THE ROLE OF DNA LIGASE I IN MOUSE DEVELOPMENT

DARREN JAMES BENTLEY

Thesis Presented for the Degree of Doctor of Philosophy

Institute of Cell and Molecular Biology
University of Edinburgh

April 1999

DECLARATION

The composition of this thesis, and the work presented in it are my own, unless otherwise stated. The experiments were designed in collaboration with my supervisor Dr. David W. Melton.

[redacted]

Darren James Bentley
April 1999

CONTENTS

ABSTRACT

Four distinct DNA ligase activities (I-IV) have been identified within mammalian cells. Evidence has indicated DNA ligase I to be central to lagging strand DNA replication, as well as being involved in DNA repair processes. It has also been reported that DNA ligase I is essential for cell viability. A patient with altered DNA ligase I displayed a phenotype similar to Bloom's Syndrome, being immunodeficient, growth retarded, and predisposed to cancer. Fibroblasts isolated from this patient (46BR) exhibited abnormal lagging strand DNA synthesis, and repair deficiency. Gene inactivation was the first step of a 'double replacement' strategy to introduce analogous mutations into the mouse DNA ligase I gene (*Lig1*). Using gene targeting in HPRT-deficient embryonic stem (ES) cells, the last 5 exons of the endogenous *Lig1* gene were replaced by an *HPRT* minigene, and subsequently DNA ligase I-deficient mice were produced.

Embryos lacking DNA ligase I developed normally to mid-term, when haematopoiesis usually switches to the foetal liver. Thereupon severe disruption to foetal liver erythropoiesis occurred, leading to acute anaemia and prenatal death. *In vitro* assays indicated that erythroid-committed progenitor cells were present in the liver, but in reduced numbers. Apart from the developing liver, organogenesis was not perceptibly affected and non-erythroid haematopoietic lineages were not reduced.

Injection of cell suspensions from single foetal livers were unable to rescue lethally-irradiated mice. However, injection of multiple livers into a single recipient was able to effect long term repopulation of the haematopoietic system. Animals rescued with cells lacking DNA ligase I displayed splenomegaly, anaemia coupled with reticulocytosis, and hypoplasia in all haematopoietic tissues.

Fibroblasts isolated from embryos lacking DNA ligase I displayed an abnormal profile of replication intermediates, but showed no increase in the level of sister chromatid exchanges. Northern and western analysis demonstrated that DNA ligase I expression was completely abrogated, but that expression of DNA ligases III and IV remained unaltered.

A second round of gene targeting in ES cells was used to introduce two flanking *loxP* sites into the endogenous *Lig1* locus, as the basis of a 'conditional knockout' strategy. Mice heterozygous for the 'floxed' *Lig1* allele were crossed with mice carrying a mammary-specific Cre recombinase transgene, and their offspring were analysed to assess the efficacy of Cre-mediated recombination *in vivo*. Deletion of the 'floxed' *Lig1* allele was detected at high levels in the mammary gland of lactating females.

Although DNA ligase I is required for normal development, we conclude that it is not essential for cell viability, or DNA replication, and the absence of DNA ligase I has no discernible consequence for most embryonic tissues. We propose that the phenotype observed reflects the particularly high replicative requirements during foetal liver erythropoiesis and *in vivo* haematopoietic repopulation. Exactly which other ligase activity replaces DNA ligase I still remains unclear, but a previously unsuspected redundancy must exist among the mammalian DNA ligases.

ACKNOWLEDGEMENTS

I am deeply indebted to a large number of people who have contributed to this project in a variety of ways. Wherever possible I have tried to give credit at the relevant point in the text, but I will repeat them here and elaborate.

Firstly thanks must go to my predecessor Kirsty Jessop, now Kirsty Millar. Kirsty initially cloned the *Lig1* cDNA and genomic fragments, did the mapping and generally laid the groundwork for the project. This included providing an extremely thorough thesis which has been an invaluable "reference" text. Carolanne McEwan constructed the targeting vectors, sequenced the 3' end of the *Lig1* locus, and has genotyped a number of the floxed mice, as well as looking after me when I first arrived in the lab. Angela Pow and Ann-Marie Ketchen taught me cell culture (the way it *should* be done), and have run a highly efficient tissue culture support system. Furthermore they did all the ES cell work involved during the two targeting steps, cultured primary embryonic fibroblasts, and subsequently established and maintained the immortalised fibroblast lines. Following Angela's departure, Ann-Marie was joined by Natalie Wilson, and both continue to sterling work, including being able to provide me with flasks of cells at a moments notice.

Jim Selfridge has been central in the smooth progression of the project over the years, and it is impossible to overstate the size of his influence (allegedly). He did all the blastocyst injections, the initial molecular biology, lots of animal husbandry and numerous other onerous tasks along the way. He guided me through almost all of the techniques I have learned, even managing to get me to isolate "snoopy dogs" instead of placentas. Through all of this he has been extremely helpful, kind, and dare I say funny, always ready with a smutty joke or innuendo. Tracey White and Caroline Holmes were also involved in the maintenance of stocks.

Our collaborators in ICAPB Kay Samuel, Helen Taylor, Nick Hole and John Ansell were vital to all of the immunology work done. I am especially grateful to Kay for her help with all of the practical aspects, and the FACS analysis in particular. In respect of the latter, a mention must also go to Martin Waterfall for his assistance in finally getting the data off the machine. John Lauder did the histology sectioning, and David Harrison kindly cast his expert eye over the sections. Alan Clarke provided the Cre transgenic mice that he had produced along with Christine Watson, Stefan Selbert, Dominic Rannie and Paula Lourenço.

Graham, Frank, Dave, and John in photography must be sick of the sight of me by now, but without their repeated efforts the pictures presented here would be a lot less pretty. I am still amazed at the job they did to make photographs of SCEs

actually visible, and am highly grateful. In addition, I would also like to acknowledge Jean and Joan for their unstinting services in the provision clean glassware.

Thanks must go to the all members of the Melton lab (past and present). Those not mentioned previously are: Daniella, Fatima, Kan-Tai, Liz, Niki, Rebecca, Richard, Sabine, and Thomas. Everyone has contributed to make my three (and a bit) year stay in Edinburgh an enjoyable one, and I will miss working with them. Life will be less exciting without the daily sparring with Richard, and I am thankful to him for showing me numerous shortcuts, and also the way not to do certain things. Niki has taken over the project, so she will be kept busy for the foreseeable future.

The one person left to pay tribute to is my supervisor David Melton, without whom none of this would have been possible. He has been extremely generous to me in so many ways, and his expert guidance has kept me on the straight and narrow. I feel extremely lucky to have worked with him, and will be eternally grateful.

Finally, I must mention Craig, Jo, Anthony and Theo who have all been admirable drinking companions. However, I would not have made it through the darkest hours without the support of my parents Eric and Rosemary, brother Miles, and especially my partner Simone. Simone is the best thing to have come out of my time in Edinburgh, and without her things not be the same.

ABBREVIATIONS

'	(prime) denotes a gene truncated at the indicated side
3AB	3-aminobenzamide
6-TG	6-thioguanine
A	adenosine
A	amps
ADP	adenosine diphosphate
AMP	adenosine monophosphate
AP	apurinic/apyrimidinic
Arg	arginine
AT	ataxia telangiectasia
ATP	adenosine triphosphate
BAP	bacterial alkaline phosphatase
BFU-E	blast forming unit-erythroid
BLG	β-lactoglobulin
bp	base pair(s)
BrdU	5-bromodeoxyuridine
BS	Bloom's syndrome
BSA	bovine serum albumin
c	centi-
C	cytidine
C-	carboxyl-
cDNA	DNA complementary to RNA
CFU	colony forming unit
CS	Cockaynes syndrome
d	deoxyribo
DBA	Diamond-Blackfan anaemia
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	DL-dithiothreitol
E	day of embryonic development
EDTA	ethanediaminotetraacetic acid
EMS	ethylmethane sulphonate
ENU	*N*-ethyl-*N*-nitrosourea
Epo	erythropoietin

ERCC	excision repair cross-complementing
ES cells	embryonic stem cells
FA	Fanconi anaemia
FACC	Fanconi anaemia cross-complementing
FACScan	fluorescence activated cell scanning
FCS	foetal calf serum
FITC	fluorescein isothiocyanate
g	gram(s)
G	guanosine
Glu	glutamate
HPRT	hypoxanthine -guanine phosphoribosyl transferase
HSC	haematopoietic stem cell
HSV	Herpes simplex virus
Ig	immunoglobulin
IL-3	interleukin-3
IPTG	isopropyl-ß-D-thiogalactopyranoside
l	litres
LB	Luria broth
LPS	lipopolysaccharide
Lys	lysine
m	metre(s)
μ	micro-
m	milli-
M	molar
MLTV	mouse DNA ligase I targeting vector
MNNG	*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine
MNU	*N*-methyl-*N*-nitrosourea
MOPS	3-(*N*-morpholino) propane-sulphonic acid
N	any nucleoside
n	nano-
N-	amino-
NAD	nicotinamide adenine dinucleotide
NADH	reduced form of NAD
NCS	new-born calf serum
nt	nucleotide(s) (number of sequence)
OD	optical density
p	pico-
p	plasmid

p	chromosome short arm
PARP	Poly(ADP-ribose) polymerase
PCR	polymerase chain reaction
PE	phycoerythrin
PHA	phytohaemagglutinin
P_i	inorganic phosphate
PMSF	phenylmethylsulphonyl fluoride
PP_i	inorganic pyrophosphate
q	chromosome long arm
r	(superscript) resistance/resistant
r	ribo
rpm	revolutions per minute
RF	replicative form
RNA	ribonucleic acid
RNase	ribonuclease
S	sedimentation constant
SCE	sister chromatid exchange
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEB	*Staphlococcus* enterotoxin B
Ser	serine
ss	single-stranded
SV40	Simian virus 40
T	Thymidine
TB	terrific broth
TEMED	N,N,N',N'-tetramethylethylenediamine
tk	thymidine kinase
Trp	tryptophan
UTR	untranslated region
UV	ultraviolet
V	volts
W	watts
X-Gal	5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside
XP	xeroderma pigmentosum

Single letter amino acid abbreviations:

A	alanine
C	cysteine
D	aspartate
E	glutamate
F	phenylalanine
G	glycine
H	histidine
I	isoleucine
K	lysine
L	leucine
M	methionine
N	asparagine
P	proline
Q	glutamine
R	arginine
S	serine
T	threonine
W	tryptophan
Y	tyrosine

CHAPTER ONE:

INTRODUCTION

1.1 Foreword

In order for data to be faithfully preserved, the medium on which it its stored has to be stable, and genetic information is no different in this respect. DNA, the genetic material of almost all organisms, must be maintained in an informationally active form in spite of any molecular changes. Modification to the structure of DNA can arise spontaneously, from normal cellular processes, or as a result of the action of environmental agents. Whatever the cause, damage that remains unrepaired has the potential to alter the information encoded by the DNA, or result in the loss of information. Such changes can have a variety of consequences, but most will be deleterious to an organism's survival or reproductive capacity. In response to this, living organisms have evolved a number of mechanisms to maintain genomic stability.

As had been true for all areas of molecular biology, study of model prokaryotic organisms, such as *Escherichia coli*, provided a powerful tool for studying the cellular mechanisms involved in the repair or tolerance of DNA damage. However, whilst many pathways are conserved across evolutionary time, the eukaryotic organism is far more complex. Not only do eukaryotic cells contain far larger genomes and multiple chromosomes, but they also possess a higher order DNA structure not present in prokaryotes. Furthermore the development of multicellular organisation requires tight control of cell division and differentiation. Accumulation of DNA damage is known to disrupt and deregulate both of these processes. The challenge is therefore to understand how DNA repair and damage tolerance mechanisms act in the context of the mammalian system.

1.2 DNA repair mechanisms

DNA damage can be divided into two classes on the basis of origin. "Spontaneous" damage arises as the result of normal cellular processes such as DNA replication, while the "environmental" damage is produced by exogenous agents. The variety of spontaneous alterations and environmental damaging agents are reviewed in Friedberg , Walker and Siede 1995, but are summarised in Table 1.2.1 and Table 1.2.2 respectively. It can be seen from these tables that the total number of different lesions is large. Rather than maintaining numerous repair pathways specific for individual damage products, eukaryotic cells have evolved a limited number of general mechanisms which are each able to deal with multiple types of lesion. Although thorough discussion of all cellular responses to DNA damage is beyond the scope of this thesis, the basic DNA repair pathways operating in mammalian cells will be described briefly.

Table 1.2.1: Types of DNA lesions arising spontaneously from cellular processes

Type of lesion	Cause
Mismatched base	Polymerase infidelity during DNA replication
Misincorporation in daughter strand	Spontaneous alteration to chemistry of DNA bases. e.g. tautomeric shift, deamination of bases
Incorporation of uracil	Hydrolysis of cytosine, perturbation of TMP biosynthesis
Depurination and depyrimidination	Spontaneous hydrolysis
Oxidative damage	Attack by reactive oxygen species (free radicals)

Table 1.2.2: Environmental DNA damaging agents

Agent	DNA damage caused
Ionising radiation	Oxidative damage by radicals, base damage, sugar damage, single and double strand breaks
UV radiation	Cyclobutane pyrimidine dimers, pyrimidine-pyrimidone (6-4) photoproducts, other photoproducts, DNA cross-links, strand breaks
Alkylating agents	Alkylation of the phosphodiester backbone and bases
Cross-linking agents	Interstrand, intrastrand and DNA-protein cross-links
Psoralens (intercalating agents)	Interstrand cross-links, helix distortion, kinking, and unwinding
Base analogues	Misincorporated bases following replication
Electrophilic reactants	Various, depending on agent

1.2.1 Direct reversal of DNA damage

Mechanisms which involve enzyme-catalysed return of damaged DNA directly to its original state are the simplest example of DNA repair. The classic example of direct reversal is the monomerisation of *cis-syn* cyclobutane pyrimidine dimers caused by UV irradiation. Enzymatic photoreactivation by DNA photolyase has been extensively studied in *E. coli*, but there is no definitive evidence that such a mechanism exists in higher eukaryotes. Similarly the rejoining of single strand DNA breaks by DNA ligase has been observed in *E. coli*, but the relative importance of this mechanism is not certain in mammalian cells. Ligation requires free ends of duplex DNA with 3'OH and 5'P termini. In most cases radiation induced single strand breaks are accompanied by damage to strand ends, and processing by other enzymes has to occur before ligation.

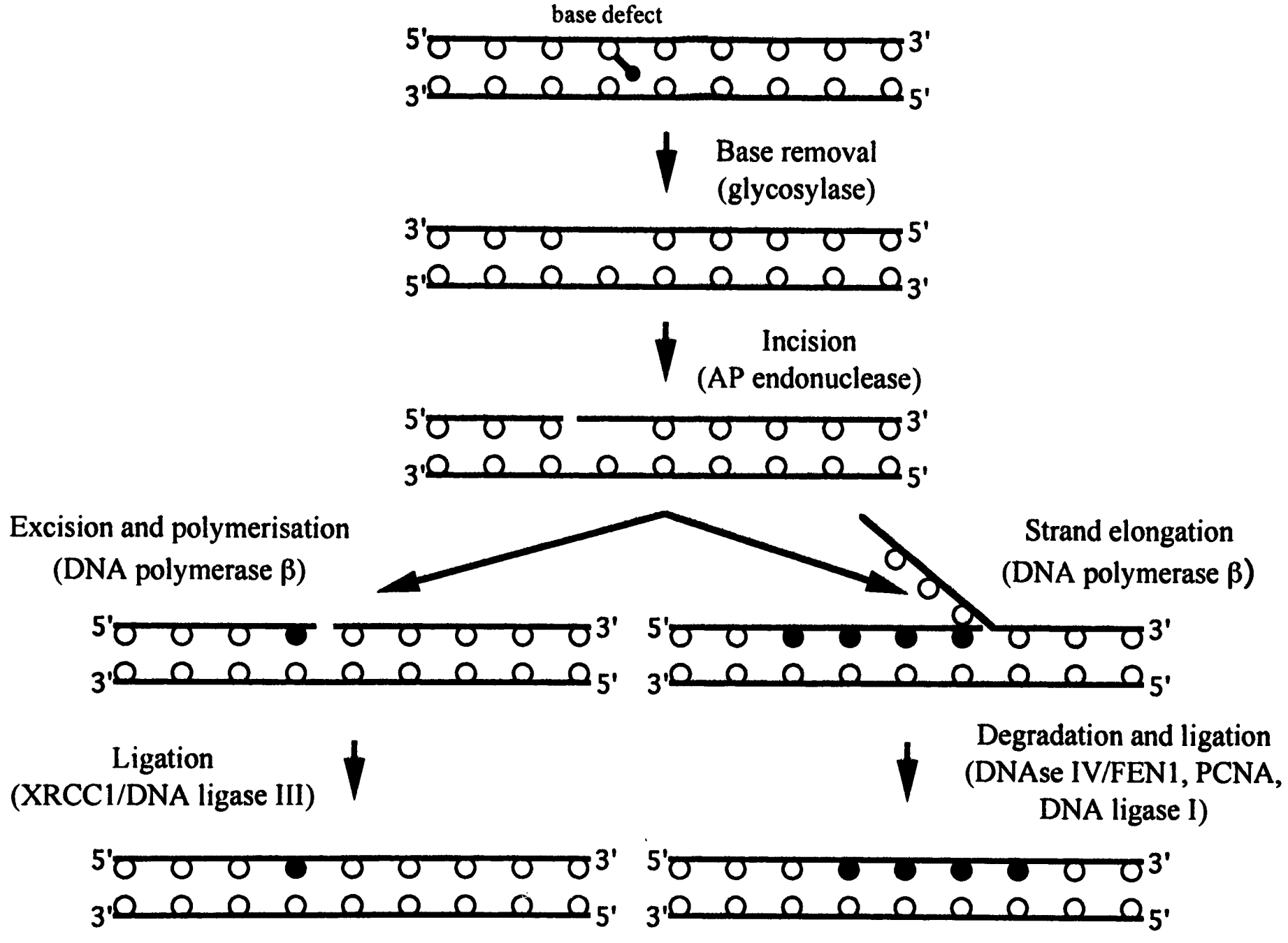
Repair of alkylation damage by direct means has been observed in mammalian cells. O^6-alkylguanine is a major product of the reaction between DNA and alkylating agents such as methylnitro-nitrosoguanidine (MNNG) and methyl-nitrosourea (MNU). O^6-alkylguanine lesions are thought to be highly mutagenic, and are removed by O^6-methylguanine transferase (O^6-MGT). The methyl group is transferred from the DNA to a cysteine residue on the enzyme, forming S-methyl cysteine and resulting in permanent inactivation of the enzyme.

1.2.2 Base excision repair

The main strategy for correcting DNA lesions arising from spontaneous hydrolysis and attack by free radicals is base excision repair (BER) (reviewed by Lindahl 1995, and Lindahl, Karran and Wood 1997). The damaged base is recognised by a DNA glycosylase, which hydrolyses the glycosylic bond and releases the modified base, thereby generating an abasic site (see Figure 1.2.1). Thereafter the apurinic or apyrimidinic (AP) site is recognised by an AP endonuclease which hydrolyses the phosphodiester bond 5' of the AP site. Recent research has shown that subsequent processing can occur by either one of two possible pathways. It has been proposed that the majority of AP sites are filled by DNA polymerase β which adds one nucleotide to the 3' terminus and excises the 5'-terminal deoxyribose phosphate. The two termini are joined together by the α form of DNA ligase III in a complex with XRCC1. Alternatively polymerase β (or polymerases δ or ε, and along with several accessory factors) can add several nucleotide residues, displacing the parental strand in the process. Subsequently the 'overhang' is removed by the structure-specific nuclease DNase IV/FEN1, and repair is completed by DNA ligase I.

Figure 1.2.1: Schematic representation of possible BER strategies

A model for base excision repair, adapted from Lindahl, Karran and Wood 1997. The sugar-phosphate backbone of DNA is depicted by lines, while bases are represented by circles. Open circles indicate bases present prior to repair, and solid circles bases introduced by the repair process. See text for details.


base defect
5'
3'
3'
5'
Base removal
(glycosylase)
3'
5'
5'
3'
Incision
(AP endonuclease)
5'
3'
3'
5'
Excision and polymerisation
(DNA polymerase β)
Strand elongation
(DNA polymerase β)
Ligation
(XRCC1/DNA ligase III)
Degradation and ligation
(DNAse IV/FEN1, PCNA,
DNA ligase I)

1.2.3 Nucleotide excision repair

Nucleotide excision repair (NER) is the most flexible of the repair pathways as it can act on a wide variety of lesions. DNA damage is recognised by local distortion of the normal helical structure of DNA, and as such any alteration that causes distortion can potentially be repaired by this mechanism. Despite this broad specificity, the main function of NER is to remove photoproducts arising from exposure of cells to UV radiation (reviewed by Lindahl, Karran and Wood 1997).

Helix-altering lesions are recognised by a complex of XPA and RPA proteins. Thereafter the transcription factor TFIIH is recruited. The helicase activity of TFIIH allows the region around the damage to be unwound ('opened up') in an ATP-dependent manner, whereupon the structure-specific nucleases ERCC1-XPF and XPG make dual incisions on the damaged strand. An oligonucleotide 24-32 residues long containing the damaged nucleotide is released, and repair synthesis occurs to fill the resulting gap. Finally the repair patch is ligated to the pre-existing strand by a DNA ligase (possibly DNA ligase I) (Figure 1.2.2).

The products of up to 30 genes could be involved in the pathway in mammalian cells (reviewed in Lehmann 1995), but a number of these do not seem to be required for the basic repair reaction. The precise role of each protein remains to be delineated, but some certainly act to refine the function of the basic NER machinery (Friedberg, Walker and Seide 1995, and references therein). The organisation of DNA into chromatin has been shown to impair NER efficiency, and it is probable that NER is linked to nucleosome rearrangement. It is known that the efficiency of NER varies across the genome. Repetitive DNA is repaired more slowly than the rest of the genome, and repair synthesis appears to localise preferentially to sites of DNA replication and transcription. Comparison of expressed genes to bulk genomic DNA revealed that actively transcribed sequences are repaired preferentially (transcription-dependent repair). Further examination showed that the coding strand of RNA polymerase II-transcribed genes was also repaired preferentially to the non-transcribed strand (transcription-coupled repair). The mechanism of strand-specific repair remains to be understood, but most probably is brought about through the involvement of TFIIH in both NER and RNA polymerase II-dependent transcription.

1.2.4 Mismatch repair

Although DNA polymerases possess a proof-reading activity, mispairing of bases during DNA replication (replication errors) still remains the main source of alterations arising from normal DNA metabolism. Repair of such mismatches in

Figure 1.2.2: Schematic representation of the basic NER reaction in mammalian cells

A model for nucleotide excision repair, adapted from Lindahl, Karran and Wood 1997. The sugar-phosphate backbone of DNA is depicted by lines, while bases are represented by open circles. Open circles indicate bases present prior to repair, and solid circles bases introduced by the repair process. Proteins are shown as shaded objects. See text for details.

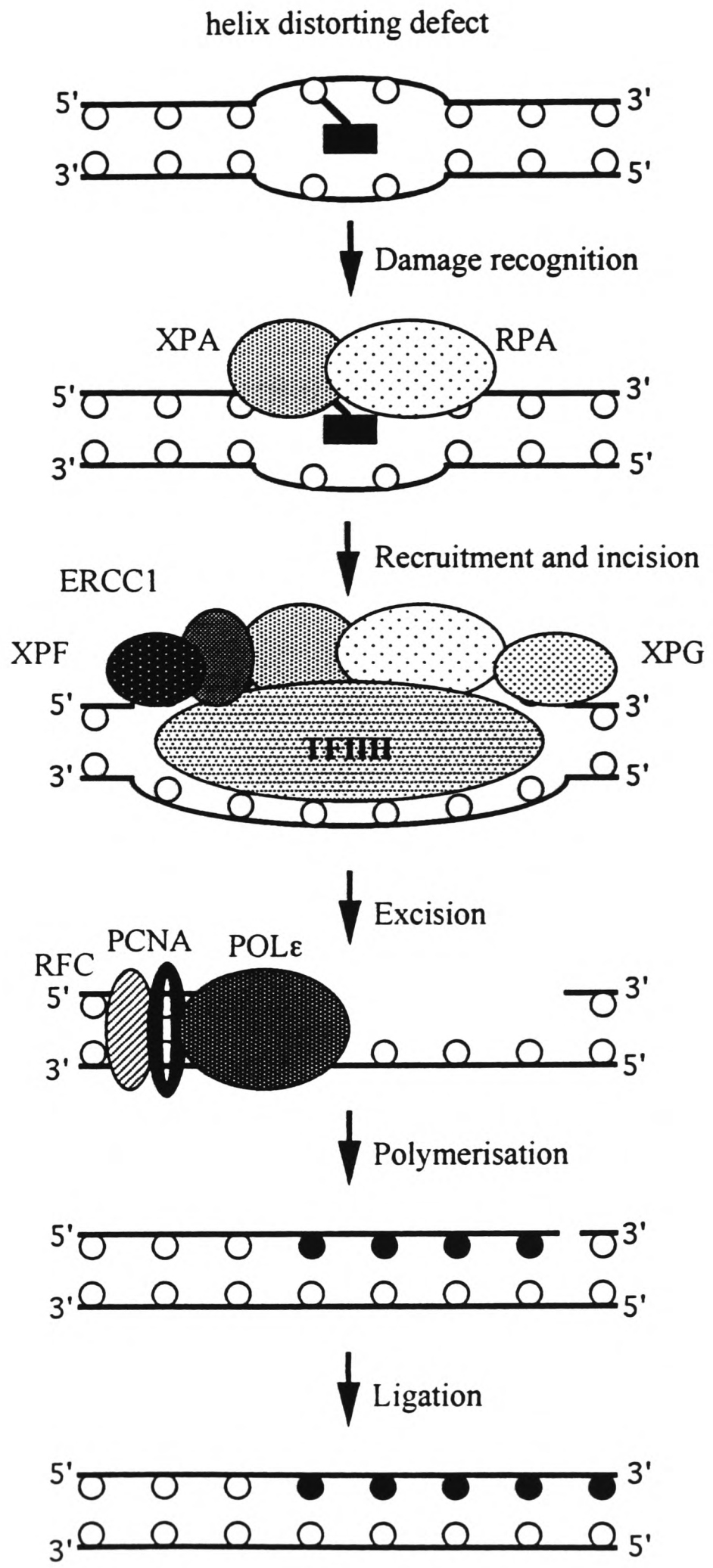
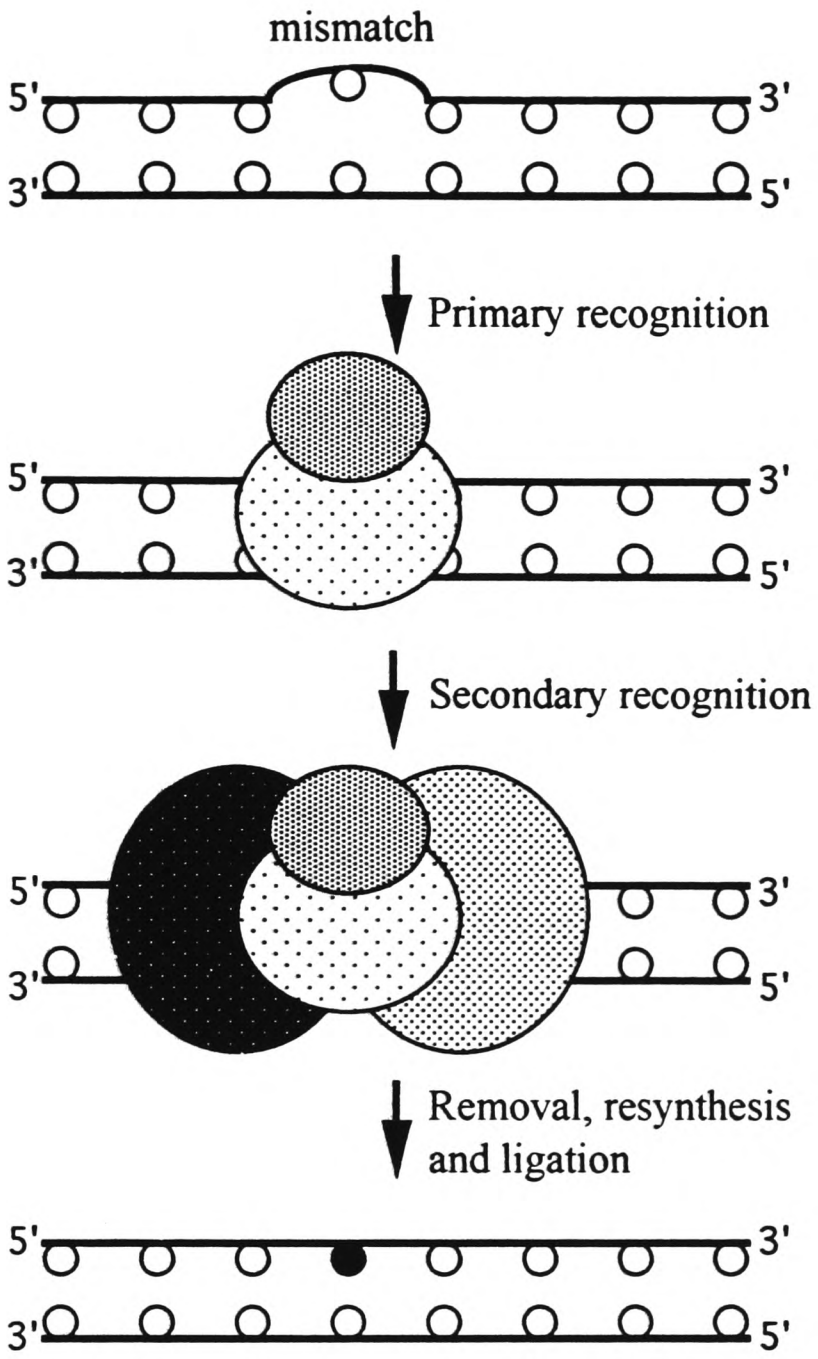

helix distorting defect
5'
3'
3'
5'
Damage recognition
XPA
RPA
5'
3'
3'
5'
Recruitment and incision
ERCC1
XPF
XPG
5'
3'
TFIIH
3'
5'
Excision
PCNA
POLε
RFC
5'
3'
3'
5'
Polymerisation
5'
3'
3'
5'
Ligation
5'
3'
3'
5'

Figure 1.2.3: Schematic representation of mismatch repair in human cells

A model for mismatch repair, adapted from Lindahl, Karran and Wood 1997. The sugar-phosphate backbone of DNA is depicted by lines, while bases are represented by circles. Open circles indicate bases present prior to repair, and solid circles bases introduced by the repair process. Proteins are shown as shaded objects. See text for details.


mismatch
5'
3'
3'
5'
Primary recognition
5'
3'
3'
5'
Secondary recognition
5'
3'
3'
5'
Removal, resynthesis
and ligation
5'
3'
3'
5'

human cells is achieved via a mechanism which closely parallels that present in *E. coli* (reviewed in Lindahl 1994). Recognition of mismatches is carried out by heterodimers of three homologues of the *E. coli* MutS protein. The protein hMSH2 can complex with either hMSH6(GTBP) or hMSH3 to form hMutSα and hMutSβ respectively. hMutSα preferentially recognises single base mispairs, single base loops and two base loops in repeated sequences, while hMutSβ favours binding to larger loops (three or four bases). The bound protein-DNA complexes then recruit hMutLα (secondary recognition). hMutLα is also a heterodimer, composed of two homologues of the *E. coli* MutL protein (hMLH1 and hPMS2), and this complex acts as an adapter molecule to initiate repair. The details of this process remain unclear, but it involves excision of a long patch of DNA from the mismatched strand, resynthesis of the oligonucleotide and finally ligation (Figure 1.2.3).

1.2.5 DNA strand break repair

Double-strand breaks (DSBs) can be created by ionising radiation or during recombination reactions, and are lethal to eukaryotic cells. It has been calculated that a single double-strand break within its genome is sufficient to kill a yeast cell (Frankenberg *et al.* 1981). In order to repair DSBs, eukaryotic cells employ two mechanisms. Lower eukaryotes, including yeasts, preferentially repair broken ends by homologous recombination, whereas higher eukaryotes predominantly use processes that do not rely on homology (illegitimate/non-homologous end joining).

1.2.5.1 Homologous recombination

Homologous recombination in *Saccharomyces cerevisiae* is mediated by genes in the *RAD52* epistasis group (*RAD50*, *RAD51*, *RAD52*, *RAD53*, *RAD54*) (Friedberg, Walker and Siede 1995). Evidence has accumulated for the functional conservation of the RAD52 pathway in mammals (Essers *et al.* 1997 and references therein). It is conceivable that one of the proposed models of DSB repair in *E. coli* will be applicable to mammalian cells. Single-stranded DNA from the damaged chromosome invades a homologous DNA duplex. Following initiation of strand exchange, DNA is synthesised from the damaged ends using the intact strands as templates. Thereupon the two Holliday junctions which have been formed are resolved by cleavage to separate the two duplexes. Depending on the orientation of the cleavage events, the resulting DNA molecules may or may not exchange flanking markers (crossing over).

1.2.5.2 Non-homologous end joining (NHEJ)

In addition to homologous recombination, DSBs can also be repaired by a second independent pathway, termed non-homologous end joining (NHEJ). Although

it plays an auxiliary role in wild-type yeast cells, NHEJ is the predominant DSB repair mechanism in mammalian cells. Study of ionising radiation-sensitive cell lines has revealed that the repair of DSBs employs a number of proteins also involved in V(D)J recombination. V(D)J recombination is a site-specific recombination event which occurs in developing B lymphocytes. It is a programmed DNA rearrangement to generate a diverse repertoire of immunoglobulin genes from germline gene segments, and brings together variable (V), diversity (D) and joining (J) subexonic gene segments in various combinations. A similar mechanism is also believed to generate diversity in T-cell receptor (TCR) genes in T lymphocytes. The action of DNA dependent protein kinase (DNA-PK) is now known to be central to both V(D)J recombination and NHEJ (reviewed in Bogue and Roth 1996).

DNA-PK is a nuclear protein-serine/threonine kinase that is activated by binding to DNA ends (Jackson and Jeggo 1995 and references therein). It is a multiprotein complex comprising of a large catalytic subunit (DNA-PK_{CS}) and a DNA binding subunit (Ku). Ku is itself a heterodimer of 70kDa and 80kDa proteins (Ku70 and Ku80 respectively). Ku binds to DNA ends and recruits the catalytic subunit, which subsequently becomes enzymatically active.

The precise chain of molecular events following DNA-PK binding to DNA remains unclear. A number of possible mechanisms have been proposed for DNA-PK action (summarised in Weaver 1995). Binding of the protein to DNA ends apparently acts to protect termini from nucleolytic degradation and promotes end joining. It is possible that DNA-PK may play an active role in the recruitment of enzymes to the site of damage, phosphorylating, and thereby activating, repair proteins. The kinase activity may also inhibit the transcriptional machinery and control the cell cycle, cellular processes that might otherwise interfere with repair. Alternatively the protein may provide a structural framework to juxtapose DNA ends and hold them together while interacting with other components of the repair machinery. It was recently reported that Ku stimulates ligation by DNA ligases I, III, and IV, and that Ku acts as a 'bridging' or 'alignment factor' between DNA molecules, stabilising intermolecular associations (Ramsden and Gellert 1998). Exactly which DNA ligase is involved will be discussed further in section 1.4.1.7, but existing evidence points towards DNA ligase IV.

1.3 Human diseases arising from genomic instability

Deficiency in the repair of DNA damage has been implicated in a disparate group of rare human diseases; ataxia telangiectasia, Bloom's syndrome, Cockayne's syndrome, Fanconi's anaemia, trichothiodystrophy, and xeroderma pigmentosum. All of these diseases are inherited in an autosomal recessive manner, and are characterised

by hypersensitivity to one or more genotoxic agents, increased mutation rates and (with exception of CS and TTD) a predisposition to cancer (reviewed in Friedberg, Walker and Siede 1995). Although it is now known that not all of the genes implicated are directly involved in the repair of DNA damage, study of this group of diseases has been important for identifying those processes responsible for the maintenance of genomic stability.

1.3.1 Xeroderma pigmentosum, Cockayne's syndrome and trichothiodystrophy

Xeroderma pigmentosum (XP), Cockayne's syndrome (CS) and trichothiodystrophy (TTD) are three hereditary diseases with strikingly different clinical features. However, it was noted that both CS and TTD could be associated with XP. Investigation into the molecular basis of each disease has revealed the surprising finding that patients of all three diseases have defective NER. Information in this section comes from reviews by Lehmann 1995, Tanaka and Wood 1994 and Friedberg, Walker and Siede 1995 with references therein.

1.3.1.1 Xeroderma pigmentosum

XP patients display extreme sensitivity to sunlight. The skin freckles very easily, and areas of the skin exposed to solar radiation show an extremely high incidence of skin cancers (2000-fold increased compared to normal), with the average age of cancer diagnosis being 8 years. Other areas of the body exposed to the sun, such as the eyes and tongue, are also prone to neoplasia, and the rate of internal cancers is also increased.

The XP phenotype is caused by a defect in the NER mechanism, first shown by a substantially reduced rate of DNA synthesis outside S-phase (unscheduled DNA synthesis) in response to irradiation with ultraviolet light. As a consequence, cells from patients are unable to repair DNA lesions caused by UV radiation, and hence are prone to formation of chromosomal abnormalities and neoplastic transformation. Cells are also sensitive to killing by a wide range of other DNA damaging agents, but levels of sensitivity vary between patients. Complementation analysis by fusion of cultured cells confirmed that this variability was the result of heterogeneity at the molecular level. Seven complementation groups were identified, designated A to G (plus a "variant" form - typical clinical symptoms, but apparently with normal NER activity). Genes corresponding to most of these complementation groups have been cloned. However, some of these genes have proven to be genes previously identified from studies of rodent cell lines. Excision repair genes were cloned on the basis of their ability to complement the repair deficiency of 11 groups of UV-sensitive mutant

cell lines, and hence the genes were designated excision repair cross-complementing (ERCC).

Table 1.3.1 summarises the current knowledge of the relationship between XP and ERCC genes, and the proposed function of each gene. Evidence suggests that in XP groups A to G the incision step of NER is defective, while XP variant cells are unable to tolerate UV damage during DNA replication but that the basic NER reaction functions normally.

Table 1.3.1: The relationship between XP, ERCC, CS and TTD genes

Complementation Group			Function
XP	ERCC	CS	
A			Damage recognition
B	3	XPB/CS	DNA helicase, part of TFIIH
C			Global repair
D	2	XPD/CS	DNA helicase, part of TFIIH
E			Not cloned
F	4		Structure-specific nuclease
G	5	XPG/CS	Endonuclease
V			Not cloned. Unknown
	6	B	Transcription-repair coupling factor
	1		Structure-specific nuclease
	8	A	Transcription-repair coupling factor

1.3.1.2 Cockayne's syndrome

XP patients in the complementation groups B, D and G have occasionally presented with the additional clinical features of CS. CS is a multi-system disorder and patients show arrested growth and development, characteristically having sunken eyes and disproportionately long limbs. Other symptoms include sun-sensitivity and deafness, optic atrophy, and severe mental retardation. The average age of death is 12 years.

Cultured CS cells are extremely sensitive to killing by UV irradiation, but measurements of NER activity proved to be normal. However, it has been shown that CS cells are deficient in transcription-dependent NER. Wild-type cells exhibit preferential repair of transcribed regions of the genome after UV-irradiation, but in CS cells both transcribed and non-transcribed regions are repaired at the same rate. Some CS cells are also defective in the transcription-coupled repair of the coding strand of transcribed genes. The result is that RNA synthesis is inhibited by UV irradiation for longer than normal.

Using cell fusion techniques two CS genetic complementation groups were identified (designated CSA and CSB), although three additional groups are defined by overlap with XP (referred to as XP-B/CS, XP-D/CS, and XP-G/CS). The corresponding genes have been cloned, and both CSA and CSB proteins are thought to function as transcription factors which allow the coupling of repair to transcription by interaction with TFIIH. In those cases where CS overlaps with XP, the primary mutations are within those XP genes that make up subunits of TFIIH and the XPG endonuclease.

1.3.1.3 Trichothiodystrophy

Trichothiodystrophy, or TTD, is the term coined to describe a syndrome of brittle sulphur-deficient hair, fish-like scales on the skin (ichthyosis), mental and physical retardation. However, photosensitivity is present in about half of TTD cases, and cells from these patients were found to be NER-deficient. In these cases it was found that the underlying molecular cause was mutation of either the *XPB* or *XPD* genes.

In order to explain how mutations in *XPB* and *XPD* can cause three different diseases a "transcription syndrome" hypothesis has been proposed. As components of the TFIIH transcription factor, these genes play roles in both NER and RNA polymerase II-dependent transcription. Certain mutations inactivate just the NER function of the protein, resulting in typical XP. Other mutations inactivate the NER function and/or alter the efficiency of transcription, which specifically affects certain proteins whose synthesis is dependent on maximal transcription. Finally another class of mutations can prevent the coupling of transcription to repair.

1.3.2 Fanconi anaemia

Fanconi anaemia (FA) is a pleiotropic disease whose main characteristic is bone marrow failure (reviewed by Digweed and Sperling 1996). The clinical picture is, however, highly variable. Patients may be affected by any number of congenital deformities such as skeletal defects, microcephaly, and ear, heart, kidney, intestinal and hand abnormalities, but 25% of patients display no symptoms other than pancytopenia (depression of all types of blood cells). Bone marrow failure usually begins in the first few years of life, and can be accompanied by stunted growth, and learning disabilities. There is a marked predisposition to develop acute myeloid leukaemia and other solid tumours, which is apparently linked to an elevated level of spontaneous chromosomal breakage.

Levels of chromosome aberrations increase after treatment of cells with mitomycin. Whilst some evidence supports the hypothesis that FA is caused by

defective repair of DNA cross-links, other molecular changes have been observed in FA cells, and it has been proposed that the primary defect is in cell cycle control. Failure to arrest DNA replication following cross-link formation will ultimately result in the generation of chromatid breaks and chromosome damage. Cell fusion experiments indicated that there are 8 complementation groups, designated A to H, and genes corresponding to 3 of these have been cloned (*FAA*, *FAC*, and *FAE*) (On-line Mendelian Inheritance in Man 1997). However, the molecular mechanism by which the disease is manifested has remained elusive.

1.3.3 Ataxia telangiectasia

Interest in ataxia telangiectasia (AT) stems from the fact that the gene responsible may also be the single largest cause of breast cancer. AT affects up to one in 40,000 people, and is pleiotropic in nature (reviewed in Friedberg, Walker and Siede 1995). The most striking characteristic is cerebellar ataxia, resulting in a staggering gait, lack of co-ordination and progressive mental retardation, but telangiectasia (dilation of the small blood vessels in the eye and skin) also occurs. Immune dysfunction means that patients are exceptionally susceptible to infections, and this may also contribute to the high incidence of cancer observed. In most AT individuals neoplasia occurs in the lymphoreticular system before the age of 20, and it has been shown that this predisposition stems from problems cells experience in response to DNA damage.

The first indication of a DNA repair defect came from observation of the severe, sometimes fatal, response of AT patients to radiotherapy. Subsequently it was demonstrated that AT cells are sensitive to killing by ionising radiation, but cells exhibit a high number of chromosomal breaks even prior to irradiation. A number of cell cycle anomalies have been recorded in AT cell lines, including failure to arrest at the G1-S or G2-M checkpoints after irradiation, and the lack of normal inhibition of DNA synthesis following irradiation accounts for the extreme radiosensitivity.

Five AT complementation groups were defined from cell fusion studies, designated AB, C, D, E and V_1. However, it was subsequently found that all AT patients have mutations in single gene, named *ATM*. Positional cloning of *ATM* was achieved in 1995, and analysis of the coding sequence immediately gave a clue to the function of the protein. ATM is a member of the PI3-kinase-related protein kinase superfamily (along with DNA-PK) and has protein kinase activity. It is believed that ATM acts as a damage sensor and checkpoint protein at the head of the p53 signal transduction pathway. In response to DNA damage, ATM phosphorylates cellular proteins, including p53, thereby arresting the cell cycle until the damage can be repaired.

1.3.4 Bloom's syndrome

Bloom's syndrome (BS) was first recognised in 1954 by paediatrician Dr David Bloom following study of three patients (Bloom 1954.), and was described as "congenital telangiectatic erythema resembling lupus erythematosus in dwarfs" (Bloom 1966). Along with the initial report, Dr Bloom also initiated a registry of BS patients to collect clinical data about the disease. Since then approximately 170 cases have been reported world-wide, and the disease has been determined to have an autosomal recessive mode of inheritance (German 1969.) Although very rare in most populations (cases tending to arise from consanguineous unions), the incidence of the disease is higher in Japan and among the Ashkenazim, where 1 in 100 are heterozygous carriers.

1.3.4.1 The Bloom's syndrome phenotype

The clinical features of BS are reviewed in Passarge 1983, and German 1993. See also On-line Mendelian Inheritance in Man. All cases are similar in clinical appearance, and show complete penetrance and expressivity (German *et al.* 1994). The main characteristics are as follows:

i) Telangiectatic erythema. Dilation of the capillaries of the face produces a characteristic butterfly rash, visible from infancy.

ii) Skin photosensitivity. Exposure to sunlight results in hyperpigmentation and worsening of areas of erythema.

iii) Growth retardation. Growth is stunted from birth, meaning that adults are dwarfed, though with normal body proportions.

iv) Moderate to severe immunodeficiency. This leads to repeated bacterial infections during infancy and early childhood. Although the situation improves with age, patients tend go on to develop severe chronic lung disease as a result. Studies revealed the basis of the immunodeficiency to be an aberrant immunoglobulin serum profile, with low IgM concentrations, and mildly decreased IgG and IgA levels. The IgM deficiency apparently arises from dysfunction during the maturation of B cells, since although the percentage of cells bearing surface IgM is not reduced, numbers of IgM secreting cells are. Further explanation may come from the observation of preferential damage to IgM production by UVB irradiation. Peripheral blood mononuclear cells (B cells, T cells and monocytes) recover their proliferative response to pokeweed mitogen stimulation more slowly than normal after UVB irradiation (Ozawa *et al.* 1993). Patients also showed reduced reactivity to skin test antigens *in vivo*, and a low proliferative response by lymphocytes to mitogen stimulation *in vitro*. Reduced numbers of IgG bearing lymphocytes accounts for the

lowered levels of IgG. In addition, natural killer cell activity is diminished, as is helper T cell function (Steihm 1993).

v) Minor clinical conditions are common. Mild mental retardation, and hypogonadism may be present (male BS patients all being sterile), and there is a higher than normal incidence of diabetes among older patients.

vi) Elevated risk of cancer. BS patients show a predisposition to develop almost any type of cancer at an earlier age than the rest of the population. On average, leukaemia becomes manifest at 22 years, while solid tumours are detected at age 35 in surviving patients, compared to the norms of 55 and 65 respectively. Despite modern medical techniques, no BS patient has survived beyond 48 years of age (German 1993).

1.3.4.2 Molecular characteristics of Bloom's syndrome

At a cellular level, BS cells show a striking level of genomic instability. Between 5 and 15% of all metaphases looked at have more than one cytogenetic abnormality, and the numbers of micronuclei formed *in vivo* are increased. It is believed that the alterations observed arise as a result of an increase in the rate of exchanges between DNA strands in BS cells (somatic recombination). This can occur between homologous chromosomes at homologous sites to produce symmetrical quadriradials or telomeric associations. Other chromosomal defects, such as chromosomal translocations, are also increased. Alternatively reciprocal interchanges may occur between chromatids of one chromosome after DNA replication (exchange of two homologous DNA helices between sister chromatids during mitosis). These type of exchanges are known as sister chromatid exchanges (SCEs). The frequency of SCE formation is raised to a level 10-fold higher than normal (Chaganti, Schonberg, and German 1974).

A number of groups were able to demonstrate increased mutation and somatic recombination rates. Spontaneous 6-thioguanine resistance was shown to be raised 4- to 15-fold in fibroblasts (Warren *et al.* 1981) and 8-fold in peripheral blood lymphocytes (Vijayalaxmi *et al.* 1983). Passage of Herpes simples virus-1 indicated that viral recombination was abnormal in Bloom's syndrome cells, and Langlois *et al.* found a 50- to 100-fold increase in NM blood group variant erythrocytes (Langlois *et al.* 1989).

BS cells display sensitivity to some DNA damaging agents (Table 1.3.3). This observation, combined with the predisposition of patients to cancer, led to BS being designated a DNA repair deficiency. Exposure of cells to UV irradiation was accompanied by altered unscheduled DNA synthesis (UDS) and increased chain breakage, taken to indicate an abnormality in a postincision step of DNA excision repair, such as ligation (Willis *et al.* 1987). Study of the *in vivo* joining ability, by

transfection with linearized replicating shuttle vector plasmid, revealed that DNA end joining was reduced and error prone in BS cells (Rünger and Kraemer 1989). It was subsequently proposed that these defective ligation processes could be the cause of not only elevated mutation, and chromosome breakage levels, but also the impaired immune function found in Bloom's syndrome patients, since normal immunoglobulin gene processing involves the joining of DNA segments (Rünger and Kraemer 1989). A defect in ligation was also suggested by observation of a retarded rate of replication fork progression during semi-conservative DNA synthesis (Hand and German 1975). Furthermore, pulse labelling of replicating DNA with [^{3}H]tritium revealed a novel class of 20kb intermediates, not seen in normal cells, to be present in BS cells (Lönn *et al.* 1990).

1.3.4.3 Genetic basis of Bloom's syndrome

Consistent with this data was the finding that the activity of DNA ligase I is reduced in BS cells. Immunoblotting showed enzyme levels to be normal (Chan and Becker 1988), but when purified, the enzyme proved to be unusually heat labile, or to form anomalous dimers (Willis and Lindahl 1987). However, when the human gene encoding DNA ligase I (*LIG1*) was cloned (Barnes *et al.* 1990), comparison of cDNA sequences revealed DNA ligase I from BS cells to contain no amino acid alterations. *LIG1* cDNA from BS cells was able to rescue *S. cerevisiae cdc9* mutants, and Northern blot analysis showed *LIG1* message levels to be normal (Petrini, Huwiler and Weaver 1991). It was therefore concluded that Bloom's syndrome does not arise primarily from an alteration in the *LIG1* gene.

Numerous investigations have revealed altered structure and/or activity levels of enzymes involved in DNA metabolism in BS cells. Nevertheless, all of these were ruled out as the primary BS defect because of their chromosomal location (summarised in Table 1.3.2). BS is represented by a single complementation group, cell fusion of BS cell lines from different ethnic backgrounds all failing to complement each other (Weksberg *et al.* 1988). McDaniel and Schultz used microcell-mediated chromosome transfer of human chromosomes to complement the SCE phenotype of BS cells, and thereby map the Bloom's syndrome locus (*BLM*) to chromosome 15q14-qter (McDaniel and Schultz 1992).

The location was further refined by linkage analysis of polymorphic loci in the progeny of consanguineous marriages (homozygosity mapping). 25 out of 26 BS patients were homozygous for a tetranucleotide repeat in intron V of the proto-oncogene *FES*. *FES* had previously been mapped to 15q26.1 by fluorescence *in situ* hybridisation (FISH), and hence it was possible to also assign *BLM* to 15q26.1

Table 1.3.2: Chromosomal locations of BS candidate genes

Gene Symbol	Chromosomal Location
LIG1	19q13.3
LIGIII	17q11.2-q12
LIGIV	13q33-q34
UNG	12
TOP1	20q12-q13.1
TOP2A	17q21-q22
TOP2B	3
MGMT	10q26
AAG/MPG	16
TYMS	18p11.31
SOD1	21q22.1
SOD2	6q21
SOD3	4p16.3-q21
p53	17p13
c-myc	8q24

Summary of the chromosomal locations of the human genes whose products show alterations in BS cells. Gene symbol abbreviations: *LIG1*, DNA ligase I. *LIG3*, DNA ligase III. *UNG*, uracil DNA-glycosylase. *TOP1*, topoisomerase I. *TOP2A*, topoisomerase IIA. *TOP2B*, topoisomerase IIB. *MGMT*, O6-methylguanine methyltransferase. *AAG/MPG*, 3-methyladenine DNA glycosylase. *TYMS*, thymidylate synthetase. *SOD1*, superoxide dismutase I. *SOD2*, superoxide dismutase II. *SOD3*, superoxide dismutase III. Chromosomal locations taken from On-line Mendelian Inheritance in Man 1997.

(German *et al.* 1994). As shown in Table 1.3.2, none of the other genes map to this region.

Subsequently, mapping of *BLM* exploited a phenomenon that been observed in some BS lymphoblastoid cell lines (LCLs) (Ellis *et al.* 1995a). High SCE frequencies are unique to BS, and are routinely used to confirm clinical diagnoses. However, it had been noted that certain cell lines established from lymphoid cells of BS patients with non-consanguineous parents exhibited low SCE rates, representing a minor population of cells *in vivo* (mosaicism). It was hypothesised that this was caused by somatic intragenic recombination between different *BLM* alleles generating a functionally normal allele, thereby correcting the SCE phenotype. One consequence of this event would be that loci distal to *BLM* would become homozygous in low SCE

cells while proximal loci would remain heterozygous. Analysis of 12 polymorphic markers around *FES* in LCLs from 11 patients displaying mosaicism narrowed the location of *BLM* to 250kb proximal of *FES* (Ellis *et al.* 1995b). There was only one non-repetitive single copy transcribed sequence within the region, and the identity of this novel gene as *BLM* was confirmed by the detection of mutations in a number of BS patients.

The BLM protein has a predicted molecular mass of 159kDa, and showed homology to the RecQ subfamily of DExH box-containing helicases. It has been suggested the BS phenotype arises from mutations which cause the loss of helicase activity, and that consequently BS cells have difficulty resolving specific DNA structures generated during replication (Ellis *et al.* 1995b). It is hypothesised that an error-prone recombination process subsequently becomes activated, leading to excessive crossing over observed as SCEs and quadriradials. Theoretically an SCE has no consequence, but imprecise or unequal exchanges could lead to base substitutions, deletions, duplication or even breaks. This would account for the spontaneous mutability of BS cells, while somatic crossing over will also cause loss of heterozygosity, thereby promoting tumour formation (German 1993).

1.3.5 '46BR' patient

1.3.5.1 The '46BR' phenotype

The so called '46BR' patient was a young female with a number of clinical symptoms suggestive of Bloom's syndrome (Webster *et al.* 1992). At birth the patient was slightly underweight and anaemic. Normal developmental milestones were delayed, and growth was retarded, leading to an "elf-like" appearance. Although there was no mental retardation, secondary sexual characteristics did not develop. Repeated ear and chest infections occurred, leading to poor hearing and lung function. Exposure of the skin to sunlight caused hypersensitivity, and, at age 17, patches of telangiectasia appeared. Shortly after this, a liver lymphoma was diagnosed, but the patient died, aged 19, from pneumonia.

Thyroid, pituitary and adrenal function were normal, but immunodeficiency was noted. IgM levels were normal, but there was moderate IgG depression, and a deficiency of IgA and isohaemagglutinins. Erythrocyte, lymphocyte and neutrophil counts were normal, but T and B lymphocyte cells showed a poor proliferative response to mitogens *in vitro*. There were abnormally large numbers of macrocytes in the peripheral blood, but no anaemia. Bone marrow culture at age 17 revealed depressed red cell colony formation, at which point the patient was found to have

become lymphopenic, with no lymphocyte proliferative response *in vitro* to phytohaemagglutinin or Ca^{2+} ionophore stimulation.

1.3.5.2 Cellular characteristics

Repeated attempts to establish a lymphoblastoid cell line by immortalisation with Epstein-Barr virus were unsuccessful, but skin fibroblasts, designated 46BR, were established in culture. These fibroblasts proved to be sensitive to a wide range of DNA damaging agents (Table 1.3.3), but particularly to monofunctional methylating agents. In contrast to BS, where numbers of SCEs are characteristically 10-fold higher, 46BR cells exhibited baseline levels only marginally higher than normal (Henderson *et al.* 1985). A single cell gel electrophoresis ('comet') assay has also been used on 46BR cells to analyse DNA breakage following treatment with DNA damaging agents. As expected from cell survival data, the number of breaks in 46BR cells was much higher than wild-type after UV irradiation, and these breaks persisted for a long time. Unexpectedly however, treatment with γ radiation and agents to induce superoxide anion formation revealed no differences from wild-type cells (Nocentini 1995).

Because DNA lesions generated by ionising radiation, UV light, mitomycin C and alkylating agents are chemically unrelated, and corrected by different modes of excision repair , it was concluded that the underlying defect in 46BR could not lie within any one particular repair pathway. Rather, the deficiency must be at a step common to all pathways, such as polymerisation or ligation (Teo and Arlett 1982). More detailed study of replicative DNA synthesis, UDS, and daughter strand postreplicative repair following UV irradiation revealed no abnormality. Similarly the removal of 3-methyladenine after DMS treatment was normal. However, strand breaks persisted, as they did after UV irradiation (Squires and Johnson 1983). Given the fact that DNA polymerisation, as measured by UDS, was apparently normal, it was therefore concluded that the raised level of DNA strand breaks was the consequence of an alteration in the rate of ligation. Using the rate of gap sealing as a measure of ligase activity, it was thus calculated that normal cells ligate induced gaps 3 times faster than 46BR.

Further support for this conclusion came from the effect of 3-aminobenzamide (3AB) on 46BR cells. 3AB is a specific inhibitor of poly(ADP-ribose) polymerase (PARP/PADPRP)[i]. PARP is an abundant zinc-finger DNA binding nuclear protein which uses NAD^+ as a precursor to synthesise chains of poly(ADP-ribose), and transfers them to a number of proteins, as well as to the enzyme itself. Activity is

[i] EC 2.4.2.30 Also known as NAD^+:protein (ADP-ribosyl)transferase (ADPRT), and poly(ADP-ribose) synthetase

Table 1.3.3: The sensitivity of BS and 46BR cells to DNA damaging agents

DNA damaging agent	Bloom's syndrome	46BR
UV	n/+	+
γ	n	+
Solar	ND	n
Mitomycin C	+	+
4NQO	n	n
DMS	ND	+
MMS	n	+
EMS	+	+
MNNG	ND	+
MNU	n	+
ENU	+	n
AAF	n	ND

The response of cell strains from BS patients and the 46BR patient to DNA damaging agents. n indicates a wild-type response; +, sensitivity to killing by that agent; ND, not determined. Abbreviations: UV, ultraviolet radiation; γ, gamma radiation; 4NQO, nitroquinoline oxide; DMS, dimethyl sulphate; MMS, methyl methanesulphonate; EMS, ethyl methanesulphonate; MNNG, *N*-methyl-*N*-nitro-*N*-nitrosoguanidine; MNU, *N*-methyl-*N*-nitrosourea; ENU, *N*-ethyl-*N*-nitrosourea; AAF, acetylaminofluorine; Adapted from Teo *et al.* 1983 .

dependent on DNA strand breaks, and although a number of nuclear functions have been ascribed to PARP, it appears to be a molecular sensor of DNA strand breaks (Masson *et al.* 1998 and references therein). In normal cells inhibition of PARP by 3AB results in the accumulation of strand breaks, and increased cell death, following DNA damage. However, 46BR cells proved to be hypersensitive to the lethal effects of 3AB, even in the absence of any other agent (Lehmann *et al.* 1988).

