCl, 90 mM Tris-HCl pH 7.5, 9 mM EDTA, 0.5% (v/v) Nonidet P40, 5 x Denhardt's and 100 µg/ml sonicated salmon sperm DNA. Following prehybridisation mutagenesis oligonucleotides which have been ^{32}P-labelled by kinasing (2.3.8 (i)) was added to a concentration of 10 ng/ml and incubation at 42°C overnight.

(iv) Washing filters with TMACl

Mutant were distinguished from the parental plaques by selective washing in 3M TMACl. In a 3 M solution of this compound G/C and A/T base pairs have equal binding energies allowing base independent melting temperatures which can be estimated from a predetermined melting curve (Wood *et al.*, 1985).

Following overnight hybridisation at 42°C, filters were rinsed twice in 6 x SSC at room temperature and then once with TMACl solution (3 M TMACl,

50 mM Tris-HCl pH 8.0, 0.1% (w/v) SDS and 2 mM EDTA) at room temperature. Two 30 minutes washes were then done at the selection temperature determined from the melting profile presented by Wood *et al.* (1985). The filters were then sealed in plastic bag and autoradiographed overnight. Positive plaques were picked off the original plate, grown up and sequenced to confirm the base changes.

2.3.15 Dideoxy chain-termination sequencing (Sanger *et al.*, 1977)

(i) Sequencing reaction

Sequencing was performed using either single-stranded M13 DNA, prepared as described (2.3.13 (iv)), or double-stranded plasmid DNA as a template.

Plasmid DNA (2 µg) dissolved in 13.5 µl of TE buffer was treated with 1.5 µl of Ribonuclease A (10 mg/ml) at 37°C for 15 minutes, then 3.75 µl of 1 M NaOH/1 mM EDTA was added and incubated at 37°C for 15 minutes. The DNA was then purified by centrifugation through a Sepharose CL-6B column before using as a template in sequencing reaction.

Sequencing reaction was performed using a sequencing kit according to the manufacturers' instructions.

(ii) Sequencing gels

Dideoxy chain-termination sequencing reaction products were electrophoresed on 6% polyacrylamide gel containing 8 M urea and TBE buffer. The gels used were 20 cm x 20 cm x 0.25 mm and were pre-electrophoresed for 20 minutes before loading the sample. The gels were run in TBE buffer under the constant current of 20 mA. After electrophoresis was complete the gel was attached to a sheet of 3 MM Whatman paper and covered with plastic-wrap before drying on a gel drier for 30 minutes. The dried gels were autoradiographed overnight at room temperature.

2.3.16 Polymerase chain reaction (PCR) analysis (Saiki *et al.*, 1985)

PCR was carried out using a Perkin-Elmer/Cetus DNA thermal cycler (model 0993-8412). PCR reactions were performed in a total of 50 µl volume containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 0.001% (w/v) gelatin, 2 mM $MgCl_2$, 200 µM dNTP's, 100 ng of template DNA, 1 µg each primer and 1 unit *Taq* polymerase. Thirty cycles were performed with denaturation at 95°C for 1 minute, annealing for 45 seconds at the temperature appropriate for each primer set and extension for 2 minutes at 72°C.

PCR cycles for PHZ.Nco (2.2.15) and pHZ.Hd (2.2.15) are as follows:

cycle 1-10: 95°C for 1 minute, 45°C for 1 minute, 72°C for 2 minutes,

cycle 11-40: 95°C for 1 minute, 54°C for 1 minute, 72°C for 2 minutes.

An aliquot (10 µl) of the reaction was analysed by electrophoresis on 1%-2% agarose gel depending on size of the PCR product.

2.3.17 Preparation of plasmid DNA for electroporation into ES cells

Plasmid DNA, prepared by CsCl ultracentrifugation (2.3.1), was digested with the appropriate enzymes using the conditions recommended by the supplier. Normally 100-150 µg DNA was incubated with 150-200 units of restriction enzyme for at least 5-6 hours in the reaction volume of 400 µl. After this time an aliquot of the reaction was examined by agarose gel electrophoresis to ensure that the digestion was completed. The reaction was extracted once with phenol/chloroform, once with chloroform. The aqueous layer containing DNA was taken and extracted twice with an equal volume of water-saturated ether in order to removed traces of phenol and chloroform from the DNA solution. Traces of ether were removed by heating the DNA solution at 50°C for 30 minutes. The DNA was ethanol precipitated and washed three times with 70% ethanol before dissolving in appropriate volume of water. Usually DNA was resuspended to a final concentration of 2 µg/µl.

2.4 RNA Methods

2.4.1 Isolation of total RNA from mouse ES cells

Total RNA was extracted from ES cells following the method described by Edwards *et al.* (1985). ES cells were harvested by trypsinisation and washed by centrifugation in PBS. The cell pellet was resuspended in 2 ml of ice-cold solution containing 30 mM Tris-HCl pH 7.6, 150 mM NaCl, 15 mM $MgCl_2$ and 0.4% Nonidet P40 and the cells were lysed by vigorously mixing. The solution was kept on ice for 5 minutes and then centrifuged at 3000 rpm (Sorvall, SS-34 rotor) for 20 minutes at 4°C. The supernatant was collected into the tube containing 2 ml of TUNES (10 mM Tris-HCl pH 8.0, 7 M urea, 350 mM NaCl, 1 mM EDTA and 2% SDS). The solution was mixed and extracted with phenol and chloroform, and then RNA precipitated by the addition of one tenth volume of 3 M sodium acetate (pH 5.5) and 2 volumes of nuclease-free absolute ethanol. After centrifugation the supernatant was collected, and RNA precipitated by mixing in one volume of isopropanol and incubating at -20°C for 15 minutes. After centrifugation (SS-34 rotor; 10,000 rpm for 20 minutes at 4°C), the RNA pellet was dissolved in 400 µl of DEPC-treated water and the solution was transferred into an Eppendorf tube. The RNA was ethanol precipitated again before dissolving in the appropriate volume of DEPC-treated water. The RNA was stored at -20°C. The amount of RNA was determined by UV spectroscopy.

2.4.2 Northern transfer of RNA

Electrophoresis of RNA was performed on 1.2% (w/v) agarose gels, containing 2.2 M formaldehyde and 1 x formaldehyde gel-running buffer (20 mM MOPS pH 7.0, 8 mM sodium acetate and 1 mM EDTA). RNA samples were ethanol precipitated, and resuspended in 4.5 µl DEPC-treated water, 2 µl 5 x formaldehyde gel-running buffer, 3.5 µl formaldehyde, and 10 µl formamide.

Samples were heated to 65°C for 15 minutes prior to loading. The gel was submerged in 1 x formaldehyde gel-running buffer and run at 4 volts/cm.

RNA was transferred onto Zeta-Probe membranes after agarose gel electrophoresis by capillary transfer for 12-18 hours, using 50 mM NaOH as the blotting solution. After transfer, the membrane was rinsed in 2 x SSC and RNA was fixed onto the membrane by baking in a vacuum oven at 80°C for 30 minutes. The filter was prehybridised and hybridised as described (2.3.10). The filter was washed in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 5% (w/v) SDS at 65°C for 30 minutes and then in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 1% (w/v) SDS at 65°C for 15-30 minutes depending on the signal. The filter was then autoradiographed.

2.5 Protein Methods

2.5.1 Preparation of protein from *E. coli*

Protein extracts were prepared from *E. coli* cultures by collecting the cells (Eppendorf centrifuge, 1 min) and resuspending in one twentieth volume x A_{600} of 2 % (w/v) SDS, 10% (v/v) β-mercaptoethanol. The cells were vortexed and then boiled for 5 minutes. Extracts were stored at -20°C. This concentration was suitable for 10 µl to be run on a 0.5 cm thick SDS-polyacrylamide gel, with 1 cm wide gel slots.

2.5.2 SDS/polyacrylamide gel electrophoresis of proteins

Protein was electrophoresed on vertical 14 x 14 x 0.05 cm SDS/polyacrylamide gels (Laemmli, 1970). The gel mix contained 15% (w/v) acrylamide (38:1 acrylamide: bisacrylamide), 375 mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.2% (w/v) APS and 0.1% (v/v) TEMED. The gels also contained a 2 cm

4% (w/v) acrylamide stacking gel. Protein samples were treated before loading onto the gel by addition of a half volume of 2 x SDS loading buffer (10% (v/v) glycerol, 375 mM Tris-HCl pH 8.8, 5% (w/v) SDS, 0.1% bromophenol blue, 5% β-mercaptoethanol (added just prior to use) and incubation at 100°C for 3 minutes.

Gel tanks contained approximately 1 L of running buffer (50 mM Tris-HCl pH 8.3, 380 mM glycine and 0.1% (w/v) SDS), and gels were run at 30 mA. The protein in the gel was visualised by staining with 0.1% (w/v) Coomassie brilliant blue in 50% (v/v) methanol, 10% (v/v) acetic acid overnight at room temperature. The gel was destained in several changes of 5% (v/v) methanol, 10% (v/v) acetic acid by diffusion at room temperature.

2.5.3 Western transfer

Protein was transferred to PVDF membrane (Immobilon-P) by the method of Matsudaira (1987), using a Hoefer transfer electrophoresis unit. Briefly, protein to be transferred was fractionated on a 15 % SDS/polyacrylamide gel (2.5.2). The gel was transferred to an Immobilon-P membrane in 10 mM CAPS and 10% (v/v) HPLC-grade methanol at 200 mA for 10 minutes.

The membrane was blocked in PBST (PBS containing 0.05% (v/v) Tween 20) at room temperature for at least 1 hour before antibodies diluted in PBST was added and incubated at 37°C for 1 hour. The filter was washed in 3 changes of PBST and then the second antibody, conjugated to alkaline phosphatase, was added. This was incubated for a further hour at 37°C and again washed in PBST before rinsing in substrate solution (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$). The membrane was left in a minimal volume of substrate solution.

The detection system comprised addition of 80 µl NBT (75 mg/ml in 70% (v/v) DMF) and 80 µl BCIP (50 mg/ml in DMF) per 20 ml of the substrate

solution. A blue precipitate was allowed to develop for 5-10 minutes before the reaction was terminated by addition of 10 mM EDTA.

2.6 Tissue Culture Methods

Tissue culture cell lines were maintained at 37°C in an atmosphere of 10% CO_2 in air.

Tissue culture dishes used for ES cells without feeder layer were gelatinised by treatment with 0.1% gelatin in PBS for a minimum of 30 minutes at room temperature.

2.6.1 Preparation of STO feeder layer for ES cell culture

G418-resistant or hygromycin-resistant STO cells (2.2.8) were used as a feeder layer for growing ES cells.

The STO feeder layer was prepared as described by Robertson (1987). STO cells were grown in DMEM supplemented with 10% FBS. The cells were passaged before they reached confluence, in order to prevent the accumulation of cells which had lost contact inhibition.

STO cells, treated with mitomycin C to prevent cell division, were used as feeder layers. When the cells were ready to passage, the culture medium was discarded and replaced with fresh medium with 10 μg/ml of mitomycin C. The flask was incubated at 37°C for 2 hours before the STO cells were gently washed three times with PBS. Then, the cells were collected by trypsinisation with 0.1% trypsin, followed by centrifugation at 1200 rpm for 5 minutes. The cells were resuspended in culture medium at 3 x 10^5 cells/ml, then seeded onto plates at a dilution of 5 x 10^4 cells/cm^2. These mitotically inactive STO cells were used as feeder layers for ES cell culture .

2.6.2 Passaging and maintenance of ES cells

E14 ES cells were grown on STO feeder cells in DMEM supplemented with 15% FBS and 1000 U/ml LIF, 50 µg/ml gentamycin and 0.1mM β–mercaptoethanol. ES cells were passaged every 3 days. When the cells were ready to passage, they were washed with 5 mM EGTA in PBS and then treated with 0.1% trypsin-EDTA for 2-3 minutes and disaggregated by repeated pipetting. The cells were washed by centrifugation at 1200 rpm for 7 minutes, resuspended in complete ES medium and counted. The E14 ES cells were reseeded on the feeder cell layer at a 1:10 or 1:20 dilution. The dilution which gave the best cell distribution and morphology was selected for the next passage. GOOD COLONY MORPHOLOGY MEANS THE ES CELL COLONY WHICH CONTAINS A TYPICAL ES CELL MORPHOLOGY (DOMED SHAPE) AND STILL IN AN UNDIFFERENTIATED STATE.

2.6.3 Electroporation of ES cells

Three protocols for electroporation of ES cells were described in these studies. The protocol used in each electroporation experiment are stated in the text. They are as follows:

Protocol 1: HBS (20 mM HEPES pH 7.0, 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 6 mM glucose and 0.1 mM β-mercaptoethanol) was used as an electroporation medium and the setting was 960 µF capacitance, 210 volts (L. Williams, pers. comm.);

Protocol 2: PBS (Dulbecco's PBS (without Ca^{2+} and Mg^{2+})) was used as an electroporation medium and the settings were 25 µF capacitance, 350 volts (J Beall, pers. comm.);

Protocol 3: MT PBS (16 mM Na_2HPO_4, 4 mM $NaH_2PO_4.2H_2O$ and 150 mM NaCl) was used as an electroporation medium and the setting was 500 µF capacitance, 220 volts (I. Lyons, pers. comm.).

E14 ES cells were maintained and harvested by trypsin as described (2.6.2). ES cells were resuspended in cold electroporation medium at the concentration of 1×10^8 cells/ml. 5×10^7 ES cells in 0.5 ml (electroporation medium) were placed in a 0.4 cm cuvette and mixed with DNA. The mixture was incubated at

4°C for 10 minutes before electroporation using Bio-Rad Gene Pulser. After receiving the pulse, the sample were placed at room temperature for 10 minutes. The cells were plated at the density of 5 x 10^6 per 100 mm plate on either the G418-resistant or hygromycin B-resistant STO feeder layer 10 minutes after electroporation. Selection either with G418 or hygromycin was begun 24 hours after plating. The selective medium was change every 2-3 days.

2.6.4 Picking and freezing ES cells colonies

The strategy was to identify transformed colonies and then each individual colony was divided in half. Half of the colony was frozen and the other half was cultured in BRL medium without STO cells for DNA preparation.

After 8-12 days of selection, the 100 mm dishes which contained G418-resistant or hygromycin-resistant ES cells colonies were washed once with PBS/0.5 mM EGTA and replaced with 10 ml of the same solution. Individual colonies were harvested in 50 µl volume using a Pipetteman P200 and then transferred into a unique well of a 96-well tray containing 50 µl of 0.1% trypsin. The colonies were dispersed by repeated pipetting and then 50 µl of this cell suspension was transferred into the wells of gelatinised 24-wells tray containing the medium (60% of BRL-CM and 40% of complete ES cell medium, see section 2.2.10) without feeder layer. The plates were incubated at 37°C for 3-5 days before the cells were lysed and DNA was isolated for Southern analysis as described (2.3.3 (iii)). The remainder of the cell suspension (50 µl) was transferred to a well of a 24-well plate containing 200 µl of feeder cell suspension. Following addition of 250 µl of 2 x freezing medium (20% DMSO in FBS), the plates were stored (in a polystyrene box) at -80°C.

2.6.5 Processing of the positive ES cell clones

The positive clones identified by Southern analysis were thawed from the

master plate kept at -80°C. The clones were expanded for verification by Southern analysis, mycoplasma testing and karyotyping before injection into host blastocysts. To thaw the master plate, the 24-wells trays were taken from -80°C and 2 ml of warm complete ES medium (37°C) was added to the well(s) that contained the required clone(s). The trays were incubated at 37°C for 5 minutes and then the cell suspension was transferred into a new well of a 24-well tray containing a fresh feeder layer. The plate was incubated at 37°C for 2 hours, during which the ES cells attached, and then the medium was replaced with fresh medium. The ES cells were expanded and an aliquot of the cells was stored in freezing medium (10% DMSO in FBS) and kept in a -80°C freezer. The ES cells were passaged cultured on the gelatinised plate using BRL medium. After expansion, the ES cell clones which were verified by Southern analysis, mycoplasma-free, normal karyotype and good morphology were chosen for injection into blastocyst.

2.6.6 Blastocyst injection and production of chimaeric mice

Blastocyst injection and production of chimaeric mice were carried out essentially as described (Bradley, 1987). Ten to fifteen ES cells were injected into CBAxC57Bl/6J F_2 blastocysts, and then transferred to CBAxC57B1/6J F_1 pseudopregnant females. Pups were born and chimaeras were identified by the coat colour.

2.7 Autoradiography

Radioactively labelled DNA, which had been electrophoresed on polyacrylamide gels, was visualised by autoradiography. Gels were either covered by a thin sheet of plastic wrap before exposure, or dried onto a piece of 3 MM Whatman paper. A sheet of Fuji X-ray film or Kodak X-OMAT film was

placed over the gel enclosed in an Ilford autoradiography cassette, and exposed at room temperature for the required amount of time. For detection of low levels of radioactivity, autoradiography was carried out either in a cassette with a tungsten intensifying screen at -80°C. After exposure, the X-ray film was developed, fixed, washed, and dried automatically using an Agfa Curix 60 model X-ray film developer.

2.8 UV Spectrophotometry

The concentrations of nucleic acid samples were determined by measuring their UV absorbance at a wavelength of 260 nm using a Shimadzu UV-160A recording spectrophotometer.

2.9 Computer Programs

The comparison of nucleic acid sequences described in this thesis was done using the program Bestfit in the Genetics Computer Group sequence analysis software package version 6.1, 1989 (Devereux *et al.*, 1984).

2.10 Containment Facilities and Animal Ethics

All manipulations involving recombinant DNA were carried out in accordance with the regulations and with the approval of the Australian Academy of Science Committee on Recombinant DNA and the University Council of the University of Adelaide.

All procedures involving animals were carried out with the approval of the University of Adelaide Animal Ethics Committee.

Chapter 3

Isolation and Characterisation of the Mouse Histone H2A.Z Gene

3.1 Introduction

A cDNA encoding the chicken histone H2A.F protein was first isolated and characterised in this laboratory (Harvey *et al.*, 1983) and subsequently the entire genomic clone containing five exons and four introns was sequenced (Dalton *et al.*, 1989). H2A.Z gene homologues have been isolated from many species, e.g., *Tetrahymena* (White *et al.*, 1988), sea urchin (Ernst *et al.*, 1987), *Drosophila* (van Daal *et al.*, 1988), human, rat and cow (Hatch and Bonner, 1988; Hatch and Bonner, 1990). However, the mouse histone H2A.Z gene had not been isolated and characterised at the commencement of this project.

The genes encoding H2A.Z protein are present in only one copy per haploid genome, except in sea urchin which probably has four copies of this gene. The presence of several pseudogenes was reported in the human genome (Hatch and Bonner, 1990).

Preliminary work, using chicken histone H2A.F cDNA as a probe to screen a mouse genomic library was performed in this laboratory by P. Wait (Hons. thesis, 1990). Two subclones (pBS14-2.9, pBS17-3) were isolated and preliminary Southern analysis indicated these as possible candidates for mouse H2A.Z genomic clones. However, as indicated below, DNA sequence analysis of these putative H2A.Z clones at the beginning of this project demonstrated that one corresponded to a portion of the major H2A.1 gene and the other was likely to be a processed pseudogene derived from H2A.Z.

To isolate an authentic mouse histone H2A.Z gene, two mouse genomic libraries (I and II; 2.2.7), derived from mouse strain BALB/c, were screened.

This chapter describes the isolation and characterisation of the mouse histone H2A.Z gene from the genomic libraries. Southern analysis of mouse genomic DNA and Northern analysis of total RNA isolated from ES cells, are also described. In addition, a clone corresponding to mouse histone H2A.Z cDNA was also isolated during the course of this work and used to produce

histone H2A.Z protein in bacteria.

3.2 Results

3.2.1 Analysis of previous isolated clones, pBS14-2.9 and pBS17-3

It was shown that a 0.55 kb *Pvu* II fragment derived from pBS14-2.9 and pBS17-3 hybridised with the 5' end of chicken histone H2A.F cDNA (Wait, Hons. thesis, 1990). However, the sequences of these fragments had not yet been obtained.

(i) Sequencing analysis of pBS14-2.9

pBS14-2.9 contained a 2.9 kb insert from which a 0.55 kb *Pvu* II fragment was isolated and subcloned into *Eco* RV-digested pBluescript KS+ (Fig. 3-1). After transformation, the recombinant plasmid pBS14-0.55 was obtained.

Sequence analysis was performed by the dideoxy chain termination method of Sanger *et al.*, (1977) using double-stranded DNA as template with USP and RSP as sequencing primer. 135 bases of sequence were read using USP (data not shown) and 185 bases of sequence were read using RSP primer (Fig. 3-1).

The nucleotide sequence of pBS14-0.55 was compared to chicken H2A.F cDNA (Harvey *et al.*, 1983) and the coding sequence of mouse histone H2A.1 (Gruber *et al.*, 1990). It was found that the sequence of pBS14-0.55 obtained from RSP primer was highly homologous (approximately 89%) to the coding sequence of mouse histone H2A.1 (Fig. 3-1). Therefore, this clone was not studied further.

(ii) Sequencing analysis of pBS17-3

pBS17-3, which contained a 3 kb insert fragment in pBluescript KS+, was digested with *Pvu* II. The 0.55 kb *Pvu* II fragment was isolated and subcloned

Figure 3-1

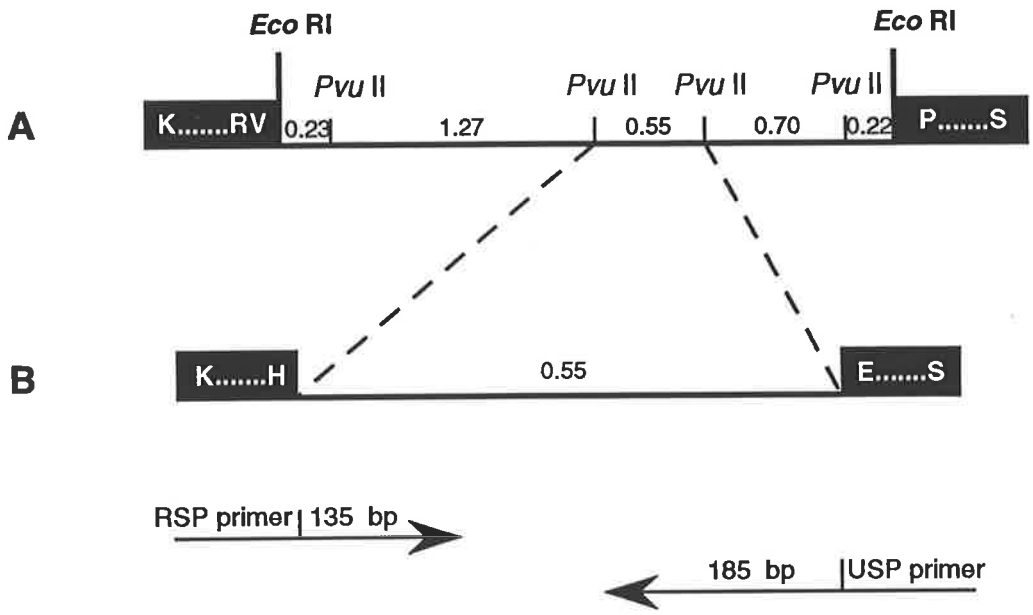
A : Restriction map of pBS14-2.9 containing a 2.9 kb *Eco* RI insert in pBluescript KS+. The internal fragment size is indicated by the numbers in kilobases. The polylinker restriction sites are boxed in black and abbreviations used are as follows:

K	*Kpn I*
RV	*Eco* RV
P	*Pst* I
S	*Sac* I

B : The map of pBS14-0.55 containing 0.55 kb *Pvu* II fragment in pBluescript KS+ derived from pBS14-2.9 is shown, together with the strategy used to obtain the sequence as indicated by the arrows. The polylinker restriction sites are boxed in black and abbreviations used are as follows:

H	*Hin*d III
E	*Eco* RI

C : The comparison of pBS14-0.55 sequence obtained using RSP primer to mouse histone H2A.1 coding sequence (Gruber *et al.*, 1990) is illustrated. The numbers indicate the base sequence from 5' end to 3' end.



C

pBS 14-0.55 (RSP primer) sequence.....

Mouse histone H2A.1 sequence.....

```
    1                                                                  64
5'  CTAC-CG-AGCGCGTGGGTGCGGGAGCGCCC--TGTACCTGGCGGCGGTGCTGGAGTACCTGACGGCC
    |||| || |||||||||| || || || ||| |||||||||||| ||||||||||||||||||||||
5'  CTACTCGGAGCGCGTGGGCGCCGGCGC-CCCGGTGTACCTGGCGGCTGTGCTGGAGTACCTGACGGCC

                                                                      135
GAGATCCTGGAGCTGGCGGGCAACGCGGCGAGGGACAACAAGAAGACGCGCATCATCCCT-GCCACCTGCAG  3'
|||||||||||||||||||||| ||||| | |||||||||||||||||||||||||| ||||||||||
GAGATCCTGGAGCTGGCGGGCAATGCGGCCCGCGACAACAAGAAGACGCGCATCATCCCGCGCCACCTGCAG  3'
```

into *Eco* RV-digested pBluescript KS+ vector. The recombinant plasmid was named pBS17-0.55.

Using double-stranded DNA of pBS17-0.55 as a template, the total sequence of 573 nucleotides was obtained. In order to obtain sequence located at the 5' end of 0.55 kb *Pvu* II fragment, DNA from the parent clone, pBS17-3 was also sequenced using the RSP primer. The maps of pBS17-3 and pBS17-0.55 are shown in Fig. 3-2, together with the strategies used to obtain sequences at the 5' end of pBS17-3 clone. As a result, a 760 nucleotide sequence (*Eco* RI to *Pvu* II restriction site) was obtained (Fig. 3-3).

Comparison of the pBS17-3 sequence (*Eco* RI to *Pvu* II restriction site) to chicken histone H2A.F cDNA (Harvey *et al.*, 1983) and human H2A.Z cDNA (Hatch and Bonner, 1988) showed approximately 76% and 85% homology respectively (Fig. 3-4). Inspection of the homologies indicated that this second clone isolated by P. Wait is certainly related to H2A.Z and the sequence did not appear to contain any intron donor or acceptor sites. Therefore, it is most likely to be a processed pseudogene present in the mouse genome. It is referred to as mouse histone H2A.Z pseudogene 1.

3.2.2 Mouse genomic library I screening

The mouse genomic library I (2.2.7) containing approximately $2x10^6$ independent recombinants, was screened using the ~~567~~ 572 bp *Pvu* II fragment of mouse histone H2A.Z pseudogene 1 (Fig. 3-5), referred to above. This probe contained H2A.Z homologous sequence and a highly repeated T, non-specific sequence. The library was plated as described (2.3.11). The probe was labelled with α-^{32}P dATP (2.3.8 (ii)) and then hybridised to the filters. After hybridisation the filters were washed at a stringency of 0.5x SSC, 0.1% SDS at 65°C and set up for autoradiography.

The autoradiography result showed a high background, probably due to

Figure 3-2

A: Restriction map of pBS17-3 containing a 3 kb *Eco* RI insert in pBluescript KS+. The internal fragment size is indicated in kilobases. The polylinker restriction sites are boxed in black and abbreviations used are as follows:

K	*Kpn* I
RV	*Eco* RV
P	*Pst* I
S	*Sac* I

B: The map of pBS17-0.55 containing 0.55 kb *Pvu* II fragment in pBluescript KS+ derived from pBS17-3. The strategy used to obtain the sequence is indicated by the arrows. The polylinker restriction sites are boxed in black and abbreviations used are as follows:

H	*Hin*d III
E	*Eco* RI

C: The position of the mouse histone H2A.Z pseudogene 1 in pBS17-3 is shown. The 760 bp *Eco* RI/*Pvu* II restriction fragment was sequenced as indicated by the arrows. The homology of this sequence to chicken histone H2A.F cDNA and to human histone H2A.Z cDNA is represented by the striped box. The numbers indicate the distances in basepairs from the *Eco* RI restriction site on the left side of the diagram.

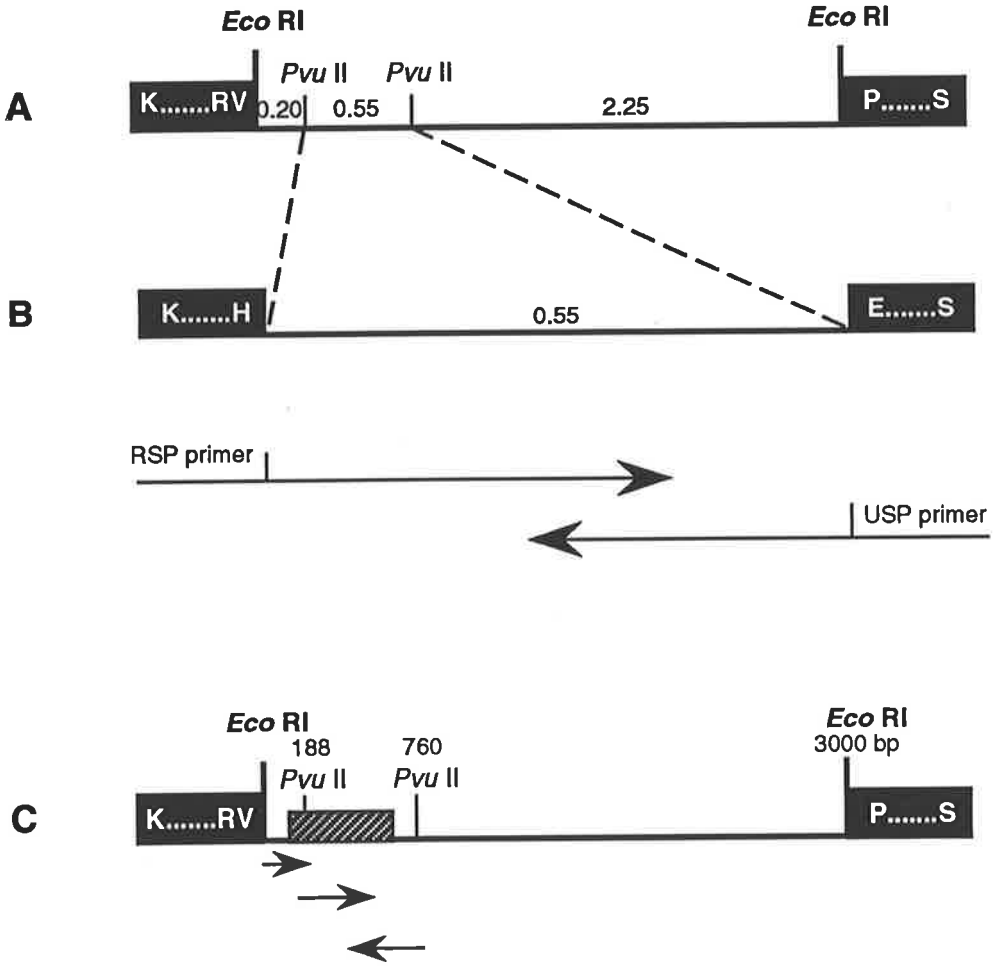

A
Eco RI
Pvu II
Pvu II
Eco RI
K.....RV
0.20
0.55
2.25
P.....S
B
K.....H
0.55
E.....S
RSP primer
USP primer
C
Eco RI
188
760
Eco RI
3000 bp
Pvu II
Pvu II
K.....RV
P.....S

Figure 3-3

Nucleotide sequence of mouse histone H2A.Z pseudogene 1

The sequence of the 760 bp *Eco* RI to *Pvu* II restriction fragment derived from the plasmid pBS17-3 (Fig. 3-2) is shown. This sequence is measured in basepairs, relative to the *Eco* RI site located at the 5' end of the region, and the position of restriction enzyme sites used for probe preparation for further library screening (Fig. 3-5) are shown.

Mouse histone H2A.Z pseudogene 1

Eco RI

1 GAATTCTGCCATGATGGAGTTTACAAATTGCTCCTGAAAAACGACAGAGAGGAGATTCTA 60

61 TTAACTATTTGTGTTATAAAATAAATTGGCAGTTTGAATTGCGGTCCGACAGAGGAGTGG 120

121 GCGCCGGGATCTCGCTGAGCGTCCGCCTGGCTTCGTCTCTTCCTCGCTTGTCGGAGCGAG 180

Pvu II

181 ATGGCTGCAGCTGGAAAGGACTCCGGAAAGGCCAAGTCAAAGGCGGTTTCCCCAGTCGCA 240

Alu I

241 GTTCCCTGGGGCCCGGTATTCATCGACACCTGAAATCTAGGAGAATCAGCCACCGACTGT 300

301 GTGCGACCGCCGCTGTGTACAGCGCAGCCATCCTGGAGTACCTCACCGCAGAGGTACTTG 360

361 AGTTGGCAGGAAATGCATCAAAAGACTTAAAGGTAATGCGTATCACCCCTCGTCACTTGC 420

Alu I

421 GGCTTGCTATACGTGGAGATGAAGAATTGGATTCTCTGATCAAAGCTACCATTGCTGGTG 480

481 GTGGTGTCATCCCACACATCCACAAATCGCTGACTGGGAAGAAAGGACAACAGAAGACTG 540

541 TCTAAGGATGTCTGGATTCCTTATTATCTCAGGACTCTAAATTTTTTGTTTTGTTTTGCT 600

601 TTATTTGGTTTGGTTTGGTTTTTTGAGACAGGGTTTCTCTGTGTAGCCCTGGCTGTCCTG 660

661 GAACTCACTCTGTAGACCAGGCTGGCCTCGAACTCAGAAATCCACCTGCCTCTGCCTCCC 720

Pvu II

721 AAGTGCTGGGATTAAAGGCATGTGCCACGAGGCCCAGCTG 760

Alu I

Figure 3-4

Comparison of nucleotide sequences for the mouse histone H2A.Z pseudogene 1, the human histone H2A.Z cDNA and chicken histone H2A.F cDNA

The alignment of DNA sequences of mouse histone H2A.Z pseudogene 1 to human histone H2A.Z cDNA (Hatch and Bonner, 1988) (top) and chicken histone H2A.F cDNA (Harvey *et al.*, 1983) (bottom) is illustrated. The homology region spans from ATG through TAA (indicated by ***) in human histone H2A.Z cDNA, to 37 bases 3' of the stop codon. The identity of this region is 85%. Comparison with chicken histone H2A.F cDNA, from ATG to TAG (indicated by ***) indicates 76 % nucleotide homology to the mouse sequence.

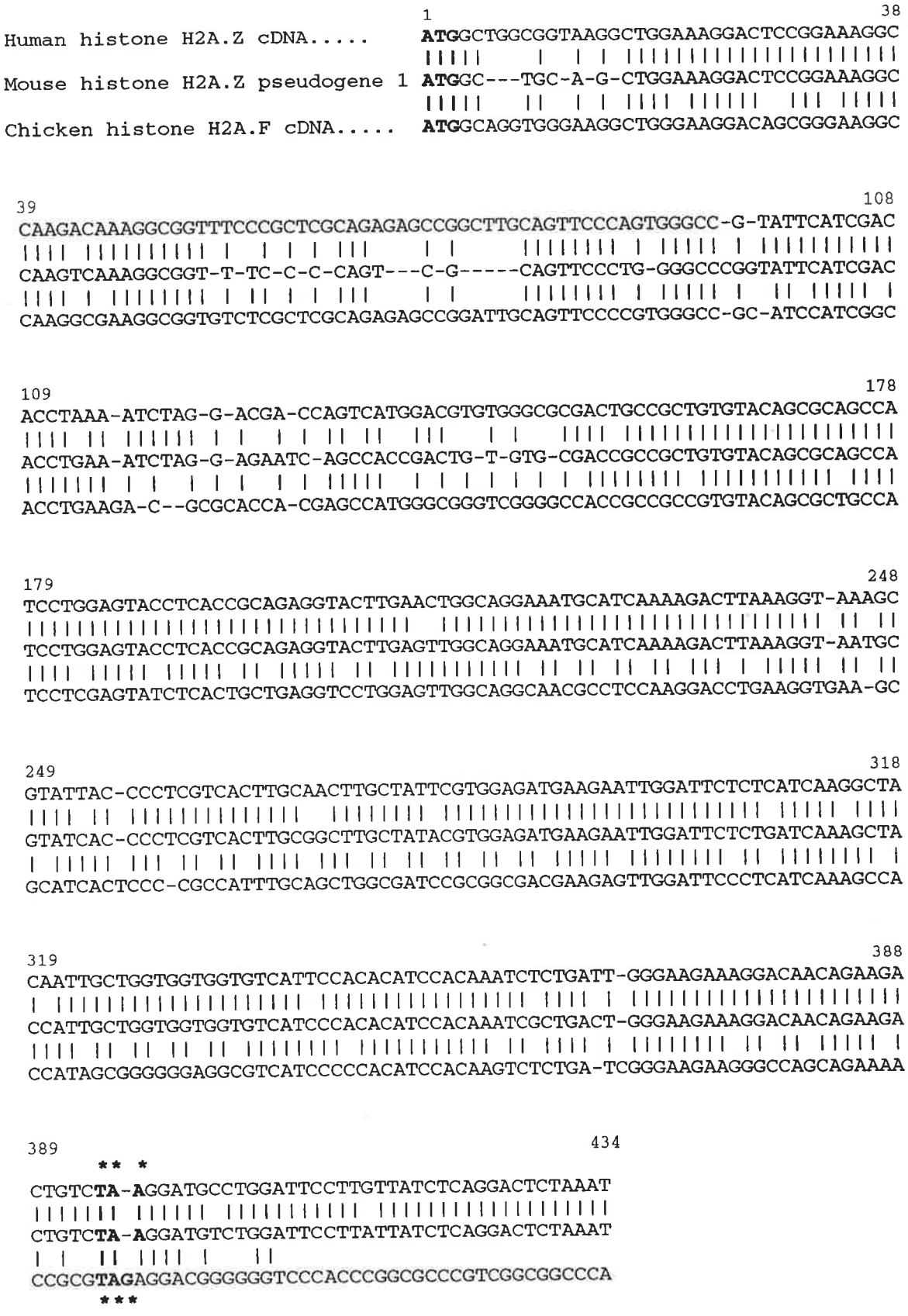
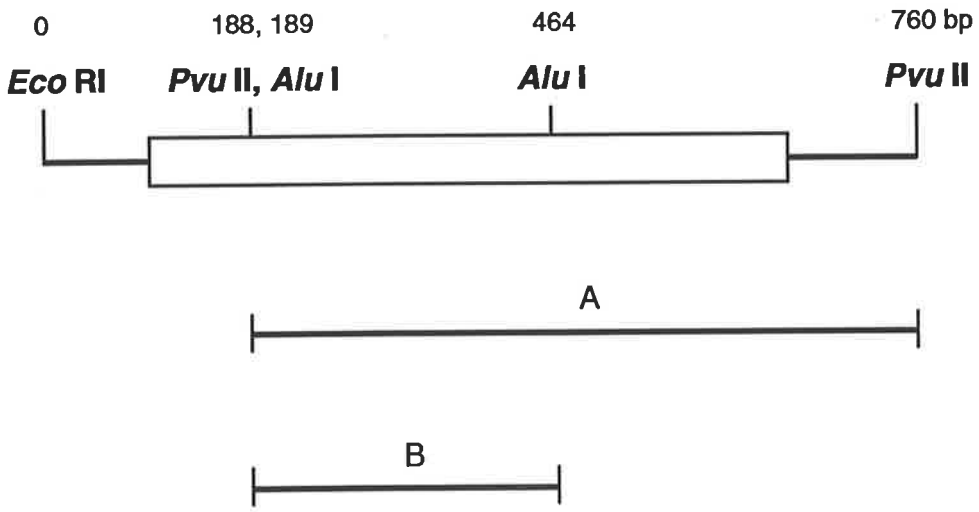

1
38
Human histone H2A.Z cDNA.....
ATGGCTGGCGGTAAGGCTGGAAAGGACTCCGGAAAGGC
Mouse histone H2A.Z pseudogene 1
ATGGC---TGC-A-G-CTGGAAAGGACTCCGGAAAGGC
Chicken histone H2A.F cDNA.....
ATGGCAGGTGGGAAGGCTGGGAAGGACAGCGGGAAGGC
39
108
CAAGACAAAGGCGGTTTCCCGCTCGCAGAGAGCCGGCTTGCAGTTCCCAGTGGGCC-G-TATTCATCGAC
CAAGTCAAAGGCGGT-T-TC-C-C-CAGT---C-G-----CAGTTCCCTG-GGGCCCGGTATTCATCGAC
CAAGGCGAAGGCGGTGTCTCGCTCGCAGAGAGCCGGATTGCAGTTCCCCGTGGGCC-GC-ATCCATCGGC
109
178
ACCTAAA-ATCTAG-G-ACGA-CCAGTCATGGACGTGTGGGCGCGACTGCCGCTGTGTACAGCGCAGCCA
ACCTGAA-ATCTAG-G-AGAATC-AGCCACCGACTG-T-GTG-CGACCGCCGCTGTGTACAGCGCAGCCA
ACCTGAAGA-C--GCGCACCA-CGAGCCATGGGCGGGTCGGGGCCACCGCCGCCGTGTACAGCGCTGCCA
179
248
TCCTGGAGTACCTCACCGCAGAGGTACTTGAACTGGCAGGAAATGCATCAAAAGACTTAAAGGT-AAAGC
TCCTGGAGTACCTCACCGCAGAGGTACTTGAGTTGGCAGGAAATGCATCAAAAGACTTAAAGGT-AATGC
TCCTCGAGTATCTCACTGCTGAGGTCCTGGAGTTGGCAGGCAACGCCTCCAAGGACCTGAAGGTGAA-GC
249
318
GTATTAC-CCCTCGTCACTTGCAACTTGCTATTCGTGGAGATGAAGAATTGGATTCTCTCATCAAGGCTA
GTATCAC-CCCTCGTCACTTGCGGCTTGCTATACGTGGAGATGAAGAATTGGATTCTCTGATCAAAGCTA
GCATCACTCCC-CGCCATTTGCAGCTGGCGATCCGCGGCGACGAAGAGTTGGATTCCCTCATCAAAGCCA
319
388
CAATTGCTGGTGGTGGTGTCATTCCACACATCCACAAATCTCTGATT-GGGAAGAAAGGACAACAGAAGA
CCATTGCTGGTGGTGGTGTCATCCCACACATCCACAAATCGCTGACT-GGGAAGAAAGGACAACAGAAGA
CCATAGCGGGGGGAGGCGTCATCCCCCACATCCACAAGTCTCTGA-TCGGGAAGAAGGGCCAGCAGAAAA
389
434
CTGTCTA-AGGATGCCTGGATTCCTTGTTATCTCAGGACTCTAAAT
CTGTCTA-AGGATGTCTGGATTCCTTATTATCTCAGGACTCTAAAT
CCGCGTAGAGGACGGGGGGTCCCACCCGGCGCCCGTCGGCGGCCCA

Figure 3-5

Map of the probes using for mouse genomic library I screening and Southern analysis.

The map of the 567 bp *Pvu* II fragment probe (probe A) derived from pBS17-0.55 (Fig. 3-2) and the 275 bp *Alu* I fragment probe (probe B) derived from the 760 bp *Eco* RI to *Pvu* II fragment are shown. Sequences homologous to histone H2A.Z are boxed. The sequence is measured in base pairs, relative to the *Eco* RI site located at the 5' end of the region, and the positions of the restriction enzyme sites are also indicated. When probe A was used for library screening there was high non-specific hybridisation. Probe B, derived from probe A, contained only sequences homologous to H2A.Z and was preferred for Southern analysis.


0
188, 189
464
760 bp
Eco RI
Pvu II, Alu I
Alu I
Pvu II
A
B

non-specific binding of probe to the filters. Since the probe contained a highly repeated T sequence (Fig. 3-3), this may have contributed to the high background. Nevertheless hybridising plaques could be seen and a total of fourteen positive plaques were picked and underwent two more rounds of screening in order to obtain pure positive plaques. Thirteen plaques were obtained after the three rounds of screening and of these, five positives (λ5w, λ12w, λ17w, λ17s, λ18) were chosen for further analysis.

(i) Analysis of five positive clones with the 573 bp *Pvu* II probe

λ DNA was prepared from the above five positives as described (2.3.2). *Eco* RI-digested fragments were analysed on a 1% agarose gel. The *Eco* RI digestion patterns indicated that all five were independent clones and different from λ17 from which the mouse histone H2A.Z pseudogene 1 was derived (Fig. 3-6 A). In order to characterise these clones, Southern analysis with *Eco* RI and *Pvu* II restriction digestion was performed using the 567 bp *Pvu* II fragment from histone H2A.Z pseudogene 1 as probe. *Eco* RI-digested pseudogene λ17 DNA was included on the same filter as a positive control. This probe gave some non-specific hybridisation with all clones (Fig. 3-6 B). However, when the same filter was rehybridised and probed with a 275 bp *Alu* I fragment, derived from the 567 *Pvu* II probe (Fig. 3-5), only λ17w hybridised with this pseudogene 1 "coding" probe. As shown in Fig. 3-6 C, putative coding regions as defined by 275 bp *Alu* I probe were contained within a 6.5 kb *Eco* RI fragment and a 2.8 kb *Pvu* II fragment of the clone λ17w.