Defective ligation in 46BR was demonstrated more directly by measuring the accumulation of replication intermediates. Okazaki fragments in mammalian cells are approximately 200nt in length, and exist only transiently before being incorporated into larger 10kb molecules, which are themselves subsequently united into high molecular weight DNA. Pulse labelling revealed that low molecular mass replication intermediates persist in 46BR cells, thereby indicating abnormally slow joining of Okazaki fragments (Lehmann *et al.* 1988, Lönn *et al.* 1989, Prigent *et al.* 1994,

Mackenney, Barnes and Lindahl 1997). Demonstration that the 46BR phenotype is a direct consequence of mutation of DNA ligase I came with the cloning of *LIG1*.

Initial suggestions as to the cause of the immunodeficiency of the 46BR patient centred upon the joining of gene segments during V(D)J recombination in developing lymphoid cells. Using 46BR fibroblasts transfected with RAG1 and RAG2 expression plasmids, the efficiency of V(D)J recombination on plasmid substrates was assayed Sequencing of the coding junctions and signal joints formed did not reveal changes in the fidelity or efficiency of V(D)J recombination in 46BR cells (Hsieh, Arlett and Lieber 1993, Petrini *et al.* 1994). It was thus concluded that the immunodeficiency observed was not simply due to failure of immunoglobulin and TCR gene rearrangement.

1.4 DNA ligases

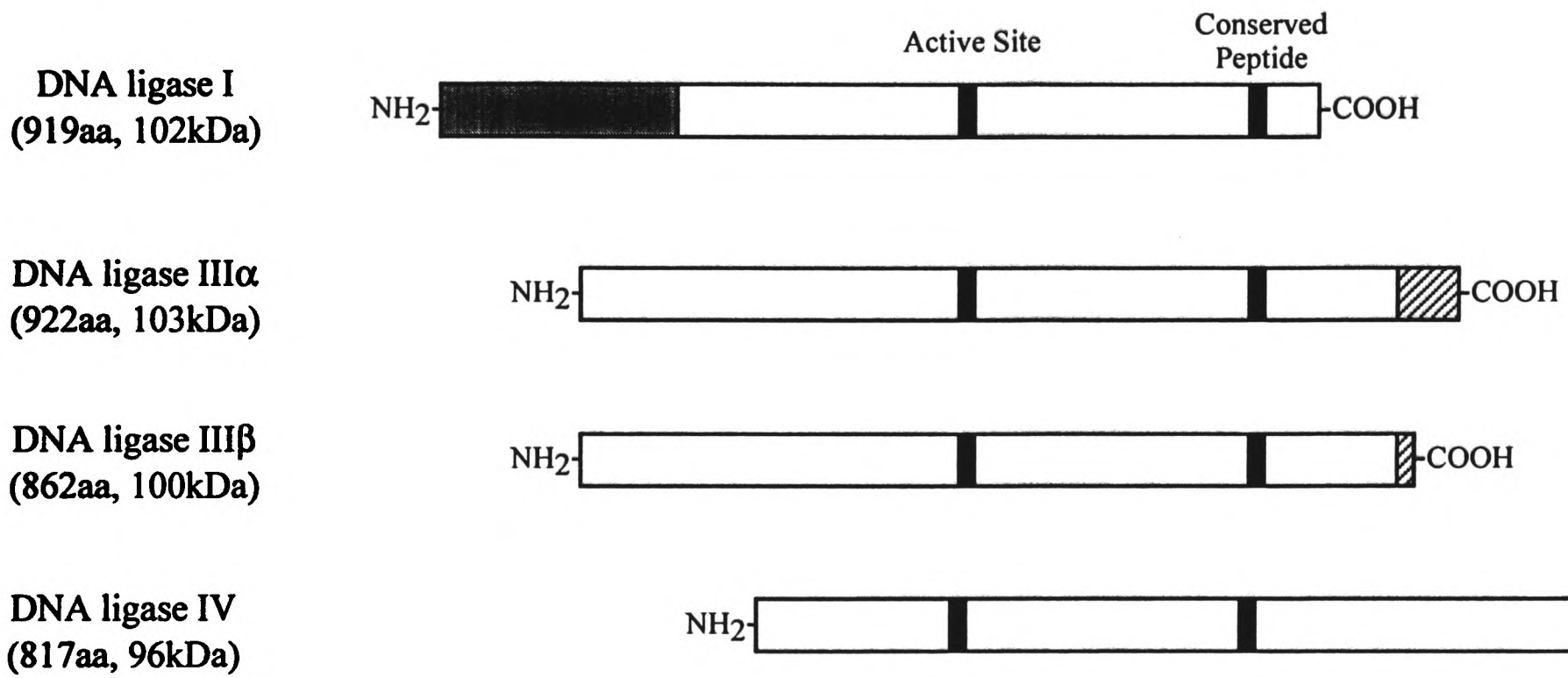
DNA ligases catalyse the formation of phosphodiester bonds, sealing breaks in double-stranded DNA molecules. As such, DNA ligases are required for a variety of cellular processes, including DNA replication, excision repair and recombination. DNA ligase activity in eukaryotes was first described in rabbit tissues by Lindahl and Edelman in 1968 (Lindahl and Edelman 1968), and soon after it was shown that there are [at least] 2 different ligases in mammals (Söderhäll and Lindahl 1973) . There are now known to be at least four biochemically distinct DNA ligases present in mammalian cells, of which three have been cloned (Table 1.4.1) (reviewed by Lindahl and Barnes 1992, Lindahl *et al.* 1993, Tomkinson and Levin 1997). Figure 1.4.1 depicts an alignment of the DNA ligases I, the two splice-variant forms of DNA ligase III, and DNA ligase IV.

Table 1.4.1: Cloned human DNA ligases

Enzyme	Gene symbol	Chromosomal location	Protein size		Cellular function
			Amino acids	Mass (kDa)	
DNA ligase I	*LIG1*	19q13.2-13.3	919	102	Replication, NER, BER
DNA ligase IIIα	*LIG3*	17q11.2-12	922	103	BER
DNA ligase IIIβ	*LIG3*	17q11.2-12	862	100	Meiotic recombination
DNA ligase IV	*LIG4*	13q33-34	817	96	DSB repair

Figure 1.4.1: Schematic representation of DNA ligases I, IIIα, IIIβ, and IV

Alignment of the three cloned human DNA ligases based on conserved motifs. Adapted from Wei *et al.* 1995 and Tomkinson and Levin 1997. The active site and conserved peptide motifs are represented by solid bars. Shaded box, N-terminal regulatory domain of DNA ligase I. Hatched box, alternatively spliced regions of DNA ligase III. Predicted molecular masses are indicated. It was noted from alignment of the predicted amino acid sequences of the three DNA ligases that the distance between the active site and conserved C-terminal peptide is practically invariant.


DNA ligase I
(919aa, 102kDa)
Active Site
Conserved
Peptide
NH2
COOH
DNA ligase IIIα
(922aa, 103kDa)
NH2
COOH
DNA ligase IIIβ
(862aa, 100kDa)
NH2
COOH
DNA ligase IV
(817aa, 96kDa)
NH2
COOH

1.4.1 DNA ligase I

1.4.1.1 Cloning and gene structure

Using oligonucleotides deduced from partial amino acid sequences of purified bovine DNA ligase I, and testing for the ability to complement a temperature sensitive DNA ligase mutant of Saccharomyces cerevisiae (*cdc9*), Barnes *et al.* were able to clone the *LIG1* cDNA from a human cDNA library (Barnes *et al.* 1990). The cDNA was found to encode a 919aa (102kDa)[ii] protein with 37% identity to the yeast Cdc9 protein. Northern blot analysis revealed *LIG1* to produce a single 3.2kb transcript, while hybridisation to a panel of rodent-human somatic cell hybrids retaining subsets of human chromosomes localised the gene to chromosome 19. Subsequently, Southern analysis of rodent-human somatic cell hybrids and FISH mapped *LIG1* to 19q13.2-13.3 (Barnes *et al.* 1992a).

Analysis of gene structure revealed the human gene to consist of 28 exons (54 to 173 bp in length) spanning 53kb (Noguiez *et al.* 1992). The first exon (-1) is untranslated, and utilises a GC dinucleotide instead of the usual GT at the donor splice site. The 400bp upstream of the translation initiation methionine was sequenced, and showed the promoter to have a high GC content, and no Goldberg-Hogness (TATA) or CCAAT boxes. Primer extension indicated that transcription initiation occurs at multiple sites, all of which is typical of a putative 'housekeeping' gene. In addition, four inverted SP1 binding sites were identified, along with a motif recognised by p53 protein-containing complexes, and a possible ATF binding site (Noguiez *et al.* 1992).

1.4.1.2 Expression pattern

DNA ligase I has been detected in all cell types looked at, even terminally differentiated neurones (Teraoka and Tsukada 1982), but a correlation has been demonstrated between *LIG1* expression and the rate of cell proliferation. DNA ligase I is the major polynucleotide joining activity in proliferating cells, contributing 85% of the total activity in calf thymus (Tomkinson *et al.* 1990). There is an increase in DNA ligase I activity upon stimulation of quiescent cells (Elder and Rossignol 1990, Söderhäll 1976). The increase is 100-fold in human lymphocytes stimulated by phytohaemagglutinin (with slight delay with respect to the onset of DNA synthesis), with levels remaining high thereafter. Correspondingly, when the human

[ii] Estimates from western blotting give the size of the protein as around 125kDa. This difference is though to reflect reduced mobility arising from proline and phosphoserine-rich regions in the N-terminus of the protein (see Lindahl and Barnes 1992)

premyelocytic leukaemia cell line HL-60 is induced to differentiate, *LIG1* mRNA levels drop (Montecucco *et al.* 1992).

DNA ligase I activity is induced 15-fold in the regenerating rat liver (Chan and Becker 1985), and mRNA levels are considerably higher in proliferating cells, but drop when cells reach confluence or are starved (Montecucco *et al.* 1992). Following starvation, the basal mRNA level increases 12 hours after serum stimulation, reaching a plateau 6hr later (maximum expression coinciding with the first S-phase after serum stimulation), and remaining high even when S phase is completed. Because *LIG1* mRNA is stable regardless of state of cell, dropping only $^1/_5$ after 3hrs incubation with transcription inhibitor actinomycin D, it was concluded that the elevation of mRNA levels are due to raised transcription rather than increased stability.

The expression pattern of *LIG1* is similar to that of DNA polymerase α. The *pola* gene, and other genes involved in DNA replication (e.g. *tk*, *TOPIIA*, *PCNA*) all encompass SP1, and ATF sites within their promoters, and lack TATA boxes (but do contain CCAAT boxes). Taken together these data suggest a role for DNA ligase I in replication. However, constitutive expression of DNA ligase I, even in resting cells, resembles the expression pattern of polymerase β, a DNA polymerase known to participate in DNA repair (Montecucco *et al.* 1992). After UV-C irradiation of confluent fibroblasts, DNA ligase I activity and mRNA levels both increase 6-fold (Montecucco *et al.* 1995a). Induction is serum independent, and results in an increase in steady state mRNA proportional to the UV dose. However, expression is raised only after 24 hours, when UDS is no longer detectable and nucleotide excision repair (NER) can no longer be operating. Thus, if DNA ligase I has a role in NER, the basal cellular level must be sufficient to accomplish this. Since activation occurs at a lower UV dose in XP cells that show UV hypersensitivity (as a result of abnormal NER), it was proffered that DNA ligase I induction could be part of a response to damage not removed by NER. This proposed DNA repair system would be activated late after UV irradiation, as a consequence of unrepaired DNA damage, in a similar manner to the SOS response in *E. coli*. Given that p53 protein has been suggested to play a role as a checkpoint protein, inducing growth arrest in cells with DNA damage, it was hypothesised that p53 may play a role in the induction of the DNA ligase I gene (Montecucco *et al.* 1995a).

1.4.1.3 Protein localisation

Although it was reported that a substantial amount of DNA ligase I was present in the cytoplasmic fraction of preparations from calf thymus (Söderhäll and Lindahl 1975), recent evidence has indicated DNA ligase I to be localised to sites of DNA replication in the nucleus. Immunohistochemistry and immunofluorescence showed a

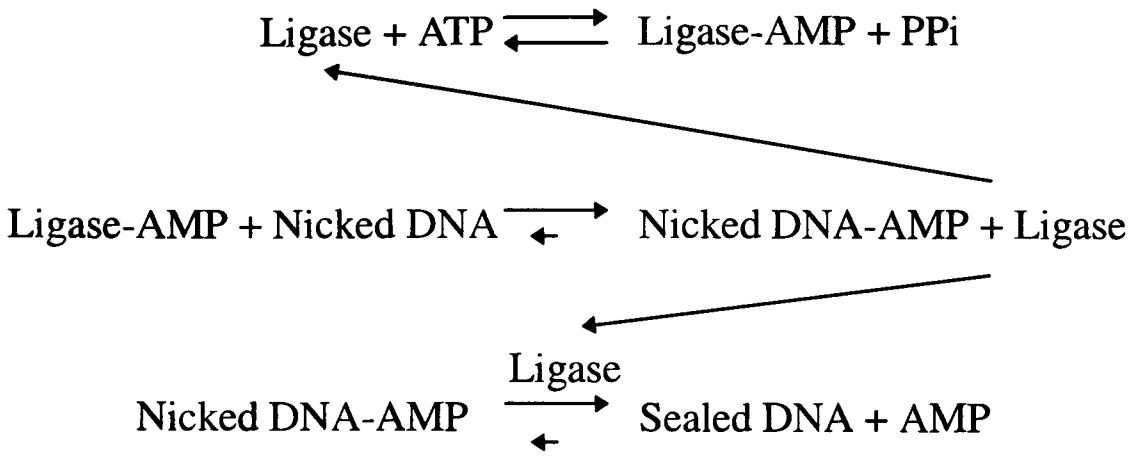
punctate pattern of staining in the nuclei of S phase cells. When cells entered mitosis, breakdown of the nuclear membrane allows DNA ligase I to become distributed throughout the cell (Lasko, Tomkinson and Lindahl 1990). This was confirmed by other groups. Intranuclear immunofluorescence staining revealed DNA polymerase α and DNA ligase I to both localise to the same granular-like structures. At the G1/S boundary these structures are distributed throughout nucleoplasm, whereas at the S/G2 boundary they are constrained adjacent to the nuclear envelope. In contrast polymerase β was distributed throughout nucleoplasm at all times (Li *et al.* 1993). Montecucco *et al.* found that in early S phase staining for DNA ligase I produced a fine, evenly distributed punctate pattern. By mid-S phase, dots became more concentrated round the nucleoli and nuclear periphery, while in late S phase homogeneous staining was seen. These changes closely paralleled alterations to the location of DNA replication (Montecucco *et al.* 1995b).

More recently, amino acid sequences which function to target the DNA ligase I to the nucleus and replication foci have been identified. Fusion of regions of DNA ligase I to a marker protein allowed visualisation of subcellular localisation by immunostaining. A 13 amino acid peptide (PKRRTARKQLPKR, 119-131 in DNA ligase I) was found to be required for targeting of the fusion protein to the nucleus. This nuclear localisation signal (NLS) is conserved between human and mouse protein, and when deleted, DNA ligase I does not enter the nucleus (Montecucco *et al.* 1995b). Deletion of 115 amino acids upstream of the NLS allowed accumulation of the protein into the nucleus, but not punctate staining during S phase (Montecucco *et al.* 1995b). Deletion of only residues 2-6 has a similar effect, as does deletion of 1-11 or 12-20, and fusion of the first 20aa to a fluorescent reporter protein will result in targeting to replication foci. Comparison of sequences between proteins involved in DNA replication has indicated that the motif responsible for the targeting of proteins to replication factories is composed of two boxes. Box 1 (1-11 in the DNA ligase I sequence) includes the sequence IxxFF, while Box 2 (12-20 in the DNA ligase I sequence) is rich in positively charged residues (Montecucco *et al.* 1998). Similar results were also achieved by Cardoso *et al.* (Cardoso *et al* 1997).

1.4.1.4 Enzyme activity

DNA ligase I catalyses the formation of phosphodiester bonds between adjacent DNA fragments by the same general mechanism delineated for T4 DNA ligase (Lehman 1974) (Figure 1.4.2). The enzyme requires Mg^{2+}, and is activated by the cleavage of ATP to generate a covalent DNA ligase-AMP intermediate, with a lysine-adenylate phosphoamide bond. AMP is then transferred to the 5' terminus of a DNA interruption, forming a transient DNA-AMP complex. In the final step, un-adenylated

Figure 1.4.2: The reaction mechanism of DNA ligase I



A representation of the reaction mechanism of DNA ligase I. Abbreviations: ATP, adenosine triphosphate, AMP, adenosine monophosphate. A hyphen (-) indicates a covalent link. Adapted from Lehman 1974.

DNA ligase I catalyses nucleophilic attack of the 3' OH group on the activated 5' phosphate to generate a phosphodiester bond. The crystal structure of DNA ligase from the bacteriophage T7 has been extrapolated to suggest that DNA ligase I also consist of two domains with a deep DNA binding cleft between, with an ATP-binding pocket present in domain 1 (Subramanya *et al.* 1996).

DNA ligase I efficiently joins both single and double strand breaks in DNA, including blunt ends, but is also active on the synthetic substrates oligo(dT)/poly (dA), and oligo(rA)/poly(dT) (Arrand *et al* 1986, Tomkinson *et al* 1991a). Examination of the efficiency of ligation of DNA nicks with mismatched termini indicated that DNA ligase I is severely inhibited by the presence of 3' mismatches, leading Husain *et al.* to conclude that the enzyme is more faithful than the other mammalian DNA ligases (Husain *et al.* 1995).

Recently it was reported that DNA ligase I can also seal 5'-monoribonucleotides into DNA, although with only 10% of the efficiency of sealing monodeoxyribonucleotides. During lagging strand DNA synthesis, cleavage of the RNA primer in Okazaki fragments by RNase H1 leaves a single ribonucleotide attached to the 5' end of the downstream DNA fragment. This is normally removed by DNase IV/FEN-1, but these data indicate that it can alternatively be sealed onto the DNA molecule by DNA ligase I (Rumbaugh *et al.* 1997) .

Human DNA ligase I has 84% homology to mouse DNA ligase I, and approximately 40% with *Schizosaccharomyces pombe cdc17+* and *S. cerevisiae CDC9* gene products (Barnes *et al.* 1990). However, identity is greatest within the C-terminus, the N-termini being poorly conserved. The ~28kDa amino-terminal region is highly susceptible to proteolysis *in vitro*, and, when removed, the remaining

~74kDa polypeptide (encoded by exons 8 to 27) is still enzymatically active (Tomkinson *et al.* 1990). The C-terminal 'catalytic domain' is also sufficient to complement the *cdc9* mutation in *S. cerevisiae* and *E. coli lig* mutant (Barnes *et al.* 1990, Kodama, Barnes and Lindahl 1991). Sequencing a radioactively labelled tryptic peptide-AMP complex (Tomkinson *et al.* 1991), determined that the adenylated lysine residue is encoded by exon 17, and located within the catalytic domain. It was also shown that the active site lysine is bracketed by hydrophobic amino acids on each side within a sequence (E-K-DG-R) common to ligases from many different organisms. Replacement of the lysine by another residue caused complete loss of enzyme activity (Kodama, Barnes and Lindahl 1991). Comparison of human DNA ligase I to ligases from vaccinia virus, *S. cerevisiae* and *S. pombe* identified an additional highly conserved sequence near the C-terminal of the catalytic domain (Smith, Chan and Kerr 1989). The 16 amino acid sequence (GISLRFPRFTRIREDK) has also been subsequently found to be present in DNA ligases III and IV (Wei *et al.* 1995), but is of unknown function.

Measurement of enzyme stability revealed DNA ligase I to have a half-life of 8 hours (Lasko *et al.* 1990), suggesting that transcriptional and translational regulation may only occur over long periods. It has been proposed that acute regulation is instead effected by post-translational mechanisms (Prigent *et al.* 1992). DNA ligase I is a phosphoprotein, with phosphorylation being essential for the formation of enzyme intermediates. Although the amino acid sequence contains putative phosphorylation sites for casein kinase II (CKII), p34cdc2, protein kinases-A and -C, only CKII, a serine/threonine kinase present in nuclei and cytoplasm of proliferating cells, affects enzyme activity *in vitro*. Seven possible CKII sites are present in the N-terminus of the human protein, although intriguingly the two 'strongest' candidates are absent from the murine enzyme (Savini *et al.* 1994, Jessop and Melton 1995). Nevertheless, a model for enzyme activation has been proposed; the primary translation product of DNA ligase I is enzymatically inactive due to the N-terminal "regulatory" domain interfering with intermediate formation. Post-translational modification, by phosphorylation of a serine residue in the N-terminal region, induces a conformational change, thereby relieving inhibition.

1.4.1.5 Interaction with other proteins

Further levels of control of enzyme activity may be achieved by interaction with other proteins. Yang *et al.* purified a 55-75kDa heat labile protein from HeLa cells that inhibits DNA ligase I (Yang, Becker and Chan 1993), while Kenne and Ljungquist isolated a heat stable protein from a BS cell line that stimulates DNA ligase I activity (Kenne and Ljunquist 1988) . DNA ligase I activity in joining double

strand breaks has been found to be dependent on the presence of another cellular protein, REP-1 (Fairman, Johnson and Thacker 1992). The roles of these proteins in controlling ligation activity have not yet been elucidated, but it has been shown that both Ku and PCNA modulate DNA ligase I activity.

Addition of Ku stimulates the activity of DNA ligases I, III, and IV *in vitro*, probably by acting as a 'bridging' or 'alignment factor' to stabilise intermolecular associations between DNA molecules (Ramsden and Gellert 1998). PCNA has been observed to inhibit the ligation activity of DNA ligase I *in vitro* (Jonsson, Hindges and Hubscher 1998). PCNA can interact with DNA ligase I via an 118aa region in the N-terminus of DNA ligase I (Levin *et al.* 1997). Construction of PCNA deletion mutants enabled Jonsson, Hindges and Hubscher to identify a hydrophobic pocket on PCNA involved in the interaction with polymerase δ, RF-C, DNase IV/FEN1, p21 and DNA ligase I (Jonsson, Hindges and Hubscher 1998). Comparison of these proteins identified a PCNA binding motif at the very N-terminus of DNA ligase I (residues 2-9) which may interact with the hydrophobic pocket of PCNA. This region of DNA ligase I was also identified as necessary to target the protein to replication foci (Montecucco *et al.* 1998). Residues 1-20 are sufficient to effect targeting, and are also required for interaction with PCNA. Mutation of the two phenylalanines in this sequence (see section 1.4.1.3) removed PCNA binding, while a peptide of amino acids 1-23 was able to compete for binding to PCNA. Taken together these data suggest that DNA ligase I is targeted to the replication foci by binding to PCNA (Montecucco *et al.* 1998).

It has also been reported that DNA ligase I affects the activity of other proteins *in vitro* (Mossi, Ferrari and Hubscher 1998). Addition of DNA ligase I stimulated RF-C-dependent loading of PCNA onto DNA. RF-C independent DNA elongation by polymerase δ was inhibited by addition of DNA ligase I, but it was demonstrated that this effect occurs through interaction with PCNA. Strand displacement by polymerase δ was also inhibited by DNA ligase I, the polymerase pausing after gap filling to allow ligation. Conversely polymerase ε activity was stimulated by addition of DNA ligase I, and further stimulated by supplementation with PCNA. It was concluded that DNA ligase I influences pol δ and ε in different ways through PCNA, and that this may reflect different roles of the two polymerases in DNA replication and repair (Mossi, Ferrari and Hubscher 1998). The inhibitory effect of DNA ligase I on DNA polymerisation by polymerase δ was also observed by Levin *et al.*, and they suggested that DNA ligase I may only interact with PCNA after the release of polymerase δ (Levin *et al.* 1997).

1.4.1.6 Function of DNA ligase I in DNA replication

A wealth of biochemical evidence has implicated DNA ligase I in lagging strand DNA synthesis. A 21S complex of enzymes known to participate in DNA synthesis was purified from HeLa cell homogenates (Malkas *et al.* 1990, Li, Goodchild and Baril 1994). Approximately 80% of the total ligase activity in HeLa extracts is associated with this 21S complex, which is composed of approximately 30 polypeptides, and functions to replicate SV40 *in vitro*. The observed ligase activity was affiliated with a ~125kDa protein that reacted with anti-DNA ligase I antibodies, and also had the catalytic properties of DNA ligase I (Li *et al.* 1993, Applegren *et al.* 1995).

Wu *et al.* purified a 17S multiprotein complex from the murine mammary carcinoma cell line FM3A (Wu *et al.* 1994). This complex can replicate polyoma virus DNA *in vitro*, and is apparently independent of non-specific interactions with other cellular proteins. Proteins present include; polymerase α and δ, DNA primase, PCNA, DNA helicase, topoisomerases I and II, and DNA ligase I, all of which were proposed to be part of the 'core' of a multiprotein DNA replication complex.

Using a synthetic DNA substrate and purified calf proteins, the enzymatic reactions for Okazaki fragment processing, and completion of lagging strand DNA replication were reconstituted *in vitro*. Purified DNA ligase I was sufficient to complete the reaction (Turchi and Bambara 1993, Turchi *et al.* 1994). Similarly it was possible to replicate SV40 *in vitro* using only purified proteins (Waga, Bauer and Stillman 1994). SV40 large tumour antigen (TAg), replication protein A (RPA), DNA topoisomerases I and II, DNA polymerase α/primase, replication factor C (RFC), proliferating cell nuclear antigen (PCNA), DNA polymerase δ, maturation factor I, and DNA ligase I. Partially purified DNA ligase III was not able to substitute for DNA ligase I, while T4 ligase was only a partially effective replacement.

Mackenney, Barnes and Lindahl used *in vitro* SV40 replication to further investigate the role of DNA ligase I (Mackenney, Barnes and Lindahl 1996). When using cell extracts from 46BR cells to support viral replication they found that, in keeping with other published data, newly replicated DNA was predominantly present as small fragments. However, the fragments were apparently unstable, not being incorporated into high molecular weight DNA with time, but instead simply being lost. Supplementing 46BR extracts with full length DNA ligase I corrected the replication defect, but DNA ligase III and T4 DNA ligase did not. The catalytic domain of DNA ligase I (lacking the N-terminal domain) worked only when present in excess, suggesting that the N-terminal region is required for efficient lagging strand DNA replication. Moreover, an excess of the N-terminal domain was able to inhibit replication in extracts from wild-type cells. From this it was concluded that the N-

terminal region may be important in the interaction of DNA ligase I with other replication proteins (Mackenney, Barnes and Lindahl 1996).

1.4.1.7 Functions in DNA repair and recombination

Other than studies of cells from the 46BR patient, direct evidence of the roles of DNA ligase I in the repair of DNA damage is limited. Expression of DNA ligase I is induced by damaging agents, and a mammalian cell variant with raised resistance to 3AB and DMS correlating to raised DNA ligase I activity has been documented (Murray *et al.* 1986). DNA ligase I has been reported to be present in a purified complex of proteins involved in NER (He and Ingles 1997), and has been used in the reconstitution of NER *in vitro* (Aboussekhra *et al.* 1995). Similarly DNA ligase I interacts with polymerase β within a multiprotein complex that catalyses the repair of a uracil-containing duplex DNA substrate by BER (Prasad *et al.* 1996), and the purified protein has been used in reconstitution of BER *in vitro* (Nicholl, Nealon and Kenney 1997) However, published data indicates the major BER pathway in mammalian cells to be DNA ligase III-dependent (see sections 1.2.2 and 1.4.2).

Reconstitution of V(D)J recombination in a cell-free system suggested that DNA ligase I may be preferred in V(D)J recombination (Ramsden, Paull and Gellert 1997). Recombination of artificial plasmid substrates by partially fractionated HeLa cell extract, RAG1 and RAG2 protein, and cofactors was assisted by further addition of purified DNA ligase I. Augmentation increased the diversity of coding joins formed, apparently by reducing the reliance on microhomologies between DNA ends, and also increased the efficiency and accuracy of signal joint formation. These effects were not observed on addition of either DNA ligases III or IV (Ramsden, Paull and Gellert 1997).

1.4.1.8 '46BR' mutations

Sequencing of cDNAs from the '46BR patient' revealed that she had two mis-sense mutations within the coding region of *LIG1* (Barnes *et al.* 1992b). By cloning and sequencing individual cDNA PCR products, it was demonstrated that the two alterations were on different alleles. Both changes are CpG→TpG transitions, possibly arising from inefficient repair after spontaneous hydrolytic deamination of 5-methylcytosine. One mutation (GAA→AAA) occurs within the active site of the enzyme, resulting in the glutamate at position 556 being modified to a lysine. This change produces an enzymatically inactive protein (Kodama, Barnes, and Lindahl 1991). The other alteration (CGG→TGG), inherited from the mother, is further towards the carboxy-terminus, and produces an Arg 771→Trp change.

46BR.1G1 is an SV40 transformed subline of 46BR (Webster et al. 1992). This cell line contains only cDNA sequences corresponding to the Arg771→Trp allele, either as a result of mitotic recombination leading to homozygosity, or chromosome loss leading to hemizygosity. Purified DNA ligase I from 46BR.1G1 showed a tendency to aggregate into dimers (Lehmann *et al.* 1988) and while the enzyme retains normal DNA binding, and intermediate formation properties, the ability to generate a phosphodiester bond is impaired. Hence, the defect of the 46BR enzyme lies in the conversion of DNA-AMP into the ligated DNA product. The final catalytic step occurs at only 3-5% of normal rate, and, as a consequence, DNA ligase I accumulates as an activated adenylate form *in vivo* (Prigent *et al.* 1994). Additional proof that the defect in 46BR lies at the DNA ligase I locus, came with the correction of the EMS and 3AB sensitivity phenotype of 46BR.1G1 by transfection of wild type *LIG1* sequences (Somia, Jessop and Melton 1993).

Although there have been reports that DNA ligase activity is diminished in T-cell acute lymphoblastic leukaemia cells (Rusquet, Feon and David 1988), no other DNA ligase I-deficient individuals have been identified. The 46BR patient is the only documented case in which a DNA sequence alteration can be directly correlated with a deficiency in ligation activity. In this respect, she appears to be 'a distinct genetic entity', and the first example of an inherited mutation in a mammalian replication enzyme (Webster *et al.* 1992).

1.4.1.9 "DNA ligase I mediates essential functions in mammalian cells"

During the course of work for this thesis, another group reported the use of gene targeting in murine embryonic stem cells to inactivate *Lig1* (Petrini, Xiao and Weaver 1995). Given the relevance to the work reported here, and the fact that strikingly different conclusions have been drawn from the two experiments, this paper will be described in detail. Further information on gene targeting is given in section 1.5.

Homologous recombination was used to introduce a partially disabled *neo* selectable marker gene into the endogenous mouse *Lig1* locus. The *neo* cassette replaced half of exon 17, and all of exons 18 and 19, in effect deleting sequences encoding the active site of the enzyme. This region had previously been demonstrated to be essential for enzyme activity (Kodama, Barnes and Lindahl 1991). ES cells heterozygous for the targeted allele were isolated, and used in an attempt to isolate ES cells homozygous for the targeted allele.

It has been demonstrated that it is possible to isolate ES cells homozygous for a targeted mutation by increasing the selection on cells heterozygous for a partially disabled selectable marker gene (Mortensen *et al.* 1992). Only those cells which have undergone a spontaneous gene conversion event to duplicate the marker gene (and

hence become homozygous) are able to survive elevated levels of selection. When ES cells heterozygous for the targeted *Lig1* allele were placed under high G418 selection, no viable clones arose. Nevertheless, it was found possible to isolate ES cells homozygous for the targeted allele if a *LIG1* cDNA expression construct was introduced into heterozygous cells prior to increasing selection. It was thus concluded that cells homozygous for the targeted mutation were inviable, and that DNA ligase I must "mediate essential functions in mammalian cells".

It was further reported that certain mutant forms of human DNA ligase I were unable to rescue homozygous ES cells. In addition to wild-type *LIG1*, *lig32B2* (an alternative splice variant of DNA ligase I isolated from a human tonsillar cDNA library lacking 30aa encoded by exon 2), and *lig231* (a 72kDa truncation mutant of DNA ligase I lacking the 231 N-terminal amino acids) were used. Although proteins produced from these two variant cDNAs had lower enzymatic activity (measured as 16 and 18% of wild-type levels respectively), both were able to complement the *S. cerevisiae cdc9* mutant. Nevertheless, it proved possible to only rescue homozygous ES cells with the *lig32B2* construct. The truncated *lig231* cDNA lacked sequences coding for the 'regulatory' N-terminal domain of the protein, and the failure of this construct to complement for DNA ligase I deficiency was taken to indicate that the N-terminus is required for function *in vivo*.

While cells rescued with the full length cDNA were indistinguishable from wild-type, cells rescued with the *lig32B2* construct showed DNA replication and DNA repair defects. In this case, although cell cycle distribution was indistinguishable from the wild-type, cells displayed a marginal increase in the percentage of newly replicated DNA present as small molecules (i.e. delayed Okazaki fragment joining), and were sensitive to killing by 3AB. The differential effects on replication and repair functions were hypothesised to reflect different protein-protein interactions being mediated by different regions of the N-terminus.

1.4.2 DNA ligase III

DNA ligase III was first identified during purification of DNA ligase II from calf thymus (Tomkinson *et al* 1991a). This ligase has different substrate specificity, immunological and purification properties compared to both DNA ligases I and II (see Lindahl and Barnes 1992) and is a polypeptide of 100kDa in size. Although in calf thymus DNA ligase III only contributes 5-10% of the total DNA ligation activity, in SV40-transformed human fibroblasts DNA ligase III contributes >85% of the high molecular weight DNA joining activity (Tomkinson, Starr and Schultz 1993). The enzyme is apparently not induced in proliferating cells (Elder and Rossignol 1990) and, unlike DNA ligase I, associates with condensed chromatin during metaphase of

the cell cycle (Mackey *et al.* 1997) suggesting that it is not involved in DNA replication.

A possible role for DNA ligase III in DNA repair was indicated by the isolation of a protein complex that repairs double stranded breaks and deletions by recombination (Jessberger *et al* 1993). The complex, deemed RC-1, contained at least five proteins including DNA polymerase (ε). A 5'-3' exonuclease and a DNA ligase activity were also identified. The ligase activity had a number of characteristics of DNA ligase III (although because this work was performed prior to the discovery of DNA ligase IV there remains the possibility that it was DNA ligase IV).

There is now a large accumulation of data indicating that DNA ligase III is a central component ot the BER pathway. It was noted during the purification of DNA ligase III that it associated with a second smaller polypeptide. This protein proved to be XRCC1, the product of a gene already known to be involved in the repair of DNA base damage (Caldecott *et al.* 1994). *XRCC1* had originally been cloned on the basis of the ability to complement the elevated SCE phenotype and sensitivity to ionizing radiation and alkylating agents of a mutant rodent cell line (EM9) (Thompson *et al.* 1990). Subsequent studies have shown that the activity of DNA ligase III is diminished (Ljungquist *et al.* 1994), and levels of the XRCC1-DNA ligase III complex are reduced in EM9 cells (Caldecott *et al.* 1995). With the discovery that the XRCC1-DNA ligase III complex interacts with DNA polymerase β and PARP it was hypothesised that the four proteins act in a multiprotein complex in the detection of strand breaks and repair by BER (Caldecott *et al.* 1996, Masson *et al.* 1998). Although it has also been reported that DNA ligase I associates with pol β, this apparently conflicting data can be explained by the presence of two alternative BER pathways (see section 1.2.2). The XRCC1-DNA ligase III complex functions solely in the single nucleotide insertion pathway, while DNA ligase I is active in the patch replacement (PCNA-dependent) pathway, EM9 cells only being deficient in the former (Cappelli *et al.* 1997).

1.4.2.1 Two isoforms of DNA ligase III

The DNA ligase III cDNA was cloned from a Hela cell cDNA library on the basis of homology to other DNA ligases (Wei *et al.* 1995). All eukaryotic DNA ligases show partial sequence homology with DNA ligase I, particularly around the active site. Expressed sequence tags (ESTs) displaying homology to this motif were identified from database searches, and these used to clone full length cDNAs. The *LIG3* cDNA contained an open reading frame of 2766bp, encoding a 922aa (103kDa) protein. The gene mapped to chromosome 17q11.2-12, and produced a 4.2kb transcript that was most abundant in testis, and thymus. Sequence analysis revealed

that within the N-terminus of the protein was a zinc finger domain similar to 'nick sensors' identified in PARP (Wei *et al.* 1995). When presented with a nicked DNA duplex, it was found that DNA ligase III was indeed able to detect nicked DNA, thereby potentially acting to target the whole XRCC1/DNA ligase III/polymerase β BER complex to the site of damage (Caldecott *et al.* 1996).

Soon after isolation of the *LIG3* cDNA by Lindahl and coworkers, another group reported isolation of a *LIG3* cDNA clone from a testis cDNA library using degenerate PCR based on peptide sequences from purified bovine DNA ligase III (Husain *et al.* 1995, Chen *et al.* 1995). Comparison of the testis cDNA with the HeLa cDNA sequence revealed a difference at the 3' end of the open reading frame, apparently arising from alternative splicing of the transcripts. The Hela cDNA (α) encoded a peptide of 103kDa, while the testicular form of the protein (β) was only 100kDa. Mapping of the 3' end of the *Lig3* locus showed that either one of two mutually exclusive final exons could be spliced onto the 3' end of the transcript. The α exon encodes 77aa, whereas the β exon codes only 17aa, thereby accounting for the protein size difference (Mackey *et al.* 1997), but also explaining the different characteristics of the β form to interact with XRCC1. It was determined that XRCC1 binds to the C-terminal 148aa of DNA ligase IIIα. The β isoform, which lacks most of this C-terminal region, does not interact with XRCC1, hence implying that only DNA ligase IIIα participates in BER.

Although DNA ligase IIIα is constitutively expressed in all tissues, the β form is specific to the mature testis, and is expressed at much higher levels (Chen *et al.* 1995, Mackey *et al.* 1997). Examination of the localization of DNA ligase IIIβ expression within the germ cell showed that elevated expression is restricted to the pachytene spermatocytes, beginning in the latter stages of meiotic prophase and ending in the round spermatid stage. Indirect immunofluorescence with anti-DNA ligase III antibodies illustrated that the protein associates with condensed chromatin incorporated into synaptonemal complexes. It was thus proposed that DNA ligase IIIβ may function specifically in meiotic recombination (Mackey *et al.* 1997).

1.4.3 DNA ligase IV

During the EST database search for human DNA ligase cDNAs, a third DNA ligase was identified. Surprisingly this proved not to encode DNA ligase II. A full length cDNA clone from a prostate library contained an open reading frame of 2532bp, encoding a 844aa (96kDa) protein. The gene mapped to chromosome 13q33-34, and produced a 4kb transcript that was most abundant in testis, thymus and prostate. In addition to the presence of ligase-characteristic motifs, and homology to the catalytic domains of other DNA ligases, the protein was demonstrated to have the

ability to form an enzyme-AMP complex and to act on an oligo(polydT)/poly(dA substrate). It was thus concluded that this gene encodes a novel, previously unidentified, DNA ligase (DNA ligase IV) (Wei *et al.* 1994).

Alignment of the three cloned DNA ligases on the basis of positions of conserved motifs (Figure 1.4.1) revealed that DNA ligase IV has an extended C-terminal region. Within this region two BRCA1 C-terminus (BRCT) modules have been identified. BRCT modules are autonomous folding units of 90-100aa, first described in BRCA1 but widespread in DNA-damage responsive and checkpoint proteins, that are involved in protein-protein interactions (Callebaut and Mornon 1997). It was therefore hypothesised that DNA ligase IV may be complexed with other proteins, and indeed the behaviour of the DNA ligase IV protein during purification procedures indicated that this was true (Robins and Lindahl 1996).

Other than genes encoding components of DNA-PK, one other gene has been directly implicated in V(D)J recombination and DSB repair: *XRCC4*. *XRCC4* was isolated on the basis of complementation of defective V(D)J recombination and DSB repair in a mutant cell line. The XRCC4 protein has no homology to any other known protein, but has been shown to interact with, and stimulate activity of, DNA ligase IV (Grawunder *et al.* 1997, Critchlow, Bowater and Jackson 1997). It has been proposed that and XRCC4-DNA ligase IV complex carries out the final ligation steps of both processes (Grawunder *et al.* 1997). This hypothesis is supported by demonstration that the *S. cerevisiae* homologue of DNA ligase IV is directly involved in the NHEJ pathway (Wilson, Grawunder and Lieber 1997, Teo and Jackson 1997, Schär *et al.* 1997, Ramos *et al.* 1997). Following the completion of the yeast genome sequencing project, database searches identified a second putative DNA ligase in yeast. This gene is not essential, and deletion confers no apparent growth phenotype, or sensitivity to DNA damaging agents. Nevertheless, reduced non-homologous end joining was observed in plasmid joining assays, and was accompanied by sensitivity to γ radiation when the *RAD52* pathway was also inactivated, indicating that DNA ligase IV functions in Ku-mediated repair of double strand breaks. Mutants sporulated less efficiently, and diploid cells were observed to be delayed in prophase I of meiosis, from which it was suggested that the protein may also play a role in specific recombination processes (Schär *et al.* 1997).

1.4.4 Other reported mammalian DNA ligase activities

1.4.4.1 DNA ligase II

A second DNA ligase was identified when it was discovered that two DNA ligase activities from calf thymus are separable by hydroxyapatite chromatography

and gel filtration (Söderhäll and Lindahl 1973b). Western blotting with an antibody specific for calf thymus DNA ligase II indicated that this enzyme consisted of a single 68kDa polypeptide (Teraoka and Tsukada 1986) while other workers estimated its size as 67kDa (Yang *et al* 1990) or 72kDa (Tomkinson *et al* 1991a).

Like DNA ligase I, DNA ligase II is a nuclear enzyme (Söderhäll and Lindahl 1975), but the two are immunologically and qualitatively distinct (Söderhäll and Lindahl 1975, Söderhäll and Lindahl 1973b, Arrand *et al* 1986, Tomkinson *et al* 1991a). Levels of DNA ligase II were reported to be maximal in the liver, and although increases in the level of DNA ligase II have been observed in response to DNA damage, the enzyme was apparently not induced in response to cell proliferation (Wang *et al.* 1994 and references therein).

However, it now appears that DNA ligase II is not derived from a novel gene. Rather it seems certain that DNA ligase II is somehow derived from DNA ligase III (Husain *et al.* 1995). Database search strategies which were used to clone cDNAs encoding DNA ligases III and IV failed to identify any ESTs corresponding to DNA ligase II (Wei *et al.* 1995). Purification of DNA ligases II and III followed by proteolytic mapping, and amino acid sequencing suggested that both enzymes are encoded by the same gene (Roberts, Nash and Lindahl 1994, Wang *et al.* 1994, Husain *et al.* 1995). 10 identical peptide sequences (136aa) were obtained from DNA ligases II and III, and all other peptide sequences from DNA ligase II could be aligned with the predicted amino acid sequence of *LIG3*. It was therefore concluded that DNA ligase II and III are encoded by the same gene, with DNA ligase II being an N-terminal truncated form of DNA ligase III. Exactly how DNA ligase II is derived is still to be determined. The N-terminus of DNA ligase II is resistant to Edman degradation, indicating the protein to be a primary translation product, and it has not been possible to convert purified DNA ligase III to DNA ligase II *in vitro*. However, alternatively spliced *LIG3* transcripts have not been observed in tissues where DNA ligase II is a predominant enzyme product (Husain *et al.* 1995).

1.4.4.2 DNA ligase V

A DNA ligase activity was identified during purification of proteins involved in double strand break repair in human cells (Fairman, Johnson and Thacker 1992). Further purification determined activity to be associated with 44kDa protein which was adenylated in the presence of DNA breaks. This protein, designated DNA ligase V, could ligate double strand breaks with a similar efficiency to DNA ligase I, but had little activity on single strand breaks (Johnson and Fairman 1997).

1.4.4.3 NHR ligase

In order to assay for the ability to repair chromosomal DSBs, joining of plasmid substrates by nuclear extracts from human cells was measured. Components of the end joining activity identified were purified further, but the ligase activity identified did not correspond to DNA ligases I, II or III. This non-homologous recombination (NHR) ligase could oligomerise monomer substrates and required both Mg^{2+} and ATP. The protein did not cross-react with anti-DNA ligase I antisera, had purification characteristics dissimilar to DNA ligase II and III, and enzymatic activity was stimulated by another component of the end joining activity (HPP-1) (Derbyshire *et al.* 1994). There does remain the possibility that NHR ligase corresponds to DNA ligase IV which had not been discovered at that time.

1.4.4.4 Mitochondrial DNA ligase

There are only two published reports of mitochondrial DNA ligases. Levin and Zimmerman identified a DNA ligase activity in rat liver mitochondria, but the enzyme responsible was not characterised (Levin and Zimmerman 1976). However, recently the purification of a DNA ligase from *Xenopus laevis* mitochondria has been detailed. Ligase activity was detected in a crude mitochondrial fraction, and further purification of the protein determined that the enzyme had different properties to the one known *Xenopus* ligase. Antibodies raised against a DNA ligase III-specific peptide detected both human DNA ligase III and the purified mitochondrial ligase, suggesting that it is a homologue of DNA ligase III.

1.5 Gene targeting

A number of animal diseases which mimic human conditions have been identified either arising spontaneously in animal populations or from random mutagenesis screens. However, factors such as time, cost and complexity constrain large scale screening of higher eukaryotes, so that until recently these 'animal models' have been too sparse to play a significant role in mammalian genetic studies. This problem has now been eliminated by the advent of gene targeting, whereby it is now possible to generate mice with defined genomic alterations (reviewed by Melton 1994, and Moore and Melton 1997). Numbers of gene targeted mice are growing rapidly and it has been estimated that by the year 2001, 3000 strains will be available (Simpson *et al.* 1997).

Gene targeting can be defined as "homologous recombination between chromosomal DNA sequences and newly introduced exogenous sequences, enabling the transfer of any desired alteration into the genome of a cell" (Capecchi 1989). Upon introduction of a targeting vector bearing DNA homologous to a gene of

interest, homologous recombination may initiate exchange of DNA sequences between the vector and the genomic locus. Formerly it was possible only to delete or "knock-out" genes, however a growing number of targeted alterations are now possible. Such alterations include conditional gene deletion (Gu *et al.* 1993), subtle alterations (e.g. single base pair changes, Moore *et al.* 1995) and targeted chromosomal translocations (Smith *et al.* 1995).

When DNA is introduced into cells it can have a number of possible fates, including nucleolytic degradation, random integration and homologous recombination with genomic DNA (for a review see Bollag, Waldman and Liskay, 1989). The limiting factor in gene targeting in higher eukaryotes is the marked preference for random integration of targeting vectors over homologous recombination. As noted in section 1.2.5, it is probable that gene targeting occurs through a *RAD52*-like recombination pathway (Essers *et al.* 1997), but homologous recombination mechanisms play only a minor role in higher eukaryotes compared to non-homologous end joining events. The ratio of random integration to homologous recombination events in mouse cells has been estimated at 1000:1 (Thomas and Capecchi, 1987).

In order to maximise the chances of identification and retrieval of targeted clones, gene targeting is performed with large numbers of cells, and utilising selection strategies which enrich for rare homologous recombination events. The most widely used selection scheme is the positive-negative strategy (Mansour, Thomas and Capecchi, 1988). Since both the transfection frequency of cells is relatively low, a positive selection scheme is required to select for cells which have integrated the targeting vector. The most common positive selection scheme in use is the bacterial neomycin phosphotransferase gene (*neo*) used in combination with G418, but a scheme based on the use of an *HPRT* minigene in combination with HAT selection has also been used successfully (Reid *et al.* 1990, Selfridge *et al.* 1992).

Positive selection schemes select both for cells which have undergone the desired gene targeting events and those which have undergone random integration events. It is therefore necessary to use a negative selection step to enrich for homologous recombinants. Selection against non-homologous recombinants with replacement vectors is achieved by introduction of negative selectable markers, such as the herpes simplex virus thymidine kinase gene (HSV-*tk*), at the termini of regions of homology (see Figure 1.5.1).

1.5.1 Production of "knockout" mice

Whilst gene targeting can be used to manipulate the genome of cells in culture, use of embryonic stem (ES) cells makes it additionally possible to transmit targeted

alterations to the mouse germline (Figure 1.5.2). ES cells are derived from the inner cell mass of 3.5 day mouse blastocysts and contribute to all cell lineages in the developing embryo (pluripotency). Moreover they retain this potential even after prolonged culture *in vitro* (Evans and Kaufman, 1981). Once gene targeting has been performed, the ES cells are injected into a blastocyst stage embryo, which is re-implanted into the uterus of a pseudopregnant foster female. Because ES cells are pluripotent, they can subsequently contribute to all tissues of the developing mouse, resulting in the production of 'chimaeric' animals. Providing the injected ES cells carry coat colour alleles different from that of the recipient embryo, any chimaeras can be identified by virtue of having a coat of mixed colours. Where injected ES cells contribute to the germ line, chimaeras will be able to pass markers derived from injected cells to their progeny (also identifiable on the basis of coat colour). Thus by carrying out gene targeting in ES cells and then using these cells for blastocyst injections, it is possible to produce mice carrying a defined genomic alteration (Thompson *et al.* 1989).

1.5.1.1 Gene targeting in HM-1, an HPRT deficient mouse ES line

The male embryonic stem cell line HM-1 is an HPRT (hypoxanthine-guanine phosphoribosyltransferase HPRT; IMP pyrophosphate phosphoribosyltransferase; EC 2.4.2.8) deficient male ES cell line which can contribute to the mouse germline with high frequency (Magin *et al.* 1992b). HM-1 cells have undergone spontaneous deletion of approximately 10kb of sequence 5' of the *HPRT* locus, resulting in loss of *HPRT* expression (Thompson *et al.* 1989). Because *HPRT* is X-linked, and male ES cells are hemizygous for this gene, loss of function of the one *HPRT* allele gives rise to HPRT-deficient cells.

HPRT catalyses an early step of the purine salvage pathway in mammals. In humans loss of HPRT function results in Lesch Nyhan disease (Lesch and Nyhan, 1964), but HPRT activity is non-essential in cell lines *in vitro* (provided that the main purine synthesis pathway is functional). It is possible to select for HPRT activity *in vitro* with medium supplemented with hypoxanthine, aminopterin and thymidine (HAT). Aminopterin is an inhibitor of *de novo* synthesis of purines and pyrimidines. Only when HPRT activity is present can cells bypass the aminopterin-induced metabolic block by utilising hypoxanthine and thymidine in the purine and pyrimidine salvage pathways respectively. Hence by incorporating *HPRT* minigenes into gene targeting vectors as selectable markers, HPRT-deficient ES cells can be used to produce "knockout" mice (Selfridge *et al.* 1992). Figure 1.5.1 and 1.5.2 outline a typical gene targeting strategy employed in HM-1 ES cells.

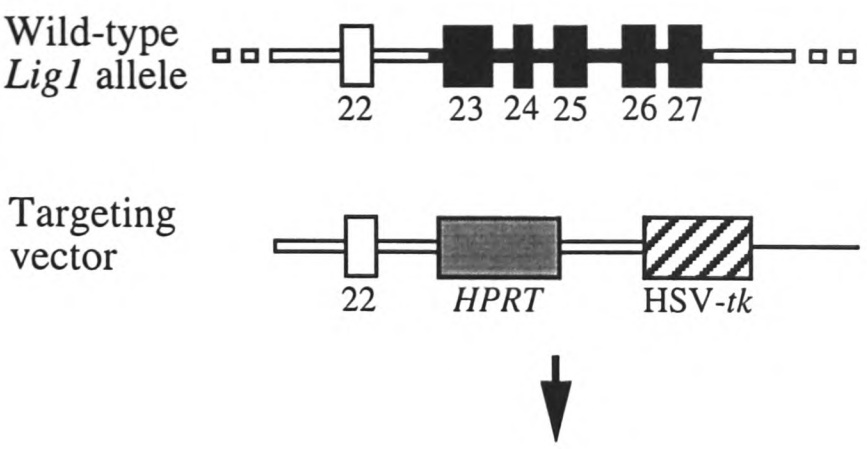
Figure 1.5.1: A gene targeting strategy employed in HM-1 ES cells

A positive-negative selection procedure used in HPRT-deficient cells to enrich for homologous recombination events. Schematic representation of 13.5kb at the 3' end of the wild-type mouse DNA ligase I locus, linear targeting vector, and gene structure following a) homologous recombination or b) random integration. Boxes, exons (numbered according to the human convention); thick lines, *Lig1* introns or flanking regions; thin lines, plasmid vector DNA. Solid boxes and lines, regions of *Lig1* replaced by targeting; stippled box, *HPRT* minigene; hatched box, Herpes simplex virus thymidine kinase (HSV-*tk*) cassette.

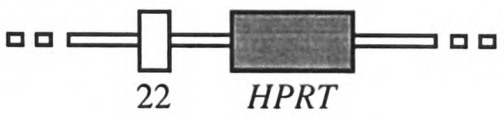
The targeting vector consist of 5' and 3' flanking sequence homologous with the target locus, an *HPRT* minigene, and an HSV-*tk* cassette outwith the region of homology (plus plasmid vector DNA). The *HPRT* minigene has the dual role of replacing essential exons within the target gene, and functions as a positive selectable marker for stable transfectants.

a) Replacement gene targeting vectors integrate following a double reciprocal recombination event resulting in the complete replacement of a region of the target locus with vector sequence. Negative selection against non-homologous recombinants is used to enrich for homologous recombinants. This is achieved by placing an HSV-*tk* negative-selectable marker outside the region of vector homology with the target gene. Regions at the vector termini which are not homologous to the target gene will be excluded from the final structure of the correctly targeted locus. This renders the cells resistant to the toxic viral thymidine kinase substrate gancyclovir and HAT selection.

b) The majority of random integrants will be stably transfected with the whole vector including the HSV-*tk* marker as well as the *HPRT* minigene. As a consequence, although cells will be resistant to HAT selection, they will be selectively killed by gancyclovir.



a) Homologous recombination



b) Random integration

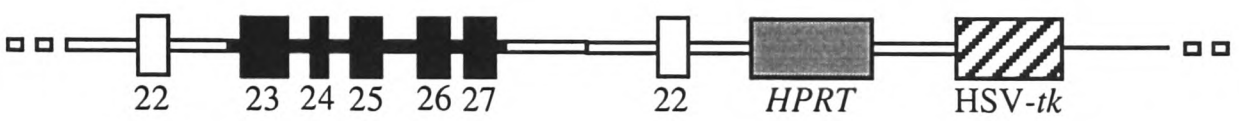


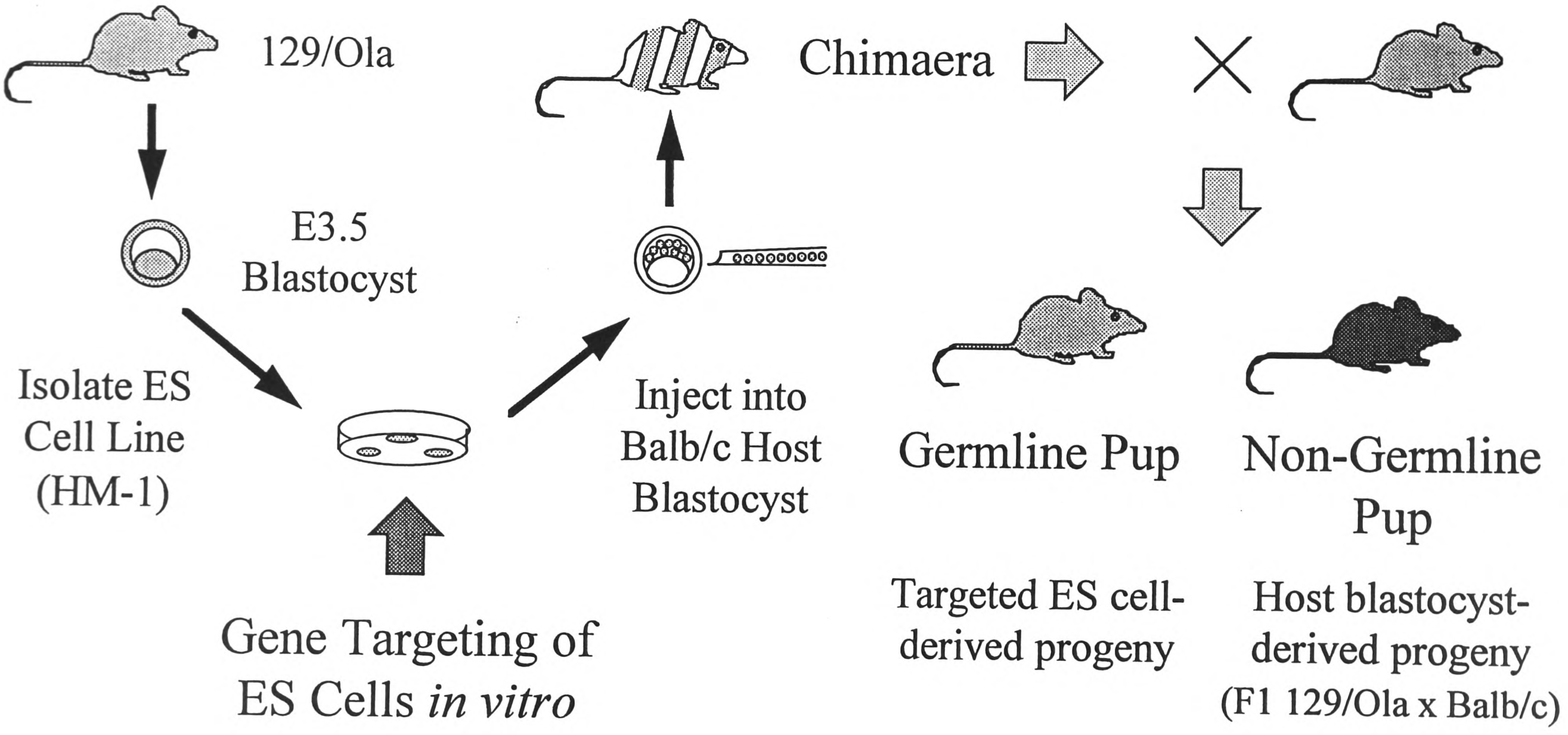
Figure 1.5.2: Generation of mice from gene targeted embryonic stem cells

Schematic representation of the procedure to generate mice heterozygous for a targeted alteration. Adapted from Moore 1997.