(ii) Subcloning and sequencing analysis of clones derived from λ17w

λ17w DNA was prepared as described (2.3.2). DNA was digested with *Eco* RI and the 6.5 kb *Eco* RI fragment was isolated and subcloned into

Figure 3-6

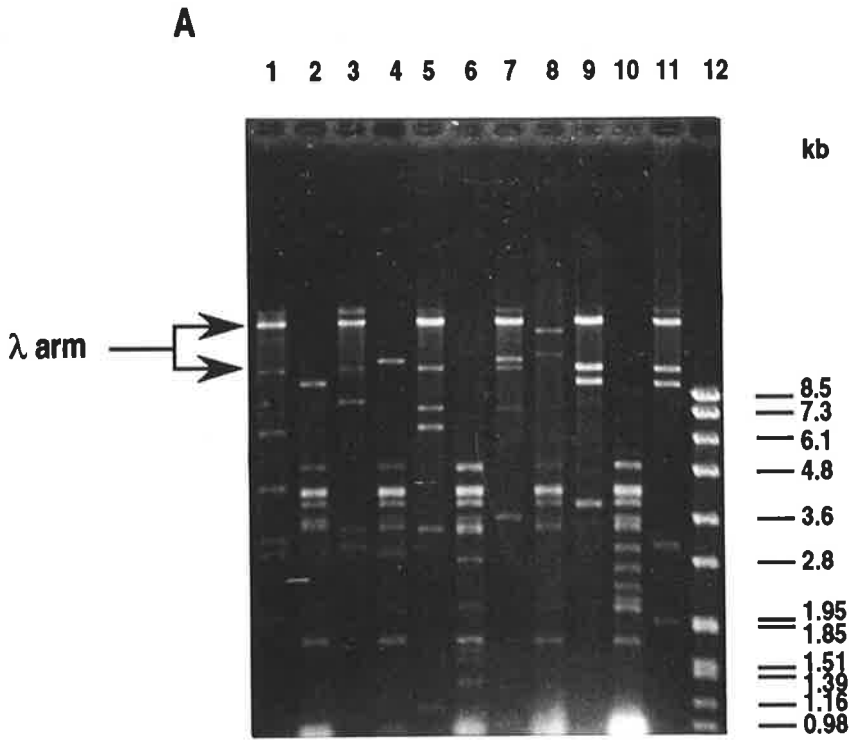
Southern analysis of λ positive clones derived from mouse genomic library I

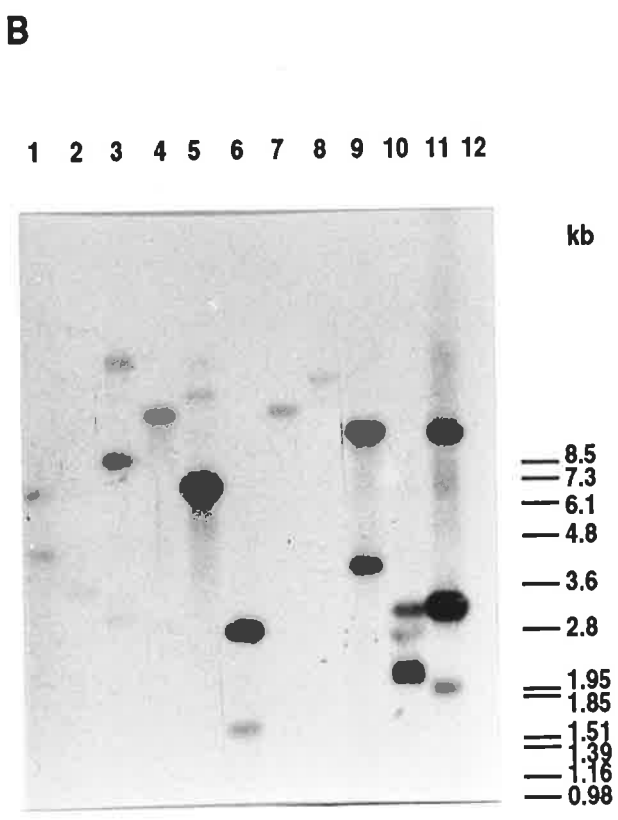
A: λ DNA from each of the λ5w, λ12w, λ17w, λ17s, λ18 clones was digested with *Eco* RI and *Pvu* II, and run with *Eco* RI-digested DNA of the λ17 clone, from which the mouse histone H2A.Z pseudogene 1 was derived, as a positive control. The digested DNA was electrophoresed on 1% agarose gel along with DNA size markers. DNA was visualised by staining with ethidium bromide. Lanes are as follows:

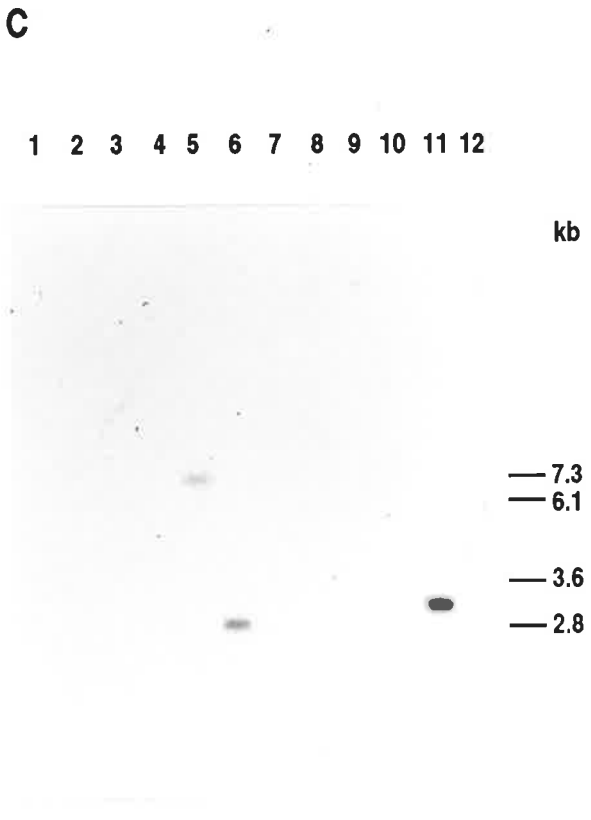
1) *Eco* RI-digested λ17s DNA
2) *Pvu* II-digested λ17s DNA
3) *Eco* RI-digested λ18 DNA
4) *Pvu* II-digested λ18 DNA
5) *Eco* RI-digested λ17w DNA
6) *Pvu* II-digested λ17w DNA
7) *Eco* RI-digested λ5w DNA
8) *Pvu* II-digested λ5w DNA
9) *Eco* RI-digested λ12w DNA
10) *Pvu* II-digested λ12w DNA
11) *Eco* RI-digested λ17 DNA
12) *Eco* RI-digested SPP1 phage DNA size markers

B: DNA was transferred from the above gel to a GeneScreen membrane, probed with ^{32}P-labelled 567 bp *Pvu* II fragment, containing the homologous sequences to H2A.Z (Fig. 3-5), and the hybridising bands were detected by autoradiography. Sizes of the markers are shown. Lanes are the same as in **A**.

C: The same membrane as in **B** after stripping and re-probing with ^{32}P-labelled 275 bp *Alu* I fragment, the internal fragment of 567 bp *Pvu* II fragment (Fig. 3-5). The hybridising bands were detected by autoradiography. Sizes of the markers are shown. Lanes are the same as in **A**.


A
1 2 3 4 5 6 7 8 9 10 11 12
kb
λ arm
8.5
7.3
6.1
4.8
3.6
2.8
1.95
1.85
1.51
1.39
1.16
0.98


B
1 2 3 4 5 6 7 8 9 10 11 12
kb
8.5
7.3
6.1
4.8
3.6
2.8
1.95
1.85
1.51
1.39
1.16
0.98


C
1 2 3 4 5 6 7 8 9 10 11 12
kb
7.3
6.1
3.6
2.8

Eco RI-digested pBluescript KS+ vector. The recombinant plasmid, p17w-RI was obtained (Fig. 3-7 A).

p17w-RI was digested with *Pvu* II and a 2.8 kb fragment was isolated and subcloned into *Eco* RV-digested pBluescript KS+ vector. This subclone was named p17w-Pvu II (Fig. 3-7 B).

Sequence analysis of p17w-Pvu II was initially performed at both ends of the 2.8 kb *Pvu* II fragment using double-stranded DNA as template with USP and RSP sequencing primers. The sequence was compared to mouse histone H2A.Z pseudogene 1 and it was found that the sequence obtained from the RSP primer had significant homology to mouse histone H2A.Z pseudogene 1. In an attempt to obtain sequences at the 3' end of the clone, sequence analysis was performed using the PHZ2 (2.2.15), (a primer complementary to a specific sequence within the insert) and a total sequence of 503 bases was obtained.

The 503 nucleotide sequence from p17W-PvuII was compared to mouse histone H2A.Z pseudogene 1 and was found to be highly homologous (approximately 90 %) (Fig. 3-8). Since the 503 bp region covers the entire H2A.Z coding region (Fig. 3-8) and given that chicken and human H2A.Z genes contain introns at identical positions (Dalton *et al.*, 1989; Hatch and Bonner, 1990) it is highly likely that the 503 bp fragment represents another mouse histone H2A.Z pseudogene. It is referred to as mouse histone H2A.Z pseudogene 2. Comparison of p17w-Pvu II sequence to chicken histone H2A.F cDNA (Harvey *et al.*, 1983) and human H2A.Z cDNA (Hatch and Bonner, 1988) showed 80% and 91% homology respectively (Fig. 3-9).

If mouse histone H2A.Z pseudogene 2 codons are compared with those from human H2A.Z cDNA (Hatch and Bonner, 1988), it can be seen that they encode identical proteins despite the fact that there are 20 single nucleotide changes between the two (Fig. 3-10).

Figure 3-7

A : Restriction map of p17w-RI containing a 6.5 kb *Eco* RI insert in pBluescript KS+. The internal fragment size is indicated in kilobases. The polylinker restriction sites are boxed in black and abbreviations used are as follows:

K	*Kpn I*
RV	*Eco* RV
P	*Pst* I
S	*Sac* I

B : The map of p17w-Pvu II, containing a 2.8 kb *Pvu* II fragment derived from p17w-RI in pBluescript KS+ is shown. The numbers indicate the distances in basepairs from the *Hind* III restriction site on the left side of the diagram. The position of the mouse histone H2A.Z pseudogene 2 (represented by the striped box) is shown. The thick arrow indicates the orientation of the mouse histone H2A.Z pseudogene 2. The strategy used to obtain the sequence is indicated by the arrows. A total sequence of 503 bp was obtained. The polylinker restriction sites are boxed in black and abbreviations used are as follows:

H	*Hind* III
E	*Eco* RI

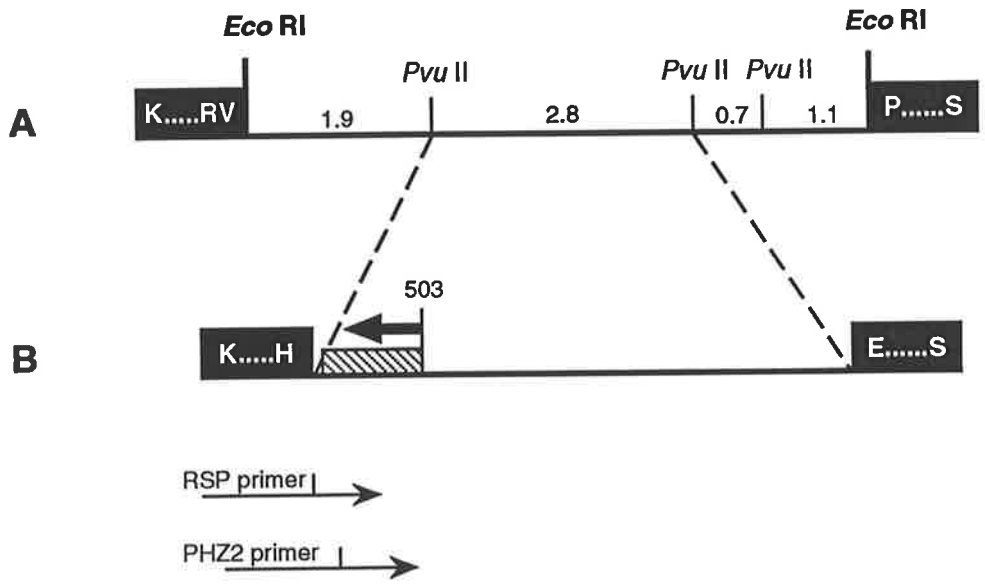

A
B
Eco RI
Pvu II
Pvu II
Pvu II
Eco RI
K.....RV
1.9
2.8
0.7
1.1
P.....S
503
K.....H
E.....S
RSP primer
PHZ2 primer

Figure 3-8

Comparison of nucleotide sequences for mouse histone H2A.Z pseudogene 1 and mouse histone H2A.Z pseudogene 2

The alignment of mouse histone H2A.Z pseudogene 1 (top) to mouse histone H2A.Z pseudogene 2 (bottom) is illustrated. The homology region spans 498 nucleotides and starts from 65 bases 5' of ATG (indicated by ###) through TAA (indicated by ***) to 43 bases 3' of the stop codon and the identity of this region is 89%.

Mouse histone H2A.Z Pseudogene 1.....

Mouse histone H2A.Z Pseudogene 2.....

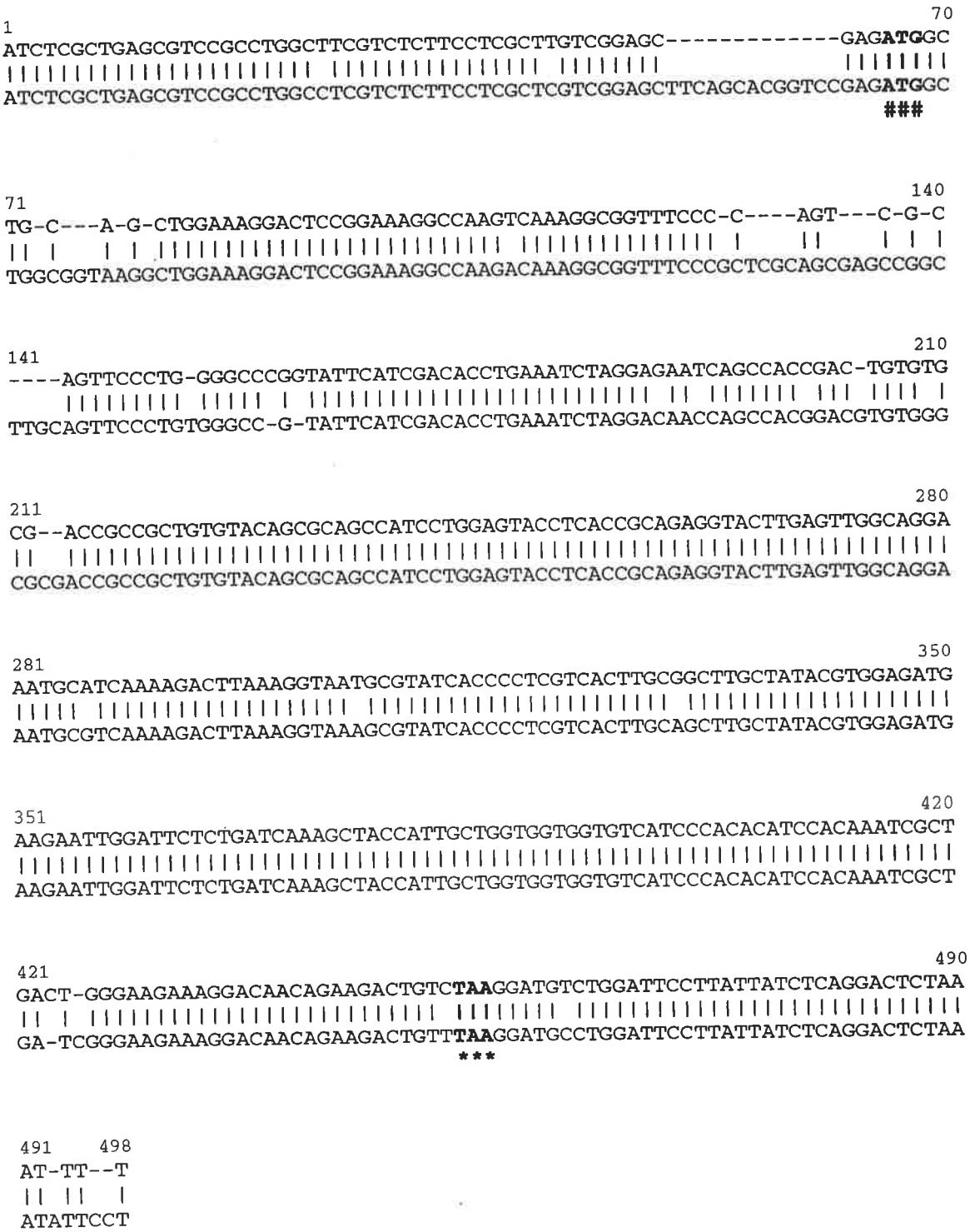


Figure 3-9

Comparison of nucleotide sequences for mouse histone H2A.Z pseudogene 2, human histone H2A.Z cDNA and chicken histone H2A.F cDNA

The alignment of mouse histone H2A.Z pseudogene 2 to human histone H2A.Z cDNA (Hatch and Bonner, 1988) (top) and chicken histone H2A.F cDNA (Harvey *et al.*, 1983) (bottom) is illustrated. The homology region spans from ATG through TAA (indicated by ***) in human histone H2A.Z cDNA to 49 bases 3' of TAA, and the identity of this region is 91%. Comparison with chicken histone H2A.F cDNA, from ATG to TAG (indicated by ***) indicates 80 % nucleotide homology to the mouse sequence.

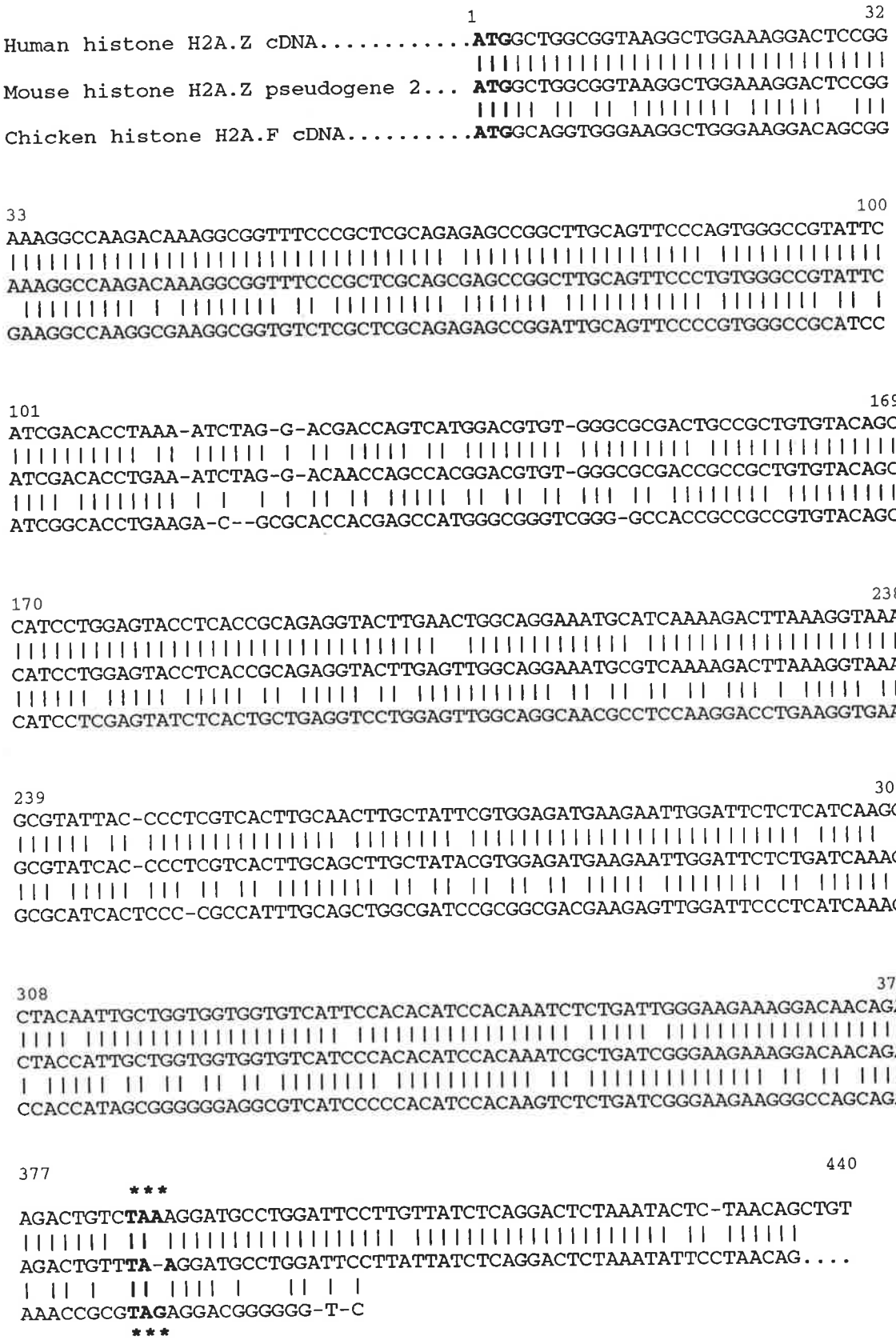

1
32
Human histone H2A.Z cDNA.....ATGGCTGGCGGTAAGGCTGGAAAGGACTCCGG
Mouse histone H2A.Z pseudogene 2... ATGGCTGGCGGTAAGGCTGGAAAGGACTCCGG
Chicken histone H2A.F cDNA.....ATGGCAGGTGGGAAGGCTGGGAAGGACAGCGG
33
100
AAAGGCCAAGACAAAGGCGGTTTCCCGCTCGCAGAGAGCCGGCTTGCAGTTCCCAGTGGGCCGTATTC
AAAGGCCAAGACAAAGGCGGTTTCCCGCTCGCAGCGAGCCGGCTTGCAGTTCCCTGTGGGCCGTATTC
GAAGGCCAAGGCGAAGGCGGTGTCTCGCTCGCAGAGAGCCGGATTGCAGTTCCCCGTGGGCCGCATCC
101
169
ATCGACACCTAAA-ATCTAG-G-ACGACCAGTCATGGACGTGT-GGGCGCGACTGCCGCTGTGTACAGC
ATCGACACCTGAA-ATCTAG-G-ACAACCAGCCACGGACGTGT-GGGCGCGACCGCCGCTGTGTACAGC
ATCGGCACCTGAAGA-C--GCGCACCACGAGCCATGGGCGGGTCGGG-GCCACCGCCGCCGTGTACAGC
170
238
CATCCTGGAGTACCTCACCGCAGAGGTACTTGAACTGGCAGGAAATGCATCAAAAGACTTAAAGGTAAA
CATCCTGGAGTACCTCACCGCAGAGGTACTTGAGTTGGCAGGAAATGCGTCAAAAGACTTAAAGGTAAA
CATCCTCGAGTATCTCACTGCTGAGGTCCTGGAGTTGGCAGGCAACGCCTCCAAGGACCTGAAGGTGAA
239
307
GCGTATTAC-CCCTCGTCACTTGCAACTTGCTATTCGTGGAGATGAAGAATTGGATTCTCTCATCAAGG
GCGTATCAC-CCCTCGTCACTTGCAGCTTGCTATACGTGGAGATGAAGAATTGGATTCTCTGATCAAAG
GCGCATCACTCCC-CGCCATTTGCAGCTGGCGATCCGCGGCGACGAAGAGTTGGATTCCCTCATCAAAG
308
376
CTACAATTGCTGGTGGTGGTGTCATTCCACACATCCACAAATCTCTGATTGGGAAGAAAGGACAACAGA
CTACCATTGCTGGTGGTGGTGTCATCCCACACATCCACAAATCGCTGATCGGGAAGAAAGGACAACAGA
CCACCATAGCGGGGGGAGGCGTCATCCCCCACATCCACAAGTCTCTGATCGGGAAGAAGGGCCAGCAGA
377
440
***
AGACTGTCTAAAGGATGCCTGGATTCCTTGTTATCTCAGGACTCTAAATACTC-TAACAGCTGT
AGACTGTTTA-AGGATGCCTGGATTCCTTATTATCTCAGGACTCTAAATATTCCTAACAG....
AAACCGCGTAGAGGACGGGGGG-T-C
***

Figure 3-10

Comparison of the coding sequence of human histone H2A.Z cDNA and mouse histone H2A.Z pseudogene 2

The alignment of coding sequence (ATG to TAA), for human histone H2A.Z protein (Hatch and Bonner, 1988) (top) to mouse histone H2A.Z pseudogene 2 (bottom) is illustrated. An asterisk (*) indicates a base substitution in mouse histone H2A.Z pseudogene 2. The two nucleotide sequences encode the same amino acid sequence (shown in the single letter code).

Human histone H2A.Z cDNA .....

Mouse histone H2A.Z pseudogene 2.....

```
    A   G   G   K   A   G   K   D   S   G   K   A   K   T   K
ATG GCT GGC GGT AAG GCT GGA AAG GAC TCC GGA AAG GCC AAG ACA AAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
ATG GCT GGC GGT AAG GCT GGA AAG GAC TCC GGA AAG GCC AAG ACA AAG

 A   V   S   R   S   Q   R   A   G   L   Q   F   P   V   G   R
GCG GTT TCC CGC TCG CAG AGA GCC GGC TTG CAG TTC CCA GTG GGC CGT
||| ||| ||| ||| ||| |||  || ||| ||| ||| ||| ||| ||  ||| ||| |||
GCG GTT TCC CGC TCG CAG CGA GCC GGC TTG CAG TTC CCT GTG GGC CGT
                         *                       *

 I   H   R   H   L   K   S   R   T   T   S   H   G   R   V   G
ATT CAT CGA CAC CTA AAA TCT AGG ACG ACC AGT CAT GGA CGT GTG GGC
|||  ||  ||| ||| ||  ||| ||| ||| ||  ||| ||  ||  ||| ||| ||| |||
ATT CAT CGA CAC CTG AAA TCT AGG ACA ACC AGC CAC GGA CGT GTG GGC
                 *                   *       *   *

 A   T   A   A   V   Y   S   A   A   I   L   E   Y   L   T   A
GCG ACT GCC GCT GTG TAC AGC GCA GCC ATC CTG GAG TAC CTC ACC GCA
|||  ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
GCG ACC GCC GCT GTG TAC AGC GCA GCC ATC CTG GAG TAC CTC ACC GCA
      *

 E   V   L   E   L   A   G   N   A   S   K   D   L   K   V   K
GAG GTA CTT GAA CTG GCA GGA AAT GCA TCA AAA GAC TTA AAG GTA AAG
|||  ||  ||| ||   ||  ||| ||| ||| ||  ||| ||| ||| ||| ||| ||| |||
GAG GTA CTT GAG TTG GCA GGA AAT GCG TCA AAA GAC TTA AAG GTA AAG
              * *                   *

 R   I   T   P   R   H   L   Q   L   A   I   R   G   D   E   E
CGT ATT ACC CCT CGT CAC TTG CAA CTT GCT ATT CGT GGA GAT GAA GAA
|||  ||  ||| ||| ||| ||| ||| ||  ||| ||| ||  ||| ||| ||| ||| |||
CGT ATC ACC CCT CGT CAC TTG CAG CTT GCT ATA CGT GGA GAT GAA GAA
      *                       *               *

 L   D   S   L   I   K   A   T   I   A   G   G   G   V   I   P
TTG GAT TCT CTC ATC AAG GCT ACA ATT GCT GGT GGT GGT GTC ATT CCA
|||  ||  ||| ||  ||| ||  ||| ||  ||| ||| ||| ||| ||| ||  ||| |||
TTG GAT TCT CTG ATC AAA GCT ACC ATT GCT GGT GGT GGT GTC ATC CCA
              *       *       *                           *

 H   I   H   K   S   L   I   G   K   K   G   Q   Q   K   T   V
CAC ATC CAC AAA TCT CTG ATT GGG AAG AAA GGA CAA CAG AAG ACT GTC TAA
|||  ||  ||| ||| ||  ||| ||  ||| ||| ||| ||| ||| ||| ||| ||  ||| |||
CAC ATC CAC AAA TCG CTG ATC GGG AAG AAA GGA CAA CAG AAG ACT GTT TAA
                  *       *                               *
```

3.2.3 Isolation of an authentic mouse histone H2A.Z genomic clone

(i) Mouse genomic library II screening

The attempt to isolate the mouse histone H2A.Z gene from the first library was unsuccessful. This result may be due to quality of the library and specificity of the probe. It was possible that this library did not contain the DNA sequence representative of the entire genome. In the second attempt, a new mouse genomic library (II) (2.2.7) was screened using a 275 bp *Alu* I fragment, containing only sequences homologous to H2A.Z (see Fig. 3-5). The library was plated as described (2.3.11). The probe was labelled with α-^{32}P dATP (2.3.8 (ii)) and then hybridised to the filters. After hybridisation the filters were washed at a stringency of 0.5x SSC, 0.1% SDS at 65°C and set up for autoradiography.

A total of sixteen positive plaques were picked and these underwent two more rounds of screening in order to obtain independent plaques. Eight positives (λ4.1, λ5.1, λ5.2, λ8, λ10, λ12, λ13, λ14) were analysed further.

(ii) PCR analysis of eight positive clones

Preliminary characterisation was carried out by PCR. Based on the fact that the histone H2A.Z gene from chicken and human contains introns, PCR primers was designed to anneal at the exon-intron boundaries. By inspection of mouse histone H2A.Z pseudogene 1 and 2 coding regions and the conserved exon-intron boundary positions in chicken and human H2A.Z genomic clones, it was possible to design exon-annealing primers to analyse the eight putative mouse genomic clones. Two primers, PHZ1 and PHZ2 (2.2.15), were designed from sequences within exon 4 and 5 which contained 100% sequence identity between the two pseudogenes. It was hoped that this conserved sequence would be identical with the mouse histone H2A.Z genomic sequence. In order to ensure PCR with these two primers was satisfactory, the PCR reaction using *Taq* DNA polymerase was first tested using mouse genomic DNA as a template (see Fig. 3-

11, lane 10). The predicted approximately 300 bp amplified product, containing intron 4 and part of exons 4 and 5, was visualised in a 2% agarose gel after 35 cycles of PCR reaction (95°C for 1 minute, 54°C for 1 minute and 45 seconds, 72°C for 2 minutes).

λ DNA was prepared from the eight positives (λ4.1, λ5.1, λ5.2, λ8, λ10, λ12, λ13, λ14) as described (2.3.2) and amplified under the same conditions as for mouse genomic DNA. The result showed that only λ13 DNA gave the predicted amplified product, the same size as the mouse genomic DNA fragment (Fig. 3-11, lane 8), and this clone was analysed further.

(iii) Southern analysis and subclone of clone λ13

The insert was cleaved from the vector by restriction enzyme digestion using *Sal* I.

To ensure that clone λ13 contained the mouse H2A.Z gene, Southern analysis was performed by digestion of DNA with *Sal* I, *Eco* RI, and the combination of *Sal* I and *Eco* RI, and a 275 bp *Alu* I fragment was used as a probe (see Fig. 3-5). The result showed that clone λ13 contained an insert of approximately 14.3 kb with two internal *Eco* RI sites. The probe hybridised with a 4.8 kb *Eco* RI/*Sal* I fragment (Fig. 3-12).

Further characterisation of DNA from clone λ13 was carried out on the 4.8 kb *Eco* RI/*Sal* I fragment after subcloning into an *Eco* RI/*Sal* I-digested pBluescript KS+ vector. The recombinant plasmid was named pES4.8.

(iv) Sequencing of the mouse histone H2A.Z gene

In an attempt to define the map of pES4.8 in more detail and determine the fragment containing the H2A.Z sequences for subcloning and sequencing analysis, pES4.8 DNA was digested with several restriction enzymes (and combinations) and Southern blotted using the 275 bp *Alu I* fragment as

Figure 3-11

PCR analysis of eight positive clones

A : PCR reactions were performed with DNA template (10 ng) from eight positive clones, derived from mouse genomic library II and from mouse genomic DNA (100 ng) as a positive control. Samples (7 µl) from each of the reactions, along with DNA size markers, were electrophoresed on a 2% agarose gel. DNA was visualised by staining with ethidium bromide. Lanes are as follows:

1) *Eco* RI-digested SPP1 phage DNA size markers (500 ng)
2) λ4.1 DNA as a template
3) λ5.1 DNA as a template
4) λ5.2 DNA as a template
5) λ8 DNA as a template
6) λ10 DNA as a template
7) λ12 DNA as a template
8) λ13 DNA as a template
9) λ14 DNA as a template
10) mouse genomic DNA as a template
11) negative control (sterile water)
12) *Hpa* II-digested pUC19 DNA size markers (1 µg)

The PCR product, approximately 300 bp in size, from lane 8 indicated that λ13 contains intron 4, which is part of the mouse histone H2A.Z gene.

B: The diagram shows the sequence and the annealing sites of the two primers, PHZ1 (top) and PHZ2 (bottom), used in the PCR reaction. The arrow represents the primer annealing site and the direction of newly synthesised DNA from the PCR reaction.

A


1 2 3 4 5 6 7 8 9 10 11 12
PCR product
331 bp
242 bp

B

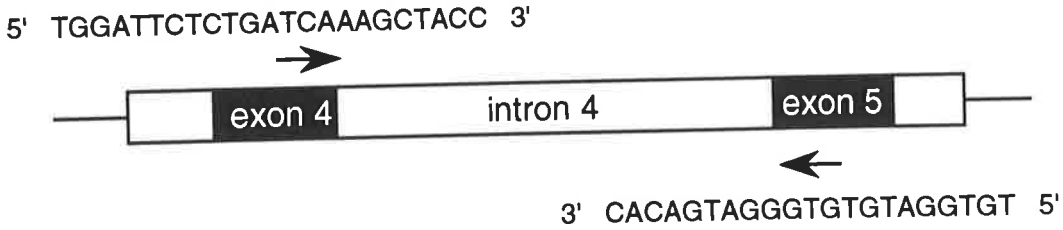

5' TGGATTCTCTGATCAAAGCTACC 3'
exon 4
intron 4
exon 5
3' CACAGTAGGGTGTGTAGGTGT 5'

Figure 3-12

Southern analysis of λ13 derived from mouse genomic library II

A: λ13 DNA, derived from mouse genomic library II (2.2.7), was digested with *Sal* I, *Eco* RI and the combination of *Sal* I and *Eco* RI. The digested DNA was electrophoresed on a 1% agarose gel along with DNA size markers. DNA was visualised by staining with ethidium bromide. Lanes are as follows:

1) Hi-Lo mixed DNA size markers (Bresatec)
2) undigested λ13 DNA
3) *Sal* I-digested λ13 DNA
4) *Eco* RI-digested λ13 DNA
5) *Sal* I and *Eco* RI-digested λ13 DNA

B: DNA was transferred from the above gel to a GeneScreen membrane, probed with ^{32}P-labelled 275 bp *Alu* I fragment, containing homologous sequences to H2A.Z (Fig. 3-5), and the hybridising bands were detected by autoradiography. Size of the markers are shown. Lanes are the same as in **A**.

C: Restriction map of λ13 clone containing a 14.3 kb *Sal* I insert in Lambda EMBL3 SP6/T7 vector is shown. The internal fragment size is indicated in kilobases and Lambda arms are boxed in black. The 4.8 kb *Sal* I/*Eco* RI fragment containing the sequence hybridising with the probe is shown below.

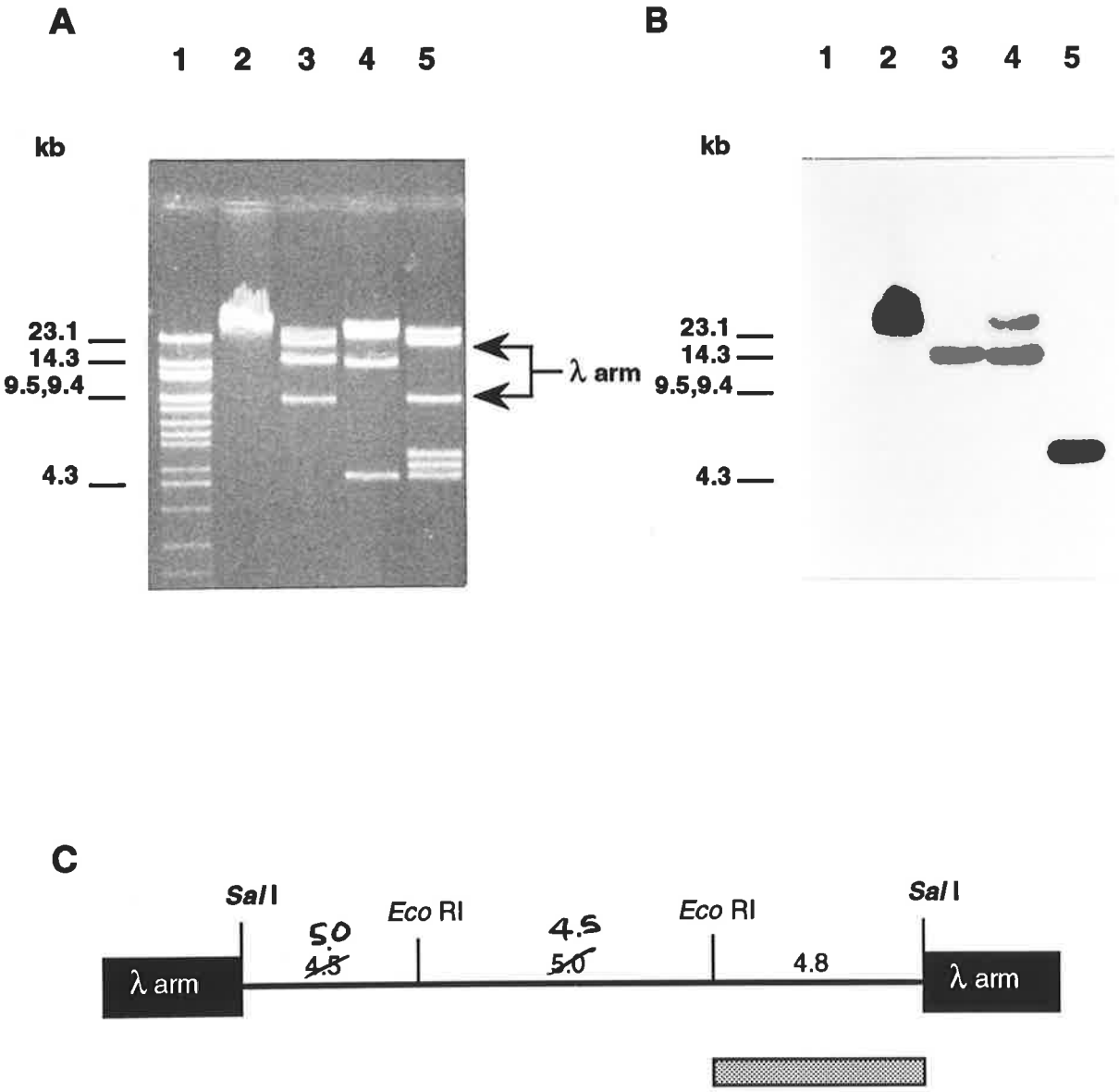

A
1 2 3 4 5
kb
23.1
14.3
9.5,9.4
4.3
λ arm
B
1 2 3 4 5
kb
23.1
14.3
9.5,9.4
4.3
C
Sal I
Eco RI
Eco RI
Sal I
5.0
4.8
4.5
5.0
4.8
λ arm
λ arm

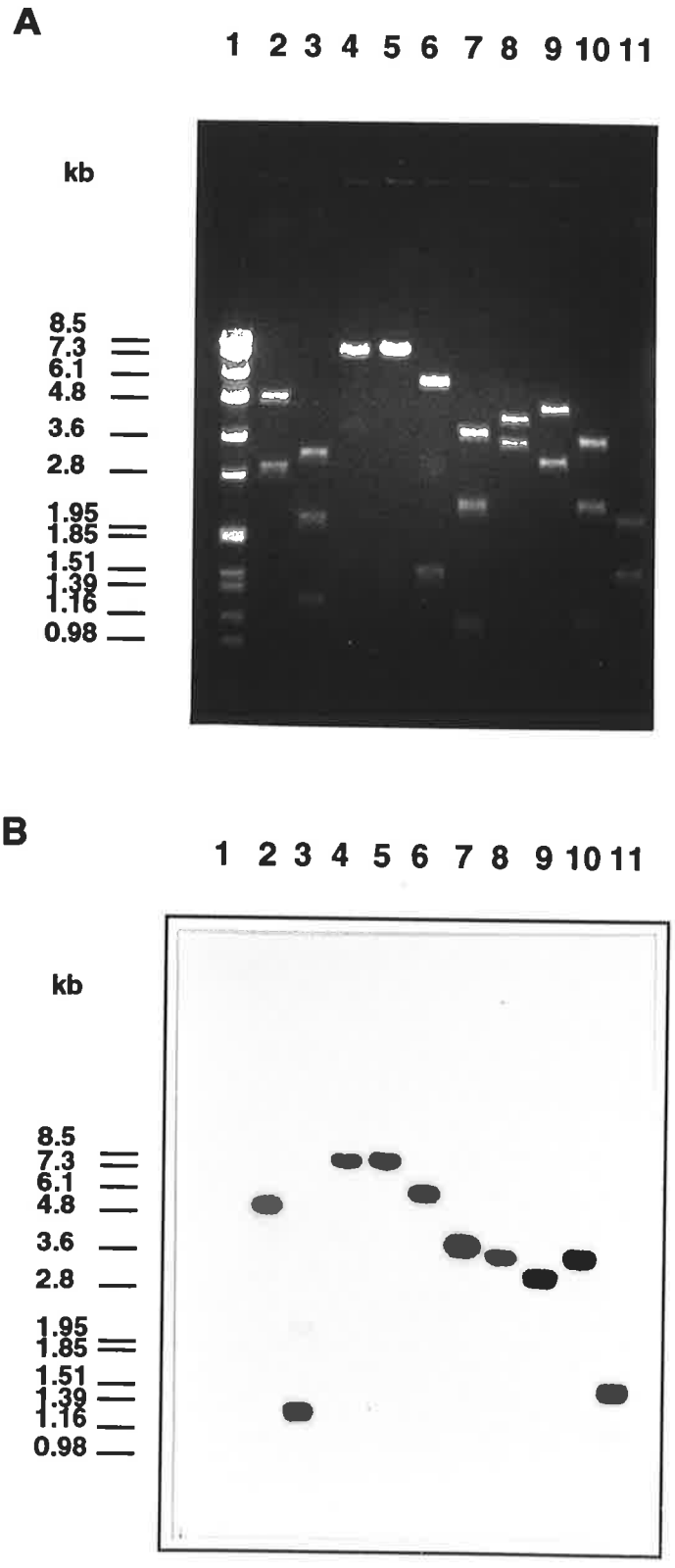
Figure 3-13

Southern analysis of subclone pES4.8 derived from λ13

A : pES4.8, containing a 4.8 kb *Eco* RI/*Sal* I insert fragment in pBluescript KS+, was digested with *Hinc* II, *Hind* III, *Kpn* I, *Pst* I, *Pvu* II, *Xba* I, *Xho* I, and combinations of *Eco* RI and *Sal* I, *Eco* RI and *Pvu* II, *Pst* I and *Pvu* II. The digested DNA was electrophoresed on a 1% agarose gel along with DNA size markers. DNA was visualised by staining with ethidium bromide. Lanes are as follows:

1) Eco RI-digested SPP1 phage DNA size markers
2) *Eco* RI and *Sal* I-digested DNA
3) *Hinc* II-digested DNA
4) *Hind* III-digested DNA
5) *Kpn* I-digested DNA
6) *Pst* I-digested DNA
7) *Pvu* II-digested DNA
8) *Xba* I-digested DNA
9) *Xho* I-digested DNA
10) *Eco* RI and *Pvu* II-digested DNA
11) *Pst* I and *Pvu* II-digested DNA

B : DNA was transferred from the above gel to a GeneScreen membrane, probed with ^{32}P-labelled 275 bp Alu I fragment, containing homologous sequences to H2A.Z (Fig. 3-5), and the hybridising bands were detected by autoradiography. Size of the markers are shown. Lanes are the same as in **A**.



TRACK 11. THE APPARANT 1.5 kb PstI-PvuII FRAGMENT, PICKED UP BY THE PROBE, IS 1.77 kb BY SEQUENCE. SEE FIG. 3,6 PstI +143 CTGCAG TO PvuII +1918 CAGCTG.

probe (Fig. 3-13). From this analysis, the map was drawn as shown in Fig. 3-14 and it indicated that the fragment which hybridised to the probe spanned a 2.0 kb *Pst* I/*Hinc* II fragment. Therefore, a 1.8 kb *Pst* I/*Pvu* II and a 0.5 *Xba* I/*Hinc* II fragment overlapping this region was subcloned into *Pst* I/*Eco* RV-digested pBluescript KS+ vector and *Xba* I/*Hinc* II pBluescript KS+ vector respectively. The sequence of these subclones was obtained by subcloning the fragments further and generating nested deletions with Exonuclease III using an Erase-a-base kit (2.2.3). The sequence of the 2.0 kb *Pst* I/*Hinc* II fragment was compared to the human H2A.Z gene (Hatch and Bonner, 1990). It contained stretches of homology within the coding sequence for exon 2, 3, 4 and 5. Each of exon is flanked by consensus sequences for splice acceptor and donor sites. In addition, the exon-intron boundaries occurred at the same position as in the human sequence. This indicated that the pES4.8 clone contained part of the mouse H2A.Z gene.

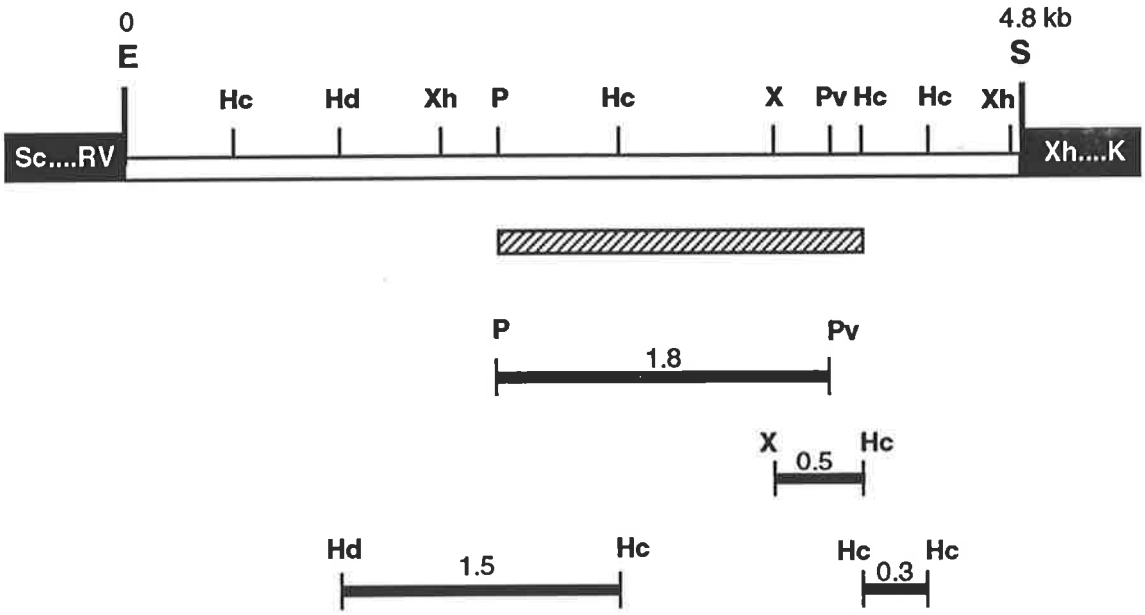
In an attempt to obtain the entire H2A.Z gene sequence, a 1.5 kb *Hind* III/*Hinc* II and a 300 bp *Hinc* II fragment were isolated from pES4.8 and subcloned (see Fig. 3-14). In order to obtain the sequence of the 1.5 kb *Hind* III/*Hinc* II fragment, the subclone was mapped further and two other *Pst* I sites were found, located upstream from the site that was used in subcloning the 1.8 kb *Pst* I/*Pvu* II fragment. These *Pst* I sites were used in subcloning and sequence was obtained. The sequence of the 1.5 kb *Hind* III/*Hinc* II fragment was obtained in both directions using the PHZseq1 and PHZseq2 primers (2.2.15) complementary to the sequence within the insert. As a result, the sequence of 3223 bp, from *Hind* III to *Hinc* II, was obtained.

The strategy used to obtain the 3223 bp of sequence is summarised in Fig. 3-15, and the sequence is shown in Fig. 3-16. The mouse H2A.Z transcript was predicted by comparison of the 3223 bp sequence containing the mouse H2A.Z gene to the human H2A.Z gene (Hatch and Bonner, 1990). The alignment

Figure 3-14

Restriction map of pES4.8

The restriction map of pES4.8 containing a 4.8 kb *Eco* RI/*Sal* I insert in pBluescript KS+ is shown. The polylinker restriction sites are boxed in black and the abbreviations used are shown below. Sequences hybridising to the probe (275 bp *Alu* I fragment containing the H2A.Z homologous sequence) is indicated by the striped box. The fragments subcloned into a plasmid vector for sequence analysis are shown in thick lines and the fragment sizes are indicated in kilobases.



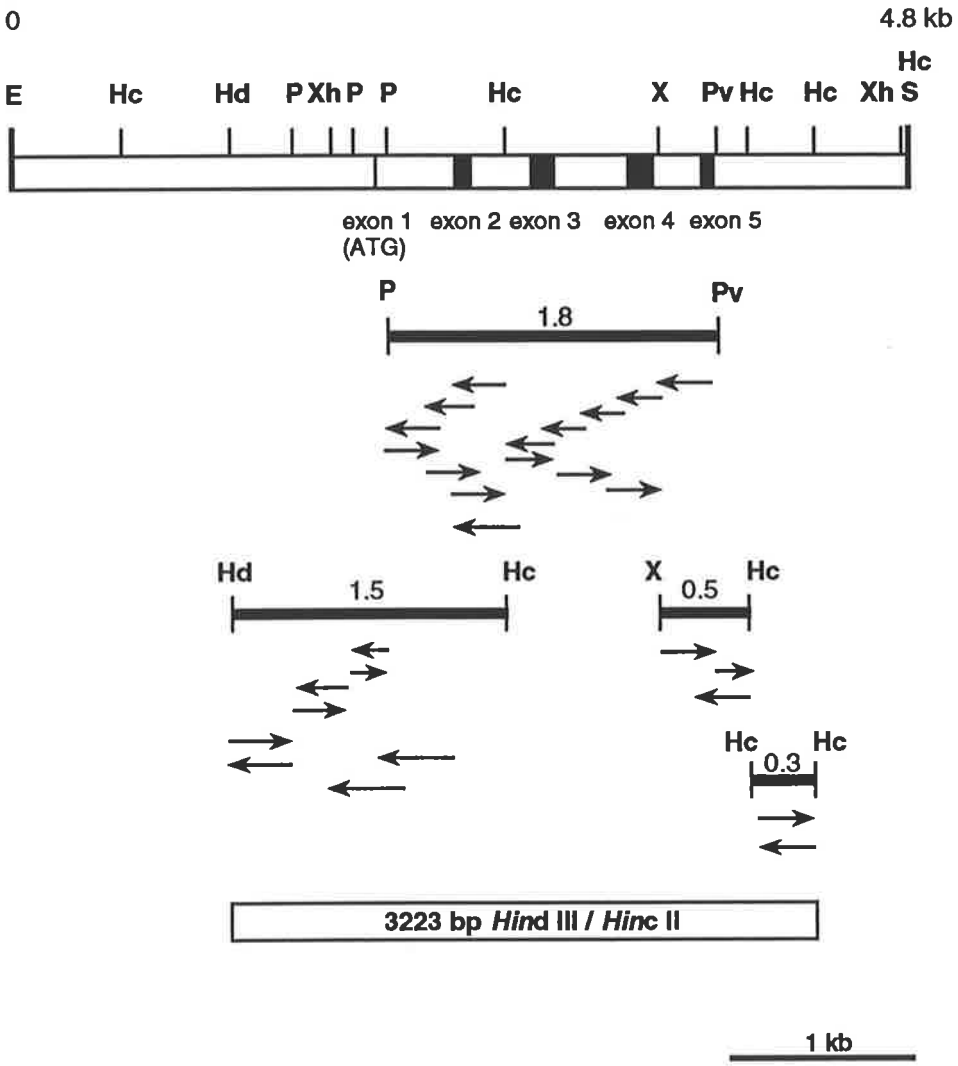
E	*Eco* RI	Hc	*Hinc* II
Hd	*Hind* III	K	*Kpn* I
P	*Pst* I	Pv	*Pvu* II
RV	*Eco* RV	S	*Sal* I
Sc	*Sac* I	X	*Xba* I
Xh	*Xho* I		

Figure 3-15

Sequencing strategy

The sequencing strategy used to obtain the 3223 bp sequence between a *Hin*d III and *Hinc* II site derived from the plasmid pES4.8 is summarised.

The position of the mouse histone H2A.Z gene is shown on the map. The exons are boxed in black. The subcloned fragments are shown in thick lines and the fragment sizes are indicated in kilobases. The directions of the sequence obtained from the subclones are shown by the arrows below each subclone fragment. The abbreviations used for the restriction enzymes are also shown.



E	*Eco* RI	Hc	*Hinc* II
Hd	*Hind* III	P	*Pst* I
Pv	*Pvu* II	S	*Sal* I
X	*Xba* I	Xh	*Xho* I

Figure 3-16

Nucleotide sequence of the mouse histone H2A.Z gene

The sequence of 3223 bp flanked by *Hin*d III and *Hin*c II restriction sites derived from the plasmid pES4.8 is shown.

Nucleotides in the predicted transcript are in capital letters. The sequence is numbered in basepairs relative to the origin of the transcript. The first exon begins at position 1 and the fifth exon ends at position 2249. The mouse H2A.Z gene has five exons interrupted by four introns. The coding regions show the amino acid sequence in single letter code above the nucleotide sequence. The consensus sequences are outlined as follows:

The ATG codon (at position 107) and the termination codon (TAA, at position 1871) are underlined.

The (putative) TATA box at position -30 is shown in green

The (putative) CCAAT boxes at position -78, -123 and -180 are shown in orange.

The (putative) GC boxes at position -93 and -153 are shown in yellow.

The polyadenylation signal (AATAAA) at position 2228-2233 is shown in pink.

Mouse histone H2A.Z sequence

-757 aagcttcctgtgtgtagacaaaattacttgttatccc

-720 cacggtaacttggctccactgttgttatgagacaggatgaggggggtgaggggtaacaag

-660 aggtgaggacagaagtggaggaagatgaaaaaactaccgggggtggggtaggtgacactt

-600 ttatcagggaagtggttcaccaagtagaggccgagctgagctcttcttgccggacgccct

-540 tcgagccacctgtcagaaggcaggactctgaaggaaaccgagatctgcggacacgaaagg

-480 ccaggccagagggcaaccgagaggcgccggttgtcagcgagggggccaggctagaccggg

-420 agggaagctcacgtggtctccggagcaccgcgggggagtggaggggggaggctgcagcgg

-360 ggaggacgccagcgcgcgctcctccccagccttccagctcctccctccaggcacgtgaca

-300 ccaccgtccactcgctcggtggcgcagcgctctgttgctgccctcccgccgcggcgcccc

-240 ttcttctcccggtcctcgtcccgccgcatctcctccccccctaactcatccccacgcg

-180 ccaatcatcgctcgagctcccgagcgccccgccccgccactccgctgtgcgttctcccatt

-120 ggctggagcctcaaggacgcgtcccgagggcgggaggcaaccattggtggccgaaccggc

-60 cgagttcccggatgagggaacattctgcagtataaagggcgcgaggaaggcgggagacgg

+1 CGCAGTTTGAATCGCGGTCCGACGGAGGAGTGGGCGCTGGGATCTCGCTGAGCGTCCGCC 60

TGGCCTCGTCTCTTCCTCGCTCGTCGGAGCTTCAGCACGGTCCGAGATGgtaacgtgaag 120

intron 1

agtgacggcgggcggccgcgggctgcaggctccgagcgggcggggccgccgtccccgggg 180

ctgggagaggcggcggcggcgagtgggagcgtgcggggacggggtcgggacgagcgg 240

aggggaggtgttttcggagccccggagctccaggcgggaggcgacggttgggggggcgcg 300

cacggcgcggctgggagccggcgcgggggcgcgcagccgcggcgggctcgcggcaggggt 360

cggcgggggagaagggcgggcgggggcgagcggcggcccggggtccccggcggggcggg 420

cggcgggcggccgagcgcgcgctcgctcgcctccaggacgcgctgtggtgggggaggggc 480

gcgccgccttggtaattctatcttctccgtcctcccgggctcgcgggcgtcggcggcgcc 540

A G G K A G K D S G K A K T K A V S R

gcagGCTGGCGGTAAGGCTGGAAAGGACTCCGGAAAGGCCAAGACAAAGGCGGTTTCCCG 600

intron 1

S Q R A G L Q

CTCGCAGCGAGCCGGCTTGCAGgtgggtagcaactgccgcagtcgggtgaccgggccttt 660

intron 2

atgtttgcacccccttgcttgtttgtctgcaccctccatgggtgttccatcttagcatg 720

agtgatgtgtcgcgacttggccctgcggagacagtgaatactggggtttggggaggggggg 780

gcgtcggatcacatgtgtgttaactggcggggcgcgttggtgcaccaccaccgtccttgg 840

ctttggcgcgcacatcttatcttaacgaatttcctatctttgtgcatttttttttcttgtt 900

```
                                                F  P  V  G  R  I  H  R  H  L
tgttcgtttgtttgtgtgcataaaaatcaagTTCCCTGTGGGCCGTATTCATCGACACCT    960
                              intron 2 ◄┘

 K  S  R  T  T  S  H  G  R  V  G  A  T  A  A  V  Y  S  A  A
GAAATCTAGGACAACCAGCCACGGACGTGTGGGCGCGACCGCCGCTGTGTACAGCGCAGC   1020

 I  L  E  Y  L  T  A  E
CATCCTGGAGTACCTCACCGCAGAGgtaagcgggagtacgcctgtgatctggaaaagctc   1080
                          └► intron 3
tgggggggcgggggaaagtggcgagagagaaggaggaaagtgatgcaagaaaactcgcag   1140

gccctaagctaagagcggcgcgcagaccgagggagcgctagagggagctcgcgttcagcc   1200

gaggaggctgctgggccagagtgtgctgctcctgcaaatgcaattgcgtgccccttcggg   1260

ttttttgggtggtttttttttttttttttgcataattctgcataattttatcattcagg   1320

tagtagggaatgacaacagttggtaaataacaccattgaggttcaaagatcttgattgaa   1380

                                                        V  L
agtaagttggaatattgatcatcagattttttatttatttattttttgccttttagGTACT   1440
                                                  intron 3 ◄┘

 E  L  A  G  N  A  S  K  D  L  K  V  K  R  I  T  P  R  H  L
TGAGTTGGCAGGAAATGCGTCAAAAGACTTAAAGGTAAAGCGTATCACCCCTCGTCACTT   1500

 Q  L  A  I  R  G  D  E  E  L  D  S  L  I  K  A  T  I  A  G
GCAGCTTGCTATACGTGGAGATGAAGAATTGGATTCTCTGATCAAAGCTACCATTGCTGG   1560

 G  G
TGGTGgtatgtcatccctaaactcctaacattctatttaagaagaaaattcctacactcc   1620
      └► intron 4
cattgtttctagaagagatttcagtacagggtgctttcatggggtttggctttccactta   1680

gtttttgctaagatcattgttccatccctcaagttttcactttgtatatgaattttacct   1740

gctttgagttcatgtattgtttttttaaatgtttatccagaaggtagacttaatctgtgt   1800

                  V  I  P  H  I  H  K  S  L  I  G  K  K  G  Q  Q
ccctctcctagGTGTCATCCCACACATCCACAAATCGCTGATCGGGAAGAAAGGACAACA   1860
  intron 4 ◄┘

 K  T  V  End
GAAGACTGTTTAAGGATGCCTGGATTCCTTATTATCTCAGGACTCTAAATATTCCTAACA   1920

GCTGTCCAGTGTTGGTGATTCCAGTGGACTGTATCTCTGTGAAAAACACAATTTTGCCTT   1980

TTTGTAATTCTATTTGAGCAAGTTGGAGGCTTAATTAGCCTTCCAACCAACCAAATTTCT   2040

GCATTCGAGTCTTAACCATATTTAAGTGTTACTGTGGCTTCAAAGAAGCTATTGATTCTG   2100

AAGTAGTGGGTTTTGATTGAGTTGACTGTTTTTAAAAAACTGTTTGGATTTTAATTGTGA   2160

TGCAGAAGTTATAGTAACAAGCATTTGGTTTTGTACAGACATTGTTTCCACTCTGGTGGA   2220

TAAGCTCAATAAAGGTCATATCCCAAACTagctttaaacttgcttaataatcgggtctta   2280

ccttagatctcactcagcaacaagtacattctctgcttactaattaaacagtgcatctgt   2340

agtcataaccccgtggtacttttgttgtttctttgtatcatatgacctcttagggtctta   2400

gaacttgggttgttgctggaggtaactgcttttggcgtccccccccccccccaatggaca   2460

gttaac 2466
```

of the transcript sequence of the two genes showed that the sequence identity of the 5' untranslated region and coding sequence are 85% and 91% respectively. The strongest homology (98%) is observed for the 3' untranslated region. The features of the nucleotide sequence presented in Fig. 3-16 can be summarised as follows:

(1) Starting from the predicted 5'-cap nucleotide to the polyadenylation site, the mouse H2A.Z gene contains 2249 nucleotides.

(2) The mouse H2A.Z gene has five exons that are interrupted by four introns, predicting a transcript of 869 nucleotides prior to polyadenylation.

(3) Between the 5'-cap and ATG, the 5'-untranslated region is 109 nucleotides long. The first exon contains the 5'-untranslated region and the initiation codon (ATG). The second exon and the third exon are 78 and 114 nucleotides encoding 26 and 38 amino acids, respectively. The fourth exon is the largest, encoding 44 amino acids. The fifth exon is 62 nucleotides encoding 19 amino acids. Thus, the mouse histone H2A.Z protein is predicted to be 127 amino acids long.

(4) Between the chain termination TAA (nucleotide 1871-1873) and the predicted polyadenylation site (nucleotide 2249) there are 376 nucleotides, including the polyadenylation signal AATAAA (nucleotide 2228-2233).

(5) Inspection of the 5' flanking region identified a (putative) TATA box (located 30 bp upstream from the transcription start site), three (putative) CCAAT boxes (located 78, 123 and 180 bp upstream from the transcription start site) and two (putative) GC boxes (located 93 and 153 bp upstream from the transcription start site). These were found at similar positions as in human H2A.Z (Hatch and Bonner, 1990).

These features strongly indicated that the 3223 bp sequence contained the functional mouse H2A.Z gene.

As previously mentioned (3.2.2 (ii)), mouse histone H2A.Z pseudogene 2

encodes an identical protein to human H2A.Z. To determine whether mouse histone H2A.Z pseudogene 2 represented mouse histone H2A.Z cDNA, the coding sequence of the mouse histone H2A.Z gene was compared with the mouse histone H2A.Z pseudogene 2 nucleotide sequence. The result indicated that they were identical. A useful adjunct to this finding was that mouse histone H2A.Z pseudogene 2 could be used in an *Escherichia coli* expression system to obtain H2A.Z protein, as described in section 3.2.6.

3.2.4 Southern analysis of mouse genomic DNA

The aim of this work was to determine the copy number of the murine H2A.Z gene per haploid genome. Therefore, genomic DNA obtained from the C57xCBA F1 mice, which was available at that time, and was used for Southern analysis.

Mouse genomic DNA was isolated from liver of a C57xCBA F1 mouse as described (2.3.3 (i)). The DNA was digested with different restriction enzymes and Southern analysis was performed as described (2.3.10). The filter was initially probed with a 0.8 kb *Hin*c III/*Xba* I fragment isolated from pES4.8 (probe A in Fig. 3-17). This probe contained exon 3 and exon 4 and the intron sequences. After hybridisation, the filter was washed at high stringency. The result is presented in Fig. 3-18 A. This suggests that this probe not only hybridised to fragments containing H2A.Z gene but also hybridised to other fragments, probably representing the H2A.Z pseudogenes. Two pseudogenes, containing sequence homologous to the coding sequence of the H2A.Z gene, were isolated during the course of this work (3.2.1 (ii)) and 3.2.2 (ii)). Therefore the probe containing the coding sequence of H2A.Z also hybridises to the fragments containing pseudogene sequence. It was likely that the more intense bands represented the H2A.Z gene, whereas the less intense bands were the fragments containing the pseudogenes. To verify this interpretation, the same filter was

stripped and with a 0.5 kb *Hinc* II fragment (probe B in Fig. 3-17). This probe is located at the 3' flanking region of the mouse H2A.Z gene and does not contain any sequence homologous to histone H2A.Z pseudogene 1 or 2. The result is shown in Fig. 3-18 B. Only single hybridising fragments were observed for all restriction enzyme digestions except in the *Pvu* II digestion (lane 14 in Fig. 3-18 B), the hybridising fragment was undetectable; it may be smaller than 1.0 kb and therefore it is not present on the filter). This indicated that only a single copy of the H2A.Z gene exists in the mouse haploid genome. In ten enzyme digestions (lanes 2, 4-8, 10, 12, 13, 15 in Fig. 3-18 B), the band matched with the more intense band obtained with the initial probe, indicating that the sequences homologous to the two probes used in this analysis reside on the same fragment. However, in three enzyme digestions (lane 1, 9, 11 in Fig. 3-18 B) the band obtained from probe B in Fig. 3-18 did not match with the more intense band obtained from the initial probe, indicating that the sequences homologous to the first probe and the second probe resided on different fragments. This was due to the presence of the restriction enzyme sites in the region between probes A and B (Fig. 3-17). This results in the different hybridising fragment detected by these two probes.

3.2.5 Northern analysis of total RNA from ES cells

The aim of this work was to determine if the mouse H2A.Z gene is expressed in embryonic stem (ES) cells. Therefore, total RNA isolated from mouse E14 ES cells was used for Northern analysis.

RNA isolation and Northern analysis was performed as described (2.4). The filter was probed with a 397 bp *Nco* I/*Hin*d III fragment isolated from pET.H2AZ, containing the sequence corresponding to the mouse H2A.Z cDNA (described in section 3.2.6 (i)). After hybridisation, the filter was washed at high stringency. A single hybridising band was detected as shown in Fig. 3-19.

Figure 3-17

Probes for Southern analysis of mouse genomic DNA

The probes used in Southern analysis of genomic DNA are shown in striped boxes relative to the 4.8 kb *Eco* RI/*Sal* I fragment in pES4.8 (3.2.3 (iv)). The position of the mouse histone H2A.Z gene is shown in the map. The exons are boxed in black. Probe A is a 0.8 kb *Hin*c III/*Xba* I fragment containing the exon and the intron sequences of the H2A.Z gene. Probe B is a 0.5 kb *Hin*c II fragment located in the 3' flanking region of the H2A.Z gene.

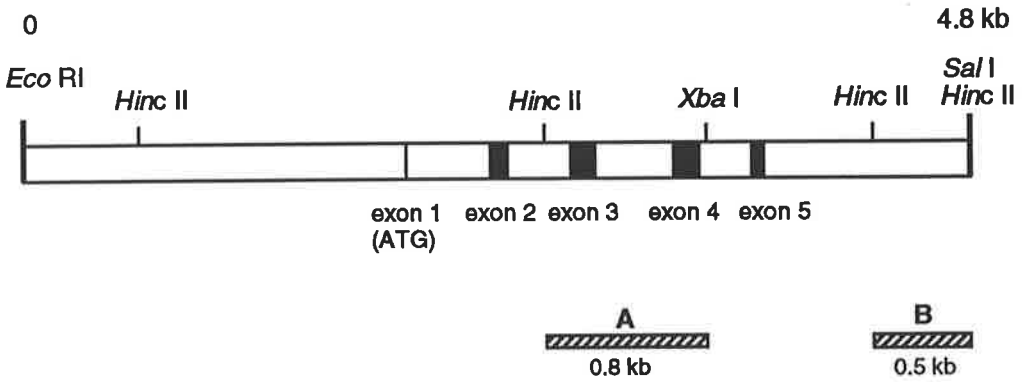

0
4.8 kb
Eco RI
Hinc II
Hinc II
Xba I
Hinc II
Sal I
Hinc II
exon 1
(ATG)
exon 2
exon 3
exon 4
exon 5
A
0.8 kb
B
0.5 kb

Figure 3-18

Southern analysis of genomic DNA isolated from C57xCBA mouse

A : Genomic DNA isolated from a C57xCBA mouse was digested with individual restriction enzymes and combinations. The digested DNA was electrophoresed on a 0.8% agarose gel along with DNA size markers. DNA was transferred from the gel to a Zeta-Probe membrane, probed with ^{32}P-labelled 0.8 kb *Hin*c III/*Xba* I fragment, containing the exon 3 and exon 4 and the intron sequences of the H2A.Z gene (probe A in Fig. 3-17). After hybridisation at 65°C overnight, the filter was washed at 65°C with 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 5% (w/v) SDS for 30 minutes, and then in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 1% (w/v) SDS, 65°C for 30 minutes and the hybridising fragments were detected by autoradiography. Lanes are as follows:

1) *Bgl* II-digested DNA
2) *Bam* HI-digested DNA
3) *Hin*d III-digested λ DNA marker
4) *Eco* RI-digested DNA
5) *Eco* RV-digested DNA
6) *Eco* RI/*Eco* RV-digested DNA
7) *Eco* RV/*Hin*d III-digested DNA
8) *Hin*d III-digested DNA
9) *Hin*d III/*Xba* I-digested DNA
10) *Kpn* I-digested DNA
11) *Xba* I-digested DNA
12) *Nco* I-digested DNA
13) *Pst* I-digested DNA
14) *Pvu* II-digested DNA
15) *Sca* I-digested DNA (partial digestion)
16) *Spe* I-digested DNA (partial digestion)

B : The same membrane as in **A** after stripping and rehybridising with ^{32}P-labelled 0.5 kb *Hin*c II fragment (probe B in Fig. 3-17). The fragments were detected by autoradiography after hybridisation and washing as described above. Lanes are the same as in **A**.

The result shows that the hybridising fragments in lanes 2, 4-8, 10, 12, 13, 15 correspond to the strongly hybridising fragments shown in **A**.

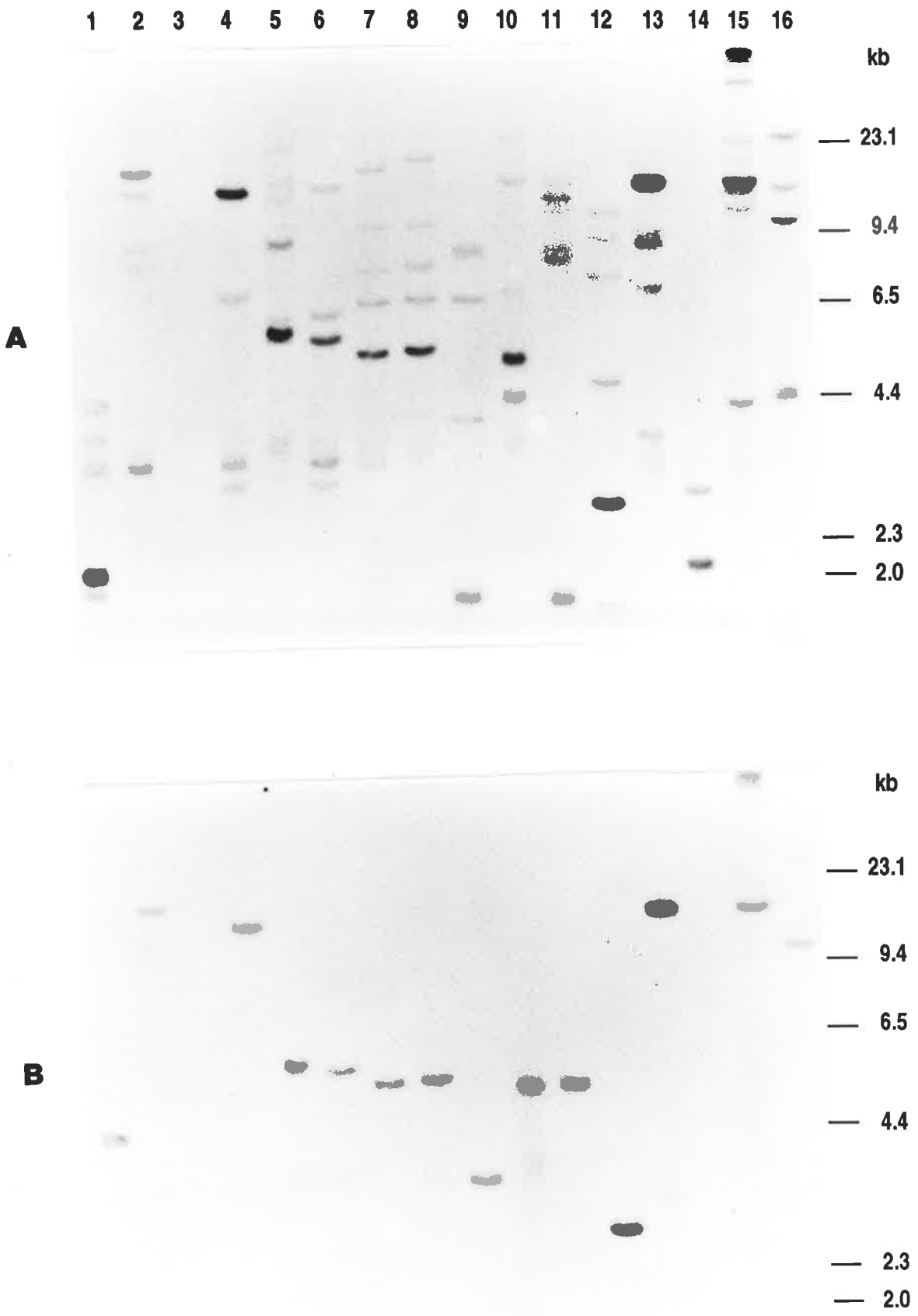
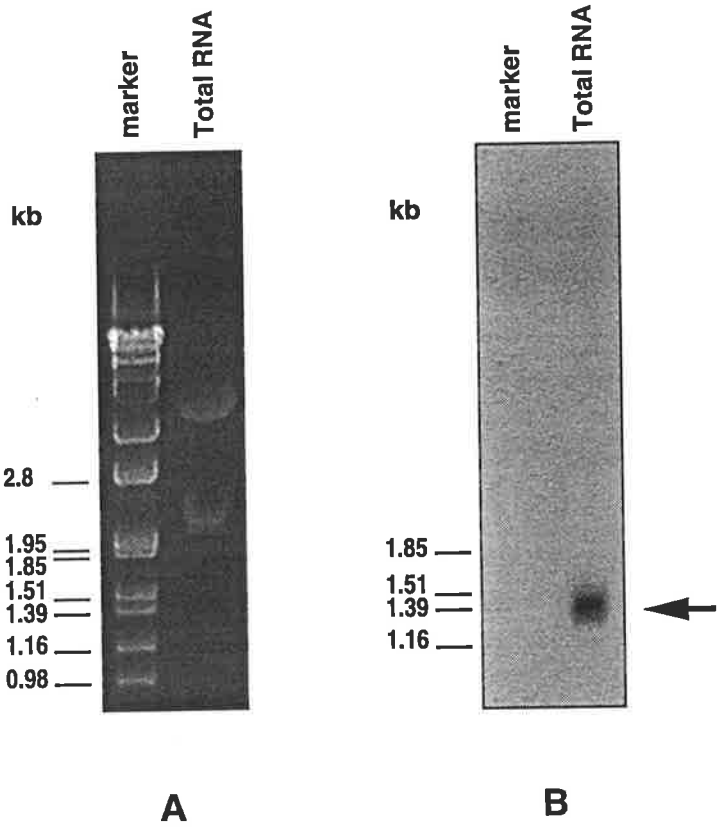

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
kb
23.1
9.4
6.5
4.4
2.3
2.0
A
kb
23.1
9.4
6.5
4.4
2.3
2.0
B

Figure 3-19

Northern analysis of total RNA isolated from mouse ES cells

10 μg of total RNA isolated from mouse E14 ES cells, along with DNA size markers, was run on a 1.2% formaldehyde-agarose gel. The RNA was transferred to Zeta-Probe membrane. The filter was probed with ^{32}P-labelled 397 bp *Nco* I/*Hin*d III fragment isolated from pET.H2AZ, containing the sequence corresponding to the mouse H2A.Z cDNA (described in section 3.2.6 (i)). After hybridisation at 65°C overnight, the filter was washed at 65°C with 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 5% (w/v) SDS for 30 minutes, and then in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 1% (w/v) SDS, 65°C for 15 minutes and hybridisation was detected by autoradiography. The ethidium bromide stained gel (**A**) and the autoradiograph (**B**) are shown, with the size of the DNA markers. The hybridising band is indicated by an arrow.


marker
Total RNA
kb
2.8
1.95
1.85
1.51
1.39
1.16
0.98
A
marker
Total RNA
kb
1.85
1.51
1.39
1.16
B

The band corresponding to H2A.Z mRNA is approximately 1.3 kb in size, as judged by the DNA markers (predicted size is 1-1.1 kb). This size may not be accurate since DNA and RNA migrate at different rates through agarose gels containing formaldehyde (Wicks, 1986). However, this experiment shows that the H2A.Z is expressed in mouse ES cells.

3.2.6 Expression of mouse histone H2A.Z in bacteria

The aim of this work was to produce a monoclonal antibody against mouse H2A.Z protein, to use as a tool to investigate the distribution of H2A.Z protein at different times during mouse development and in different tissues. It was important to have sufficient amounts of mouse H2A.Z protein, without contamination from other proteins, for use as an antigen in the immunisation process. The best way to achieve this was to produce this protein in bacteria, since an H2A.Z cDNA was available.

A T7 RNA polymerase-mediated expression system was chosen for the production of mouse H2A.Z protein in *Escherichia coli*. The RNA polymerase of bacteriophage T7 is very selective for specific promoters that are rarely encountered in DNA unrelated to T7 DNA (Chamberlin *et al.*, 1970; Dunn and Studier, 1983). Efficient termination signals of T7 RNA polymerase are also rare, so that the enzyme is able to make complete transcripts of almost any DNA that is placed under control of a T7 promoter (Studier *et al.*, 1990). T7 RNA polymerase elongates RNA transcript five times faster than *Escherichia coli* RNA polymerase (Chamberlin and Ring, 1973; Golomb and Chamberlin, 1974). These properties make T7 RNA polymerase attractive for directing high level expression of a cloned gene in *Escherichia coli* (Studier and Moffatt, 1986).

The bacterial host for expression was the *Escherichia coli* strain BL21(DE3). The features of this strain relevant to the expression system are as follows:

(1) As a B strain, it is naturally deficient in the *lon* protease

(Bukhari and Zipser, 1973; Philips *et al.*, 1984). It also lacks the *omp*T outer membrane protease which can degrade proteins during purification (Grodberg and Dunn, 1988).

(2) It contains bacteriophage lysogen DE 3. This phage carries a single copy of the T7 RNA polymerase in the chromosome under the control of inducible *lac* UV5 promoter. Addition of 0.4 mM IPTG to a culture of BL21(DE3) induces T7 RNA polymerase, which in turn initiates high level expression of a target gene in the plasmid (Studier *et al.*, 1990).

The vector used for expression of the H2A.Z gene was pET8.H3 (Fig. 3-20). This expression plasmid was derived from plasmid pBR322 as described by Studier *et al.*, (1990). This vector carries a T7 promoter, translation initiation signals and a transcription terminator. The *Nco* I site, containing an ATG initiation sequence, was created to facilitate the cloning of the DNA fragment into this vector. The DNA to be expressed can be subcloned into pET8.H3 by using the *Nco* I site at the 5' end and the *Hin*d III site at the 3' end.

The following section describes the expression of H2A.Z gene in *Escherichia coli* using a T7 expression system as described by Studier *et al.*, (1990) and the use of this protein as an antigen for antibody production and purification.

(i) Cloning of mouse histone H2A.Z cDNA sequences into the expression vector

Mouse histone H2A.Z pseudogene 2 carries sequences (from the ATG initiation codon to the TAA termination codon) corresponding to mouse histone H2A.Z cDNA sequence. To clone the H2A.Z cDNA sequence (ATG to TAA) into pET8.H3 plasmid (Fig. 3-20), it was necessary to create the *Nco* I site at the ATG initiation codon, and the *Hin*d III site at the TAA termination sequence.

The strategy used for constructing the expression vector is summarised in

Fig. 3-21. Two PCR primers, PHZ.Nco (2.2.15) and PHZ.Hd (2.2.15), were designed to introduce flanking *Nco* I and *Hind* III restriction sites to 5' and 3' ends of a PCR amplified cDNA, respectively, while maintaining the reading frame of the protein coding region. H2A.Z coding sequence was amplified from a clone containing the mouse histone H2A.Z pseudogene 2 (p17w-Pvu II, described in 3.2.2. (ii)) as a template. The amplified product of 397 bp was obtained after 40 cycles of PCR (2.3.16). The 397 bp fragment was digested with *Nco* I and *Hind* III and subcloned into *Nco* I and *Hind* III-digested pET8.H3 vector. After transformation into BL21(DE3) the recombinant plasmid was named pET.H2AZ. To ensure that no mutation occurred during the PCR step, pET.H2AZ was sequenced. No nucleotide substitution or deletion was observed in the H2A.Z cDNA sequence.

(ii) Production of mouse histone H2A.Z protein

A single colony of BL21(DE3) containing pET.H2AZ was inoculated in 25 ml of L-broth plus ampicillin and cultured at 37°C for 16 hours. One millilitre of the overnight culture was used to inoculate 100 ml of fresh L-broth containing ampicillin. When the A_{600} of this culture reached 0.6, one millilitre of culture was taken (0 hour) before the cells were induced with 0.4 mM IPTG. Another millilitre was taken 1, 2 and 3 hours after IPTG induction. Protein extracts were prepared from these aliquots as described (2.5.1) and electrophoresed on a 15% SDS/polyacrylamide gel (2.5.2).

The protein gel shown in Fig. 3-22 demonstrated that pET.H2AZ is highly expressed in BL21(DE3) after induction. The induced protein had an approximate molecular mass of 16 kDa, which corresponds to the deduced molecular mass of mouse histone H2A.Z (15,674 Da) (Fig. 3-16).

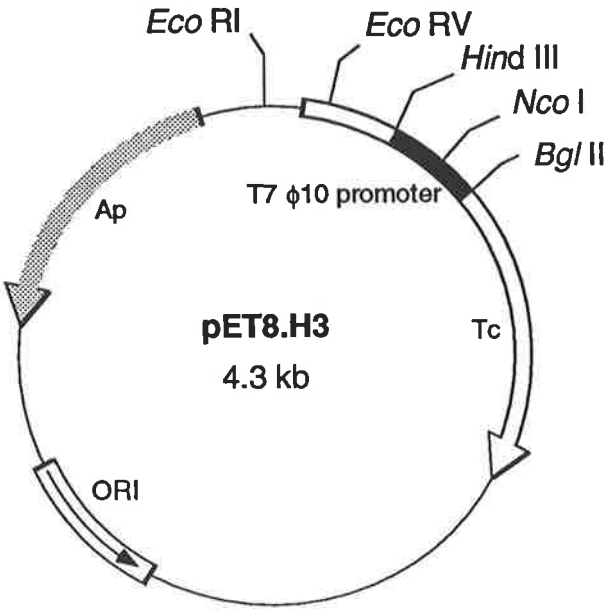
Figure 3-20

Relevant restriction map of expression vector pET8.H3

A The relevant restriction map of expression vector pET8.H3 is shown. This plasmid is derived from pBR322 and contains T7 ϕ10 promoter sequences inserted in the opposite orientation to the tetracycline resistance (Tc) gene.

B The T7 ϕ10 promoter and upstream sequences of pET8.H3 are shown, together with the relevant restriction sites.

A



B

T7 ϕ10 promoter

GGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGAGACCACA

Bgl II

Translation start

Shine Dalgarno

ACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGG

Nco I

Figure 3-21

Construction of pET.H2AZ

A flow chart illustrating the construction of pET.H2AZ is shown. The plasmid pET8.H3 was digested with *Nco* I and *Hin*d III. The sequences corresponding to H2A.Z cDNA were obtained from PCR reaction using the plasmid p17w-Pvu II as a template with primer PHZ.Nco and PHZ.Hd (2.2.15). The PCR product was then digested with *Nco* I and *Hin*d III and the fragment was isolated and ligated to pET8.H3 vector. As a result, the plasmid pET.H2AZ was obtained.

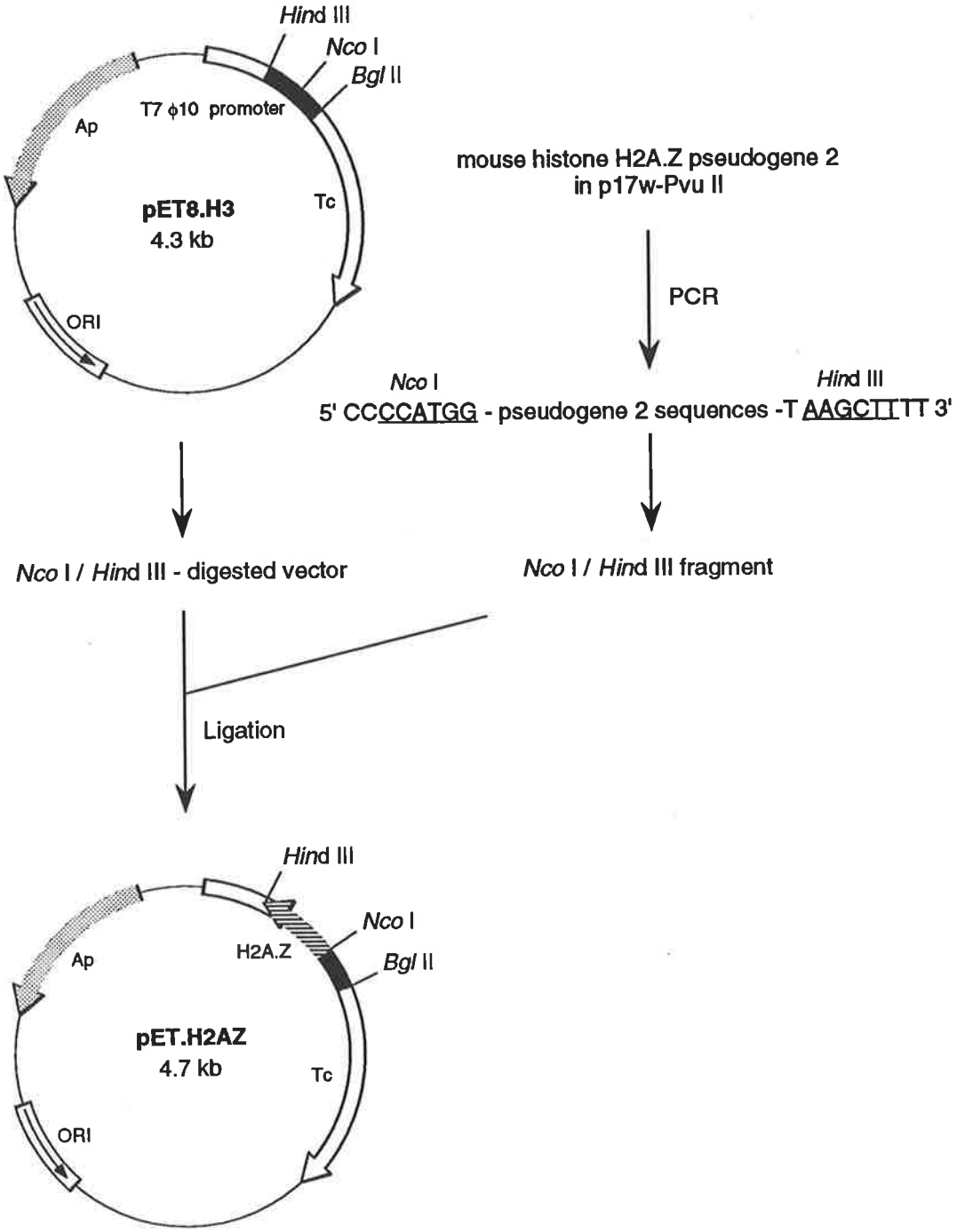
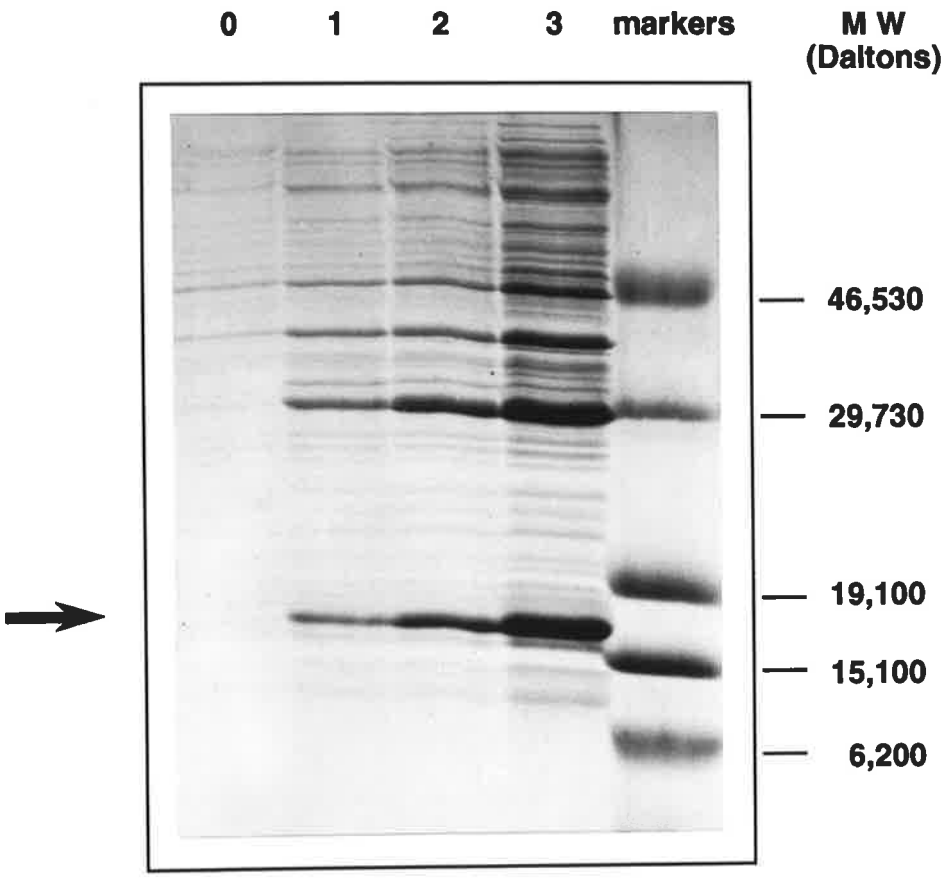

Hind III
Nco I
Bgl II
T7 ϕ10 promoter
Ap
pET8.H3
4.3 kb
Tc
ORI
mouse histone H2A.Z pseudogene 2
in p17w-Pvu II
PCR
Nco I
Hind III
5' CCCCATGG - pseudogene 2 sequences -T AAGCTTTT 3'
Nco I / Hind III - digested vector
Nco I / Hind III fragment
Ligation
Hind III
Nco I
H2A.Z
Bgl II
Ap
pET.H2AZ
4.7 kb
Tc
ORI

Figure 3-22

Expression of mouse histone H2A.Z from plasmid pET.H2AZ

BL21(DE3) cells containing pET.H2AZ were induced with 0.4 mM IPTG for 1, 2 and 3 hours. Protein extracts prepared from the cells before (0 hour) and after induction were electrophoresed on a 15% SDS/polyacrylamide gel, along with molecular weight markers, and stained with Coomassie blue. A band of protein corresponding to the expected molecular weight of mouse histone H2A.Z was observed after the BL21(DE3) cells were induced for at least 1 hour.


0
1
2
3
markers
M W
(Daltons)
46,530
29,730
19,100
15,100
6,200

(iii) Western analysis

To ensure that this expressed protein was mouse H2A.Z, Western blot analysis was conducted. Protein extracts prepared from uninduced BL21(DE3) containing pET.H2AZ (negative control), a 3 hour induced BL21(DE3) containing pET.H2AZ, along with protein markers were electrophoresed on a 15% SDS/polyacrylamide gel (2.5.2). The proteins were transferred onto the membrane. The filter was incubated with polyclonal antibodies against chicken H2A.F produced by J. Whiting (Ph.D. thesis, 1988) and subsequently the protein bound to these antibodies was detected as described (2.5.3). The result indicated that the antibodies bound to the band corresponding to the position at which expressed H2A.Z protein was located (data not shown). WESTERN BLOT SHOULD HAVE BEEN SHOWN AS INCREASE IN 16KD BAND INTENSITY, RELATIVE TO OTHER PROTEINS, IN FIG 3.22 NOT CONVINCING.

(iv) The use of histone H2A.Z expression product

(1) <u>As an antigen for monoclonal antibody production</u>

Having established a protocol to obtain high expression of mouse histone H2A.Z, a large scale preparation (5 litre culture) was performed. It was found that the highly expressed protein formed inclusion bodies within *Escherichia coli* cells. This facilitated purification of H2A.Z away from most of the other bacterial protein by extraction of inclusion bodies.

An attempt to produce a monoclonal antibody in mice against the recombinant H2A.Z protein was conducted by J. Wrin. The inclusion body material was dissolved in a solution containing 50 mM Sodium Acetate and 6 M urea pH 4.5 and applied to a Mono S cation exchange column. H2A.Z was eluted from the column at 400 mM NaCl in 50 mM Sodium Acetate and 6 M urea pH 4.5. A series of immunisations with the purified protein were performed. None of these mice developed a detectable antibody titre in ELISA assays. This may be due to this protein being an endogenous protein.

(2) Affinity purification of specific H2A.Z antibodies

Since the production of a monoclonal antibody against H2A.Z protein was unsuccessful, polyclonal antibodies against the synthetic peptides (designed to the C terminal end of H2A.Z protein by J. Wells) were consigned and produced by Chiron mimotopes. This sera contained high titre cross reactivity with H2A.Z in mouse liver total histone extracts (M. Clarkson, pers. comm.). The H2A.Z expression product was used for affinity purification of specific H2A.Z antibody from immunised sheep sera.

3.3 Summary and Discussion

3.3.1 Isolation and characterisation of the mouse histone H2A.Z gene

This chapter describes the isolation and characterisation of a mouse histone H2A.Z gene from the BALB/c strain. The mouse H2A.Z transcript was predicted on the basis of homology to sequence from the human H2A.Z gene (Hatch and Bonner, 1990).

The mouse H2A.Z gene encodes a protein of 127 residues. The gene consists of four introns and five exons and has an exon-intron organisation similar to that of chicken H2A.F (Dalton *et al.*, 1989) and human H2A.Z (Hatch and Bonner, 1990). The exon-intron boundaries of mouse H2A.Z occur in the same positions in the coding sequence and the intron lengths are similar to human H2A.Z but are quite different from chicken. The four mouse H2A.Z introns are 435, 309, 390, and 246 bp long, compared to 427, 276, 332, and 287 bp in the human gene and 1536, 246, 1370, and 2681 bp in the chicken gene. The polyadenylation signal (AATAAA) of the mouse H2A.Z is identical to the signal in H2A.Z mRNA from cow, rat and human (Hatch and Bonner, 1988). However, it is slightly different in chicken which is GATAAA (Dalton *et al.*, 1989).