ES cells can be genetically modified *in vitro*, then introduced into blastocysts and, following embryo transfer and successful implantation, can generate germline chimeras. Coat and eye colour markers aid the identification of ES cell contribution to chimeras and to germline pups.

For autosomal loci, only 50% of germline pups carry targeted alleles because ES cells are usually heterozygous for targeted alterations. Germline transmission of 129/Ola HM-1-ES cell derived gametes is determined by the transmission of 129/Ola coat and eye colour markers (Selfridge et al., 1992): 129/Ola are homozygous for the white bellied agouti allele (A^w) and the chinchilla allele (c^{ch}) at the albino locus which results in a light brown coat colour. 129/Ola are also homozygous for the recessive allele, pink-eyed dilute (p) and have pink eyes.

When used in conjunction with recipient blastocysts which differ from 129/Ola at these loci, it is possible to determine the contribution of ES derived tissue in chimeras. For example, introduction of 129/Ola ES cells into Balb/c recipient blastocysts (Balb/c mice have a white coat and pink eyes: *c/c*, *P/P*, *A/A*) results in white mice (blastocyst coat colour) and chimaeric, 129/Ola derived tissue, shows up against this coat colour as light brown. Following test crosses between chimeras and inbred 129/Ola mice, the transmission of ES gametes from germline chimeras is indicated by a chinchilla coat and pink eyes in progeny. The targeted allele of a gene will be transmitted to 50% of the offspring, and these germline pups will be heterozygous for that allele. Non-germline pups have a grey coat and black eyes as they are F1(Balb/c x 129/Ola) i.e. they are c^{ch}/c, p/P, A^w/A.


129/Ola
E3.5
Blastocyst
Isolate ES
Cell Line
(HM-1)
Gene Targeting of
ES Cells *in vitro*
Inject into
Balb/c Host
Blastocyst
Chimaera
Germline Pup
Targeted ES cell-
derived progeny
Non-Germline
Pup
Host blastocyst-
derived progeny
(F1 129/Ola x Balb/c)

1.5.2 Double replacement gene targeting

In many cases, mutations that are the root of a human disease do not disrupt gene expression completely. It is therefore more desirable to be able to make subtle alterations to a gene. To this end, a method of 'double replacement' gene targeting has been developed to exploit the ability to select both for and against the presence of *HPRT* (Stacey *et al.* 1994, Melton 1997). This method was first suggested by Reid *et al.* (Reid *et al.* 1990). An initial round of gene targeting is used to introduce an *HPRT* selectable marker into the desired locus, enabling a "knockout" allele to be generated if required. Subsequently, a second round of gene targeting is initiated. The second step employs a replacement vector consisting of sequence homologous to the target gene sufficient to reconstruct the disrupted allele, but bearing an engineered mutation. Homologous recombination between the "knockout" allele and the second step targeting vector results in the replacement of the selectable marker by sequences bearing the engineered mutation (Figure 1.5.3). It is possible to enrich those cells which have homologously recombined with the alteration vector by selecting for the loss of the *HPRT* minigene. Selection against HPRT activity *in vitro* is achieved by supplementing culture media with the cytotoxic base analogue 6-thioguanine.

1.5.3 Cre/*loxP*-mediated gene targeting

The ability to modify the mouse genome has been further extended by the use of site-specific recombinases in conjunction with homologous recombination. The strategy most commonly used exploits the *loxP*/Cre recombinase system from the bacteriophage P1 (reviewed by Kühn and Schwenk 1997). Cre is a site-specific DNA recombinase which recognises a 34bp target sequence (*loxP* site) (Kilby, Snaith and Murray 1993). Cre-mediated recombination between two *loxP* sites results in the reciprocal exchange of DNA flanking the two sites, but the products formed by the recombination reaction will depend on the location and orientation of the two *loxP* sites.

When two *loxP* sites are in the same orientation on the same DNA molecule, recombination will excise the intervening DNA segment as a circular molecule, leaving a single *loxP* site on each reaction product. By flanking a gene of interest with two *loxP* sites (termed 'floxing'), it is thus possible to excise the genomic locus, and hence delete the gene by Cre-mediated recombination. Conversely, if the recognition sites are in the opposite orientation, inversion of the intervening region will occur. Intermolecular recombination between *loxP* sites on different DNA molecules can also take place. By selecting for the desired recombinant products it has proved possible to produce site directed chromosomal translocations (Smith *et al.* 1995),

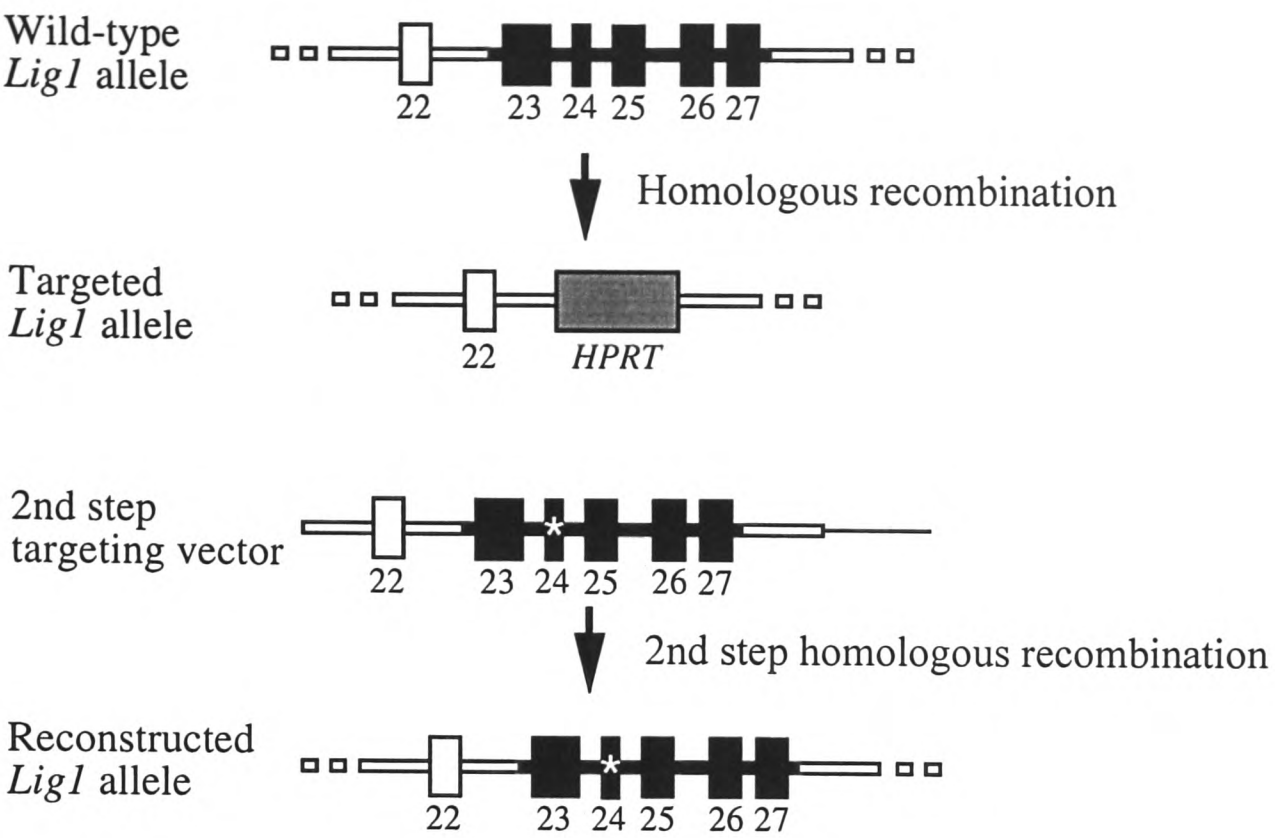
inversions and duplications (Ramírez-Solis, Liu and Bradley 1995). Further refinement has allowed site directed mutagenesis (Torres *et al.* 1996), gene insertion ("knock-in") (Kitamoto *et al.* 1996) and large chromosomal deletions (Li *et al.* 1996).

In each of these strategies, modification of the genome involves two steps; introduction of the *loxP* sites into the genome followed by Cre-mediated recombination. Whilst it is possible to carry out both of these steps in ES cells prior to blastocyst injection, because recombination efficiency is high it is possible to carry out the Cre-mediated recombination step *in vivo*. Cre recombinase is not dependent on cofactors, and so Cre-mediated recombination between two *loxP* sites requires only expression of the protein. Control of Cre expression can thus be used to limit recombination to a particular tissue, cell lineage, or timepoint (termed conditional gene targeting). One way of doing this is to cross mice harbouring a 'floxed' gene to a strain of transgenic mice expressing Cre under the control of a tissue-specific promoter. The 'floxed' gene will only be deleted in those cells expressing Cre, and hence other tissues will be unaffected. Conditional gene targeting therefore offers the potential to permit analysis of gene knockouts that are otherwise lethal.

Figure 1.5.3: Introduction of subtle alterations by double replacement gene targeting

A two step 'double replacement' strategy for the introduction of a subtle alteration into an endogenous gene. Schematic representation of the 3' end of the wild-type mouse DNA ligase I locus, the 'knockout' *Lig1* allele, linear second step targeting vector, and gene structure following homologous recombination. Boxes, exons; thick lines, *Lig1* introns or flanking regions; thin lines, plasmid vector DNA. Solid boxes and lines, regions of *Lig1* replaced by targeting; stippled box, *HPRT* minigene; asterisk, engineered mutation.

The second step targeting vector consists of *Lig1* sequences sufficient to reconstruct the disrupted locus, but bearing an engineered mutation. Following a double reciprocal recombination event between regions of homology between the "knockout" allele and the introduced vector, the *HPRT* minigene will be lost. Selection against the *HPRT* minigene with 6-thioguanine is sufficient to enrich for targeted ES cells as random integration of the vector will not cause the loss of HPRT activity.


Wild-type
Lig1 allele
22
23 24 25
26 27
Homologous recombination
Targeted
Lig1 allele
22
HPRT
2nd step
targeting vector
*
22
23 24 25
26 27
2nd step homologous recombination
Reconstructed
Lig1 allele
*
22
23 24 25
26 27

1.6 Project Aim

A number of distinct DNA ligase activities have been identified within mammalian cells, and three of these have been cloned (DNA ligases I, III, and IV). Various lines of evidence have indicated DNA ligase I to be involved in DNA replication, repair and recombination processes. As a result it was suggested that DNA ligase I is essential for cell viability, and it has been reported to be impossible to isolate cells lacking the enzyme.

One patient ('46BR' patient) has been identified with reduced DNA ligase I activity correlating to missense mutations within the protein coding sequence. This individual displayed a phenotype similar to Bloom's Syndrome, being immunodeficient, growth retarded, and predisposed to cancer. Fibroblasts exhibited abnormal lagging strand DNA synthesis and repair deficiency. Despite extensive screening, no other DNA ligase I-deficient individuals have been identified. Further study of the roles of DNA ligase I within the mammalian system would therefore be greatly facilitated by an experimental model of DNA ligase I-deficiency.

The work described in this thesis is part of an ongoing project to produce a mouse model of the '46BR' patient by gene targeting. Gene inactivation was the first step of a 'double replacement' strategy to introduce analogous mutations into the mouse DNA ligase I gene (*Lig1*). Utilising established techniques, the 3' end of the endogenous *Lig1* gene was replaced with an *HPRT* minigene, and this alteration transferred into the mouse germline. I report here investigation of the developmental and molecular consequences of targeting of *Lig1*. In addition, efforts to modify *Lig1* *in vivo* via the Cre/*loxP* system are also described.

CHAPTER TWO:
MATERIALS AND METHODS

2.1 Materials

2.1.1 Suppliers of equipment, laboratory reagents and consumables

Amersham International plc, Lincoln Place, Green End, Aylesbury, Buckinghamshire, HP20 2TP

Becton Dickinson Labware, Becton Dickinson and Company, Franklin Lakes, NJ, USA 87417-1886

BDH Ltd., Burnfield Avenue, Thornliebank, Glasgow, G46 7TP

Biogenesis Ltd., 7 New Fields, Stinsford Road, Poole, England, BH17 0NF

Bio-Rad Laboratories Ltd., Bio-Rad House, Maylands Avenue, Hemmel Hempstead, Hertfordshire, HP2 7TD

Boehringer Mannheim UK Ltd., Bell Lane, Lewes, East Sussex, BN7 1LG

The Boots Company plc, 1 Thane Road West, Nottingham NG2 3AA

Caltag Ltd., 1849 Old Bayshore Highway, Suite 200, Burlingame, CA 94010

Canberra Packard, Brook House, 14 Station Road, Pangbourne, Berkshire, RG8 7DT

Coulter Electronics Ltd., Northwell Drive, Luton, Bedfordshire, LU3 3RH

Difco Laboratories, PO Box 14B, Central Avenue, West Molesey, Surrey, KT8 2SE

Du Pont (UK) Ltd., Diagnostics and Biotechnology Systems, NEN Products, Wedgwood Way, Stevenage, Hertfordshire, SG1 4QN

Fisher Scientific UK Ltd., and Fisons Scientific Equipment, Bishop Meadow Road, Loughborough, Leicestershire, LE11 0RG

Forma Scientific, Life Sciences International (UK) Ltd., Unit 5, The Ringway Centre, Edison Road, Basingstoke, Hampshire, RG21 6YH

Gibco BRL Life Technologies Ltd., PO Box 35, Trident House, Renfrew Road, Paisley, Renfrewshire, PA3 4EF

Hoefer, 654 Minnesota Street, San Francisco, CA 94107, USA

Hybaid Ltd., Life Sciences International (UK) Ltd., Unit 5, The Ringway Centre, Edison Road, Basingstoke, Hampshire, RG21 6YH

ICN Biomedicals Ltd., Thame Park Business Centre, Wenman Road, Thame, Oxfordshire, OX9 3XA

Millipore (UK) Ltd., The Boulevard, Blackmore Lane, Watford, Hertfordshire, WD1 8YW

Molecular Dynamics Ltd., 5 Beech House, Chiltern Court, Asheridge Road, Chesham, Buckinghamshire, HP5 2PX

National Diagnostics, 305 Patton Drive, Atlanta, Georgia 30336, USA

New England Biolabs (UK) Ltd., 67 Knowl Piece, Wilbury Way, Hitchin, Hertfordshire, SG4 0TY

New England Nuclear Life Science Products, 549 Albany Street, Boston, MA 02118, USA

Oswel DNA Service, Lab 5005, Medical and Biological Sciences Building, University of Southampton, Boldrewood, Bassett Crescent East, Southampton, SO16 7PX

Packard Instrument Company, 800 Research Parkway, Meriden, CT 06450, USA

Perkin Elmer Applied Biosystems, Warrington, WA1 4SR

Pharmacia Biotech, 23 Grosvenor Road, St. Albans, Hertfordshire, AL1 3AW

Pharmingen, 10975 Torreyana Road, San Diego, CA 92121, USA

Pierce, PO Box 117, Rockford, Illinois 61105, USA

Promega UK Ltd., Delta House, Chilworth Research Centre, Southampton SO16 7NS

Qiagen Ltd., Unit 1 Tillingbourne Court, Dorking Business Park, Station Road, Dorking, Surrey, RH4 1HJ

Scotlab Ltd., Kirkshaws Road, Coatbridge, Strathclyde, ML5 8AD

Sigma-Aldrich Company Ltd., Fancy Road, Poole, Dorset, BH17 7NH

Stratagene Ltd., 140 Cambridge Innovation Centre, Cambridge Science Park, Milton Road, Cambridge, CB4 4GF

2.1.1.1 Laboratory reagents

BDH Ltd.: dimethyldichlorosilane, DPX mountant, hydrochloric acid, Sarkosyl NL30, xylene, xylene cyanol.

Boehringer Mannheim UK Ltd.: deoxyribonucleotides, proteinase K.

The Boots Company plc: nail varnish, skimmed milk powder.

Difco Laboratories: agar, bacto-tryptone, bacto-yeast extract.

Fisher Scientific UK Ltd./Fisons Scientific Equipment: 3M blotting paper, boric acid, butanol, calcium chloride, citric acid, disodium hydrogen orthophosphate, disodium hydrogen orthophosphate, EDTA, filter paper, glacial acetic acid, hydroxyquinolin, isoamyl alcohol, isopropanol, potassium dihydrogen orthophosphate, sodium acetate, sodium chloride, sodium citrate, sodium hydroxide, sodium dihydrogen orthophosphate, potassium acetate, trichloroacetic acid.

Gibco BRL Life Technologies Ltd.: 100bp DNA marker ladder, agarose, formamide, guanidine hydrochloride, *Hin*d III digested λ DNA marker ladder, phenol, Tris, low melting point agarose, urea.

ICN Biomedicals Ltd.: Dulbecco's PBS tablets.

Millipore (UK) Ltd.: disposable sterile filters, Immobilon P.

National Diagnostics: ecoscint A.

New England Biolabs (UK) Ltd.: prestained broad range molecular weight marker proteins, 3X reducing SDS sample buffer.

New England Nuclear Life Science Products: Genescreen *Plus*.

Pierce: BCA* protein assay reagents A and B.
Promega (UK) Ltd.: BCIP, NBT, oligo dT, random hexamers, rRNasin.
Qiagen Ltd.: tip 500 columns.
Scotlab Ltd.: acrylamide/bis-acrylamide solutions.
Sigma-Aldrich Company Ltd.: ammonium persulphate, bovine serum albumin, bromophenol blue, dextran sulphate, DTT, ethidium bromide, ficoll, formaldehyde, Giemsa stain, glycerol, glycine, Harris eosin Y alcoholic solution, Harris eosin Y aqueous solution, Harris hematoxylin solution, herring sperm DNA, Hoechst 33258, IPTG, mineral oil, MOPS, new methylene blue, NP-40, parafilm M, PMSF, Ponceau S solution, sodium bicarbonate, TEMED, thymidine, Triton X-100, Tween 20, X-Gal.

2.1.1.2 DNA/RNA modifying enzymes

Boehringer Mannheim UK Ltd.: Klenow polymerase, M-MuLV reverse transcriptase, restriction endonucleases.
Gibco BRL Life Technologies Ltd.: bacterial alkaline phosphatase, restriction endonucleases, T4 DNA kinase, *Taq* DNA polymerase.
New England Biolabs (UK) Ltd.: restriction endonucleases, T4 DNA ligase.

2.1.1.3 Mammalian cell culture reagents

Becton Dickinson Labware: plasticware.
Boehringer Mannheim UK Ltd.: colcemid.
Difco Laboratories: trypsin.
Gibco BRL Life Technologies Ltd.: foetal calf serum, Glasgow's Modified Eagle Medium (BHK21), horse serum, L-glutamine, new-born calf serum, non-essential amino acids, sodium bicarbonate, sodium pyruvate.
Sigma-Aldrich Company Ltd.: β-mercaptoethanol, ES cell trypsin.

2.1.1.4 Radioactive reagents

Amersham International plc.: Redivue [α-^{32}P]-dCTP (~3000 Ci/mmole, 10mCi/ml), [methyl-^{3}H]-thymidine (25 Ci/mmole, 1mCi/ml).

2.1.1.5 Antibiotics

Sigma-Aldrich Company Ltd.: ampicillin, penicillin G, streptomycin.

2.1.1.6 Antisera

Caltag Ltd.: Anti-mouse IgM F(ab')2 PE-conjugated affinity purified goat IgG, anti-mouse CD4 PE-conjugated rat monoclonal IgG2a (clone CT-CD4), anti-mouse CD8a FITC-conjugated rat monoclonal IgG2a (clone CT-CD8a), anti-mouse B220 PE-

conjugated rat monoclonal IgG2a (clone RA3-6B2), anti-mouse Gφ FITC-conjugated rat monoclonal IgG2b (clone RB6-8C5).

Pharmingen: Anti-mouse T200 PE-conjugated rat monoclonal IgG2b (clone 30F 11.1), anti-mouse TCRαβ FITC-conjugated hamster monoclonal IgG (clone H57-597), anti-mouse TER 119 FITC-conjugated rat monoclonal IgG2b (clone TER-119), anti-mouse H-2Kd FITC-conjugated rat monoclonal IgG2a (clone SF1-1.1), anti-mouse H-2Kb PE-conjugated rat monoclonal IgG.

Promega UK Ltd.: Anti-rabbit IgG alkaline phosphatase-conjugated mouse monoclonal IgG

Sigma-Aldrich Company Ltd.: Anti-mouse CD5 FITC-conjugated rat monoclonal IgG (clone 53-7.3).

Antisera raised against DNA ligases I, III, and IV were kindly provided by Dr. Tomas Lindahl and co-workers. TL5 is a rabbit polyclonal antibody raised against purified full length bovine DNA ligase I, and TL6 is a rabbit polyclonal antibody raised against a conserved peptide present in the C-terminus of DNA ligase I (Tomkinson *et al.* 1990). TL18 is a rabbit polyclonal antibody raised against a peptide in the C-terminus of human DNA ligase IV, whilst TL25 is an affinity purified rabbit polyclonal antibody raised against recombinant full length human DNA ligase III (T. Lindahl, personal communication).

2.1.1.7 Oligonucleotides

All oligonucleotides were synthesised by the Oswel DNA service. The sequences of all primers used are given in Table 2.1.1.

2.1.2 Media

2.1.2.1 Bacterial media

Luria Broth (LB): 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, and 0.5% (w/v) NaCl, adjusted to pH 7.2 with NaOH.

Luria Agar: 1.5% agar, 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, and 0.5% (w/v) NaCl, adjusted to pH 7.2 with NaOH.

Terrific Broth (TB): 1.2% (w/v) bacto-tryptone, 2.4% (w/v) bacto-yeast extract, 0.4% (v/v) glycerol, 170mM potassium dihydrogen orthophosphate, and 72mM dipotassium hydrogen orthophosphate.

When required, ampicillin was added immediately prior to use, to a final concentration of 100μg/ml.

Table 2.1.1: PCR primers

Name	Sequence (5'→3')	Description*
033M	CCAGTGTTGAAGTTTGTGCG	*Ercc-1* exon 4 (429-448) (McWhir *et al.* 1993)
035M	CGAAGGGCGAAGTTCTTCCC	*Ercc-1* exon 5 (598-579) (McWhir *et al.* 1993)
262W	AGCCTACCCTCTGGTAGATTGTCG	*HPRT* exon 9 (960 to 984) (Stacey *et al.* 1994)
348M	GATGTCAAGGACCTTCAG	*Prn-p* intron 2 (38-57) (Moore *et al.* 1995)
B188	GCCTAGACCACGAGAATGCG	*Prn-p* exon 3 (916-896) (Moore *et al.* 1995)
G1022	GGAAAACACCGGTAAGTCAGG	Binds 3' to *Lig1* (D. W. Melton, unpublished data)
G4607	CTCAGAAGCCAGGTACATTG	*Lig1* exon 13 (1606 to 1625)
G4608	CTGCCAACATGCGAGAAAGT	*Lig1* exon 1 (297 to 316)
G8320	TCCTTCACTGACTGCTCAAG	*Lig1* exon 21 (2448 to 2428)
H3123	CAGTTCGACGCTTTGGGAAT	*Lig1* exon 4 (674 to 655)
HLIG33'	GGGAAGGTTAGTGGCAGATTTCCA	*LIG3* cDNA (2542 to 2519)
HLIG35'	ATGAAGAAATGTCCCAATGGCATG	*LIG3* cDNA (1558 to 1581)
HLIG43'	TTGTTCTAGGTCGTCCAGGGTCAT	*LIG4* cDNA (2076 to 2053)
HLIG45'	GAGCACATTGAGAAGGATATG	*LIG4* cDNA (1042 to 1062)
LIG43'	CAAGGTGCAGCCAGTTTTATACAT	*LIG4* cDNA (1998 to 1975)
LIG45'	ATTCTTGATGGTGAGATGATGGC	*LIG4* cDNA (1252 to 1274)
M5175	CGCGGATCCGCTGAAGAAGGACTACCTTGA	*Lig1* exon 23 (2531 to 2550)
M5176	CCGGAATTCCGGCTTGCATATAGCCTGAAGCTCTTC	*Lig1* exon 23 (2683 to 2660)

Table 2.1.1: PCR primers (continued)

Name	Sequence (5'→3')	Description*
N2468	CTGCAGGCTCTGGTGGTGTTCCTCCAGC	*Lig1* exon 24 (2737 to 2710)
N2469	CCTGTGAGGGCCTGATGGTGAAGACCTTGG	*Lig1* exon 22 (2451 to 2479)
G7853	ATTAACCCTCACTAAAGGGA	T3 promoter in pBluescript II KS (795-775) (Thummel *et al.* 1988)
G7854	AATACGACTCACTATAGGGC	T7 promoter in pBluescript II KS (622-642) (Thummel *et al.* 1988)
V0727	GTCAGATTCAGAACCAACAAAG	*Lig1* exon 27 (3000 to 3021)
V2014	AGTGTTCCCATGGCACAAGTGGCTGAAGC	Binds 3' to *Lig1* (D. W. Melton, unpublished data)
Zfy1	GACTAGACATGTCTTAACATCTGTCC	Binds within *Zfy-1* (Kunieda *et al.* 1992)
Zfy2	CCTATTGCATGGACTGCAGCTTATG	Binds within *Zfy-1* (Kunieda *et al.* 1992)

* Figures in brackets indicate the binding site of the primer within the appropriate cDNA sequence (quoted 5'→3'). Genbank database accession numbers of the cDNA sequences are as follows: *Lig1*; U19604, *LIG3*; X84740, *LIG4*; X83441, *HPRT*; J00423, *Prn-p*; M18070, Accession numbers of primers Zfy1 and Zfy2 are A20479 and A20478 respectively.

2.1.2.2 Mammalian tissue culture media

Glasgow's Modified Eagles Medium (GMEM) (McPherson, 1962, with modifications by W. House, Medical Research Council Institute of Virology, University of Glasgow, 1964) was used as the base culture medium for mouse and human fibroblast cell lines, ES cells, and haematopoietic cells.

Freezing medium for long term storage of cells consisted of the appropriate supplemented culture medium with additional DMSO and FCS to final concentrations of 10% (v/v) and 20% (v/v) respectively.

Bottom agarose was prepared by addition of equal volumes of 2X alpha medium and molten 1.2% (w/v) agarose. The mixture was equilibrated to 42°C before $^{1}/_{10}$ volume of L929 and $^{1}/_{40}$ volume of AF1 factors were added. L929 and AF1 were a gift from G. Graham (Beatson Institute, Glasgow) to J. D. Ansell.

Top agarose was prepared by addition of equal volumes of 2X alpha medium and molten 1.2% (w/v) agarose.

Methylcellulose medium was prepared by mixing equal volumes of 2X alpha medium with 18% (w/v) methylcellulose, and supplemented with growth factors required for erythroid differentiation (1ng/ml interleukin-3, and 2units/ml erythropoietin).

ES cell culture medium: 0.1mM β-mercaptoethanol, 5% (v/v) foetal calf serum (FCS), 5% (v/v) new-born calf serum (NCS), 1X non-essential amino acids, 1mM sodium pyruvate, leukaemia inhibitory factor, in GMEM.
Fibroblast culture medium: 10% (v/v) FCS, 1X non-essential amino acids, 25U/ml penicillin, 1mM sodium pyruvate, 25μg/ml streptomycin, in GMEM.
Serum free fibroblast culture medium: 1X non-essential amino acids, 25U/ml penicillin, 1mM sodium pyruvate, 25μg/ml streptomycin, in GMEM.
2X alpha medium: 50% (v/v) horse serum, 2mM L-glutamine, 0.45% (w/v) sodium bicarbonate, 1mM sodium pyruvate in GMEM.
Lymphocyte proliferation medium: 10% (v/v) FCS, 1X non-essential amino acids, 25U/ml penicillin, 1mM sodium pyruvate, 25μg/ml streptomycin, in GMEM.

2.1.3 Bacterial strains

The *E. coli* strain DH5α was used for the propagation and purification of all plasmids. DH5α has the genotype *supE44, ÆlacU169 (ϕ 80lacZÆM15), hsdR17, RecA1, endA1, gyrA96, thi-1*, and was first described by Hanahan 1983.

2.1.4 Mammalian cell lines

The human fibroblast cell lines and mouse ES cell lines utilised are outlined in Table 2.1.2.

2.1.5 Plasmids and cloning vectors

The plasmids and cloning vector employed are described in Table 2.1.3.

2.1.6 Buffers

alkaline phosphatase buffer: 100mM Tris-HCl pH 9.5, 100mM NaCl, 5mM $MgCl_2$
alkaline running buffer: 50mM NaOH, 1mM EDTA
6X alkaline sample buffer: 300mM NaOH, 6mM EDTA, 18% (w/v) Ficoll (Type 400), 0.15% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF
bacterial alkaline phosphatase buffer: 530mM Tris-HCl pH 7.5, 95mM $MgCl_2$, 50mM DTT
CA: 24 parts chloroform, 1 part isoamyl alcohol
citrate saline: 0.6% (w/v) sodium citrate, 0.72% (w/v) sodium chloride
Coomassie stain: 0.025% (w/v) Coomassie blue R-250, 40% (v/v) methanol, 7% (v/v) glacial acetic acid
Coomassie destain: 40% (v/v) methanol, 7% (v/v) glacial acetic acid
denaturation buffer: 0.5M NaOH, 1.5M NaCl
electroblotting buffer: 48mM Tris-HCl, 39mM glycine, 20% (v/v) methanol, pH 9.2
electroporation buffer: 21mM HEPES, 5mM D-glucose, 8mM Na_2HPO_4, 5mM KCl, 140mM NaCl
eosin: 75% (v/v) Harris' eosin Y aqueous solution, 25% (v/v) Harris' eosin Y alcoholic solution, 0.5% (v/v) glacial acetic acid
FACS PBS: 0.1% BSA, 0.1% sodium azide in PBS
formalin: 10% (v/v) formaldehyde in phosphate buffered saline
formamide sample buffer: 2.3X MOPS, 50% deionized formamide, 11% formaldehyde
gel fix: 0.6% (w/v) trichloroacetic acid, 20% (v/v) methanol, 7% (v/v) ethanol
Giemsa stain: 5% (v/v) Giemsa's stain improved R66 solution in phosphate buffer pH 6.8
guanidine hydrochloride: 6M GnHCl, 10mM DTT, 25mM EDTA
haematoxylin: 4.6% (v/v) glacial acetic acid in Harris hematoxylin solution
Hoechst stain: 5μg/ml Hoechst 33258 stain in phosphate buffer pH 6.8
10X kinase buffer: 50mM Tris-HCl pH 7.5, 100mM $MgCl_2$, 100mM DTT, 10mM ATP, 0.5μg/ml BSA
L buffer: 0.1M EDTA, 0.01M Tris-HCl pH 8.0, 0.02M NaCl

Table 2.1.2: Human and mouse cell lines

Cell Line	Description	Reference
46BR.1G1	SV40 transformed human fibroblast cell line with aberrant DNA ligase I activity	Webster *et al.* 1992
HM-1	HPRT-deficient embryonic stem cell line isolated from blastocysts of HPRT-deficient strain 129/Ola mice	Selfridge *et al.* 1992, Magin *et al.* 1992b
HM-1/*Lig1 PGK-HPRT* #53	HM-1-derived ES clone in which exons 23-27 of one allele of the *Lig1* gene has been replaced by a *PGK-HPRT* minigene	Jessop 1995, Bentley *et al.* 1996
MRC5V1	SV40 transformed wild-type human fibroblast cell line	Huschtscha and Holliday 1983
$PF20^T$	Immortalised wild-type mouse embryonic fibroblast cell line	Unpublished data, D. W. Melton

Table 2.1.3: Plasmids

Name	Description	Reference
pBluescript II KS+	General purpose cloning vector	Thummel 1988
pMC/CreN	Cre recombinase expression vector	Gu *et al.* 1993
pML	Mouse DNA ligase I cDNA cloned into pBluescript II KS+	Jessop 1995
p91	Mouse actin cDNA cloned into pBR322	Minty *et al.* 1981

10X T4 DNA ligase buffer: 200mM Tris-HCl pH 7.6, 75mM $MgCl_2$, 50mM DTT, 250μg/ml BSA, 5mM ATP

2X lysis buffer: 100mM Tris-HCl pH 8.3, 250mM NaCl, 2% NP-40, 4mM EDTA, 200mM Na_2HPO_4, 2mM PMSF

10X MOPS: 200mM 3-(*N*-morpholino) propane-sulphonic acid (MOPS), 50mM sodium acetate, 10mM EDTA, pH 7.0

neutralising buffer: 3M NaCl, 0.5M Tris-HCl, pH 7.0

Northern blot stripping solution: 0.01% (w/v) SDS, 0.01X SSC

oligo labelling buffer: 50μl solution A, 125μl solution B, 75μl solution C

Solution A: 1.25M Tris-HCl pH 8.0, 0.125M $MgCl_2$, 25mM ß-mercaptoethanol, 0.5mM each of dGTP, dATP and dTTP.

Solution B: 2M HEPES buffer adjusted to pH 6.6 with NaOH.

Solution C: Random hexadeoxyribonucleotides OD_{260nm}=90 units/ml in TE

PI: 50mM Tris-HCl pH 8.0, 10mM EDTA, 100μg/ml RNase A

PII: 20mM NaOH, 1% (w/v) SDS

PIII: 2.55M potassium acetate pH 4.8

PCA: 25 parts redistilled phenol, 24 parts chloroform, 1 part isoamyl alcohol

10X PCR: 50mM KCl, 1.5mM $MgCl_2$, 10mM Tris-HCl pH 8.3, 0.01% (w/v) gelatin, 0.45% (v/v) Triton X-100, 0.45% (v/v) Tween 20

phosphate buffer pH 6.8: 49mM Na_2HPO_4, 51mM NaH_2PO_4

phosphate buffer pH 8.0: 95mM Na_2HPO_4, 5mM NaH_2PO_4

phosphate buffered saline (PBS): 8mM K_2HPO_4, 1.5mM KH_2PO_4, 150mM NaCl

QBT: 750mM NaCl pH 7.0, 50mM MOPS, 15% (v/v) ethanol, 0.15% (v/v) Triton X-100

QC: 1M NaCl pH 7.0, 50mM MOPS, 15% (v/v) ethanol

QF: 1.25M NaCl pH 8.2, 50mM MOPS, 15% (v/v) ethanol

new methylene blue: 1% (w/v) new methylene blue in citrate saline

3X reducing SDS sample buffer: 187mM Tris-HCl pH 6.8, 6% (w/v) SDS, 30% (v/v) glycerol, 0.03% phenol red, 125mM DTT

RNA solution: 0.1M Tris-HCl pH 8.9, 0.1M NaCl, 1mM EDTA, 1% (w/v) SDS

10X RT buffer: 500mM Tris-HCl pH 8.3, 60mM $MgCl_2$, 400mM KCl, 10mM DTT

5X sample buffer: 20% (v/v) glycerol, 100mM EDTA, 0.1% (w/v) bromophenol blue

Scott's tap water: 1% (w/v) potassium hydrogen carbonate, 10% (w/v) magnesium sulphate

SDS-PAGE running buffer: 250mM Tris-HCl, 1.92M glycine, 2% (w/v) SDS

sequencing sample buffer: 0.3% (w/v) bromophenol blue, 0.3% (w/v) xylene cyanol FF, 10mM EDTA pH 7.5, 97.5% (v/v) deionised formamide

20X SSC: 3M NaCl, 0.3M tri-sodium citrate, pH 7.0

Southern blot stripping solution: 1% (w/v) SDS, 0. 1X SSC
10X TBE: 0.9M Tris-HCl, 0.9M boric acid, 20mM EDTA
tail buffer: 300M sodium acetate, 10mM Tris-HCl, 1mM EDTA, 1% (w/v) SDS
TBST: 50mM Tris-HCl pH 7.5, 150mM NaCl, 0.05% (v/v) Tween 20
TE buffer: 10mM Tris-HCl, 1mM EDTA, pH 8.0
transformation buffer: 50mM $CaCl_2$,10mM Tris-HCl

2.2 Methods

2.2.1 Animal procedures

2.2.1.1 Animal husbandry

Animals were maintained according to established animal care guidelines (Poole 1989), and all procedures subject to Home Office licence under the Animals (Scientific Procedures) Act 1986. Briefly, mice were kept under a 12 hour light/12 hour dark cycle, with standard mouse chow and tap water *ad libitum*. At 4 weeks of age young animals were weaned into single sex cages, 1cm of distal tail tissue removed for genotyping, and ears marked to permit identification.

For timed matings, individually housed male stud mice were presented with 2-4 sexually mature females overnight. The morning after, females were checked for the presence of a copulation plug in the vagina, the day on which a vaginal plug was found being designated as E0.5.

2.2.1.2 Collection of tissue samples

Peripheral blood was obtained from anaesthetised mice by bleeding from the eye orbit through heparinized capillary tubes. Haematocrit values were determined by plugging one end of a full capillary tube and centrifuging at 1,500rpm for 5 minutes. The length of the pellet of packed red cells was measured and expressed as a percentage of the total length of the sample within the capillary. For reticulocyte staining and genomic DNA preparation, two drops of whole blood were collected into a microfuge tube. To produce blood smears, one drop was placed onto a clean microscope slide, thinly spread across the surface using a coverslip, and allowed to air dry.

To isolate the white cell fraction for FACScan analysis, 15 drops were collected into 10mls of PBS. Cells were pelleted by centrifugation at 1,300rpm for 6 minutes, the supernatant discarded, and 9mls of water added to lyse the red cells. After 10 seconds 1ml of 10X PBS was added to arrest lysis, and the remaining cells pelleted by centrifugation at 1,300rpm for 6 minutes. The supernatant was aspirated, the white

cell pellet resuspended in 10mls of PBS, then spun down again, and finally resuspended in 200μl of FACS PBS.

Other tissues were dissected from mice killed by cervical dislocation. Cells were collected from the peritoneal cavity by repeated flushing with a large volume of PBS and centrifugation of the eluate at 1,300rpm for 6 minutes. Bone marrow was collected from the femur. Surrounding muscle was stripped away, and the bone severed at either end as close to the knee and pelvis as possible. The cells in the marrow cavity were flushed out with 1ml of PBS using a syringe and 23G needle, and then pelleted by brief centrifugation at 1,000rpm. Tissues for RNA preparation were snap frozen, and stored at -70°C, whilst samples to be analysed by FACScan were teased apart in a small volume of FACS PBS, and a single cell suspension created by successive passage through a 25G needle.

2.2.1.3 Dissection of conceptuses

Plugged females were killed on the required day of embryonic development, and the uteri were dissected out of the body cavity. The muscle layer of the uterus was removed, along with the Reichert's membrane and placenta, to leave the embryo and visceral yolk sac.

The yolk sac was removed for genomic DNA preparation, and embryonic peripheral blood obtained from the umbilical vessels. 10μl blood samples were diluted in 1ml of PBS to prevent clotting. Embryos were treated in a number of possible ways:

a) photographed under a Zeiss photomicroscope using Kodak film (100 ASA).

b) fixed whole in formalin for histologic analysis.

c) snap frozen, and stored at -70°C for RNA and protein preparation.

d) foetal liver isolated for injection into lethally irradiated recipients, or colony forming assays

e) foetal liver and thymic rudiment isolated for FACS analysis.

f) incubated in trypsin and disaggregated for embryonic fibroblast isolation.

Alternatively, if the yolk sacs were to be analysed, the embryo carcass was used for genomic DNA preparation.

2.2.1.4 Rescue of lethally-irradiated animals with foetal liver cells

Embryos were isolated from plugged females at E13.5 of development. The embryonic yolk sacs were removed for rapid sexing of the embryos by male-specific PCR. Foetal livers were dissected out into 400μl of sterile PBS supplemented with foetal calf serum to a final concentration of 2% (v/v), and stored on ice until the sex of each embryo had been determined. Single cell suspensions were then derived from

foetal livers by snipping with scissors and passage through a 25G needle three times. A 10μl sample of each cell suspension was taken for a cell count, and the remainder injected into lethally-irradiated recipients.

Strain 129/Ola or 129/CBA F1 mice were used as recipients of foetal liver cell suspensions. Female mice aged 6 to 8 weeks were exposed to 10.5 Gy of γ radiation. Foetal liver single cell suspensions were subsequently injected into the recipients' tail vein, and the animals ear marked for identification. In most cases suspensions from single livers were injected into single recipients. Alternatively suspensions from multiple embryos were injected into a single recipient, in which case suspensions were mixed, the cells pelleted by brief centrifugation at 1,000 rpm, and resuspended in 400μl of sterile PBS prior to injection. As a negative control for survival following irradiation, a small number of irradiated animals received no injection of cells. All animals received neomycin in their drinking water for the first week following irradiation, and their weight was monitored daily.

Blood samples were taken from those animals which survived five weeks. Further samples were taken two weeks later. Two months after irradiation, surviving animals were killed, and haematopoietic tissues removed for analysis.

2.2.2 Bacterial culture

2.2.2.1 Growth of *E. coli*

Cells were grown at 37°C either in suspension in LB or TB with shaking, or on the surface of Luria agar plates. Cells grown on Luria agar plates supplemented with ampicillin were incubated at 30°C to prevent growth of satellite colonies.

2.2.2.2 Storage of *E. coli*

For short term storage (4 to 6 weeks) bacteria were streaked onto Luria agar plates, and stored at 4°C. For longer term storage, 900μl of a fresh culture, grown overnight in LB, was mixed with 100μl DMSO, and frozen at -70°C. Cells were recovered by scraping the surface of the frozen culture with a sterile inoculating loop and streaking onto a Luria agar plate.

2.2.2.3 Transformation of *E. coli*

Bacteria were transformed with plasmid DNA by the method of Mandel and Higa (1970) with modifications by Dagert and Ehrlich (1974). 0.5ml of an overnight culture of *E. coli* was added to 50ml of LB supplemented with $MgCl_2$ to a final concentration of 20mM. Cells were grown with vigorous shaking at 37°C until the OD_{600nm} of the suspension equalled 0.2, and the cells were in log phase growth. The cells were chilled on ice for 5 minutes, and then spun down at 1,500rpm for 15

minutes at 4°C. The pellet was gently resuspended in 20ml ice-cold transformation buffer, incubated for 30 minutes on ice, subsequently re-pelleted at 1,500rpm for 15 minutes at 4°C, and finally resuspended in 2ml ice-cold transformation buffer. The suspension of competent cells was then incubated on ice for a minimum of 2 hours prior to use.

Approximately 10ng of plasmid DNA in a volume of 5-10μl was added to 100μl of competent cells, which were then incubated on ice for 30 minutes. Following heat shock at 37°C for 5 minutes, 400μl of warm LB was added, and the mixture incubated at 37°C for 1 hour to allow expression of antibiotic resistance. Thereafter, the cells were spread onto Luria agar plates with ampicillin selection, and incubated overnight at 30°C.

Where appropriate, 30μl of 2% (w/v in dimethylformamide) X-Gal and 20μl of 100mM IPTG were also spread over the surface of the plate to enable blue-white colour selection of colonies containing recombinant plasmids.

2.2.3 Mammalian cell culture

2.2.3.1 Growth of embryonic stem cells

HM-1, a feeder independent line of murine ES cells (Magin *et al.* 1992b) were maintained in supplemented GMEM at 37°C in a 5% CO_2 atmosphere in a humidified tissue culture incubator (Forma Scientific). Cells were prompted to adhere to flasks and dishes by coating the plastic surfaces with 0.1% (v/v) gelatin.

To passage cells, the medium was aspirated, the bottom of the flask or dish rinsed with 1-2mls of trypsin, then a further 1-2mls of trypsin added and incubated at 37°C until the cells had become detached. Thereupon 5 volumes of medium were added, the cell suspension transferred to a sterile 50ml tube, and cells pelleted by 5 minutes centrifugation at 1,300rpm. The supernatant was discarded, and the cells resuspended in an appropriate volume of medium, and an aliquot transferred into a tissue culture flask.

For long term storage, cells were kept in liquid nitrogen. Cultures to be frozen were trypsinised, spun down, the supernatant removed, and cells resuspended in ice cold freezing medium. 1ml aliquots were placed at -20°C for 2 hours, then transferred to -70°C overnight, before removal into liquid nitrogen. Frozen aliquots were thawed rapidly at 37°C and diluted in 9mls of medium, before being pelleted by centrifugation at 1,300rpm for 5 minutes. The supernatant was discarded, the cells resuspended in an appropriate volume of medium, and transferred into a flask.

2.2.3.2 Electroporation

In order to achieve transient expression of cloned DNA sequences in ES cells, a circular expression vector was introduced by electroporation Cells to be electroporated were trypsinized, pelleted by centrifugation, and $2.0X10^7$ cells were resuspended in 0.8ml of electroporation buffer. The suspension was mixed with 200μg of plasmid DNA in 100μl sterile distilled water, and transferred into an electroporation cuvette (0.4cm electrode, Bio-Rad). The cuvette was slotted into a Gene Pulser (Bio-Rad), and the cells pulsed at 900V, 3μFd capacitance. After 10 minutes incubation at room temperature, the cells were added to 20mls supplemented GMEM and transferred to a flask. The pool of cells were allowed to grow for 48 hours, whereupon the process was repeated.

2.2.3.3 Growth of mammalian fibroblasts

Human and mouse fibroblasts were maintained, passaged and stored in the manner outlined in section 2.2.3.1, except that plasticware was not gelatinised prior to use.

2.2.3.4 Isolation of mouse embryonic fibroblast strains

Primary fibroblasts were isolated from E12.5 embryos. Following dissection, embryos were incubated in 0.5ml trypsin for 30 minutes at 37°C. 15mls of medium were added, the embryo mechanically disaggregated with a pipette, and cells plated out into 90mm dishes. In order to derive established fibroblast cell lines, primary fibroblast cultures were passaged until cells lost growth potential and underwent "crisis". Thereupon the medium was changed periodically until an immortalised cell line emerged.

2.2.3.5 Pulse labelling of DNA replication intermediates

Newly synthesised DNA was labelled by incorporation of [methyl-^{3}H]-thymidine. Fibroblasts were plated out into 30mm dishes in supplemented GMEM so that cultures were about 80% confluent 48 hours later. Dishes were rinsed twice with warm serum free GMEM, and then incubated at 37°C for 10 minutes in 2mls of warm serum free GMEM supplemented with [methyl-^{3}H]-thymidine to a final concentration of 2μCi/ml. The medium was aspirated, and the dishes rinsed with ice cold PBS to stop DNA replication. A further 1ml of ice cold PBS was added and the cells harvested by scraping, whereupon they were transferred to a microfuge tube and pelleted by brief centrifugation. The supernatant was removed and the pellet resuspend in 20μl of ice cold L buffer. 60μl of molten 1.5% (w/v) low melting point agarose was added to the suspension, mixed, and left to set on ice for 5 minutes.

Alternatively, cells were chased with unlabelled thymidine after being pulsed. Following 10 minutes incubation with [methyl-^{3}H]-thymidine, the medium was aspirated, and the dishes rinsed with warm supplemented GMEM plus 2mM thymidine. A further 2mls warm supplemented GMEM plus 2mM thymidine was added and the dishes incubated at 37°C for the required chase period (2, 5, 10, 15, or 30 minutes). Subsequently, the dishes were rinsed with ice cold PBS and processed as above.

2.2.3.6 CFU-A assay

Numbers of primitive pluripotent haematopoietic progenitors in the foetal liver were estimated by an *in vitro* colony forming assay (Lorimore *et al.* 1990). $1X10^5$ cells from foetal liver single cell suspensions were mixed with 1ml of molten top agarose equilibrated to 37°C. The mixture was transferred to 30mm dishes into which 1ml of bottom agarose had already been distributed, and allowed to set. Dishes were incubated at 37°C in a humidified atmosphere of 10% CO_2, 5% O_2 in nitrogen for 11 days, whereupon colony numbers and sizes were scored.

2.2.3.7 BFU-E assay

Numbers of erythroid burst forming units in the foetal liver were estimated from an *in vitro* methylcellulose colony forming assay (Freshney 1994). $1.5X10^5$ cells from foetal liver single cell suspensions were mixed with 2mls of methylcellulose medium in 30mm dishes, and incubated at 37°C in a humidified atmosphere of 10% CO_2, 5% O_2 in nitrogen. Colony numbers and sizes were scored after 8 days.

2.2.3.8 Lymphocyte proliferation assays

The ability of B and T lymphocytes to proliferate in response to mitogens was measured by incorporation of [methyl-^{3}H]-thymidine as an indicator of total cell numbers. Cells from haematopoietic tissue were resuspended in supplemented GMEM to a concentration of $5X10^6$ cells per ml. 100μl aliquots were distributed into round bottomed 96 well microtitre plates, and to each well a further 100μl of supplemented GMEM plus mitogen was added. The mitogens employed were lipopolysaccharide (LPS), phytohaemagglutinin (PHA), *Staphlococcus* enterotoxin B (SEB). Plates were incubated at 37°C in a 5% CO_2 atmosphere in a humidified tissue culture incubator for 72 hours, after which period 20μl of GMEM supplemented with 5μCi/ml [methyl-^{3}H]-thymidine was added, to a final concentration of 1μCi per well. Following incubation for a further 18 hours, the plates were harvested and assayed for ^{3}H incorporation using a filtermate system (Canberra Packard). Briefly, cells from each well were transferred onto GF/C membrane, which was washed and then dried

prior to detection of bound ^{3}H by scintillation. A measure of the ability of lymphocytes to proliferate in response to mitogens was derived by comparing the amount of radiolabel incorporated in mitogen-stimulated samples with the level incorporated into control samples that had not received mitogen.

2.2.4 Nucleic acid procedures

2.2.4.1 Nucleic acid isolation

2.2.4.1.1 Small-scale preparation of plasmid DNA

A modification of the method described by Ish-Horowicz and Burke (1981) was used to prepare amounts of plasmid DNA up to 10μg. Bacteria were grown overnight at 37°C in 5ml LB or TB supplemented with ampicillin to a final concentration of 50μg/ml. 1.5ml of the culture was transferred to a microfuge tube, and the cells pelleted by centrifugation for 5 minutes. The pellet was resuspended in 100μl of solution P1, and the cells lysed by addition of 200μl of solution P2 with 5 minutes incubation on ice. 150μl of cold solution P3 was added, with gentle mixing, and the tube incubated on ice for a further 15 minutes. Precipitated chromosomal DNA, SDS and protein were sedimented by centrifugation for 10 minutes in a microcentrifuge, and the supernatant, containing the plasmid DNA, removed to a fresh tube. The supernatant was centrifuged again for a further 10 minutes to remove residual debris, and then plasmid DNA was precipitated by addition of 1ml cold ethanol. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, vacuum dried, resuspended in 50μl sterile distilled water and stored at -20°C.

2.2.4.1.2 Large-scale preparation of plasmid DNA

For preparation of larger amounts of plasmid DNA, up to 500μg, a column purification method was employed. Bacteria were grown for 24 hours at 37°C in 50ml of TB supplemented with ampicillin to a final concentration of 100μg/ml. Cells were pelleted by spinning at 4,000rpm for 20 minutes at 4°C. The pellet was resuspended in 10ml of solution P1 supplemented with RNase to a final concentration of 0.5μg/ml, and the cells lysed by addition of 10ml of solution P2 with 5 minutes incubation at room temperature. 10ml of cold solution P3 was added, with gentle mixing, and the tube incubated on ice for a further 20 minutes. Precipitated chromosomal DNA, SDS and protein were sedimented by centrifugation at 16,000rpm for 10 minutes at 4°C. The supernatant was removed promptly, filtered and applied to a Qiagen-tip 500 column (Qiagen), previously equilibrated with 10ml of QBT buffer. After washing twice with 30ml QC buffer, the plasmid DNA was eluted from the column with 15ml QF buffer. 12ml of isopropanol was added to precipitate the DNA, which was then

spun down at 10,000rpm for 10 minutes at 4°C. The DNA pellet was resuspended in 400μl of sterile distilled water, prior to transfer into an microfuge tube, whereupon the DNA was reprecipitated with 40μl of P3 and 1ml of cold absolute ethanol. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, vacuum dried, resuspended in 200μl sterile distilled water and stored at -20°C.

2.2.4.1.3 Preparation of mammalian genomic DNA

DNA was prepared from mouse tail samples, embryonic yolk sacs, whole blood, tissue samples, ES cells and fibroblasts by the same method. Cultured cells were harvested by scraping in ice cold PBS, and pelleted by centrifugation at 1,300rpm for 5 minutes. Cell pellets and tissue samples were digested overnight at 37°C in 700μl tail buffer supplemented with proteinase K to a final concentration of 280μg/ml. The supernatant was extracted twice with 700μl PCA and vigorous shaking, and subsequently once with 700μl CA to remove traces of phenol. Thereupon the DNA was precipitated for 10 minutes at room temperature by addition of 25μl 3.0M sodium acetate pH 5.0 and 700μl of isopropanol. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, vacuum dried, resuspended in 200μl sterile distilled water and stored either at 4°C in the short term or at -20°C for longer periods.

2.2.4.1.4 Preparation of RNA from mammalian cells

RNA was prepared using a modification of the methods described by Strohman *et al.* (1977) and MacDonald *et al.* (1987). Cultured cells were harvested by scraping in ice cold PBS, and pelleted by centrifugation at 1,300rpm for 5 minutes. The cells were resuspended in 5mls of guanidine hydrochloride, transferred to a glass homogenizer and macerated prior to being transferred into a centrifuge tube. Frozen embryo and tissue samples were homogenized in 10mls guanidine hydrochloride with an Ultra-Turrax T25 homogenizer (IKA-Laboratechnik). $^1/_{30}$ volume of 3M potassium acetate, pH 5.0 and $^1/_2$ volume of cold absolute ethanol were added to the homogenate, mixed, and the solution left at -20°C for at least 4 hours to precipitate the RNA. The RNA was pelleted by centrifugation at 12,000rpm for 20 minutes at 4°C, the supernatant discarded, and then the RNA re-precipitated twice more under the same conditions. The resulting pellet was suspended in 3ml of RNA solution. An equal volume of PCA was added and the sample extracted by vigorous shaking for 10 minutes, followed by centrifugation at 4,000rpm for 10 minutes. RNA was precipitated from the aqueous layer by addition of $^1/_5$ volume of 3M sodium acetate, pH 5.0 and 2 volumes of cold absolute ethanol. The RNA was left to precipitate for 4

hours at -20°C and then pelleted by spinning at 12,000rpm for 20 minutes at 4°C. The pellet was resuspended in 0.5ml sterile distilled water, transferred to an microfuge tube, and reprecipitated by addition of $^1/_{10}$ volume 3M sodium acetate, pH 5.0, 2 volumes cold absolute ethanol and incubation at -70°C for 15 minutes. The nucleic acids were pelleted by centrifugation for 10 minutes in a microcentrifuge, washed twice with 70% (v/v) ethanol, dried, resuspended in 50μl of sterile distilled water and stored at -70°C.

2.2.4.1.5 Processing of agarose plugs

Cells which had been embedded in agarose plugs were processed by an adaptation of the method outlined in Sambrook *et al.* (1989). L buffer was supplemented with sarkosyl NL30 and proteinase K to final concentrations of 2% (v/v) and 400μg/ml respectively. 1ml of this mixture was added to the tube containing the plug and allowed to soak at 4°C for 2 hours. The tube was then transferred to 37°C for 24 to 48 hours, after which time the plug was removed into 5mls of cold L buffer, and stored at 4°C.

2.2.4.1.6 Quantification of nucleic acids

2.2.4.1.6.1 Estimation of DNA concentration

DNA samples were diluted in 1ml of distilled water and the absorbency at 260 and 280nm measured in a spectrophotometer (Perkin-Elmer, Lambda 15, UV/VIS Spectrophotometer). Double stranded DNA of concentration 50μg/ml has an OD_{260nm} = 1.0. The ratio OD_{260nm}/OD_{280nm} gives an estimate of nucleic acid purity. For DNA, a value of around 1.8 indicates that the preparation is not significantly contaminated with protein or phenol..

2.2.4.1.6.2 Estimation of RNA concentration

RNA concentrations were measured as for DNA. OD_{260nm} = 1.0 represents an RNA concentration of 40μg/ml, and OD_{260nm}/OD_{280nm} = 2.0 indicates a pure preparation of RNA.

2.2.4.2 DNA manipulation

2.2.4.2.1 Fragment purification

The DNA to be purified was made up to a volume of at least 100μl with 1X TBE, 0.2M NaCl. 1μl of glycogen was added to act as a carrier, and thereby increase the amount of DNA ultimately recovered. An equal volume of redistilled phenol equilibrated in Tris-HCl pH 8.0 was added, the mixture vortexed until an emulsion

was formed, then spun down in a microfuge for 5 minutes. The aqueous phase was removed into a fresh tube, and 100μl of butanol added to remove traces of phenol. The mixture was vortexed, then spun down for 1 minute and subsequently the butanol layer discarded. $^1/_{10}$ volume 3M sodium acetate pH 5.0 and $2^1/_2$ volumes of cold absolute ethanol were added, and the DNA precipitated at -70°C for 20 minutes. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, vacuum dried, and resuspended in 10μl sterile distilled water.

2.2.4.2.2 Restriction of DNA with endonucleases

DNA was digested with 5 units of the desired endonuclease per μg of DNA, using buffer and temperature conditions recommended by the manufacturer. For double digests involving enzymes with different buffer requirements, reactions were either carried out under intermediate buffer conditions, or DNA was digested in two sequential steps separated by a DNA purification procedure. Reactions were halted by addition of 5X sample buffer or by heating at 65°C for 10 minutes.

2.2.4.2.3 Dephosphorylation

Bacterial alkaline phosphatase (BAP) removes 5' phosphates from DNA and RNA by hydrolysis. It was used to dephosphorylate linearized plasmid vectors, prior to insert ligation, thereby preventing vector recircularization and hence increasing the frequency of recombinant plasmids produced. Approximately 1ng of DNA, in 90μl of sterile distilled water, was mixed with 10μl of 10X BAP buffer and 66 units of BAP. The reaction mix was incubated for one hour at 65°C, following which the DNA was purified to remove all traces of the enzyme.

2.2.4.2.4 Phosphorylation of PCR products and ligation

In order to facilitate the cloning of PCR products it was necessary to phosphorylate their 5' termini by means of a kinase reaction. Reactions were performed on gel purified PCR products in 10μl volumes containing 1μl 10X kinase buffer, and 1μl of T4 DNA kinase. Following incubation at 37°C for $1^1/_2$ hours, a further 1μl of 10X kinase buffer was added, along with linearized plasmid vector DNA and 2 units of T4 DNA ligase. The total volume was made up to 20μl with sterile distilled water, and the ligation reaction incubated overnight at 15°C, after which it was used to transform competent *E. coli*.

2.2.4.3 Electrophoresis of nucleic acids

2.2.4.3.1 Electrophoresis of DNA in agarose gels

DNA fragments were separated on 0.8-2% (w/v) electrophoresis grade agarose gels containing 0.5μg/ml ethidium bromide and 1X TBE. DNA samples were mixed with $^1/_5$ volume of 5X sample buffer prior to loading. Electrophoresis was carried out horizontally at 1-5V/cm in a 1X TBE buffer system. *Hin*d III digested bacteriophage λ DNA marker ladder and a 100bp DNA ladder were used as size markers. After electrophoresis, DNA was visualised by ultra-violet illumination.

2.2.4.3.2 Recovery of DNA from agarose gels

If DNA was to be recovered from agarose gels, low melting point agarose was used, and the desired section of gel excised following electrophoresis. An equal volume of 1X TBE, 0.2M NaCl was added, and the gel melted at 65°C. 1μl of glycogen was added to act as a carrier, and increase the proportion of DNA ultimately recovered. Next, an equal volume of redistilled phenol equilibrated in Tris-HCl pH 8.0 was added, the mixture vortexed until an emulsion was formed, then spun down in a microfuge for 5 minutes. The aqueous phase was removed into a fresh tube, and 100μl of butanol added to remove traces of phenol. The mixture was vortexed, then spun down for 1 minute and subsequently the butanol layer discarded. $^1/_{10}$ volume 3M sodium acetate pH 5.0 and $2^1/_2$ volumes of cold absolute ethanol were added, and the DNA precipitated at -70°C for 20 minutes. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, vacuum dried, and resuspended in 10μl sterile distilled water.

2.2.4.3.3 Electrophoresis of DNA in denaturing alkaline agarose gels

Single-stranded DNA replication intermediates were separated on 1% (w/v) electrophoresis grade agarose gels containing 50mM NaCl, 50mM NaOH, and 1mM EDTA, according to the method outlined in Sambrook *et al.* (1989). Agarose plugs were soaked in 40μl 6X alkaline loading buffer at 37°C for 90 minutes to denature the DNA. Plugs and buffer were transferred into wells cut in the alkaline gel, which were then sealed with molten agarose prior to submersion of the gel in alkaline running buffer. Electrophoresis was carried out horizontally at 2V/cm in a 50mM NaOH, 1mM EDTA buffer system for 6 hours. *Hin*d III digested bacteriophage λ DNA marker ladder and a 100bp DNA ladder were used as size markers. After electrophoresis, the gels were neutralised by soaking in neutralisation buffer for 1 hour, and following a further 1 hour incubation in 0.5μg/ml ethidium bromide solution, DNA was visualised

by ultra-violet illumination. For subsequent detection of ^{3}H by scintillation, the gel was sliced into chunks of approximately $1cm^3$ volume.

2.2.4.3.4 Electrophoresis of DNA in polyacrylamide gels

For enhanced resolution of DNA fragments, electrophoresis was carried out in a 0.2mm thick 6% (w/v) polyacrylamide gel using a Sequi-Gen sequencing cell (Bio-Rad). Gels were made up of 50% (w/v) of urea, 6% (w/v) acrylamide/0.3% (w/v) bis-acrylamide in 1X TBE buffer, and polymerised by addition of ammonium persulphate and TEMED to final concentrations of 0.1% (w/v) and 0.04% (v/v) respectively.

DNA samples were prepared for loading by addition of $^{1}/_{3}$ volume of sequencing sample buffer, and heated to 80°C for 2 minutes. Gels were run in a 1X TBE buffer system at 50W, at a temperature of 55°C to give maximal resolution. After electrophoresis, gels were laid on filter paper, covered with cling film and dried for two hours at 80°C under vacuum. Ladders of radiolabelled bands were visualised by autoradiography.

2.2.4.3.5 Recovery of DNA from polyacrylamide gels

DNA was purified from polyacrylamide gels by a modification of the 'crush and soak' method outlined in Sambrook *et al.* (1989). The relevant band was cut from the gel and transferred to an microfuge tube, whereupon it was crushed against the side of the tube. 2 volumes of elution buffer was added, and the tube incubated at 37°C overnight with shaking. Polyacrylamide debris was removed by centrifugation for 1 minute, and then passing the supernatant through a large pipette tip packed with a Whatman GF/C filter. The DNA was precipitated by addition of 2 volumes of cold absolute ethanol and incubation on ice for 30 minutes. Following 10 minutes centrifugation, the nucleic acid pellet was resuspended in 200μl of sterile distilled water, and reprecipitated by addition of $^{1}/_{10}$ volume 3M sodium acetate pH 5.0 and $2^{1}/_{2}$ volumes of cold absolute ethanol. After 20 minutes incubation at -70°C, the DNA was pelleted by 10 minutes centrifugation, washed twice with 70% (v/v) ethanol, vacuum dried, and resuspended in 10μl sterile distilled water.

2.2.4.3.6 Electrophoresis of RNA in agarose gels

RNA samples were separated on denaturing 1% (w/v) agarose gels containing 0.5μg/ml ethidium bromide, 1X MOPS and 0.66M formaldehyde. 20μg of total RNA in 20μl of sterile distilled water was added to an equal volume of formamide sample buffer and $^{1}/_{4}$ volume of 5X sample buffer. Samples were heated for 5 minutes at 65°C and snap chilled on ice immediately prior to loading. Electrophoresis was carried out in a 1X MOPS buffer system at 2.5V/cm for 3-4 hours.

2.2.4.4 Transfer of nucleic acids from agarose gels to membranes

2.2.4.4.1 DNA transfer (Southern blotting)

This procedure was originally developed by Southern (1975) and modified by Smith and Summers (1980). For Southern analysis of genomic DNA, 10μg of DNA was cut with an appropriate restriction enzyme, and the resulting fragments were separated by agarose gel electrophoresis. Similarly, for quantification of PCR reactions, products also separated on a 1% agarose gel. The gel was soaked in denaturation buffer for 30 minutes preceding DNA transfer onto Genescreen *Plus* nylon membrane (New England Nuclear) by capillary action using denaturation buffer as the transfer medium (Reed and Mann 1985). A wick made of wet blotting paper was placed on a platform, with both ends of the blotting paper immersed in denaturing buffer in a reservoir underneath the platform. The gel was laid on top of the wick, bordered with cling film. A sheet of nylon membrane (previously equilibrated in denaturing buffer) was placed on top, overlaid by three sheets of moistened blotting paper and a stack of dry paper towels. A weighted glass plate was put on top of the paper towels. Transfer was allowed to proceed for 12-24 hours, after which time the membrane was neutralised for 30 minutes in neutralising buffer, and air dried.

2.2.4.4.2 RNA transfer (northern blotting)

Following electrophoresis the gel was soaked for two 20 minute periods in 10X SSC. Transfer was as for DNA, except that 10X SSC was used as the transfer medium. When transfer was complete, the membrane was rinsed in 2X SSC and baked for 2 hours at 80°C under a vacuum.

2.2.4.5 Nucleic acid hybridisation

2.2.4.5.1 Labelling DNA by random priming with hexadeoxyribonucleotide primers

Radioactively labelled DNA was obtained using the randomly primed DNA labelling method (Feinberg and Vogelstein 1983). This method is based on the hybridisation of a mixture of hexanucleotides to the DNA, and allows small amounts of DNA to be labelled to high specific activities. Many sequence combinations are represented in the hexanucleotide primer mixture, which leads to binding of primer to template in a statistical manner. The complementary strand is synthesised from the 3' hydroxyl termini of the hexanucleotide primers using Klenow polymerase, simultaneously incorporating radiolabelled dCTP into the newly synthesised DNA strand.

DNA fragments to be labelled were purified from agarose gels. Approximately 5ng of DNA dissolved in 32μl of sterile distilled water was denatured by boiling for 5 minutes, followed by incubation at 37°C for 10 minutes. 10μl of oligo labelling buffer, 20μg of BSA, 50μCi of α-^{32}P-dCTP and 2 units of Klenow polymerase were added, and the mixture incubated overnight at room temperature. 4.5mg of sonicated herring sperm DNA was added to the probe, which was then denatured prior to hybridisation by boiling for 5min.

2.2.4.5.2 Hybridisation

Membranes onto which DNA had been transferred were blocked by incubation in 30ml of 6X SSC, 1% (w/v) SDS, and 10% (w/v) dextran sulphate with 3.5mg of denatured sonicated herring sperm DNA for two hours at 65°C in a hybridisation oven (Hybaid Ltd.). Hybridisation was performed by addition of denatured radiolabelled probe to the prehybridisation mixture, and incubation for a further 12-24 hours at 65°C. Membranes to which RNA was bound were treated identically, except that incubations were performed at 60°C.

Following hybridisation, non-specifically bound DNA molecules were removed by washing. The membrane was first immersed in 2X SSC at room temperature for 5 minutes with agitation, and then twice in 2X SSC, 1% (w/v) SDS for 30 minutes at either 65°C or 60°C. Finally the membrane was rinsed in 0.1X SSC for 10 minutes at room temperature, and sealed in a plastic bag to prevent drying out. Radioactive DNA molecules bound to the membrane were then visualised by autoradiography or phosphorimagery. If a high background to signal ratio was observed, the membrane was rewashed in more stringent conditions of 1X SSC, 1% (w/v) SDS for 15 minutes at 65°C or 60°C.

2.2.4.5.3 Stripping probes from filters

Radiolabelled DNA probes hybridised to nucleic acids immobilised on nylon were removed by the method described in the manual supplied with Genescreen *Plus* membranes. Southern filters were boiled three times for 10 minutes in 0.1X SSC, 1% (w/v) SDS. Northern filters were washed five times for 3 minutes in hot 0.1X SSC, 0.01% (w/v) SDS. Filters could then be dried and hybridisation repeated as above.

2.2.4.6 Detection of radioactivity

2.2.4.6.1 Autoradiography

Autoradiography was performed using Cronex (Du Pont) X-ray film in a cassette with intensifying screens (Cronex Lightning Plus, Du Pont). Cassettes were stored at -70°C during exposure to slow the reversal of activated bromide crystals to their stable form and give an enhanced signal.

2.2.4.6.2 Phosphorimagery

Phosphorimagery was performed using a storage phosphor screen (Molecular Dynamics). After exposure of the screen, signals were visualised through a Molecular Dynamics phosphorimager, and analysed with ImageQuant software.

2.2.4.6.3 Scintillation of agarose gel pieces

$1cm^3$ chunks of gel were transferred into scintillation vials containing 1.5mls of 0.1M HCl, and left for 1 hour. The agarose in each vial was melted by heating in a microwave at 600W for 40 seconds, and subsequently 10mls of Ecoscint A scintillant was added with vigorous shaking to form an even emulsion. A Tri-Carb 2100TR liquid scintillation analyser (Packard Instrument Company) was used to detect ^3H disintegration by scintillation.

2.2.4.7 DNA sequencing

2.2.4.7.1 Dye terminator cycle sequencing

DNA sequencing was performed using an ABI PRISM dye terminator cycle sequencing reaction ready kit (Perkin-Elmer Corporation), on an Omnigene thermal cycler (Hybaid Ltd.) and extension products separated on an ABI PRISM 377 DNA sequencer (Perkin-Elmer Corporation).

Approximately 0.4μg of double stranded DNA template was mixed with 3.2pmoles of an appropriate oligonucleotide primer and 8μl of terminator ready reaction mix in a final volume of 20μl. The mixture was overlaid with a drop of mineral oil, and subjected to thermal cycling as follows: 96°C for 30 seconds, 50°C for 15 seconds, 60°C for 4 minutes, repeated for 25 cycles.

Extension products were purified away from unincorporated terminators by addition of 2.0μl 3M sodium acetate pH 4.8, 50μl absolute ethanol, and incubation on ice for 10 minutes followed by 15 minutes centrifugation. The resulting pellet was washed with 70% (v/v) ethanol, vacuum dried, and stored dry at 4°C prior to electrophoresis.

2.2.4.7.2 Sequence analysis

Sequence analysis was carried out on a UNIX mainframe computer using the Genetics Computer Group (GCG) package of programs, version 9 (Genetics Computer Group 1996). Alignment of nucleotide sequences was performed using the Bestfit, Lineup, and Pileup programs, while nucleotide and protein database searches were carried out using Fasta and Blast functions.

2.2.4.8 Amplification of DNA by the polymerase chain reaction

All PCR reactions were carried out in 50μl volumes in 1X PCR buffer supplemented with EDTA and dNTPs, to a final concentration of 0.4mM and 0.1mM respectively, and 2.5 units of *Taq* DNA polymerase. Approximately 100ng of DNA was used per reaction, along with 1ng of each primer. Cycle conditions varied for different primers pairs, but denaturation was generally carried out at 94°C for 1 minute, followed by 1 minute at an appropriate annealing temperature, and synthesis at 72°C for 1 to 2 minutes. 35-40 cycles were carried out for all genotyping PCR reactions. All experiments were carried out using an Omnigene thermal cycler (Hybaid Ltd.).

2.2.4.8.1 PCR genotyping

In order to determine the genotype of an individual, PCR reactions were performed on genomic DNA samples from tail or yolk sac. The sex of embryos was also determined by this method. Cycle conditions employed for different primer pairs are outlined in Table 2.2.1. 20μl of each reaction was subjected to agarose gel electrophoresis, and products visualised by ultra-violet illumination.

2.2.4.8.2 Quantitative PCR

Quantitative PCR was performed in a manner similar to genotyping PCR reactions, except that only 20 cycles were performed. Cycle conditions employed for different primer pairs are described in Table 2.2.2. *Lig1* deleted and *Lig1* targeted (B) PCR products were each amplified as part of duplex reactions, with an independent reaction acting as a control for the initial amount of DNA. The *Lig1* deleted product was amplified along with a product from the wild-type mouse *Ercc-1* gene, whilst the *Lig1* targeted (B) product was amplified with a product from the wild-type mouse *Prn-p* gene. The male-specific reaction was run in parallel with the wild-type mouse *Ercc-1* reaction, although the two reactions had to be performed separately. PCR products were separated by agarose gel electrophoresis, transferred onto nylon membrane, and hybridised to ^{32}P-labelled probes specific for the two PCR products. Levels of bound radiolabel were analysed by phosphorimagery, and the ratio of the

Table 2.2.1: Genotyping PCR conditions

Reaction	Primer A	Primer B	Denaturation		Annealing		Synthesis		Cycle Number	Product Size (kb)
			°C	Min.	°C	Min.	°C	Min.		
Lig1 wild-type	N2469	M5176	94	1	66	1	72	2	40	1.5
Lig1 targeted	262W	G1022	94	1	66	1	72	2	40	1.8
Lig1 deleted	N2469	V2014	94	1	70	1	72	1	35	0.5
male-specific	Zfy-1	Zfy-2	95	0.1	65	0.5	72	0.5	30	0.15
Lig1 wild-type (B)	V0727	V2014	94	1	65	1	72	1	35	0.3
Lig1 targeted (B)	262W	V2014	94	1	65	1	72	1	35	0.6

Table 2.2.2: Quantitative PCR conditions

Reaction	Primer A	Primer B	Denaturation		Annealing		Synthesis		Cycle Number	Product Size (kb)
			°C	Min.	°C	Min.	°C	Min.		
Lig1 deleted	N2469	V2014	94	1	70	1	72	1	20	0.5
Ercc-1 wild-type	033M	035M	94	1	70	1	72	1	20	0.6
Lig1 targeted (B)	262W	V2014	94	1	65	1	72	1	20	0.6
Prn-p wild-type	348M	B188	94	1	65	1	72	1	20	0.9
male-specific	Zfy-1	Zfy-2	95	0.1	65	0.5	72	0.5	20	0.15

Table 2.2.3: RT-PCR conditions

Reaction	Primer A	Primer B	Denaturation		Annealing		Synthesis		Cycle Number	Product Size (kb)
			°C	Min.	°C	Min.	°C	Min.		
Lig3	HLIG35'	HLIG33'	94	1	50	1	72	1	40	1.0
Lig4 external	HLIG45'	HLIG43'	94	1	50	1	72	1	40	1.0
Lig4 internal	LIG45'	LIG43'	94	1	50	1	72	1	40	0.75

two signals calculated. In order to estimate the percentage of an allele present in each sample, calculated ratios were compared against a standard curve. The standard curve was produced by spiking varying amounts of DNA positive for a particular allele into a wild-type DNA sample, and subjecting to PCR quantification in parallel to experimental samples.

2.2.4.8.3 Rapid screening using the polymerase chain reaction

In order to rapidly screen colonies of cultured mammalian cells, a PCR-based strategy that bypassed the need for DNA purification was employed (McMahon and Bradley 1990). When colonies were picked from dishes to be transferred into multiwell plates, half of each colony (circa 1,000 cells) was transferred into a 1.5ml microfuge tube. The cells were pelleted for by centrifugation for 30s, and the medium removed using a pipette. The pellet was resuspended in 40μl 1X PCR buffer to which proteinase K had been added to a final concentration of 0.1mg/ml, and incubated at 65°C for two hours. The proteinase K was inactivated by heating at 90°C for 15 minutes, and 10μl of sample used as a template for PCR.

Rapid sexing of embryos was also achieved by a modification of this protocol. The embryonic yolk sac was digested by incubation at 65°C for 15 minutes in 200μl 1X PCR buffer supplemented with proteinase K to a final concentration of 1mg/ml. The proteinase K was inactivated by heating at 90°C for 15 minutes, and 10μl used in male-specific PCR reactions.

2.2.4.8.4 RT-PCR

Reverse transcription was performed on 10-20μg of total RNA. Samples were made up to a volume of 10μl with sterile distilled water, heated to 70°C for 2 minutes to denature the RNA, then snap cooled on ice. Reactions were carried out at 43°C for 1 hour in 1X RT buffer supplemented with 0.5μl of rRNasin, dNTPs to a final concentration of 1.25mM, 0.1μg of random hexamers, 25μg of oligo dT, and 30 units of reverse transcriptase. A second round of reverse transcription was then carried out. The mixture was heated to 75°C for 1 minute, snap cooled on ice, and an additional 0.5μl of rRNasin and 30 units of reverse transcriptase added. After incubation at 43°C for a further 1 hour, the mixture was diluted with 200μl of TE, and stored at 4°C.

Specific cDNA sequences were amplified by PCR, under the buffer conditions outlined in section 2.2.4.8, using 10μl of the reverse transcription reaction product as a template. Cycle conditions employed for different primer pairs are outlined in Table 2.2.3. 20μl of each reaction was subjected to agarose gel electrophoresis, and products visualised by ultra-violet illumination. To amplify rare cDNA sequences, nested sets of primers were used in two sequential rounds of PCR. 5μl of the initial reaction was

seeded into a second PCR reaction employng a pair of internal primers which bound within the region amplified by the first reaction.

2.2.4.8.5 Low stringency RT-PCR using sequence specific primers

Reverse transcription was performed as described in section 2.2.4.8.4, and 10μl of the reverse transcription reaction product was used as a template for PCR reactions. Buffer conditions were as outlined in section 2.2.4.8, except that 0.5μCi α-^{32}P-dCTP was additionally included in each reaction. The cycle conditions employed were identical for each pair of primers: 30 seconds at 94°C, 1 minute at 42°C, 1 minute at 72°C, 35 cycles. Details of the primer combinations tried are given in Table 5.5.1.

PCR reaction products were precipitated by addition of $^{1}/_{10}$ volume 3M sodium acetate pH 5.0 and $2^{1}/_{2}$ volumes of cold absolute ethanol, and incubation at -70°C for 20 minutes. Following 10 minutes centrifugation, the nucleic acid pellet was washed twice with 70% (v/v) ethanol, vacuum dried, and resuspended in 10μl sterile distilled water. Samples were then separated by electrophoresis through polyacrylamide gels.

Once differentially expressed bands had been identified, one further polyacrylamide gel was run, and exposed to autoradiography film whilst still wet. The desired bands were excised, and DNA purified by the 'crush and soak' method. In order to obtain sufficient quantities of DNA to clone, 1μl of each purified fragment was subsequently reamplified by PCR. The same primer combination was used as employed in the initial amplification, but under higher stringency conditions: 30 seconds at 94°C, 1 minute at 50°C, 1 minute at 72°C, 35 cycles.

2.2.5 Protein procedures

2.2.5.1 Protein isolation

Whole cell protein extracts were prepared from fibroblasts and embryos by a method adapted from Prigent *et al.* 1994. Cell pellets were incubated for 30 minutes on ice in an equal volume of 2X lysis buffer. Frozen embryos were homogenized briefly in an equal volume of 2X lysis buffer and then incubated for 30 minutes on ice. The lysate was centrifuged at 8,000rpm at 4°C for 15 minutes, and the supernatant removed into a fresh microfuge tube. Glycerol was added to a final concentration of 30%, and the extract stored at -20°C.

2.2.5.2 Quantification of proteins

Protein concentrations were assayed using BCA* protein assay reagents according to manufacturers instructions. 1 volume of reagent A was mixed with 50 volumes of reagent B, and 2mls of the mixture added to 100μl of the sample to be assayed. After incubation at 37°C for 30 minutes, the absorbency of the solution at

562nm was measured by spectrophotometry. The concentration of protein was then estimated by comparison with the absorbencies of a range of standards of known concentration.

2.2.5.3 Electrophoresis of proteins

Proteins were separated by electrophoresis through 8% SDS-polyacrylamide gels according to the method of Laemmli 1970, using an SE250 Mighty Small II slab gel electrophoresis unit (Hoefer). Gels consisted of a 5% stacking gel upon an 8% separating gel. The separating gel was poured first and was composed of 375mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 7.8% (w/v) acrylamide/0.3% (w/v) bis-acrylamide polymerised by addition of ammonium persulphate and TEMED to final concentrations of 0.1% (w/v) and 0.06% (v/v) respectively. The stacking gel was composed of 125mM Tris-HCl pH 6.8, 0.1% (w/v) SDS, 4.9% (w/v) acrylamide/0.2% (w/v) bis-acrylamide polymerised by addition of ammonium persulphate and TEMED to final concentrations of 0.1% (w/v) and 0.1% (v/v) respectively.

Protein samples were mixed with $^1/_2$ volume of 3X reducing SDS sample buffer and heated to 90°C for 5 minutes prior to loading. Prestained broad range protein markers were used as molecular weight markers, and were prepared analogously. Electrophoresis was carried out vertically at 10mA in a SDS-PAGE running buffer system for 2 hours or until the dye front reached the end of the gel.

Gels were either blotted onto nylon membranes or stained. Staining was achieved by immersing the gel in a large volume of Coomassie stain for 1 hour, and then in Coomassie destain overnight. Gels were subsequently laid on filter paper, covered with cling film and dried for two hours at 80°C under vacuum.

2.2.5.4 Transfer of proteins from polyacrylamide gels to membranes (western blotting)

Proteins were transferred from gels onto Immobilon P nylon membrane by electroblotting using a Trans-Blot SD semi-dry transfer cell (Bio-Rad). The gel and membrane were equilibrated in electroblotting buffer for 15 minutes. A piece of blotting paper, soaked in transfer buffer, was placed on the bottom electrode plate of the unit. On top of this was laid the nylon membrane sheet, the gel, and a second piece of blotting paper soaked in transfer buffer. The other electrode plate was positioned on top to complete the circuit, and proteins were transferred at 15V for 1 hour. In order to confirm that complete transfer had occurred, the membrane was immersed in Ponceau S solution for 1 minute, after which the colouration was washed off by rinsing in TBST.

2.2.5.5 Detection of blotted proteins with antibodies

Membranes carrying transferred proteins were blocked overnight in 5% (w/v) skimmed milk powder in TBST at 4°C with shaking. After briefly rinsing in TBST, the membrane was incubated with primary antibody diluted in TBST for 1 hour at room temperature with shaking. The four antisera employed for western blotting were TL5, TL6, TL18, and TL25, all of which were raised in rabbits, and which were used at $^{1}/_{1000}$, $^{1}/_{500}$, $^{1}/_{1000}$, and $^{1}/_{500}$ dilutions respectively. The membrane was then washed three times for 5 minutes in TBST, following which it was incubated with an anti-rabbit IgG alkaline phosphatase-conjugated mouse IgG monoclonal antibody. The membrane was again washed three times for 5 minutes in TBST, before being immersed in alkaline phosphatase colour substrate solution. This consisted of alkaline phosphatase buffer supplemented with NBT and BCIP to final concentrations of 330μg/ml and 167μg/ml respectively. Colour was allowed to develop in the dark without agitation until the signal was of the required intensity, whereupon the reaction was stopped by rinsing with copious amounts of water.

2.2.6 Staining of cells and tissue sections

2.2.6.1 Peripheral blood samples

2.2.6.1.1 Giemsa staining of dry samples

Embryonic peripheral blood cells were distributed onto the surface of a microscope slide using a cytospinner. 200μl aliquots of diluted samples were spun at 700rpm for 5 minutes, and the resulting slides allowed to air dry.

Air-dried cytospin slides and adult blood smears were fixed by brief immersion in absolute methanol. The slides were rinsed with copious amounts of water, and again left to air dry, before being bathed in Giemsa stain for 10 minutes. After rinsing in water, the slides were left to air dry. Permanent preservation was achieved using DPX mounting medium, and a coverslip.

2.2.6.1.2 Vital staining of reticulocytes

Reticulocytes were distinguished from mature erythrocytes by staining of wet preparations with new methylene blue. 2 to 4 volumes of 1% new methylene blue were added to fresh blood samples, and incubated at 37°C for 20 minutes. One drop of the mixture was spotted onto the middle of a clean microscope slide, and a coverslip positioned over the liquid. The edges of the coverslip were sealed with a thin layer of nail varnish to prevent evaporation, and the wet preparation examined under a microscope and photographed.

2.2.6.2 Tissue histology

Embryos were fixed in formalin, embedded in paraffin and sectioned at 5μm onto microscope slides. Sections were stained with haematoxylin and eosin by a modification of Harris' method whereby a 1% acid alcohol wash is omitted. Slides were immersed in xylene for 5 minutes, then dipped into absolute ethanol for 20 seconds, 95% ethanol for 20 seconds, and 70% ethanol for 20 seconds. After rinsing in water, the slides were stained with haematoxylin for 10 minutes, and rinsed again. Immersion in Scott's tap water for 30 seconds, was followed by another rinse in water, and then staining in eosin for 5 minutes. After one final rinse, the sections were dehydrated by sequential passage through 70% ethanol for 10 seconds, 95% ethanol for 10 seconds, and ultimately absolute ethanol for 10 seconds. Following immersion in absolute alcohol, the slides were cleared by dipping in xylene, and then left to air dry. Permanent preservation was achieved using DPX mounting medium, and a coverslip. Photography was performed under a Zeiss photomicroscope using Kodak slide film (200 ASA).

2.2.6.3 Fluorescence activated cell scanning

Fluorescence activated cell scanning (FACScanning) was used to identify individual cells by detection of expressed cell surface markers. FACScanning was performed by direct staining of single cell suspensions from haematopoietic tissues. Cells numbers were electronically counted (Coulter counter ZM with Channelyzer 256, Coulter Electronics), and all samples diluted to a concentration of $1X10^7$ cells/ml with FACS PBS. 100μl aliquots (approximately $1X10^6$ cells) were distributed into round bottomed 96-well microtitre plates, and 20μl of an optimal dilution of the desired fluorochrome-linked antibody added to each. After incubation at 4°C for 30 minutes, the volume of liquid in each well was made up to 200μl with FACS PBS. The cells were pelleted by centrifugation at 800rpm for 30 seconds, the supernatant thrown off, and the cells resuspended in 200μl of FACS PBS. This washing procedure was repeated once more to remove all unbound antibody, before the cells were finally resuspended in 200μl of FACS PBS, and transferred to 3ml FACS tubes.

For each sample, 25,000 to 50,000 cells were acquired using a Becton Dickinson FACScan, and the resulting data subsequently analysed using Lysis II software. FACScan analysis is based on the use of laser light (488nm wavelength) to excite fluorochrome molecules, which subsequently emit light of a longer wavelength. Two different fluorochromes were employed, each with different peak emission wavelengths: fluorescein isothiocyanate (FITC) (550nm, "green"), and phycoerythrin (PE) (630nm, "orange"). By linking the fluorochromes to antibodies raised against cell surface molecules, it is possible to attach the fluorochrome to the surface of cells. Cell

suspensions pass into the FACScanner under pressure, such that a stream of single cells pass through a beam of laser light. Cells to which fluorochrome-conjugated antibody is bound will fluoresce strongly, and the emitted light is detected by the FACScan. Using a combination of different fluorochrome-conjugated antibodies specific for different cell surface marker molecules it was possible to identify which markers were expressed on the surface of an individual cell, and hence to determinate the phenotypic identity of that cell.

Cells passing through the beam of laser light will also scatter that beam. Light which strikes the surface of the cell and deviates very little ("forward scatter") is used to measure cell size. Light which passes through the cell will be refracted and emerge at a different angle ("side scatter"), and is used as an indicator of internal cell complexity. The FACScan identifies and records the different parameters of each cell that passes through the laser beam.

Mature erythrocytes and dead cells were distinguished from live nucleated cells on the basis of their small size, and were excluded from data acquisition. Unstained cells can fluoresce even though no antibody is bound (autofluorescence). However, because the intensity of autofluorescence is weaker than emissions from bound fluorochromes it was possible to differentiate between the two signals. To this end, unstained samples were always included as a measure of the level of autofluorescence. It was hence possible to exclude autofluorescing cells from final analyses by only including data points above a determined fluorescence threshold.

Optimal dilutions of the fluorochrome-linked antibodies were deduced by staining cells suspensions from an appropriate control tissue. The tissue was processed as described, incubated with serial dilutions of the antibody, washed, and subjected to FACScan analysis. From the results, the optimal dilution was determined that gave maximal sensitivity of detection with minimal background fluorescence.

2.2.6.4 Sister chromatid exchange rate assay

The number of exchanges between sister chromatids occurring during mitosis was assessed by differential staining of sister chromatids, using a modified fluorescence plus Giemsa technique after Perry and Wolff 1974, with modifications by Goto *et al.* 1978. Conditions were optimised for the study of mouse embryonic fibroblasts.

2.2.6.4.1 BrdU labelling of newly synthesised DNA

Fibroblasts were plated into 60mm dishes in supplemented GMEM plus BrdU to a final concentration of 10μM, and grown at 37°C for 36-96 hours. The precise length of incubation time depended on the cell strain or line used, but had to be

sufficient for two rounds of DNA replication to occur. Dishes were kept in the dark to prevent light-induced chromosome breakage.

2.2.6.4.2 Metaphase chromosome preparation

Next, colcemid was added to a final concentration of 0.135μM, and the dishes incubated at 37°C for an additional 4 hours. Cells were detached from the dishes using trypsin, pelleted by centrifugation at 1,000rpm for 5 minutes, and the supernatant poured off. The cells were gently resuspended in 5mls of hypotonic KCl solution and incubated at 37°C for 15 minutes. After being pelleted by spinning at 1,000 rpm for 5 minutes, the cells were gently resuspended in 5mls of Carnoy fixative. Following incubation at room temperature for 20 minutes, the cells were again pelleted, resuspend in 5mls of Carnoy fixative, and incubated at room temperature for 20 minutes as before. Finally, the suspension was centrifuged again, the cells resuspended in 0.5mls of Carnoy fixative, and then dropped onto a clean microscope slide. The slides were left to air dry before staining.

2.2.6.4.3 Differential staining of sister chromatids

Slides were immersed in Hoechst stain for 15 minutes at room temperature, rinsed in phosphate buffer pH 8.0, then mounted in phosphate buffer pH 8.0. Nail varnish was used to temporarily fix a coverslip onto the slide and prevent evaporation of the buffer. Spreads were exposed to a "blacklight" (18W Blacklight blue special fluorescent lamp, Philips. Peak emission: 365nm) at a distance of 5cm for 30 minutes at 55°C. The coverslip was removed, and the slide incubated in 2X SSC for 2 hours at 65°C. Subsequently, the slide was rinsed in phosphate buffer pH 6.8, and stained for 30 minutes in Giemsa stain. After rinsing in water, and air drying, slides were mounted with DPX mounting agent under a coverslip. Ultimately, metaphase spreads were photographed, and the number of sister chromatid exchanges counted. Photography was performed with a Zeiss photomicroscope using Kodak Technical Pan film (25 ASA) and a green filter to enhance contrast.

CHAPTER THREE:

DNA LIGASE I IS REQUIRED FOR FOETAL LIVER ERYTHROPOIESIS

One human patient has been identified with reduced DNA ligase I enzyme activity correlating to missense mutations within the protein coding sequence. Gene inactivation was the first step of a 'double replacement' strategy to introduce analogous mutations into the mouse DNA ligase I gene. Utilising positive-negative selection in the HPRT-deficient ES cell line HM-1, exons 23 to 27 of *Lig1* were replaced with an *HPRT* minigene. Targeted ES cells were injected into host Balb/c blastocysts. The resulting germ line chimaeras were outbred with Balb/c mice to produce mice heterozygous for the targeted allele (Jessop 1995, Bentley *et al.* 1996). Heterozygotes were phenotypically normal. This suggested that a single active allele of DNA ligase I is sufficient for all cellular requirements, and was consistent with the recessive nature of DNA ligase I-deficiency in humans. However, when interbred, no pups homozygous for the targeted allele were found. It was thus concluded that the absence of DNA ligase I is developmentally lethal. This chapter describes the phenotype of embryos homozygous for the targeted *Lig1* allele, and work to characterise the developmental defect arising from the absence of DNA ligase I.

This work was carried out in association with Jim Selfridge, Kay Samuel, Helen Taylor, Nicholas Hole and John Ansell and published in *Nature Genetics* (**13:** 489-491). A reprint of the paper is included at the rear of this thesis. Kay Samuel carried out all FACScan cell acquisition procedures.

3.1 PCR genotyping of embryos

In order to identify which embryos carried the targeted *Lig1* allele, genotyping PCR reactions were carried out on genomic DNA prepared from embryonic yolk sacs. Two reactions were carried out on each sample: one specific for the targeted *Lig1* allele, and one specific for the wild-type *Lig1* allele.

The PCR reaction used to screen for the presence of the targeted allele was the same reaction that had been initially used to screen ES cell colonies following gene targeting, and which had subsequently been used to follow the allele during breeding. One primer (262W) was specific for the 3' end of the *HPRT* minigene, and the other (G1022) specific for the 3' flanking region of the *Lig1* gene. This second primer was specific for endogenous sequence outwith the targeted region, and consequently, only the correctly targeted allele could give rise to a 1.8kb PCR product (Figure 3.1.1).

A second PCR reaction was used to screen for the presence of the wild-type allele. One primer (N2469) was specific for exon 22 of the *Lig1* gene, and the other (M5176) specific for exon 23. Because exon 23 is replaced by the *HPRT* minigene during targeting, the targeted allele will not give a product with this primer pair. Hence, the 1.5kb product from this reaction is diagnostic of the wild-type allele (Figure 3.1.1).

Figure 3.1.1: Schematic representation of the *Lig1* gene targeting strategy

Schematic representation of 13.9kb at the 3' end of the wild-type mouse DNA ligase I locus, targeting vector and gene structure following homologous recombination. Recombination occurs between the wild-type *Lig1* allele and the two regions of *Lig1* homology contained within the targeting vector (MLTV). The result is that a region of the endogenous DNA ligase I gene, containing the last five exons, is replaced with an *HPRT* minigene. In correctly targeted cells, no other part of the vector should be inserted into the genome, and therefore the thymidine kinase gene is lost along with other vector sequences. The targeted gene lacks the final five exons, but still retains 3' flanking sequences

Drawn to scale: 1cm = 2kb. Solid boxes indicate exons (numbered according to the human convention); hatched box, *HPRT* minigene (PGK-*HPRT*); stippled box, Herpes simplex virus thymidine kinase (PGK-*tk*) cassette; thin lines, *Lig1* introns or flanking regions; thick line, plasmid vector DNA. Arrows indicate the positions at which PCR primer pairs used for genotyping bind. *Bam*HI restriction sites are marked by B, *Eco*RI sites by E, and *Hin*dIII sites by H.

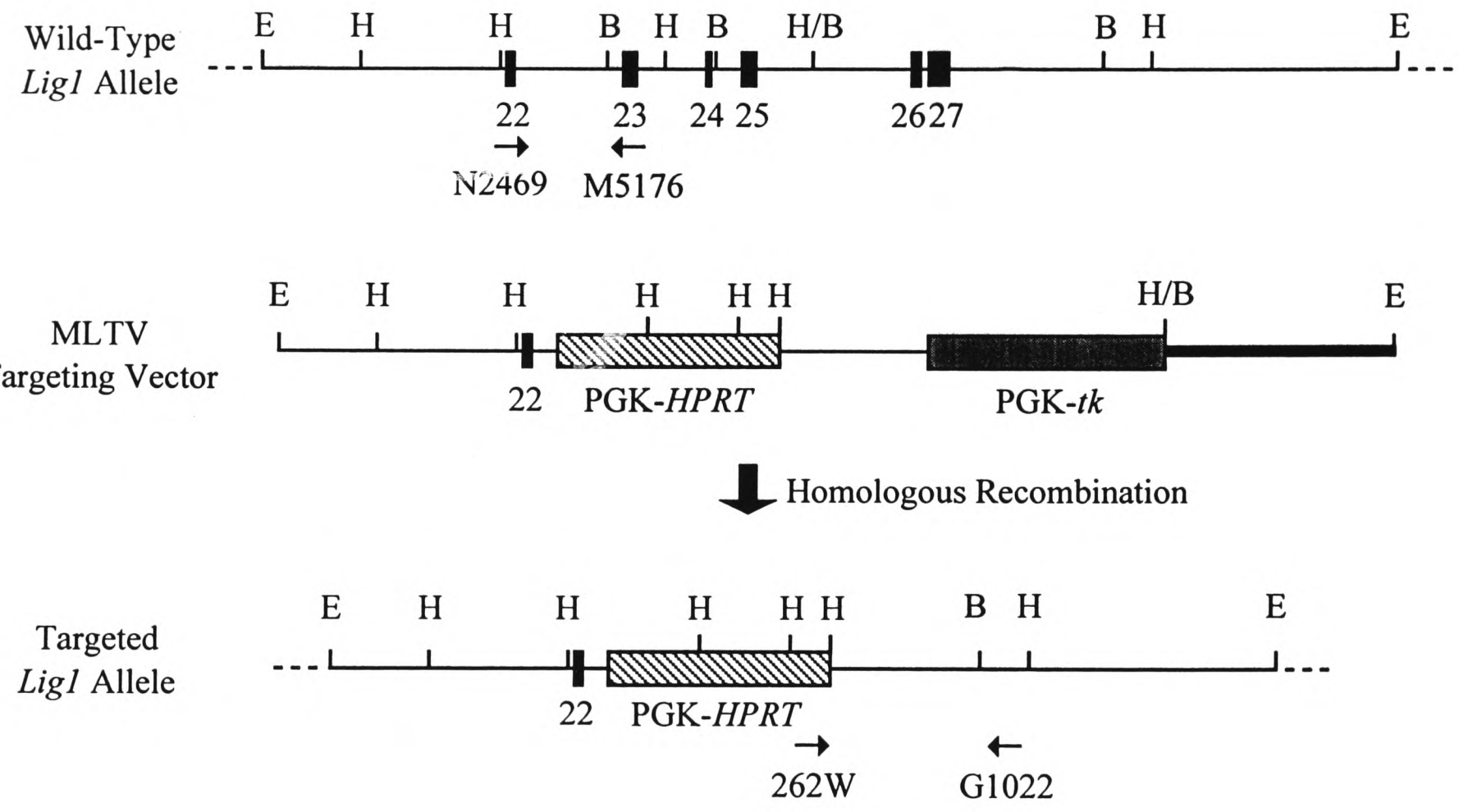
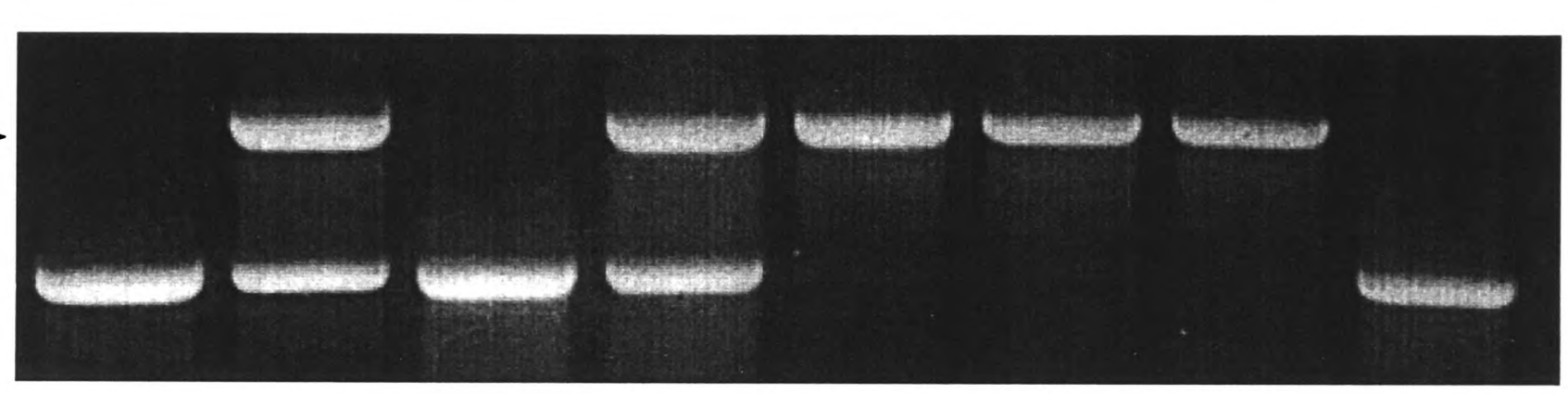

Wild-Type
Lig1 Allele
E H H B H B H/B B H E
22 23 24 25 2627
N2469 M5176
MLTV
Targeting Vector
E H H H H H H/B E
22 PGK-HPRT PGK-tk
Homologous Recombination
Targeted
Lig1 Allele
E H H H H H B H E
22 PGK-HPRT
262W G1022

Figure 3.1.2: PCR genotyping reaction products

The products of PCR genotyping reactions for the wild-type and targeted *Lig1* alleles. Genomic DNA prepared from ES cells, embryo yolk sacs and isolated fibroblast lines was used as a template for PCR reactions carried out using primer pairs N2469+M5176, and 262W+G1022, and the products separated by agarose gel electrophoresis. The 1.8kb product indicated by an arrow is specific for the targeted allele, whilst the 1.5kb product indicated by an asterisk is specific for the wild-type allele. Samples used were as follows: HM-1, wild-type ES cell line; #53, HM-1/*Lig1*/*PGK-HPRT* #53 ES cell clone; +/+, +/-, and -/- representative midgestation embryos; $PFL10^T$, $PFL13^T$, and $PF20^T$, embryonic fibroblast lines (see Chapter 5). Deduced *Lig1* genotypes are also indicated.

ES Cell Clones | Mid-gestation Embryos | Embryonic Fibroblast Lines

HM-1 | #53 | +/+ | +/- | -/- | PFL10T | PFL13T | PF20T



1.8kb product from targeted *Lig1* allele

1.5kb product from wild-type *Lig1* allele *

Lig1 genotype +/+ +/- +/+ +/- -/- -/- -/- +/+

By performing PCR assays for both the wild-type and targeted *Lig1* alleles, it was possible to determine the genotype of each embryo. Wild-type mice (+/+), homozygous for the wild-type allele, gave a signal only from the wild-type *Lig1* reaction. Those embryos homozygous for the targeted allele (-/-) would only produce a positive result from the targeted *Lig1* reaction, whilst heterozygotes, carrying both alleles, were positive for both reactions (+/-). Figure 3.1.2 shows the results of PCR genotyping analysis performed on samples of the three different genotypes.

3.2 DNA ligase I is not essential for cell viability

In order to define the stage of embryonic development at which the absence of DNA ligase I resulted in lethality, embryos ranging from day 10.5 to day 16.5 postcoitum (E10.5-E16.5) were collected from matings between heterozygous mice, and genotyped. Contrary to expectation, embryos homozygous for the targeted *Lig1* allele were observed most of the way through development (Table 3.3.1).

Live embryos of the three genotypes were obtained in the expected ratios at E10.5 [χ^2_2 = 0.74 (NS)], indicating that there was no significant loss of any one genotype before day 10 of development. Furthermore, there was no significant loss of any genotype between E11.5 and E15.5. Even at day 15 of development numbers of live embryos of the three genotypes did not depart significantly from the expected ratios [χ^2_2 = 3.8 (NS)]. However, whilst the majority of -/- embryos survived up to E15.5, all nine -/- embryos isolated at E16.5 were dead, as determined by the absence of a detectable foetal heartbeat.

Heterozygous embryos were indistinguishable from mice homozygous for the wild-type allele, and throughout the investigations reported in this thesis no differences were observed between +/- and +/+ samples. Hence, heterozygotes were regarded as 'wild-type', and, unless otherwise stated, the category 'wild-type' includes data from heterozygous samples.

3.3 Phenotype of -/- embryos

At E10.5 -/- embryos were visually indistinguishable from their wild-type and heterozygous siblings. Examination of embryonic carcasses revealed no detectable abnormalities or differences from wild-type, and no significant abnormalities were observed in the yolk sac of -/- embryos (data not shown).

Table 3.3.1: Summary of embryo genotypes from heterozygote timed matings

Embryonic Age (Days)	Total Conceptuses	Genotypes		
		+/+	+/-	-/-
10.5	77	20	35	22
11.5	34	11	14	9
12.5	18	5	8	5
13.5	124	36	59	29
14.5	44	13	20	11
15.5	78	27	34	17
16.5	19	6	13	0

Embryos were isolated from matings between heterozygous mice and genomic DNA prepared from the embryonic yolk sacs for PCR genotyping. *Lig1* genotypes were determined as indicated in section 3.1. To determine if the ratios of the three genotypes deviated significantly from the expected Mendelian ratios, chi-squared tests were carried out. Calculated values were compared against a table of χ^2 distribution to determine the level of significance.

However, at E11.5 all -/- embryos were paler in colour and slightly smaller than normal (Figure 3.3.1a). As development proceeded, these differences became exacerbated. By day 15 of development -/- embryos displayed a strikingly abnormal phenotype (Figure 3.3.1c and e), having a pallid appearance, and being smaller than their wild-type siblings.

Although small, -/- embryos appeared grossly normal (Figure 3.3.1e). Histological examination confirmed that, with the exception of the developing liver, organogenesis was not perceptibly affected (data not shown). The embryonic liver, appeared disproportionately reduced in size when compared to controls.

The pallor observed indicated severe anaemia, and this was further supported by the observation of reduced numbers of circulating red blood cells (erythrocytes) visible in the yolk sac blood vessels of -/- embryos. Cell counts performed on samples of blood taken from embryos between E11.5 and E15.5 revealed a dramatic reduction in density of cells in the peripheral circulation of -/- embryos (Table 3.3.2). By the sixteenth day, concentrations of cells in samples from -/- embryos was down to 4% of those in age-matched controls. Assuming the total volume of plasma to be similar in both genotypes, there were therefore significantly fewer total erythrocytes in the peripheral circulation of -/- embryos at E15.5.

Table 3.3.2: Cell densities in embryonic peripheral blood

Genotype	E11.5 $X10^{-8}$	E12.5 $X10^{-8}$	E13.5 $X10^{-8}$	E14.5 $X10^{-8}$	E15.5 $X10^{-8}$
wild-type	1.65 ±0.13	1.41 ±0.35	1.61 ±0.46	1.87 ±0.19	2.59 ±0.77
-/-	1.18 ±0.77	0.94±.0.13	0.25 ±0.15	0.15 ±0.05	0.1 ±0.03
-/- % of wt	71.4%	67.0%	15.2%	7.7%	3.9%

Blood samples were taken from the peripheral circulation of wild-type and -/- embryos between E11.5 and E15.5 and diluted into PBS. Cell numbers were counted using a haemocytometer, and the concentration of cells per millilitre of blood calculated. All values are expressed as cell concentrations per ml of blood, and are the mean values derived from at least three independent samples (± standard error). The density of -/- cells was also expressed as a percentage of the age-matched control.

3.3.1 Foetal liver erythropoiesis is severely disrupted

In the mouse embryo, blood cell production (haematopoiesis) begins in the yolk sac at day 7 of development, but is limited to the production of primitive, nucleated erythrocytes. Shortly afterwards haematopoietic stem cells (HSCs) from the AGM region colonise the foetal liver (Medvinsky *et al.* 1996), and haematopoiesis is initiated, with a concurrent decline in yolk sac erythropoiesis. The result is that at E10.5 the focus of erythropoiesis in the developing embryo shifts from the yolk sac to the foetal liver. From E11.5 onwards, the liver is the major site of erythropoiesis, and remains so until birth, during which period erythroid cells can comprise up to 70% of the total liver mass (Metcalf and Moore 1971).

Livers from mid-gestation -/- embryos were disproportionately reduced in size, and contained consistently fewer total cells (Table 3.3.3). Differences in total cell numbers between wild-type and -/- livers increased over the course of development, so that by E15.5 -/- livers contained less than a third of the wild-type value. At this stage sections through the developing liver also appeared radically different. Normally, cells of the erythroid lineage cluster into 'blood forming islands', which could be identified in cross-sections through wild-type livers (Figure 3.3.2a). E15.5 -/- livers contained far fewer identifiable erythroid cells, with no discernible blood forming islands, giving an overall appearance of acellularity (Figure 3.3.2b). Nevertheless, liver hepatocytes appeared normal, as did megakaryoctes.

Figure 3.3.1: External morphology of mid-gestation embryos

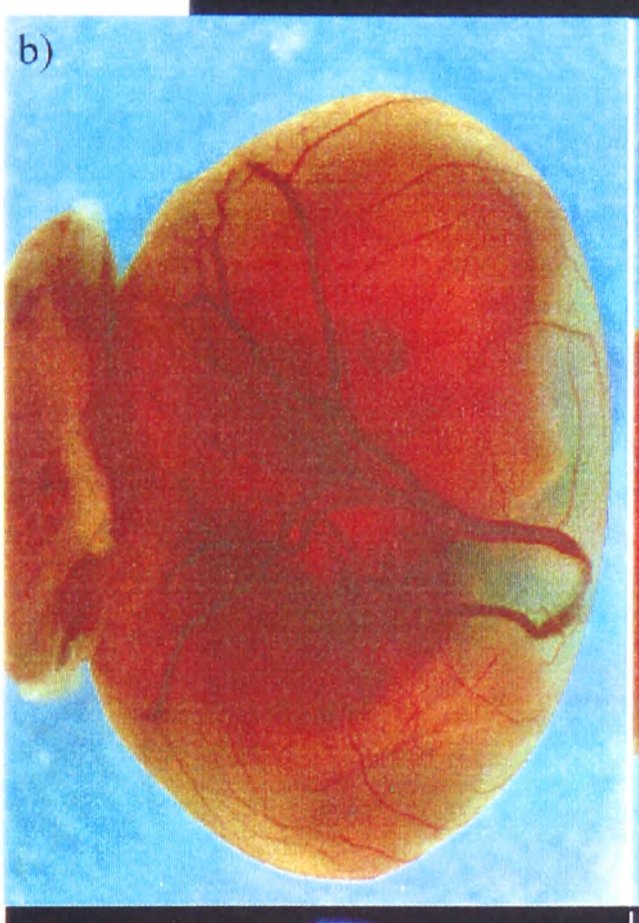
Comparison of the phenotype of -/- embryos with age-matched wild-type embryos. Embryos were isolated from matings between heterozygous mice and photographed prior to PCR genotyping.

a) Morphology of wild-type (left) and -/- (right) siblings at E11.5. The -/- embryo was paler and slightly smaller than its sibling.

b, c) Morphology of wild-type (b) and -/- (c) conceptuses at E15.5 before dissection of embryos from the yolk sac, and placenta. -/- conceptuses exhibited a profound pallor, and reduced numbers of red blood cells were visible in the yolk sac blood vessels.

d, e) Morphology of wild-type (d) and -/- (e) embryos at E15.5 after dissection from the yolk sac. The -/- embryos were pale, smaller than their wild-type siblings, and the liver was disproportionately reduced in size, but were otherwise developmentally normal.


a)


b)


c)


d)


e)

Figure 3.3.2: Foetal liver histology

Histology of livers from wild-type and -/- E15.5 embryos. Embryos were fixed in formalin, sectioned, and then stained with haematoxylin and eosin. The wild-type liver contained numerous islands of erythropoietic cells (bracketed), and megakaryocytes (asterisked). The -/- liver lacked extensive cellularization and contained no readily identifiable erythropoietic islands. There was no evidence of increased apoptosis, hepatocytes appeared normal, and megakaryocytes (asterisked) were present in normal numbers (D. Harrison, data not shown).

wild-type -/-

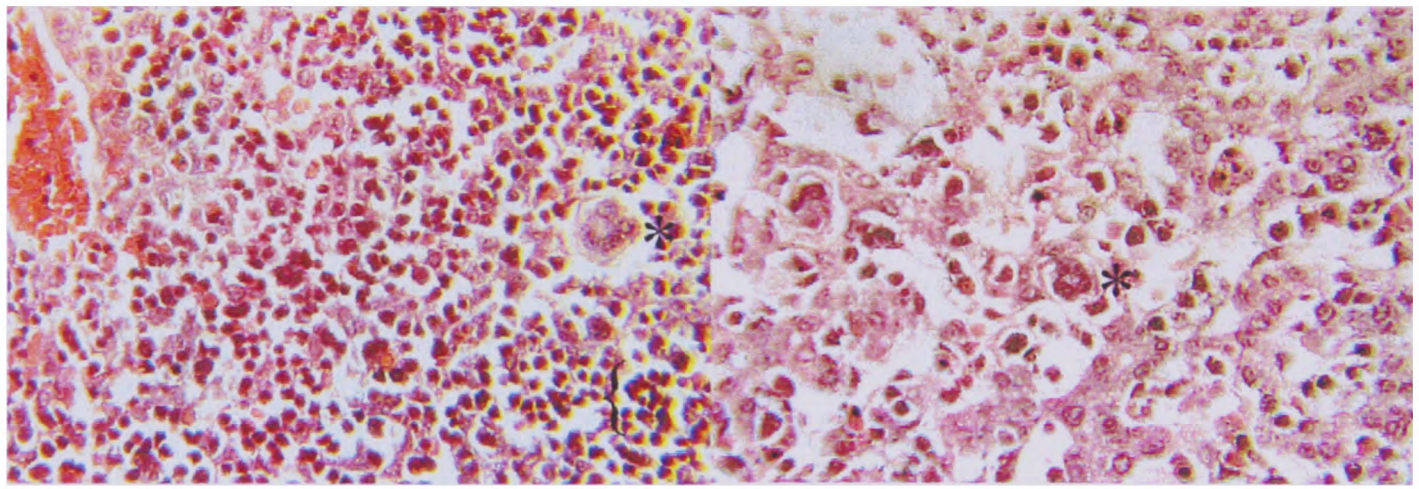


Figure 3.3.3: Incidence of nucleated erythrocytes in the embryonic circulation

The percentage of nucleated erythrocytes in wild-type and -/- embryonic blood. Cytospin preparations were made from blood samples from the peripheral circulation of wild-type and -/- embryos between E11.5 and E16.5, and stained with Giemsa. The numbers of nucleated cells were scored, and this expressed as the percentage of total cells. The percentage was then plotted against developmental age for both -/- and wild-type samples. For each point, a minimum of 500 cells from at least two separate embryos were scored. Error bars indicate the standard deviation at each sample point.

■ , wild-type; ▲ , -/-.