Two mouse histone H2A.Z pseudogenes were isolated and characterised during this work. One of these pseudogenes, designated mouse histone H2A.Z pseudogene 2, contains sequence identical to mouse histone H2A.Z cDNA. The presence of H2A.Z pseudogenes in the genome was also found in human (Hatch and Bonner, 1990).

3.3.2 Southern analysis of mouse genomic DNA

Southern analysis indicated that there is only a single copy of the mouse H2A.Z gene per mouse haploid genome. There are also several pseudogenes in the mouse genome, as demonstrated by Southern analysis using a probe containing the coding sequence for the H2A.Z gene.

3.3.3 Northern analysis of total RNA from ES cells

Northern analysis of total RNA isolated from ES cells demonstrated that the H2A.Z gene is expressed in these cells. This result was expected since it is known that H2A.Z is expressed at a low level throughout the cell cycle, in all cell types.

3.3.4 Expression of mouse histone H2A.Z in bacteria

High level expression of mouse histone H2A.Z protein in *Escherichia coli* cells was obtained by using a T7 RNA polymerase-mediated expression system. The coding sequence used was obtained from mouse histone H2A.Z pseudogene 2. The expression vector was constructed by cloning a PCR fragment, amplified from mouse histone H2A.Z pseudogene 2, into pET8.H3. The vector was then transformed into *Escherichia coli* BL21(DE3) and the protein was expressed, and formed inclusion bodies, in the cells after induction with IPTG for at least 1 hour.

Several results indicated that the expressed protein was mouse histone H2A.Z. Firstly, the molecular mass was the same as histone H2A.Z. Secondly,

Western blot analysis using polyclonal antibodies against chicken H2A.F cross reacted with recombinant H2A.Z. Finally, this protein was successfully used for purification of H2A.Z specific antibodies.

In collaboration with D. Tremethick at the John Curtin School of Medical Research, Australian National University, the expressed H2A.Z protein is being used for chromatin assembly studies.

Chapter 4

Histone H2A.Z Gene "Knock-Out" by Homologous Recombination in Mouse Embryonic Stem Cells

4.1 Introduction

The biological function of histone H2A.Z is unknown but recently it has been shown to be essential in *Drosophila* development (van Daal and Elgin, 1992). Null mutant flies only survive until the third larval stage, presumably utilising maternal supplies of the histone up to this stage. The function of this gene has not been demonstrated in mammals.

To investigate the role of the mouse H2A.Z gene, homologous recombination of this gene in embryonic stem (ES) cells was performed, to create ES cells and mice lacking a functional H2A.Z gene. This chapter describes targeted disruption of the histone H2A.Z gene in ES cells. The aim was to create heterozygous H2A.Z gene "knock-out" ES cell clones. Subsequently, these cells could be used to generate homozygous H2A.Z gene "knock-out" ES cell lines and mice; these experiments will be described in Chapter 5.

Three different targeting vectors (pHZHR1, pHZHR2 and pHZHR3) were constructed, as described in this chapter. The first vector, pHZHR1, was prepared from the clones containing the H2A.Z gene derived from the BALB/c mouse genomic DNA library, which were isolated and characterised as described in Chapter 3. During the course of this work, emerging evidence in the literature showed that vectors prepared from isogenic DNA are more effective in homologous recombination experiments than vectors prepared from non-isogenic DNA (Deng and Capecchi, 1992; te Riele *et al.*, 1992 and van Deursen and Wieringa, 1992). The E14 ES cells, to be used for gene targeting, are derived from the 129/Ola mouse strain. Therefore, mouse genomic library III (2.2.7), derived from mouse strain 129/Sv, was screened for the isogenic H2A.Z gene. Clones containing the isogenic H2A.Z gene were isolated and used for constructing the targeting vectors pHZHR2 and pHZHR3. The three vectors were classified as replacement type (Thomas and Capecchi, 1987). They contained the neomycin phosphotransferase gene (*neo*^{r}) flanked by nucleotide sequences

homologous to the H2A.Z gene locus. The purposes of using the *neo*^{r} gene in these vectors were: (1) to disrupt the H2A.Z gene; (2) to act as a selectable marker (conferring resistance to the antibiotic drug G418) for cells containing an integrated copy of the targeting vector. DNA was introduced into E14 ES cells by electroporation; cells which had integrated the targeting construct were selected by G418 treatment. Homologous recombination events were identified by Southern analysis.

4.2 Results

4.2.1 Electroporation of E14 ES cells with targeting vector pHZHR1

(i) Construction of pHZHR1

The targeting vector pHZHR1 was constructed using the H2A.Z gene isolated from a BALB/c mouse genomic DNA library. pHZHR1 was derived from a 4.0 kb *Eco* RI/*Hinc* II fragment of pES4.8 (Fig. 3-14). It contains 3.7 kb of DNA homologous to the H2A.Z locus. To inactivate the H2A.Z gene, 257 bp (from nucleotide 1302 to 1558, in Fig. 3-16) were deleted, removing exon 2 and part of intron 2, and were replaced with the *neo*^{r} gene (isolated from pgkNEOpA, 2.2.13). The *neo*^{r} gene was inserted in the same transcriptional orientation as H2A.Z. For further enrichment of the clones in which homologous recombination events occurred, a strategy using the *neo*^{r} gene without polyadenylation signals was used. Therefore, pHZHR1 was designed without the polyadenylation signals from either the *neo*^{r} gene or the H2A.Z gene. The protein product which has a functional *neo*^{r} resistant activity can be produced only when the vector is integrated in the genome in such a way that stable transcripts of the *neo*^{r} gene are produced using polyadenylation signals from the

genome. Therefore G418-resistant ES cell clones were expected to be enriched for homologous recombination events. The map of pHZHR1 is presented in Fig. 4-1.

(ii) Electroporation and Southern analysis

Prior to electroporation pHZHR1 was digested with *Eco* RI and *Sal* I in order to release the insert and processed as described (2.3.17). E14 ES cells were maintained as described (2.6.2). For electroporation, $5x10^7$ E14 ES cells were mixed with 25 μg of *Eco* RI/*Sal* I-digested pHZHR1 in 0.5 ml HBS and electroporated at 220 volts and 960 μF capacitance with a Bio-Rad Gene Pulser as described (2.6.3). Transformed cells were selected with G418 (200 μg/ml) for 10 days. G418-resistant colonies were expanded as described (2.6.4).

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Bam* HI and *Kpn* I and filters were probed with a 1.5 kb *Hinc* II fragment which was isolated from p129ES (described in section 4.2.2 (ii)) (see Fig. 4-2). This probe did not contain sequences present in the targeting vector and hence will only hybridise to endogenous H2A.Z locus fragments, not random integration fragments derived from the targeting vector. The endogenous H2A.Z allele was detected as a 14 kb hybridising fragment with this probe. If a homologous recombination event occurred, the *Bam* HI restriction site in the vector would be integrated in the genome. Thus the H2A.Z disrupted allele was expected to generate a 3.6 kb hybridising fragment. No homologous recombination was observed in any of the 120 G418-resistant ES clones.

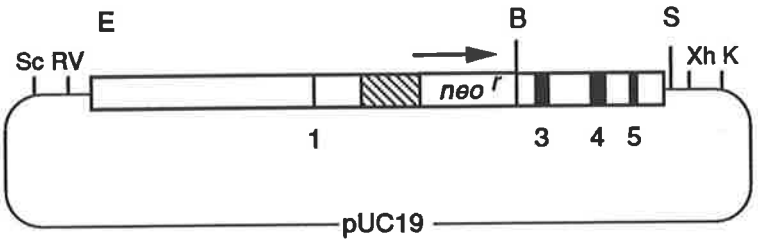
4.2.2 Isolation of a mouse histone H2A.Z genomic clone derived from mouse strain 129

The first targeting construct, pHZHR1, was constructed before the

Figure 4-1

Map of targeting vector pHZHR1

The restriction map of pHZHR1 containing a 5.1 kb *Eco* RI/*Sal* I fragment in pUC19. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The *neo*^{r} cassette disrupted H2A.Z gene is indicated with the arrow above indicating the direction of the transcription. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated.



pHZHR1

1 kb

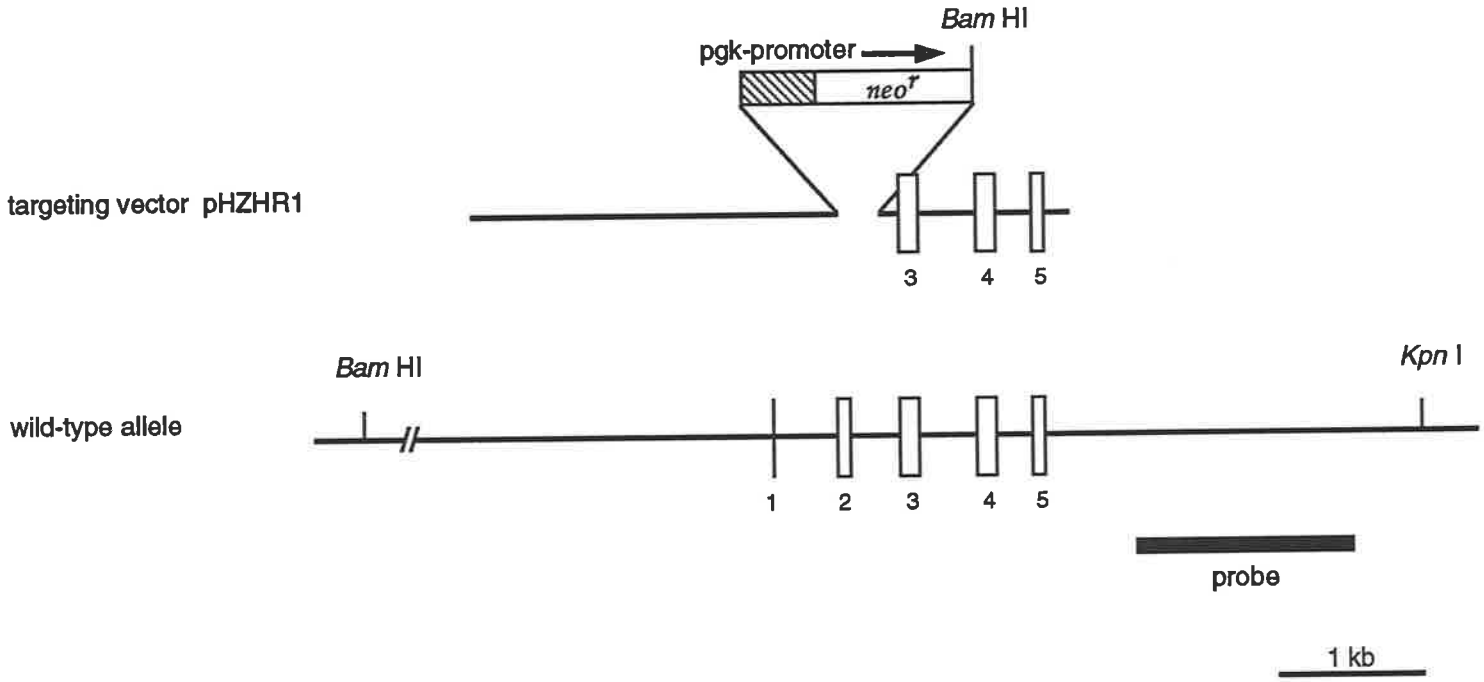
B	*Bam* HI	E	*Eco* RI
K	*Kpn* I	RV	*Eco* RV
S	*Sal* I	Sc	*Sac* I
Xh	*Xho* I		
▨	pgk-promoter		

Figure 4-2

Southern analysis of ES cells transfected with pHZHR1

The targeting vector pHZHR1 is shown at the top in alignment with the restriction map of the H2A.Z locus. Exons of the H2A.Z gene are represented by open boxes and are numbered. The arrow above the *neo*^{r} cassette indicates the direction of the transcription of *neo*^{r} gene.

ES cell DNA was digested with *Bam* HI and *Kpn* I restriction enzymes and probed with a 1.5 kb *Hinc* II fragment represented by the black box. This probe is located 3' and external to pHZHR1. The endogenous and the disrupted H2A.Z alleles generated from an homologous recombination event were expected to give a 14 kb and a 3.6 kb hybridising fragment respectively.


Bam HI
pgk-promoter
neo^r
targeting vector pHZHR1
3
4
5
Bam HI
Kpn I
wild-type allele
1
2
3
4
5
probe
1 kb

importance of utilising isogenic DNA was appreciated. In order to improve the targeting efficiency, vectors were prepared from isogenic DNA. To achieve this, it was necessary to obtain H2A.Z genomic clones from isogenic DNA. E14 ES cells, to be used for the gene targeting experiments, are derived from 129/Ola. Therefore, mouse genomic library III (derived from 129/Sv mouse strain; 2.2.7) was screened for isogenic H2A.Z clones.

(i) Screening mouse genomic library III

The results from previous library screenings (described in Chapter 3) indicated that there are at least two histone H2A.Z pseudogenes in the mouse genome. In order to isolate the mouse histone H2A.Z gene and avoid isolation of pseudogenes, mouse genomic library III (2.2.7) was screened using probes flanking the transcribed region. Therefore, a 0.5 kb *Hinc* II/*Hind* II, a 0.6 kb *Hind* III/*Xho* I and a 0. 5 kb *Hinc* II fragment (probes A, B and C in Fig. 4-3) derived from the 4.8 kb *Eco* RI/*Sal* I fragment in pES4.8 were used in this screening. The mouse genomic library III was plated as described (2.3.11). The probe was labelled with a α-^{32}P dATP (2.3.8 (ii)) and then hybridised to the filters. After hybridisation the filters were washed at a high stringency of 0.1x SSC, 0.1% SDS at 65°C and set up for autoradiography.

A total of six positive plaques were picked and these underwent two more rounds of screening in order to obtain clonal positive plaques. Three plaques were obtained after the three rounds of screening and all of these positives (λ1, λ2, λ10) were analysed further.

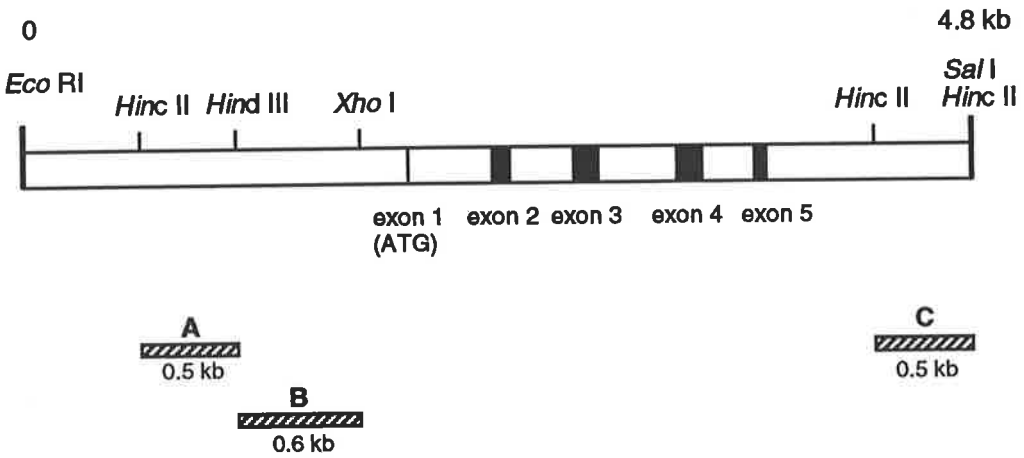
(ii) Characterisation of the three positive lambda clones (λ1, λ2, λ10)

DNA was prepared from λ1, λ2, λ10 (2.3.2) and Southern analysis was performed, using DNA digested with *Sal* I and *Eco* RI, and probed with the 0.6 kb

Figure 4-3

Probes for screening of mouse genomic DNA library III

The probes used in library screening are shown in striped boxes relative to the 4.8 kb *Eco* RI/*Sal* I fragment in pES4.8. Fragments A and B are the 5' flanking region probes, fragment C is the 3' flanking region probe. The restriction enzymes used to prepare these probes from pES4.8 are indicated on the map.


0
4.8 kb
Eco RI
Hinc II
Hind III
Xho I
Hinc II
Sal I
Hinc II
exon 1
(ATG)
exon 2
exon 3
exon 4
exon 5
A
0.5 kb
B
0.6 kb
C
0.5 kb

Hind III/*Xho* I fragment (probe B in Fig. 4-3), which is located 5' of the mouse H2A.Z gene. The result is shown in Fig. 4-4 and the alignment of the clones is shown in Fig. 4-5. It was found that: (a) all of the clones contained the H2A.Z gene and are overlapping, and (b) the DNA obtained from these overlapping clones spanned 20.1 kb with approximately 10.2 kb flanking the 5' end of the H2A.Z gene region and 7 kb flanking the 3' end of the H2A.Z gene region.

The insert fragments in the overlapping lambda clones were subcloned into plasmid vectors for further analysis. The map of the subclones, p129S3.6, p129SK and p129ES, relative to the lambda clones, is shown in Fig. 4-5. p129S3.6 contained a 3.6 kb *Eco* RI/*Sal* I fragment, isolated from λ10, in pBluescript KS+ vector. p129SK contained a 14.1 kb *Sal* I/*Kpn* I fragment, isolated from λ2, in pBluescript KS+ vector. p129ES contained a 12 kb *Eco* RI/*Sal* I fragment, isolated from λ1, in pBluescript KS+ vector. Partial sequencing of p129ES using a primer, pHZseq3 (2.2.15), which annealed to an internal sequence of H2A.Z, confirmed that it contained the H2A.Z gene as expected. The restriction map and the generation of the targeting constructs are described in the following section.

4.2.3 Electroporation of E14 ES cells with targeting vector pHZHR2

(i) Construction of pHZHR2

The second targeting vector pHZHR2 was constructed using a region of the H2A.Z locus cloned from the 129/Sv mouse DNA described above. This vector contained 5.8 kb of DNA homologous to the H2A.Z locus and a *neo*^{r} gene driven by the mouse phosphoglycerate kinase promoter (pgk-promoter) and a pgk-polyadenylation signal (pgk-polyA). The *neo*^{r} was chosen as a selectable marker since the enrichment strategy described above for homologous recombination using the *neo*^{r} gene without a polyadenylation signal was unsuccessful.

The strategy used to create the targeting vector pHZHR2 is presented in

Fig. 4-6. To construct pHZHR2, p129Xb and p129E/RV were generated by subcloning the fragment isolated from p129ES (Fig. 4-5). A fragment containing a *neo*^{r} cassette was introduced into *Mlu* I/*Bst* EII-digested p129Xb. The resulting plasmid, designated p129Xb.neo, contains the pgk *neo*^{r} cassette transcribed in the opposite orientation to H2A.Z. H2A.Z is therefore rendered nonfunctional by the loss of part of the H2A.Z promoter region, exon 1, intron 1, exon 2 and part of intron 2.

(ii) Electroporation of pHZHR2 into ES cells and Southern analysis

Prior to electroporation pHZHR2 was digested with *Eco* RI and *Kpn* I in order to release the insert and processed as described (2.3.17). E14 ES cells ($5x10^7$) were mixed with 25 µg of *Eco* RI/*Kpn* I-digested pHZHR2 in 0.5 ml HBS and electroporated at 220 volts and 960 µF capacitance as described (2.6.3). Transformed cells were selected with G418 (200 µg/ml) for 10 days. G418-resistant colonies were expanded as described (2.6.4).

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Bam* HI and probed with a 0.6 kb *Bam* HI/*Sma* I fragment (isolated from p129S3.6; probe B in Fig. 4-10) which is an external probe not present in the vector. The endogenous H2A.Z allele was detected as a 27 kb hybridising fragment with this probe. If a homologous recombination event occurred, the *Bam* HI restriction site in the vector would be introduced into the H2A.Z locus. Thus a disrupted H2A.Z allele was expected to generate a 10 kb hybridising fragment. Homologous recombination was not observed in any of the 273 G418-resistant ES clones.

Figure 4-4

Southern analysis of λ1, λ2 and λ10 derived from mouse genomic library III

A : DNA of λ1, λ2 and λ10, derived from mouse genomic library III, was digested with *Sal* I and *Eco* RI. The digested DNA was electrophoresed on a 1% agarose gel along with DNA size markers. DNA was visualised by staining with ethidium bromide. Lanes are as follows:

1) Hi-Lo mixed DNA size markers (Bresatec)
2) *Sal* I and *Eco* RI-digested λ1 DNA
3) *Sal* I and *Eco* RI-digested (*Eco* RI with star activity) λ2 DNA
4) *Sal* I and *Eco* RI-digested λ10 DNA

B : DNA was transferred from the above gel to a Zeta-Probe membrane, probed with ^{32}P-labelled 0.6 kb *Hin*d III/*Xho* I fragment located at the 5' end of H2A.Z (probe B in Fig. 4-3), and the hybridising bands were detected by autoradiography. Sizes of the markers are shown. Lanes are the same as in **A**

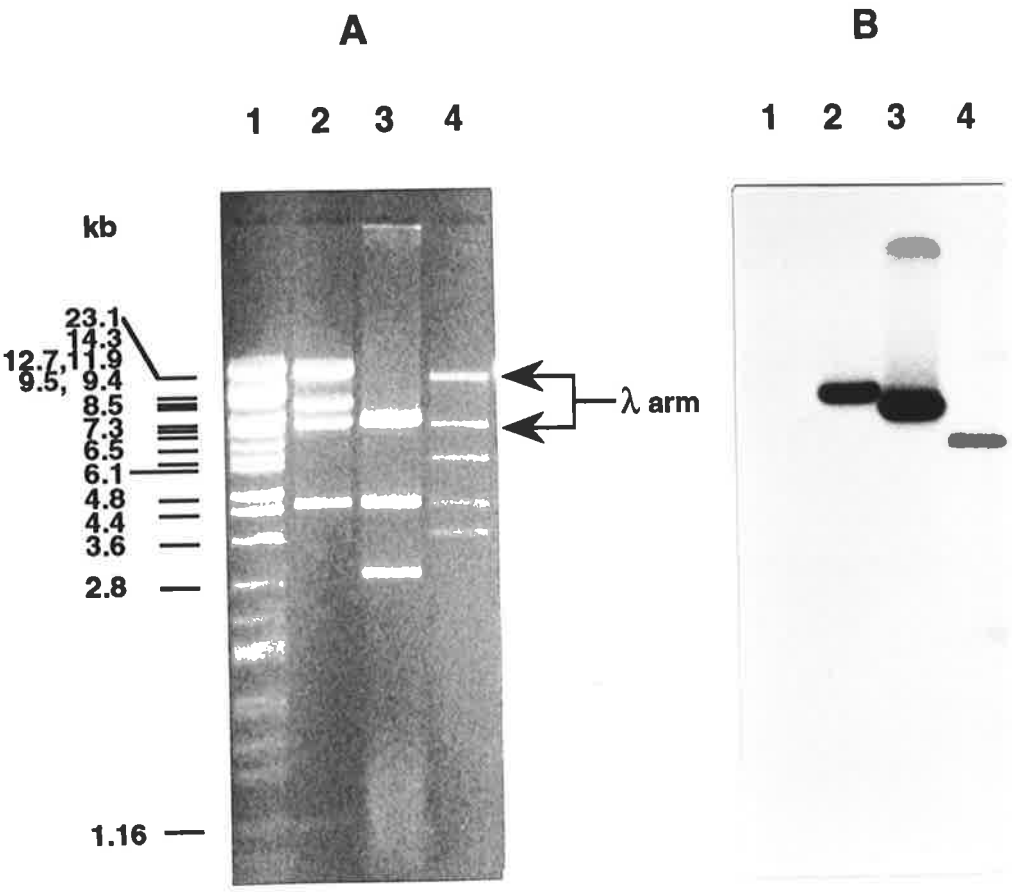

A
B
1 2 3 4
1 2 3 4
kb
23.1
14.3
12.7,11.9
9.5, 9.4
8.5
7.3
6.5
6.1
4.8
4.4
3.6
2.8
1.16
λ arm

Figure 4-5

Alignment of λ1, λ2 and λ10 and subclones p129S3.6, p129SK and p129ES

The overall restriction map of the region of genomic DNA covered by the inserts of λ1, λ2 and λ10 is shown below the maps of individual clones. The numbers above the map indicate the distances in kilobases from the 5' end of the genomic region. The location of the H2A.Z gene is shown, and the arrow indicates the orientation of the gene. The thick lines represent the insert fragments subcloned into plasmids and the names of the recombinant plasmid are indicated.

p129S3.6 contains a 3.6 kb *Eco* RI/*Sal* I fragment (isolated from λ10) subcloned into a *Eco* RI/*Sal* I-digested pBluescript KS+ vector. p129SK contains a 14.1 kb *Sal* I/*Kpn* I fragment (isolated from λ2) subcloned into a *Sal* I/*Kpn* I-digested pBluescript KS+ vector. p129ES contains a 12 kb *Eco* RI/*Sal* I fragment (isolated from λ1) subcloned into a *Eco* RI/*Sal* I-digested pBluescript KS+ vector.

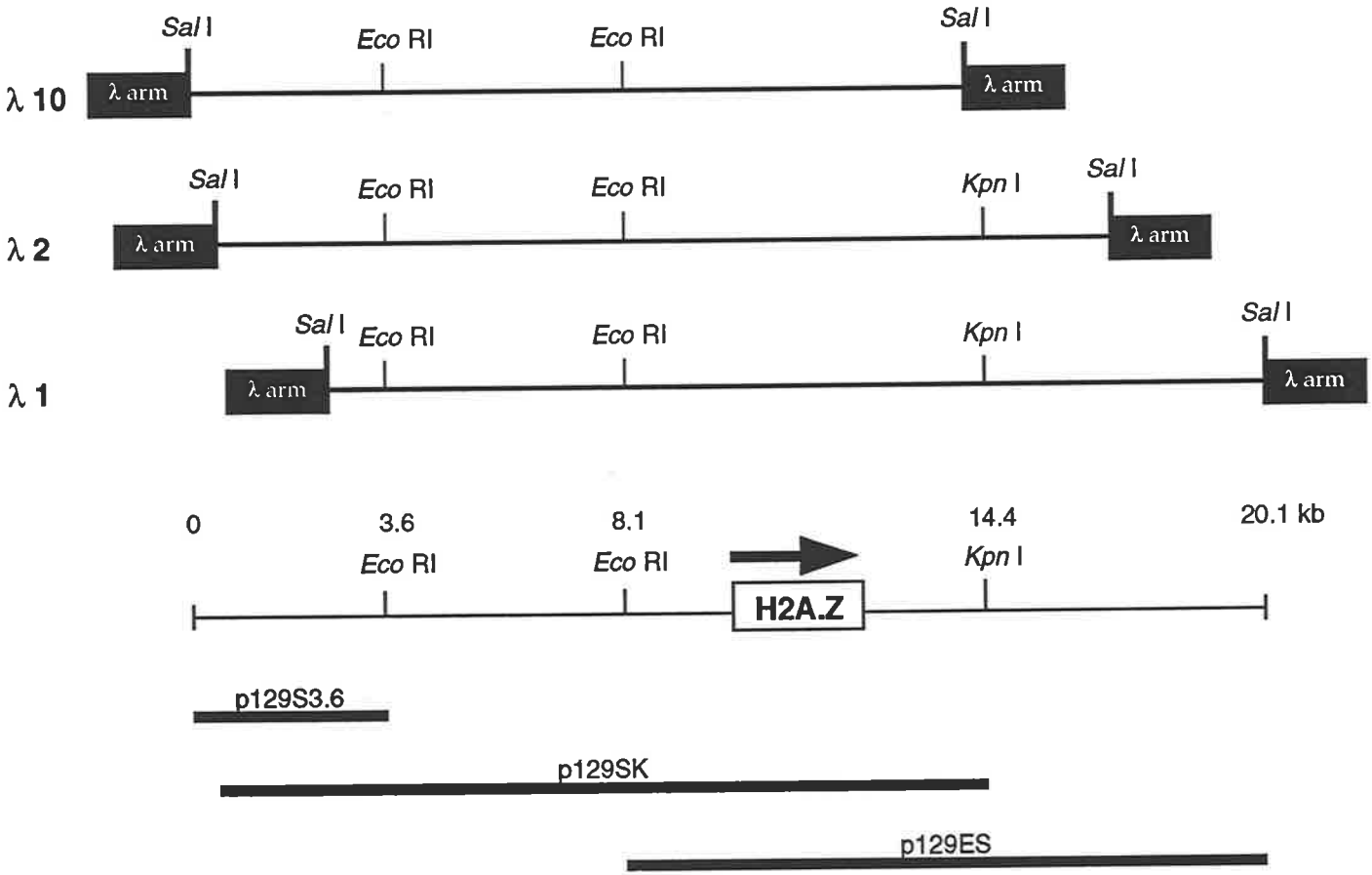

λ 10
Sal I
Eco RI
Eco RI
Sal I
λ arm
λ arm
λ 2
Sal I
Eco RI
Eco RI
Kpn I
Sal I
λ arm
λ arm
λ 1
Sal I
Eco RI
Eco RI
Kpn I
Sal I
λ arm
λ arm
0
3.6
8.1
14.4
20.1 kb
Eco RI
Eco RI
Kpn I
H2A.Z
p129S3.6
p129SK
p129ES

Figure 4-6

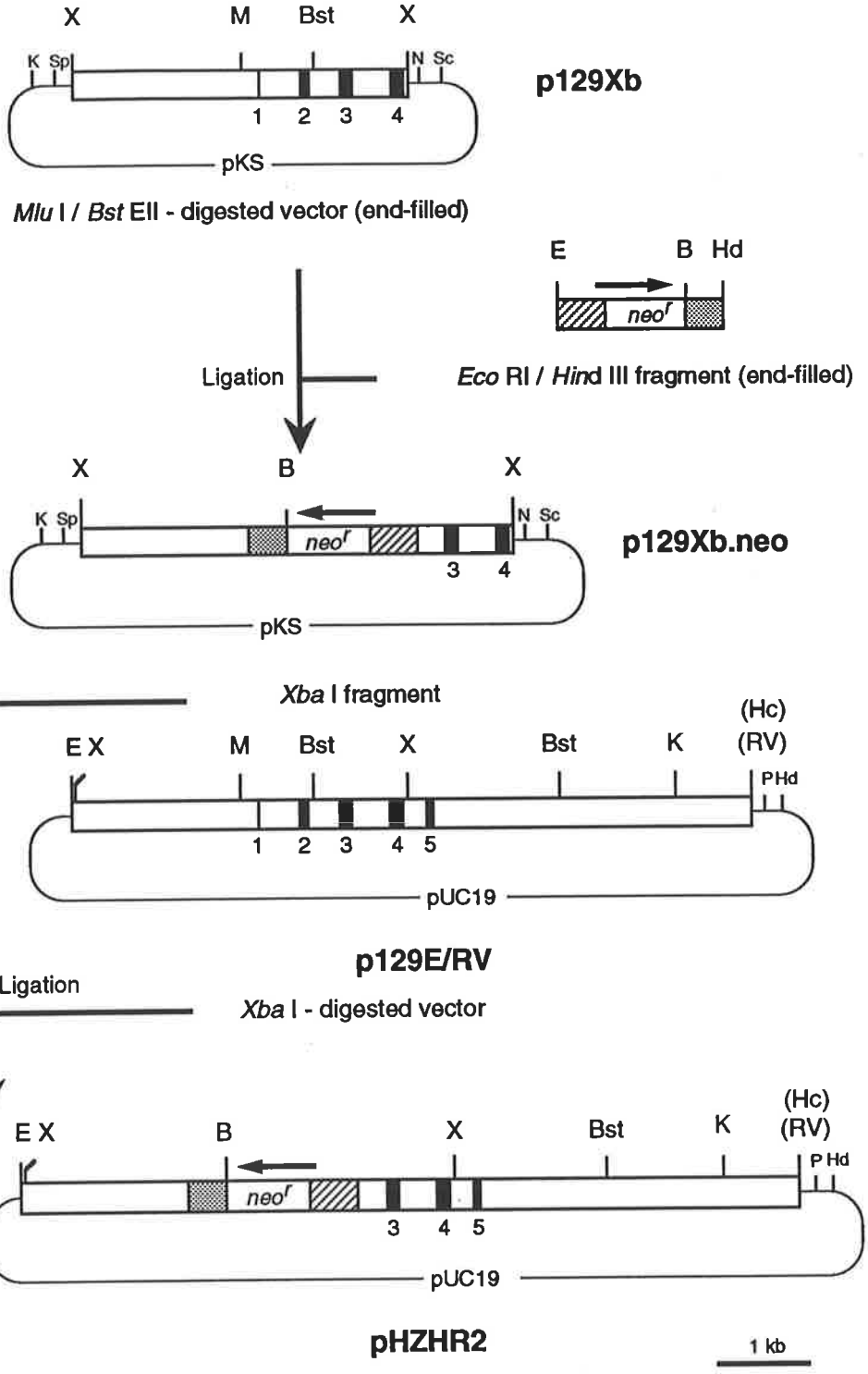
Construction of targeting vector pHZHR2

The strategy used for constructing the targeting vector pHZHR2 is shown. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The pgk *neo*^{r} cassette is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated. The restriction sites which were lost as a result of the ligation reaction are shown in brackets.

p129Xb contained a 3.6 kb *Xba* I fragment subcloned into the *Xba* I-digested pBluescript KS+ vector. A 0.7 kb *Mlu* I/*Bst* EII fragment containing part of the H2A.Z promoter region, exon 1, intron 1, exon 2 and part of intron 2 was replaced with a 1.8 kb *Eco* RI/*Hin*d III pgkNEOpA (2.2.13) fragment. The resulting plasmid containing the *neo*^{r} gene inserted in the opposite transcriptional direction to H2A.Z gene was named p129Xb.neo.

p129E/RV contained a 7.3 kb *Eco* RI/*Eco* RV fragment, carrying the entire H2A.Z gene, subcloned into a *Eco* RI/*Hin*c II-digested pUC19 vector. To create pHZHR2, a 4.7 kb *Xba* I fragment isolated from p129Xb.neo was replaced by a 3.6 kb *Xba* I fragment in p129E/RV.

For electroporation, pHZHR2 was digested with *Eco* RI and *Kpn* I restriction enzymes to release the insert.


X
M
Bst
X
K
Sp
N
Sc
1
2
3
4
pKS
p129Xb
Mlu I / Bst EII - digested vector (end-filled)
E
B
Hd
neor
Eco RI / Hind III fragment (end-filled)
Ligation
X
B
X
K
Sp
N
Sc
neor
3
4
pKS
p129Xb.neo
Xba I fragment
E X
M
Bst
X
Bst
K
(Hc)
(RV)
P
Hd
1
2
3
4
5
pUC19
p129E/RV
Ligation
Xba I - digested vector
E X
B
X
Bst
K
(Hc)
(RV)
P
Hd
neor
3
4
5
pUC19
pHZHR2
1 kb
B
Bam HI
Bst
Bst EII
E
Eco RI
Hc
Hinc II
Hd
Hind III
K
Kpn I
M
Mlu I
N
Not I
P
Pst I
RV
Eco RV
Sc
Sac I
Sp
Spe I
X
Xba I
pgk-promoter
pgk-polyA

4.2.4 Homologous recombination in ES cells using pHZHR3

(i) Construction of pHZHR3

pHZHR3 was constructed from pHZHR2 to extend the homology sequence flanking the *neo*^{r} cassette from 5.8 kb to a total of 18.6 kb. The efficiency of homologous recombination is proportional to the length of the homology up to 14 kb (Hasty *et al.*, 1991), ALTHOUGH DENG & CAPECCHI (1992) FIND NONLINEAR,

The features of pHZHR3 were essentially the same as pHZHR2. By using a fragment isolated from pHZHR2, p129ES.neo was generated as a precursor for constructing pHZHR3 (Fig. 4-7). The strategy used to construct pHZHR3 is shown in Fig. 4-8.

(ii) Isolation of an overlapping clone containing sequences downstream from clone λ1

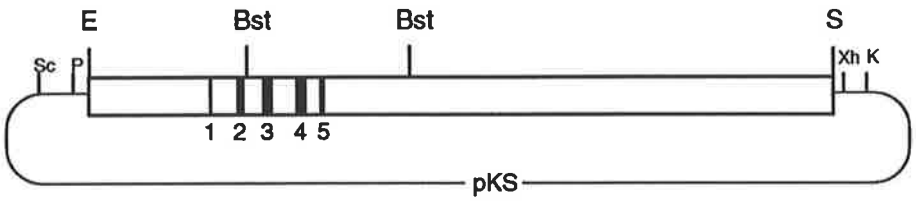
The aim of this work was to obtain a clone containing sequences downstream from pHZHR3 for use as 3'-external probe for detection of the homologous recombination event. Therefore, mouse genomic library III (2.2.7) was screened using the 0.3 kb *Xba* I/ *Sal* I fragment, located at the 3' end of λ1 from which pHZHR3 was derived (Fig. 4-9), as a probe. One positive phage (λ6.1) was obtained from the first round. This phage was purified by three subsequent rounds of screening at low density. DNA was prepared from the phage and digested with *Eco* RI, *Sal* I, and *Eco* RI plus *Sal* I. Southern analysis using the 0.3 kb Xba I/ Sal I fragment allowed the map to be drawn in alignment with clone λ1 and λ10 obtained from the previous screening (Fig. 4-9). In total, 31.9 kb surrounding the H2A.Z gene was obtained from the mouse genomic library III (2.2.7). The insert fragment of λ6.1 was isolated and subcloned into a plasmid vector for further analysis in order to search for a specific probe. It was found that a 0.6 kb *Eco* RI/*Hin*d III fragment, located 4.2 kb downstream from λ1 (Fig. 4-9) was a suitable probe for Southern analysis to identify the ES clones in which homologous recombination had occurred.

Figure 4-7

Construction of p129ES.neo

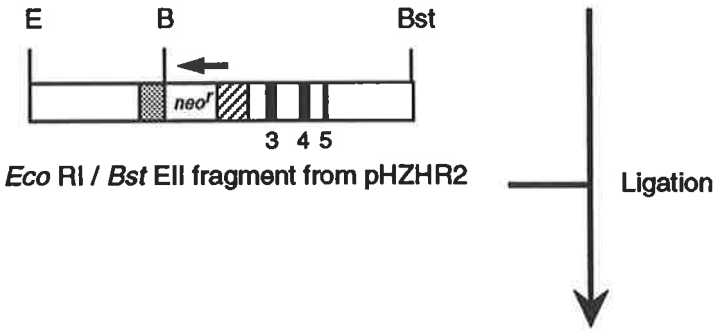
The strategy used for constructing p129ES.neo plasmid, as a precursor for pHZHR3, is shown. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The position of the *neo*^{r} cassette within the H2A.Z fragment is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated.

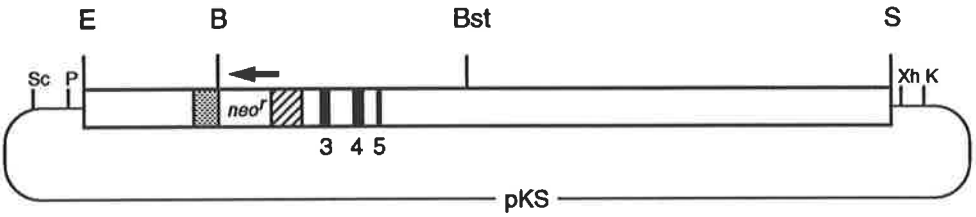
p129ES contained a 12 kb *Eco* RI/*Sal* I fragment subcloned into the *Eco* RI/*Sal* I-digested pBluescript KS+ vector. *Eco* RI/*Bst* EII-digested p129ES vector was ligated to the *Eco* RI/*Bst* EII fragment isolated from pHZHR2. The resulting plasmid was designated p129ES.neo.



p129ES

Eco RI / *Bst* EII - digested vector





p129ES.neo

1 kb

B	*Bam* HI	Bst	*Bst* EII
E	*Eco* RI	K	*Kpn* I
P	*Pst* I	S	*Sal* I
Sc	*Sac* I	Xh	*Xho* I

pgk-promoter

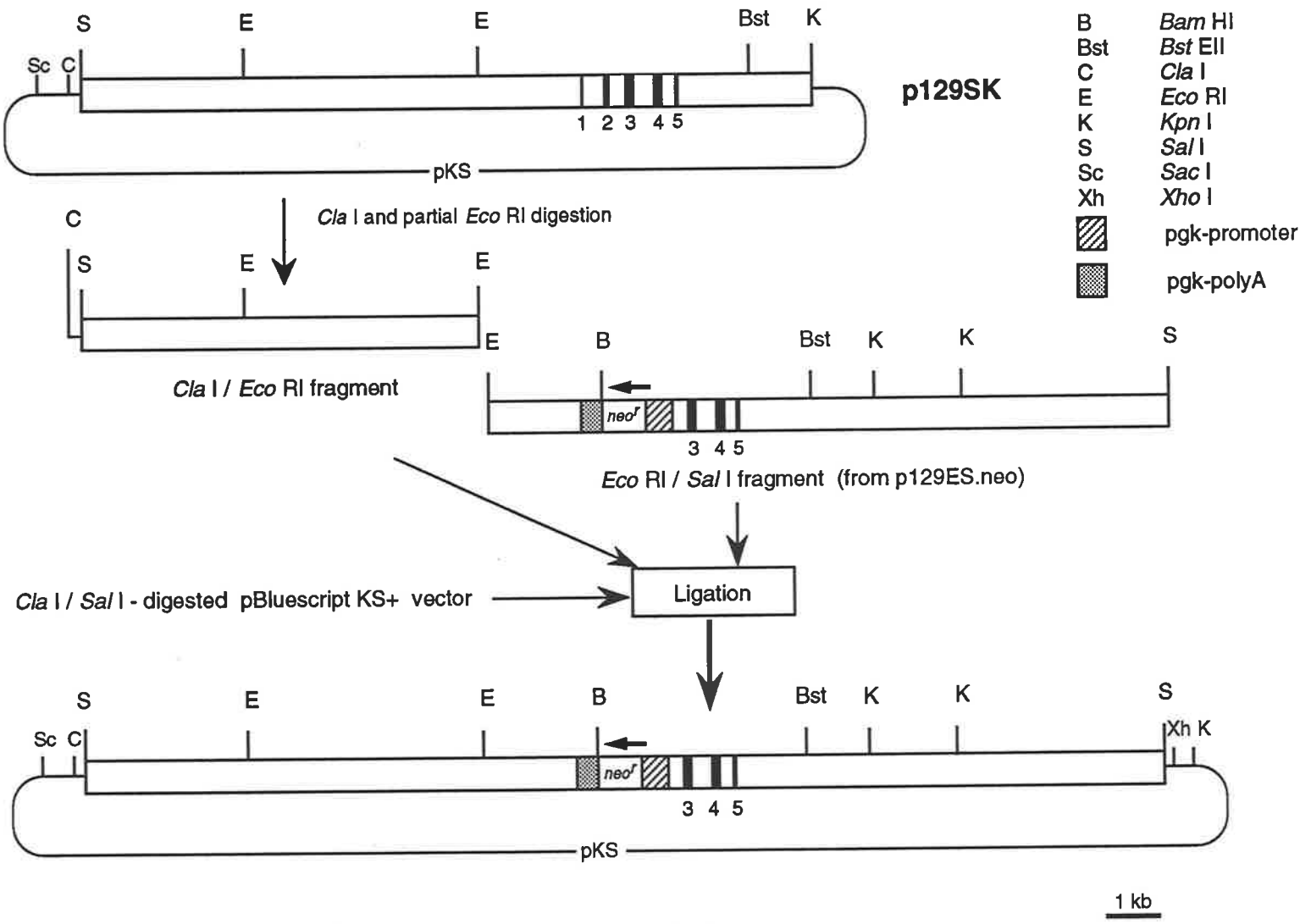
pgk-polyA

Figure 4-8

Construction of the targeting vector pHZHR3

The strategy used for constructing the targeting vector pHZHR3 is shown. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The position of the *neo*^{r} cassette within the H2A.Z fragment is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated.

pHZHR3 (23.4 kb) was generated by ligation of a 7.7 kb *Cla* I/*Eco* RI fragment of p129SK, a 12.7 kb *Eco* RI /*Sal* I fragment derived from p129ES.neo and *Cla* I/*Sal* I -digested pBluescript KS+ vector.

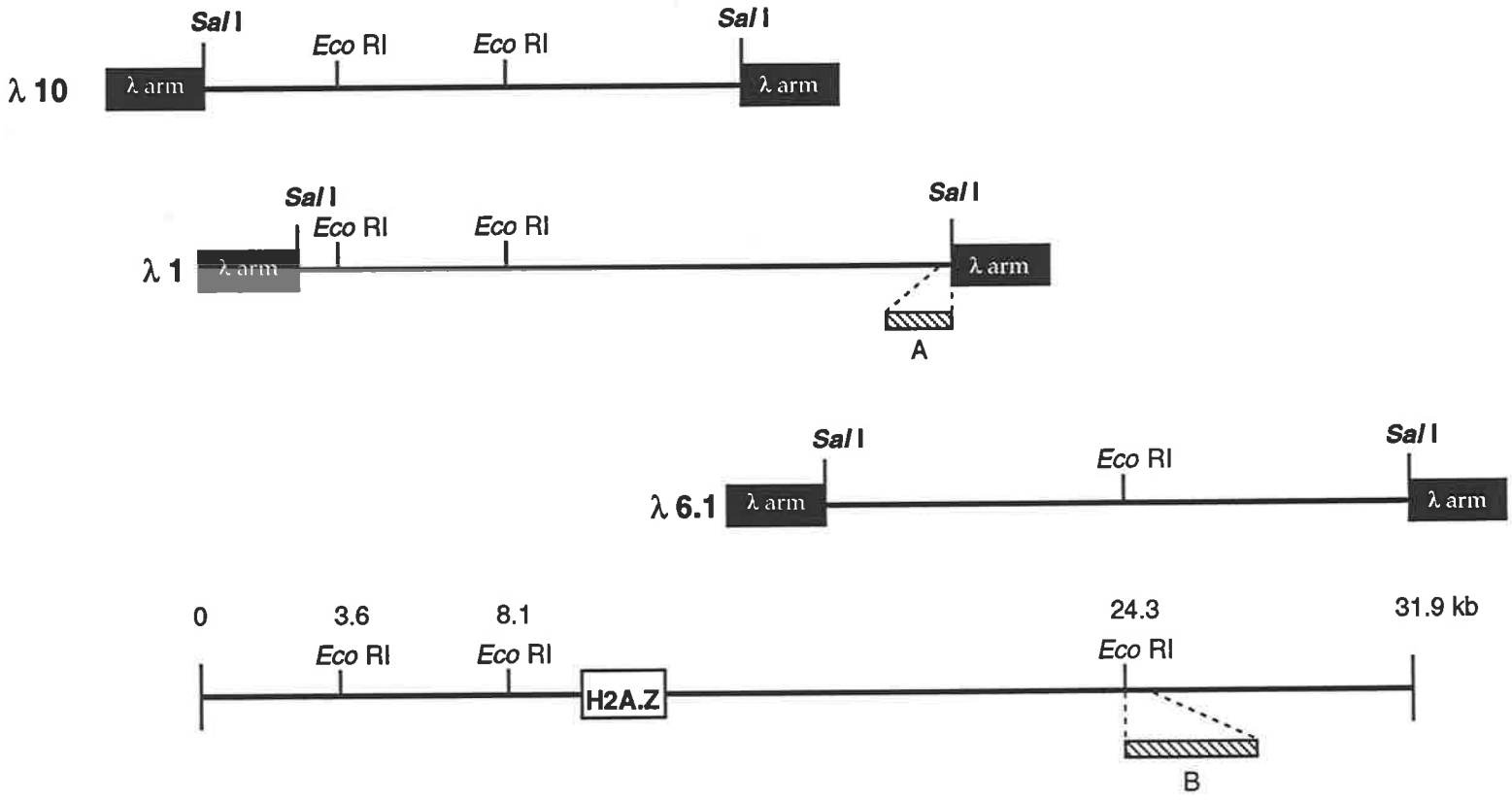

S
E
E
Bst
K
Sc
C
1 2 3 4 5
pKS
p129SK
B Bam HI
Bst Bst EII
C Cla I
E Eco RI
K Kpn I
S Sal I
Sc Sac I
Xh Xho I
pgk-promoter
pgk-polyA
Cla I and partial Eco RI digestion
Cla I / Eco RI fragment
B
neo^r
3 4 5
Eco RI / Sal I fragment (from p129ES.neo)
Cla I / Sal I - digested pBluescript KS+ vector
Ligation
Xh
1 kb

pHZHR3

Figure 4-9

Alignment of λ1, λ10 and λ6.1

The overall restriction map of the region of the genomic DNA covered by the inserts of λ1, λ10 and λ6.1 is shown below the maps of individual clones. The numbers above the map indicate the distances in kilobases from the 5' end of genomic region. The location of the H2A.Z gene is indicated. A 0.3 kb *Xba* I/*Sal* I probe used to screen mouse genomic library III (2.2.7) in order to obtain λ6.1 is shown in a striped box (A). The position of a 0.6 kb *Eco* RI/*Hin*d III fragment, used as a probe for detection of homologous recombination, is shown as a striped box (B) relative to the overall restriction map.