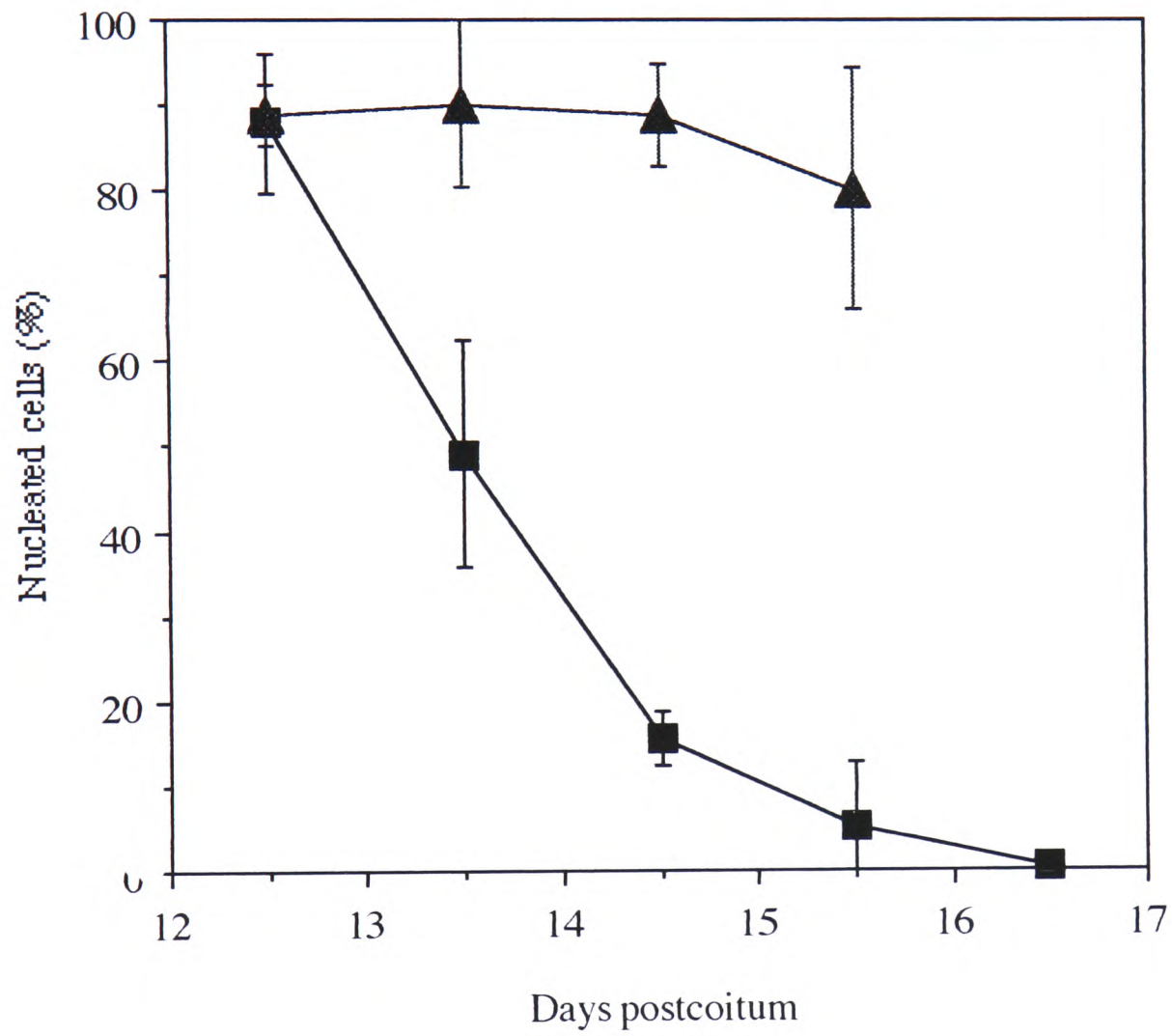
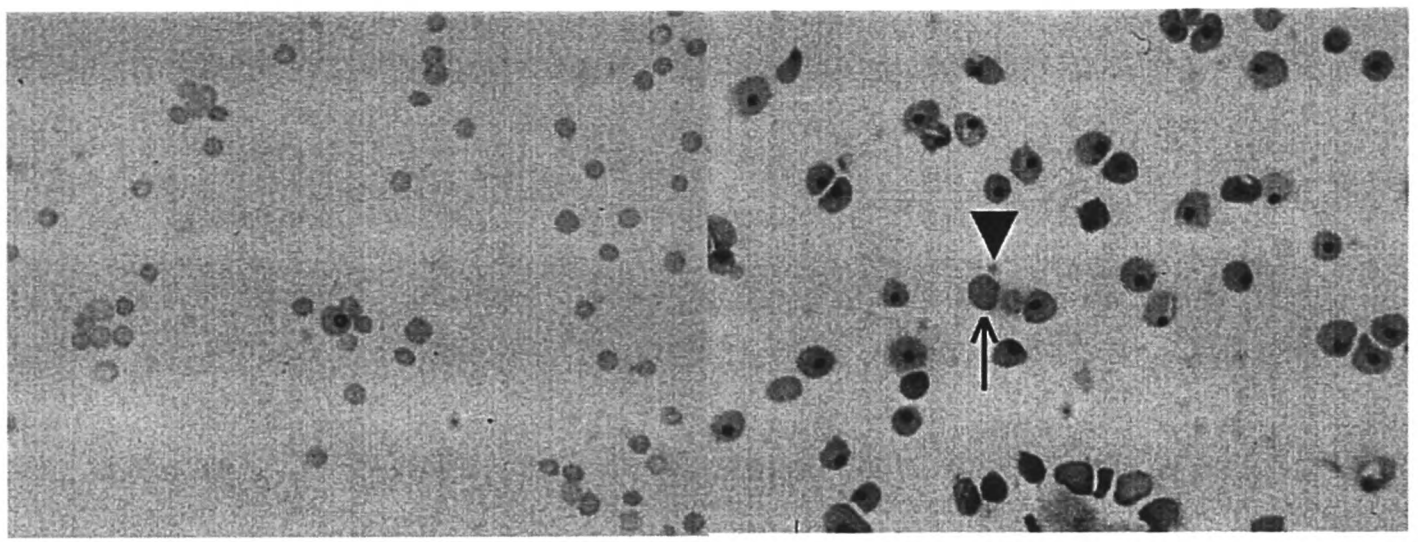

100
80
60
40
20
0
Nucleated cells (%)
12
13
14
15
16
17
Days postcoitum

Figure 3.3.4: Cytospin preparations of embryonic peripheral blood

Cytospin preparations of embryonic blood samples were made from wild-type and -/- E15.5 embryos, and stained with Giemsa. Blood from wild-type embryos consists predominantly of mature, enucleated liver-derived erythrocytes, although a few residual nucleated yolk sac-derived erythrocytes were present. In contrast, samples from -/- embryos displayed an elevated ratio of nucleated to enucleated erythrocytes. Those enucleated cells present display poikilocytosis and anisocytosis, with microcytes (arrowhead) and macrocytes (arrow) evident.

wild-type -/-



Haematopoiesis in the foetal liver is distinct from that occurring in the yolk sac in several ways; blood cell production is not limited to erythropoiesis, the erythrocytes do not mature in a synchronous manner, and production is responsive to stimulation by erythropoietin. However, the most visible difference is the production of mature, enucleated erythrocytes as opposed to nucleated erythrocytes. As development proceeds and erythropoiesis shifts from the yolk sac to the foetal liver, there is a resulting change in the proportion of nucleated cells in the peripheral blood.

Emergence of the anaemia phenotype correlated very closely with the shift of the major site of haematopoiesis from the yolk sac to the foetal liver. Therefore, in order to investigate the nature of the anaemia further, we examined the peripheral blood from embryos derived from heterozygous matings at days 11 to 16 of development. Examination of cytospin preparations revealed a deficiency of mature enucleated erythrocytes in -/- embryos. The percentage of nucleated erythrocytes in the peripheral circulation of wild-type embryos showed a decrease characteristic of the shift in erythropoiesis from the yolk sac to the foetal liver. However, samples from -/- embryos did not show the normal decrease, and the percentage of nucleated erythrocytes remained high even until E15.5 (Figure 3.3.3). The high percentage of nucleated cells could be seen clearly in blood samples from E15.5 embryos stained with a Romanowsky stain (Figure 3.3.4). Wild-type blood consisted predominantly of mature, enucleated liver-derived erythrocytes, although a few residual nucleated yolk sac-derived erythrocytes were present. In contrast, samples from -/- embryos were predominantly composed of nucleated cells, with very few enucleated erythrocytes. Many of those enucleated cells present also displayed abnormal features. There was variability in shape (poikilocytosis) and size (anisocytosis), with microcytes (erythrocytes smaller than normal) and macrocytes (erythrocytes larger than normal) evident. Some red cells were also observed to include nuclear remnants in the form of Howell-Jolly bodies.

These results demonstrated a severe disruption to foetal liver erythropoiesis, resulting in severe anaemia. Because erythrocytes are crucial to transport of oxygen around the body, acute anaemia will result in reduced oxygen supply to respiring tissues (hypoxia). In the absence of any other visible abnormalities it is postulated that -/- embryos die as a result of hypoxia secondary to the anaemia.

3.3.2 Numbers of haematopoietic progenitors in the foetal liver are reduced

The observed disruption to foetal liver erythropoiesis in -/- embryos could have been due to a number of factors, including failure of HSCs to migrate efficiently to the liver. To test for the presence of haematopoietic precursors in the foetal liver, we

Table 3.3.3: Haematopoietic cell numbers per foetal liver

Genotype	E11.5			E13.5			E15.5		
	Total Cells $x10^{-6}$	CFU-A progenitors $x10^{-4}$	++:+ Ratio	Total Cells $x10^{-6}$	CFU-A progenitors $x10^{-4}$	++:+ Ratio	Total Cells $x10^{-6}$	CFU-A progenitors $x10^{-4}$	++:+ Ratio
wild-type	1.5 ±0.1	0.023 ±0.006	1:2.4 ±0.3	10.6 ±1.2	4.1 ±0.6	1:4.6 ±0.9	59.0 ±11.5	12.3 ±1.0	1:1.5 ±0.1
-/-	1.2 ±0.5	0.014 ±.0.009	1:8.7 ±0.5	3.3 ±0.5	1.1 ±0.4	1:12.6 ±4.1	16.0 ±3.7	2.8 ±0.6	1:14.6 ±3.1
P	>0.05	>0.05	<0.01	<0.01	<0.02	<0.01	<0.05	<0.01	<0.02

Estimates of the numbers of pluripotent haematopoietic progenitors in wild-type and -/- foetal livers derived from in vitro colony forming assays. Cell suspensions were produced from wild-type and -/- foetal livers at E11.5, E13.5 and E15.5, and the number of single cells counted. Equal numbers of single cells (cell equivalents) were plated out in conditions optimised for the proliferation and differentiation of primitive multipotent haematopoietic progenitors (CFU-A assay). Numbers of colonies were scored after 11 days, and the total number of colony forming units per foetal liver calculated. All cell and colony numbers are expressed per foetal liver, and are the mean values derived from at least three independent samples (± standard error). Qualitative scoring was based on a subjective assessment of colony size. The number of 'large' (++) colonies was scored, and expressed as a ratio to the number of 'normal' (+) colonies. Statistical comparisons of the sample means were carried out using the Student's t-test (two tailed). Calculated values were compared against a table of the distribution of t to determine the level of significance (P).

performed an *in vitro* colony-forming assay. Single-cell suspensions from -/- and wild-type livers were plated under growth conditions optimised for primitive multipotent (CFU-A) haematopoietic progenitors. Colonies were formed by suspensions from -/- livers (Table 3.3.3), indicating that haematopoietic progenitor cells were present. At E11.5, numbers of total cells and multipotent progenitors in -/- foetal livers were not significantly reduced. However, as development proceeded, both total cell numbers and numbers of CFU-A progenitors were significantly reduced in -/- embryos compared to age-matched controls. At E13.5 there were approximately 3-fold fewer total cells and 4-fold fewer CFU-A progenitors, and by E15.5 these differences had increased with 4-fold fewer total cells and 5-fold fewer CFU-A progenitors in -/- livers. Qualitative scoring at each developmental stage also revealed CFU-A colonies derived from -/- livers to be smaller on average than wild-type colonies, although cell morphology appeared normal (data not shown).

3.3.3 Numbers of erythroid progenitors in the foetal liver are reduced

To test for the presence of erythroid precursors in the foetal liver, we performed another *in vitro* colony-forming assay. Single-cell suspensions from -/- and wild-type E13.5 livers were plated under growth conditions optimised for erythroid (BFU-E) progenitors. BFU-E colonies were formed by suspensions from -/- livers (Table 3.3.4), indicating that cells committed to erythroid differentiation were present. However, significantly reduced numbers of colonies were formed from -/- embryos compared to age-matched controls, there being approximately 3-fold fewer BFU-E progenitors in -/- livers. Qualitative scoring again revealed colonies derived from -/- livers to be smaller on average than wild-type colonies, although cell morphology once more appeared normal (data not shown).

3.3.4 Non-erythroid haematopoietic lineages are not diminished

Although numbers of pluripotent progenitors were depleted in -/- livers, numbers of megakaryocytes observed in histology sections were not reduced. Hence it was possible that, despite a drop in the number of CFU-A progenitors, numbers of terminally differentiated non-erythroid haematopoietic cells were not diminished. To this end samples from E15.5 foetal livers were subjected to FACScan analysis using fluorochrome-conjugated antibodies raised against T200, B220, IgM, IgD, CD3, CD4, CD8, TCR$\alpha\beta$, Gϕ, and Ter-119 cell surface markers. Each molecule is expressed by a defined subset of haematopoietic cells, and it is thus possible to categorise any cell by the combination of cell surface markers which it expresses. Cell suspensions from wild-type and -/- foetal livers were incubated with combinations of fluorochrome-conjugated antibodies raised against the different markers. After washing to remove

unbound antibodies, samples were analysed by FACScan. However, the presence of large numbers of erythroid cells in foetal liver samples (especially wild-type samples) resulted in high levels of autofluorescence and non-specific binding of antibodies. This precluded analysis of small populations of terminally differentiated cells. Only populations positive for the either B220 or T200 markers were present in sufficient numbers to be analysed. B220 is expressed on all cells committed to the B lymphoid lineage, while T200, which is equivalent to CD45 in humans, is present on all non-erythroid haematopoietic cells. In each case, an electronic gate was used to isolate the population of positive staining cells from unstained. The number of cells within this gate was expressed as a percentage of the total sample, and was also used to estimate the total number of positive cells per foetal liver (Table 3.3.5).

Percentages of T200- and B220-positive cells in samples from -/- livers were higher than in the corresponding wild-type samples. However, when differences in the total cell count per foetal liver were taken into account, the total numbers of T200-positive non-erythroid haematopoietic cells and B220-positive B lineage cells were approximately normal. These data suggest that, in contrast to the erythroid lineage, members of non-erythroid lineages were relatively unaffected by the lack of DNA ligase I.

Table 3.3.4: Erythroid progenitor cell numbers per foetal liver at E13.5

Genotype	BFU-E progenitors $x10^{-4}$	++:+ Ratio
wild-type	1.2 ±0.1	1:6.9 ±1.1
-/-	0.4 ±0.1	1:15.6 ±1.0
P	<0.02	<0.01

Estimates of the numbers of erythroid committed progenitors in wild-type and -/- E13.5 foetal livers. Cell suspensions were produced from wild-type and -/- foetal livers at E13.5, and the number of single cells counted. Equal numbers of single cells (cell equivalents) were plated out in conditions optimised for the proliferation and differentiation of erythroid progenitors. Numbers of colonies were scored after 8 days, and the total number of burst forming units per foetal liver calculated. All numbers are expressed per foetal liver, and are the mean values derived from at least three independent samples (± standard error). As a qualitative assessment of colony size, the number of 'large' (++) colonies was scored, and expressed as a ratio to the number of 'normal' (+) colonies. Statistical comparisons of the sample means were carried out using the Student's t-test (two tailed). Calculated values were compared against a table of the distribution of t to determine the levels of significance (P).

Table 3.3.5: Non-erythroid haematopoietic cell numbers per foetal liver

	B220 +ve Staining Cells		T200 +ve Staining Cells	
Genotype	% of sample	Total Cells $x10^{-6}$	% of sample	Total Cells $x10^{-6}$
wild-type	4.7 ±2.0	1.3 ±0.6	3.7 ±1.5	1.0 ±0.4
-/-	28.7 ±8.8	1.6 ±0.5	22.4 ±10.3	1.3 ±0.6

Estimates of the numbers of non-erythroid (T200+ve) and B lineage (B220+ve) cells per E15.5 foetal liver. Single cell suspensions were produced from wild-type and -/- foetal livers and subjected to FACScan analysis. Samples were stained either with a PE-conjugated rat monoclonal antibody raised against the mouse cell surface marker T200, or with a PE-conjugated rat monoclonal antibody raised against mouse B220 cell surface marker. For each sample 25,000 to 50,000 cells were analysed. The percentages of cells stained positive in each sample were multiplied by the total number of cells per foetal liver to estimate the total numbers of cells expressing T200 or B220 in the whole foetal liver. Figures represent mean values from at least three independent samples (± standard error).

CHAPTER FOUR:

RESCUE OF LETHALLY-IRRADIATED RECIPIENTS WITH HAEMATOPOIETIC CELLS FROM -/- EMBRYOS

Embryos lacking DNA ligase I developed normally to mid-term. Thereupon disruption to foetal liver erythropoiesis resulted in acute anaemia and prenatal death. Haematopoietic progenitor cells were present, as were cells committed to the erythroid lineage, but there was a deficiency of terminally differentiated erythrocytes. To determine whether this problem arose from a specific block in erythropoiesis we employed *in vivo* repopulation analysis.

A controlled dose of gamma (γ) radiation ablates the haematopoietic system of adult mice. Without further treatment, the mice will die approximately 14 days after irradiation. However, it is possible to rescue lethally-irradiated mice by introduction of haematopoietic stem cells (HSCs). HSCs are, by definition, able to differentiate into any one of the 8 haematopoietic lineages and hence are able to repopulate the haematopoietic system of irradiated mice. Furthermore, by looking at the contribution of donor-derived cells to the haematopoietic system of rescued mice it is possible to study haematopoietic differentiation in an adult environment. Using foetal livers from *Lig1* -/- embryos as a source of HSCs, it would thus be possible to investigate the effects of the absence of DNA ligase I on mature haematopoietic cells, and determine whether the embryonic phenotype observed was due to an intrinsic defect in haematopoietic development.

This chapter describes efforts to rescue lethally-irradiated mice by injection of cell suspensions from DNA ligase I-deficient foetal livers, and characterisation of the surviving animals. Work was carried out in collaboration with Kay Samuel, Martin Waterfall and John Ansell. Kay Samuel performed all eye bleeding and γ-irradiation procedures, and, together with Martin Waterfall, assisted FACScan analysis. Elements of this work have been published in *Nature Genetics* (**13:** 489-491), a reprint of which is included at the rear of this thesis.

4.1 Foetal liver cell suspensions from single *Lig1* -/- embryos were unable to rescue lethally-irradiated mice

Cell suspensions were produced from livers of E13.5 embryos, and injected intravenously into γ-irradiated recipients. Donor embryos were isolated from timed matings between heterozygotes from either an outbred stock, where 129/Ola founder chimaeras had been crossed to strain Balb/c females and the offspring interbred, or from an inbred strain where the mutation was maintained on a 129/Ola background. In each case individual recipients received cells from a single donor liver (liver equivalents). In total, suspensions from wild-type livers were injected into 13 recipients, and seven mice received cells from -/- embryos. Five irradiated animals received no donor cells, and hence acted as negative or irradiation controls. After one month 85% of the animals which had received wild-type cells remained alive, but

none of the recipients of liver suspensions from -/- embryos survived (Table 4.1.1). Thus liver equivalent numbers of -/- cells were unable to rescue lethally-irradiated mice, indicating that disruption to foetal liver erythropoiesis is not simply the consequence of an inappropriate foetal hepatic microenvironment.

Table 4.1.1: Summary of repopulation analysis using liver equivalents

Lig1 Genotype of Donor Cells	Initial Number of Recipients	Number of Recipients Surviving 4 Weeks	Percentage of Recipients Surviving
wild-type	13	11	85%
-/-	7	0	0%
irradiation controls	5	0	0%

Numbers of γ-irradiated mice surviving after injection of foetal liver cell suspensions. Single cell suspensions were produced from individual wild-type and -/- E13.5 livers, and injected into the tail vein of mice exposed to 10.5 Gy of γ radiation. Cell suspensions from inbred donors were injected into strain 129/Ola recipients, while suspensions from outbred donors were injected into 129/CBA F1 recipients. Because numbers of inbred -/- embryos were limited, the majority of data (10/13 wild-type and 5/7 -/- recipients) is derived from outbred donors. There was no apparent difference between inbred and outbred results. Numbers of mice surviving after 4 weeks were recorded. All irradiated but uninjected control mice died within 3 weeks.

4.2 *H-2K* haplotype analysis of heterozygous mice

Although liver equivalent cell numbers from -/- embryos were unable to rescue lethally-irradiated mice, it was hypothesised that cell equivalent numbers would be able to effect rescue. Undertaking such an experiment required injection of cell suspensions from multiple -/- livers into a single recipient. However, because of the problems associated with breeding inbred 129/Ola mice it proved impossible to generate the large numbers of mice demanded. It was therefore necessary to employ outbred donors, and immunologically compatible recipients. Outbred donor heterozygote stocks had a heterogeneous heritage, being a mix of 129/Ola and Balb/c of uncertain proportions. A preliminary experiment utilising 129/Ola X Balb/c F1 mice as recipients unfortunately did not produce the expected number of survivors; only 3 out of 8 recipients of wild-type cells, and 1 out of 3 recipients of -/- cells survived four weeks (data not shown). The reason for this failure is discussed further

in Chapter 7. In order to avoid this problem and maximise survival, we determined to use either inbred 129/Ola or Balb/c mice as recipients.

Rejection of tissue transplants between mouse strains is controlled predominantly by cell surface antigens encoded by the major histocompatability complex locus (termed *H-2* in mice). The *K* gene within the *H-2* locus has a number of different possible alleles, and Balb/c strain mice possess the "*d*" allele, while 129/Ola mice carry the "*b*" allele. The outbred heterozygote stock population would therefore be comprised of mice of *bb*, *dd* and *bd* genotypes. To ascertain which mice were homozygous for one or other of the alleles, and which could hence be mated to produce only *bb* or *dd* offspring, we undertook *H-2K* haplotype analysis of the heterozygous stock mice.

Haplotyping was performed by FACScanning. *H-2K* encodes an MHC class I molecule expressed on the surface of most cells, and cells expressing this antigen will be recognised by antibodies raised against the *H-2K* gene product. Blood samples were taken from adult mice under anaesthetic, the red blood cells lysed, and the remaining white cell fraction collected. Cells were incubated with an FITC-conjugated antibody raised against the H-2Kb cell surface molecule and a PE-conjugated antibody raised against the H-2Kd product. After washing to remove unbound antibodies, samples were examined for "green" and "orange" fluorescence by FACScan analysis. Mice homozygous for the *b* allele only express that allele, so cells will only bind antibodies against the H-2Kb product, and will fluoresce green. Similarly, samples from mice homozygous for the *d* allele will only bind the orange fluorescent anti-H-2Kd antibodies. Heterozygous mice express both alleles, and hence cells stain both green and orange. In this manner it was possible to infer the *H-2K* genotype of each mouse.

In total 12 male and 37 female *Lig1* +/- mice were typed on the basis of *H2K* expression. 2 males and 6 females were positive for expression of both b and d alleles, and were therefore bd heterozygotes. The remainder of the animals (10 malesand 31 females) proved to be b haplotype, and thus homozygous for the *H-2K* b allele.

None of the animals tested were d haplotype. This asymmetric distribution of genotypes within the population was presumed to arise from a 'founder effect'. The stock population was derived from a small number of progenitor animals. 129/Ola founder chimaeras were crossed to strain Balb/c females, and offspring carrying the targeted *Lig1* allele interbred. Subsequent generations were produced by selection and successive interbreeding of *Lig1* +/- mice (although without a defined inbreeding strategy), which could readily have given rise to the distribution observed. Linkage between the *H-2* locus and *Lig1* was ruled out because the former is located on

chromosome 17, whilst the latter maps to chromosome 7 (Mouse Genome Database 1997).

4.3 Liver suspensions from multiple *Lig1* -/- embryos were able to rescue lethally-irradiated mice

The 10 males and 31 females found to be homozygous for the *H-2Kb* allele were set up in timed matings. Plugged females were killed 13 days later and embryos isolated as donors for *in vivo* repopulation analysis. All embryos were theoretically homozygous for the *H-2Kb* allele, and thus immunologically compatible with the use of strain 129/Ola mice as recipients (b haplotype). *Lig1* -/- embryos were differentiated from wild-type embryos on the basis of visible anaemia.

Foetal livers were dissected into sterile PBS supplemented with FCS and stored on ice, while embryonic yolk sacs were used for rapid sexing of embryos by PCR. The yolk sacs were digested in 1X PCR buffer supplemented with proteinase K, then the enzyme heat inactivated, and a small sample used as a template for the male-specific PCR reaction. The male-specific PCR reaction amplifies part of the coding sequence of the *Zfy-1* gene (Kunieda *et al.* 1992), which is located on the Y chromosome, and hence only samples from male embryos should give a product. PCR sexing revealed 16 out of 31 embryos to be male (data not shown).

When the sexes of embryos had been determined, stored foetal livers were disaggregated to produce single cell suspensions, and cell counts were performed. The single cell suspensions were injected intravenously into γ-irradiated female 129/Ola recipients. In total, wild-type cells were injected into 9 recipients, and 5 mice received cells from -/- embryos. In the case of the former, each recipient received cells from single donors, with preference for cells from male embryos. Cell suspensions from trios of same sex -/- livers were amalgamated for injection into individual recipients, so that approximately equal number of cells were injected for both genotypes (cell equivalents). 6 irradiated animals received no donor cells. In addition one irradiated mouse received cells from both wild-type and -/- embryos. Details of donor samples and the ultimate fate of the recipients are summarised in Table 4.3.1.

The mass of each mouse was monitored daily (Figure 4.3.1). All irradiation control mice died within two weeks, preceded by a drop in bodyweight. Two out of nine recipients of wild-type cells survived for at least 10 weeks, as did two out of five recipients of -/- cells. In addition, the one mouse which had received cells from both wild-type and -/- embryos also survived. All recipients which died did so concomitantly with the irradiation controls, indicating that the cause of death was failure of the introduced HSCs to repopulate the recipient haematopoietic system.

Table 4.3.1: Summary of donor and recipient data

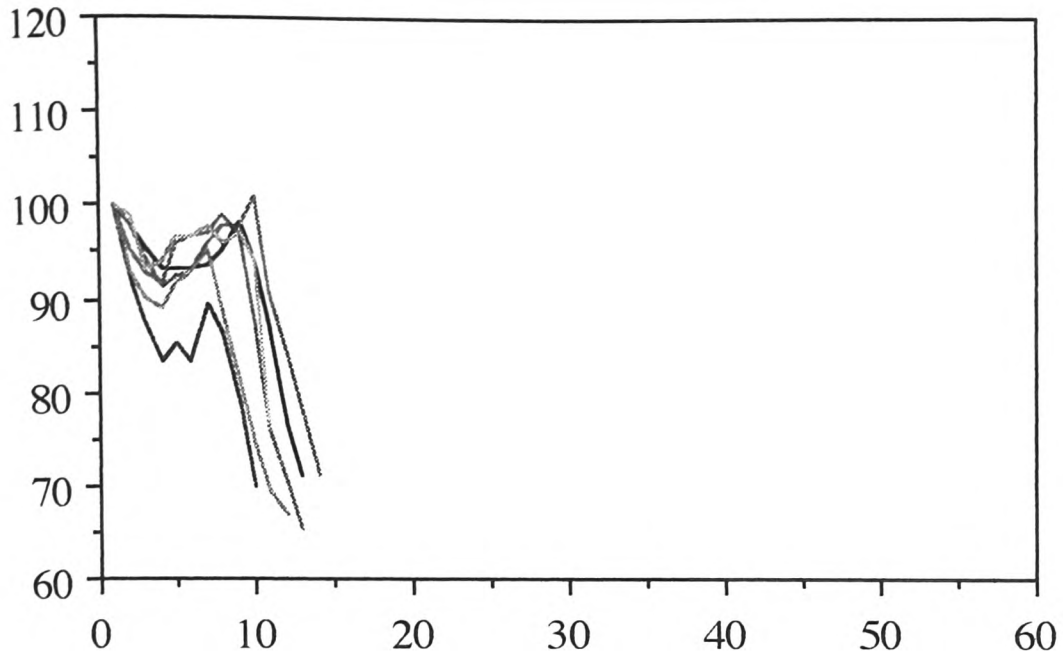
Summary of data from γ-irradiated mice injected with foetal liver cell suspensions. E13.5 embryos were isolated from heterozygote timed matings. The embryonic yolk sac was used for rapid sexing by PCR, and subsequently genomic DNA prepared for *Lig1* PCR genotyping. The foetal livers were disaggregated in sterile PBS supplemented with FCS, and numbers of single cells in suspensions counted. Approximately $1X10^8$ cells were injected into the tail vein of wild-type female 129/Ola mice exposed to 10.5 Gy of γ radiation. Samples of the same sex from -/- embryos were amalgamated to equalise numbers (cell equivalents). Male samples were employed preferentially. Mice were subsequently monitored daily. All irradiated but uninjected control mice died within 14 days.

Mouse Identifier	Donor Genotype	Donor Sex	Cells per Liver X10^{-7}	Total Cells Injected X10^{-7}	Fate of Recipient
DB2 L0R0	-/-	M	1.6	9.4	Survived
	-/-	M	4.5		(>2 months)
	-/-	M	3.3		
DB2 L1R0	-/-	F	2.9	10.4	Survived
	-/-	F	3.8		(>2 months)
	-/-	F	3.7		
DB2 L0R1	-/-	F	3.6	7.9	Died (14 days)
	-/-	F	1.8		
	-/-	F	2.5		
DB2 L1R1	-/-	M	3. 5	10.8	Died (14 days)
	-/-	M	4.4		
	-/-	M	3.0		
DB3 L0R0	+/+	F	20.2	10.1	Died (14 days)
DB3 L1R0	+/-	F	10.3	10.3	Died (10 days)
DB3 L0R1	+/+	M	14.4	14.4	Died (7 days)
DB3 L1R1	+/-	F	11.5	11.5	Died (13 days)
DB3 L2R0	+/-	M	9. 6	9.6	Died (14 days)
DB5 L0R0	-/-	F	4.4	13.1	Died (13 days)
	-/-	F	5.0		
	-/-	F	3.7		
DB5 L1R0	-/-	F	3.1	13.0	Survived
	-/-	F	4.0		(>2 months)
	+/+	M	11.9		
DB6 L0R0	+/-	M	9.9	9.9	Survived
					(>2 months)
DB6 L1R0	+/-	M	16.1	8.0	Died (14 days)
DB6 L0R1	+/+	M	15.6	7.8	Survived
					(>2 months)
DB6 L1R1	+/-	M	14.2	14.2	Died (13 days)

Figure 4.3.1: Bodyweight profile of recipients following irradiation

Bodyweights of γ-irradiated mice injected with foetal liver cell suspensions. Recipients were weighed daily as an indicator of general health, and the results expressed as a percentage of the initial mass of each animal. Percentage bodyweight was then plotted against length of time elapsed since irradiation, measured in days. Lines terminate where the subject died. Daily monitoring of the surviving animals ceased after 50 days.

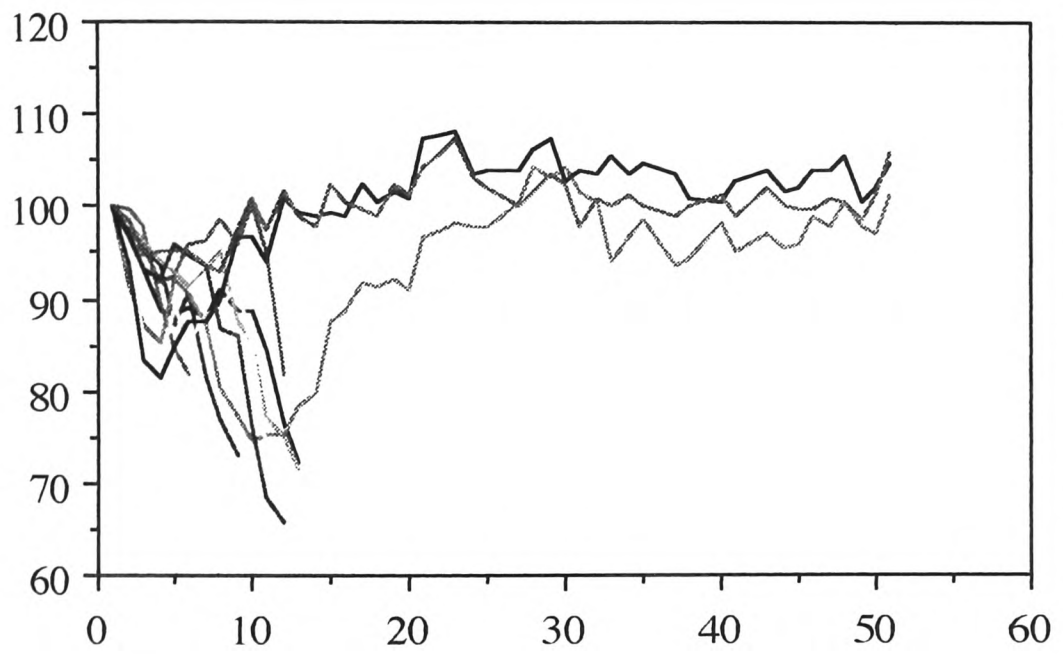
a) Irradiated but uninjected controls



Key:

- DB1 L0R0
- DB1 L1R0
- DB1 L0R1
- DB4 L0R0
- DB4 L1R0
- DB4 L0R1

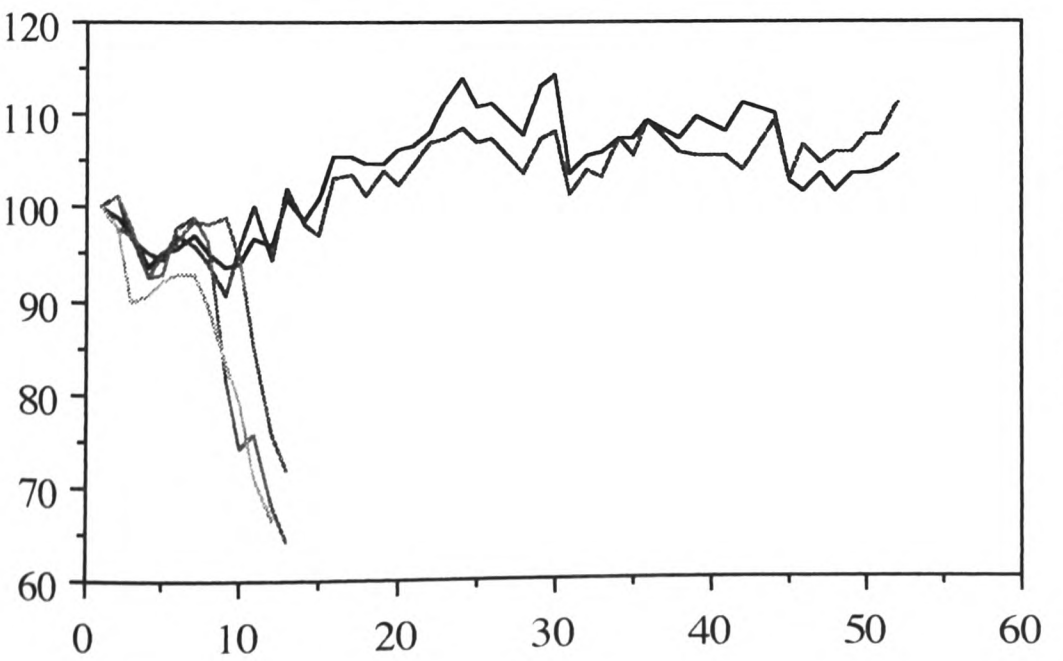
b) Recipients of wild-type cells



Key:

- DB3 L0R0
- DB3 L1R0
- DB3 L0R1
- DB3 L1R1
- DB3 L2R0
- DB6 L0R0
- DB6 L1R0
- DB6 L0R1
- DB6 L1R1
- DB5 L1R0

c) Recipients of -/- cells



Key:

- DB2 L0R0
- DB2 L1R0
- DB2 L0R1
- DB2 L1R1
- DB5 L0R0

After an initial drop, the bodyweights of all surviving animals recovered to their original levels and were stable. These animals were outwardly healthy, and showed no symptoms of graft versus host disease (data not shown). Thus *Lig1* -/- HSCs were able to rescue lethally-irradiated mice, and to effect long-term haematopoietic repopulation when present in sufficient numbers.

Peripheral blood samples were taken from all surviving recipient mice five and seven weeks after irradiation for FACScan, red blood cell staining and genomic DNA preparation. At 10 weeks two mice (DB6 L0R1 and DB2 L0R0) were killed for detailed analysis of the haematopoietic tissues. Peripheral blood was sampled, cells flushed from the peritoneal cavity, and the spleen, thymus, bone marrow, peripheral and mesenteric lymph nodes dissected. Single cell suspensions were prepared from each tissue, cell counts performed, and then samples processed for DNA preparation, FACScan analysis and mitogen stimulation. Two weeks later (12 weeks following irradiation) animals DB6 L0R0 and DB2 L1R0 were subject to the same treatment. The remaining animal DB5 L1R0, which had received a mixture of wild-type and -/- cells, was killed at a later date, and tissue samples taken for DNA preparation.

4.4 *Lig1* -/- cells were able to contribute to all haematopoietic tissues

In order to prove that cells in the haematopoietic tissues of rescued animals are derived from donor cells, it was necessary to be able to differentiate between cells of different origins. Typically this is achieved using isozyme markers, but given the mixed background of the donor embryos, it was not possible in this situation. Instead we were forced to use a PCR-based strategy to follow donor cells.

Because the recipients were all female, it was possible to use the presence of a Y chromosome as an indicator of donor-derived material. The male-specific PCR reaction which amplified part of the coding sequence of the Y chromosome-specific *Zfy-1* gene, had originally been used to sex the donor embryos. By performing the same PCR reaction on DNA prepared from haematopoietic tissues of the rescued mice, it was possible to detect the presence of donor-derived cells (data not shown).

Although preference had been given to male donor material, in order to increase the size of the experiment, female samples had also been used. To be able to study these animals as well, we employed an additional PCR marker. By definition, -/- donor samples carried two copies of the targeted *Lig1* allele per cell, whilst +/- samples carried one. Wild-type recipient mice only possessed wild-type *Lig1* alleles. Therefore, it was also possible to use the presence of the targeted *Lig1* allele as a marker for donor-derived material. A PCR assay specific for the targeted *Lig1* allele had been used to screen for targeting events in ES cells, track the passage of the allele

in breeding populations, and genotype embryos. Performing the *Lig1* targeted PCR reaction on DNA from rescued mice, it was possible to confirm the presence of donor-derived cells (data not shown).

In order to quantify the levels of donor contribution to haematopoietic tissues we performed quantitative PCR on genomic DNA samples from the rescued animals. Quantitative PCR using the male-specific PCR reaction had previously been carried out by H. Taylor and N. Hole (N. Hole, personal communication), and we followed their protocol, using a reaction specific for the wild-type mouse *Ercc-1* gene as an independent control for the initial amount of DNA. 20 cycles of PCR were carried out under conditions appropriate to the primer pairs used. Products were separated by agarose gel electrophoresis, transferred onto nylon membrane, and hybridised to α-^{32}P-labelled probes specific for the two PCR products. Levels of radiolabel bound to each PCR product were analysed by phosphorimagery, and the ratio of the two signals calculated.

A standard curve was produced by spiking varying amounts of genomic DNA from a male animal into a female DNA sample, and subjecting to PCR quantitation in parallel to the experimental samples. The ratios between the two signals were calculated and plotted against the percentage 'maleness' of the standard samples. In order to estimate the percentage of male cells present in each unknown sample, experimental signal ratios were compared against this standard curve.

Quantitative PCR using a reaction specific for the targeted *Lig1* allele proved more complicated than initially anticipated. The original reaction specific for the targeted allele was unsuitable for the purposes of quantification, and it was consequently necessary to design an alternative PCR assay. Details of the new reaction [*Lig1* targeted (B)] are given in section 4.5. Quantitative PCR was carried out as above, except that the *Lig1* targeted (B) product was amplified as part of a duplex reaction with a pair of primers specific for the wild-type mouse *Prn-p* gene acting as an independent control for the initial amount of DNA. The standard curve was produced by spiking varying amounts of genomic DNA from a -/- embryo into a wild-type DNA sample.

The standard curve was produced from a positive control with two copies of the targeted *Lig1* allele per cell, but heterozygous cells only have a single copy of the allele. To compensate for this difference, calculated percentages from the animal rescued with a *Lig1* +/- sample (DB6 L0R1) were multiplied twofold to give more accurate estimates of the contribution donor cells made to haematopoietic tissues.

By combining results of PCR quantification of the two markers it was possible to estimate the contribution donor-derived cells made to the haematopoietic system of rescued recipients (Table 4.4.1). Donor specific markers were detected in all

Table 4.4.1: Estimates of donor contribution to haematopoietic tissues of rescued animals

Recipient Tissue	Percentage Donor Contribution	
	Wild-Type	-/-
Peripheral Blood (4 weeks)	53.8 [33.6, 74.0]	35.5 [40.0, 29.5]
Peripheral Blood (6 weeks)	58.4 [87.0, 29.8]	21.8 [25.6, 17.9]
Peripheral Blood (10/12 weeks)	51.3 [17.4, 85.2]	40.9 [29.1, 52.7]
Bone Marrow	54.1 [14.0, 94.2]	86.3 [78.0, 94.7]
Mesenteric Lymph	45.7 [26.6, 64.8]	12.5 [12.0, 13.0]
Peripheral Lymph	33.8 [27.6, 40.0]	6.5 [3.0, 10.3]
Peritoneal Cells	29.3 [9.2, 49.4]	-
Spleen	63.4 [39.2, 87.6]	64.6 [56.9, 72.3]
Thymus	49.4 [30.0, 68.8]	10.2 [8.9, 11.5]

Levels of donor contribution to the haematopoietic tissues of γ-irradiated mice rescued with foetal liver cell suspensions. Genomic DNA was prepared from the haematopoietic tissues of the four irradiated mice repopulated with wild-type and -/- cells. PCR reactions for donor cell markers were carried out in a quantitative manner to approximate the proportion of cells in each sample that were derived from injected cells. Figures represent the mean values from pairs of animals rescued with cells of each genotype [individual mean value of each animal sampled], and take into account differences in the number of targeted *Lig1* alleles between samples.

haematopoietic tissues of all recipients. Significant levels were detected in all tissues from animals rescued with wild-type cells. Donor-derived cells were also detected in all haematopoietic tissues from animals rescued with -/- cells, with exception of the peritoneal cavity sample, which also failed to give a signal from the control *Prn-p* wild-type PCR reaction. Apart from the bone marrow and spleen, the contributions of -/- cells to repopulated tissues were lower than the equivalent wild-type control values.

The animal that had received both wild-type and -/- cells (DB5 L1R0) was also analysed by quantitative PCR to assess the contribution of donor cells to haematopoietic tissues. Because male +/+ and female -/- donor cells had been injected into a female recipient, it proved possible to differentiate between cells of the three different origins. Cells of -/- donor origin could not be detected in peripheral blood samples even though high levels of wild-type donor cells were present. Similarly other haematopoietic tissues displayed virtually no contribution

from -/- cells, although a low level (9.8% ±2.0) was detected in the spleen (data not shown).

4.5 An alternative PCR genotyping assay

Alternative PCR genotyping assays for both wild-type and targeted *Lig1* alleles were designed to utilise a common primer specific for the 3' flanking region of the *Lig1* gene (V2014) (C. McEwan, D. W. Melton, unpublished data). This primer could be used in conjunction either with a primer specific for exon 27 of the *Lig1* gene (V0727), or with a primer specific for the 3' end of the *HPRT* minigene (262W). The former pair (V2014+V0727) produced a 0.3kb product diagnostic of the wild-type allele, while the latter pair (V2014+262W) was used to identify the targeted *Lig1* allele (Figure 4.5.1). Because exon 27 is replaced during targeting, the targeted allele will not give a product from the wild-type reaction. Similarly only the targeted *Lig1* allele would give rise to a 0.6kb product from the targeted reaction. Although both reactions could be performed independently, the two were found to work under identical cycle parameters, and could be carried out as a duplex without appreciable bias towards either product. In other words, the two reactions could be carried out in a single tube, with both wild-type and targeted products produced in a manner seemingly independent of one another. This enabled *Lig1* PCR genotyping to be performed as a single reaction involving the three primers (V2014+V0727+262W). Furthermore both reactions were highly sensitive and robust, functioning on extremely small amounts of initial template DNA even under suboptimal conditions (data not shown). In order to differentiate the alternative genotyping assays from the original PCR reactions, the new reactions were designated *Lig1* wild-type (B), and *Lig1* targeted (B).

4.6 Animals rescued with *Lig1* -/- cells displayed anaemia coupled with macrocytosis and reticulocytosis

Although all surviving animals were outwardly healthy and bodyweights were stable, when blood samples were taken at 4 weeks following irradiation it was noted that animals rescued with -/- cells were anaemic. Measurements of haematocrit values (the volume of packed cells per unit of blood) of wild-type control animals were within the normal range (42-50%). However, haematocrits from animals rescued with -/- cells were significantly lower at each sample point (Table 4.6.1), indicating the number of red cells present in the peripheral blood of these animals to be lower than normal.

The severity of anaemia displayed by the recipients of -/- cells also proved to be variable. Whereas the haematocrits of the two control animals were stable up to

Figure 4.5.1: Schematic representation of the *Lig1* gene targeting strategy

Schematic representation of 13.9kb at the 3' end of the wild-type mouse DNA ligase I locus, targeting vector and gene structure following homologous recombination. Drawn to scale: 1cm = 2kb. Solid boxes indicate exons (numbered according to the human convention); hatched box, *HPRT* minigene (PGK-*HPRT*); stippled box, Herpes simplex virus thymidine kinase (PGK-*tk*) cassette; thin lines, *Lig1* introns or flanking regions; thick line, plasmid vector DNA. Arrows indicate the positions at which PCR primer pairs used for genotyping bind.

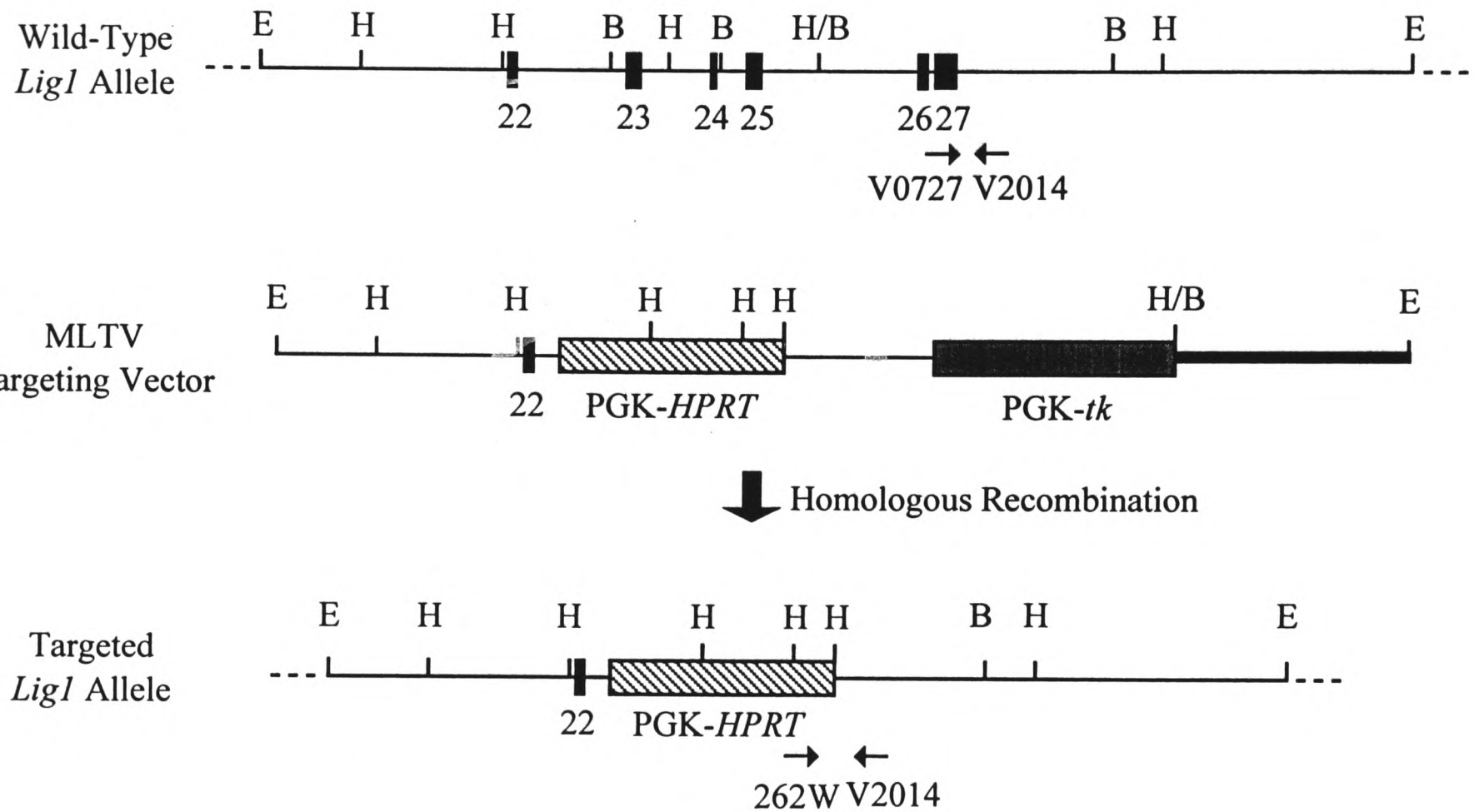

Wild-Type
Lig1 Allele
E H H B H B H/B B H E
22 23 24 25 2627
V0727 V2014
MLTV
Targeting Vector
E H H H H H H/B E
22 PGK-*HPRT* PGK-*tk*
Homologous Recombination
Targeted
Lig1 Allele
E H H H H H B H E
22 PGK-*HPRT*
262W V2014

Table 4.6.1: Haematocrit values of peripheral blood from rescued animals

	Wild-Type (%)	-/- (%)
Haematocrit (4 weeks)	48.6 [50.8, 46.4]	26.4 [26.2, 26.5]
Haematocrit (6 weeks)	48.7 [51.0, 46.3]	23.0 [16.7, 29.4]
Haematocrit (10/12 weeks)	45.8 [45.4, 46.2]	20.8 [15.9, 25.7]

The percentage of packed cells in blood samples from γ-irradiated mice rescued with foetal liver cell suspensions. Blood samples from the peripheral circulation of the four irradiated mice repopulated with wild-type and -/- cells were centrifuged, and the volume of the resulting packed cell pellet expressed as a percentage of the total volume of each sample. Figures represent the mean values from pairs of animals rescued with cells of each genotype [individual mean value of each animal sampled].

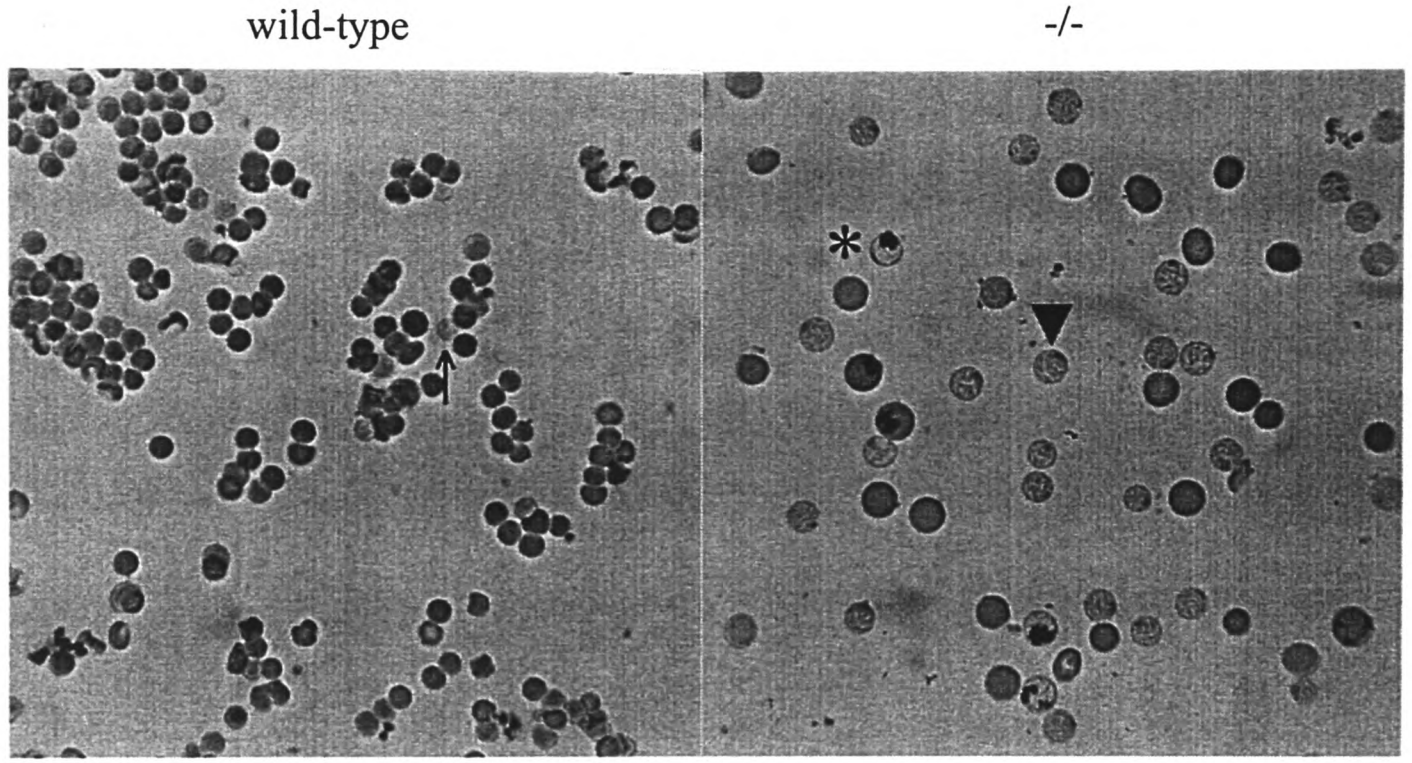
10/12 weeks, haematocrits from the animals rescued with -/- cells dropped with successive sampling. Measured values across the 8 week period ranged from 29.4% down to 15.9%, and mouse DB2 L0R0 was always more anaemic than DB2 L1R0 (data not shown). Whether this effect was an artefact of the sampling procedure, or truly reflected an underlying decline in erythrocyte production with time is unclear. Nevertheless, the results suggested that erythropoiesis was not occurring normally in mice rescued with -/- cells.

Examination of blood samples revealed that the anaemia was accompanied by macrocytosis and reticulocytosis. Blood samples were spread across the surface of a microscope slide, allowed to dry, then fixed and stained with Giemsa. The mean corpuscular volume of enucleated red cells (as indicated by cellular diameter) was greater in samples taken from mice rescued with -/- donor cells compared to wild-type controls. Although some ovalocytes (oval erythrocytes) and polychromatic macrocytes (erythroid cells larger than normal tending to stain intensely) were observed, red cell morphology was not otherwise abnormal.

When vital staining of wet blood samples was performed to distinguish reticulocytes from mature erythrocytes, a marked difference was observed. Reticulocytes are erythrocytes that have been released into the circulation before the maturation process has been completed. Unlike mature erythrocytes, the cytoplasm of reticulocytes still contains ribosomal RNA, which will stain with new methylene blue to give a stippled or reticular pattern. Samples from irradiated animals rescued with wild-type cells contained a low level of reticulocytes, but animals rescued with -/- cells displayed elevated numbers (Figure 4.6.1). Most reticulocytes displayed limited basophilia, but some early reticulocytes, containing large amounts of

Figure 4.6.1: Blood from animals repopulated with -/- cells displayed an elevated level of reticulocytes

Staining for the presence of reticulocytes in blood samples from γ-irradiated mice rescued with foetal liver cell suspensions. Blood samples from the peripheral circulation of irradiated mice repopulated with wild-type (a) and -/- (b) cells were stained with new methylene blue. Blood from wild-type controls consisted predominantly of mature, evenly stained erythrocytes. A small number of basophilic staining reticulocytes were present (arrow). In contrast, samples from animals repopulated with -/- cells displayed a high percentage of basophilic staining reticulocytes (arrowhead). Also present were early "stress" reticulocytes containing inclusions of reticulin (asterisk). The average diameter of erythrocytes in samples from mice rescued with -/- cells was increased (macrocytosis) and occasionally accompanied by ovalocytosis.


wild-type
-/-

reticulin, were observed. These cells are known to produce polychromatic staining with Romanowsky stains, thereby accounting for the polychromatic macrocytes or "stress reticulocytes" observed after Giemsa staining. These data indicated that in animals repopulated with -/- cells a high proportion of erythrocytes were being released into the circulation prematurely.

The severity of reticulocytosis was found to correlate to the severity of the anaemia observed. The level of reticulocytosis was quantified by scoring the number of reticulocytes as a percentage of the total number of cells. Samples with haematocrit values at the lowest end of the range observed (<20%) exhibited a high percentage of reticulocytes. Samples with haematocrit values at the upper end of the range (>25%) contained fewer reticulocytes, although still above the level in control samples (data not shown). Therefore, as the total number of erythrocytes in the blood of mice rescued with -/- cells decreased, the proportion of reticulocytes increased.

4.7 Animals rescued with *Lig1* -/- cells displayed splenomegaly but hypoplasia of other haematopoietic tissues

When repopulated animals were killed 10 or 12 weeks after irradiation it was observed that the spleens of recipients of -/- cells were enlarged compared to those mice which had received wild-type cells (splenomegaly). This was reflected by an increase in the average mass of the spleen (± standard deviation): Wild-type; 100.5mg (±16.5), -/-; 327.0mg (±55.0). When single cell suspensions were prepared, the number of cells in spleens from animals rescued with -/- cells was also increased compared to wild-type controls (Table 4.7.1). However, reduced numbers of cells were recovered from all other haematopoietic tissues of animals repopulated with -/- cells compared to controls (hypoplasia). The hypoplasia was most severe in the bone marrow, where the number of single cells recovered was approximately 5-fold lower. This was also accompanied by a difference in visual appearance. Bone marrow from animals rescued with wild-type cells was dark red in colour, but marrow from animals rescued with -/- cells was pale pink, suggesting reduced numbers of haemoglobinised cells.

Because the percentage contribution made by donor cells to each haematopoietic tissue had been calculated (Table 4.4.1), it was subsequently possible to estimate the number of cells of donor origin in each tissue (Table 4.7.1). With exception of the spleen, the number of donor-derived cells in the haematopoietic tissues of mice rescued with -/- cells was reduced compared to wild-type cells. The extent to which estimated cell numbers were reduced varied between tissues, but total numbers were lower even when the percentage donor contribution was higher

Table 4.7.1: Numbers of cells in suspensions from haematopoietic tissues of rescued animals

Recipient Tissue	Total Cell Numbers $X10^{-7}$		Approximate Donor Cell Numbers $X10^{-7}$	
	Wild-Type	-/-	Wild-Type	-/-
Peripheral Blood (4 weeks)	2.9 [2.2, 3.6]	1.3 [1.0, 1.6]	1.6 [0.4-3.5]	0.5 [0.2-0.8]
Peripheral Blood (6 weeks)	3.7 [3.6, 3.7]	2.1 [1.8, 2.4]	2.1 [1.1-3.2]	0.5 [0.3-0.6]
Peripheral Blood (10/12 weeks)	2.5 [2.1, 2.8]	1.9 [0.8, 3.0]	1.3 [0.4-2.4]	0.8 [0.2-1.6]
Bone Marrow	3.1 [2.4, 3.8]	0.7 [0.4, 1.1]	1.7 [0.1-4.0]	0.6 [0.3-1.0]
Mesenteric Lymph Nodes	2.8 [2.5, 3.1]	2.3 [1.6, 3.0]	1.3 [0.7-2.0]	0.3 [0.2-0.4]
Peripheral Lymph Nodes	1.8 [1.5, 2.1]	1.0 [0.6, 1.3]	0.6 [0.4-0.8]	0.1 [0.1-0.1]
Peritoneal Cells	0.7 [0.4, 1.1]	0.4 [0.2, 0.6]	0.2 [0.1-0.5]	-
Spleen	22.0 [33.1, 13.6]	41.0 [29.0, 53.0]	13.9 [2.6-32.7]	26.5 [16.5-38.3]
Thymus	7.8 [6.0, 9.6]	2.4 [1.1, 3.8]	3.9 [1.8-6.6]	0.2 [0.1-0.4]

Numbers of cells in single cell suspensions from the haematopoietic tissues of γ-irradiated mice rescued with foetal liver cells. Single cell suspensions were prepared from the haematopoietic tissues and peripheral blood of the four long-term repopulated mice, and cell counts performed. Figures of total cells represent the mean values from pairs of animals rescued with cells of each genotype [individual mean value of each animal sampled]. The calculated total number of cells in each cell suspension (assumed to represent the total number of cells in each tissue), was multiplied by the corresponding percentage donor contribution from Table 4.4.1 to produce an estimate of the number of donor-derived cells in each tissue. Figures of approximate donor cell numbers represent the mean value from pairs of animals rescued with cells of each genotype [range]

(e.g. bone marrow). As noted previously, the spleen of animals rescued with -/- cells displayed hyperplasia. In this case, the level of donor contribution was equivalent to controls and hence estimation of the number of donor-derived cells gave a figure double the control value.

The general reduction in cell numbers in the haematopoietic tissues of rescued mice indicated that -/- HSCs were deficient in some aspect of haematopoietic cell production. To determine whether this defect affected all 8 mature haematopoietic cell lineages or was limited to particular cell types, we analysed haematopoietic tissues from the four rescued mice by FACScan. Single cell suspensions from the haematopoietic tissues of mice rescued with wild-type and -/- cells were incubated with combinations of fluorochrome-conjugated antibodies raised against T200, B220, IgM, IgD, CD3, CD4, CD8, TCRαβ, Gϕ, and Ter-119 cell surface markers. Each molecule is expressed by a defined subset of haematopoietic cells, although not necessarily in a mutually exclusive manner. The correlation of markers to cell lineage is as follows: B220 is expressed on all cells committed to the B lymphoid lineage, whilst IgM and IgD are present only on those mature B lymphocytes which have undergone V(D)J recombination. CD4 and CD8 are found on the surface of thymocytes, and the TCRαβ molecule on those thymocytes which have undergone rearrangement of the T cell receptor gene. Ter-119 is a marker of erythroid lineage cells, and Gϕ expressed only on granulocytes. Antibodies raised against CD3 will bind to both T cells and a subset of B cells. T200, which is equivalent to CD45 in humans, is present on all non-erythroid haematopoietic cells. It is thus possible to categorise any cell by the combination of cell surface markers which it expresses.

After washing to remove unbound antibodies, samples were analysed by FACScanning. In each case, electronic gates were used to subdivide populations of cells on the basis of size and complexity, and to differentiate numbers of positively stained cells from autofluorescent signals (e.g. Figure 4.7.2). The number of stained cells was expressed as a percentage of the total sample, which was then used to estimate the total number of positive cells within the whole tissue (Table 4.7.2).

Whilst accurate interpretation of the data was difficult because of small sample sizes and high variability in estimated cell numbers, certain trends could be identified. The total numbers of cells positive for all non-erythroid markers were reduced in mice rescued with -/- cells (calculated mean values between 25 and 56% of wild-type controls), and numbers of T200+ve haematopoietic cells were generally reduced in all haematopoietic tissues (Table 4.7.2a). These data suggested that the haematopoietic deficiciency causing the hypoplasia affected all non-erythroid haematopoietic cell types in all tissues.

Although mice rescued with -/- cells were anaemic, the total number of cells positive for the erythroid marker Ter-119 was increased (Table 4.7.2j). Large numbers of Ter-119+ cells were present in the spleen. Given that numbers of all other haematopoietic cell types were reduced compared to wild-type, it is probable that the splenomegaly observed was a direct result of the increase in erythroid cell numbers. Increased Ter-119+ cell numbers were also noted in peripheral blood samples, peripheral lymph nodes, and thymi. However, because of the relatively small numbers of cells involved, the variability in estimated numbers, and the high potential for contamination of the tissues samples with peripheral blood, no inferences were made.

Examination of the FACScan plots revealed that staining profiles of samples from mice rescued with -/- cells were similar to, and in many cases indistinguishable from, wild-type controls (data not shown). Ratios between single and double positive cells were also apparently normal in the majority of cases. Only when total numbers of cells per tissue were calculated did a difference emerge between wild-type and -/- samples. It was therefore concluded that those haematopoietic cells present had seemingly matured normally, just in fewer numbers.

However, whilst all haematopoietic cells were depleted, there did also appear to be quantitative differences between the mature lineages. B lymphocytes (as represented by B220, IgD, and IgM+ve cells) were less well represented than T cells (CD4, CD8, and TCRαβ+ve cells), and other granulocytic white cells (Gϕ+ve) (Table 4.7.2, and data not shown). Although there were tissue specific variations, this disparity was particularly apparent in samples of peripheral blood (Figure 4.7.2). Very few circulating IgD+IgM+ cells were present in the sample from the animal repopulated with -/- cells compared to the wild-type control (Figure 4.7.2b). By contrast the populations of CD4+ and CD8+ cells are almost normal (Figure 4.7.2c). The difference is further highlighted by staining for TCRαβ and CD3 markers. The CD3+TCRαβ+ T cell population was present in samples of both genotypes, but CD3+TCRαβ- cells (i.e. B cells) were practically absent in samples from mice repopulated with -/- cells. Similarly the relative proportions of B220+ and Gϕ+ cells were altered in a number of tissues (data not shown). Whereas total B cell numbers were markedly reduced, granulocyte numbers approached, and even exceeded, wild-type values (Table 4.7.2b and i).

To draw further conclusions about the abililty DNA ligase I-deficient cells to function normally, it was necessary to determine whether the lymphocytes observed by FACScan analysis were able to respond to mitogens. To this end we carried out *in vitro* proliferation assays. Single cell suspensions from haematopoietic tissues were incubated in the presence of [methyl-^{3}H]-thymidine and mitogens to stimulate B and

T lymphocyte proliferation. By measuring amount of radiolabel incorporated it was possible to determine the relative numbers of cells with and without stimulation, and hence to gauge the ability of lymphocytes to proliferate in response to mitogens. We found that samples from all four rescued mice were able to proliferate in response to mitogens *in vitro* (data not shown). To determine the origins of these proliferating cells, quantitative PCR reactions specific for donor cell markers were carried out on stimulated samples (data not shown). Whilst significant numbers of donor cells were present in stimulated samples from mice rescued with wild-type cells (39-94%), we were unable to detect significant levels of donor markers cells in corresponding samples from mice rescued with -/- cells (0-15%). It was thus concluded that, in the latter case, proliferating lymphocytes were of recipient origin, and as such not DNA ligase I-deficient. Whether these results were due to the absence of terminally differentiated -/- cells from haematopoietic tissue samples, or because -/- cells present could not proliferate in response to mitogens, could not be determined.

Table 4.7.2: Estimates of haematopoietic cell numbers in rescued mice by FACScan analysis

Estimates of the numbers of cell surface marker-positive cells in the haematopoietic tissues of γ-irradiated mice rescued with foetal liver cells. Single cell suspensions were produced from the haematopoietic tissues and peripheral blood of the four long-term repopulated mice, and cell counts performed. Cell suspensions were subjected to FACScan analysis using fluorochrome-conjugated antibodies raised against T200 (a), B220 (b), IgM (c), IgD (d), CD3 (e), CD4 (f), CD8 (g), TCRαβ (h), Gϕ (i), and Ter-119 (j) cell surface markers. For each sample 10,000 cells were analysed. The percentages of cells stained positive in each tissue sample (inclusive of double-stained cells) were multiplied by the total number of cells in that tissue to estimate the total numbers of cells expressing that marker. Where multiple electronic gates were employed, numbers quoted are the sum of the estimates from the different gates. Figures represent the mean values from pairs of animals rescued with cells of each genotype (± standard error). The numbers of -/- cells were also expressed as percentages of the matched wild-type control values. NA, not applicable.

a) T200-positive cells

Recipient Tissue	Number of cells staining positive for T200 surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.4 ±0.2	0.5 ±0.3	120
BM	6.2 ±2.2	1.2 ±1.0	20
MLN	21 ±3.6	16 ±6.2	78
PEC	1.3 ±0.4	0.5 ±0.1	35
PLN	12 ±02	7.2 ±3.0	58
Spleen	110 ±50	55 ±12	51
Thymus	73 ±21	19 ±11	26
Total Cells	220 ±60	99 ±18	45

b) B220-positive cells

Recipient Tissue	Number of cells staining positive for B220 surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.2 ±0.1	0.02 ±0.02	15
BM	1.1 ±0.2	0.08 ±0.01	8
MLN	7.2 ±3.7	3.0 ±2.1	42
PEC	0.8 ±0.3	0.2 ±0.1	18
PLN	2.0 ±0.6	1.0 ±0.7	53
Spleen	60 ±45	13 ±10	22
Thymus	0.4 ±0.1	0.5 ±0.2	110
Total Cells	71 ±45	18 ±10	25

Table 4.7.2: Estimates of haematopoietic cell numbers in rescued mice by FACScan analysis (continued)

c) IgM-positive cells

Recipient Tissue	Number of cells staining positive for IgM surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.1 ±0.1	0.02 ±0.02	15
BM	0.6 ±0.1	0.1 ±0.1	22
MLN	7.9 ±2.4	2.2 ±1.7	29
PEC	0.7 ±0.4	0.2 ±0.1	32
PLN	2.7 ±0.3	1.1 ±0.2	41
Spleen	68 ±34	19 ±6.0	28
Thymus	0.9 ±0.4	0.2 ±0.1	27
Total Cells	81 ±34	23 ±06	29

d) IgD-positive cells

Recipient Tissue	Number of cells staining positive for IgD surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.1 ±0.1	0.03 ±0.02	24
BM	0.3 ±0.1	0.2 ±0.1	49
MLN	7.4 ±2.4	2.1 ±1.7	28
PEC	0.5 ±0.4	0.1 ±0.1	13
PLN	2.8 ±0.3	1.2 ±0.2	42
Spleen	61±34	15 ±06	25
Thymus	0.6 ±0.4	0.3 ±0.1	43
Total Cells	73 ±33	19 ±06	26

Table 4.7.2: Estimates of haematopoietic cell numbers in rescued mice by FACScan analysis (continued)

e) CD3-positive cells

Recipient Tissue	Number of cells staining positive for CD3 surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.4 ±0.2	0.3 ±0.3	92
BM	3.7 ±1.6	0.9 ±0.5	24
MLN	20 ±03	16 ±06	78
PEC	1.2 ±0.5	0.2 ±0.1	18
PLN	13 ±01	7.3 ±3.4	55
Spleen	100 ±30	49 ±05	48
Thymus	12 ±01	4.9 ±0.8	38
Total Cells	160 ±30	79 ±08	51

f) CD4-positive cells

Recipient Tissue	Number of cells staining positive for CD4 surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.2 ±0.1	0.3 ±0.2	13
BM	0.4 ±0.2	0.2 ±0.2	45
MLN	11 ±02	11 ±04	100
PEC	0.4 ±0.2	0.1 ±0.1	28
PLN	9.2 ±1.0	5.2 ±3.1	57
Spleen	42 ±14	19 ±05	46
Thymus	71 ±18	17 ±10	24
Total Cells	130 ±20	54 ±13	40

Table 4.7.2: Estimates of haematopoietic cell numbers in rescued mice by FACScan analysis (continued)

g) CD8-positive cells

Recipient Tissue	Number of cells staining positive for CD8 surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.1 ±0.1	0.1 ±0.1	86
BM	0.2 ±0.1	0.06 ±0.04	37
MLN	3.3 ±0.7	4.4 ±1.0	131
PEC	0.1 ±0.1	0.03 ±0.01	24
PLN	2.9 ±0.5	1.8 ±0.4	61
Spleen	15 ±04	5.7 ±1.3	39
Thymus	66 ±18	15 ±10	23
Total Cells	87 ±18	27 ±10	32

h) TCRαβ -positive cells

Recipient Tissue	Number of cells staining positive for TCRαβ surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.2 ±0.1	0.4 ±0.3	160
BM	1.7 ±1.3	0.5 ±0.3	29
MLN	16 ±03	15.78 ±5.68	97
PEC	0.8 ±0.4	0.3 ±0.1	34
PLN	11 ±01	7.2 ±3.4	62
Spleen	68 ±29	37 ±06	53
Thymus	23 ±08	8.4 ±3.8	36
Total Cells	120 ±30	69 ±10	56

Table 4.7.2: Estimates of haematopoietic cell numbers in rescued mice by FACScan analysis (continued)

i) Gϕ -positive Cells

Recipient Tissue	Number of cells staining positive for Gϕ surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	0.1 ±0.1	0.03 ±0.02	39
BM	4.5 ±2.3	0.9 ±0.8	19
MLN	0.04 ±0.02	0.04 ±0.02	100
PEC	0.02 ±0.01	0.01 ±0.00	52
PLN	0.02 ±0.01	0.01 ±0.01	70
Spleen	2.4 ±0.2	4.0 ±1.5	160
Thymus	0.05 ±0.03	0.07 ±0.03	150
Total Cells	7.1 ±2.3	5.0 ±1.7	70

j) Ter119-positive cells

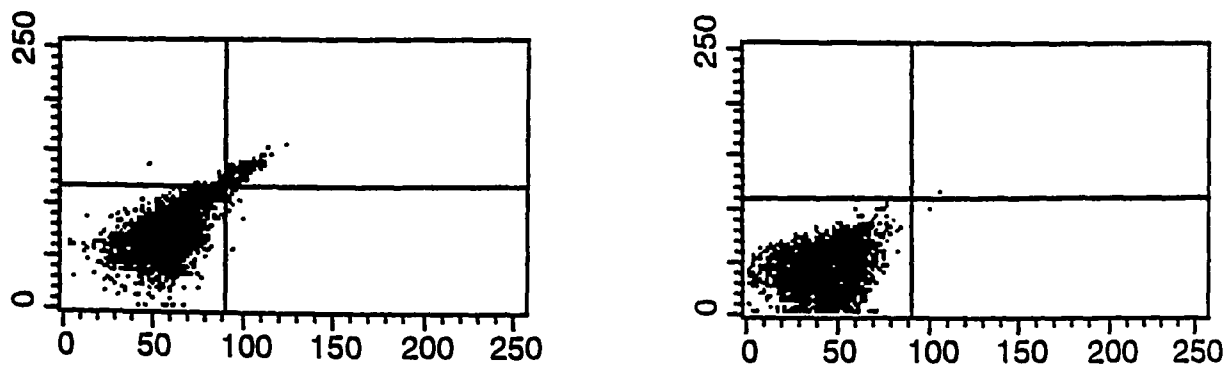
Recipient Tissue	Number of cells staining positive for Ter119 surface marker $X10^{-6}$		-/- % of wild-type (%)
	Wild-Type	-/-	
Blood	NA	NA	-
BM	0.9 ±0.1	0.5 ±0.1	52
MLN	1.7 ±1.5	1.6 ±1.4	94
PEC	0.5 ±0.3	0.1 ±0.1	16
PLN	0.7 ±0.6	1.0 ±1.0	140
Spleen	18 ±09	39 ±05	210
Thymus	0. 8 ±0.1	1.2 ±0.2	150
Totals	23 ±09	44 ±06	190

Figure 4.7.2: FACScan analysis of peripheral blood from rescued animals

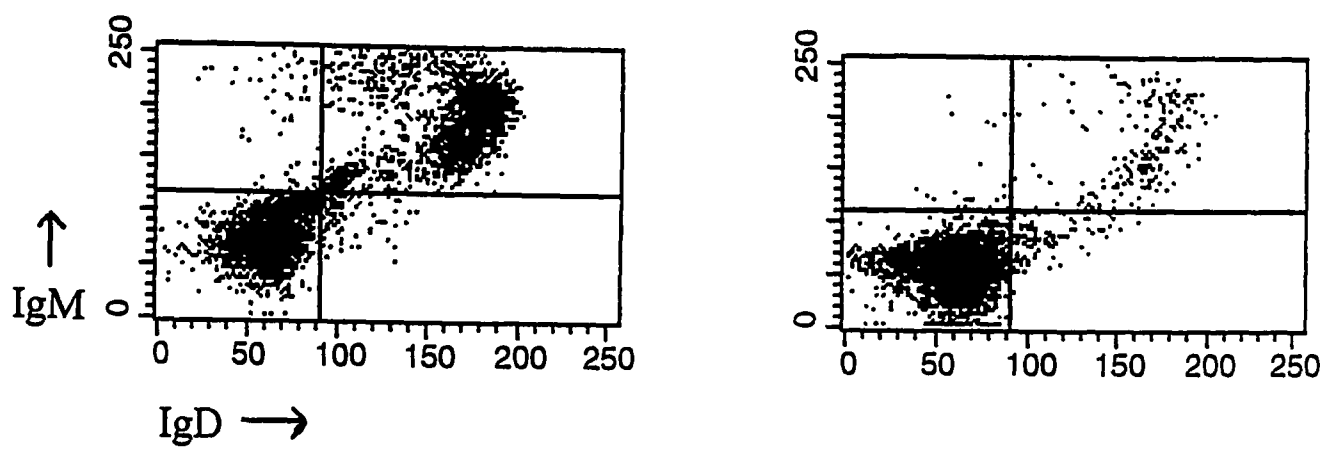
FACScan data plots of cell staining profiles of peripheral blood samples from mice rescued with wild-type and -/- cells. Erythrocytes were lysed and the remaining white cell fraction collected from the peripheral blood of the four long-term repopulated mice, and cell counts performed. Cell suspensions were subjected to FACScan analysis using pairs of fluorochrome-conjugated antibodies raised against (b) IgM and IgD, (c) CD4 and CD8, (d), CD3 and TCR$\alpha\beta$, (e) T200 and Ter-119, (f and g) B220 and Gϕ cell surface markers. Panel a shows an unstained control sample. For each sample 10,000 cells were analysed. Panels a, b, c, d, e and f show events contained within electronic gate R1, while panel g shows cells falling within region R3. The relative locations of electronic gates subdividing populations of cells on the basis of size (forward scatter) and internal complexity (side scatter) are depicted in panel h. Gate R1 was positioned to encompass the normal lymphocyte population, while R3 was intended to capture those cells with greater internal complexity e.g. granulocytes and actively proliferating lymphocytes. Gates R2, R4 and R5 were set in order to provide information on those populations of cells which were not contained within gates R1 and R3.

wild-type -/-

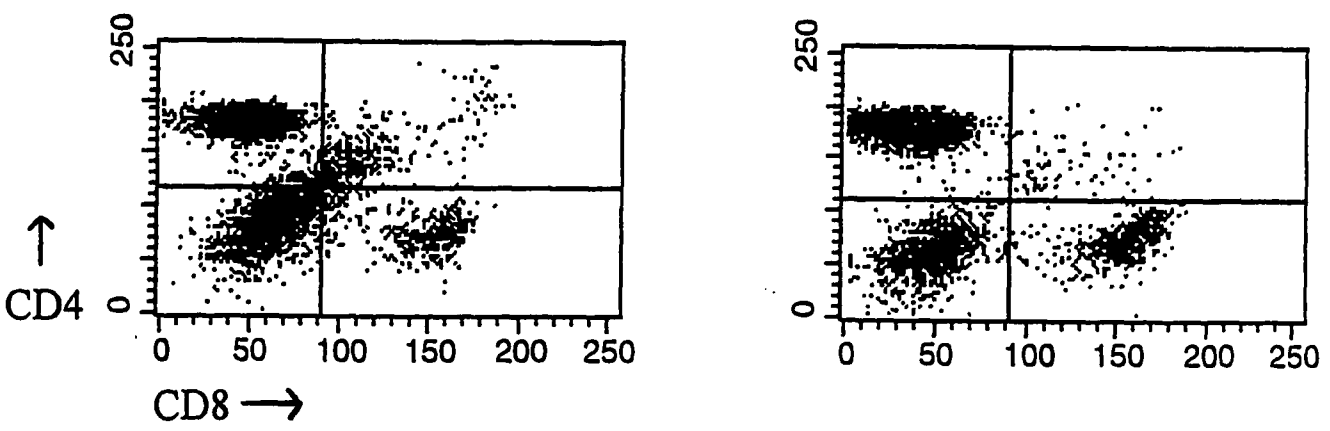
a) Unstained



b) IgM and IgD



c) CD4 and CD8



d) CD3 and TCRαβ

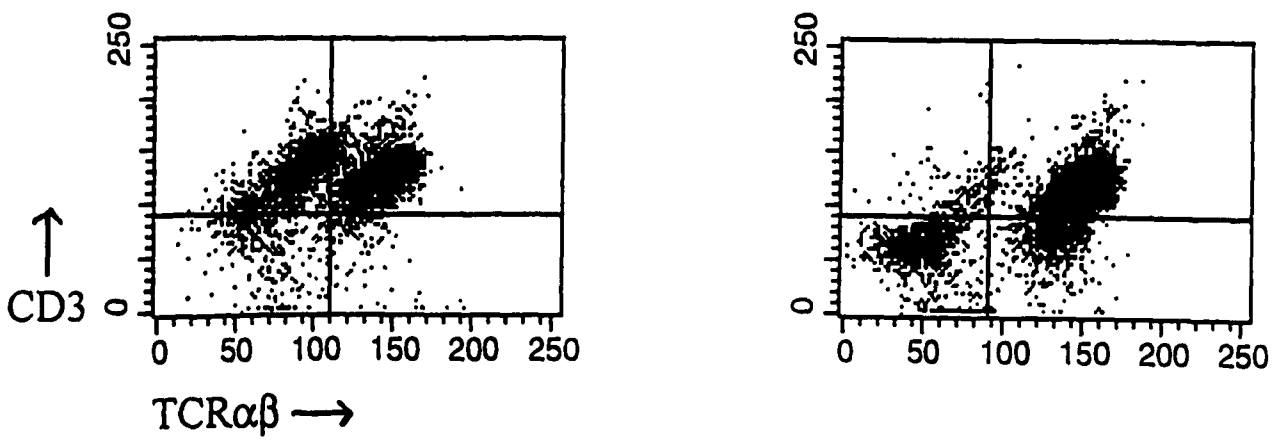
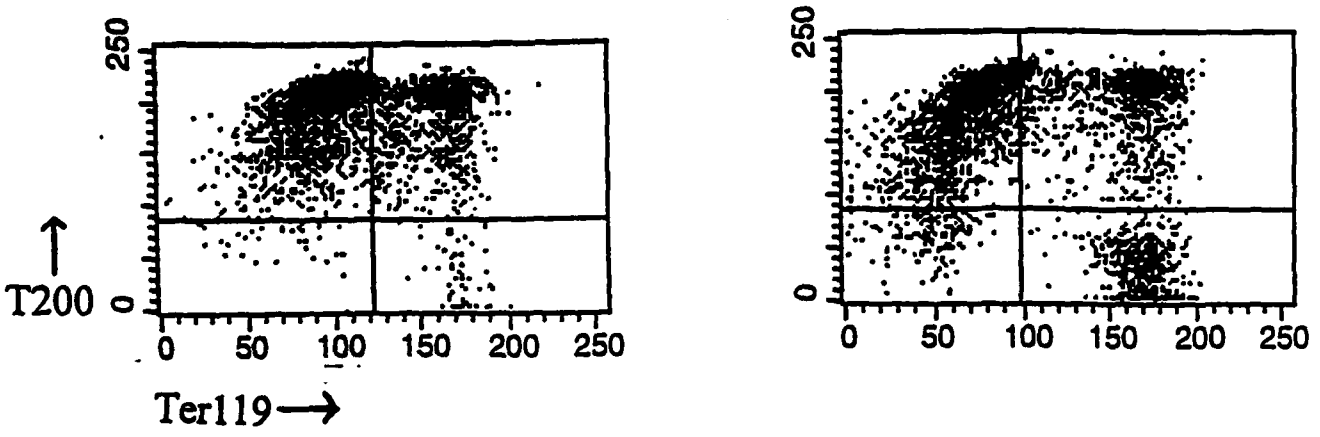
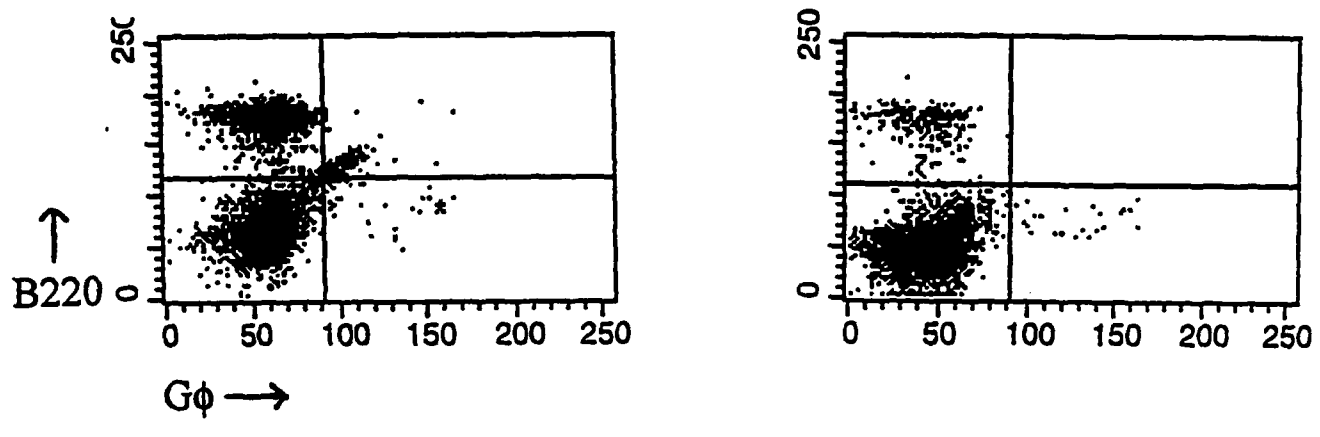


Figure 4.7.2: FACScan analysis of peripheral blood from rescued animals (continued)

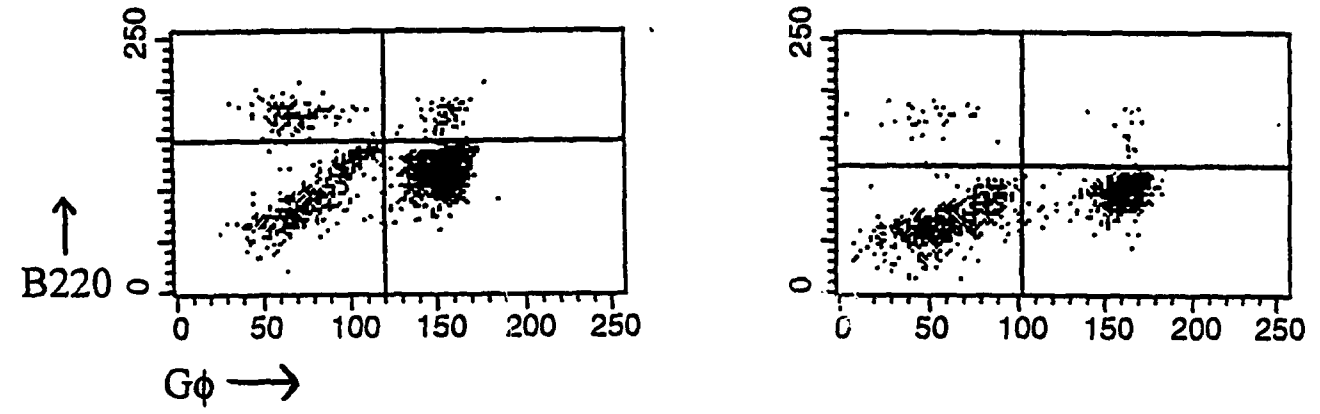
e) T200 and Ter119



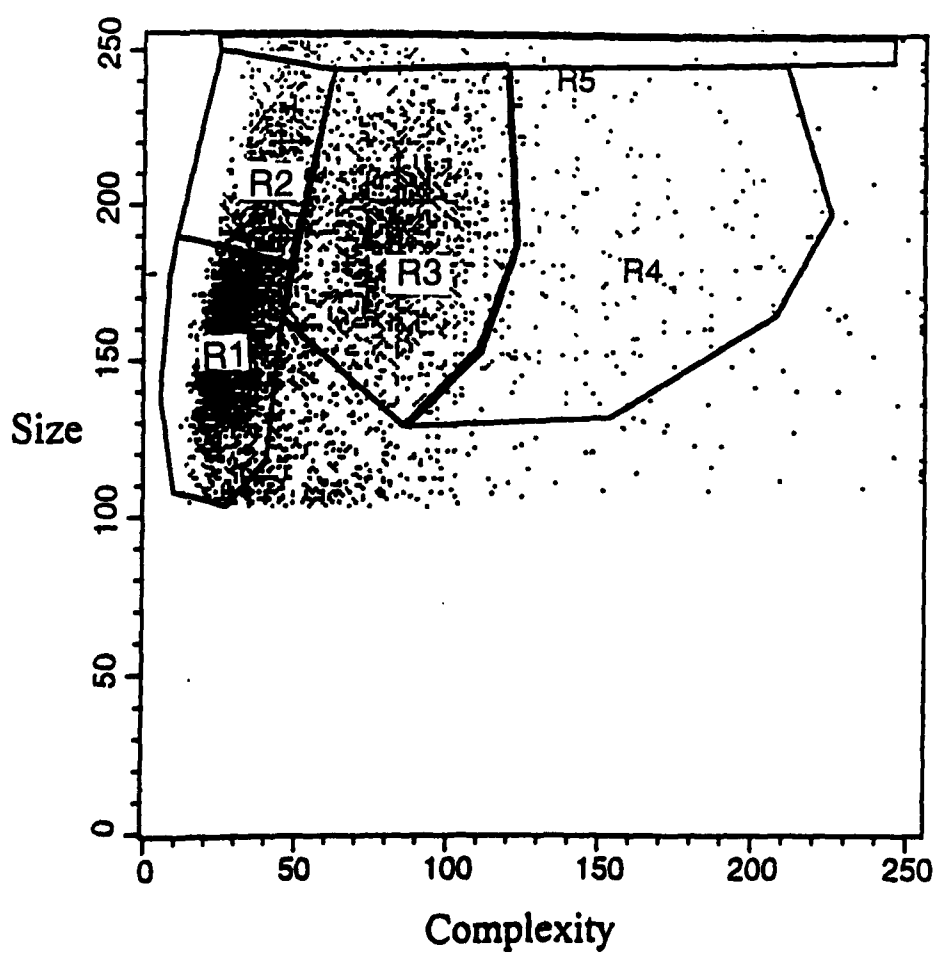
f) B220 and Gϕ (R1)



g) B220 and Gϕ (R3)



h) Location of electronic gates



CHAPTER FIVE:

MOLECULAR ANALYSIS OF CELLS DERIVED FROM -/- EMBRYOS

Embryos homozygous for the targeted *Lig1* allele displayed severe anaemia due to disruption of foetal liver erythropoiesis. The phenotype cosegregated with inheritance of two copies of the targeted allele, and showed complete penetrance. In addition, an independently derived line of *Lig1* targeted mice displayed a phenotype identical to that described in Chapter 3 (D. W. Melton, unpublished data). Taken together these data suggested that the observed phenotype was a result of DNA ligase I-deficiency. This chapter details further analysis to prove that samples PCR genotyped as being homozygous for the targeted allele lacked DNA ligase I.

A fibroblast cell line (46BR.1G1) derived from a patient with reduced DNA ligase I activity displayed abnormal lagging strand DNA synthesis. DNA ligase I had also been reported to be altered in cells from Bloom's Syndrome patients. These cells exhibited changes in the profile of replication intermediates, and high rates of SCEs. This chapter also describes investigation of lagging strand DNA replication and SCE rates in fibroblasts derived from -/- embryos. Furthermore, the effect of DNA ligase I-deficiency on the levels of other DNA ligases is reported, along with a pilot experiment to identify changes in expression of other genes.

5.1 Production of immortalised fibroblast lines

To facilitate investigation of the properties of cells lacking DNA ligase I, fibroblast cell lines were established from -/- embryos. Primary fibroblasts were isolated from E11.5 embryos and passaged until cells lost growth potential and underwent "crisis". Thereupon the medium was changed periodically until immortalised cell lines emerged. Two lines established proved to be homozygous for the targeted *Lig1/PGK-HPRT* allele (designated PFL10^T and PFL13^T), whilst a third wild-type line (PF20^T) had been previously derived by the same method. All three lines were apparently capable of being propagated indefinitely (A.-M. Ketchen and D. W. Melton, unpublished data).

5.2 Confirmation of the structure of the targeted *Lig1/PGK-HPRT* allele

The PCR reaction product specific for the targeted allele was designed to be amplified only when homologous recombination had inserted an *HPRT* minigene into the endogenous *Lig1* locus. However, generation of a PCR product did not guarantee that the locus had been targeted in precisely the desired manner, as targeting events can often be accompanied by other DNA rearrangements. In order to confirm that the correct homologous recombination event had occurred, Southern blot analysis was performed.

Southern analysis had been carried out on the original targeted heterozygous ES cell clone prior to injection into blastocysts (Jessop 1995), but interpretation of

the banding pattern was complicated by the presence of a wild-type allele in addition to the targeted *Lig1/PGK-HPRT* allele. Using *Eco*RI and *Hin*dIII restriction digests, targeting was marked only by alterations in relative band intensities.

To confirm that the targeted *Lig1/PGK-HPRT* allele did indeed have the predicted structure, and to verify the PCR genotyping of embryos, Southern blot analysis was carried out on genomic DNA samples representative of the three different *Lig1* genotypes (Figure 5.2.1). The results are summarised in Table 5.2.1, and the structures of the wild-type and targeted *Lig1/PGK-HPRT* alleles are shown in Figure 5.2.2 and Figure 5.2.3, along with predicted DNA fragments arising from restriction digestion. Those fragments highlighted in bold contain exons of the gene, and should thus be detected by hybridisation to a cDNA probe. The 13.9kb *Eco*RI genomic region corresponding to the 3' end of the wild-type allele contains four *Bam*HI fragments that include exons (>4.2, 1.3, 1.0 and 3.7kb respectively). Replacement of exons 22 to 27 by an *HPRT* minigene deletes the 1.3, 1.0 and 3.7kb fragments, but also results in the extension of the >4.2kb fragment to >7.9kb. Thus targeting should result in the disappearance of the >4.2, 1.3, 1.0, and 3.7kb fragments, and the appearance of a novel >7.9kb band.

When wild-type genomic DNA (HM-1, +/+ embryo, and $PF20^T$) was digested with *Bam*HI, and subject to Southern blotting using a cDNA probe corresponding to exons 8 to 27 of the *Lig1* gene, six bands were detected (1.0, 1.3, 1.9, 3.6, 3.9, and 9.4kb). Previous work had shown that the 3.6 and 1.9kb bands were derived from regions of the *Lig1* locus unaffected by targeting, and that the wild-type 9.4kb band corresponded to the predicted >4.2kb band (Jessop 1995). In samples heterozygous for the targeted allele (#53, +/- embryo), a novel 13.1kb fragment was observed and the 9.4, 3.9, 1.3 and 1.0kb bands were reduced in intensity compared to the 3.6 and 1.9kb bands. Lanes corresponding to -/- samples (-/- embryo, $PFL10^T$, and $PFL13^T$) exhibited only 3 bands: 13.1, 3.6 and 1.9kb. The 9.4, 3.9, 1.3 and 1.0kb fragments were all completely absent. Therefore the pattern of restriction fragments observed changed in the expected manner after targeting, and was entirely consistent with the proposed structure of the targeted allele. It was also concluded that Southern blot analysis corroborated the *Lig1* genotypes determined by PCR analysis.

Figure 5.2.1: Southern blot analysis of samples representative of the different *Lig1* genotypes

Confirmation of the structure of the targeted *Lig1/PGK-HPRT* allele and PCR genotyping by Southern analysis. Genomic DNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the *Lig1/PGK-HPRT* targeted allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). DNA was then digested with *Bam*HI, electrophoresed through an agarose gel, blotted onto nylon membrane and probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA, corresponding to exons 8 to 27. *Lig1* genotypes of the different samples, derived by PCR, are shown, along with approximate sizes of the restriction fragments detected.

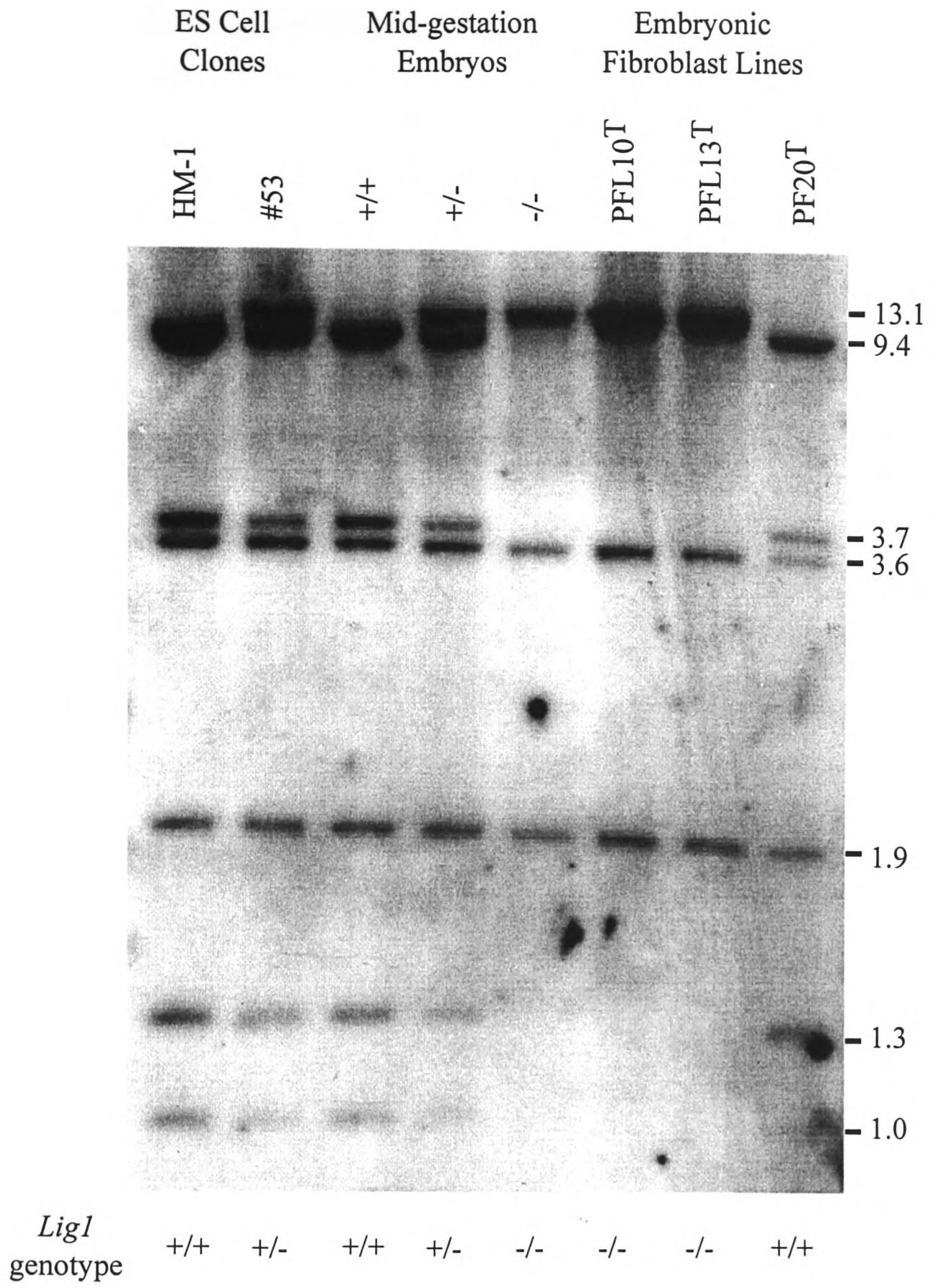

ES Cell Clones
Mid-gestation Embryos
Embryonic Fibroblast Lines
HM-1
#53
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
13.1
9.4
3.7
3.6
1.9
1.3
1.0
Lig1 genotype
+/+ +/- +/+ +/- -/- -/- -/- +/+

Figure 5.2.2: Restriction maps of the wild-type *Lig1* allele

Predicted restriction map of the 3' end of the wild-type *Lig1* allele. The wild-type map was based on restriction mapping and Southern hybridisation data from Jessop 1995, sequences of subcloned genomic fragments, and the sizes of certain PCR products (D. W. Melton, unpublished data). The locations of exons 25 and 26 are not certain beyond that neither are present within the 1.3kb *Kpn*I fragment, and at least one must lie within the 1.0kb *Bam*HI fragment. However, because exon 26 is known to lie only 140bp upstream of exon 27 in humans, it has been provisionally assigned to the 0.6kb *Kpn*I/*Sst*I fragment.

Drawn to scale: 1cm = 2kb. Black boxes indicate exons; lines, introns or flanking DNA. Restriction sites of *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I are indicated, along with the arrangement of resulting restriction fragments for each individual enzyme. Sizes are given in kb. Fragments (and associated sizes) which contain exons, and hence will be detected on Southern blots probed with *Lig1* cDNA, are in bold.

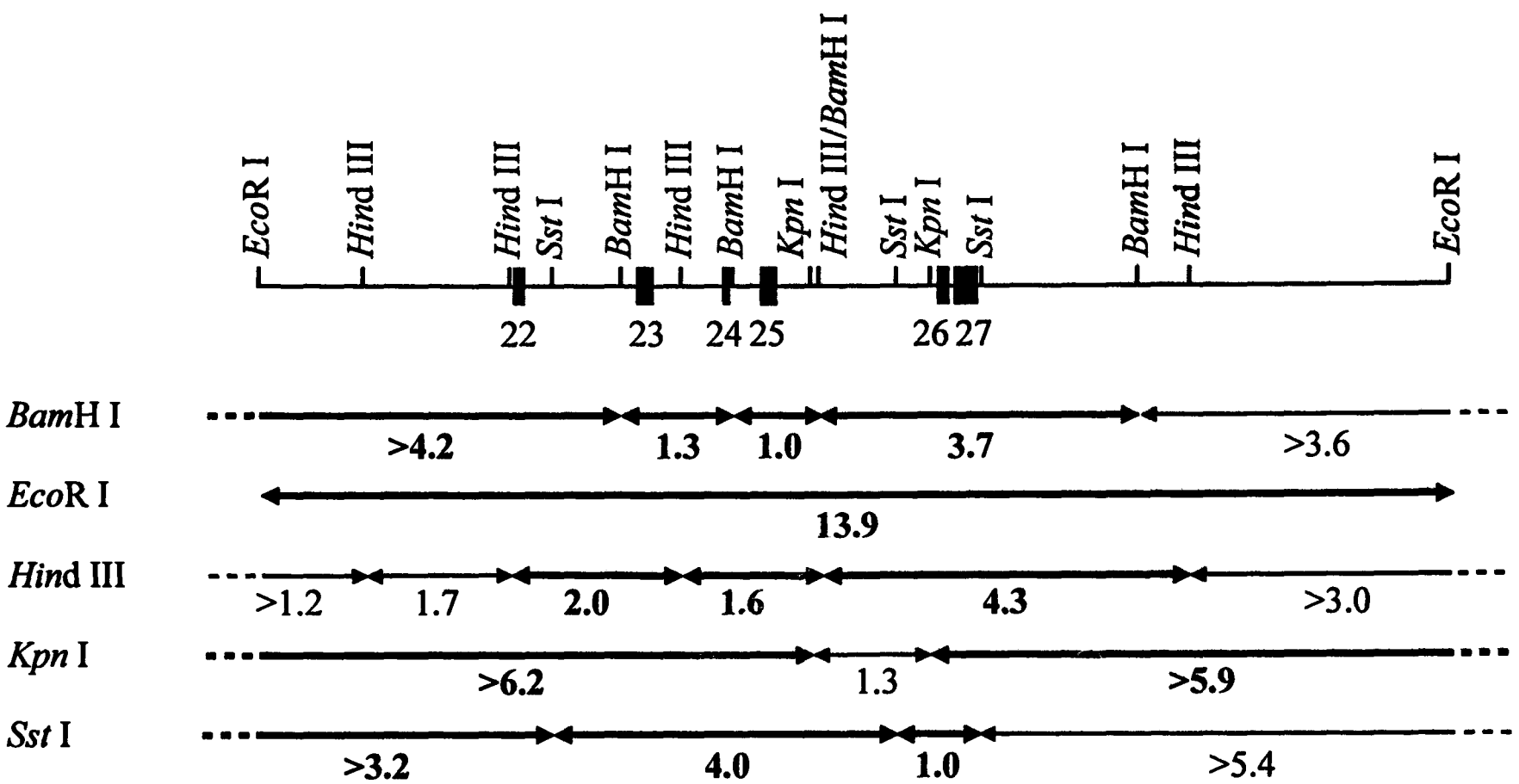

EcoR I
Hind III
Hind III
Sst I
BamH I
Hind III
BamH I
Kpn I
Hind III/BamH I
Sst I
Kpn I
Sst I
BamH I
Hind III
EcoR I
22
23
24 25
26 27
BamH I
>4.2
1.3
1.0
3.7
>3.6
EcoR I
13.9
Hind III
>1.2
1.7
2.0
1.6
4.3
>3.0
Kpn I
>6.2
1.3
>5.9
Sst I
>3.2
4.0
1.0
>5.4

Figure 5.2.3: Restriction map of the targeted *Lig1/PGK-HPRT* allele

Predicted restriction map of the 3' end of the targeted *Lig1/PGK-HPRT* allele. The map was based on the known structures of the wild-type allele and MLTV gene targeting vector (Jessop 1995). Drawn to scale: 1cm = 2kb. Black boxes indicate exons; lines, introns or flanking DNA; hatched box,*HPRT* minigene. Restriction sites of *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I are indicated, along with the arrangement of resulting restriction fragments for each individual enzyme. Sizes are given in kb. Fragments (and associated sizes) which contain exons, and hence will be detected on Southern blots probed with *Lig1* cDNA, are in bold.

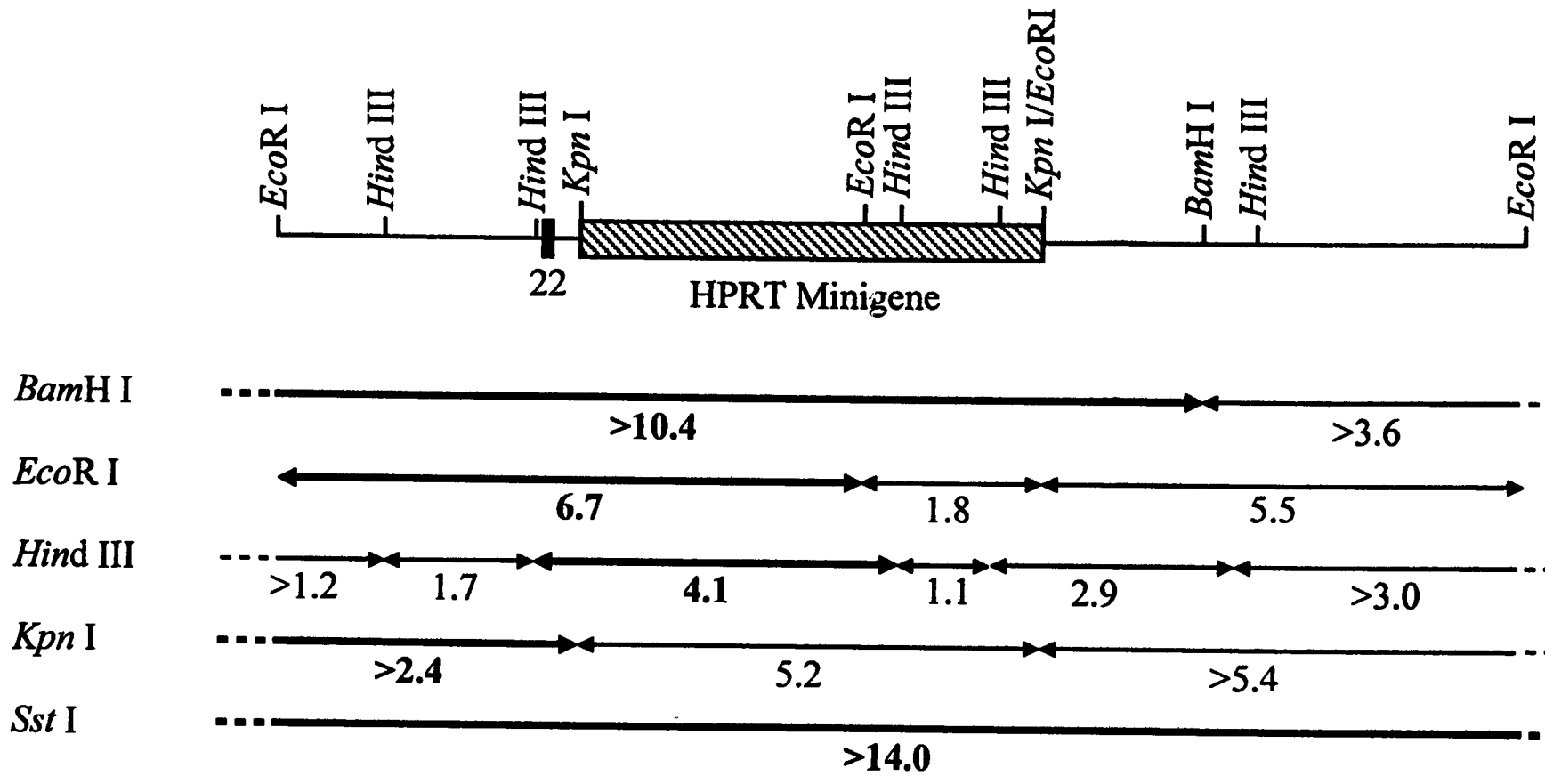

EcoR I
Hind III
Hind III
Kpn I
EcoR I
Hind III
Hind III
Kpn I/EcoRI
BamH I
Hind III
EcoR I
22
HPRT Minigene
BamH I
>10.4
>3.6
EcoR I
6.7
1.8
5.5
Hind III
>1.2
1.7
4.1
1.1
2.9
>3.0
Kpn I
>2.4
5.2
>5.4
Sst I
>14.0

Table 5.2.1: Summary of the predicted *Bam*HI restriction fragments from wild-type and targeted *Lig1/PGK-HPRT* alleles, and bands detected by Southern blotting

Predicted *Bam*HI Fragments		Observed Banding Pattern		
Wild-Type Allele	Targeted Allele	+/+ Samples	+/- Samples	-/- Samples
>4.2	>7.9	9.4	13.1	13.1
3.7		3.7	9.4*	(3.6)
1.3		(3.6)	3.7*	(1.9)
1.0		(1.9)	(3.6)	
		1.3	(1.9)	
		1.0	1.3*	
			1.0*	

Predicted and actual *Bam*HI restriction fragments arising from the 3' end of the mouse DNA ligase I locus when probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA. Fragment sizes in brackets are those arising from parts of the gene outwith the 13.9kb *Eco*RI fragment used in the targeting experiments. Fragments marked with an asterisk are those which were present at a reduced intensity compared to +/+ samples.

5.3 Expression from the targeted *Lig1/PGK-HPRT* allele is abolished

The effect that a targeted alteration has on gene expression can vary between loci. Commonly insertion of a marker gene results in production of a fusion transcript and subsequently a truncated or fusion protein. Introduction of the *HPRT* marker into the *Lig1* locus is accompanied by deletion of exons 23 to 27 which should have effectively disrupted the integrity of the gene. However, to ascertain the nature of any product expressed from the targeted *Lig1/PGK-HPRT* allele, northern and western analysis was carried out.

Total RNA was isolated from the same representative panel of samples as employed for Southern analysis. ES cell and fibroblast cell cultures were passaged one day prior to harvesting to ensure that cells were actively dividing. The RNA was separated by electrophoresis, blotted onto nylon membrane, then hybridised to a *Lig1* cDNA probe (Figure 5.3.1). The 3.2kb *Lig1* transcript was readily detected in wild-type samples (HM-1, +/+ embryo, and PF20^{T}). Samples heterozygous for the targeted allele (#53, +/- embryo) displayed a single band of the same size as the wild-type product but of reduced intensity. This was consistent with expression only from the

single wild-type allele and not from the targeted allele. No bands of any size were observed in lanes corresponding to -/- samples (-/- embryo, PFL10^T, and PFL13^T), suggesting that no stable transcripts were produced from the targeted allele. Although Figure 5.3.1 shows the result achieved using a *Lig1* cDNA probe corresponding to exons 8 to 27, an identical result was achieved using a probe for exons -1 to 7 (data not shown)

Western blot analysis was then performed on a representative panel of embryos and fibroblast cell lines. Whole cell protein extracts were prepared by brief homogenisation of tissue or cell pellets in protein lysis buffer. Proteins were separated on the basis of size by SDS-PAGE, blotted onto nylon membrane, and probed with antibodies raised against purified bovine DNA ligase I (Figure 5.3.2). A band of ~125kDa corresponding to DNA ligase I was detected in wild-type samples. The protein was present at reduced levels in +/- samples, while -/- lanes contained no detectable DNA ligase I. Because the anti-DNA ligase I antisera employed predominantly recognised epitopes located in the N-terminus of the protein (D. Barnes, ICRF Clare Hall, personal communication) any truncated peptides should be detected. However, no DNA ligase I-derived peptide of any size was detected in samples from -/- embryos or fibroblasts. Visual inspection of bands arising from non-specific crossreactivity did not reveal any other of these 'background' bands to be upregulated in +/- or -/- samples compared to wild-type.

Figure 5.3.2b shows western analysis of varying amounts of protein extract from a wild-type embryo spiked into a -/- embryo sample. As little as 1% of the amount of DNA ligase I normally present in wild-type embryos could be detected by western analysis. It was therefore reasonable to suppose that any significant level of DNA ligase I or DNA ligase I-derived peptides would have been detected. Thus we were unable to detect either mRNA or protein expression from the targeted *Lig1* allele. We therefore concluded that expression from the targeted allele was abrogated, and that DNA ligase I was absent from -/- embryos.

Figure 5.3.1: Absence of detectable expression from the targeted *Lig1/PGK-HPRT* allele

Northern blot analysis of samples representative of the different *Lig1* genotypes. Total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the targeted *Lig1/PGK-HPRT* allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). Upper, RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA, corresponding to exons 8 to 27. Positions of the 28S and 18S ribosomal RNA transcripts are indicated flanking the 3.2kb wild-type *Lig1* transcript. *Lig1* genotypes of the different samples are shown. Lower, ethidium bromide staining to verify equal RNA loading.