λ 10
Sal I
Eco RI
Eco RI
Sal I
λ arm
λ arm
λ 1
Sal I
Eco RI
Eco RI
Sal I
λ arm
λ arm
A
λ 6.1
Sal I
Eco RI
Sal I
λ arm
λ arm
0
3.6
8.1
24.3
31.9 kb
Eco RI
Eco RI
Eco RI
H2A.Z
B

(iii) Electroporation of pHZHR3 and Southern analysis

Prior to electroporation, pHZHR3 was digested with *Sal* I to release the insert and processed as described (2.3.17).

Two electroporations were performed using *Sal* I-digested pHZHR3. In the first experiment, 50 μg of DNA was introduced into $5x10^7$ E14 ES cells using the electroporation conditions as previously described (4.2.1 (ii)). Transformed cells were selected with G418 (200 μg/ml) for 10 days. G418-resistant colonies were expanded as described (2.6.4).

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Bam* HI and filters were probed with a 0.6 kb *Eco* RI/*Hin*d III fragment (probe A in Fig. 4-10) which is a 3'-external probe not present in the vector. No homologous recombination was observed in any of the 120 G418-resistant ES clones. To determine whether the introduced DNA integrated only in a single site in the genome, the filters were reprobed with the neo^r coding probe (a 696 bp *Pst* I fragment isolated from pgkNEOpA (2.2.13)). Unexpectedly, about 60% of these clones had multiple-copies of DNA randomly integrated in the genome. This suggested that these electroporation conditions, with a high concentration of DNA (50 μg) may contribute to high copy number integration.

Therefore, the second experiment was performed using 25 μg DNA of pHZHR3 introduced into ES cells ($5x10^7$) at passage 21. Electroporation was performed using the same conditions as previously described (4.2.1 (ii)). Transformed cells were selected with G418 (200 μg/ml) for 10 days. G418-resistant colonies were expanded as described (2.6.4). Genomic DNA was isolated from individual ES clone and analysed by Southern blot. DNA was digested with *Bam* HI and filters were probed with a 0.6 kb *Eco* RI/*Hin*d III fragment (probe A in Fig. 4-10) which is a 3'-external probe.

The result showed that 31 out of 305 clones gave an extra hybridising fragment of 18 kb as expected if one allele was mutated (see Fig. 4-10).

The structure of the mutated allele was further confirmed by rehybridising the same filters with two probes: (1) probe B in Fig. 4-10, a 0.6 kb *Bam* HI/*Sma* I fragment present in the vector, (2) the *neo*^{r} coding sequence probe, a 696 *Pst* I fragment isolated from pgkNEOpA (2.2.13). A surprising result was obtained from probe B. Three out of 31 positive clones identified with probe A did not give the expected band of 10 kb. These 3 clones were classified as having undergone a "pick-up" event (as explained in discussion). Hybridisation with the *neo*^{r} probe revealed that 11 out of 31 positive clones with probe A gave an 18 kb fragment as expected plus an extra fragment(s) indicating a random integration event together with an homologous recombination event in these clones. The three outcomes from this experiment are summarised in Table 4-1. An example of Southern blot analysis is shown in Fig. 4-11.

Table 4-1 Electroporation of ES cells with pHZHR3: Southern blot analysis

Event	number of clones	Southern analysis		
		3' end probe (kb)	5' end probe (kb)	*neo*^{r} probe (kb)
HR[a]	17	27, 18	27,10	18
HR and RI[b]	11	27, 18	27, 10, F[d]	18, F[e]
" pick-up[c] "	3	27, 18	27	18

305 clones were analysed by Southern blot after *Bam* HI digestion.
[a](HR) Homologous Recombination.
[b](RI) Random integration.
[c](pick-up) Type III recombination (described by Adair *et al.*, 1989).
[d](F) Fragment(s) not 10 kb in size.
[e](F) Fragment(s) not 18 kb in size.

Figure 4-10

Disruption of the H2A.Z gene in mouse embryonic stem cells

The targeting vector pHZHR3 used to disrupt the H2A.Z gene is shown at the top. The crossovers between the genomic H2A.Z allele (middle) and the targeting vector are indicated by the dashed lines. The predicted map of the mutated allele following homologous recombination is shown below. Exons of the H2A.Z gene are represented by numbered open boxes. The *neo*^{r} cassette is represented by the arrow indicating the direction of transcription. Probes A, B, C and *neo*^{r} and the lengths of fragments recognised by these probes are indicated.

The H2A.Z gene is flanked by two *Bam* HI sites about 27 kb apart. The *neo*^{r} cassette within pHZHR3 contains a *Bam* HI site approximately 10 kb from the 5' end of the vector. As a result of homologous recombination, a *Bam* HI site is inserted between the endogenous sites to generate two *Bam* HI fragments of about 10 kb and 18 kb in size.

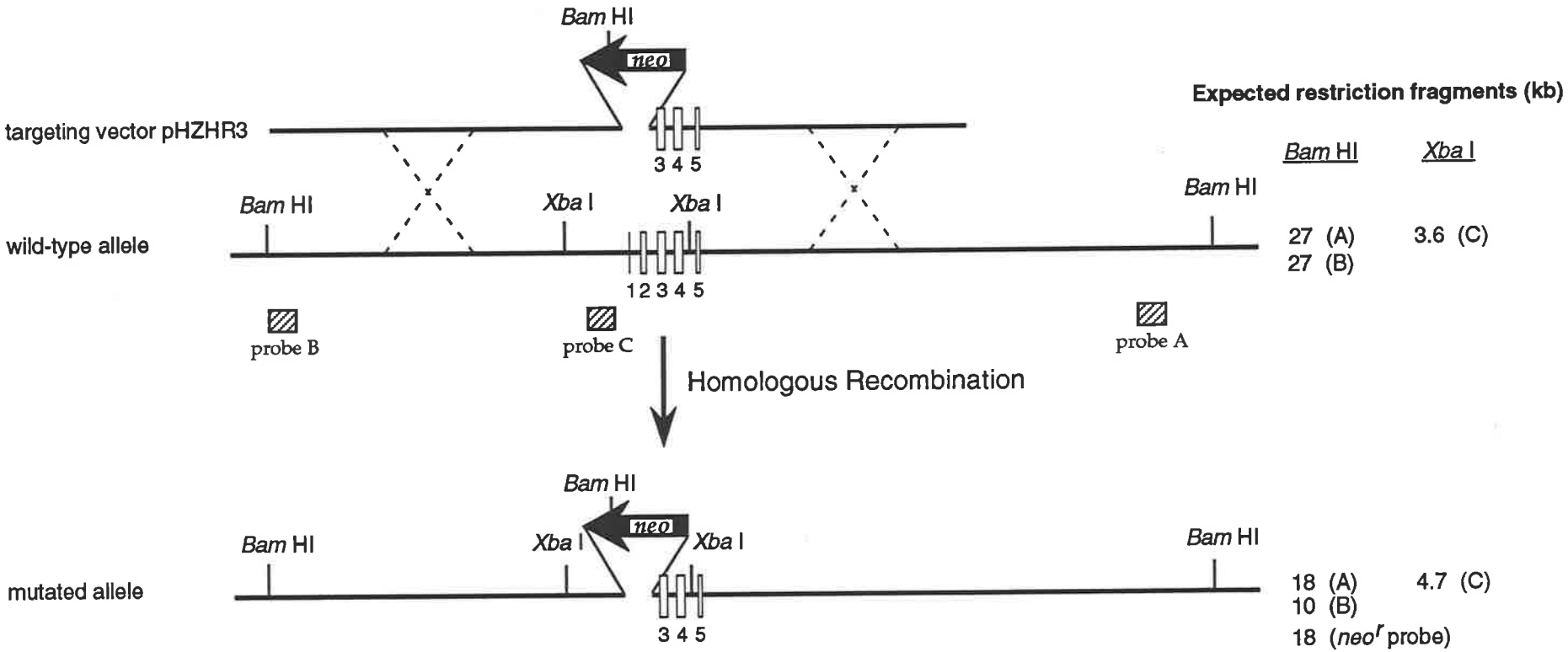
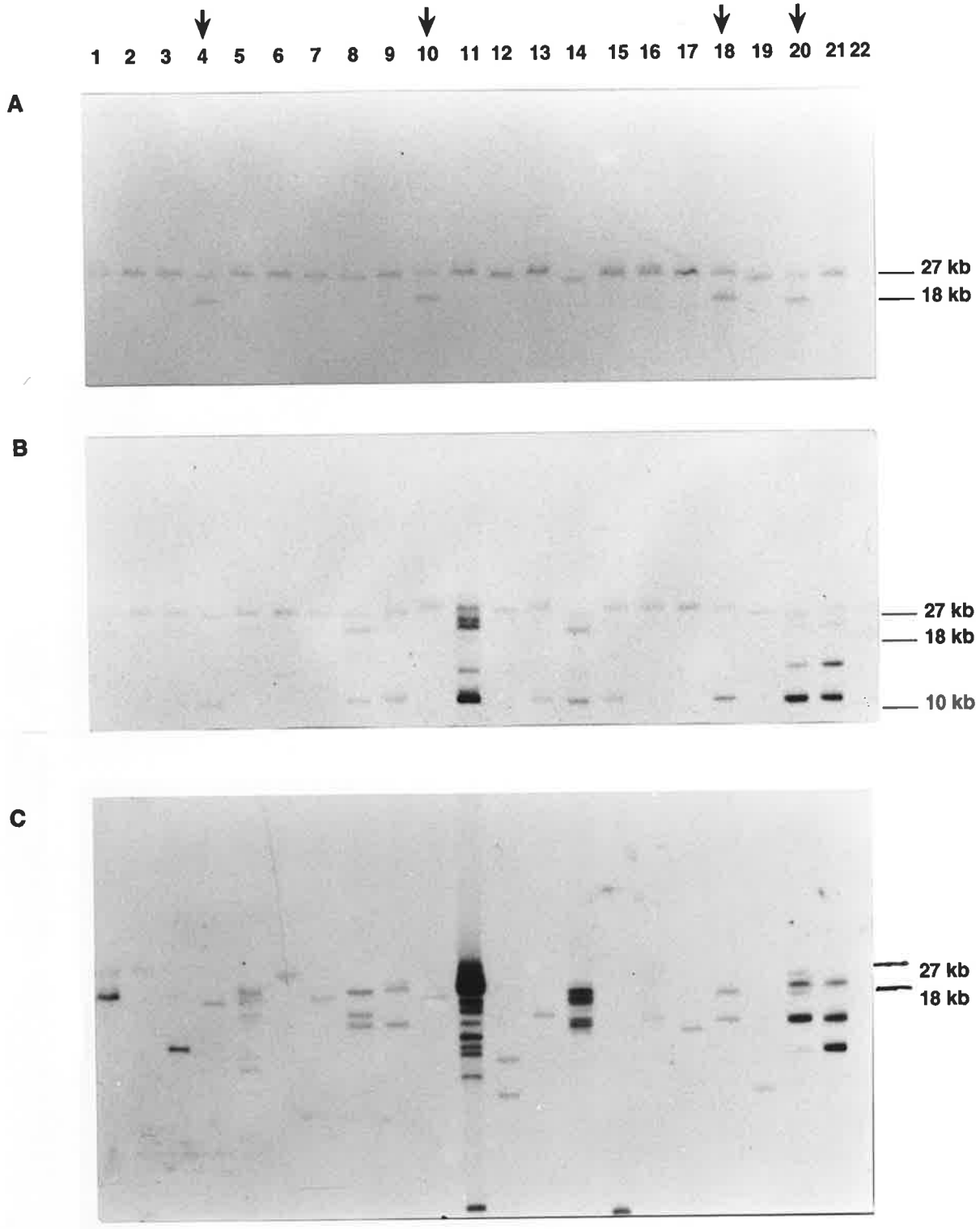

Bam HI
neo
targeting vector pHZHR3
3 4 5
Bam HI
Xba I
Xba I
Bam HI
wild-type allele
1 2 3 4 5
probe B
probe C
probe A
Homologous Recombination
Bam HI
neo
Bam HI
Xba I
Xba I
Bam HI
mutated allele
3 4 5
Expected restriction fragments (kb)
Bam HI
Xba I
27 (A)
27 (B)
3.6 (C)
18 (A)
10 (B)
18 (neo^r probe)
4.7 (C)

Figure 4-11

Southern analysis of selected ES cell clones after electroporation with pHZHR3

DNA isolated from 21 individual ES cell clones (numbers 1-21) was digested with *Bam* HI. The digested DNA (lanes 1-21) was electrophoresed on a 0.7% agarose gel along with DNA size markers (lane 22), and transferred onto Zeta-Probe membrane. The filter was initially probed with a 3' external probe (probe A in Fig. 4-10) and the result is shown in **A**. The arrows indicated the clones in which the 18 kb fragment was observed. The same filter was stripped and reprobed with a 5'-internal probe (probe B in Fig. 4-10) and neo^r probe, and the results are presented in **B** and **C** respectively.

The result obtained from these three probes indicated that ES cell clone number 4 is a homologous recombination clone. ES cell clone number 10 is classified as a "pick-up" event clone due to the absence of a 10 kb hybridising fragment when probed with probe B (Fig. 4-10). ES cell clones 18 and 20 gave a 27 kb and an 18 kb band with probe A but when these clones were probed with the 5'-internal fragment (probe B in Fig. 4-10) and neo^r probe they gave multiple bands which indicated both homologous recombination and random integration events have occurred.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
A
27 kb
18 kb
B
27 kb
18 kb
10 kb
C
27 kb
18 kb

(iv) Verification

Southern analysis of 31 positive clones was confirmed by digestion with *Bam* HI and hybridising to the three probes (probe A, B, *neo*^{r} probe in Fig. 4-10) after expanding from the frozen stock. To define the structure of the disrupted allele in detail, DNA isolated from 17 positive clones without random integration was also analysed by *Xba* I digestion using a 0.6 kb *Hin*d III/ *Xho* I probe (probe C in Fig. 4-10). This probe contained DNA sequences present in the targeting vector and located in the promoter region of the H2A.Z gene. The expected fragments of 3.6 kb (endogenous H2A.Z gene allele) and 4.7 kb (mutated allele) were observed with an equal intensity (Fig 4-12). To ~~verify the~~ "pick-up" event in the three clones, DNA was also digested with *Xba* I and probed with 0.6 kb *Hin*d III/ *Xho* I probe (probe C in Fig. 4-10). Two hybridising fragments with different intensities were observed. The 3.6 kb fragment, derived from the endogenous H2A.Z gene allele, had double the intensity of the band on autoradiography compared to the other fragment of 4.7 kb, derived from the pHZHR3 DNA (Fig. 4-12). This result indicated that the two functional copies of H2A.Z were still intact and in addition the pHZHR3 DNA was randomly integrated in the genome.

4.3 Summary and Discussion

This chapter describes targeted disruption of the histone H2A.Z gene in mouse ES cells. The successful targeting vector (pHZHR3), gave a frequency of homologous recombination of 1 in 11 of the G418-resistant clones. There are at least two factors that contribute to this high efficiency: the utilisation of isogenic DNA and the long stretch of homologous DNA (a total of 18.6 kb) between the vector and the H2A.Z locus.

Southern analysis was used to identify the homologous recombination

Figure 4-12

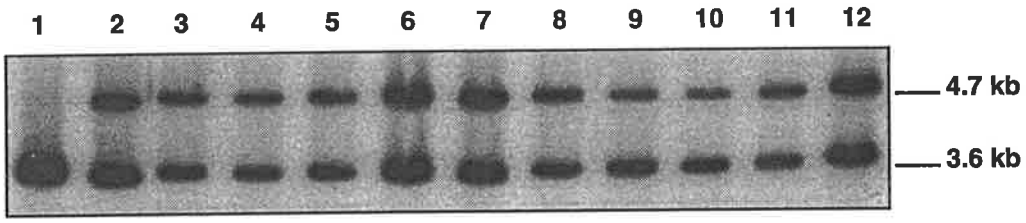
Southern analysis of positive ES cell clones

DNA was digested with *Xba* I. The digested DNA (lanes 1-12) was electrophoresed on a 1.0% agarose gel along with DNA size markers, and transferred onto Zeta-Probe membrane. The filter was probed with a 0.6 kb *Hin*d III/*Xho* I fragment (probe C in Fig. 4-10). The endogenous H2A.Z allele and the disrupted H2A.Z allele were detected as a 3.6 kb and a 4.7 kb hybridising fragment, respectively.

Lane 1: Unmanipulated E14 ES cell DNA. The H2A.Z allele was detected as a 3.6 kb hybridising fragment with the probe.

Lane 2, 9, and 10: ES cell clones classified as "pick-up" event DNA. Two hybridising fragments, a 3.6 kb (endogenous H2A.Z gene) and a 4.7 kb (randomly integrated pHZHR3 DNA), were detected. The 3.6 kb fragment has double the intensity of 4.7 kb fragment.

Lane 3-8, 11 and 12: Targeted ES cell clone DNA. Two hybridising fragments, 3.6 kb (endogenous H2A.Z gene) and 4.7 kb (disrupted H2A.Z allele) in length, were detected with equal intensity.


1
2
3
4
5
6
7
8
9
10
11
12
4.7 kb
3.6 kb

events and confirm the structure of the disrupted H2A.Z allele. As previously mentioned, 3 out of 31 clones were classified as "pick-up" or type III recombination events as described by Adair *et al.*, (1989). The frequency of this event is 1 out of 10 homologous recombinants, which is higher than previously reported by other investigators. A frequency of 1 out of 21 homologous recombinants in the *pim-1* targeting experiment was reported by te Riele *et al.* (1990), and a frequency of 1 out of 48 homologous recombinants in the retinoblastoma gene experiment was reported by te Riele *et al.* (1992). This event is probably a result of incomplete homologous recombination, with the products of incomplete reaction subsequently integrating elsewhere in the genome. Southern analysis with *Xba* I and a probe containing DNA sequence present in the targeting vector (probe C in Fig. 4-10) confirmed that two functional copies of the H2A.Z gene were still intact, with one copy of the disrupted H2A.Z gene within the targeting vector randomly integrated in the genome.

Southern analysis using the *neo*^{r} probe (4.2.4 (iii)) suggested that the amount of DNA and the electroporation conditions, used to introduce DNA into ES cells, may be factors that contribute to multiple copies of DNA integration in the genome. This speculation led to the evaluation of various electroporation conditions with different amounts of DNA, with an aim to define the best conditions resulting in a single site of DNA integration. This experiment appears in chapter 5.

Targeting attempts using pHZHR1 and pHZHR2 were unsuccessful. This may be due to the fact that pHZHR1 was derived from non-isogenic DNA and the length of homologous DNA was too short (only 2.4 kb homology at the 5' end and 1.3 kb homology at the 3' end of *neo*^{r} gene) to allow the alignment of the homologous region to occur. Although pHZHR2 was derived from isogenic DNA, the low targeting efficiency observed may have been due to it utilising only 5.8 kb of homology with the H2A.Z locus.

Chapter 5

Production of H2A.Z Null Mutants In vitro and In vivo

5.1 Introduction

The aim of this work was to establish homozygous H2A.Z gene "knock-out" ES cell lines, and create homozygous H2A.Z gene "knock-out" mice.

To generate homozygous H2A.Z gene "knock-out" ES cell lines, the heterozygous H2A.Z gene "knock-out" ES cells lines were subjected to a second electroporation to disrupt the remaining wild-type H2A.Z allele. To do this, a targeting vector containing a positive selectable marker different from the neo^r gene was required. Therefore, a new vector named pHZHR4 was constructed containing the hygromycin phosphotransferase (hyg^r) gene as a positive selectable marker. This construct was designed to improve the efficiency of homologous recombination using a long stretch (17 kb) of homologous isogenic DNA.

The electroporation protocol previously used (described in Chapter 4) may contribute to multiple copies of DNA integrating into the genome. Therefore, prior to performing the second electroporation, various electroporation conditions were evaluated. Conditions which most frequently gave integration of a single copy of the targeting construct were used.

To create homozygous H2A.Z gene "knock-out" mice, heterozygous H2A.Z gene "knock-out" ES cells lines were injected into host blastocysts and then transferred into pseudopregnant mothers to generate chimaeric mice. The heterozygous animals were obtained and the breeding program was set up in order to obtain the homozygous mutant animals.

This chapter describes the approaches used to obtain ES cell lines, and mice, which a lack the functional histone H2A.Z gene.

5.2 Results

5.2.1 Test for electroporation conditions resulting in single copy integration of introduced DNA

Experiments were designed to identify conditions for introducing a single copy of the targeting vector per transformed ES cell. Parameters explored included the amount of DNA, electroporation buffer, and capacitance and voltage settings used for electroporation. High (20 μg) and low (5 μg) amounts of pHZHR3 DNA (4.2.4 (i)) were tested. Three different electroporation protocols (see section 2.6.3) which were described by other investigators, including the original electroporation (protocol 1), were tested. They are as follows:

Protocol 1: HBS was used as an electroporation medium and the settings were 960 μF capacitance, 210 volts;

Protocol 2: PBS was used as an electroporation medium and the settings were 25 μF capacitance, 350 volts (J. Beall, pers. comm.);

Protocol 3: MT PBS was used as an electroporation medium and the setting were 500 μF capacitance, 220 volts (I. Lyons, pers. comm.).

E14 ES cells resuspended in electroporation buffer were mixed with *Sal* I-digested pHZHR3 DNA (20 μg or 5 μg) and the electroporation was performed as described (2.6.3). Forty-eight G418-resistant clones were picked from each electroporation experiment after 10 days of selection with G418 at 150 μg/ml (2.6.4). Genomic DNA was isolated from individual clones and analysed by Southern blot. DNA was digested with *Bam* HI and filters were probed with a *neo*^{r} coding sequence probe (696 bp *Pst* I fragment isolated from pgkNEOpA, see Fig. 4-10). There is only one *Bam* HI site in the pHZHR3 DNA. The size of the hybridising fragments with this probe varies depending on the next *Bam* HI site in the genome from the site of integration. Therefore the number of integration copies can be determined by the number and intensity of the hybridising fragments. The data are presented in Table 5-1.

Table 5-1 The effect of electroporation conditions and amount of introduced DNA (pHZHR3) on the number of integration copies

Electroporation condition	DNA used (μg)	number of analysed clones	% of clones with 1 copy	% of clones with >1 copy
HBS 960 μF, 210 volts	5	47	52	48
	20	48	50	50
PBS 25 μF, 350 volts	5	38	87	13
	20	44	45	55
MT PBS 500 μF, 220 volts	5	44	56	44
	20	40	50	50

The condition that gave the highest percentage of single copy integration was electroporation with the electric pulse setting at 25 μF, 350 volts and using 5 μg of DNA in PBS. This condition was then used for electroporating DNA into ES cells in the following experiments.

5.2.2 An attempt to create a homozygous H2A.Z gene "knock-out" ES cell line

(i) Construction of pHZHR4

pHZHR4 was constructed from H2A.Z clones derived from a 129/Sv mouse genomic library (described in chapter 4). It contained 17 kb of DNA homologous to the H2A.Z locus and the hygromycin phosphotransferase (*hyg*^{r}) gene as a selectable marker. The *hyg*^{r} gene was driven by the phosphoglycerate

kinase (pgk) promoter, with pgk polyadenylation signals. The cells that incorporated pHZHR4 were resistant to hygromycin B.

To create pHZHR4, three recombinant plasmids named pUChyg, p129St/Xh.hyg and p129Bst/S, were generated as precursors. pUChyg was derived from pgkHYG (2.2.13) by subcloning a fragment containing the *hyg*^{r} cassette into pUC19 vector. The map of pUChyg is presented in Fig. 5-1. p129St/Xh.hyg was derived from pUChyg. The strategies used to create p129St/Xh.hyg and p129Bst/S are presented in Fig. 5-1 and Fig. 5-2, respectively.

The strategy used for constructing pHZHR4 from the precursors is shown in Fig. 5-3. pHZHR4 contained the *hyg*^{r} cassette transcribed in the same direction as H2A.Z and located at a position almost identical to where the *neo*^{r} cassette was located in pHZHR3. H2A.Z is rendered nonfunctional by the loss of a 0.8 kb *Xho* I/*Bst* EII fragment containing part of the promoter region, exon 1, intron 1, exon 2 and part of intron 2.

(ii) Gene targeting using pHZHR4

The ES cell line designated ES-193 is one of the 18 positive clones in which one allele of H2A.Z has been disrupted as described in chapter 4 (Table 4-1). This clone was chosen on the basis of normal karyotype, good morphology and normal growth characteristics. ES-193 was subjected to second round electroporation in order to generate homozygous H2A.Z gene "knock-out" ES cell clones.

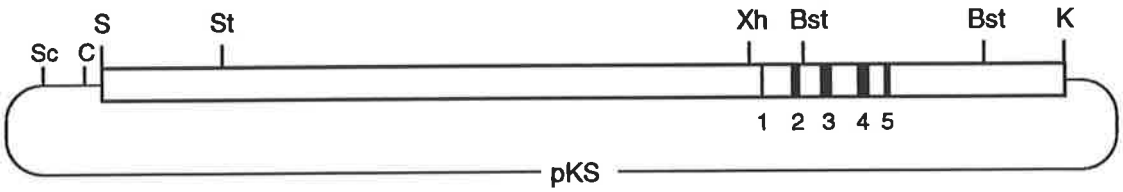
Before electroporation, pHZHR4 was linearised with *Sal* I to release the insert and processed as described (2.3.17). ES-193 cells ($5x10^7$) were transformed with 5 µg of *Sal*-digested pHZHR4 using and electroporation protocol 2 (see section 5.2.1). Transformed cells were selected in the presence of hygromycin (110 µg/ml) for 8 days. Hygromycin-resistant colonies were expanded as described (2.6.4).

Figure 5-1

Construction of pUChyg and p129St/Xh.hyg

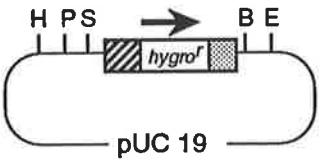
The strategy used for constructing p129St/Xh.hyg plasmid, as a precursor for pHZHR4, is shown. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The position of the *hyg*^{r} cassette within the H2A.Z fragment is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated. The restriction sites which were lost as a result of the ligation reaction are shown in brackets.

pUChyg was constructed by subcloning a 2.0 kb *Bgl* II fragment containing the *hyg*^{r} cassette isolated from pgkHYG (2.2.13) into *Xba* I-digested pUC19 vector. p129SK contained a 14.1 kb *Sal* I/*Kpn* I fragment subcloned into pBluescript KS+ vector. p129St/Xh.hyg was constructed by subcloned a 7.6 kb *Stu* I/*Xho* I fragment of p129SK into *Sal* I-digested pUChyg vector. The ligation between the blunt end fragment and blunt end vector in the correct orientation generates a unique *Sal* I site within the recombinant plasmid as shown in the map.



p129SK

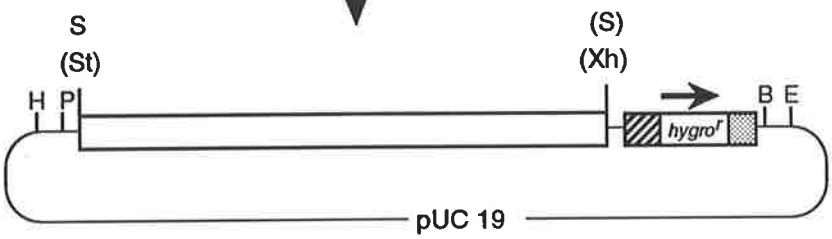
Stu I / *Xho* I - digested fragment (end - filled)



pUChyg

Ligation

Sal I - digested vector (end - filled)



p129St/Xh.hyg

1 kb

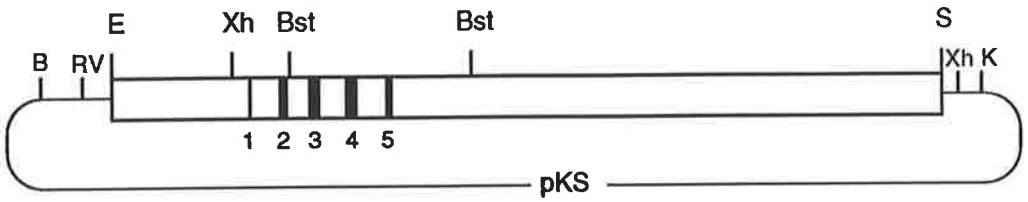
B	*Bam* HI	Bst	*Bst* EII
C	*Cla* I	E	*Eco* RI
H	*Hind* III	K	*Kpn* I
P	*Pst* I	S	*Sal* I
Sc	*Sac* I	St	*Stu* I
Xh	*Xho* I		
▨	pgk-promoter	▩	pgk-poly A

Figure 5-2

Construction of p129Bst/S

The strategy used for constructing the p129Bst/S plasmid, as a precursor for pHZHR4, is shown. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated. The restriction sites which were lost as a result of the ligation reaction are shown in brackets.

p129ES contained a 12 kb *Eco* RI/*Sal* I fragment subcloned into pBluescript KS+ vector. p129Bst/S was constructed by removing a 2.6 kb *Eco* RI/*Bst* EII fragment from p129ES.

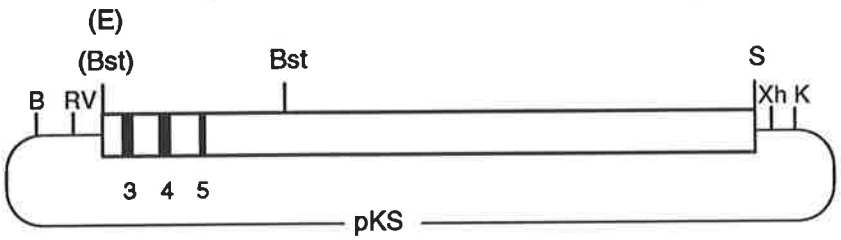


p129ES

(partially - digested) *Bst* EII and *Eco* RI digestion

End - filling

Ligation



p129Bst/S

1 kb

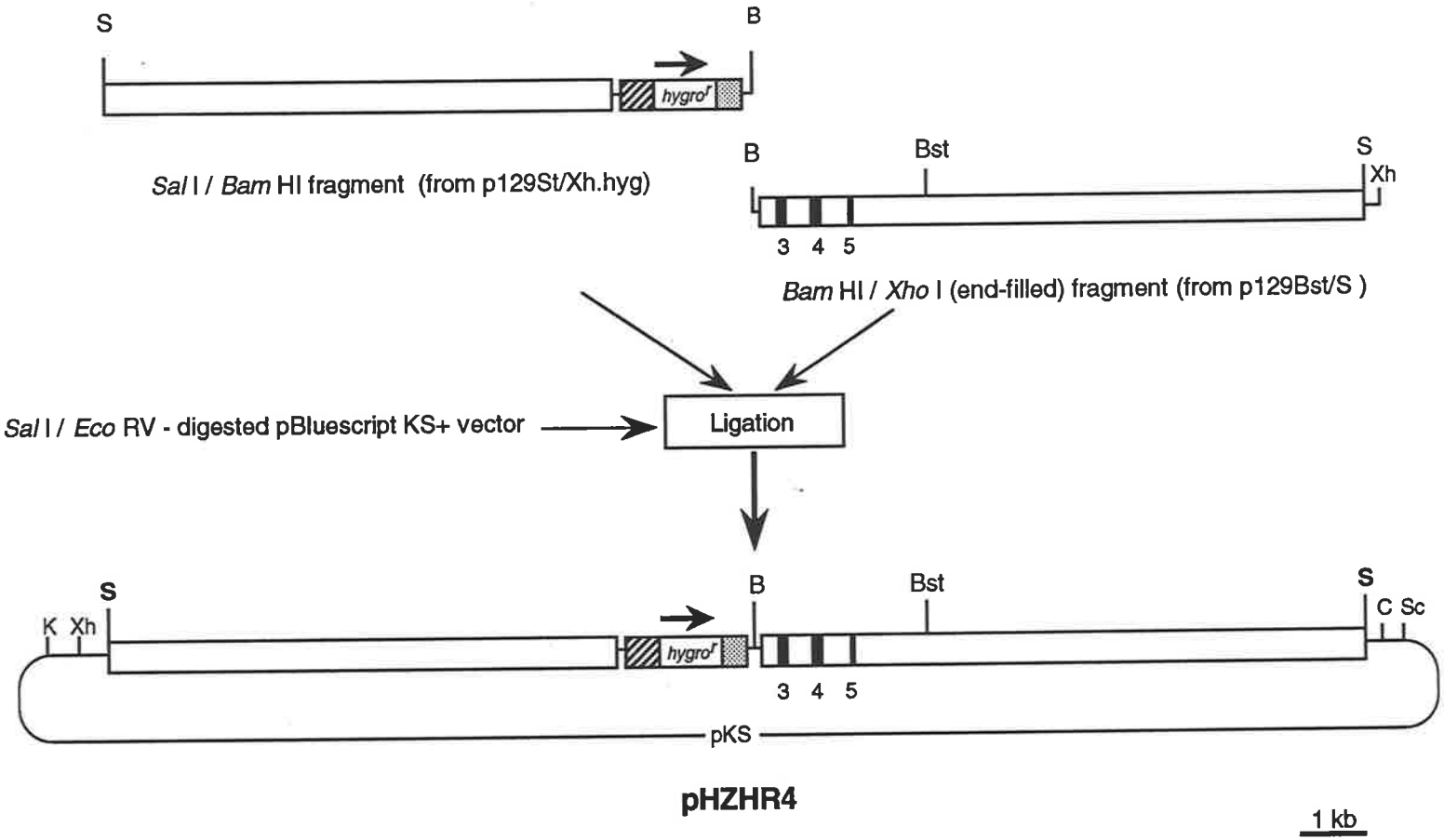
B	*Bam* HI	Bst	*Bst* EII
E	*Eco* RI	K	*Kpn* I
RV	*Eco* RV	S	*Sal* I
Sc	*Sac* I	Xh	*Xho* I

Figure 5-3

Construction of the targeting vector pHZHR4

The strategy used for constructing the targeting vector pHZHR4 is shown. The H2A.Z genomic fragment is shown as an open box with exons filled and numbered. The position of the *hyg*^{r} cassette within the H2A.Z fragment is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment and the polylinker are abbreviated as indicated.

pHZHR4 (22 kb) was generated by ligation of a 9.6 kb *Sal* I /*Bam* HI fragment isolated from p129St/Xh.hyg, a 9.4 kb *Bam* HI/*Xho* I (end-filled) fragment isolated from p129Bst/S, and *Sal* I/*Eco* RV-digested pBluescript KS+ vector.



pHZHR4

1 kb

B	*Bam* HI	Bst	*Bst* EII	C	*Cla* I
E	*Eco* RI	K	*Kpn* I	S	*Sal* I
Sc	*Sac* I	Xh	*Xho* I		

pgk-promoter

pgk-poly A

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Bam* HI and filters were probed with a 0.6 kb *Bam* HI/*Sma* I fragment, (probe A in Fig. 5-4). As this probe did not contain any sequence present in pHZHR4, it hybridised only to the endogenous H2A.Z gene allele and not to randomly integrated pHZHR4. The functional H2A.Z allele and the disrupted H2A.Z (neo^r) allele present in the genome of ES-193 cells were detected as 27 kb and 10 kb hybridising fragments, respectively. If homologous recombination occurred, the *Bam* HI site within pHZHR4 would integrate in the genome. Thus, the disrupted H2A.Z (hyg^r) allele generated by this vector was expected to generate a 11.6 kb hybridising fragment (Fig. 5-4). ES cell clones in which both alleles were disrupted were not found in any of 282 hygromycin-resistant clones. However, 5 of the analysed clones appeared to contain two disrupted alleles plus a wild-type allele. Also, 9 clones had undergone homologous recombination in such a way that the hyg^r sequence had replaced the neo^r sequence. Therefore the targeting efficiency for this experiment was approximately 14 out of 282 clones (~ 1 in 20 analysed clones). An example of the Southern analysis is presented in Fig. 5-5.

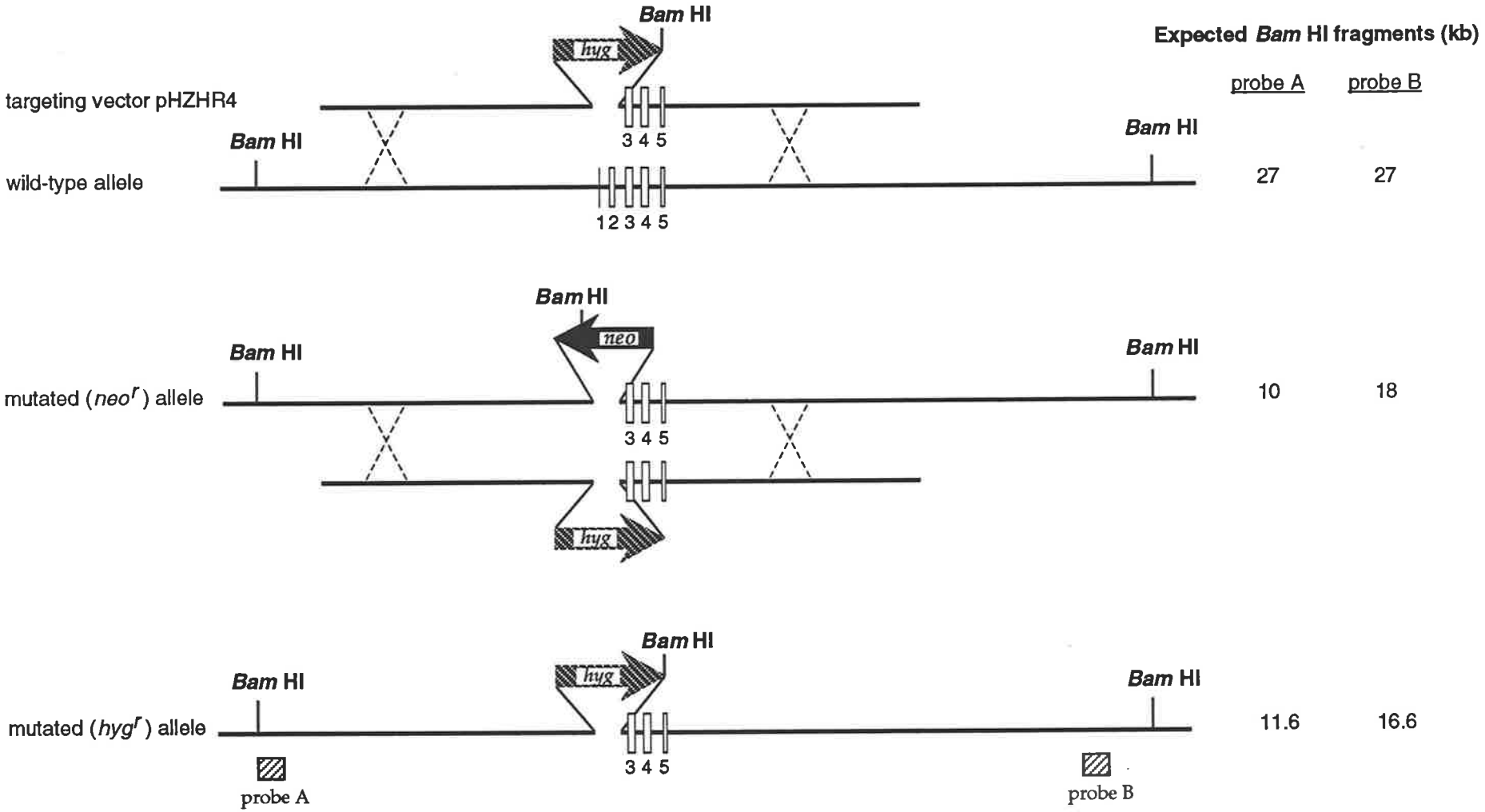
The apparent presence of three alleles in 5 of the clones could be explained in several ways. Firstly, there are three alleles in these clones. Secondly, the 27 kb hybridising fragment may be the result of incomplete digestion by *Bam* HI. Finally, it may be the result of mixed populations of cells (the original ES-193 and cells in which hyg^r replaced neo^r sequence). To rule out the latter two possibilities, one of these clones, designated ES-dko-122, was subjected to further investigation. ES-dko-122 cells were grown at low density and 24 subclones were isolated. Genomic DNA from the 24 subclones was digested with *Bam* HI and filters were probed with a 0.6 kb *Bam* HI/*Sma I* fragment (probe A in Fig. 5-4). The structure of the three hybridising fragments was investigated further by rehybridising the same filter with a 0.6 kb *Eco* RI/*Hin*d III fragment (probe B in

Figure 5-4

Targeting with pHZHR4 in heterozygous H2A.Z gene " knock-out " ES cells

The targeting vector pHZHR4 is shown. The crossovers between pHZHR4 and the two H2A.Z alleles, wild-type and mutated (*neo*^{r}), are indicated by the dashed lines. The predicted map of the mutated allele following homologous recombination is shown at the bottom. Exons of H2A.Z gene are represented by open boxes and numbered. The *neo*^{r} and *hyg*^{r} cassettes are represented by arrows indicating the direction of transcription. Probe A (0.6 kb *Bam* HI/*Sma* I fragment) and B (0.6 kb *Eco* RI/*Hin*d III fragment), and the lengths of fragments recognised by these probes are indicated.

Southern analysis of heterozygous H2A.Z gene knock-out ES cell DNA using *Bam* HI and probe A generates two fragments: one is 27 kb, representing a wild-type allele, and the other one is 10 kb, representing the mutated (*neo*^{r}) allele. Homologous recombination between pHZHR4 and the wild-type allele results in ES cells in which both alleles have been disrupted. Southern analysis of these cells using *Bam* HI digestion and probe A shows hybridising fragments of 10 kb (*neo*^{r} allele) and 11.6 kb (*hyg*^{r} allele).



Expected *Bam* HI fragments (kb)

	probe A	probe B
wild-type allele	27	27
mutated (*neo*^{r}) allele	10	18
mutated (*hyg*^{r}) allele	11.6	16.6

Figure 5-5

An example of Southern analysis of selected ES cell clones after electroporation with pHZHR4

DNA isolated from individual ES cell clones was digested with *Bam* HI. The digested DNA was electrophoresed on a 0.7% agarose gel along with DNA size markers, and transferred onto Zeta-Probe membrane. The filter was probed with a 5'-external probe (probe A in Fig. 5-4).

Lanes are as follows:

1) ES-dko-121 DNA

2) ES-dko-122 DNA

3) ES-dko-123 DNA

The result indicates that in lane 1: no homologous recombination event by pHZHR4 has occurred; lane 2: it appears that both alleles of the H2A.Z gene have been disrupted, but that a wild-type allele remains; lane 3: homologous recombination has occurred in such a way that the hyg^r sequence has replaced the neo^r sequence.

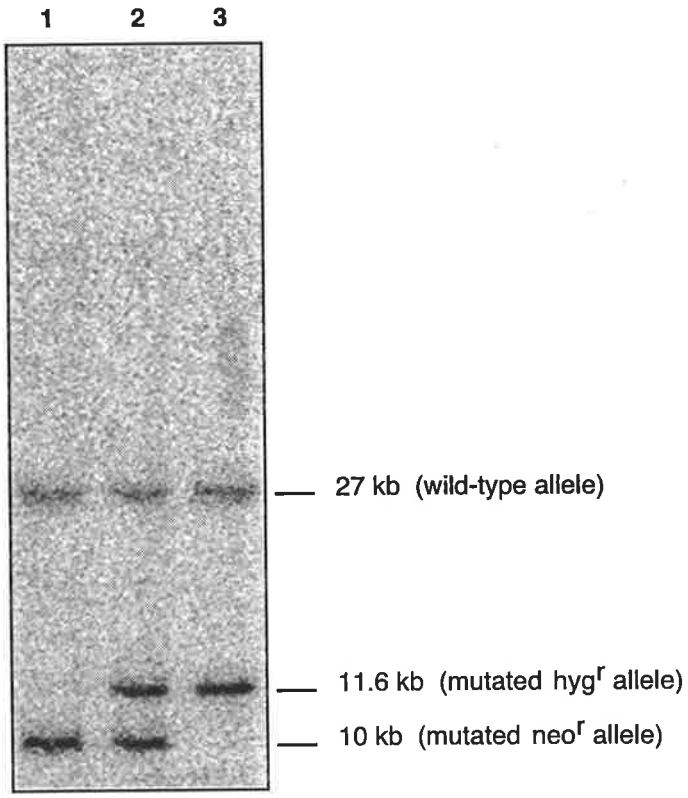

1
2
3
27 kb (wild-type allele)
11.6 kb (mutated hygr allele)
10 kb (mutated neor allele)

Fig. 5-4). All 24 clones gave the same hybridising fragments indicating that the original clone was a pure population. The results obtained from 16 clones are presented in Fig. 5-6. It can be seen that all three alleles were detectable and the size of hybridising fragments were as expected for the wild type allele, the mutated (hyg^r) allele and the mutated (neo^r) allele with both 5'-external and 3'-external probes. This result verified that there are three alleles present in the genome of the ES-dko-122, and that the cell population was clonal.