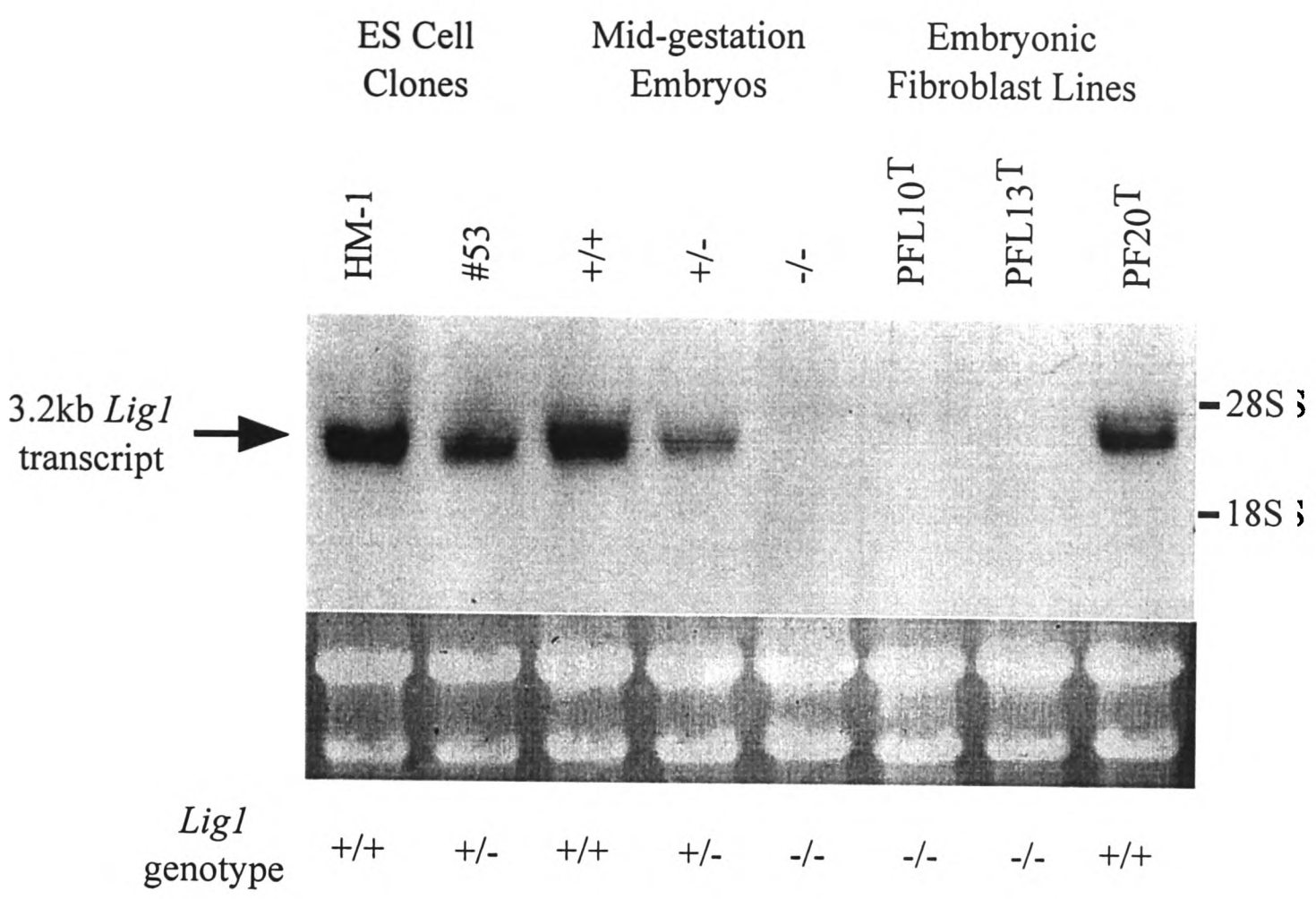
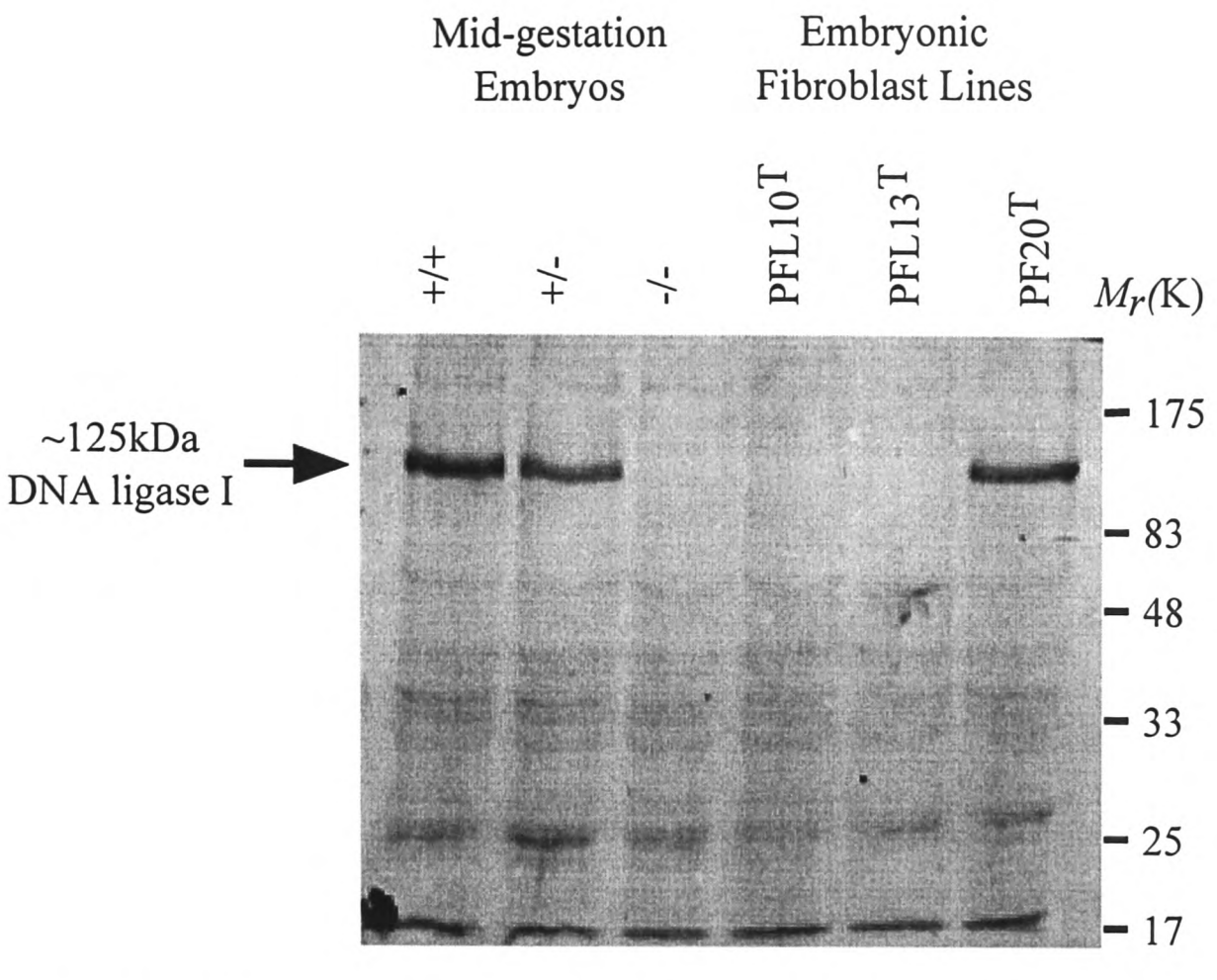
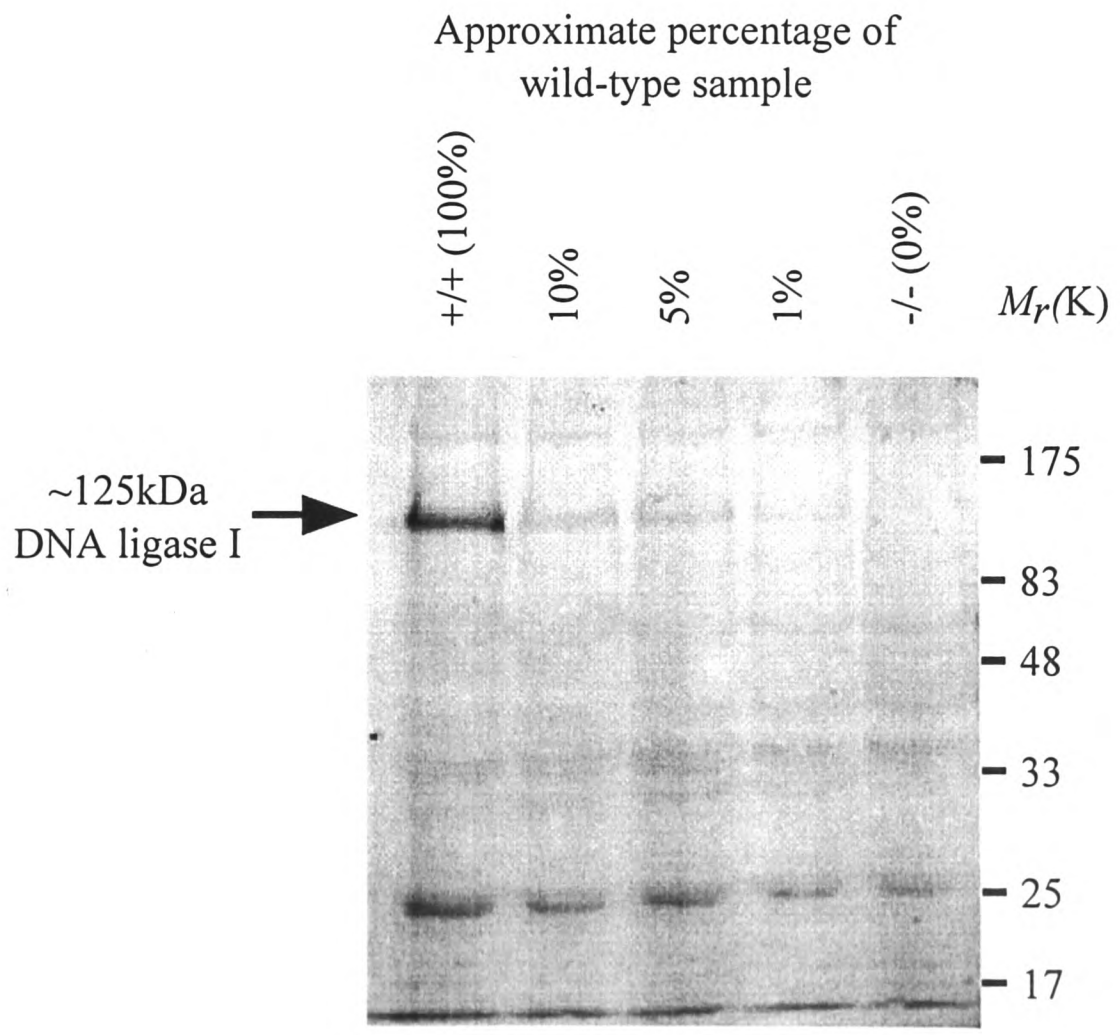

ES Cell Clones
Mid-gestation Embryos
Embryonic Fibroblast Lines
HM-1
#53
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
3.2kb Lig1 transcript
28S
18S
Lig1 genotype
+/+ +/- +/+ +/- -/- -/- -/- +/+

Figure 5.3.2: Absence of detectable DNA ligase I in -/- samples.

a) Western blot analysis of samples representative of the different *Lig1* genotypes. Total protein extracts were prepared from E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). Protein lysates were made by brief homogenisation of tissue or cell pellets in lysis buffer, and 100μg loaded onto an 8% SDS-polyacrylamide gel. Separated proteins were blotted onto a nylon membrane, blocked in TBST plus 5% non-fat milk powder, and incubated with TL5 rabbit polyclonal antibody raised against purified bovine DNA ligase I. Detection of bound antibody was achieved using alkaline phosphatase-conjugated anti-rabbit IgG antibodies. Positions of protein molecular weight markers are indicated, as is the ~125kDa wild-type DNA ligase I protein. *Lig1* genotypes of the different samples are shown.

b) Western blot detection of varying amounts of DNA ligase I. Decreasing amounts of whole cell protein lysate prepared from a +/+ E13.5 embryo were spiked into a lysate from a -/- embryo. Details as above. In total 100μg of protein extract was loaded in each lane. The approximate percentages of wild-type protein present in each lane are indicated.


Mid-gestation Embryos
Embryonic Fibroblast Lines
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
Mr(K)
~125kDa DNA ligase I
175
83
48
33
25
17
Lig1 genotype
+/+
+/-
-/-
-/-
-/-
+/+


Approximate percentage of wild-type sample
+/+ (100%)
10%
5%
1%
-/- (0%)
Mr(K)
~125kDa DNA ligase I
175
83
48
33
25
17

5.4 DNA ligases III and IV in -/- embryos and fibroblasts

The viability of cells lacking DNA ligase I was unexpected given the wealth of data published on the involvement of the enzyme in lagging strand DNA replication. We hypothesised that for DNA replication to proceed, another ligase activity must compensate for the absence of DNA ligase I. Because DNA ligase I is far more abundant in actively dividing cells than any of the other DNA ligases, it would not be unreasonable to expect the levels of any complementing ligase to be increased. To test this hypothesis, we investigated the expression levels of the other known DNA ligases in samples lacking DNA ligase I.

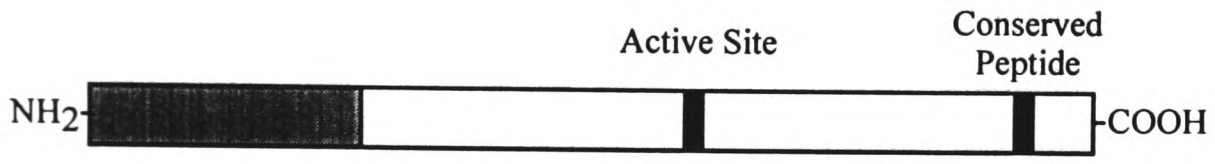
5.4.1 Cloning of cDNA fragments of mouse DNA ligases III and IV

cDNAs encoding human DNA ligases III and IV had previously been isolated (Wei *et al.* 1995), but there had been no reports of potential murine homologues. In order to clone fragments of mouse DNA ligase III and IV cDNAs for use in northern blot analysis, RT-PCR was performed on mouse total RNA. The primers used were designed from the published human sequences (Database accession numbers: *LIG3*; X84740, *LIG4*; X83441), taking advantage of homology between the catalytic domains of all ATP-dependent DNA ligases (Figure 5.4.1).

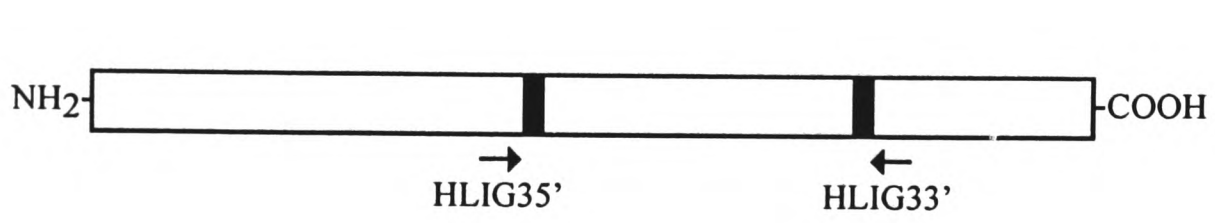
To clone a fragment of DNA ligase III cDNA, PCR primers were designed to bind a short distance upstream of (5' to) the enzyme active site (HLIG35'), and a short distance downstream of (3' to) the conserved peptide (HLIG33'). The binding sites of these primers were separated by 1.0kb on the human *LIG3* cDNA. Northern analysis of the distribution of *LIG3* mRNA had shown maximal expression in human testes (Wei *et al.* 1995). Total RNA isolated from mouse testes was reverse transcribed, and the reaction products employed as a template for PCR using the primer pair HLIG35' and HLIG33'. PCR products were separated by agarose gel electrophoresis, and a 1.0kb fragment purified. The fragment was then cloned into the *Eco*RV restriction site of pBluescriptII SK+ (designated pMLIG3) and sequenced. The predicted amino acid sequence of the insert showed high homology to the amino acid sequence of human DNA ligase III (Figure 5.4.2). When the cloned product was used as a probe for northern blot analysis of a panel of wild-type adult mouse tissues, two transcripts were detected (Figure 5.4.3) The smaller band was approximately 3kb in size, while the second transcript detected was significantly larger, and corresponded to a size of about 6kb. Both transcripts were abundant in the testis, but expressed at considerably lower levels in all other tissues. The identity of the larger band, and the relationship between the two transcripts, is discussed further

Figure 5.4.1: Schematic representation of DNA ligases I, III, and IV

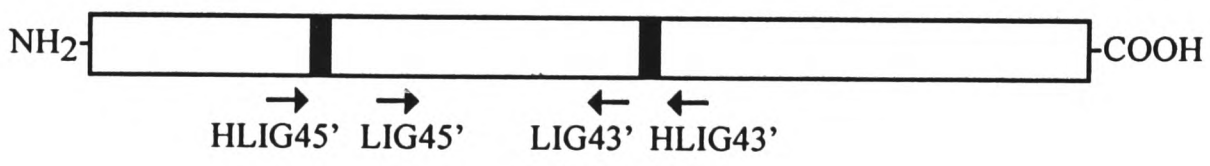
Alignment of the three cloned human DNA ligases. Adapted from Wei *et al.* 1995. The active site and conserved peptide motifs are represented by solid bars. Shaded box, N-terminal regulatory domain of DNA ligase I. Predicted molecular masses are indicated, as are the relative locations at which oligonucleotide primers bind in the *LIG3* and *LIG4* cDNA sequences.



DNA ligase I
(102kDa)



DNA ligase III
(103kDa)



DNA ligase IV
(96kDa)

Figure 5.4.2: The amino acid sequence of the putative mouse DNA ligase III homologue cDNA fragment aligned with the published human DNA ligase III sequence.

The nucleotide sequence of the PCR product cloned in pMLIG3 was translated and the predicted amino acid sequence aligned with that of human DNA ligase III using the Bestfit program from the Genetics Computer Group package, version 9 (Genetics Computer Group 1996). Vertical lines indicate identical amino acid residues; two dots, a conservative substitution; one dot, a semi-conservative change. Numbers denote the location of residues within the published *LIG3* amino acid sequence (Database accession number: X84740) (top), and the translated pMLIG3 sequence (bottom). The relative locations at which oligonucleotide primers used in RT-PCR amplification bound are shown.

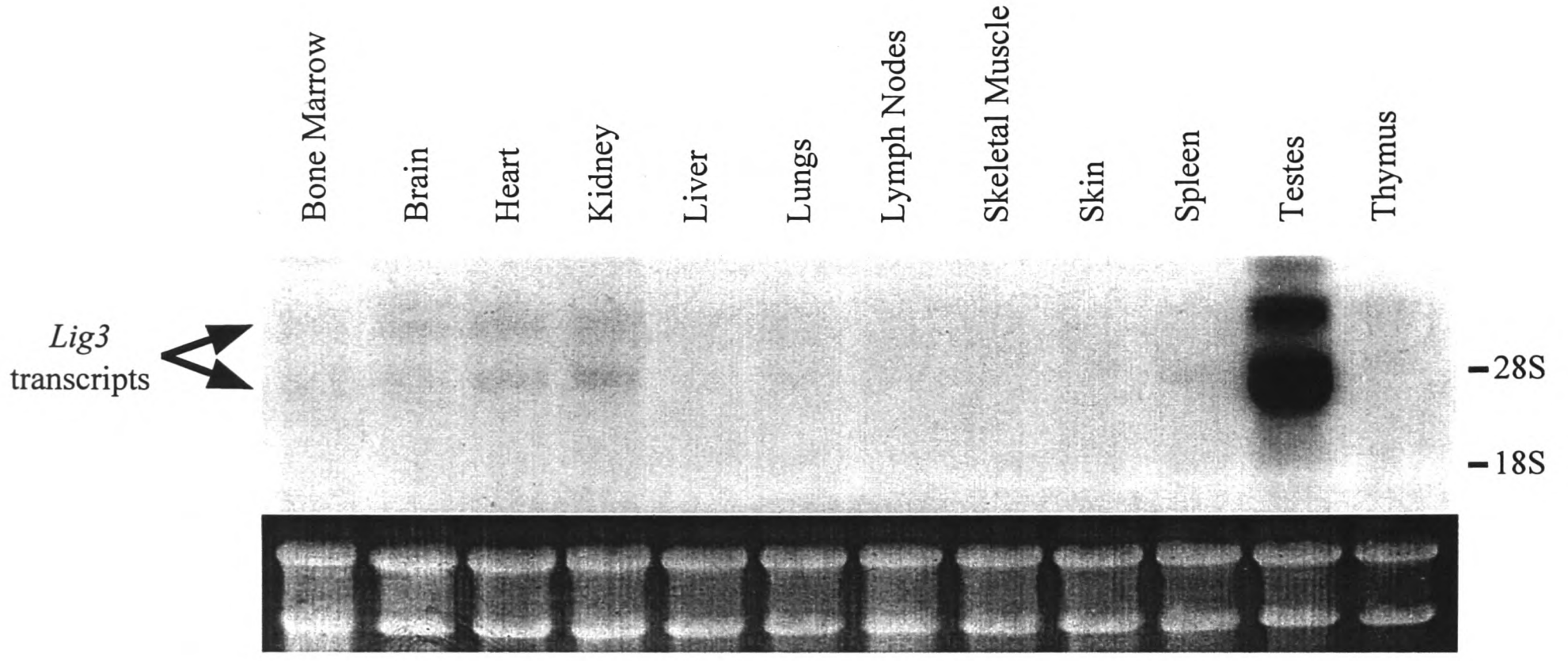
HLIG35'

```
411 KCPNGMFSEIKYDGERVQVHKNGDHFSYFSRSLKPVLPHKVAHFKDYIPQ 460
    |||.||||||||||||||||| |||||||||||||||||| |||||||.
  1 KCPHGMFSEIKYDGERVQVHKKGDHFSYFSRSLKPVLPHKVTHFKDYIPK 50

461 AFPGGHSMILDSEVLLIDNKTGKPLPFGTLGVHKKAAFQDANVCLFVFDC 510
    ||||| |||||||||||| |||||||||||||||||||||||||||||||
 51 AFPGGQSMILDSEVLLIDNNTGKPLPFGTLGVHKKAAFQDANVCLFVFDC 100

511 IYFNDVSLMDRPLCERRKFLHDNMVEIPNRIMFSEMKRVTKALDLADMIT 560
    |||||||||||||||||||||||| |||||||||.|||| ||||||
101 IYFNDVSLMDRPLCERRKFLHDNMVEIRNRIMFSEMKQVTKASDLADMIN 150

561 RVIQEGLEGLVLKDVKGTYEPGKRHWLKVKKDYLNEGAMADTADLVVLGA 610
    |||..|||||||||||||||||||||||||||||||||||||||||||||
151 RVIRKGLEGLVLKDVKGTYEPGKRHWLKVKKDYLNEGAMADTADLVVLGA 200

611 FYGQGSKGGMMSIFLMGCYDPGSQKWCTVTKCAGGHDDATLARLQNELDM 660
    |||||.|||||||||||||:| ||||||||||||||||||||| || |
201 FYGQGNKGGMMSIFLMGCYEPDSQKWCTVTKCAGGHDDATLARLQKELVM 250

661 VKISKDPSKIPSWLKVNKIYYPDFIVPDPKKAAVWEITGAEFSKSEAHTA 710
    ||||||||||||||:|||||||||||||||||||||||:||||||
251 VKISKDPSKIPSWLKINKIYYPDFIVPDPKKAAVWEITGAEFSRSEAHTA 300

711 DGISIRFPRCTRIRDDKDWKSATNLP 736
    |||||||||||||||||||||||||
301 DGISIRFPRCTRIRDDKDWKSATNLP 326
```

HLIG33'

Figure 5.4.3: *Lig3* expression is highest in testes

Northern blot analysis of *Lig3* expression in a panel of wild-type mouse tissues. Upper, total RNA was prepared from a variety of adult tissues, electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with a 1.0kb *Eco*RI/*Hin*dIII fragment of pMLIG3. Positions of the 28S and 18S ribosomal RNA transcripts are indicated, as are the two *Lig3* transcripts (3.0 and ~6kb). Lower, ethidium bromide staining to verify equal RNA loading.


Bone Marrow
Brain
Heart
Kidney
Liver
Lungs
Lymph Nodes
Skeletal Muscle
Skin
Spleen
Testes
Thymus
Lig3
transcripts
28S
18S

in Chapter 7, but the data were not inconsistent with the supposition that the cloned product was a fragment of the cDNA of the mouse homologue of DNA ligase III.

To clone a fragment of DNA ligase IV cDNA, PCR primers were designed to bind a short distance upstream of the enzyme active site (HLIG45'), and a short distance downstream of the conserved peptide (HLIG43'). The binding sites of these primers were separated by 1.0kb on the human *LIG4* cDNA sequence. Northern analysis of the distribution of *LIG4* mRNA had shown highest levels of expression in testes and thymus (Wei *et al.* 1995). Thus, total RNA isolated from mouse thymus was reverse transcribed, and the reaction products employed as a template for PCR using the primer pair HLIG45' and HLIG43'. However, no product of the appropriate size was generated from this reaction. Therefore, in order to maximise the chances of amplifying a fragment of the *Lig4* cDNA, a nested PCR strategy was employed. A second pair of primers was designed based on the human *LIG4* cDNA sequence. The binding sites of these primers (LIG45' and LIG43') were separated by 0.75kb on the cDNA sequence, and both were within the 1.0kb region of the human cDNA defined by primers HLIG45'and HLIG43'. Primers HLIG45' and HLIG43' were employed in a PCR reaction on reverse transcribed mouse thymus RNA. 5μl of this reaction was then used to seed a second sequential PCR reaction, this time utilising primers LIG45' and LIG43'. The reaction products were separated by agarose gel electrophoresis, and a 0.75kb fragment purified. The PCR product was then cloned into the *Eco*RV restriction site of pBluescriptII SK+ (designated pMLIG4) and sequenced. The predicted amino acid sequence of the insert showed high homology to the amino acid sequence of human DNA ligase IV (Figure 5.4.4). When the cloned product was used as a probe for northern blot analysis of a panel of wild-type adult mouse tissues, a single 3kb transcript was detected (Figure 5.4.5) The transcript was abundant in the testes and thymus, but also expressed at lower levels in certain other tissues. These data were consistent with the cloned product being a fragment of the cDNA of the mouse homologue of DNA ligase IV.

Figure 5.4.4: The amino acid sequence of the putative mouse DNA ligase IV homologue cDNA fragment aligned with the published human DNA ligase IV sequence.

The nucleotide sequence of the PCR product cloned in pMLIG4 was translated and the predicted amino acid sequence aligned with that of human DNA ligase IV using the Bestfit program from the Genetics Computer Group package, version 9 (Genetics Computer Group 1996). Vertical lines indicate identical amino acid residues; two dots, a conservative substitution; one dot, a semi-conservative change. Numbers denote the location of residues within the published *LIG4* amino acid sequence (Database accession number: X83441) (bottom), and the translated pMLIG4 sequence (top). The relative locations at which oligonucleotide primers used in RT-PCR amplification bound are shown.

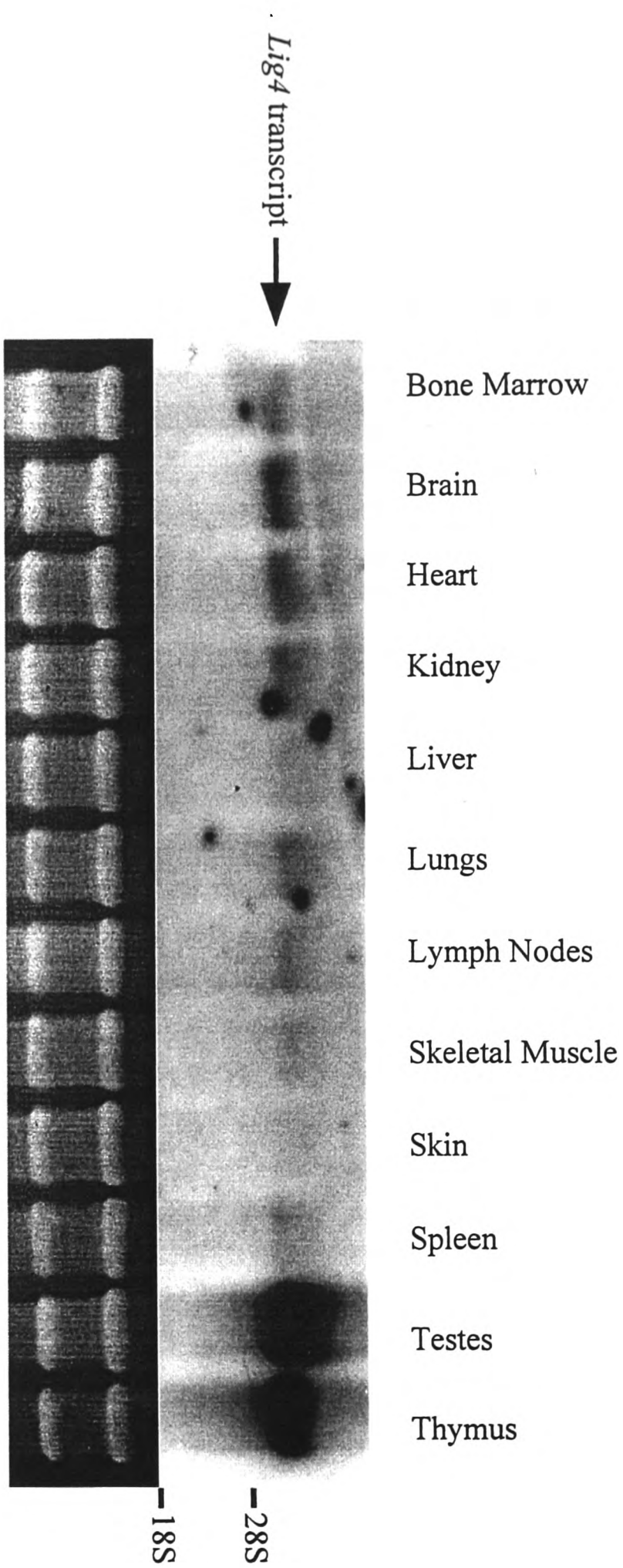
LIG45'

```
  1 ILDGEMMAYNPTTQTFMQKGVKFDIKRMVEKSCLQTCYSVFELMMGNKKK 50
    |||||||||||| |||||||| ||||||||| | ||||| ||:.:| | ||
260 ILDGEMMAYNPNTQTFMQKGTKFDIKRMVEDSDLQTCYCVFDVLMVNNKK 309

 51 PGAQTLRGGYEILCSIFTPIKGRTEIVQKIQAHTKEEVVDALNDALDKRE 100
     | :|||  |||| |||||| || ||||| ||||| ||:||||:|:||||
310 LGHETLRKRYEILSSIFTPIPGRIEIVQKTQAHTKNEVIDALNEAIDKRE 359

101 EGIMVKHXLSIHKPDKRGEGWLKIKPEYVSGLMDELDVLIVGGYWGKGSR 150
    ||||||  |||:||||||||||||||||||||||:||||||||||||||
360 EGIMVKQPLSIYKPDKRGEGWLKIKPEYVSGLMDELDILIVGGYWGKGSR 409

151 GGMMSHFLCAVAETPPPGDRPSVFHTLCRVGSGYTMKELYDLGLKLAKYW 200
    |||||||||||||| ||||::||||||| ||||| ||||||||||||||||
410 GGMMSHFLCAVAEKPPPGEKPSVFHTLSRVGSGCTMKELYDLGLKLAKYW 459

201 KPFHKKSPPSSILCGTEKPEVYIEPQNSVIVQIKAAEIVPSDMYKTGCTL 250
    ||||:|.|||||||||||||||| ||||||||||||||||||||||||||
460 KPFHRKAPPSSILCGTEKPEVYIEPCNSVIVQIKAAEIVPSDMYKTGCTL 509
```

LIG43'

Figure 5.4.5: *Lig4* expression is highest in testes and thymus

Northern blot analysis of *Lig4* expression in a panel of wild-type mouse tissues. Upper, total RNA was prepared from a variety of adult tissues, electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with a 0.75kb *Eco*RI/*Hin*dIII fragment of pMLIG4. Positions of the 28S and 18S ribosomal RNA transcripts are indicated flanking the ~3.0kb *Lig4* transcript. Lower, ethidium bromide staining to verify equal RNA loading.


Lig4 transcript
Bone Marrow
Brain
Heart
Kidney
Liver
Lungs
Lymph Nodes
Skeletal Muscle
Skin
Spleen
Testes
Thymus
18S
28S

5.4.2 Expression of *Lig3* and *Lig4* is unaffected by the absence of DNA ligase I

Northern blot analysis was performed on RNA from a panel of embryos, ES cell, and fibroblast cultures, to assess levels of *Lig3* and *Lig4* expression in the context of altered *Lig1* expression. Total RNA from ES cell, embryo and fibroblast samples representative of the different *Lig1* genotypes was separated by electrophoresis, blotted onto nylon membrane, and then hybridised to cDNA fragments corresponding to either *Lig3* or *Lig4*. Levels of the two transcripts were then assessed by phosphorimagery. *Lig3* transcripts were detected at equal levels in all samples, regardless of *Lig1* genotype (Figure 5.4.6). Similarly, levels of *Lig4* expression remained unchanged following targeting of the *Lig1* locus (Figure 5.4.7).

5.4.3 Levels of DNA ligases III and IV are unaffected by the absence of DNA ligase I

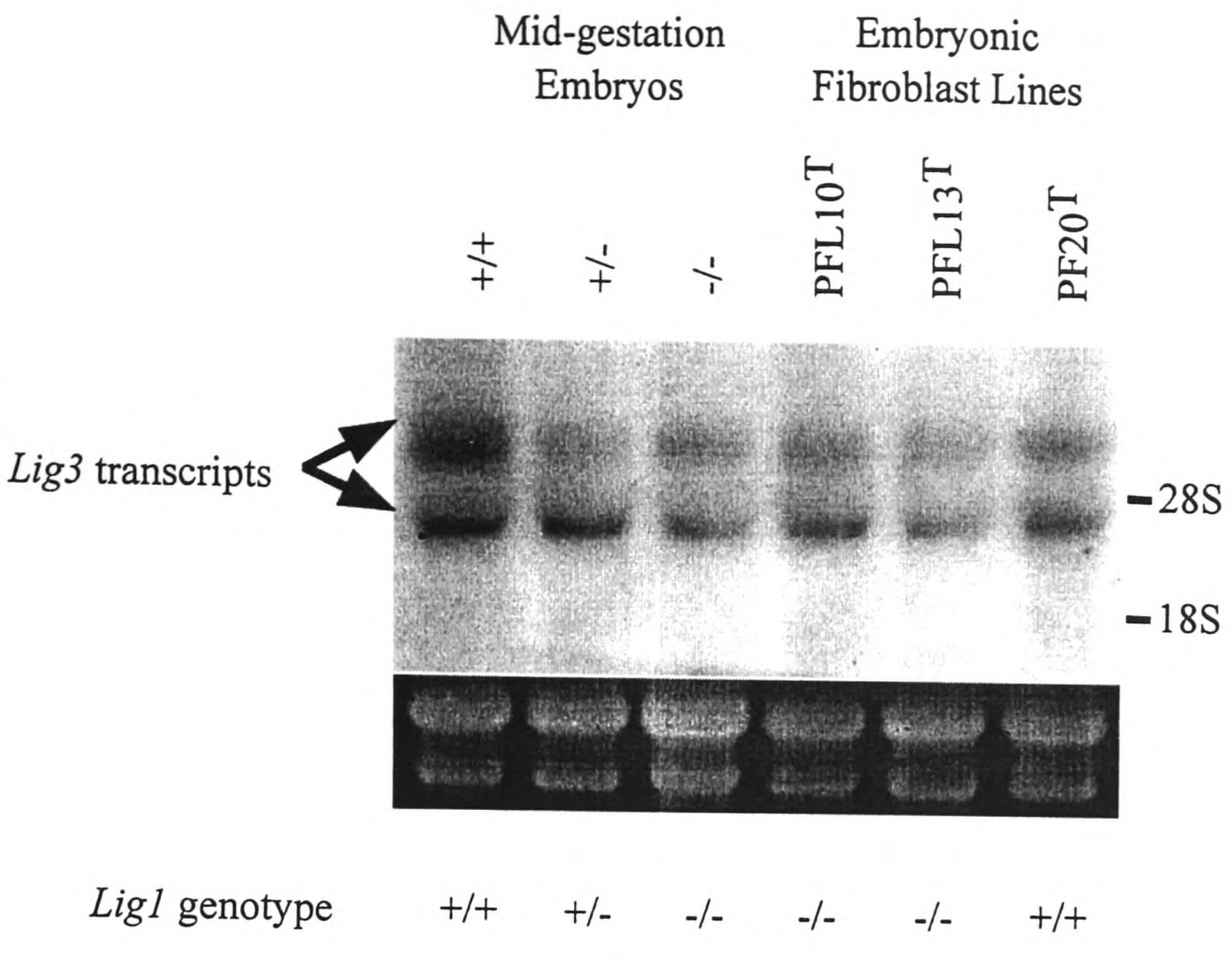
Although the levels of *Lig3* and *Lig4* mRNA appeared unaffected by the absence of DNA ligase I, to confirm that protein levels were similarly unchanged, western blot analysis was performed on protein extracts. Whole cell lysates from a representative panel of embryos and fibroblast cell lines were separated by SDS-PAGE, and blotted onto nylon membrane. Blotted proteins were then incubated with antibodies raised against purified human DNA ligase III (Figure 5.4.8) or with antibodies against DNA ligase IV (Figure 5.4.9). Despite being raised against human proteins, both antisera recognised the mouse homologues, without high levels of crossreactivity to proteins of other sizes. A band of ~100kDa corresponding to DNA ligase III was detected at equal levels in all samples, regardless of *Lig1* genotype. Similarly, a protein of ~100kDa corresponding to DNA ligase IV was detected in each lane at equivalent levels in wild-type, +/- and -/- samples. Thus we concluded that the absence of DNA ligase I did not result in alteration of the levels of either DNA ligase III or DNA ligase IV.

Figure 5.4.6: *Lig3* expression is unaltered by the absence of DNA ligase I

Northern blot analysis of *Lig3* expression in samples representative of the different *Lig1* genotypes. Upper, total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the targeted *Lig1/PGK-HPRT* allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with a 1.0kb *Eco*RI/*Hin*dIII fragment of pMLIG3. Positions of the 28S and 18S ribosomal RNA transcripts are indicated, as are the two *Lig3* transcripts (3.0 and ~6kb). Lower, ethidium bromide staining to verify equal RNA loading.

Figure 5.4.7: *Lig4* expression is unaltered by the absence of DNA ligase I

Northern blot analysis of *Lig4* expression in samples representative of the different *Lig1* genotypes. Upper, total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the targeted *Lig1/PGK-HPRT* allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with a 0.75kb *Eco*RI/*Hin*dIII fragment of pMLIG4. Positions of the 28S and 18S ribosomal RNA transcripts are indicated flanking the ~3.0kb *Lig4* transcript. Lower, ethidium bromide staining to verify equal RNA loading.


Mid-gestation Embryos
Embryonic Fibroblast Lines
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
Lig3 transcripts
28S
18S
Lig1 genotype
+/+ +/- -/- -/- -/- +/+

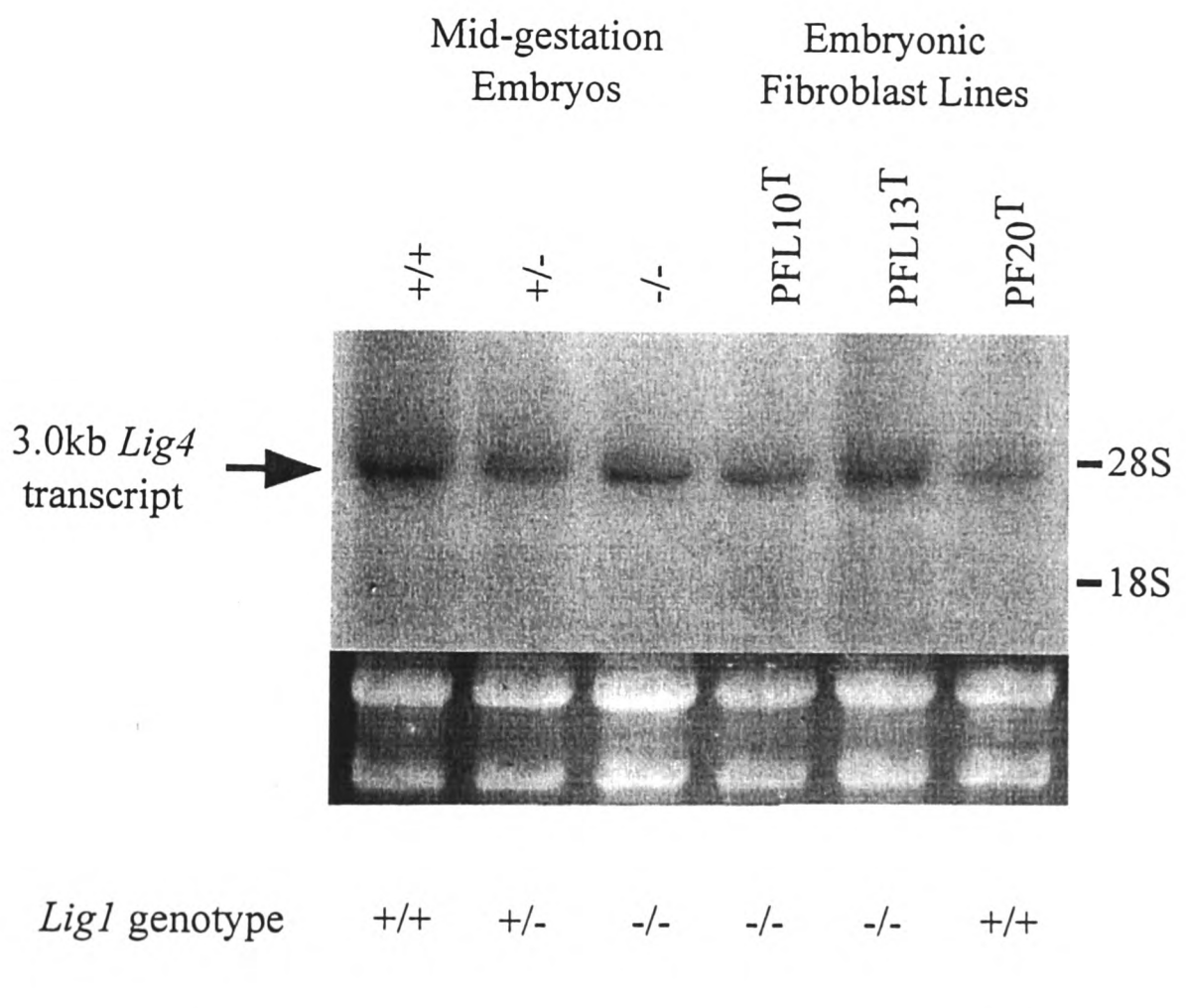
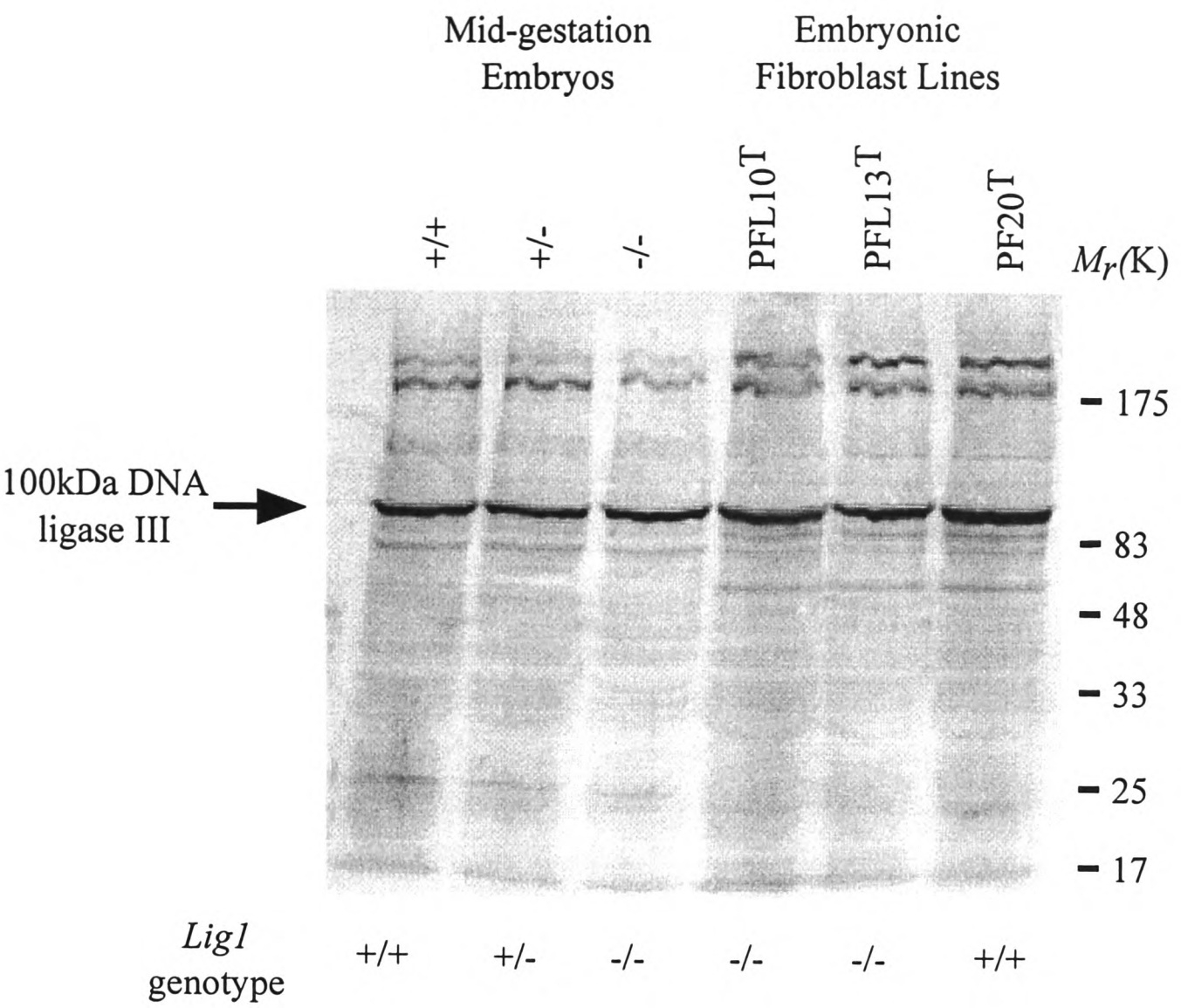

Mid-gestation Embryos
Embryonic Fibroblast Lines
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
3.0kb Lig4 transcript
28S
18S
Lig1 genotype
+/+ +/- -/- -/- -/- +/+

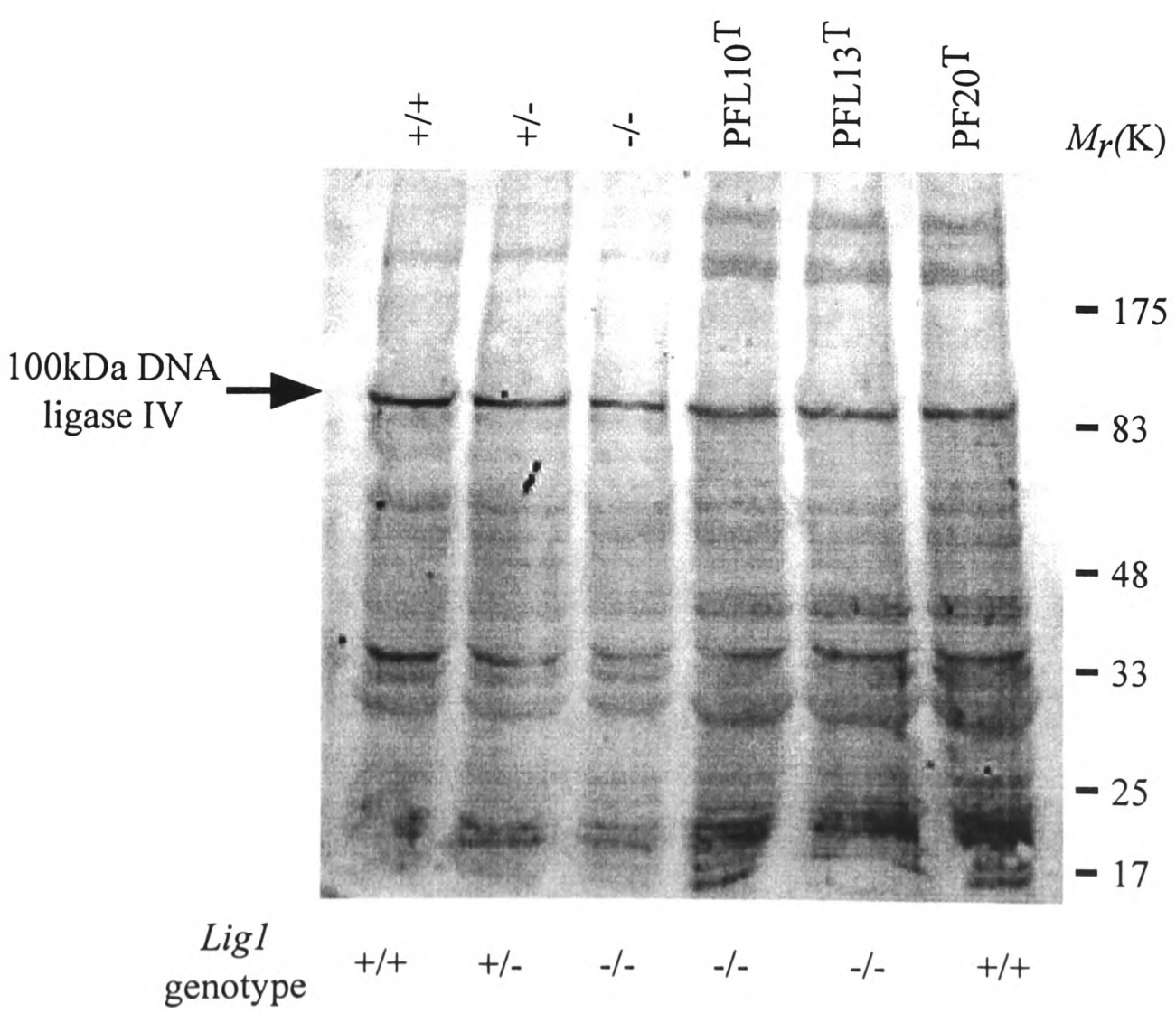
Figure 5.4.8: The level of DNA ligase III is unaltered by the absence of DNA ligase I

Western blot analysis of DNA ligase III in samples representative of the different *Lig1* genotypes. Total protein extracts were prepared from E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). Protein lysates were made by brief homogenisation of tissue or cell pellets in lysis buffer, and 100μg loaded onto an 8% SDS-polyacrylamide gel. Separated proteins were blotted onto a nylon membrane, blocked in TBST plus 5% non-fat milk powder, and incubated with TL25 rabbit polyclonal antibody raised against recombinant full length human DNA ligase III. Detection of bound antibody was achieved using alkaline phosphatase-conjugated anti-rabbit IgG antibodies. Positions of protein molecular weight markers are indicated, as is the ~100kDa wild-type DNA ligase III protein. *Lig1* genotypes of the different samples are shown.

Figure 5.4.9: The level of DNA ligase IV is unaltered by the absence of DNA ligase I

Western blot analysis of DNA ligase IV in samples representative of the different *Lig1* genotypes. Total protein extracts were prepared from E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). Protein lysates were made by brief homogenisation of tissue or cell pellets in lysis buffer, and 100μg loaded onto an 8% SDS-polyacrylamide gel. Separated proteins were blotted onto a nylon membrane, blocked in TBST plus 5% non-fat milk powder, and incubated with TL18 rabbit polyclonal antibody raised against a peptide in the C-terminus of human DNA ligase IV. Detection of bound antibody was achieved using alkaline phosphatase-conjugated anti-rabbit IgG antibodies. Positions of protein molecular weight markers are indicated, as is the ~100kDa wild-type DNA ligase IV protein. *Lig1* genotypes of the different samples are shown.


Mid-gestation Embryos
Embryonic Fibroblast Lines
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
Mr(K)
100kDa DNA ligase III
175
83
48
33
25
17
Lig1 genotype
+/+
+/-
-/-
-/-
-/-
+/+


+/+
+/-
-/-
PFL10T
PFL13T
PF20T
Mr(K)
100kDa DNA ligase IV
175
83
48
33
25
17
Lig1 genotype
+/+
+/-
-/-
-/-
-/-
+/+

5.5 RT-PCR fingerprinting of DNA ligase I-deficient embryos.

mRNA differential display (DDRT-PCR) is an RT-PCR based technique that permits analysis of changes in gene expression between two or more samples (Liang and Pardee 1992). It allows identification and cloning of differentially expressed genes by arbitrary amplification and comparative display of transcripts. mRNA is reverse transcribed using an oligo-dT primer for first strand synthesis and a short arbitrary primer for second strand synthesis. Subsequently repeated PCR cycles are carried out under low stringency conditions with the same primers, and the multiple products separated on a polyacrylamide gel to produce a ladder of bands. Using a combination of different arbitrary primers it is possible to produce a series of patterns representative of all expressed genes. If mRNAs from two different sources are analysed in parallel, it is possible to identify and clone genes which are expressed at different levels. Two variations on this technique have been described: arbitrarily primed PCR (AP-PCR), which employs short arbitrary primers in both first and second strand synthesis, and RT-PCR fingerprinting. RT-PCR fingerprinting (Boehm 1993) was designed to visualise differences in mRNA expression amongst multigene families, and has also been used successfully to identify novel members of known families. Degenerate oligonucleotide primers are designed to recognise regions known to be conserved amongst members of a gene family. mRNA is reverse transcribed using an oligo-dT primer for first strand synthesis, and subsequently repeated PCR cycles are carried out under low stringency conditions with the two specific primers. The PCR products are digested with a frequently cutting restriction enzyme such as *Hae*III, and separated on a polyacrylamide gel to produce a ladder of bands. Individual bands correspond to the different members of the gene family, and it is possible to clone and identify differentially expressed bands as before.

Northern and western analysis of DNA ligases III and IV had shown that expression of these enzymes was not significantly altered in DNA ligase I-deficient samples. To examine the possibility that expression of a further (novel) DNA ligase may alternatively be upregulated, RT-PCR fingerprinting was attempted on RNA samples from +/+ and -/- midgestation embryos, using PCR primers based on the coding regions of *Lig1*, *LIG3*, and *LIG4*.

5.5.1 Low stringency RT-PCR fingerprinting using long sequence specific primers

Total RNA samples from E13.5 *Lig1* +/+ and -/- embryos were reverse transcribed under standard conditions. 3μl of each RT reaction was then used as a template for PCR under low stringency conditions. A small quantity of ^{32}P labelled

dCTP was included in each reaction to allow subsequent visualisation of the PCR products. A number of different combinations of PCR primer pairs were tried, a summary of which is given in Table 5.5.1. All were at least 21 bases in length, and had been initially used either for sequencing of the *Lig1* cDNA (Jessop 1995), genotyping PCR reactions, or to clone fragments of *Lig3* and *Lig4*. The primer pairs employed were predicted to give products between 0.15 and 1.0kb from their respective gene specific templates.

The reaction products were precipitated and then resuspended in 20μl sterile distilled water. Half of the sample was separated on a polyacrylamide gel, whilst the remainder was digested with the restriction enzyme *Hae*III, and then subjected to electrophoresis on a polyacrylamide gel. After drying, the gels were exposed to autoradiography film overnight.

Examination of the autoradiograph revealed the ladder of bands produced to be much more complex than anticipated. Even before *Hae*III digestion, hundreds of distinct bands were visible in each track, and after digestion the complexity of the patterns increased (data not shown). Without full length cDNAs of *Lig3* and *Lig4* as positive controls, it proved impossible to determine which bands were derived from the three human DNA ligases. This unexpected complexity most probably arose as a result of the nature of the primers used. None of the primers had been designed for the purpose of RT-PCR fingerprinting. All were long sequence specific oligonucleotides not well suited for use as degenerate primers. Also, the predicted products amplified by many pairs were too large to be useful for this technique. Those few pairs that gave reasonably small products were invariably based on the *Lig1* cDNA sequence. Because of the relative abundance of *Lig1* expression in comparison to transcripts of the other two DNA ligases, these pairs exclusively amplified *Lig1* sequences. Hence, the attempt to carry out RT-PCR fingerprinting of the DNA ligase gene family was not wholly successful.

It was however noted that a few bands from certain primer pairs were of different intensity in different lanes. In most cases this manifested as the presence of a band in the +/+ sample which was absent from the -/- sample, and therefore attributable to targeted disruption of *Lig1*. However, some bands exhibited the reciprocal pattern, or were more intense in the -/- track. It was decided to investigate the origin of these fragments further.

Table 5.5.1: Summary of the primer pairs used for RT-PCR fingerprinting

Primer A		Primer B		Predicted product size (bp)	Number of Differential bands
Name	Gene	Name	Gene		
G4608	*Lig1*	H3123	*Lig1*	380	0
G4608	*Lig1*	H2066	*Lig1*	600	2
G4607	*Lig1*	G8320	*Lig1*	840	0
N2469	*Lig1*	M5176	*Lig1*	230	2
N2469	*Lig1*	N2468	*Lig1*	290	0
M5175	*Lig1*	N2468	*Lig1*	210	5
M5175	*Lig1*	M5176	*Lig1*	150	9
HLIG35'	*LIG3*	HLIG33'	*LIG3*	1000	0
HLIG45'	*LIG4*	HLIG43'	*LIG4*	1000	0
LIG45'	*LIG4*	LIG43'	*LIG4*	750	1
HLIG45'	*LIG4*	LIG43'	*LIG4*	950	1
LIG45'	*LIG4*	HLIG43'	*LIG4*	800	3
HLIG35'	*LIG3*	HLIG43'	*LIG4*	1000	0
HLIG35'	*LIG3*	LIG43'	*LIG4*	950	1
HLIG45'	*LIG4*	HLIG33'	*LIG3*	1000	1
LIG45'	*LIG4*	HLIG33'	*LIG3*	800	0

The different combinations of oligonucleotide primers tried for RT-PCR fingerprinting. Primer identifiers are given, along with an indication of the gene to which they were originally designed to bind. The predicted product sizes are based on binding locations within the respective cDNA sequences. The number of bands of greater intensity in -/- tracks compared to wild-type samples is also recorded. See text for details.

Table 5.5.2: Summary of results from low stringency RT-PCR using long sequence specific primers

The 25 bands differing in intensity between +/+ and -/- samples in RT-PCR fingerprinting were used to probe northern blots of total RNA from midgestation embryos. Approximate sizes of the purified fragments are recorded, along with the estimations of the sizes of bands on the northern blots, and the origins of the signals detected. Asterisks mark those signals more highly expressed in -/- samples. N.D., origins not determined.

Band	Primer Combination	Product Size (bp)	Signal(s) Detected by Northern Analysis	Identity
1	G4608, H2066	<200	None	-
2	G4608, H2066	200	1kb, 3kb, >15kb*	rRNA, DNA
3	N2469, M5176	<200	None	-
4	N2469, M5176	300	1kb, >15kb*	rRNA, DNA
5	LIG45', LIG43'	600	3kb, >15kb*	rRNA, DNA
6	HLIG45', LIG43'	300	>15kb*	DNA
7	LIG45', HLIG43'	<200	0.3kb	N.D.
8	LIG45', HLIG43'	250	0.5kb*, 2kb	MRP14
9	LIG45', HLIG43'	600	>15kb*	DNA
10	HLIG35', LIG43'	350	3kb, 6kb*	Unknown
11	HLIG45', HLIG33'	<200	>15kb*	DNA
12	M5175, M5176	<200	None	-
13	M5175, M5176	200	>15kb*	DNA
14	M5175, M5176	300	>15kb*	DNA
15	M5175, M5176	350	2kb, >15kb*	N.D., DNA
16	M5175, M5176	400	>15kb*	DNA
17	M5175, M5176	600	>15kb*	DNA
18	M5175, M5176	750	3kb, >15kb*	N.D., DNA
19	M5175, M5176	1000	3kb, >15kb*	N.D., DNA
20	M5175, M5176	1000	3kb, >15kb*	N.D., DNA
21	M5175, N2468	<200	None	-
22	M5175, N2468	<200	1kb	N.D.
23	M5175, N2468	450	2kb*, 5kb, >15kb*	transferrin, N.D., DNA
24	M5175, N2468	600	None	-
25	M5175, N2468	1000	>15kb*	DNA

Undigested banding patterns were examined to identify those bands differentially expressed between the +/+ and -/- samples. Once differential bands were identified from dried gels, repeat PCR reactions were carried out, and the reactions separated as before, except that the gel was exposed to autoradiography film without being dried. Areas of the gel corresponding to bands of interest were excised, and the DNA fragments contained therein purified by the 'crush and soak' method. 1μl of each purified fragment was then used to seed a further round of PCR in order to amplify the recovered DNA fragments. The primers employed were those that had been used in the initial round of low stringency PCR. For this amplification step higher stringency conditions were used.

Reamplified PCR reaction products were run out on a low melting point agarose gel, and the visible band (or bands) purified. Each recovered fragment was then used as a probe in a secondary screening procedure based on northern blot analysis. Total RNA samples from pairs of +/+ and -/- embryos were separated by formaldehyde-agarose gel electrophoresis, blotted onto nylon membrane, and hybridised to probes produced from the purified fragments. Autoradiography was used to identify those which corresponded to transcripts upregulated in -/- embryos. Results of this procedure, best described as 'low stringency RT-PCR using long, sequence specific primers', are summarised in Table 5.5.2

Of 25 fragments initially investigated, 5 did not give signals on northern blots, although in 4 of these cases the bands were under 0.2kb and the negative result probably stemmed from technical difficulties. 16 of the remaining 20 fragments picked up a very high molecular weight band (>15kb), either with or without hybridisation to smaller transcripts. This band was of greater intensity in the -/- sample, but its large size indicated the most likely source to be DNA contamination in the RNA preparations, and hence was ignored. Three fragments hybridised to ribosomal RNA, and nine detected transcripts that did not appear to be differentially expressed. However, three products, numbered #8, #10 and #23, did pick up transcripts which were more highly expressed in -/- RNA samples compared to wild-type. The sizes of the three transcripts were estimated as 0.5kb, 6kb and 2kb respectively.

Because none of the three fragments of interest detected only a single band, each probably consisted of more than one species of DNA molecule. Therefore before any further analysis could be conducted, it was necessary to isolate the individual fragments. To this end, the purified products from the second round of PCR were cloned into the *Eco*RV site of pBluescriptII SK+. In order to identify those plasmids containing fragments corresponding to differentially expressed transcripts, each clone was re-screened by northern analysis. The cloned sequence was excised

from the plasmid vector by restriction digestion, and used to probe pairs of +/+ and -/- embryo RNA samples by northern blotting as before. Subsequently, clones which detected differentially expressed transcripts were sequenced, and these sequences then used to search nucleotide sequence databases.