To determine whether the apparent presence of the third allele was due to gaining an extra chromosome, karyotyping of the ES-dko-122 cells were performed in parallel with the E14 ES cells and the parental ES-193 cells. The results show that ES-dko-122 has a normal karyotype (G. Webb, pers. comm.), indicating that the extra copy of the H2A.Z gene was not due to gross changes at the chromosome level.

The mechanism which generates the third allele remains unknown. It has been postulated that the parental ES-193 cells may have had an extra allele before the second round of gene targeting. This possibility is unlikely since Southern analysis of this clone using *Xba* I digestion and probed with 0.6 kb *Hin*d III/*Xho* I fragment (probe C in Fig. 4-10) gave two hybridising fragments, of 3.6 kb (H2A.Z wild-type allele) and 4.7 kb (the mutated neo^r allele) in length, with equal intensity on the autoradiograph (see lane 7 in Fig. 4-12). Therefore, it seems likely that the third allele was generated during the second round gene targeting experiment. It is possible that the third allele occurred as a result of type III recombination (Adair et al., 1989). However, type III events usually show "pick-up" on only one side of the locus. The result presented here is unique in that the extra band is detectable with both 5'-external and 3'-external probes (as mentioned above). The 5'-external probe is located 1.6 kb upstream from the targeting sequence and the 3'-external probe located 4.2 kb downstream from the targeting sequence. If the third allele is the result of type III recombination, it

Figure 5-6

Southern analysis result of 16 individual clones isolated from a pool of ES-dko-122 clones

DNA isolated from individual ES clones was digested with *Bam* HI. The digested DNA was electrophoresed on a 0.7% agarose gel along with DNA size markers, and transferred onto Zeta-Probe membrane. The filter was initially probed with a 5'-external probe (probe A in Fig. 5-4) and the result is shown in panel **A**. The same filter was stripped and reprobed with a 3'-external probe (probe B in Fig. 5-4), and the result is shown in panel **B**.

Lanes are as follows:

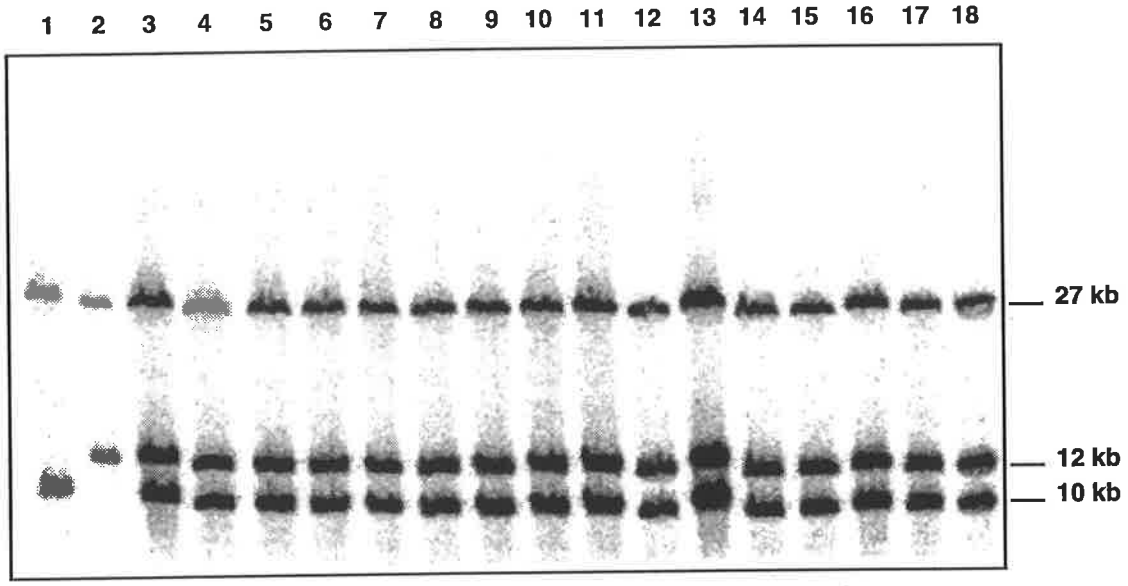
1) ES-193 DNA (as a positive control for neo^r mutated H2A.Z allele)

2) ES-dko-123 DNA (as a positive control for hyg^r mutated H2A.Z allele)

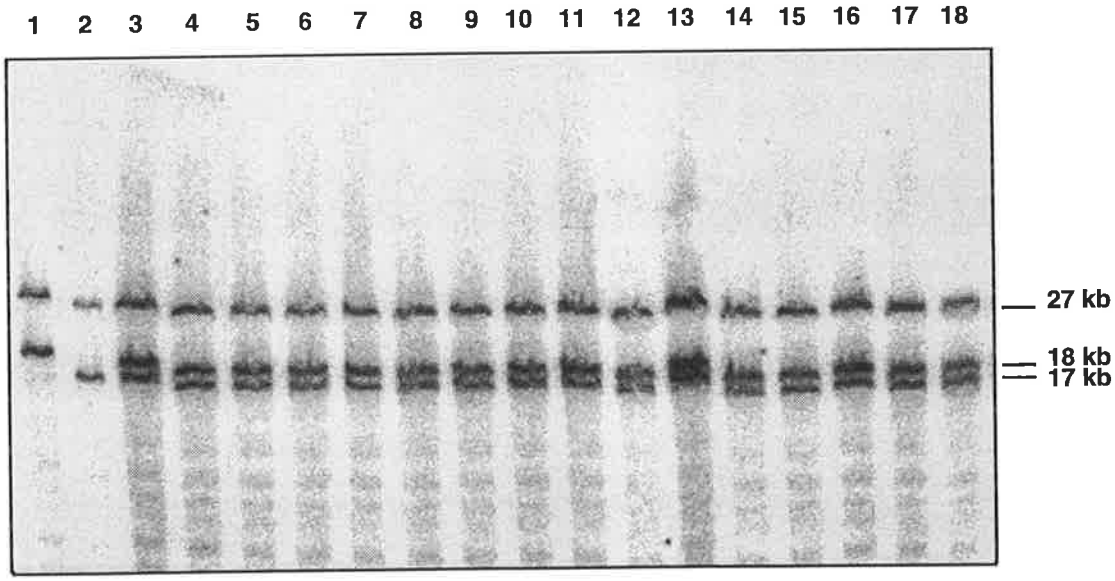
3-18) DNA of individual ES cell clones (numbers 1-16), isolated from pool of ES-dko-122 clones

The result obtained from these experiments indicated that the ES-dko-122 clone is pure. Southern analysis using probe A gave the same result as previously obtained (see lane 2 in Fig. 5-5). The result obtained from the two probes confirmed the apparent presence of the three H2A.Z alleles (wild-type, neo^r -targeted, and hyg^r-targeted).

A


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
27 kb
12 kb
10 kb

B


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
27 kb
18 kb
17 kb

must have occurred in such a way that incomplete homologous recombination "picks up" the surrounding sequence at both ends of the locus (the size of the "pick-up" sequence is at least 27 kb, which includes the two *Bam* HI sites) before random integration elsewhere in the genome. The other possibility is that duplication of the functional H2A.Z locus by another mechanism, has allowed a second round of gene targeting to occur without eliminating H2A.Z gene function.

The absence of any "double knock-out" ES cells suggests that functional H2A.Z protein is required for ES cell viability. However, this result was obtained from one experiment and only one particular ES clone. Therefore, it was necessary to verify this result, as described in the following section.

5.2.3 Verification of the double allele H2A.Z gene knock-out experiment

To confirm that ES cells lacking a functional H2A.Z gene do not survive, the double "knock-out" was repeated, but with a reversal of the order of the targeting constructs used. Therefore, E14 ES cells were initially electroporated with pHZHR4, and targeted cells were subsequently electroporated with pHZHR3 (4.2.4).

(i) One allele gene targeting using pHZHR4

E14 ES cells ($5x10^7$) were mixed with 5 µg of *Sal* I-digested pHZHR4 (5.2.2 (i)) and electroporation was performed using protocol 2 (see section 5.2.1). Transformed cells were selected in the presence of hygromycin (110 µg/ml) for 8 days. Hygromycin-resistant colonies were expanded as described (2.6.4).

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Bam* HI and filters were probed with a 0.6 kb *Bam* HI/*Sma* I fragment (probe A in Fig. 5-4), which is a 5'-external probe. The results show that 17 out of 116 hygromycin-resistant clones gave an extra

hybridising fragment of 12 kb corresponding to the mutated allele (data not shown). To confirm the structure of the mutated allele, genomic DNA was reanalysed by Southern blot using *Bam* HI and probed with a 0.6 kb *Eco* RI/*Hin*d III fragment, a 3'-external probe (probe B in Fig. 5-4) and a 1.2 kb *Aat* II/*Sca* I fragment isolated from pUChyg (5.2.2 (i)) containing the coding sequence of the hyg^r gene. The result obtained with the hyg^r coding gene probe indicated that the introduced DNA was present as a single copy in all of these clones, and supported the finding that the electroporation conditions used in this experiment result in a high percentage of single copy integration (see section 5.2.1).

The results presented here indicated that the targeting efficiency of pHZHR4 in this experiment was approximately 1 in 7 of analysed hygromycin-resistant clones. A targeted ES cell clone, designated ES-hyg-123, was chosen for the second round of gene targeting.

(ii) Second allele gene targeting using pHZHR3

ES-hyg-123 cells ($5x10^7$) were mixed with 5 µg of *Sal* I-digested pHZHR3 (4.2.4 (i)) and electroporation was performed using protocol 2 (see section 6.1). Transformed cells were selected in G418 (150 µg/ml) for 8 days. G418-resistant colonies were expanded as described (2.6.4).

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Bam* HI and filters were probed with a 0.6 kb *Eco* RI/*Hin*d III fragment, which is a 3'-external probe (probe B in Fig. 5-4). ES cell clones in which both alleles have been disrupted were not found in any of 293 G418-resistant clones. In 2 out of the analysed clones, the wild-type H2A.Z hybridising fragment was found in conjunction with both neo- and hyg-modified fragments. In 8 clones, homologous recombination had occurred in such a way that the neo^r sequence had replaced the hyg^r sequence. Therefore

the targeting efficiency in this experiment was 10 in 293 (1/30) analysed G418-resistant clones.

This result is consistent with the previous result (described in 5.2.2 (ii)) in that ES cell clones in which both alleles have been disrupted were not obtained. Therefore it is likely that ES cells require a functional H2A.Z gene for viability.

5.2.4 Generation of H2A.Z deficient mice

(i) Blastocyst injection and production of chimaeric mice

Two independent H2A.Z-targeted E14 ES cells lines (ES-193 and ES-325, described in chapter 4) were used to make chimaeras (2.6.6). These two clones were not infected with mycoplasma, had a normal karyotype and undifferentiated ES cell morphology *in vitro*. The two ES cell lines used in these injections were passage number 4. Ten to fifteen ES cells were injected into blastocysts which were transferred to the uteri of CBAxC57B1/6J F1 pseudopregnant females. Chimaeric pups were identified by coat colour. CBAxC57B1/6J F1 embryo donors are *Aa BB CC* at the agouti, B, and albino loci, and are phenotypically agouti. Their offspring are agouti (*AA* or *Aa*) or black (*aa*) in the mendelian ratio of 3:1. E14 cells are A^wA^w *bb* $c^{ch}c^{ch}$ and give rise to chinchilla coat when E14 ⟷ agouti chimaeras are generated. E14 ⟷ black chimaeras have chinchilla and black patches of coat, or regions of agouti where low levels of E14 cells provide Agouti factor to skin that would otherwise be black. The blastocyst injection was initially performed using an unmanipulated E14 ES cell clone to test whether the parental E14 ES cells could contribute to the germ line. 36, 47 and 13 pups were born from blastocysts transferred after injection with unmanipulated E14 ES cells, ES-193 ES cells and ES-325 ES cells respectively. The percentage of coat colour chimaerism in these pups is presented in Fig. 5-7. Among the chimaeras with some chinchilla coat

Figure 5-7

Coat colour of chimaeras

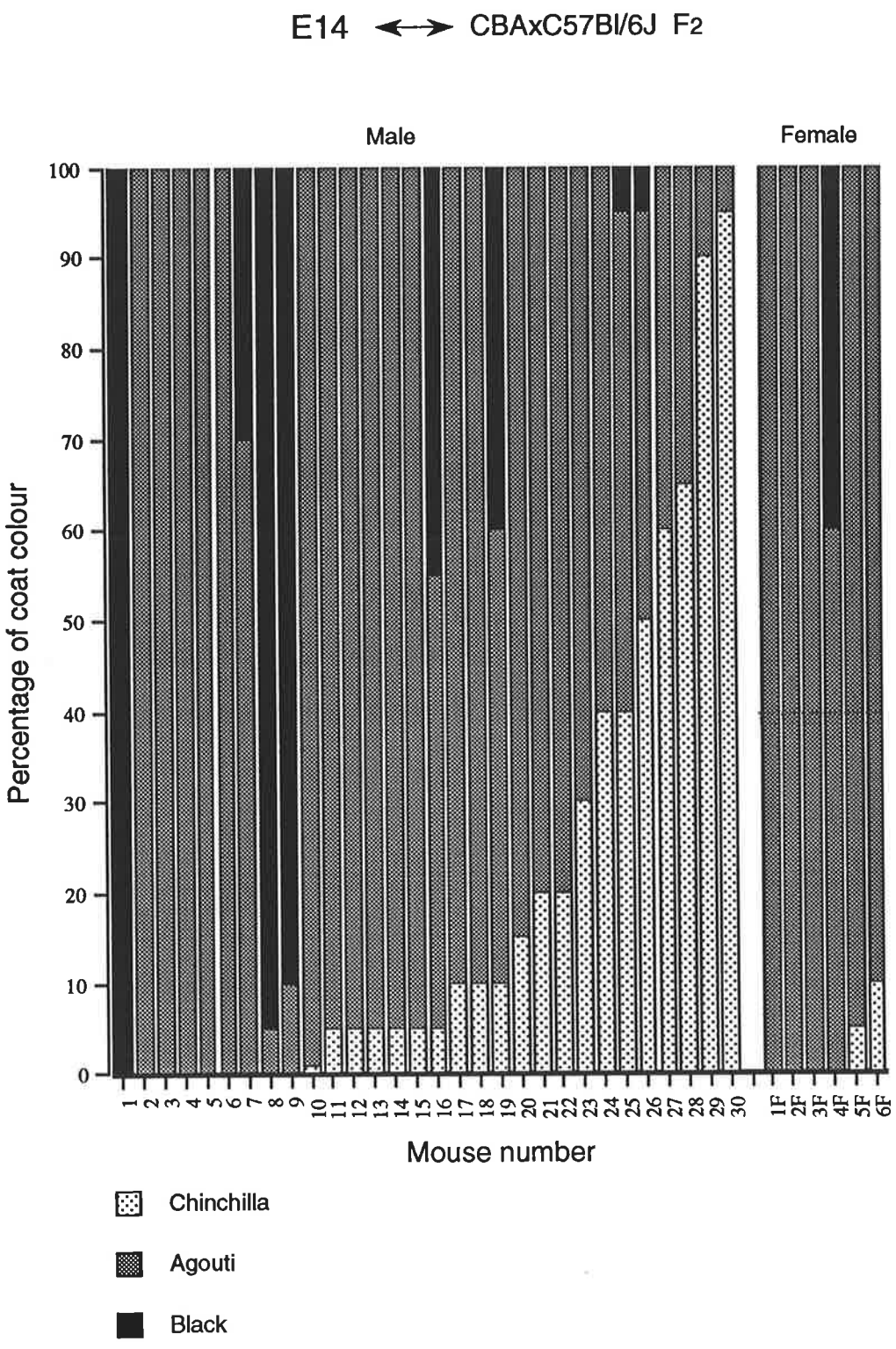
A E14 ⟷ CBAxC57Bl/6J F2

B ES-193 ⟷ CBAxC57Bl/6J F2

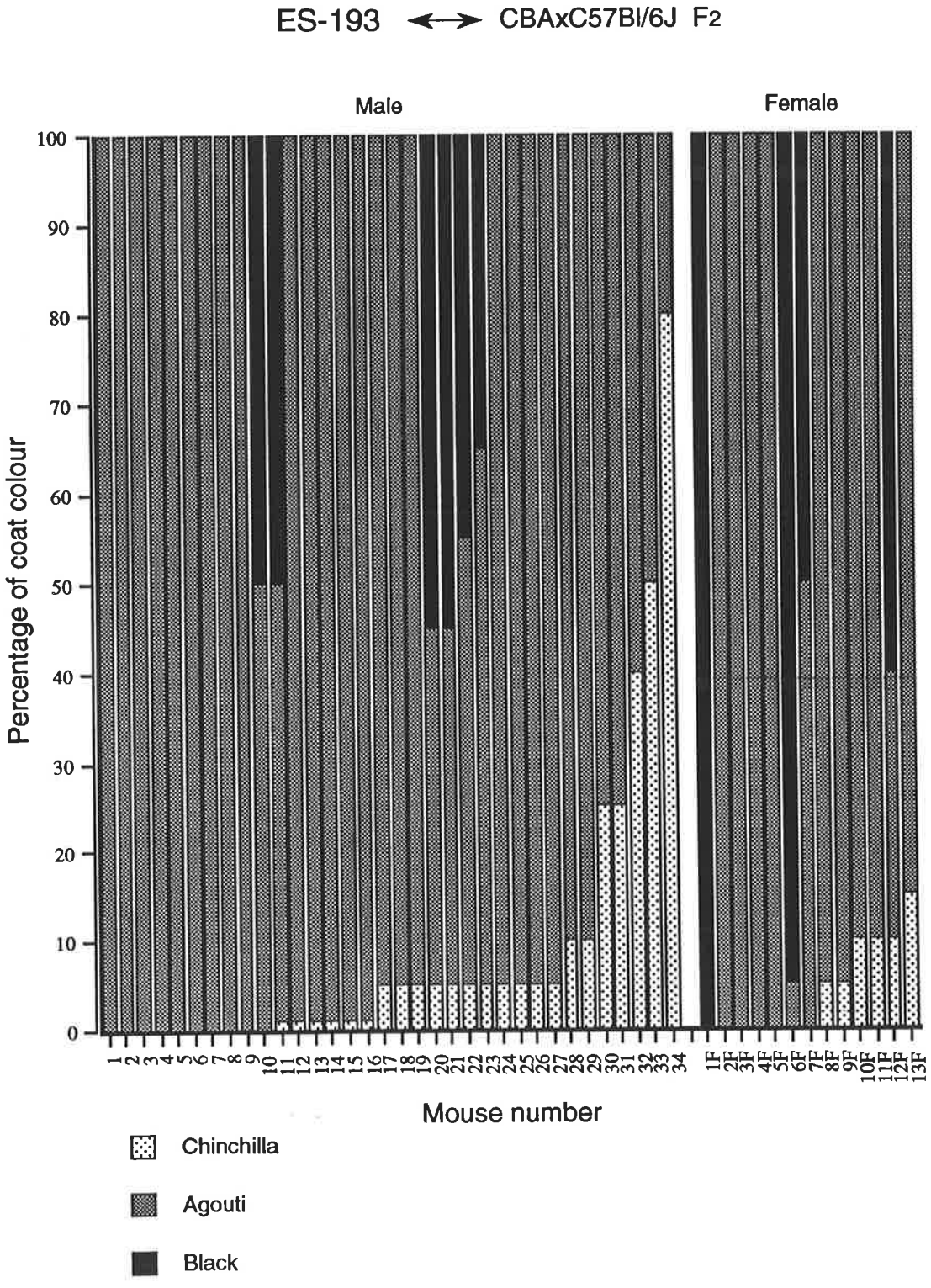
C ES-325 ⟷ CBAxC57Bl/6J F2

The F2's are agouti or black (3:1). Black is from the host blastocyst and chinchilla is from E14 while agouti may be from host or from the combination of black and E14 ES cell.

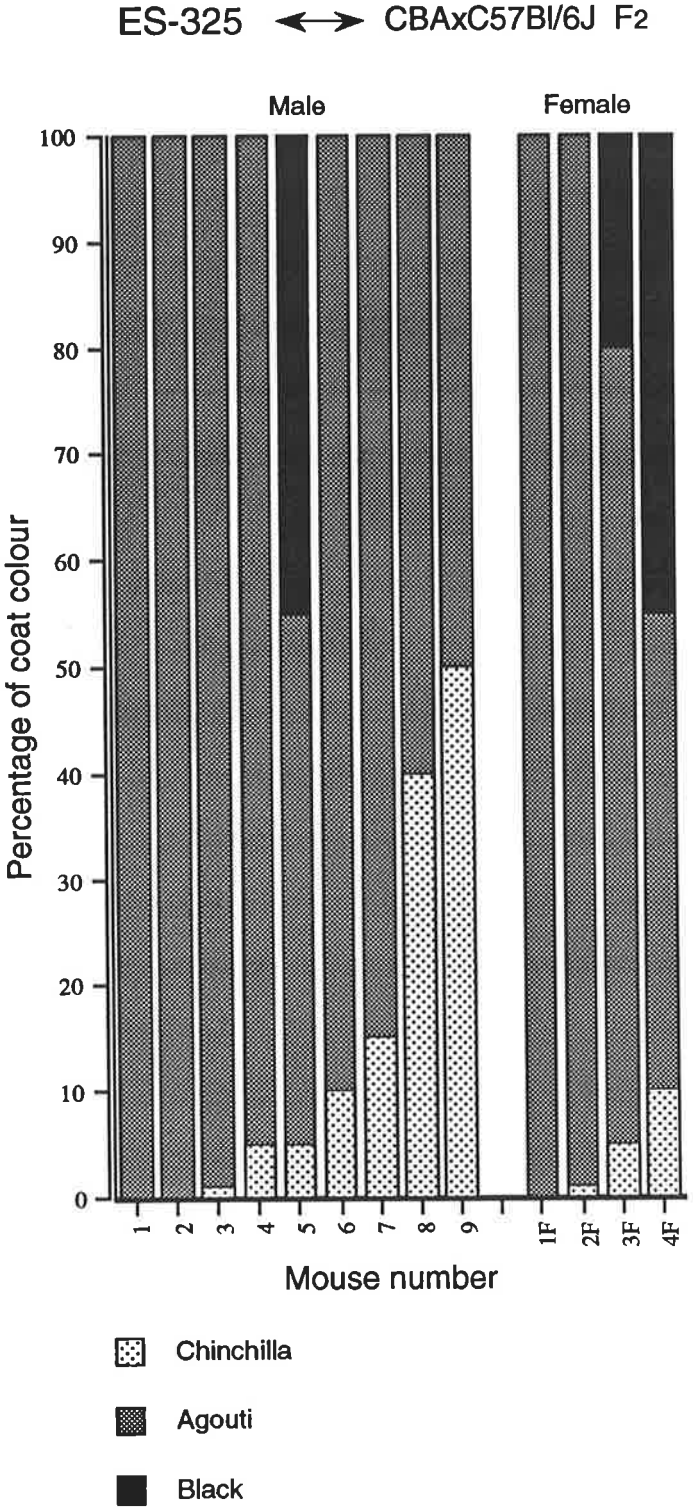
A


E14 ⟷ CBAxC57Bl/6J F2
Male
Female
Percentage of coat colour
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 1F 2F 3F 4F 5F 6F
Mouse number
Chinchilla
Agouti
Black

B


ES-193 ⟷ CBAxC57Bl/6J F2
Male
Female
Percentage of coat colour
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
1F 2F 3F 4F 5F 6F 7F 8F 9F 10F 11F 12F 13F
Mouse number
Chinchilla
Agouti
Black

C


ES-325 ⟷ CBAxC57Bl/6J F2
Male
Female
Percentage of coat colour
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
1F
2F
3F
4F
Mouse number
Chinchilla
Agouti
Black

colour, the proportion of chinchilla within individual animals is higher in chimaeras obtained from injection of unmanipulated E14 ES cells. This indicated that the unmanipulated E14 ES cells gave higher contribution to the skin of chimaeras than did the two targeted-ES cells.

As expected, the sex ratio of the chimaeras is distorted toward male because the E14 cells (XY) are able to cause some XX blastocysts to develop into males.

(ii) Germ line chimaera testing

Male chimaeras were mated to BALB/c females (*AA bb cc*); white coat. The offspring from this cross had either agouti or grey coat colour. Pups with grey coat colour ($A^{w}A$ *bb* $c^{ch}c$) indicated that E14 ES cells contribute to the germ line.

Eight male chimaeras derived from ES-325 ES cells were tested for ES contribution to the germ line. One of the chimaeras (mouse number 8 in Fig. 5-7 C) gave 100 % germ line transmission (n = 29). The other two chimaeras, mouse number 4 and 6 in Fig. 5-7 C, gave approximately 10 % germ line transmission (n = 4/30 and 5/35, respectively).

The parental E14 ES cells could transmit the ES cell genome through the germ line. Two out of four fertile chimaeras tested transmitted the ES genome to their offspring (data not shown).

Thirty-four male chimaeras derived from ES-193 clones were mated to BALB/c females. Three were sterile. At least 9 pups (average 39) were generated from each fertile chimaera and none of these pups had the grey coat colour, indicating E14 contribution to the germ line. This suggested that the ES-193 cell clone had lost the capacity to contribute to the germ line.

THE DIFFERENCE IN THE ABILITY TO CONTRIBUTE TO THE GERM LINE OF ES CELLS IS COMMON IN MANIPULATION OF ES CELLS. THE MECHANISM IS UNKNOWN.

(iii) Detection of the disrupted H2A.Z gene in F1 mice

DNA from the first seven offspring of mouse number 8 in Fig. 5-7 C

designated 325-A, 325-B, 325-C, 325-D, 325-E, 325-F, 325-G, were analysed by Southern blot to identify mice with the mutated H2A.Z allele (H2A.Z$^-$ allele). Genomic DNA was isolated from tail biopsy as described (2.3.3. ii). DNA was digested with *Xba* I and the filter was probed with a *Xho* I/ *Hind* III fragment (probe C, Fig. 4-10). The functional H2A.Z allele (H2A.Z$^+$) and the mutated allele were detected as 3.6 and 4.7 kb fragments, respectively. The result showed that five of seven offspring were heterozygous for a mutated H2A.Z locus (lane 2, 4 and 6-8 in Fig. 5-8). Among these mice only 325-A is female. Two of the heterozygous H2A.Z$^+$/H2A.Z$^-$ mice was mated to test whether H2A.Z deficiency is lethal.

A breeding program was established to generate sufficient mice to study the phenotype induced by targeted inactivation of the H2A.Z gene.

5.3 Discussion

5.3.1 Test for electroporation conditions resulting in single copy integration of introduced DNA

Results presented in Table 5-1 show that the conditions which gave the highest percentage of single copy integration involved the use of 5 μg of DNA with PBS as the electroporation medium with a pulse of 25 μF, 350 volts. This suggested that both DNA concentration and electroporation conditions were factors contributing to the copy number of integrated DNA. Using PBS as an electroporation medium, and 25 μF at 350 volts, may result in less cell permeability compared to the other two protocols described above, thereby resulting in less DNA introduced into the cells. Using lower amounts of DNA also appears to result in lower rates of multiple copy integration. Although this result was obtained from only one experiment, it was later confirmed by the

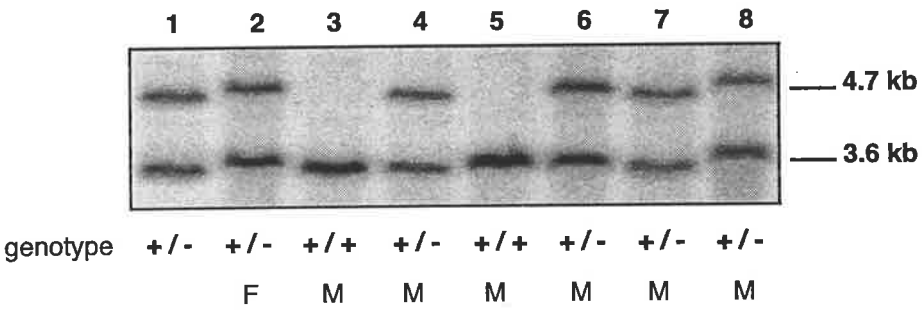
Figure 5-8

Southern analysis of seven offspring derived from the cross between chimaeric mouse number 8 (Fig. 5-7 C) and BALB/c female

Tail DNA from seven offspring, designated 325-A, 325-B, 325-C, 325-D, 325-E, 325-F and 325-G was isolated. DNA of ES-325 cells from which the chimaera was derived was also used in this analysis as a positive control. The digested DNA was electrophoresed on a 0.8% agarose gel along with DNA size markers, and transferred onto Zeta-Probe membrane. The filter was probed with a *Xho* I/ *Hin*d III fragment (probe C in Fig. 4-10). The functional H2A.Z allele and the mutated-allele were detected as 3.6 and 4.7 kb hybridising fragments respectively. Lanes are as follows:

1) ES-325 cells DNA
2) 325-A DNA
3) 325-B DNA
4) 325-C DNA
5) 325-D DNA
6) 325-E DNA
7) 325-F DNA
8) 325-G DNA

Genotype of mice at the H2A.Z locus are indicated: +/+ for wild type and +/- for heterozygous. Five heterozygous for H2A.Z locus are obtained. Of these 4 are male (M) and only one is female (F).


1
2
3
4
5
6
7
8
4.7 kb
3.6 kb
genotype
+/-
+/-
+/+
+/-
+/+
+/-
+/-
+/-
F
M
M
M
M
M
M

other electroporation performed using these conditions, i.e., all of the pHZHR4-targeted ES cells had only a single copy of integration.

5.3.2 An attempt to establish homozygous H2A.Z " knock-out " ES cells

Two independent ES cell clones were used in an attempt to generate ES cells which lack a functional H2A.Z gene. The results obtained from these two clones are consistent in that homozygous H2A.Z gene "knock-out" ES cell clones were not found in any of the analysed clones. This result was not due to low targeting efficiency since each experiment showed that homologous recombination did occur at a frequency of at least 1 in 30 of analysed transformed clones.

The clones which appeared to have two alleles disrupted always had an extra copy of a H2A.Z gene. This suggests that a functional H2A.Z gene is essential for cell viability. It should be noted here that the presence of the H2A.Z gene was demonstrated by Southern analysis, without demonstrating the presence of functional H2A.Z mRNA. The third H2A.Z allele may be due to either a type III recombination /"pick-up" event or to a duplication event (by an unknown mechanism). This supports the idea that functional H2A.Z is required for cell growth and proliferation.

5.3.3 Generation of H2A.Z deficient mice

E14 ES cells retain the capacity to contribute to the germline after maintenance in our laboratory. Therefore, it was possible that ES cells which have been subjected to electroporation and selection might still be germ line competent. However, it is also possible that electroporation and other manipulations may result in loss of this capability. It has been shown that the manipulated MBL-5 ES cell line did not give germ line transmission in contrast to unmanipulated cells (L. Williams, pers. comm.).

Two independent H2A.Z-targeted ES cell lines, ES-193 and ES-325, were used to generate chimaeras. ES-325 contributed efficiently to the germ-line of chimaeras whereas the other clone, ES-193 appeared to have lost this capacity. Both clones were chosen on the basis of good morphology and normal growth characteristics. One of chimaeras derived from ES-193 had a high degree of coat colour chimaerism (mouse number 34 in Fig. 5-7 B), indicated by the high percentage of chinchilla coat colour. Therefore, it was predicted that the ES-193 clone would have a high probability for germ line transmission. In contrast the chimaera (mouse number 8 in Fig. 5-7 C) derived from ES-325 which had a lesser degree of coat colour chimaerism gave 100% germ line transmission. From this experiment it is clear that a high percentage of chinchilla coat colour does not necessarily predict germ line contribution. Several factors may contribute to the loss of this capability such as the timing of ES cells in culture, the passage number of the ES cells and culture conditions.

Mice heterozygous for a mutated H2A.Z mice ($H2A.Z^{+}/H2A.Z^{-}$) were obtained and cross-bred to generate H2A.Z deficient ($H2A.Z^{-}/H2A.Z^{-}$) mice. Preliminary results indicated that the absence of a functional H2A.Z gene may lead to embryonic lethality. Twenty-five offspring from this cross-breeding have been analysed and none are $H2A.Z^{-}/H2A.Z^{-}$.

This result seems to agree with a requirement for functional H2A.Z protein for cell viability (described above).

In summary, the results presented in this chapter indicate an essential role for H2A.Z in both ES cells and in mice.

Chapter 6

Final Discussion

6.1 Conclusions and Discussion

6.1.1 Isolation and characterisation of the mouse histone H2A.Z gene

Mouse histone H2A.Z was initially isolated and characterised as described in chapter 3. Histone H2A.Z is highly conserved throughout evolution, indicating an essential function for this protein. To date histone H2A.Z genes have been isolated from several organisms. Histone H2A.Z genes are different from the major histone genes in that they exist as a single copy gene per haploid genome, contain introns, and encode polyadenylated mRNAs. The expression of H2A.Z genes is not tissue-specific and occurs throughout the cell cycle.

The mouse histone H2A.Z gene, isolated and characterised in the present work, contains four introns and exists as a single copy per haploid mouse genome with several pseudogenes. The presence of H2A.Z pseudogenes was also found in the human genome (Hatch and Bonner, 1990). One of the mouse pseudogenes contains sequence equivalent to mouse histone H2A.Z cDNA, and was used in this work to produce H2A.Z protein in *Escherichia coli*, using a T7 expression system. This protein was used as an antigen for monoclonal antibody production, and in purification of H2A.Z polyclonal antibodies. Expression of histone H2A.Z in mouse ES cells was demonstrated by Northern analysis.

6.1.2 Histone H2A.Z gene targeting in mouse embryonic stem cells

Gene targeting by homologous recombination in mouse ES cells has been widely used to investigate the function of many genes. This approach was used here to study the function of H2A.Z by attempting to generate ES cells and animals deficient in a functional H2A.Z gene. To generate ES cells in which one allele of the H2A.Z gene was disrupted, several targeting vectors derived from both nonisogenic (derived from BALB/c strain) and isogenic DNA (derived from 129/Sv strain) were constructed. These vectors were replacement type vectors,

and targeting utilised only a positive selection strategy. The successful targeting vectors were derived from isogenic DNA and contained a long stretch of DNA (approximately 17-18 kb) homologous to the H2A.Z chromosomal locus. The best targeting efficiency described in this thesis is 1 in 7 of the analysed clones. This result supports the finding by other investigators that utilising isogenic DNA is important for high targeting efficiencies. Targeted ES cell clones were obtained and then used in attempts to generate ES cells and germ line chimaeras which lack a functional H2A.Z gene.

To generate ES cell lines which lack a functional H2A.Z gene, targeted ES cells were subjected to a second round of gene targeting as described by te Riele *et al.*, (1990). ES cells in which both alleles were disrupted were not found amongst the 575 analysed clones, from two separate experiments, suggesting that histone H2A.Z may be required for cell viability. The clones which appeared to have both alleles disrupted always had three alleles of the H2A.Z gene. The third allele may have been generated by a form of type III recombination (Adair *et al.*, 1989).

To attempt to generate mice deficient for histone H2A.Z, targeted ES cells were used to generate germ line chimaeras. Mice which are heterozygous for the disrupted H2A.Z gene were obtained and a breeding program has been set up to attempt to obtain mice (or embryos) which are homozygous for the disrupted H2A.Z gene. Preliminary results indicate that deficiency of a functional H2A.Z gene may lead to embryonic lethality, since mice homozygous for the H2A.Z gene were not found in any of the twenty-five pups born and analysed so far.* Heterozygous mice were obtained from breeding male chimaeras and normal females, suggesting that sperm carrying the disrupted H2A.Z gene are viable. To test whether eggs carrying the disrupted H2A.Z gene are viable, crosses between female mice which are heterozygous for the disrupted H2A.Z gene and normal males were set up. Mice which are heterozygous for the disrupted H2A.Z gene

* STATISTICAL TREATMENT DESIRABLE FOR PUBLICATION.

were obtained, indicating that eggs lacking the H2A.Z gene are also viable. Therefore, the results suggest that the lack of the live-born homozygotes from the crosses between the heterozygous mice was not due to the inability to produce viable gametes. Rather they indicate that H2A.Z is required during embryogenesis. Experiments are now underway in this laboratory to determine the stage at which death occurs during embryonic development.

In summary, ES cells and animals which are homozygous for the disrupted H2A.Z gene were not obtained. These results indicate for the first time that histone H2A.Z has an essential role in mammals.

6.2 What is the function of histone H2A.Z?

There are an increasing number of studies indicating a functional role of histones in transcription. However, little is known about the function of histone variants. There has been much speculation that the presence of histone variants could generate nucleosome heterogeneity for distinct functions and have significant biological consequences.

As previously mentioned, histone H2A.Z is highly conserved throughout evolution and has been found to be essential in *Tetrahymena* (Liu and Gorovsky, unpublished observation, cited in Thatcher and Gorovsky, 1994) and in *Drosophila* (van Daal and Elgin, 1992). Studies described above indicate an essential role for H2A.Z in mammalian development.

The biological role of histone H2A.Z is still unclear. Some work indicates that histone H2A.Z is associated with active chromatin (Allis *et al*, 1980). Recent studies suggest that histone H2A.Z may have a role in establishing an active chromatin structure rather than appear as a consequence of active transcription (Stargell *et al.*, 1993). The question of how H2A.Z facilitates gene transcription still requires further investigation.

Histone H2A.Z is different from histone H2A in several regions, including

the N-terminal and C-terminal regions. These regions are known to be involved in DNA interaction. Therefore it can be predicted that histone H2A.Z may interact with DNA in a different way from normal H2A. Furthermore, H2A.Z may interact differently with other histones in the nucleosome.

The structure of the histone octamer has been solved at 3 Å resolution (Arents *et al.*, 1991) and the topology of the histone octamer surface used in docking the nucleosome DNA has also been investigated (Arents *et al.*, 1993). It may be possible to use computer modelling to predict the structural and functional consequences of incorporating H2A.Z in nucleosomes. Also, this laboratory is collaborating with the laboratory of Dr. V. Ramakrishnan of Brookhaven National Laboratory, Upton, New York. They are proposing to solve the 3-dimensional structure of H2A.Z-containing nucleosomes. This can be achieved by utilising the recombinant mouse histone H2A.Z to form nucleosomes *in vitro*.

6.3 Future work

The work presented in this thesis indicates an essential role for histone H2A.Z in mammalian development. Experiments can now be conducted in order to further investigate the function of this protein.

Current work in this laboratory is designed to investigate the stage at which death occurs during embryonic development of mice lacking a functional H2A.Z gene. Two approaches are being taken: (1) Obtaining embryos at different stages from a heterozygous cross; (2) Obtaining fertilised eggs from a heterozygote cross and attempting to culture these to different developmental stages *in vitro*. PCR will be used to identify the genotypes of the embryos.

In addition a monoclonal antibody specific for histone H2A.Z will be used to determine the abundance and distribution of H2A.Z in wild-type and mutant embryos, by immunohistochemistry.

Some studies indicate that H2A.Z is associated with active chromatin (Allis *et al*, 1980). One experiment which would test this association of H2A.Z with active chromatin is the following: Chromatin from particular cell types can be digested with micrococcal nuclease to prepare mononucleosomes (Simpson, 1978). H2A.Z-specific antibodies can be used to isolate the H2A.Z-containing nucleosomes by immunoprecipitation. The DNA from unbound and antibody-bound fractions can then be probed for gene sequences that are active or inactive in the cells. If histone H2A.Z is enriched in active chromatin, DNA that can hybridise to the probe representing an active gene will be found in antibody-bound fractions. This method has been previously described by Hebbes *et al.*, (1994) and used successfully to study the preferential distribution of acetylated histone H4.

Appendix

Towards Histone H1^{o} Gene "Knock-Out" in Mouse Embryonic Stem Cells

A.1 Introduction

The overall aim of this work was to investigate the possible role of histone H1° in mammals. Histone H1° is a variant of histone H1. It is similar to histone H5, (found in nucleated erythrocytes of birds) with respect to sequence homology (Pehrson and Cole, 1981), role in chromatin condensation and accumulation in terminally differentiated cells (Panyim and Chalkley, 1969; Gjerset *et al.*, 1982). However, H1° is found in several tissues, and expression of H1° can be regulated by hormones (Gjerset *et al.*, 1982; García-Segura *et al.*, 1993). It has been postulated that H1° and H5 may have a similar function. The studies in this laboratory have focused on several aspects of chicken histone H5, and the chicken histone H5 gene has been isolated in this laboratory (Krieg *et al.*, 1983).

The biological functions of histone H1° have not yet been definitively demonstrated. Several studies have suggested that histone H1° may have an important role in differentiation. As mentioned above, histone H1° has been found to increase in terminally differentiated tissues and is associated with non-transcribed chromatin (Gjerset *et al.*, 1982; Roche *et al.*, 1985; Mendelson *et al.*, 1986). H1° expression is increased in response to induction by differentiation inducing agents (Alonso *et al.*, 1988; Rousseau *et al.*, 1991).

To investigate the role of the mouse histone H1°, gene targeting by homologous recombination in ES cells was initiated during the course of these studies. The aim was to generate mice lacking a functional H1° gene. The H1° gene exists as a single copy per haploid mouse genome (Alonso *et al.*, 1988) and a mouse H1° genomic clone had been previously isolated and characterised (Breuer *et al.*, 1989).

A targeting vector, designated pH1°HR1, used in this experiment was derived from an H1° genomic clone which was previously isolated and characterised by Breuer *et al.*, (1989). The mouse strain from which the H1° genomic clone was derived was unknown. This vector was constructed to enrich

for cells containing the targeted disruption by using a positive/negative selection strategy as described by Mansour *et al.*, (1988). In addition, the vector was designed with the initial intention to use PCR as a screening method for targeted H1° ES cells. However, Southern analysis was later used as a screening method.

The work presented here describes an attempt to obtain ES cells in which one allele of the H1° gene has been disrupted by homologous recombination. pH1°HR1 DNA was introduced into E14 ES cells by electroporation and Southern analysis was used to screen for the targeted-ES cell clones. In order to improve the targeting efficiency the H1° gene was isolated from a 129 mouse genomic DNA library, in order to construct an isogenic targeting vector. This targeting vector, pH1°HR2, was constructed, but has not yet been used in a targeting experiment. The strategy for detection of homologous recombination with pH1°HR2 is described.

A.2 Results

A.2.1 Electroporation of E14 ES cells with targeting vector pH1°HR1

(i) Construction of pH1°HR1

The targeting vector pH1°HR1 was constructed from the recombinant plasmid named pH1° (2.2.13), containing a mouse H1° genomic clone. A 9.0 kb *Eco* RI fragment within pH1° was subcloned for further analysis. The maps of the subclones, relative to the 9.0 kb *Eco* RI fragment, are presented in Fig. A-1. Three subclones, named pMH1°.1, pMH1°.2 and pMH1°.3, were generated and used to create the targeting vector. pMH1°.1, pMH1°.2 and pMH1°.3 were generated by subcloning a 2.9 kb *Eco* RI/*Xho* I fragment, a 2.5 kb *Xho* I/*Sal* I fragment and a 3.6 kb *Sal* I/*Eco* RI fragment into *Eco* RI/*Xho* I-digested pBluescript KS+ vector, *Xho* I/*Sal* I-digested pBluescript KS+ vector and

Sal I/*Eco* RI-digested pBluescript KS+ vector, respectively.

pH1^{o}HR1 contains 7.6 kb of DNA homologous to the H1^{o} locus. As mentioned above, this vector was designed for enrichment of the clones in which homologous recombination events occurred by using a positive/negative selection strategy. Therefore, it was designed to contain both the neomycin phosphotransferase gene (*neo^{r}*) and Herpes simplex virus-thymidine kinase (HSV-*tk*) gene. The purposes of using the *neo^{r}* gene in this vector was to disrupt the H1^{o} gene and to act as a selectable marker by conferring resistance to the antibiotic drug G418. The HSV-*tk* gene was used in order to select against ES cells in which the targeting vector DNA was randomly integrated into the genome. Cells which incorporate the functional HSV-*tk* gene are sensitive to the nucleotide analogue FIAU and will be eliminated during selection. In addition, pH1^{o}HR1 was designed to include a 9 bp insertion into the targeting vector, 1.9 kb upstream from the H1^{o} gene, for use as a unique primer binding site for PCR detection.

To create pH1^{o}HR1, seven recombinant plasmids named pMH1^{o}.2RV, pMH1^{o}.2neo, pMH1^{o}neo, pMH1^{o}neo/tk, pHP1.3, pPH0.7 and pMH1^{o}.1mut were generated as precursors. The restriction maps of pMH1^{o}.2RV and pMH1^{o}.2neo, pMH1^{o}neo and pMH1^{o}neo/tk, are presented in Fig. A-2 and Fig. A-3, together with the strategies used to create these plasmids. The restriction map and the strategies used for construction of pHP1.3 and pPH0.7, and pMH1^{o}.1mut are presented in Fig. A-4 and Fig. A-5, respectively.

The strategy used for constructing pH1^{o}HR1 from the precursors is shown in Fig. A-6. pH1^{o}HR1 contained the *neo^{r}* gene transcribed in the opposite direction to H1^{o} and the HSV-*tk* gene transcribed in the same direction as H1^{o}. H1^{o} is rendered nonfunctional by the disruption of the coding sequence by the *neo^{r}* gene.

Figure A-1

Restriction map of an H1^{o} genomic fragment and subclones

Restriction map of a 9.0 kb *Eco* RI fragment, containing a mouse H1^{o} genomic clone, in pH1^{o}(2.2.13) is shown at the top. The numbers above the map indicate the distances in kilobases from the *Eco* RI site located at the 5' end. The location of the H1^{o} gene within this fragment is shown in the black box, and the arrow indicates the orientation of the gene. The open boxes below the map represent the insert fragments subcloned into plasmids, and the names of the recombinant plasmids are indicated.

pMH1^{o}.1 contains a 2.9 kb *Eco* RI/*Xho* I fragment subcloned into a *Eco* RI/*Xho* I-digested pBluescript KS+ vector. pMH1^{o}.2 contains a 2.5 kb *Xho* I/*Sal* I fragment subcloned into a *Xho* I/*Sal* I-digested pBluescript KS+ vector. pMH1^{o}.3 contains a 3.6 kb *Sal* I/*Eco* RI fragment subcloned into a *Sal* I/*Eco* RI-digested pBluescript KS+ vector.

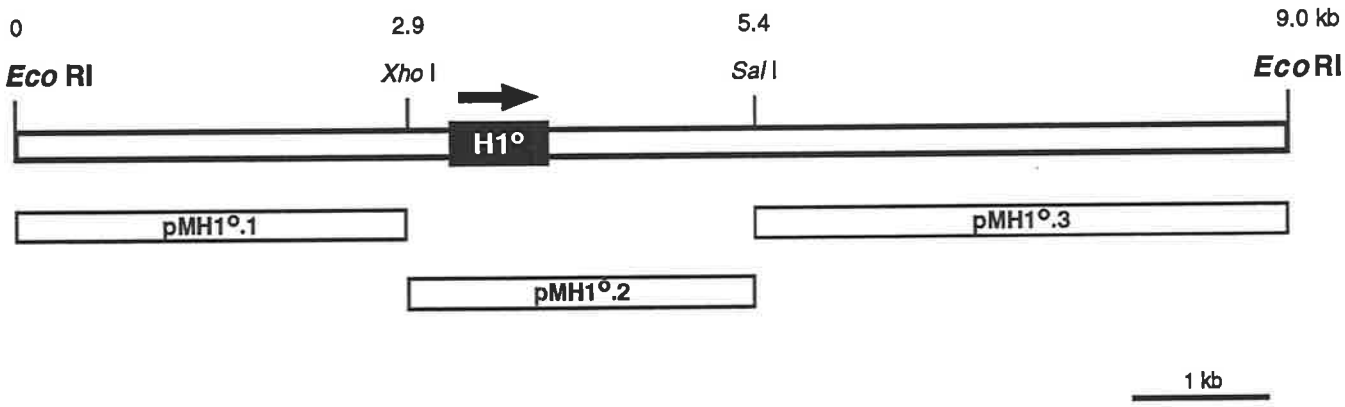

0
2.9
5.4
9.0 kb
Eco RI
Xho I
Sal I
Eco RI
H1°
pMH1°.1
pMH1°.3
pMH1°.2
1 kb

Figure A-2

Restriction maps of pMH1°.2RV and pMH1°.2neo

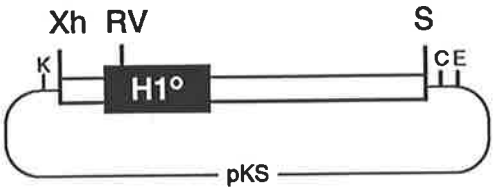
The restriction maps of pMH1°.2RV and pMH1°.2neo are shown in **A** and **B**, respectively. The H1° genomic fragment is shown as an open box with the coding sequence filled. The restriction sites within the insert fragment and polylinkers are abbreviated as indicated. The restriction sites which were lost as a result of the ligation reaction are shown in brackets. The *neo*^{r} cassette is shown with the arrow above indicating the direction of transcription.

pMH1°.2RV was derived from pMH1°.2 (Fig. A-1). It was modified from pMH1°.2 by creating an *Eco* RV restriction site within the coding sequence of the H1° gene at a position 103 bp from the ATG initiation site (Alonso *et al.*, 1988). The *Eco* RV restriction site was generated by mutagenesis and was then used for insertion of a *neo*^{r} gene into the H1° gene, in order to inactivate the H1° gene.

To create pMH1°.2RV, a 2.5 kb *Kpn* I/*Sal* I fragment was initially subcloned into M13mp18 as described (2.3.13) and the recombinant M13 was obtained. The single-stranded M13 DNA was prepared from this phage (2.3.13) and used in a mutagenesis reaction (2.3.14) in the presence of a 25 base oligomer named H1°/RV (2.2.15). This oligomer contains bases homologous to the H1° coding sequence at a position 92-116 bp from the ATG initiation site, except at position 104 (C to A). The replicative form (RF) DNA was prepared from the positive plaque and digested with *Eco* RV, *Eco* RV/*Sal* I, and *Kpn* I/*Eco* RV to verify the correct clone. pMH1°.2RV was obtained after subcloning the 2.5 kb *Kpn* I/*Sal* I fragment isolated from RF DNA of the correct clone, into *Kpn* I/*Sal* I-digested pBluescript KS+ vector.

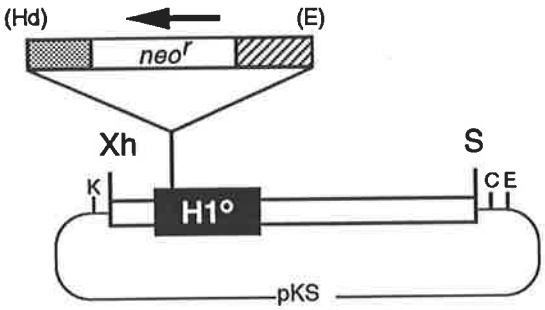
pMH1°.2neo was generated by subcloning a 1.8 kb *Eco* RI/*Hin*d III pgkNEOpA (2.2.13) fragment into a *Eco* RV-digested pMH1°.2RV vector. It contains the *neo*^{r} gene inserted in the opposite transcriptional direction to the H1° gene.

A



pMH1°.2RV

B



pMH1°.2neo

1 kb

C	*Cla* I	E	*Eco* RI
Hd	*Hind* III	K	*Kpn* I
RV	*Eco* RV	S	*Sal* I
Xh	*Xho* I		
▨	pgk-promoter	▩	pgk-polyA

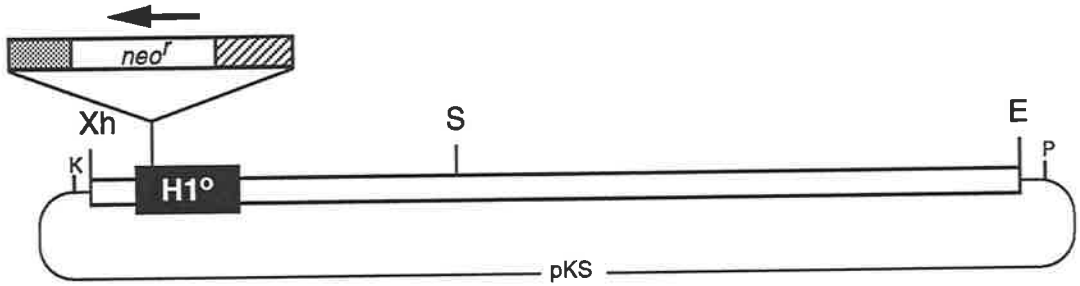
Figure A-3

Restriction maps of pMH1^{o}neo and pMH1^{o}neo/tk

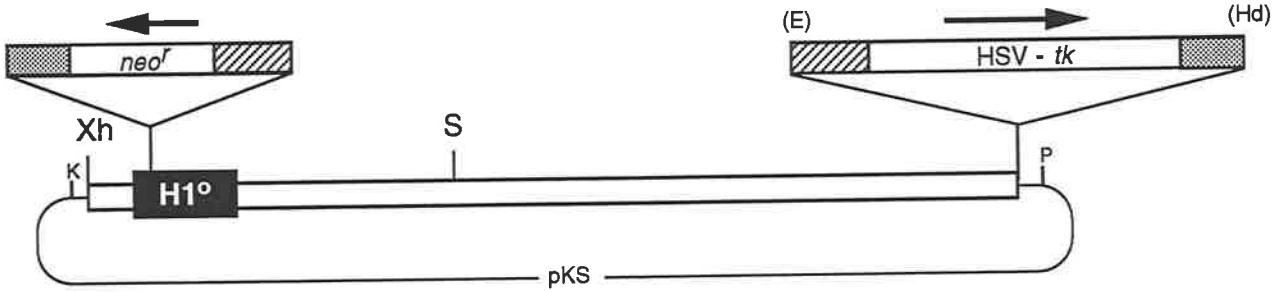
The restriction maps of pMH1^{o}neo and pMH1^{o}neo/tk are shown. The *neo^{r}* and the HSV-*tk* cassettes are shown with the arrows above indicating the directions of transcription. The restriction sites within the insert fragment and polylinkers are abbreviated as indicated. The restriction sites which were lost as a result of the ligation reaction are shown in brackets.

pMH1^{o}neo was generated by subcloning a 3.6 kb fragment isolated from pMH1^{o}.3 (see Fig. A-1) into a *Sal* I/*Eco* RI-digested pMH1^{o}.2neo (see Fig. A-2).

pMH1^{o}neo/tk was constructed by subcloning a 2.85 kb *Eco* RI/*Hin*d III (end-filled) fragment isolated from pgkTKpA (2.2.13) into the *Eco* RI-digested (end-filled) pMH1^{o}neo vector.



pMH1°neo



pMH1°neo/tk

1 kb

E	*Eco* RI	Hd	*Hind* III	K	*Kpn* I
K	*Kpn* I	P	*Pst* I	S	*Sal* I
Xh	*Xho* I		pgk-promoter		pgk-polyA

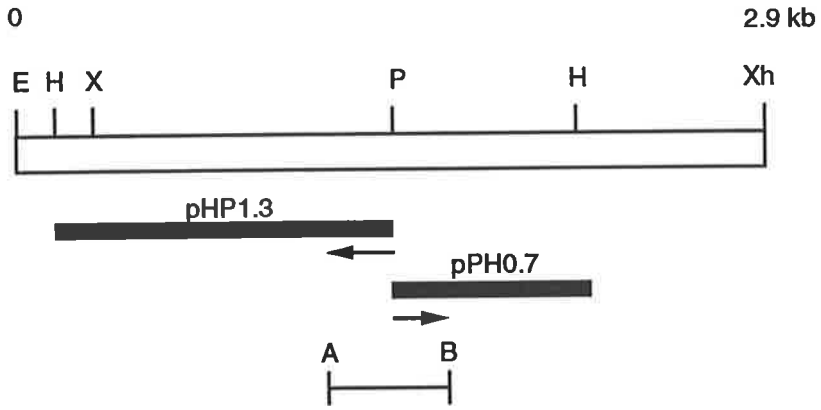
Figure A-4

Restriction maps of pHP1.3 and pPH0.7

Restriction maps of pHP1.3 and pPH0.7 (represented by thick lines) are shown relative to a 2.9 kb *Eco* RI/*Xho* I fragment within pMH1°.1 (Fig. A-1). The H1° genomic fragment is shown as an open box. The restriction sites are abbreviated as indicated.

pHP1.3 and pPH0.7 were generated in order to obtain the sequence surrounding the *Pst* I restriction site. This sequence was required for designing the oligomers for mutagenesis reactions to generate fragments containing a 9 bp insertion, and an *Xba* I restriction site, which were then used to generate pMH1°.1mut (Fig. A-5).

pHP1.3 and pPH0.7 were generated by subcloning a 1.3 kb *Hin*d III/*Pst* I fragment and a 0.7 kb *Pst* I/*Hin*d III fragment isolated from pMH1°.1 into a *Hin*d III/*Pst* I-digested pBluescript KS+ vector, respectively. Double-stranded DNA of pHP1.3 and pPH0.7 were used as templates for sequencing analysis (2.3.15) with USP primer. The direction of sequencing and the DNA sequence obtained from this analysis are indicated by arrows. 461 bp of DNA sequences was obtained (represented by the thin line spanning from A to B). The sequence is shown below with the *Pst* I site indicated. The boxes indicate the regions used to design the oligomers for mutagenesis reactions.




```
1    CCCTCATCTAAGCTGCTCCTTCTGGTGTGACAAACAGCCCCAGGCCTTAG    50

51   CAGAGGTTAGTTCATCTTGATCAGCTACAATTCCCCACTCCCAGGTTAAT   100

101  GTTTAATTAAACACTTTCGTGCCTTTTTGCTAGGGAATAAAAACACCAGG   150

151  GTCCGGGGATCCAGCTCAGGTAGGCAATACCACCAGAACCCTACGGAGCC   200

                                        PstI
201  GCTGCCAAGAGGTCCTAACATTGTGAGTAGGATGTCTGCAGCAGACAAGA   250

251  GGTTCTGGGCTCTCTGGGGGTCTCTTCGGCAAATCTCGGAGCCTCTAAAC   300

301  TCTCTCGGAGCCTAGGACTCCCAACAGCCCCAGGTCGGACTACAGATGGT   350

351  GGCGCCCCTTGCTCACTTGGTGGATTTTTGCCAAGCAGGGCGGGAAAGAT   400

401  GGATACATTTTAACAAAGACATGGCATCTTCCCAGCAAAATGGCGGTGGT   450

451  CGGGACCACGG                                          461
```

E	*Eco* RI	H	*Hin*d III
X	*Xba* I	Xh	*Xho* I
P	*Pst I*		

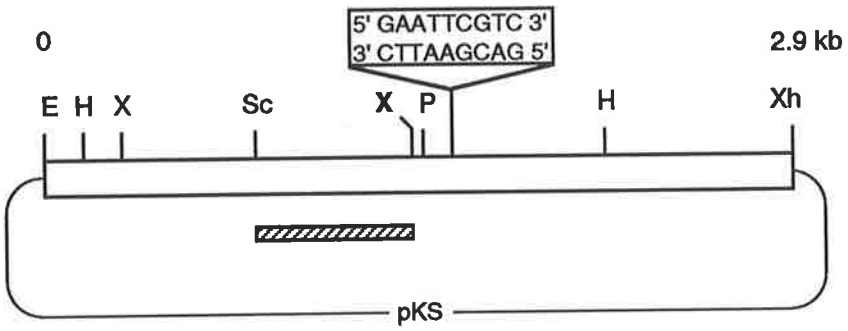
Figure A-5

Restriction map of pMH1°.1mut

Restriction map of pMH1°.1mut is shown. The H1° genomic fragment is shown as an open box. The restriction sites are abbreviated as indicated. The striped box represents a 0.6 kb *Sac* I/*Xba* I fragment which was used as a probe for detection of homologous recombination in ES cells.

pMH1°.1mut was modified from pMH1°.1 by insertion of an *Xba* I site and an extra 9 bp sequence, by mutagenesis.

To create pMH1°.1mut, a 1.3 kb *Hin*d III/*Pst* I fragment and a 0.7 kb *Pst* I/*Hin*d III fragment isolated from pMH1°.1 were initially subcloned into M13mp19 and M13mp18, respectively. Single-stranded DNA was isolated from these phages and used for mutagenesis reactions. The mutagenesis reactions were performed as described (2.3.14), using a 25 base oligomer named H1°/Xb (2.2.15) for generating an *Xba* I site within the 1.3 kb *Hin*d III/*Pst* I fragment, and a 47 base oligomer named H1°/9bp (2.2.15) for insertion of 9 bp (indicated in the box above the restriction map) within the 0.7 kb *Pst* I/*Hin*d III fragment. H1°/Xb and H1°/9bp contained bases homologous to DNA sequence positions 171-195 and 318-355 in Fig. A-4, respectively. Positive plaques were obtained and sequence analysis was used to verify the correct clones. Replicative form (RF) DNA was prepared from the correct clones. pMH1°.1mut was then generated by ligation of a 1.3 kb *Hin*d III/*Pst* I fragment containing the new *Xba* I site, a 0.7 kb *Pst* I/*Hin*d III fragment with the 9 bp insertion, and *Hin*d III-digested pMH1°.1 vector (see Fig. A-1).



pMH1°.1mut

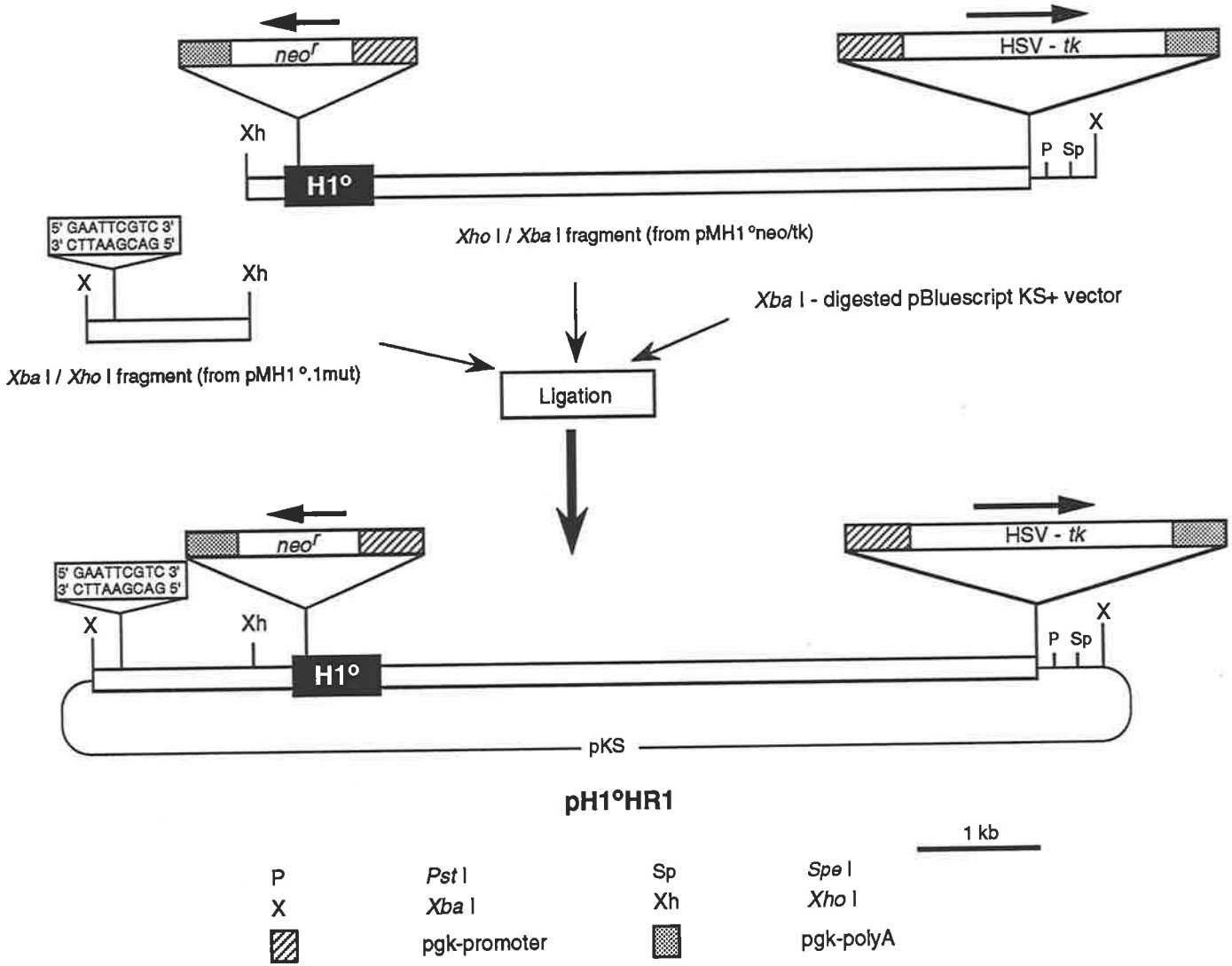
E	*Eco* RI	H	*Hin*d III
X	*Xba* I	Xh	*Xho* I
P	*Pst I*		

Figure A-6

Construction of the targeting vector pH1^{o}HR1

The strategy used for constructing the targeting vector pH1^{o}HR1 is shown. The H1^{o} genomic fragment is shown as an open box with coding sequences filled. The position of the *neo^{r}* and HSV-*tk* cassettes within the H1^{o} fragment are shown with the arrows above indicating the direction of transcription. The restriction sites within the insert fragment and polylinkers are abbreviated as indicated.

pH1^{o}HR1 (14.2 kb) was generated by ligation of a 1.5 kb *Xba* I/*Xho* I fragment from pMH1^{o}.1mut, a 10.8 kb *Xho* I/*Xba* I fragment from pMH1^{o}neo/tk and *Xba* I-digested pBluescript KS+ vector.


neo^r
HSV - tk
Xh
X
P Sp
H1°
5' GAATTCGTC 3'
3' CTTAAGCAG 5'
X
Xh
Xho I / Xba I fragment (from pMH1 °neo/tk)
Xba I / Xho I fragment (from pMH1 °.1mut)
Xba I - digested pBluescript KS+ vector
Ligation
neo^r
HSV - tk
5' GAATTCGTC 3'
3' CTTAAGCAG 5'
X
Xh
X
P Sp
H1°
pKS
pH1°HR1
1 kb
P
Pst I
Sp
Spe I
X
Xba I
Xh
Xho I
pgk-promoter
pgk-polyA

(ii) Electroporation of pH1^{o}HR1 into ES cells and Southern analysis

Prior to electroporation, pH1^{o}HR1 was digested with *Xba* I in order to release the insert, and processed as described (2.3.17). E14 ES cells passage 20 ($5x10^7$) were mixed with 20 µg of *Xba* I-digested pH1^{o}HR1 in 0.5 ml HBS and electroporated at 220 volts and 960 µF capacitance as described (2.6.3). ES cells were selected with FIAU (0.2 µM) plus G418 (200 µg/ml) for 11 days. To determine the enrichment by FIAU, ES cells were also selected with only G418 (200 µg/ml). The result showed that only two fold enrichment was observed. This result is consistent with results observed by other investigators in that the negative selection gave little enrichment. To date, results as high as the 2000 fold enrichment reported by Mansour *et al.*, (1988) have not yet been obtained by other investigators. Transformed-ES cell colonies which were resistant to both G418 and FIAU were expanded as described (2.6.4).

Genomic DNA was isolated from individual ES cell clones and analysed by Southern blot. DNA was digested with *Sac* I and probed with a 0.6 kb *Sac* I/*Xba* I fragment isolated from pMH1^{o}.1mut (see Fig. A-5 and A-7) which is an external probe not present in the vector. The endogenous H1^{o} allele was detected as a 4.4 kb hybridising fragment with this probe. If a homologous recombination event occurred, the H1^{o} disrupted allele was expected to give a 6.2 kb hybridising fragment, since *Sac* I was not present in the *neo^{r}* cassette, and therefore the size of the *Sac* I hybridising fragment was increased by the length of the *neo^{r}* cassette insertion (1.8 kb). No homologous recombination was observed in any of the 105 analysed clones.

A.3 Discussion

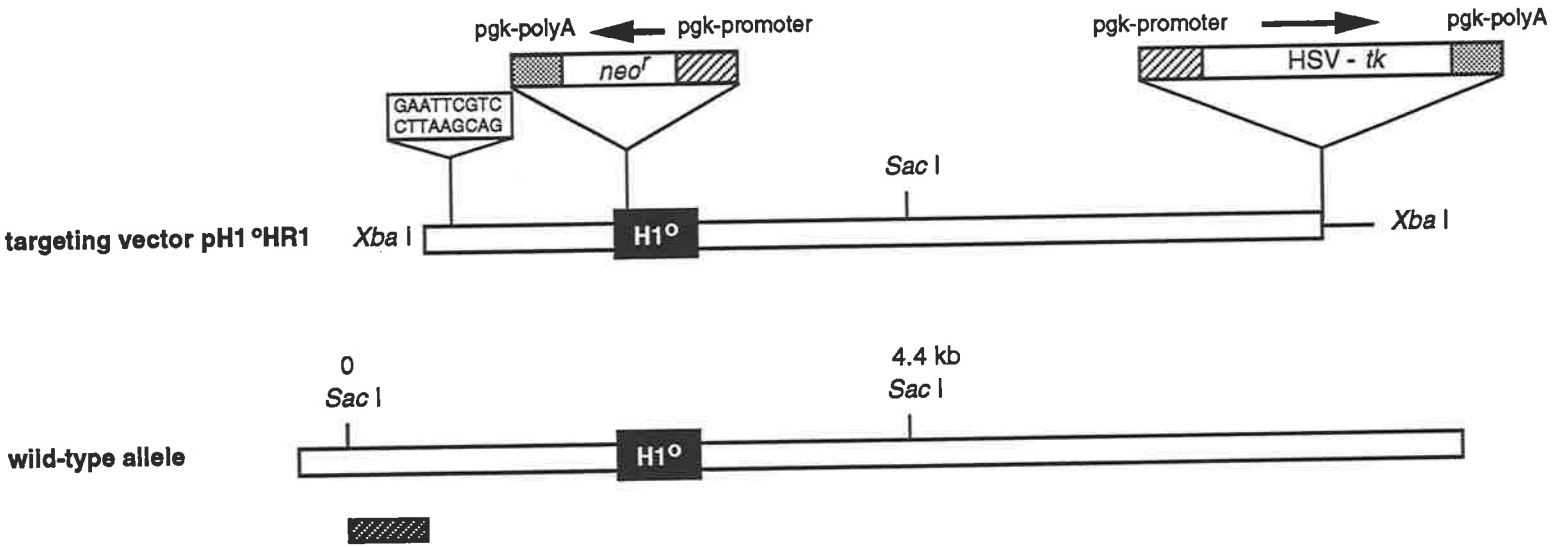
An attempt to obtain targeted H1^{o} ES cell clones using the positive/negative selection strategy described above was unsuccessful. This

Figure A-7

Southern analysis of ES cells transfected with pH1^{o}HR1 DNA

The targeting vector pH1^{o}HR1 is shown at the top in alignment with the restriction map of the H1^{o} locus. The H1^{o} coding sequence is represented by a black box. The arrows above the *neo*^{r} and HSV-*tk* cassettes indicate the direction of the transcription of these genes.

ES cell DNA was digested with *Sac* I and probed with a 0.6 kb *Sac* I/*Xba* I fragment (represented by the striped box) located 5'-external to the targeting sequence. This probe was isolated from pMH1^{o}.1mut (see Fig. A-5). The "wild-type" allele and the disrupted H1^{o} allele generated from an homologous recombination event, were expected to give a 4.4 kb and a 6.2 kb hybridising fragment, respectively.


pgk-polyA
pgk-promoter
neo^r
GAATTCGTC
CTTAAGCAG
pgk-promoter
pgk-polyA
HSV - tk
Sac I
targeting vector pH1 °HR1
Xba I
H1°
Xba I
0
Sac I
4.4 kb
Sac I
wild-type allele
H1°

result indicated that targeting efficiency at the H1^{o} locus using this vector is very low. Several factors may be responsible for this low targeting efficiency. Firstly, negative selection using FIAU gave very little enrichment, suggesting that the HSV-*tk* cassette may not be fully functional or that the concentration of FIAU used in this experiment was not high enough to eliminate cells which had the targeting vector randomly integrated in the genome. Secondly, it may be due to the low level of H1^{o} gene expression in ES cells. In post-implantation embryos, the expression of H1^{o} is very low up to day ten post coitum (García-Iglesias *et al.*, 1993), therefore it may be predicted that expression of the H1^{o} gene in ES cells is very low as well. It has been postulated that the variation in targeting efficiency at different gene loci may depend on the level of gene expression in ES cells and the chromatin environment surrounding the target gene. A gene expressed in ES cells is believed to have an open chromatin conformation which would facilitate access for recombination machinery. Finally, the vector may not be derived from isogenic DNA, and one end of the vector contained the nonhomology sequence of HSV-*tk* (2.85 kb), both of which may decrease the targeting efficiency.

In order to improve the targeting efficiency, a vector derived from isogenic DNA was made. The E14 ES cells are derived from the 129/Ola strain of mouse. Therefore, to isolate H1^{o} isogenic DNA, mouse genomic library III (2.2.7) was screened, as described in the following section. Clones containing the H1^{o} gene were isolated and used for constructing targeting vector pH1^{o}HR2.

A.4 Isolation of a mouse histone H1^{o} genomic clone derived from mouse strain 129

A.4.1 Screening mouse genomic library III

The mouse genomic library III (2.2.7) was screened using the 0.6 kb

Sac I/*Xba* I fragment isolated from pMH1^{o}.1mut as a probe (Fig. A-5). The library was plated as described (2.3.11). The probe was labelled with a α-^{32}P dATP (2.3.8 (ii)) and then hybridised to the filters. After hybridisation the filters were washed at a stringency of 0.5x SSC, 0.1% SDS at 65°C, and set up for autoradiography.

Two positive plaques were picked and these underwent two more rounds of screening in order to obtain pure positive plaques. The two plaques, designated λ2 and λ4, were obtained after the three rounds of screening and were analysed further.

A.4.2 Characterisation of the two positive lambda clones (λ2 and λ4)

DNA was prepared from λ2 and λ4 and Southern analysis was performed, using DNA digested with *Sal* I and *Eco* RI, and probed with the 0.6 kb *Sac* I/*Xba* I fragment isolated from pMH1^{o}.1mut (Fig. A-5). This result allowed the alignment of the two clones to be drawn as shown in Fig. A-8. It was found that λ4 contains DNA internal to λ2. The λ2 insert was approximately 18 kb with approximately 3.5 kb flanking at the 5' end of the H1^{o} gene region and 14.0 kb flanking at the 3' end of the H1^{o} gene region.

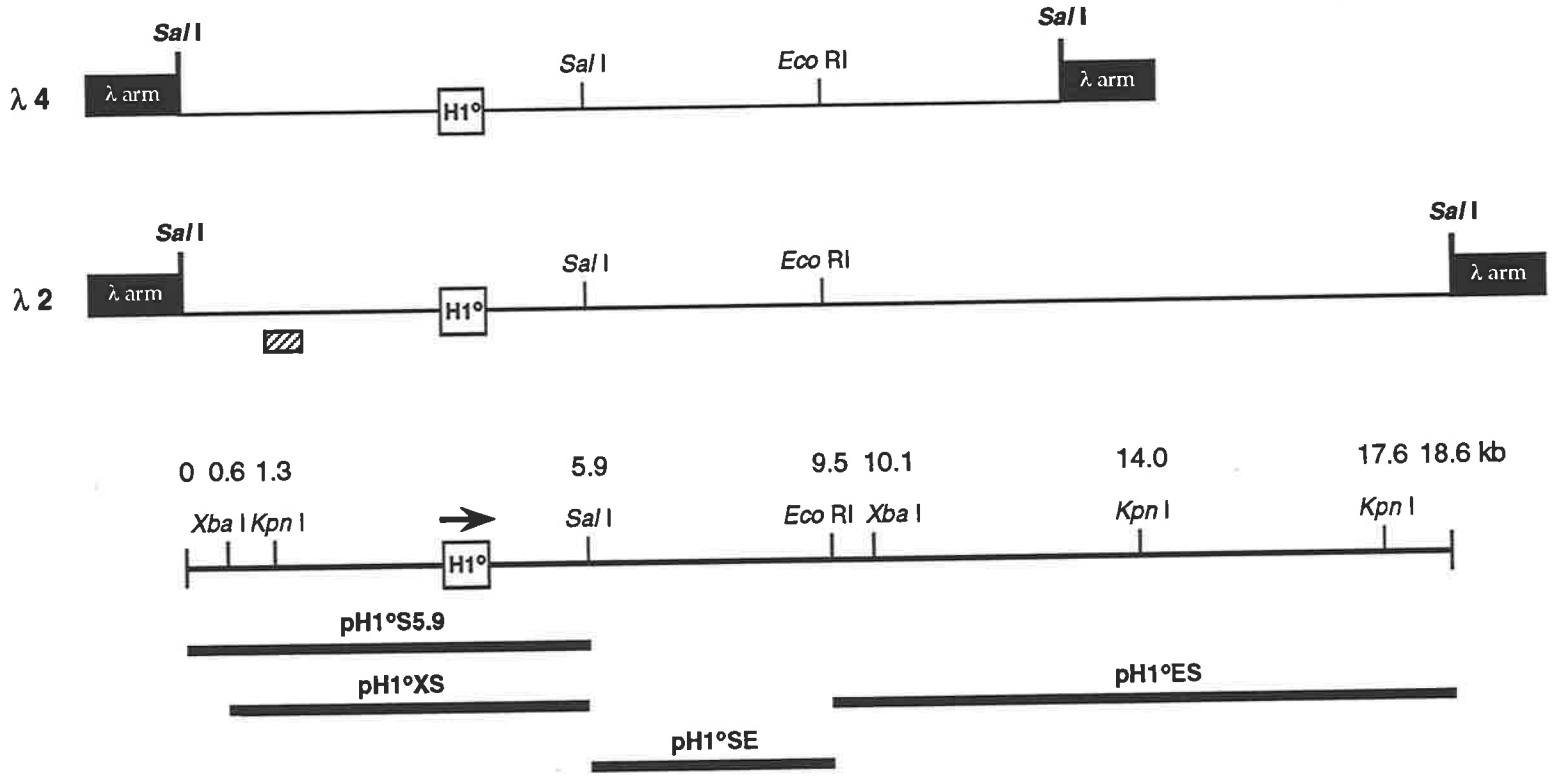
The insert fragments in the λ2 clone were subcloned into plasmid vectors for further analysis. The map of the subclones, pH1^{o}S5.9, pH1^{o}XS, pH1^{o}SE and pH1^{o}ES, relative to the lambda clones, is shown in Fig. A-8. Several restriction enzyme digestions of these subclones were performed, and showed that the restriction map surrounding the H1^{o} gene was identical to the previously isolated clone (Breuer *et al.*, 1989), indicating that the positive lambda clones are H1^{o} clones. The restriction map and the generation of the targeting construct are described in the following section.

Figure A-8

Alignment of λ4 and λ2 and subclones pH1°S5.9, pH1°XS, pH1°SE and pH1°ES

The overall restriction map of the region of genomic DNA covered by the inserts of λ4 and λ2 is shown below the maps of individual clones. The numbers above the map indicate the distances in kilobases from the 5' end of the genomic region. The location of the H1° gene is shown, and the arrow indicates the orientation of the gene. The thick lines represent the insert fragments (isolated from λ2) subcloned into plasmids, and the names of the recombinant plasmids are indicated.

pH1°S5.9 contains a 5.9 kb *Sal* I fragment subcloned into a *Sal* I-digested pTZ19R vector. pH1°XS contains a 5.3 kb *Xba* I/*Sal* I fragment subcloned into a *Xba* I/*Sal* I-digested pUC19 vector. pH1°SE and pH1°ES contain a 3.6 kb *Sal* I/*Eco* RI fragment and a 9.1 kb *Eco* RI/*Sal* I fragment subcloned into a *Sal* I/*Eco* RI-digested pBluescript KS+ vector, respectively.


λ 4
Sal I
λ arm
H1°
Sal I
Eco RI
Sal I
λ arm
λ 2
Sal I
λ arm
H1°
Sal I
Eco RI
Sal I
λ arm
0 0.6 1.3
5.9
9.5 10.1
14.0
17.6 18.6 kb
Xba I Kpn I
H1°
Sal I
Eco RI Xba I
Kpn I
Kpn I
pH1°S5.9
pH1°XS
pH1°ES
pH1°SE

A.5 Strategies for gene targeting using H1^{o} isogenic DNA

In an attempt to improve the targeting efficiency, it was decided to prepare an isogenic DNA construct using only a positive selection marker since the result described above showed that the enrichment strategy using positive/negative selection did not work. Therefore, a new targeting vector, named pH1^{o}HR2, was constructed. It contains the neomycin phosphotransferase gene (*neo^{r}*) flanked by nucleotide sequences homologous to the H1^{o} gene locus. However, pH1^{o}HR2 has not yet been used for targeting. Therefore, the strategy that can be used for screening of homologous recombination in ES cells will be described in the following section. In order to identify suitable diagnostic restriction enzyme sites for screening, Southern analysis of E14 ES cell genomic DNA was performed as described below.

A.5.1 Construction of pH1^{o}HR2

The targeting vector pH1^{o}HR2 was constructed using an H1^{o} genomic clone derived from a 129/Sv mouse genomic library (described above). This vector contained 11.8 kb of DNA homologous to the H1^{o} locus, and the *neo^{r}* gene, driven by the mouse phosphoglycerate kinase promoter (pgk-promoter) and a pgk-polyadenylation signal (pgk-poly A), as a selectable marker.

To construct pH1^{o}HR2, two recombinant plasmids named pH1^{o}XS.neo and pH1^{o}XK.neo were generated as precursors. The strategies used to create pH1^{o}XS.neo and pH1^{o}XK.neo are presented in Fig. A-9 and Fig. A-10, respectively.

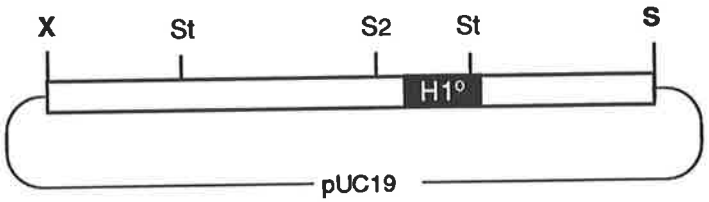
The strategy used for constructing pH1^{o}HR2 from precursors is shown in Fig. A-10. pH1^{o}HR2 contained the *neo^{r}* cassette transcribed in the opposite direction to H1^{o}. H1^{o} is rendered nonfunctional by the loss of the promoter region (45 bp upstream from the transcription start site), the 3'-untranslated region, and the coding sequence (amino acids 1-168) of the H1^{o} gene (Alonso *et al.*, 1988; Breuer *et al.*, 1989).

Figure A-9

Construction of pH1^{o}XS.neo

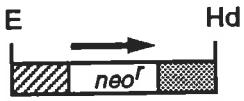
The strategy used for constructing pH1^{o}XS.neo plasmid, as a precursor for pH1^{o}HR2, is shown. The H1^{o} genomic fragment is shown as an open box with coding sequences filled. The *neo*^{r} cassette is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment are abbreviated as indicated. The restriction sites which were lost as a result of the ligation reaction are shown in brackets.

pH1^{o}XS.neo (8.9 kb) was constructed by replacing a 819 bp *Sac* II/*Stu* I fragment from pH1^{o}XS (Fig. A-8) containing the promoter region (45 bp upstream from the transcription start site), the 3'-untranslated region and coding sequence (amino acids 1-168) of the H1^{o} gene (Alonso *et al.*, 1988; Breuer *et al.*, 1989), with a 1.8 kb *Eco* RI/*Hin*d III pgkNEOpA (2.2.13) fragment. The ligation between the blunt end fragment and the blunt end vector regenerated the *Eco* RI restriction site as shown.



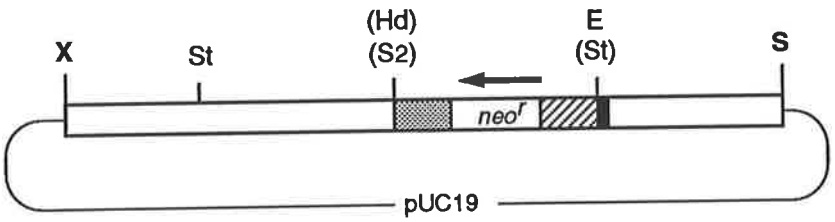
pH1°XS

partial *Stu* I / *Sac* II - digested vector (end-filled)



Ligation

Eco RI / *Hind* III fragment (end-filled)



pH1°XS. neo

1 kb

E	*Eco* RI	Hd	*Hind* III
S	*Sal* I	S2	*Sac* II
St	*Stu* I	X	*Xba* I
▨	pgk-promoter	▩	pgk-polyA

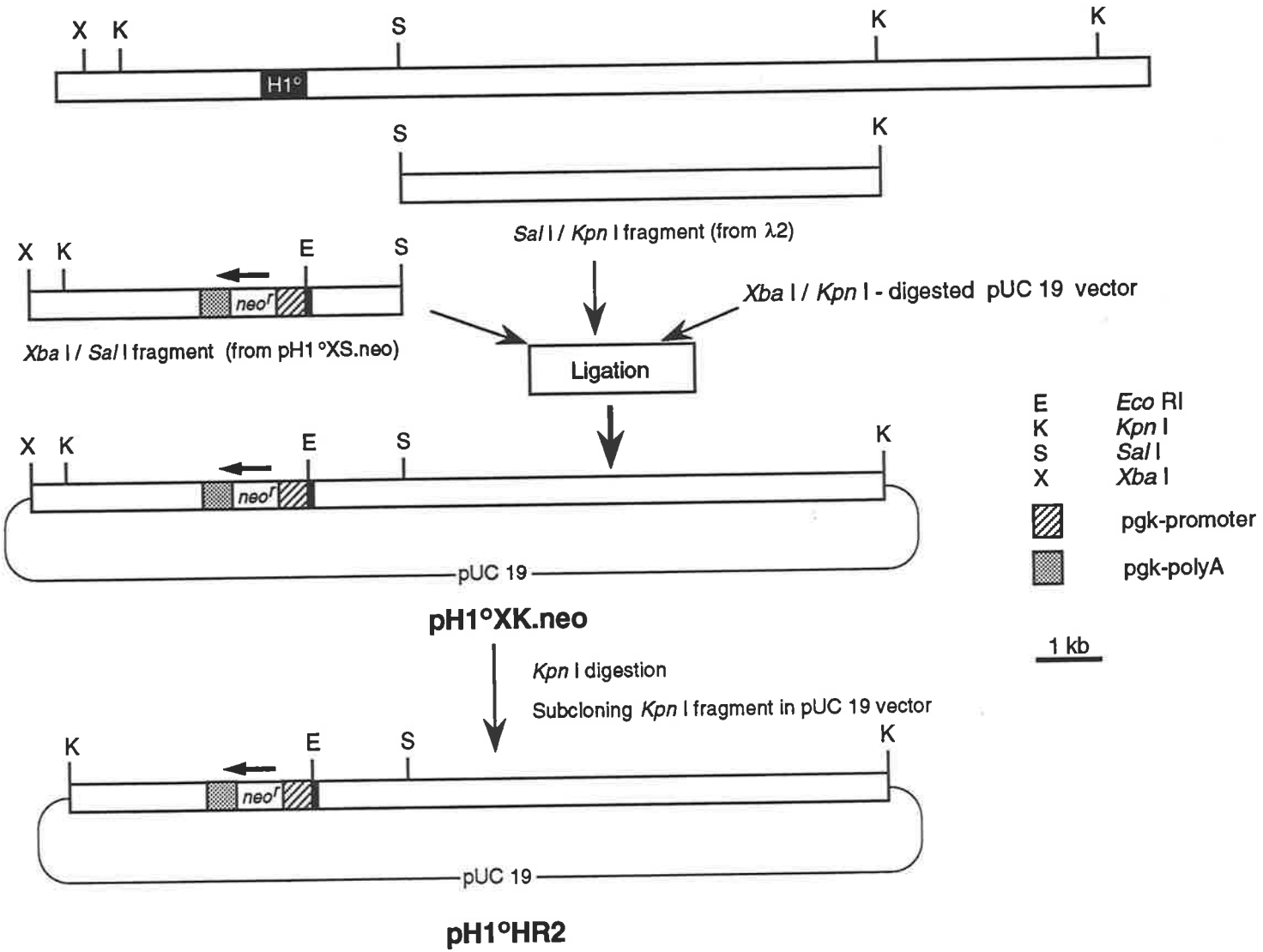
Figure A-10

Construction of pH1°XK.neo and the targeting vector pH1°HR2

The strategies used for constructing pH1°XK.neo and the targeting vector pH1°HR2 are shown. The H1° genomic fragment is shown as an open box with coding sequence filled. The position of the *neo*^{r} cassette within the H1° fragment is shown with the arrow above indicating the direction of transcription. The restriction sites within the insert fragment are abbreviated as indicated.

The H1° chromosomal locus is shown at the top. pH1°XK.neo (17.0 kb) was generated by ligation of a 6.3 kb *Xba* I/*Sal* I fragment from pH1°XS.neo (see Fig. A-9), a 8.1 kb *Sal* I/*Kpn* I fragment from λ2 (see Fig. A-8) and *Xba* I/*Kpn* I -digested pUC19 vector.

pH1°HR2 (16.3 kb) was generated by subcloning a 13.7 kb *Kpn* I fragment from pH1°XK.neo into a *Kpn* I-digested pUC19 vector.


X K
S
K
K
H1°
S
K
Sal I / Kpn I fragment (from λ2)
X K
E
S
neo^r
Xba I / Sal I fragment (from pH1 °XS.neo)
Xba I / Kpn I - digested pUC 19 vector
Ligation
X K
E
S
K
neo^r
pUC 19
pH1°XK.neo
E Eco RI
K Kpn I
S Sal I
X Xba I
pgk-promoter
pgk-polyA
1 kb
Kpn I digestion
Subcloning Kpn I fragment in pUC 19 vector
K
E
S
K
neo^r
pUC 19
pH1°HR2

A.5.2 Southern analysis of mouse genomic DNA isolated from E14 ES cells

DNA was isolated from E14 ES cells as described (2.3.3 (iii)). The DNA was digested with different restriction enzymes and Southern analysis was performed as described (2.3.10). The filter was probed with a 1.3 kb *Sal* I/*Kpn* I fragment isolated from pH1^{o}S5.9 (see Fig. A-8). This probe is located 5' to the H1^{o} gene. After hybridisation, the filter was washed at high stringency. The result is presented in Fig. A-11. Most of the restriction enzyme digestions gave only a single hybridising fragment. From this result, *Eco* RI was chosen as a diagnostic enzyme for detection of homologous recombination in ES cells. In addition, this result also showed that the probe contains a unique DNA sequence and therefore can be used as a 5' external probe for gene targeting experiments.

A.5.3 Strategy for gene targeting using pH1^{o}HR2 DNA

Prior to electroporation, pH1^{o}HR2 can be digested with *Kpn* I in order to release the insert fragment from the plasmid (see Fig. A-10) and processed as described (2.3.17). pH1^{o}HR2 DNA can then be introduced into ES cells using electroporation protocol 2 (see 2.6.3 and 5.2.1), since the studies described in this thesis show that this protocol gives the highest percentage of single copy integration (5.2.1). After selection with G418 for 8-10 days, DNA isolated from G418-resistant clones can be screened by Southern analysis using *Eco* RI and probed with a 1.3 kb *Sal* I/*Kpn* I isolated from pH1^{o}S5.9 (see Fig. A-8). This strategy is presented in Fig. A-12. The probe does not contain any sequence present in pH1^{o}HR2. The wild-type allele will give a 14 kb hybridising fragment, whereas the mutated allele resulting from the homologous recombination will give a 9.7 kb hybridising fragment with this probe. To verify the targeted ES cell clones, and to determine the number of integration sites, DNA isolated from the positive clones obtained from the initial screening can be digested with *Kpn* I and probed with a *neo^{r}* coding probe. However, it is also necessary to probe the DNA of the positive clones using a 3'-external probe in order to ensure that the

Figure A-11

Southern analysis of mouse genomic DNA isolated from E14 ES cells

Genomic DNA isolated from E14 ES cells was digested with individual restriction enzymes and combinations. The digested DNA was electrophoresed on a 0.8% agarose gel along with DNA size markers. DNA was transferred from the gel to a Zeta-Probe membrane, probed with ^{32}P-labelled 1.3 kb *Sal* I/*Kpn* I fragment isolated from pH1^{o}S5.9 (see Fig. A-8) which is located at the 5'-end of the H1^{o} gene. After hybridisation at 65^{o}C overnight, the filter was washed at 65^{o}C with 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 5% (w/v) SDS for 30 minutes, and then in 40 mM Na_2HPO_4 pH 7.2, 1 mM EDTA, 1% (w/v) SDS, 65^{o}C for 30 minutes and the hybridising fragments were detected by autoradiography.

Lanes are as follows:

1) Hi-Lo mixed DNA size markers
2) *Bam* HI-digested DNA
3) *Bst* EII-digested DNA
4) *Eco* RI-digested DNA
5) *Eco* RV-digested DNA
6) *Hinc* II-digested DNA
7) *Hind* III-digested DNA
8) *Kpn* I-digested DNA
9) *Pst* I-digested DNA
10) *Pvu* II-digested DNA
11) *Xba* I-digested DNA
12) *Bam* HI/*Eco* RI-digested DNA
13) *Bam* HI/*Eco* RV-digested DNA
14) *Bam* HI/*Hind* III-digested DNA
15) *Bam* HI/*Kpn* I-digested DNA
16) *Kpn* I/*Eco* RI-digested DNA
17) *Kpn* I/*Eco* RV-digested DNA
18) *Eco* RI-digested SPP1 phage DNA size markers

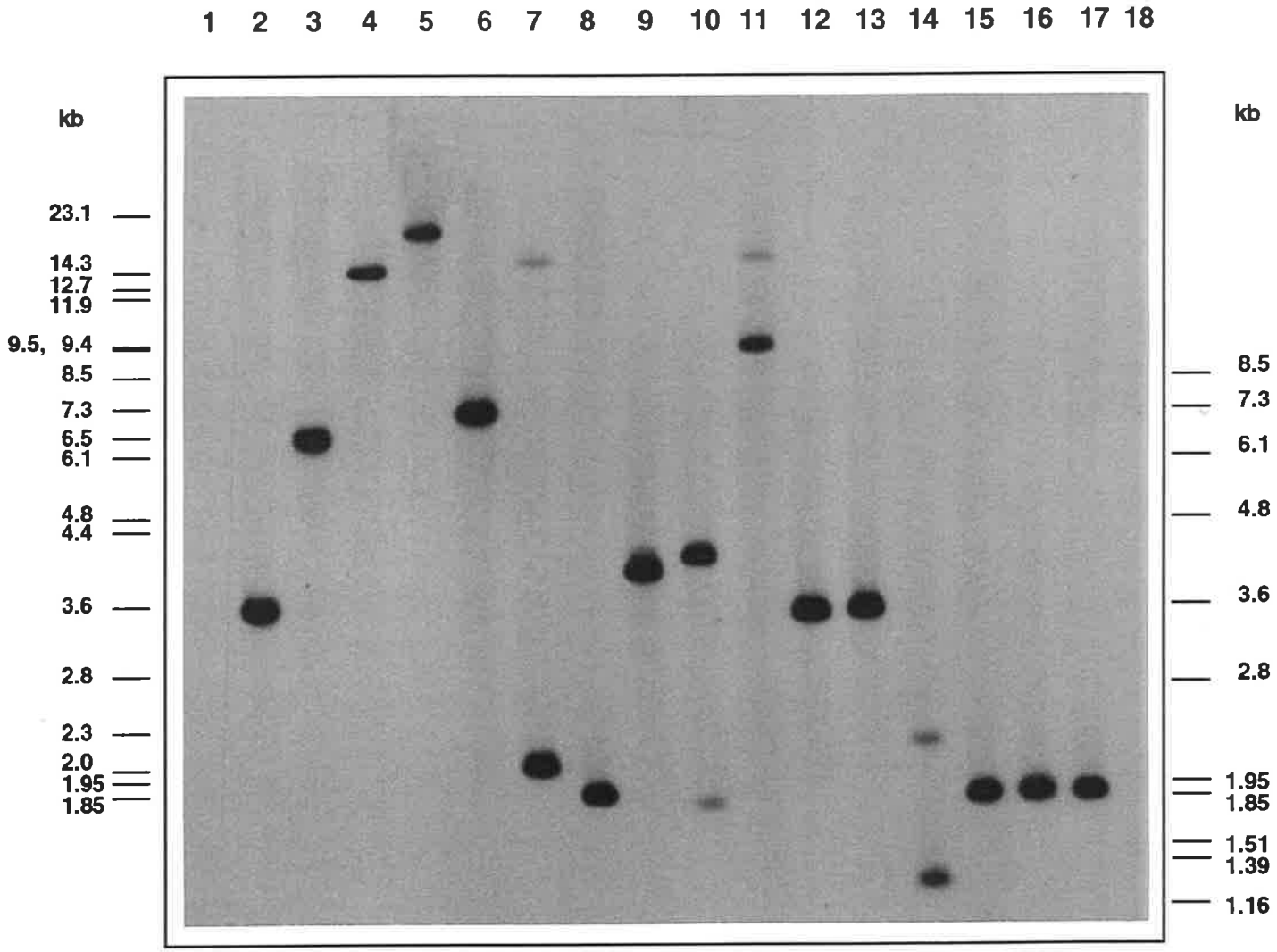
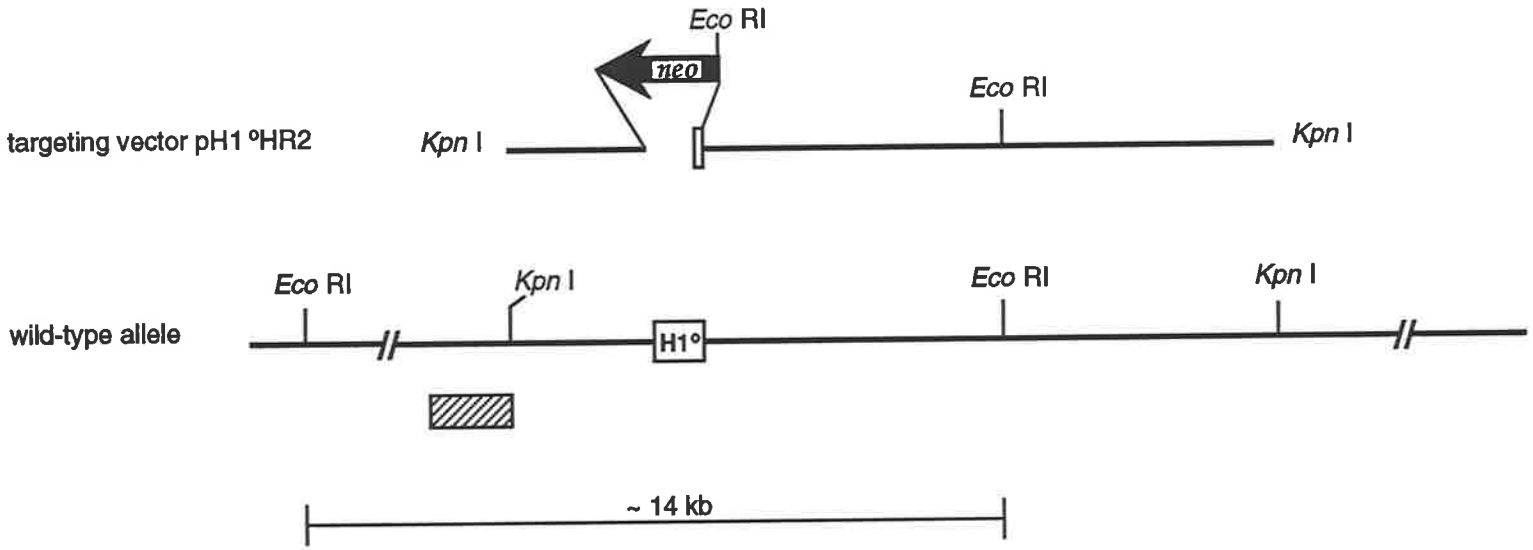

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
kb
23.1
14.3
12.7
11.9
9.5, 9.4
8.5
7.3
6.5
6.1
4.8
4.4
3.6
2.8
2.3
2.0
1.95
1.85
kb
8.5
7.3
6.1
4.8
3.6
2.8
1.95
1.85
1.51
1.39
1.16

Figure A-12

Strategy for detection of a homologous recombination event in ES cells using pH1^{o}HR2 DNA

The targeting vector pH1^{o}HR2 is shown at the top in alignment with the map of the H1^{o} chromosomal locus. The H1^{o} gene is represented by an open box. The *neo^{r}* cassette is represented by the arrow indicating the direction of transcription. A 1.3 kb *Sal* I/*Kpn* I fragment isolated from pH1^{o}S5.9 (see Fig. A-8) is represented by the striped box. This fragment is located 5'-external to the targeting sequence.

The H1^{o} gene is flanked by two *Eco* RI sites about 14 kb apart. The *neo^{r}* cassette within pH1^{o}HR2 contains a unique *Eco* RI site. If ES cells were transfected with pH1^{o}HR2 DNA and homologous recombination occurred, a new *Eco* RI site would be inserted between the endogenous sites. Therefore, it is possible to screen the ES cells transfected with pH1^{o}HR2 by Southern analysis using *Eco* RI digestion and hybridising with a 1.3 kb *Sal* I/*Kpn* I fragment as a probe. By using this strategy, the "wild-type" and the mutated alleles are expected to give hybridising fragments of 14 kb and 9.7 kb, respectively. The positive ES cell clones from the initial screening can be further analysed by digestion with *Kpn* I and probing with a *neo^{r}* coding probe to determine the number of integration sites. The mutated allele is expected to give a 13.7 kb band with the *neo^{r}* coding probe.


Eco RI
neo
Eco RI
targeting vector pH1 °HR2
Kpn I
Kpn I
Eco RI
Kpn I
Eco RI
Kpn I
wild-type allele
H1°
~ 14 kb

positive clones are not the result of a type III recombination event (Adair *et al.*, 1989).

A.6 Summary

The work presented here describes an attempt to use gene targeting in order to obtain an ES cell in which one allele of the H1^{o} gene has been disrupted. Targeting vector pH1^{o}HR1 was designed using a positive/negative selection strategy in order to enrich for the targeted ES cells. The result showed that this strategy gave only a two fold enrichment, and homologous recombination was not observed in any of 105 analysed clones.

In order to improve the targeting efficiency, a mouse genomic library derived from 129 strain was screened to allow production of a targeting vector with isogenic DNA. H1^{o} genomic clones were isolated and then used to generate the targeting vector pH1^{o}HR2. This vector contains 11.8 kb of DNA homologous to the H1^{o} locus, and the *neo^{r}* gene as a positive selectable marker. pH1^{o}HR2 has not yet been used for gene targeting experiments. The strategy for detecting homologous recombination with pH1^{o}HR2 has been described (Southern analysis of *Eco* RI-digested DNA and a 1.3 kb *Sal* I/*Kpn* I fragment isolated from pH1^{o}S5.9 as a probe). Southern analysis on ES cell genomic DNA was performed and described. The result showed that the 1.3 kb *Sal* I/*Kpn* I probe can be used for the initial screening, since it contains a unique sequence.

A.7 Future work

Gene targeting in ES cells can be performed using pH1^{o}HR2 as a targeting vector, and H1^{o} targeted-ES cells can be identified using the screening strategy described above. However, a 3'-external probe, and the appropriate restriction

enzyme digestion have not yet been identified. Further work is therefore required to search for a specific 3' probe and diagnostic enzyme(s).

If H1^{o} targeted ES cell clones were obtained and verified, the next step would be to inject the targeted ES clones which have normal karyotype, are mycoplasma-free and have normal growth characteristics, into host blastocysts and generate germ line chimaeras. A breeding program with mice heterozygous for the disrupted H1^{o} gene would then be set up to attempt to generate mice or embryos homozygous for the disrupted H1^{o} gene.

Bibliography

Adair, G. M., Nairn, R. S., Wilson, J. H., Seidman, M. M., Brotherman, K. A., MacKinnon, C., and Scheerer, J. B. (1989). Targeted homologous recombination at the endogenous adenine phosphoribosyltransferase locus in Chinese hamster cells. *Proc. Natl. Acad. Sci. USA*, **86**, 4574-4578.

Allan, J., Hartman, P. G., Crane-Robinson, C., and Aviles, F. X. (1980). The structure of histone H1 and its location in chromatin. *Nature*, **288**, 675-679.

Allegra, P., Sterner, R., Clayton, D. F., and Allfrey, V. G. (1987). Affinity chromatographic purification of nucleosomes containing transcriptionally active DNA sequences. *J. Mol. Biol.*, **196**, 379-388.

Allfrey, V. G. (1980). Molecular aspects of the regulation of eukaryotic transcription-nucleosomal proteins and their postsynthetic modification in the control of DNA conformation and template function. In *Cell Biology: A Comprehensive Treatise* ., vol. 3, eds. L. Goldstein and D. M. Prescott, pp. 348-437., New York: Academic Press.

Allis, C. D., Glover, C. V.C., Bowen, J. K., and Gorovsky, M. A. (1980). Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, *Tetrahymena thermophila*. *Cell*, **20**, 609-617.

Allis, C. D., Ziegler, Y. S., Gorovsky, M. A., and Olmsted, J. B. (1982). A conserved histone variant enriched in nucleoli of mammalian cells. *Cell*, **31**, 131-136.

Alonso, A., Breuer, B., Bouterfa, H., and Doenecke, D. (1988). Early increase in histone $H1^{o}$ mRNA during differentiation of F9 cells to parietal endoderm. *EMBO J.*, **7**, 3003-3008.

Appels, R., and Wells, J. R. E. (1972). Synthesis and turnover of DNA-bound histone during maturation of avian red blood cells. *J. Mol. Biol.*, **70**, 425-434.

Appels, R., Wells, J. R. E., and Williams, A. F. (1972). Characterization of DNA-bound histone in the cells of the avian erythropoietic series. *J. Cell. Sci.*, **10**, 47-59.

Arents, G., and Moudrianakis, E. N. (1993). Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc. Natl. Acad. Sci. USA*, **90**, 10489-93.

Arents, G., Burlingame, R. W., Wang, B.-C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. USA*, **88**, 10148-10152.

Ausio, J., Dong, F., and van-Holde, K. E. (1989). Use of selectively trypsinized nucleosome core particles to analyze the role of the histone "tails" in the stabilization of the nucleosome. *J. Mol. Biol.*, **206**, 451-463.

Aviles, F. J., Chapman, G. E., Kneale, G. G., Crane-Robinson, C., and Bradbury, E. M. (1978). The conformation of histone H5. Isolation and characterisation of the globular segment. *Eur. J. Biochem.*, **88**, 363-371.

Bauer, W. R., Hayes, J. J., White, J. H., and Wolffe, A. P. (1994). Nucleosome structural changes due to acetylation. *J. Mol. Biol.*, **236**, 685-690.

Benton, W. D., and Davis, R. W. (1977). Screening λgt recombinant clones by hybridization to single plaques *in situ*. *Science*, **196**, 180-182.

Benyajati, C., and Worcel, A. (1976). Isolation, characterization and structure of the folded interphase genome of *Drosophila melanogaster*. *Cell*, **9**, 393-407.

Bergman, M. G., Wawra, E., and Winge, M. (1988). Chicken histone H5 inhibits transcription and replication when introduced into proliferating cells by microinjection. *J. Cell. Sci.*, **91**, 201-209.

Bhatnagar, Y. M., McCullar, M. K., and Chronister, R. B. (1984). Immunocytochemical localization of a histone H2A variant in the mammalian nucleolar chromatin. *Cell. Biol. Int. Rep.*, **8**, 971-979.

Bird, A. (1992). The essentials of DNA methylation. *Cell*, **70**, 5-8.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209-213.

Boffa, L. C., Vidali, G., Mann, R. S., and Allfrey, V. G. (1978). Suppression of histone deacetylation *in vivo* and *in vitro* by sodium butyrate. *J. Biol. Chem.*, **253**, 3364-3366.

Boyes, J., and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*, **64**, 1123-1134.

Bradley, A. (1987). Production and analysis of chimaeric mice. In *Teratocarcinomas and embryonic stem cells: A Practical Approach*, ed. E. J. Robertson, pp. 113-151. Oxford: IRL Press.

Braunstein, M., Rose, A. B., Holmes, S. G., Allis, C. D., and Broach, J. R. (1993). Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev.*, **7**, 592-604.

Bresnick, E. H., Rories, C., and Hager, G. L. (1992). Evidence that nucleosomes on the mouse mammary tumor virus promoter adopt specific translational positions. *Nucleic Acids Res.*, **20**, 865-870.

Breuer, B., Fischer, J., and Alonso, A. (1989). Cloning and characterization of the mouse histone H1° promoter region. *Gene*, **81**, 307-314.

Bukhari, A. I., and Zipser, D. (1973). Mutants of *Escherichia coli* with a defect in the degradation of nonsense fragments. *Nature*, **243**, 238-241.

Butler, P. J. G. (1988). The organization of the chromatin fiber. In *Chromosome and Chromatin*, vol 1., ed. K. W. Adolph, pp. 57-82. Boca Raton, Florida: CRC Press.

Candido, E. P. M., Reeves, R., and Davie, J. R. (1978). Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell*, **14**, 105-113.

Capecchi, M. R. (1989). Altering the genome by homologous recombination. *Science*, **244**, 1288-1292.

Capecchi, M. R. (1994). Targeted gene replacement. *Sci. Am.*, **270**, 52-59.

Cary, P. D., Hines, M. L., Bradbury, E. M., Smith, B. J., and Johns, E. W. (1981). Conformation studies of histone H1 degree in comparison with histones H1 and H5. *Eur. J. Biochem.*, **120**, 371-377.

Chamberlin, M., and Ring, J. (1973). Characterization of T7-specific ribonucleic acid polymerase. *J. Biol. Chem.*, **248**, 2235-2244.

Chamberlin, M., McGrath, J., and Waskell, L. (1970). New RNA polymerase from Escherichia coli infected with bacteriophage T7. *Nature*, **228**, 227-231.

Chan, S., Attisano, L., and Lewis, P. N. (1988). Histone H3 thiol reactivity and acetyltransferases in chicken erythrocyte nuclei. *J. Biol. Chem.*, **263**, 15643-15651.

Choe, J., Kolodrubetz, D., and Grunstein, M. (1982). The two yeast histone H2A genes encode similar protein subtypes. *Proc. Natl. Acad. Sci. USA*, **79**, 1484-1487.

Chomczynski, P., and Qasba, P. K. (1984). Alkaline transfer of DNA to plastic membrane. *Biochem. Biophys. Res. Commun.*, **122**, 340-344.

Churchill, M. E., and Suzuki, M. (1989). 'SPKK' motifs prefer to bind to DNA at A/T-rich sites. *EMBO J.*, **8**, 4189-4195.

Ciechanover, A., Finley, D., and Varshavsky, A. (1984). Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell*, **37**, 57-66.

Clark, D. J., and Felsenfeld, G. (1991). Formation of nucleosomes on positively supercoiled DNA. *EMBO J.*, **10**, 387-395.

Clark, D. J., and Felsenfeld, G. (1992). A nucleosome core is transferred out of the path of a transcribing polymerase. *Cell*, **71**, 11-22.

Csordas, A. (1990). On the biological role of histone acetylation. *Biochem J.*, **265**, 23-38.

Dagert, M., and Ehrlich, S. D. (1979). Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene*, **6**, 23-28.

Dalton, S., Robins, A. J., Harvey, R. P., and Wells, J. R. E. (1989). Transcription from the intron-containing chicken histone H2A.F gene is not S-phase regulated. *Nucleic Acids Res.*, **17**, 1745-1756.

DeChiara, T. M., Efstratiadis, A., and Robertson, E. J. (1990). A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature*, **345**, 78-80.

Deng, C., and Capecchi, M. R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell. Biol.*, **12**, 3365-3371.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A Comprehensive Set of Sequence Analysis Programs for the VAX. *Nucleic Acids Res.*, **12**, 387-395.

Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C., Jr., Butel, J. S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature*, **356**, 215-221.

Dunn, J. J., and Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.*, **166**, 477-535.

Edward, D. R., Parfett, C. I. J., and Denhardt, D. (1985). Transcriptional regulation of two serum-induced RNAs in mouse fibroblasts. *Mol. Cell. Biol.*, **5**, 3280-3288.

Eisen, H., Gjerset, R., and Hasthorpe, S. (1981). Distribution and regulation of histone H1^{o} in rodents. In *Histone Genes: Structure, Organization and Regulation*, eds. G. S. Stein, J. L. Stein, and W. F. Marzluff. New York: John Wiley and Sons.

Eissenberg, J. C., Cartwright, I. L., Thomas, G. H., and Elgin, S. C. R. (1985). Selected topics in chromatin structure. *Ann. Rev. Genet.*, **19**, 485-536.

Elgin, S. C. R. (1988). The formation and function of DNase I hypersensitive sites in the process of gene activation. *J. Biol. Chem.*, **263**, 19259-19262.

Elgin, S. C. R. (1990). Chromatin structure and gene activity. *Curr. Opin. Cell. Biol.*, **2**, 437-445.

Engel, J. D., Sugarman, B. J., and Dodgson, J. B. (1982). A chicken histone H3 gene contains intervening sequences. *Nature*, **297**, 434-436.

Ernst, S. G., Miller, H., Brenner, C. A., Nocente-McGrath, C., Francis, S., and McIsaac, R. (1987). Characterization of a cDNA clone coding for a sea urchin histone H2A variant related to the H2A.F/Z histone protein in vertebrates. *Nucleic Acids Res.*, **15**, 4629-4644.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, **292**, 154-156.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6-13.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. *Nature*, **355**, 219-224.

Felsenfeld, G., and McGhee, J. D. (1986). Structure of the 30 nm chromatin fiber. *Cell*, **44**, 375-377.

Finch, J. T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA*, **73**, 1897-1901.

Finch, J. T., Lutter, L. C., Rhodes, D., Brown, R. S., Rushton, B., Levitt, M., and Klug, A. (1977). Structure of nucleosome core particles of chromatin. *Nature*, **269**, 29-36.

García-Iglesias, M. J., Ramirez, A., Monzo, M., Steuer, B., Martínez, J. M., Jorcano, J. L., and Alonso, A. (1993). Specific expression in adult mice and post-implantation embryos of a transgene carrying the histone H1^{o} regulatory region. *Differentiation*, **55**, 27-35.

García-Segura, L. M., Luquín, S., Martínez, P., Casas, M. T., and Suau, P. (1993). Differential expression and gonadal hormone regulation of histone H1^{o} in the developing and adult rat brain. *Brain Res. : Dev. Brain Res.*, **73**, 63-70.

Garrard, W. T. (1991). Histone H1 and the conformation of transcriptionally active chromatin. *Bioessays*, **13**, 87-88.

Gasser, S. M., and Laemmli, U. K. (1987). A glimpse at chromosomal order. *Trends Genet.*, **3**, 16-22.

Georgiev, G. P., Vassetzky, Y., Jr., Luchnik, A. N., Chernokhvostov, V. V., and Razin, S. V. (1991). Nuclear skeleton, DNA domains and control of replication and transcription. *Eur. J. Biochem.*, **200**, 613-624.

Gjerset, R., Gorka, C., Hasthorpe, S., Lawrence, J. J., and Eisen, H. (1982). Developmental and hormonal regulation of protein $H1^{o}$ in rodents. *Proc. Natl. Acad. Sci. USA*, **79**, 2333-2337.

Goldknopf, I. L., and Busch, H. (1977). Isopeptide linkage between nonhistone and histone H2A polypeptides of chromosomal conjugate-protein A24. *Proc. Natl. Acad. Sci. USA*, **74**, 864-868.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D., and Boyse, E. A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci. USA*, **72**, 11-15.

Golomb, M., and Chamberlin, M. (1974). Characterization of T7-specific ribonucleic acid polymerase. IV. Resolution of the major in vitro transcripts by gel electrophoresis. *J. Biol. Chem.*, **249**, 2858-2863.

Graziano, V., Gerchman, S. E., Schneider, D. K., and Ramakrishnan, V. (1994). Histone H1 is located in the interior of the chromatin 30-nm filament. *Nature*, **368**, 351-354.

Grodberg, J., and Dunn, J. J. (1988). *ompT* encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.*, **170**, 1245-1253.

Gross, D. S., and Garrard, W. T. (1988). Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.*, **57**, 159-197.

Gruber, A., Streit, A., Reist, M., Benninger, P., Böhni, R., and Schümperli, D. (1990). Structure of a mouse histone-encoding gene cluster. *Gene*, **95**, 303-304.

Grunstein, M. (1990a). Histone function in transcription. *Annu. Rev. Cell Biol.*, **6**, 643-678.

Grunstein, M. (1990b). Nucleosomes: regulators of transcription. *Trends Genet.*, **6**, 395-400.

Gurley, L. R., D'Anna, J. A., Barham, S. S., Deaven, L. L., and Tobey, R. A. (1978). Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. *Eur. J. Biochem.*, **84**, 1-15.

Hansen, J. C., and Wolffe, A. P. (1994). A role for histones H2A/H2B in chromatin folding and transcriptional repression. *Proc. Natl. Acad. Sci. USA*, **91**, 2339-2343.

Harvey, R. P., Whiting, J. A., Coles, L. S., Krieg, P. A., and Wells, J. R. E. (1983). H2A.F: an extremely variant histone H2A sequence expressed in the chicken embryo. *Proc. Natl. Acad. Sci. USA*, **80**, 2819-2823.

Hasty, P., and Bradley, A. (1993). Gene targeting vectors for mammalian cells. In *Gene Targeting: A Practical Approach,* ed. A. L. Joyer, pp. 1-31. Oxford: IRL Press.

Hasty, P., Rivera-Pérez, J., and Bradley, A. (1991). The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell. Biol.*, **11**, 5586-5591.

Hatch, C. L., and Bonner, W. M. (1988). Sequence of cDNAs for mammalian H2A.Z, an evolutionarily diverged but highly conserved basal histone H2A isoprotein species. *Nucleic Acids Res.*, **16**, 1113-1124.

Hatch, C. L., and Bonner, W. M. (1990). The human histone H2A.Z gene. Sequence and regulation. *J. Biol. Chem.*, **265**, 15211-15218.

Hebbes, T. R., Clayton, A. L., Thorne, A. W., and Crane-Robinson, C. (1994). Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken β-globin chromosomal domain. *EMBO J.*, **13**, 1823-30.

Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.*, **7**, 1395-1402.

Hentschel, C. C., and Birnstiel, M. L. (1981). The organization and expression of histone gene families. *Cell*, **25**, 301-313.

Hershko, A. (1983). Ubiquitin: roles in protein modification and breakdown. *Cell*, **34**, 11-12.

Hershko, A., Leshinsky, E., Ganoth, D., and Heller, H. (1984). ATP-dependent degradation of ubiquitin-protein conjugates. *Proc. Natl. Acad. Sci. USA*, **81**, 1619-1623.

Hill, C. S., and Thomas, J. O. (1990). Core histone-DNA interactions in sea urchin sperm chromatin. The N-terminal tail of H2B interacts with linker DNA. *Eur. J. Biochem.*, **187**, 145-153.

Hnilica, L. S., Stein, G. S., and Stein, J. L. (1989). *Histones and Other Basic Nuclear Proteins*. Boca Raton, Florida: CRC press.

Holliday, R., and Ho, T. (1991). Gene silencing in mammalian cells by uptake of 5-methyl deoxycytidine-5'-triphosphate. *Somat. Cell Mol. Genet.*, **17**, 537-542.

Huang, S. Y., Barnard, M. B., Xu, M., Matsui, S., Rose, S. M., and Garrard, W. T. (1986). The active immunoglobulin kappa chain gene is packaged by non-ubiquitin-conjugated nucleosomes. *Proc. Natl. Acad. Sci. USA*, **83**, 3738-3742.

Igo-Kemenes, T., and Zachau, H. G. (1977). Domains in chromatin structure. *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 109-118.

Isenberg, I. (1979). Histones. *Annu. Rev. Biochem.*, **48**, 159-191.

Jeannotte, L., Ruiz, J. C., and Robertson, E. J. (1991). Low level of *Hox1.3* gene expression does not preclude the use of promoterless vectors to generate a targeted gene disruption. *Mol. Cell. Biol.*, **11**, 5578-5585.

Jentsch, S. (1992). The ubiquitin-conjugation system. *Annu. Rev. Genet.*, **26**, 179-207.

Jerzmanowski, A., and Cole, R. D. (1990). Flanking sequences of *Xenopus* 5 S RNA genes determine differential inhibition of transcription by H1 histone in vitro. Mitotic phosphorylation of H1 decreases its inhibitory power. *J. Biol. Chem.*, **265**, 10726-10732.

Johnson, E. M., Sterner, R., and Allfrey, V. G. (1987). Altered nucleosomes of active nucleolar chromatin contain accessible histone H3 in its hyperacetylated forms. *J. Biol. Chem.*, **262**, 6943-6946.

Johnson, L. M., Kayne, P. S., Kahn, E. S., and Grunstein, M. (1990). Genetic evidence for an interaction between SIR3 and histone H4 in the repression of the silent mating loci in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA*, **87**, 6286-6290.

Johnson, R. S., Sheng, M., Greenberg, M. E., Kolodner, R. D., Papaioannou, V. E., and Spiegelman, B. M. (1989). Targeting of nonexpressed genes in embryonic stem cells via homologous recombination. *Science*, **245**, 1234-1236.

Joyner, A. L., Skarnes, W. C., and Rossant, J. (1989). Production of a mutation in mouse *En-2* gene by homologous recombination in embryonic stem cells. *Nature*, **338**, 153-156.

Kamakaka, R. T., and Thomas, J. O. (1990). Chromatin structure of transcriptionally competent and repressed genes. *EMBO J.*, **9**, 3997-4006.

Kim, H. S., and Smithies, O. (1988). Recombinant fragment assay for gene targetting based on the polymerase chain reaction. *Nucleic Acids Res.*, **16**, 8887-8903.

Klug, A., Rhodes, D., Smith, J., Finch, J. T., and Thomas, J. O. (1980). A low resolution structure for the histone core of the nucleosome. *Nature*, **287**, 509-516.

Koller, B. H., and Smithies, O. (1992). Altering genes in animals by gene targeting. *Annu. Rev. Immunol.*, **10**, 705-730.

Kornberg, R. D., and Lorch, Y. (1991). Irresistible force meets immovable object: Transcription and the nucleosome. *Cell*, **67**, 833-836.

Kornberg, R. D., and Lorch, Y. (1992). Chromatin structure and transcription. *Annu. Rev. Cell Biol.*, **8**, 563-587.

Krieg, P. A., Robins, A. J., D' Andrea, R., and Wells, J. R. E. (1983). The chicken H5 gene is unlinked to core and H1 histone genes. *Nucleic Acids Res.*, **11**, 619-627.

Krystal, G. W., and Poccia, D. L. (1981). Phosphorylation of cleavage stage histone H1 in mitotic and prematurely condensed chromosomes. *Exp. Cell Res.*, **134**, 41-48.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell*, **72**, 73-84.

Lee, H. L., and Archer, T. K. (1994). Nucleosome-mediated disruption of transcription factor-chromatin initiation complexes at the mouse mammary tumor virus long terminal repeat in vivo. *Mol. Cell Biol.*, **14**, 32-41.

Lennox, R. W., and Cohen, L. H. (1983). The histone H1 complements of dividing and nondividing cells of the mouse. *J. Biol. Chem.*, **258**, 262-268.

Lennox, R. W., and Cohen, L. H. (1984). The alterations in H1 histone complement during mouse spermatogenesis and their significance for H1 subtype function. *Dev. Biol.*, **103**, 80-84.

Levinger, L., and Varshavsky, A. (1982). Selective arrangement of ubiquitinated and D1 protein-containing nucleosomes within the Drosophilia genome. *Cell*, **28**, 375-385.

Lin, F.-L., Sperle, K., and Sternberg, N. (1985). Recombination in mouse L cells between DNA introduced into cells and homologous chromosomal sequences. *Proc. Natl. Acad. Sci. USA*, **82**, 1391-1395.

Lin, R., Leone, J. W., Cook, R. G., and Allis, C. D. (1989). Antibodies specific to acetylated histones document the existence of deposition- and transcription-related histone acetylation in Tetrahymena. *J. Cell Biol.*, **108**, 1577-1588.

Lorch, Y., LaPointe, J. W., and Kornberg, R. D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell*, **49**, 203-210.

Lu, Q., Wallrath, L. L., Granok, H., and Elgin, S. C. R. (1993). (CT)n (GA)n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the Drosophila hsp26 gene. *Mol. Cell Biol.*, **13**, 2802-2814.

Mann, G. B., Fowler, K. J., Gabriel, A., Nice, E. C., Williams, R. L., and Dunn, A. R. (1993). Mice with a null mutation of the TGFα gene have abnormal skin architecture, wavy hair, and curly whiskers and often develop corneal inflammation. *Cell*, **73**, 249-261.

Mannironi, C., Bonner, W. M., and Hatch, C. L. (1989). H2A.X. a histone isoprotein with a conserved carboxyl-terminal sequence, is encoded by a novel messenger RNA with both DNA replication type and polyA 3' processing signals. *Nucleic Acids Res.*, **17**, 9113-9126.

Mansour, S. L., Thomas, K. R., and Capecchi, M. R. (1988). Disruption of the proto-oncogene *int-2* in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature*, **336**, 348-352.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA*, **78**, 7634-7638.

Matsudaira, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.*, **262**, 10035-10038.

Matsui, S. I., Seon, B. K., and Sandberg, A. A. (1979). Disappearance of a structural chromatin protein A24 in mitosis: implications for molecular basis of chromatin condensation. *Proc. Natl. Acad. Sci. USA*, **76**, 6386-6390.

Matsumoto, S., and Yanagida, M. (1985). Histone gene organization of fission yeast: a common upstream sequence. *EMBO J.*, **4**, 3531-3538.

Matthews, H. R. (1988). Histone Modifications and Chromatin Structure. In *Chromosomes and Chromatin*, ed. K. W. Adolph, pp. 3-32. Boco Raton, Florida: CRC Press.

Maxson, R., Cohn, R., Kedes, L., and Mohun, T. (1983). Expression and organization of histone genes. *Ann. Rev. Genet.*, **17**, 239-277.

May, G. S., and Morris, N. R. (1987). The unique histone H2A gene of *Aspergillus nidulans* contains three introns. *Gene*, **58**, 59-66.

McGhee, J. D., and Felsenfeld, G. (1980). Nucleosome structure. *Annu. Rev. Biochem.*, **49**, 1115-1156.

McIsaac, R., Ng, E. Y. W., Nocente, M. C., and Ernst, S. G. (1992). Histone H2A.F/Z mRNA is stored in the egg cytoplasm and basally regulated in the sea urchin embryo. *Dev. Biol.*, **153**, 402-406.

Mendelson, E., Landsman, D., Druckmann, S., and Bustin, M. (1986). Immunofractionation of chromatin regions associated with histone $H1^{o}$. *Eur. J. Biochem.*, **160**, 253-260.

Messing, J. (1979). A multipurpose cloning system on single-stranded DNA bacteriophage M13. *Recomb. DNA Tech. Bull*, **2**, 43.

Messing, J. (1983). New M13 vectors for cloning. *Methods Enzymol.*, **101**, 20-78.

Miki, B. L., and Neelin, J. M. (1975). The histones of rainbow trout erythrocytes include an erythrocyte-specific histone. *Can. J. Biochem.*, **53**, 1158-1169.

Mirkovitch, J., Mirault, M. E., and Laemmli, U. K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell*, **39**, 223-232.

Mueller, R. D., Yasuda, H., Hatch, C. L., Bonner, W. M., and Bradbury, E. M. (1985). Identification of ubiquitinated histones 2A and 2B in *Physarum polycephalum:* disappearance of these proteins at metaphase and reappearance at anaphase. *J. Biol. Chem.*, **260**, 5147-5153.

Mura, C. V., and Stollar, B. D. (1981). Serological detection of homologies of $H1^{o}$ with H5 and H1 histones. *J. Biol. Chem.*, **256**, 9767-9769.

Neelin, J. M., Callahan, P. X., Lamb, D. C., and Murray, K. (1964). The histones of chicken erythrocyte nuclei. *Can. J. Biochem.*, **42**, 1743-1752.

Newrock, K. M., Cohen, L. H., Hendricks, M. B., Donnelly, R. J., and Weinberg, E. S. (1978). Stage-specific mRNAs coding for subtypes of H2A and H2B histones in the sea urchin embryo. *Cell*, **14**, 327-336.

Nickel, B. E., Allis, C. D., and Davie, J. R. (1989). Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochem.*, **28**, 958-963.

Old, R. W., and Woodland, H. R. (1984). Histone genes: not simple after all. *Cell*, **38**, 624-626.

Pantazis, P., and Bonner, W. M. (1981). Quantitative determination of histone modification. H2A acetylation and phosphorylation. *J. Biol. Chem.*, **256**, 4669-4675.

Panyim, S., and Chalkley, R. (1969). A new histone found only in mammalian tissues with little cell division. *Biochem. Biophys. Res. Commun.*, **37**, 1042-1049.

Paulson, J. R., and Laemmli, U. K. (1977). The structure of histone-depleted metaphase chromosomes. *Cell*, **12**, 817-828.

Pehrson, J. R., and Cole, R. D. (1981). Bovine $H1^{o}$ histone subfractions contain an invariant sequence which matches histone H5 rather than H1. *Biochem.*, **20**, 2298-2301.

Pfaffle, P., Gerlach, V., Bunzel, L., and Jackson, V. (1990). *In vitro* evidence that transcription-induced stress causes nucleosome dissolution and regeneration. *J. Biol. Chem.*, **265**, 16830-16840.

Philips, T. A., VanBogelen, R. A., and Neidhardt, F. C. (1984). *lon* gene product of *Escherichia coli* is a heat-shock protein. *J. Bacteriol.*, **159**, 283-287.

Pina, B., Martinez, P., Simon, L., and Suau, P. (1984). Differential kinetics of histone H1 degrees accumulation in neuronal and glial cells from rat cerebral cortex during postnatal development. *Biochem. Biophys. Res. Commun.*, **123**, 697-702.

Ramírez-Solis, R., Davis, A. C., and Bradley, A. (1993). Gene Targeting in Embryonic Stem Cells. *Methods Enzymol.*, **225**, 855-878.

Reed, K. C., and Mann, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids. Res.*, **13**, 7207-7221.

Reeves, R. (1988). Active chromatin structure. In *Chromosome and chromatin*, vol 1, ed. K. W. Adolph, pp. 109-131. Boca Raton, Florida: CRC Press.

Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature*, **311**, 532-537.

Ridsdale, J. A., Hendzel, M. J., Delcuve, G. P., and Davie, J. R. (1990). Histone acetylation alters the capacity of the H1 histones to condense transcriptionally active/competent chromatin. *J. Biol. Chem.*, **265**, 5150-5156.

Robertson, E. J. (1987). Embryo-derived stem cell lines. In *Teratocarcinomas and embryonic stem cells: A Practical Approach*, ed. E. J. Robertson, pp. 71-112. Oxford: IRL Press.

Roche, J., Gorka, C., Goeltz, P., and Lawrence, J. J. (1985). Association of histone $H1^{o}$ with a gene repressed during liver development. *Nature*, **314**, 197-198.

Romano, G. (1992). Histone variants during sea urchin development. *Cell. Biol. Int. Rep.*, **16**, 197-206.

Rousseau, D., Khochbin, S., Gorka, C., and Lawrence, J.-J. (1991). Regulation of histone $H1^{o}$ accumulation during induced differentiation of murine erythroleukemia cells. *J. Mol. Biol.*, **217**, 85-92.

Saiki, R. K., Scharf, S. J., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985). Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350-1354.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (second ed.). New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Schild, C., Claret, F. X., Wahli, W., and Wolffe, A. P. (1993). A nucleosome-dependent static loop potentiates estrogen-regulated transcription from the *Xenopus* vitellogenin B1 promoter in vitro. *EMBO J.*, **12**, 423-433.

Schümperli, D. (1988). Multilevel regulation of replication-dependent histone genes. *Trends Genet.*, **4**, 187-191.

Schwartzberg, P. L., Robertson, E. J., and Goff, S. P. (1990). Targeted gene disruption of the endogenous *c-abl* locus by homologous recombination with DNA encoding a selectable fusion protein. *Proc. Natl. Acad. Sci. USA*, **87**, 3210-3214.

Sealy, L., and Chalkley, R. (1978). The effect of sodium butyrate on histone modification. *Cell*, **14**, 115-121.

Simpson, R. T. (1978) Structure of chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochem., 17, 5524-5531.

Simpson, R. T. (1991). Nucleosome positioning: occurrence, mechanisms, and functional consequences. *Prog. Nucl. Acid Res. Mol. Biol.*, **40**, 143-184.

Smith, A. G., and Hooper, M. L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev. Biol.*, **121**, 1-9.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, **336**, 688-690.

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., and Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature*, **317**, 230-234.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, **98**, 503-517.

Stargell, L. A., Bowen, J., Dadd, C. A., Dedon, P. C., Davis, M., Cook, R. G., Allis, C. D., and Gorovsky, M. A. (1993). Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila. *Genes Dev.*, **7**, 2641-2651.

Staynov, D. Z., and Crane-Robinson, C. (1988). Footprinting of linker histones H5 and H1 on the nucleosome. *EMBO J.*, **7**, 3685-3691.

Stein, G. S., Stein, J. L., and Marzluff, W. F. (1984). *Histone Genes: Structure, Organization and Regulation.* New York: John Wiley and Sons.

Sterner, R., Ferrari, N., and Vidali, G. (1989). Posttranslational Modifications of Histones. In *Histones and Other Basic Nuclear Proteins*, eds. L. S. Hnilica, G. S. Stein, and J. L. Stein, pp. 197-250. Boca Raton, Florida: CRC Press.

Straka, C., and Hörz, W. (1991). A functional role for nucleosomes in the repression of a yeast promoter. *EMBO J.*, **10**, 361-368.

Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **189**, 113-130.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60-89.

Sun, J. M., Ali, Z., Lurz, R., and Ruiz-Carrillo, A. (1990). Replacement of histone H1 by H5 in vivo does not change the nucleosome repeat length of chromatin but increases its stability. *EMBO J.*, **9**, 1651-1658.

Sun, J. M., Wiaderkiewicz, R., and Ruiz-Carrillo, A. (1989). Histone H5 in the control of DNA synthesis and cell proliferation. *Science*, **245**, 68-71.

Svaren, J., and Chalkley, R. (1990). The structure and assembly of active chromatin. *Trends Genet.*, **6**, 52-56.

Svaren, J., and Hörz, W. (1993). Histones, nucleosomes and transcription. *Curr. Opin. Genet. Dev.*, **3**, 219-225.

Tate, P. H., and Bird, A. P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr. Opin. Genet. Dev.*, **3**, 226-231.

te Riele, H., Maandag, E. R., and Berns, A. (1992). Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Natl. Acad. Sci. USA*, **89**, 5128-5132.

te Riele, H., Maandag, E. R., Clarke, A., Hooper, M., and Berns, A. (1990). Consecutive inactivation of both alleles of the *pim-1* proto-oncogene by homologous recombination in embryonic stem cells. *Nature*, **348**, 649-651.

Thatcher, T. H., and Gorovsky, M. A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res.*, **22**, 174-179.

Thoma, F., Koller, T., and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.*, **83**, 403-427.

Thomas, K. R., and Capecchi, M. R. (1986). Targeting of genes to specific sites in the mammalian genome. *Cold Spring Harbor Symp. Quant. Biol.*, **51**, 1101-1113.

Thomas, K. R., and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*, **51**, 503-512.

Thomas, K. R., and Capecchi, M. R. (1990). Targeted disruption of the murine *int-1* proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. *Nature*, **346**, 847-850.

van Daal, A., and Elgin, S. C. R. (1992). A histone variant, H2AvD, is essential in *Drosophila melanogaster*. *Mol. Biol. Cell*, **3**, 593-602.

van Daal, A., White, E. M., Elgin, S. C. R., and Gorovsky, M. A. (1990). Conservation of intron position indicates separation of major and variant H2As is an early event in the evolution of eukaryotes. *J. Mol. Evol.*, **30**, 449-455.

van Daal, A., White, E. M., Gorovsky, M. A., and Elgin, S. C. R. (1988). *Drosophila* has a single copy of the gene encoding a highly conserved histone H2A variant of the H2A.F/Z type. *Nucleic Acids Res.*, **16**, 7487-7497.

van Deursen, J., and Wieringa, B. (1992). Targeting of the creatine kinase M gene in embryonic stem cells using isogenic and nonisogenic vectors. *Nucleic Acids Res.*, **20**, 3815-3820.

van Holde, K. E. (1988). *Chromatin.*, ed Rich A., New York: Springer-Verlag.

van Holde, K., and Zlatanova, J. (1994). Unusual DNA structures, chromatin and transcription. *Bioessays*, **16**, 59-68.

Varricchio, F. (1977). Postnatal increase in histone H1a in the rat pancreas. *Arch. Biochem. Biophys.*, **179**, 715-717.

Vogelstein, B., and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. USA*, **76**, 615-619.

Wait, P. (1990) *Isolation of the mouse H2A.Z gene for targeting by homologous recombination.* Hons. Thesis, University of Adelaide, Adelaide, South Australia.

Watt, F., and Molloy, P. L. (1988) Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for the optimal expression of the adenoviral major late promoter. *Genes Dev.*, **2**, 1136-1143.

Weintraub, H. (1978). The nucleosome repeat length increases during erythropoiesis in the chick. *Nucleic Acids Res.*, **5**, 1179-1188.

West, M. H. P. , and Bonner, W. M. (1980). Histone 2A, a heteromorphous family of eight protein species. *Biochem.*, **19**, 3238-3245.

West, M. H. P., and Bonner, W. M. (1983). Structural comparisons of mouse histones 2A.X and 2A.Z with 2A.1 and 2A.2. *Comp. Biochem. Physiol.*, **76B**, 455-464.

White, E. M., Shapiro, D. L., Allis, C. D., and Gorovsky, M. A. (1988). Sequence and properties of the message encoding *Tetrahymena* hv1, a highly evolutionarily conserved histone H2A variant that is associated with active genes. *Nucleic Acids Res.*, **16**, 179-198.

Whiting, J. A. (1988) *Structure and Distribution of Chicken Histone H2A.F.*, Ph.D. Thesis, The University of Adelaide, Adelaide, South Australia.

Whitlock, J. P. J., and Stein, A. (1978). Folding of DNA by histones which lack their NH2-terminal regions. *J. Biol. Chem.*, **253**, 3857-3861.

Wicks, R. J. (1986). RNA molecular weight determination by agarose gel electrophoresis using formaldehyde as denaturant: comparison of RNA and DNA molecular weight markers. *Int. J. Biochem.*, **18**, 277-278.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, **336**, 684-687.

Wolffe, A. P. (1992). New insights into chromatin function in transcriptional control. *FASEB J.*, **6**, 3354-3361.

Wolffe, A. P. (1994). Nucleosome positioning and modification: chromatin structures that potentiate transcription. *Trends Biochem. Sci.*, **19**, 240-244.

Wood, W. I., Gitschier, J., Lasky, L. A., and Lawn, R. M. (1985). Base composition-independent hybridization in tetra-mrthylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA*, **82**, 1585-1588.

Wu, R. S., and Bonner, W. M. (1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. *Cell*, **27**, 321-330.

Wu, R. S., Panusz, H. T., Hatch, C. L., and Bonner, W. M. (1986). Histones and their modifications. *CRC Crit. Rev. Biochem.*, **20**, 201-263.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103-119.

Zimmer A. (1992) Manipulating the genome by homologous recombination in embryonic stem cells. *Annu. Rev. Neurosci.*, **15**, 115-37.

Zimmer, A., and Gruss, P. (1989). Production of chimaeric mice containing embryonic stem (ES) cells carrying a homoeobox *Hox 1.1* allele mutated by homologous recombination. *Nature*, **338**, 150-153.

Zlatanova, J. S., and van Holde, K. E. (1992). Chromatin loops and transcriptional regulation. *Crit. Rev. Euk. Gene Expr.*, **2**, 211-24.

Zoller, M. J., and Smith, M. (1983). Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods Enzymol.*, **100**, 468-500.

Zweidler, A. (1984). Core histone variants of the mouse: primary structure and differential expression. In *Histone Genes: Structure, Organization and Regulation*, eds. G. S. Stein, J. L. Stein, and W. F. Marzluff, pp. 339-371. New York: John Wiley and Sons.