5.5.2 Analysis of the three differentially expressed transcripts

5.5.2.1 0.5kb transcript (fragment #8)

Expression of the 0.5kb transcript was approximately 3-fold higher in -/- embryos compared to wild-type and heterozygous samples (Figure 5.5.2). No expression could be detected in ES cell or fibroblast samples, irrespective of *Lig1* genotype. The PCR fragment was produced from a reaction utilising two primers based on the *LIG4* cDNA sequence (LIG45' and HLIG43'). When cloned and sequenced, the fragment proved to be 225bp in length, of which 176bp was non-primer derived (Figure 5.5.1). Analysis of this sequence revealed that it had 100% identity with a region of the murine intracellular calcium binding protein MRP14 mRNA (Database accession number: M83219)[iii]. MRP14 transcripts are 0.5kb in size (Lagasse and Weissman 1992), closely matching the transcript observed on northern blots probed with the cloned PCR fragment. MRP14 is a member of the S100 family of proteins, most of which are involved in cellular differentiation or cell cycle progression. The protein complexes with another calcium binding protein MRP8 on the inside face of the cell membrane and on intermediate filaments. The precise function of MRP14 is still unclear, but it may be involved in intracellular signalling by modulating protein kinase function. However, irrespective of function, the protein is known to be expressed predominantly in granulocytes and monocytes of the haematopoietic system, and is linked to development of the myeloid lineage and activation during the inflammation response (Murao 1994).

5.5.2.2 2kb transcript (fragment #23)

Expression of the 2kb transcript was approximately 2-fold higher in -/- embryos compared to wild-type and heterozygous samples (Figure 5.5.4). No expression could be detected in samples from either ES cells or fibroblast lines, irrespective of *Lig1* genotype. The fragment was produced from a PCR reaction utilising two primers based on the mouse *Lig1* cDNA sequence (M5175 and N2468). When cloned and sequenced, the fragment was found to be 290bp, of which 246bp was non-primer derived (Figure 5.5.3). Analysis of this sequence revealed that it had

[iii] Also called Calgranulin B or *Cagb*

Figure 5.5.1: Sequence of fragment #8 aligned with the published mouse *MRP14* cDNA sequence.

The nucleotide sequence of fragment #8 was used to search the Genbank nucleotide sequence database through the Fasta program of the Genetics Computer Group package version 9 (Genetics Computer Group 1996). A matching sequence, corresponding to the murine intracellular calcium binding protein MRP14 cDNA, was retrieved. Numbers indicate the location of nucleotides within the published *MRP14* cDNA sequence (Database accession number: M83219) (top), and the sequence of clone p8Q (bottom). Vertical lines indicate identical nucleotides in the two sequences, and PCR primers sequences are marked.

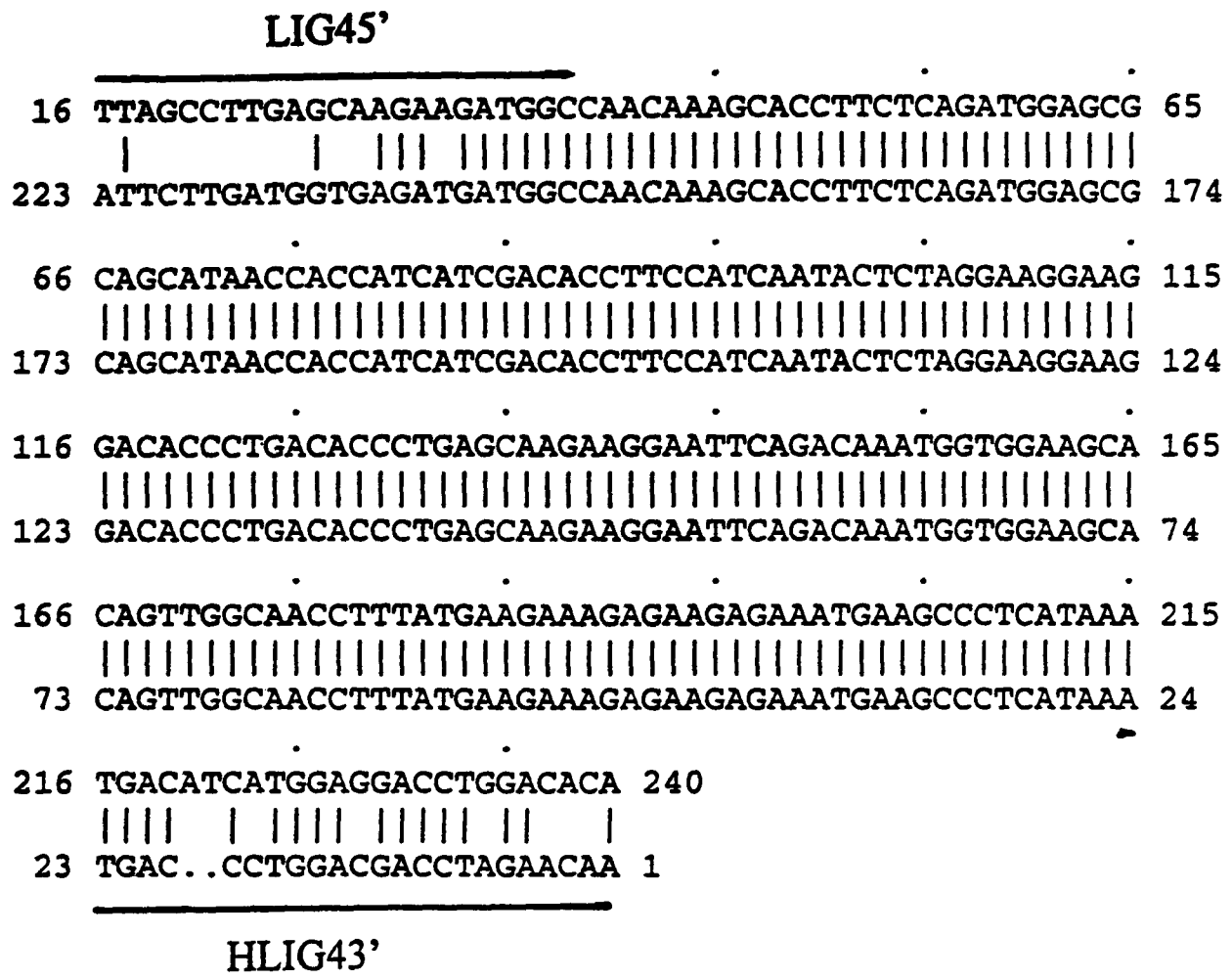

LIG45'
16 TTAGCCTTGAGCAAGAAGATGGCCAACAAAGCACCTTCTCAGATGGAGCG 65
223 ATTCTTGATGGTGAGATGATGGCCAACAAAGCACCTTCTCAGATGGAGCG 174
66 CAGCATAACCACCATCATCGACACCTTCCATCAATACTCTAGGAAGGAAG 115
173 CAGCATAACCACCATCATCGACACCTTCCATCAATACTCTAGGAAGGAAG 124
116 GACACCCTGACACCCTGAGCAAGAAGGAATTCAGACAAATGGTGGAAGCA 165
123 GACACCCTGACACCCTGAGCAAGAAGGAATTCAGACAAATGGTGGAAGCA 74
166 CAGTTGGCAACCTTTATGAAGAAAGAGAAGAGAAATGAAGCCCTCATAAA 215
73 CAGTTGGCAACCTTTATGAAGAAAGAGAAGAGAAATGAAGCCCTCATAAA 24
216 TGACATCATGGAGGACCTGGACACA 240
23 TGAC..CCTGGACGACCTAGAACAA 1
HLIG43'

Figure 5.5.2: Expression of *MRP14* is upregulated in -/- embryos

Northern blot analysis of *MRP14* expression in samples representative of the different *Lig1* genotypes. Upper, total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the *Lig1/PGK-HPRT* targeted allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with an *Eco*RI/*Hin*dIII fragment of p8Q. Positions of the 28S and 18S ribosomal RNA transcripts are indicated. Lower, ethidium bromide staining to verify equal RNA loading.

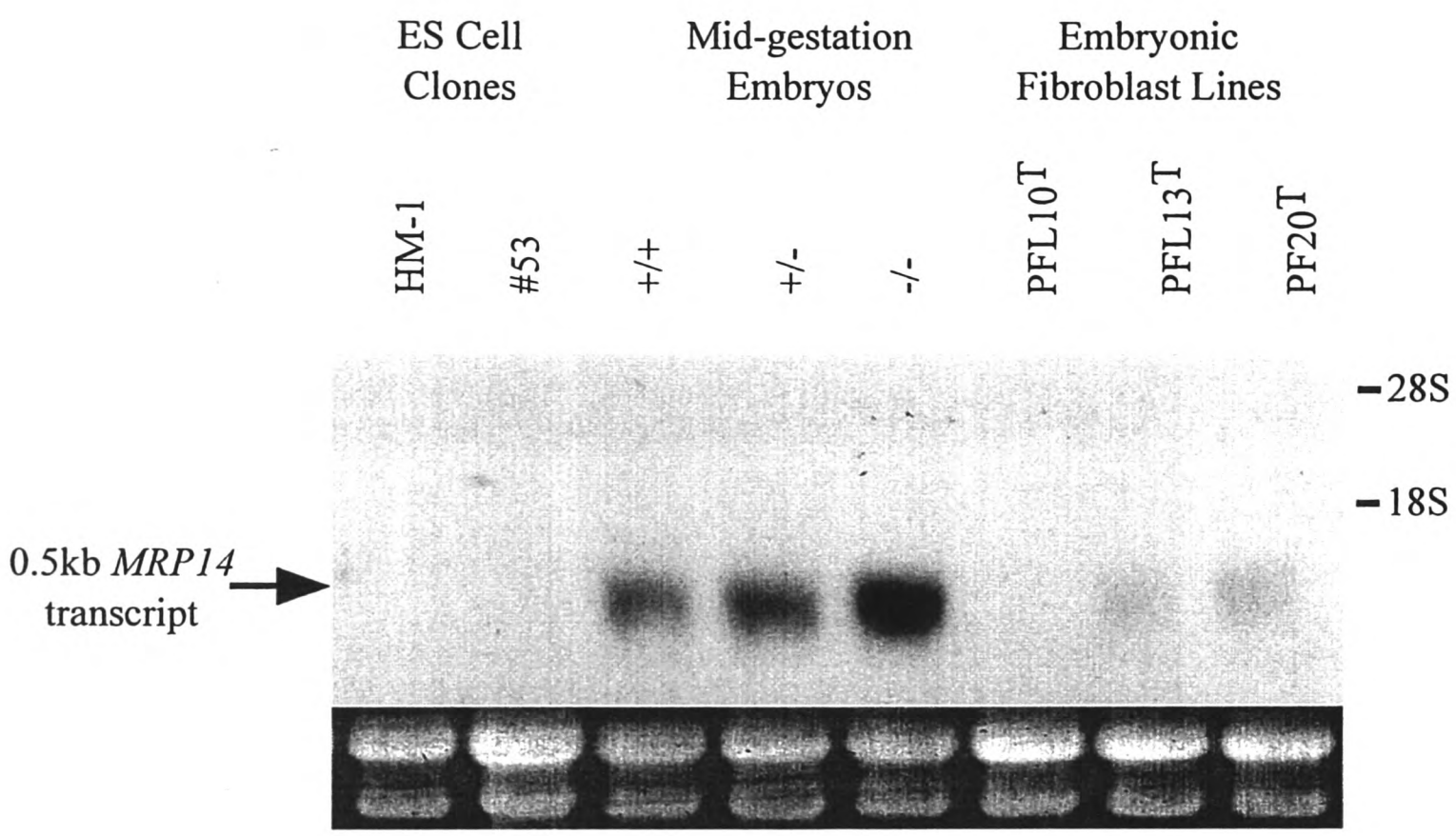

ES Cell Clones
Mid-gestation Embryos
Embryonic Fibroblast Lines
HM-1
#53
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
28S
18S
0.5kb MRP14 transcript
Lig1 genotype
+/+ +/- +/+ +/- -/- -/- -/- +/+

Figure 5.5.3: Sequence of fragment #23 aligned with the published mouse transferrin cDNA sequence.

The nucleotide sequence of fragment #23 was used to search the Genbank nucleotide sequence database through the Fasta program of the Genetics Computer Group package, version 9 (Genetics Computer Group 1996). A matching sequence, corresponding to the rat transferrin cDNA, was retrieved. Numbers indicate the location of nucleotides within the published transferrin cDNA sequence (Database accession number: D38380) (top), and the sequence of clone p23G (bottom). Vertical lines indicate identical nucleotides in the two sequences, and PCR primers sequences are marked.

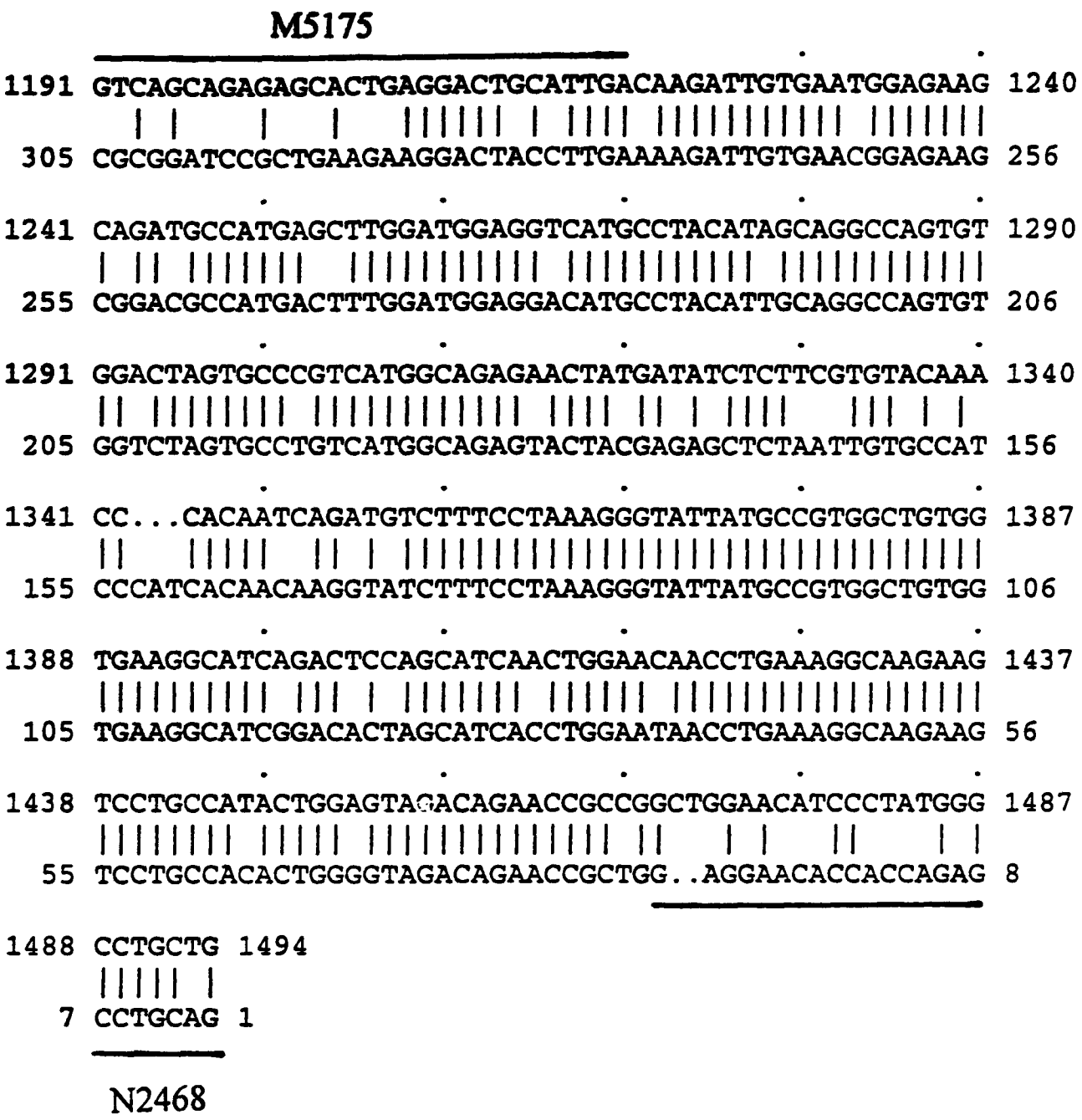
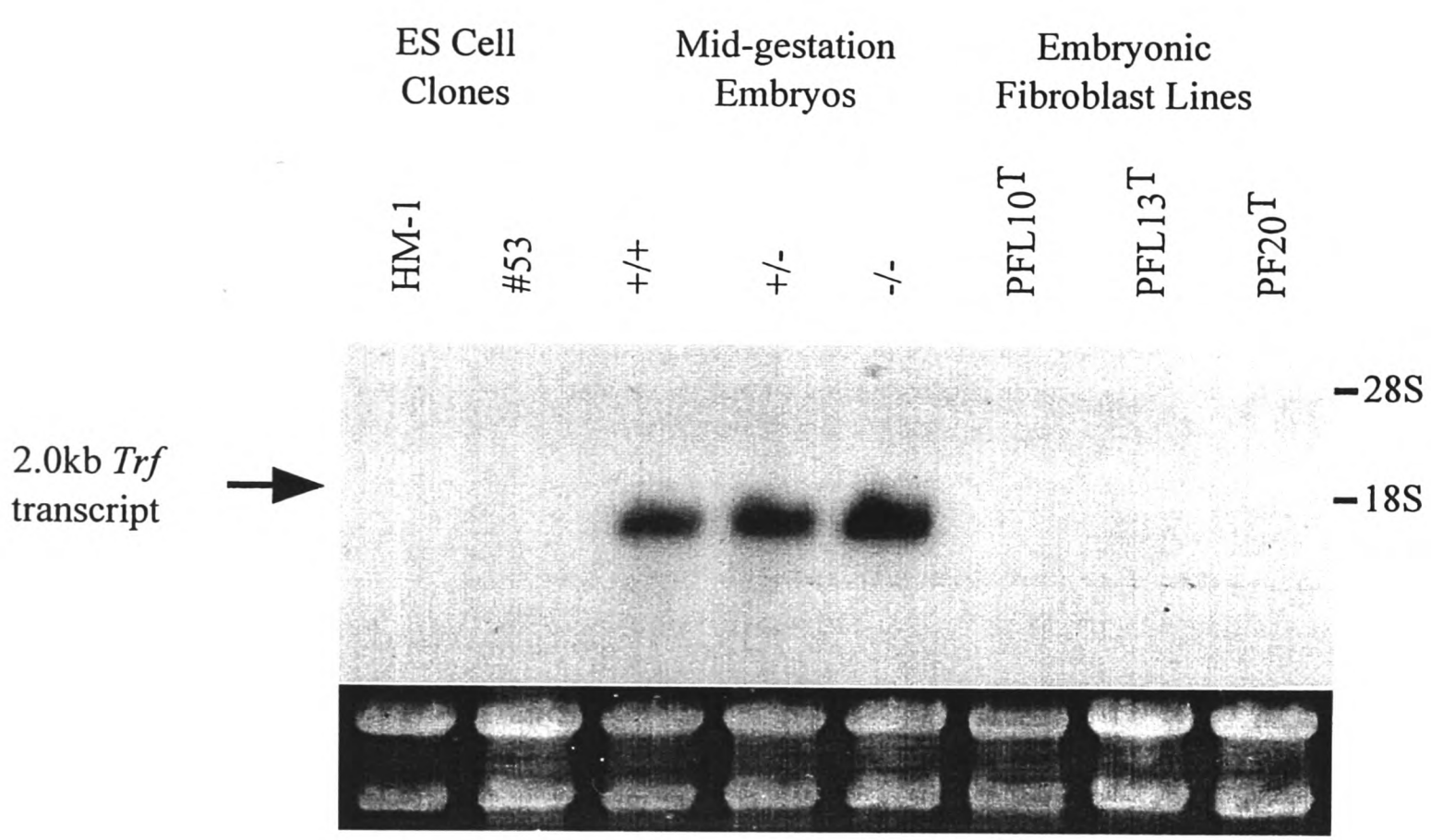

M5175
1191 GTCAGCAGAGAGCACTGAGGACTGCATTGACAAGATTGTGAATGGAGAAG 1240
305 CGCGGATCCGCTGAAGAAGGACTACCTTGAAAAGATTGTGAACGGAGAAG 256
1241 CAGATGCCATGAGCTTGGATGGAGGTCATGCCTACATAGCAGGCCAGTGT 1290
255 CGGACGCCATGACTTTGGATGGAGGACATGCCTACATTGCAGGCCAGTGT 206
1291 GGACTAGTGCCCGTCATGGCAGAGAACTATGATATCTCTTCGTGTACAAA 1340
205 GGTCTAGTGCCTGTCATGGCAGAGTACTACGAGAGCTCTAATTGTGCCAT 156
1341 CC...CACAATCAGATGTCTTTCCTAAAGGGTATTATGCCGTGGCTGTGG 1387
155 CCCATCACAACAAGGTATCTTTCCTAAAGGGTATTATGCCGTGGCTGTGG 106
1388 TGAAGGCATCAGACTCCAGCATCAACTGGAACAACCTGAAAGGCAAGAAG 1437
105 TGAAGGCATCGGACACTAGCATCACCTGGAATAACCTGAAAGGCAAGAAG 56
1438 TCCTGCCATACTGGAGTAGACAGAACCGCCGGCTGGAACATCCCTATGGG 1487
55 TCCTGCCACACTGGGGTAGACAGAACCGCTGG..AGGAACACCACCAGAG 8
1488 CCTGCTG 1494
7 CCTGCAG 1
N2468

Figure 5.5.4: Expression of transferrin is upregulated in -/- embryos

Northern blot analysis of *Trf* expression in samples representative of the different *Lig1* genotypes. Upper, total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the *Lig1/PGK-HPRT* targeted allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele ($PFL10^T$, $PFL13^T$), and a wild-type fibroblast line ($PF20^T$). RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with an *Eco*RI/*Hin*dIII fragment of p23G. Positions of the 28S and 18S ribosomal RNA transcripts are indicated. Lower, ethidium bromide staining to verify equal RNA loading.


ES Cell Clones
Mid-gestation Embryos
Embryonic Fibroblast Lines
HM-1
#53
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
28S
18S
2.0kb Trf transcript
Lig1 genotype
+/+ +/- +/+ +/- -/- -/- -/- +/+

100% identity with a stretch of sequence from a partial mouse transferrin cDNA (*Trf*), and 86%identity at the nucleotide level with a region of the rat transferrin mRNA (Database accession number: D38380). The rat *Trf* transcript is 2.3kb in size (Hosino, Hisayasu,and Shimada 1996), again matching the size of transcript observed on northern blots probed with the cloned PCR fragment. Transferrin is an iron transport protein whose expression is known to be stimulated by iron deficiency (Scriver 1995)).

5.5.2.3 6kb transcript (fragment #10)

Expression of the 6kb transcript was approximately 5-fold higher in -/- embryos compared to wild-type and heterozygous samples (Figure 5.5.5). Although no expression could be detected in ES cells or the wild-type fibroblast line $PF20^T$, a significant level of expression was observed in one of the two DNA ligase I-deficient fibroblast lines ($PFL10^T$). The PCR fragment was produced from a reaction employing one primer based on the *LIG3* cDNA sequence (HLIG35') and one primer derived from the *LIG4* cDNA sequence (LIG43'). Two independent clones were obtained (p10E and p10X), each containing around 800bp of non-vector sequence. However, when analysed, the two inserts were not identical. Homology only extended over half of the length of each clone (~400bp) (data not shown). Figure 5.5.6 indicates the way in which the two sequences were apparently related. Each clone was a doublet of two unrelated 400bp sequences, but with only one of each doublet common to both clones. Both clones detected single transcripts of the same size on northern blots (data not shown). Thus it was assumed that the 400bp sequence common to both clones was the element which hybridised to the 6kb transcript, and hence the region of interest. The consensus sequence of this element is shown in Figure 5.5.7.

Nucleotide database searches with all of the sequences contained within clones p10E and p10X failed to reveal any significant homology to existing sequences. Blast searches for potential homology amongst translated amino acid sequences also failed to pick up any good candidate sequences. The closest match from these searches was to the P-type cation translocating ATPase from *Plasmodium falciparum*, but even in this case homology was limited (data not shown). It was therefore concluded that the fragment which hybridised to the ~6kb transcript was part of a transcript from a gene not in the sequence database.

Figure 5.5.5: Expression of the transcript corresponding to fragment #10 is upregulated in -/- embryos and PFL10^T fibroblasts

Northern blot analysis of "fragment #10" expression in samples representative of the different *Lig1* genotypes. Upper, total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the *Lig1/PGK-HPRT* targeted allele (#53), E13.5 embryos of the three *Lig1* genotypes, the two immortalised fibroblast lines homozygous for the targeted *Lig1/PGK-HPRT* allele (PFL10^T, PFL13^T), and a wild-type fibroblast line (PF20^T). RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with an *Eco*RI/*Hin*dIII fragment of p10E. Positions of the 28S and 18S ribosomal RNA transcripts are indicated. Lower, ethidium bromide staining to verify equal RNA loading.

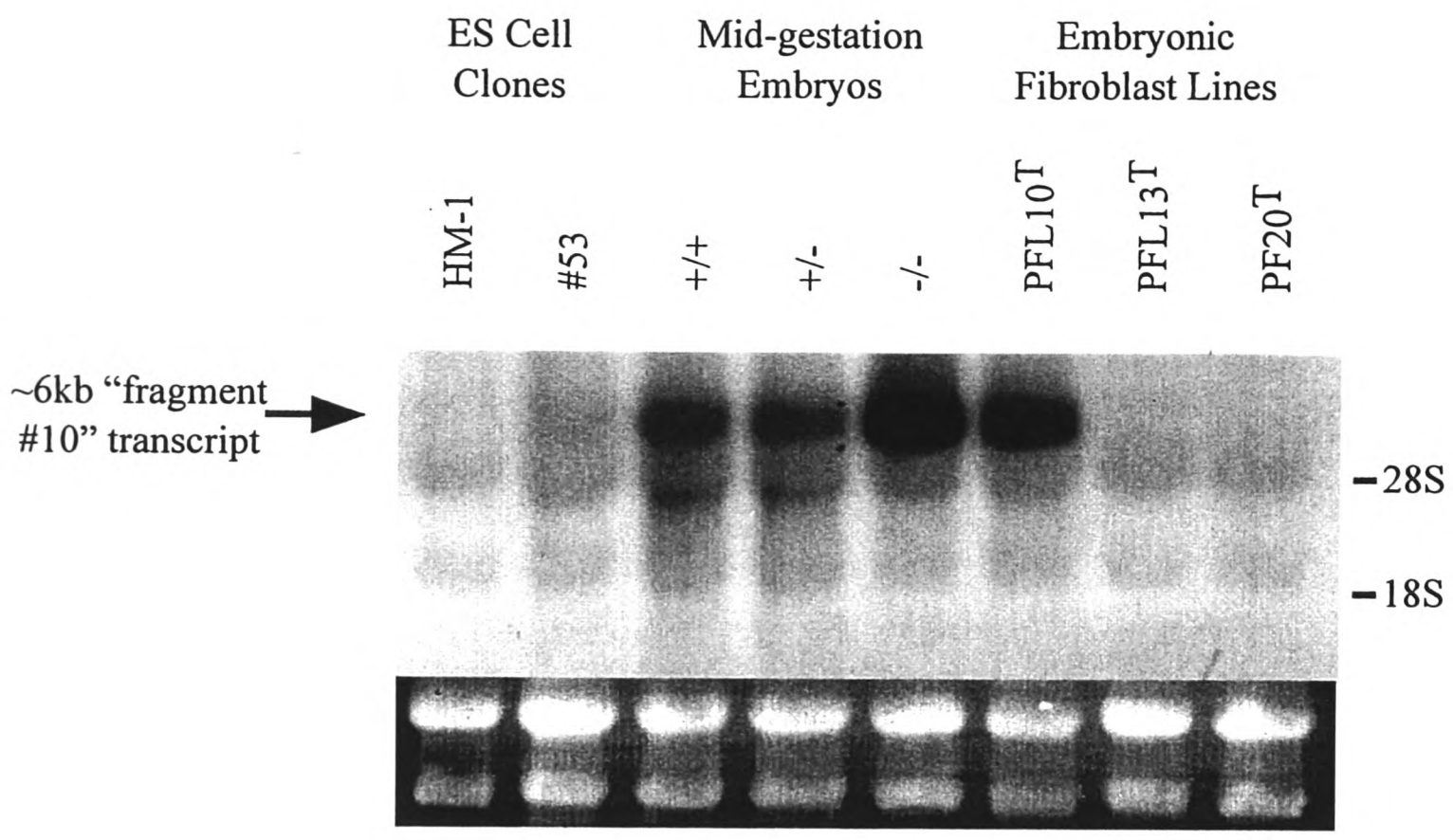
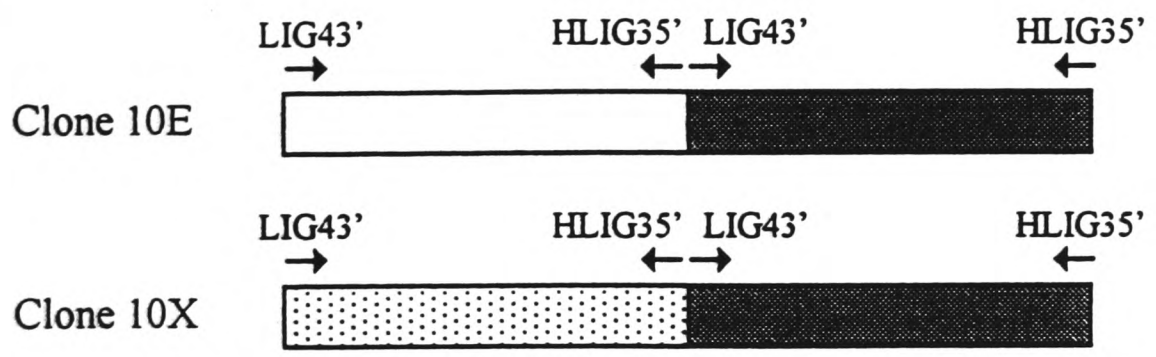

ES Cell Clones
Mid-gestation Embryos
Embryonic Fibroblast Lines
HM-1
#53
+/+
+/-
-/-
PFL10T
PFL13T
PF20T
~6kb "fragment #10" transcript
28S
18S
Lig1 genotype
+/+ +/- +/+ +/- -/- -/- -/- +/+

Figure 5.5.6: Schematic representation of the relationship between clones 10E and 10X.

Nucleotide sequences of two independent clones of fragment #10 were aligned using the Lineup and Pileup programs on the Genetics Computer Group package, version 9 (Genetics Computer Group 1996), and the resulting alignment represented schematically. Sequences common between the two clones are depicted by filled boxes, other sequences by open and shaded boxes. Arrows indicate PCR primer sequences.

Figure 5.5.7: Consensus sequence of fragment 10.

Nucleotide sequences of clones 10E and 10X were aligned using Lineup and Pileup programs of the Genetics Computer Group package, version 9 (Genetics Computer Group 1996), and the consensus sequence of the common region produced using the Pretty program. PCR primer sequences are indicated.



HLIG35'

```
  1  ATGAAGAAAT GTCCCAATGG CATGTTTGCT GTGACTTTGC AGGACTCAGT
 51  AAGCCAGTAG ATAAACATTA TTGTGTGATC TATGATTAGT AGTCTAACCA
101  TGACACAAGG CATTTGTATG CACTTCTGCT GGAGATGTGA TTCATTCTTG
151  GTTTACAAAT TCAAATTGTT CCTTCATTTT TGCTTTTTCC AAGAGTGAAA
201  CATCAACAAA TATTCATTAT TATACTAGTC CCCTTGCAGA TAATGTTATG
251  CCCTTGAGAT AGGTACCCAA TGCAAATTAA ATGAATGTAT AAAATTGGCT
301  GCACCTTG
```

LIG43'

5.6 Cells lacking DNA ligase I exhibited delayed joining of newly synthesised DNA replication intermediates into high molecular weight forms

Biochemical and genetic evidence has indicated DNA ligase I to be central to mammalian DNA replication. DNA ligase I expression and activity is induced upon cell proliferation. The protein has been isolated from mammalian replication complexes, and functions in the reconstitution of replication in vitro. Moreover, cells from the 46BR patient, with reduced DNA ligase I activity, exhibited defective joining of replication intermediates (Reviewed in Tomkinson and Levin 1997). Short nascent DNA chains (40-300 bases), termed Okazaki fragments, are formed during discontinuous 'lagging strand' DNA synthesis prior to incorporation into higher molecular weight strands. These replication intermediates are extremely transient in wild-type cells incubated at 37°C, and cannot be detected with pulse labelling periods over 1 minute. However, Okazaki-type fragments were detected in cells from the 46BR patient, and these disappeared only after long chase times (Henderson *et al.* 1985, Lönn *et al.* 1989, Prigent *et al.* 1994). We were able to reproduce these observations using incorporation of radiolabelled deoxyribonucleotides to specifically label newly replicated DNA.

Exponentially proliferating cultures of immortalised fibroblast lines were pulse labelled with [^{3}H]-thymidine for 10 minutes. Alternatively cells were pulsed with ^{3}H thymidine for 10 minutes followed by 2, 5, 10, 15, or 30 minutes incubation with an excess of unlabelled thymidine (chase). Cells were then harvested, embedded in agarose plugs, and lysed *in situ*, prior to incubation in alkaline conditions to denature the contained DNA. Thereupon single stranded DNA fragments were separated on the basis of size by electrophoresis through alkaline agarose gels. To detect the presence of ^{3}H labelled DNA, gels were cut into defined slices which were assayed by scintillation. The numbers of radioactive counts detected in samples corresponding to the different ranges of DNA fragment size were expressed as a percentage of the total counts incorporated into DNA at that timepoint. Figure 5.6.1 shows the percentage counts in the different fractions plotted against the chase time for four human and mouse fibroblast cell lines, while Figure 5.6.2 shows the cumulative percentage of counts in the four smallest fractions (DNA fragments under 2.0kb size) plotted against chase time.

Following a 10 minute pulse, the majority of ^{3}H radiolabel in wild-type human fibroblast cells (MRC5V1) was incorporated into DNA molecules over 23kb in size (Figure 5.6.1a). Very little labelled thymidine was present in fragments under 2kb (Figure 5.6.2), and that which was present rapidly disappeared when chased with unlabelled thymidine. In contrast, labelled DNA from 46BR.1G1 cells was generally

of a lower molecular weight, with less than 50% of the radiolabel in the heaviest fraction (Figure 5.6.1b). Moreover, significant amounts of label could be detected in DNA fragments of less than 2kb size (Figure 5.6.2). Thus, as reported previously there was a delay in incorporation of replication intermediates into larger molecules in 46BR.1G1 cells. However, this delay was transient. After prolonged chase periods, very little ^{3}H label remained in DNA fragments under 2kb, and the majority of radiolabelled DNA was at least 23kb in size.

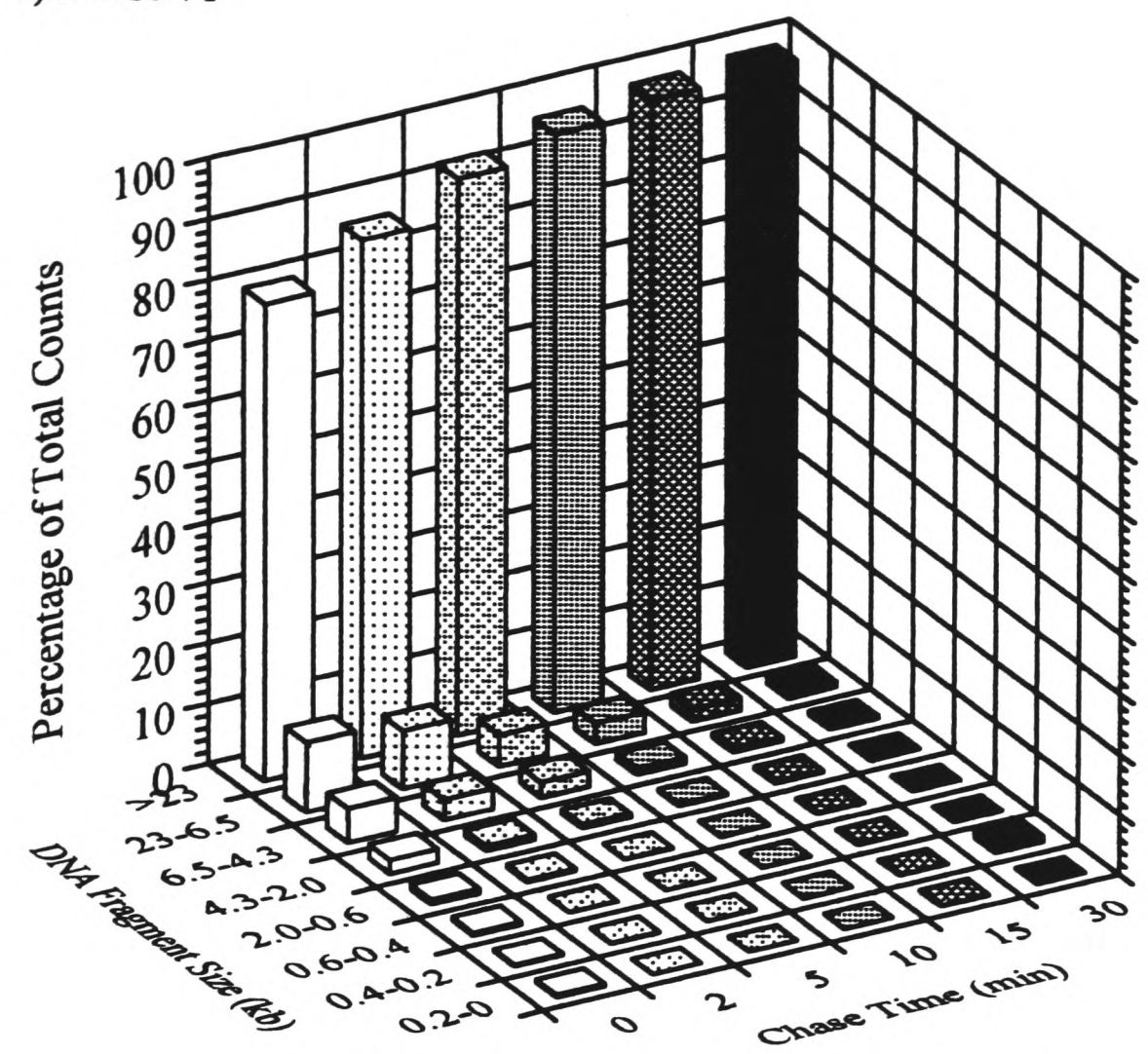
A wild-type mouse immortalised line ($PF20^{T}$) behaved in a manner very similar to its human counterpart in this assay. The majority of ^{3}H radiolabel was incorporated into large DNA molecules after a 10 minute pulse, and only small amounts were present in smaller fractions (Figure 5.6.2). The fibroblast line lacking DNA ligase I ($PFL13^{T}$) did not however display a normal distribution of nascent DNA molecules. Labelled DNA was of lower average molecular weight compared to the wild-type control, with less than 50% of the label in fractions over 23kb size (Figure 5.6.1d). Approximately 10% of ^{3}H label was detected in DNA fragments under 2kb in size after a 10 minute pulse (Figure 5.6.2), indicating a delay in the joining of replication intermediates. Again this delay proved to be only transient, as after chase periods greater than 10 minutes, all of the label was to be found in molecules over 2kb.

Although $PFL13^{T}$ fibroblasts displayed abnormal joining of replication intermediates, the profile observed was distinct from that seen in 46BR.1G1 cells. After a radioactive pulse, the average size of nascent DNA molecules was greater in $PFL13^{T}$ cells, with less radiolabel being located in small DNA fragments compared to 46BR.1G1 cells. However, these small intermediates were chased away at broadly similar rates in both 46BR.1G1 and $PFL13^{T}$ cell lines (Figure 5.6.2), and the rates of disappearance of small DNA fragments were similar across all size fractions (Figure 5.6.1b and d, and data not shown). Thus, whilst the initial rate of Okazaki-type intermediate joining in $PFL13^{T}$ cells was apparently closer to normal parameters than in 46BR.1G1 cells, subsequent joining (or degradation) of these fragments apparently occurred at equivalent rates in both cell lines.

Figure 5.6.1: Analysis of newly replicated DNA

Graphical representations of the size distribution of DNA fragments formed during DNA replication. MRC5V1 (a), 46BR.1G1 (b), $PF20^T$ (c) and $PFL13^T$ (d) cells were pulse labelled for 10 minutes with [^{3}H]-thymidine. Alternatively cells were pulsed with [^{3}H]-thymidine for 10 minutes followed by 2, 5, 10, 15, or 30 minutes chase with an excess of unlabelled thymidine. Cells were harvested, embedded in agarose plugs, and lysed *in situ*, prior to incubation in alkaline conditions. Thereupon single stranded DNA fragments were separated on the basis of size by electrophoresis through alkaline agarose gels. To detect the presence of ^{3}H labelled DNA, gels were cut into defined slices which were assayed by scintillation. The numbers of radioactive counts detected in samples corresponding to the different ranges of DNA fragment size were expressed as a percentage of the total counts incorporated into DNA at that timepoint. These percentages were then plotted on a 3 dimensional bar chart against DNA fragment size and chase time. Each point represents the mean value from two independent experiments (error bars not shown for clarity). Increasing lengths of chase times are indicated by heavier shading.

a) MRC5V1



b) 46BR.1G1

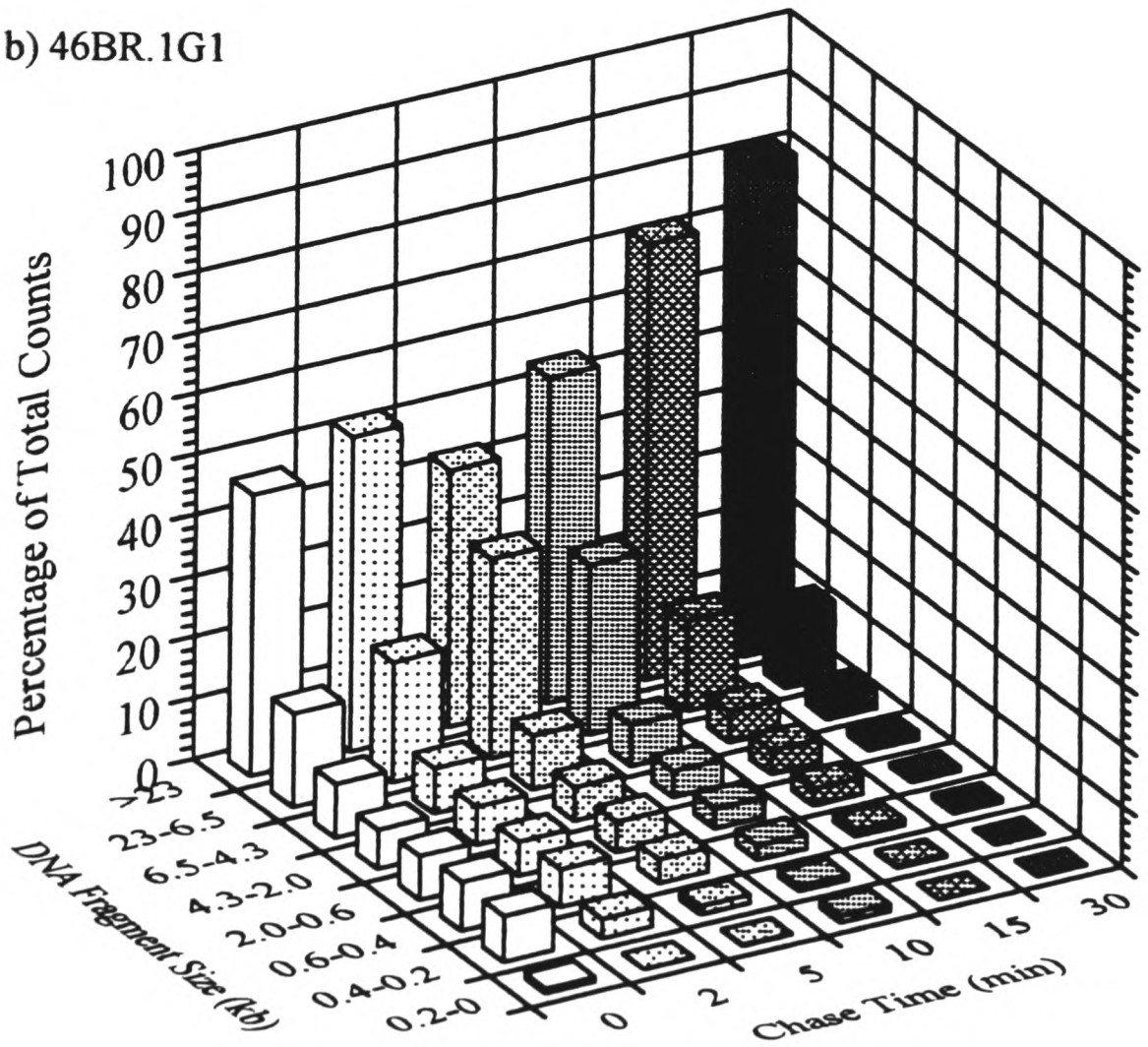
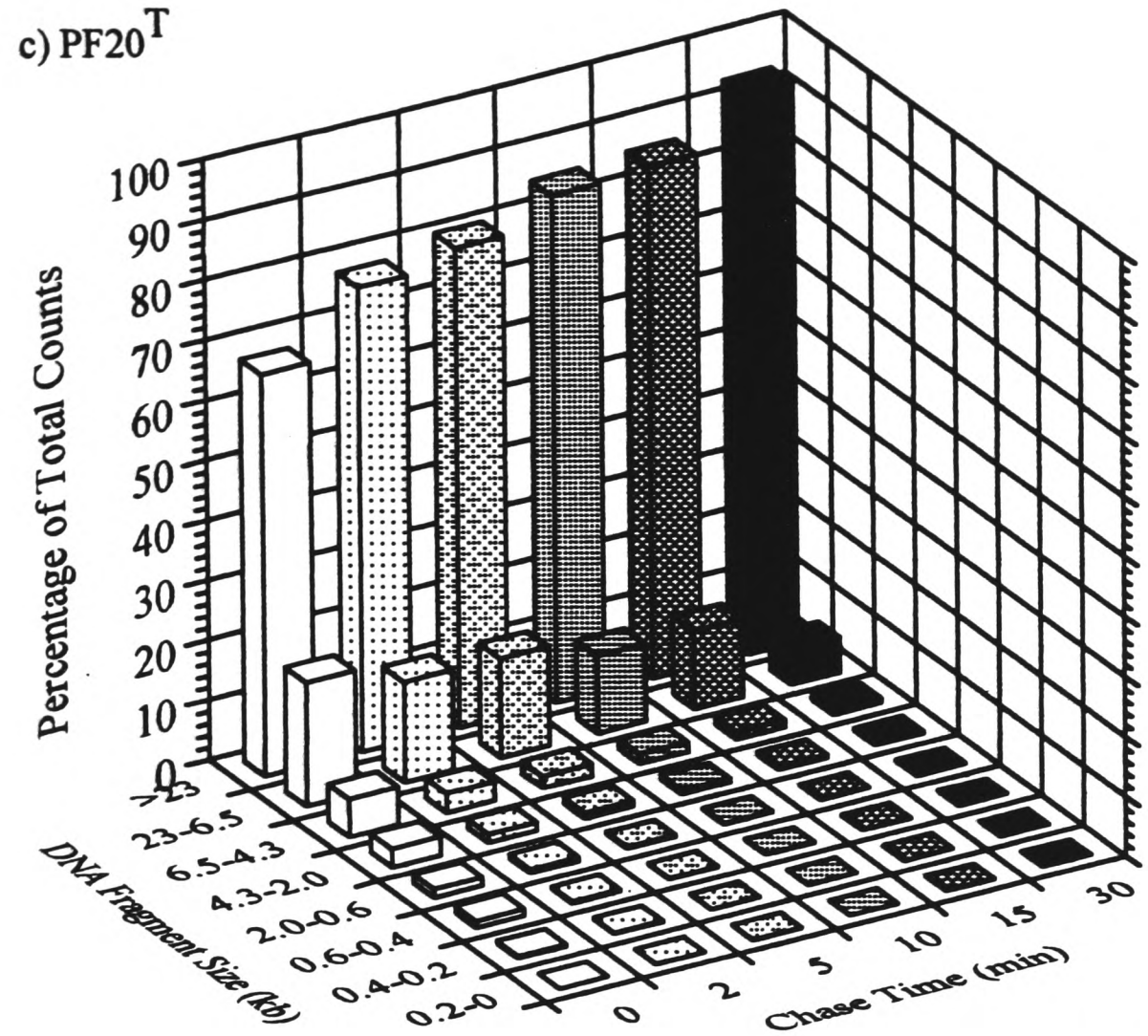


Figure 5.6.1: Analysis of newly replicated DNA (continued)

c) PF20^T



d) PFL13^T

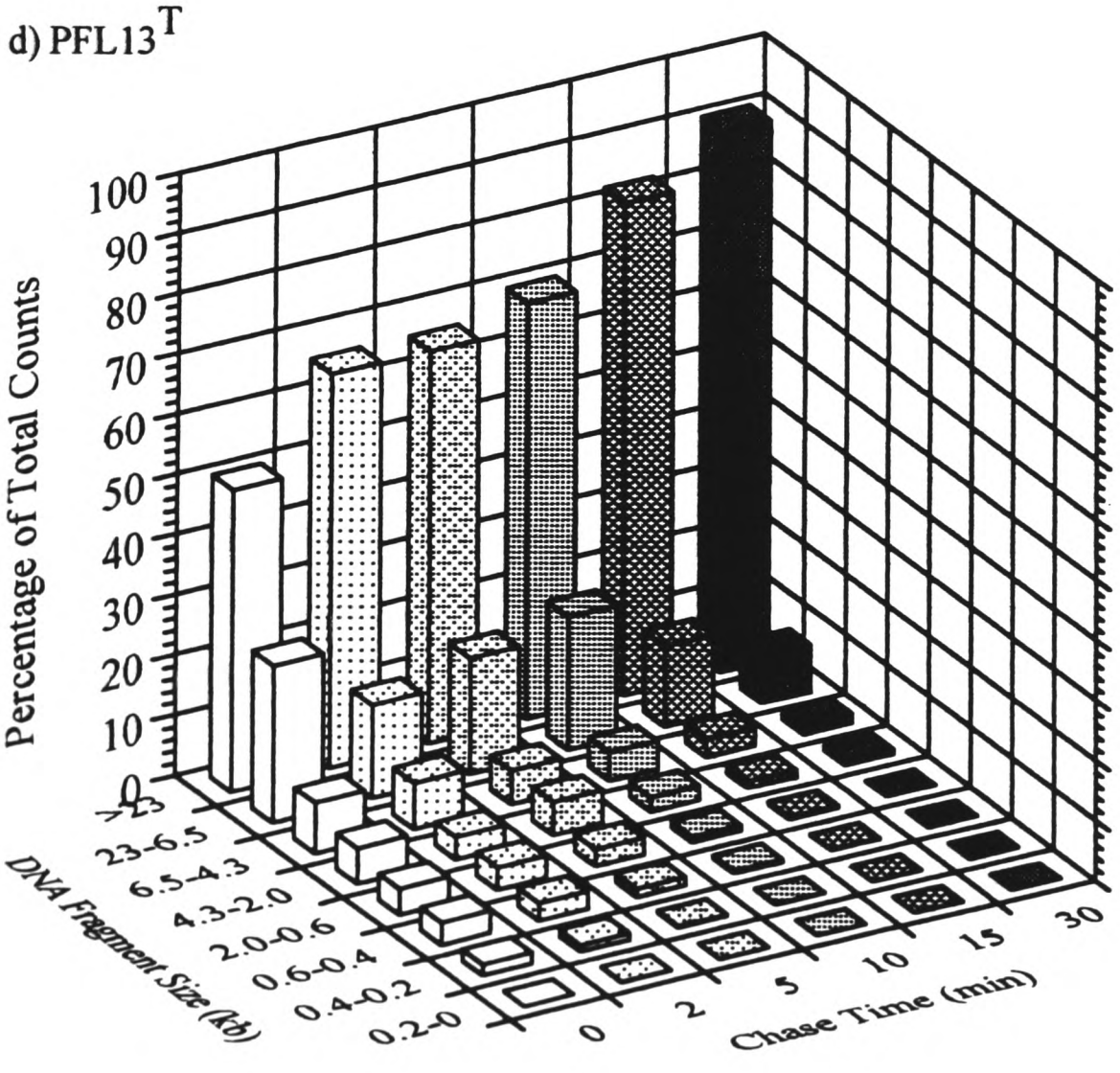
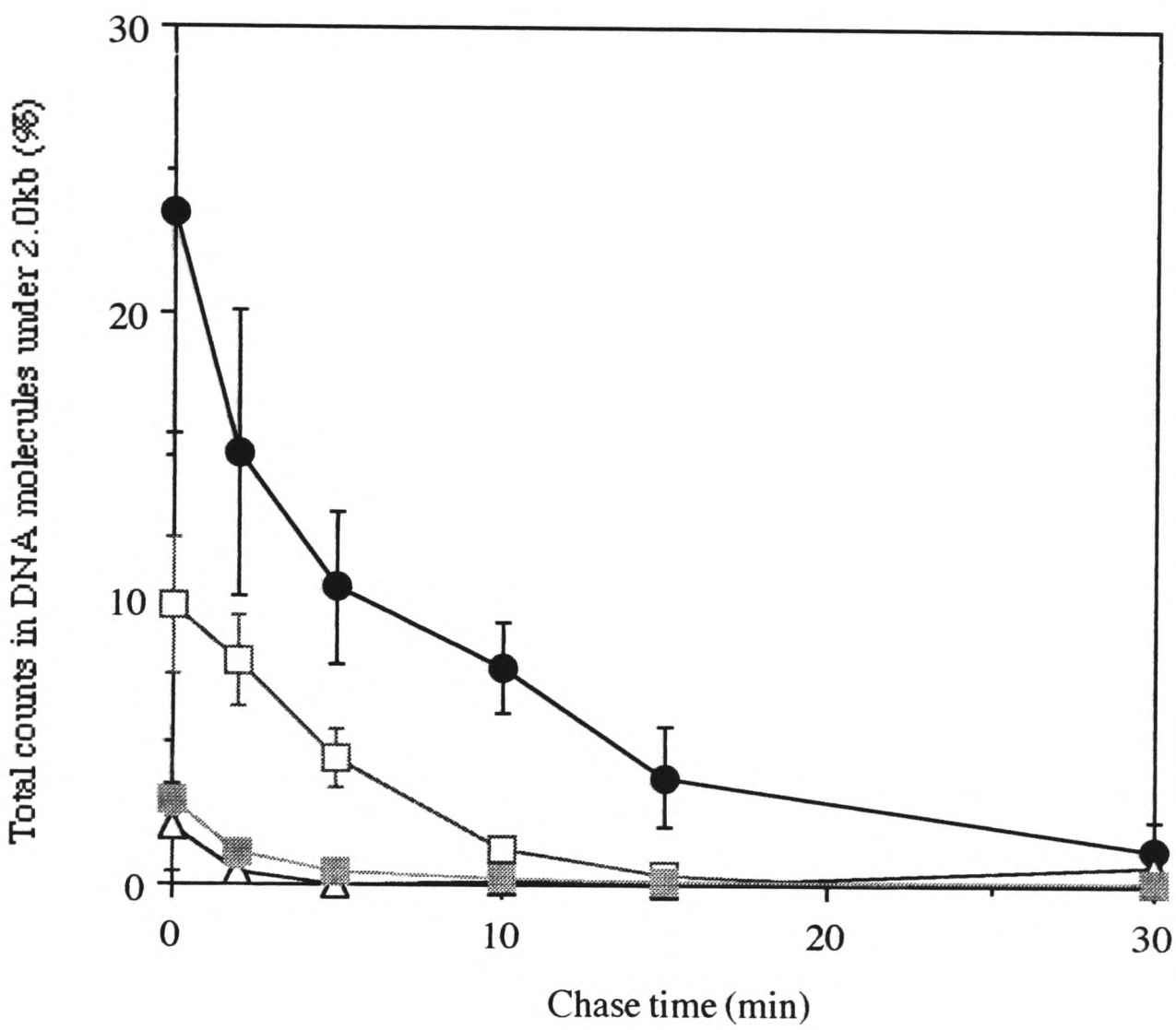


Figure 5.6.2: The percentage of total radioactive counts in DNA fragments less than 2.0kb size.

A graphical representation of the persistence of small DNA fragments formed during replication over time. As indicated previously, MRC5V1, 46BR.1G1, PF20T and PFL13T cell lines were subjected to pulse-chase analysis. The sum percentages of radiolabel present within the four fractions under 2.0kb size were calculated at each timepoint for each cell line, and these percentages plotted as a function of chase time. Each point represents the mean of two independent experiments, error bars depicting the standard deviation at each sample point.

△ ; MRC5V1, ● ; 46BR.1G1 , ■ ; $PF20^T$, □ ; $PFL13^T$


Total counts in DNA molecules under 2.0kb (%)
30
20
10
0
0
10
20
30
Chase time (min)

5.7 Sister chromatid exchanges rates are normal in cells lacking DNA ligase I

One of the characteristics of Bloom's Syndrome is a high level of chromosomal instability and somatic recombination, characterised by reciprocal interchanges between homologous chromatids (sister-chromatid exchange, SCE) (reviewed in Ray and German 1983). Cells from BS patients exhibit 60-90 SCEs per metaphase, far in excess of the normal baseline of ~10 SCE/metaphase. Although BS is not caused by mutations at the *LIG1* locus, alterations in the activity and/or structure of DNA ligase I have been observed in BS cells (Willis and Lindahl 1987, Chan *et al.* 1987). Moreover, cells from the 46BR patient were hypersensitive to the induction of SCE by DNA damaging agents (Henderson *et al.* 1985). Because cells from *Lig1* -/- embryos showed an elevated level of chromosome instability (D. W. Melton unpublished data), we were interested to determine the levels of SCEs in these cells.

It is possible to differentiate between sister chromatids in metaphase chromosomes by employing a combination of 5-bromodeoxyuridine labelling and dye staining. As a result, numbers of exchanges occurring between the two chromatids can be quantified. The method we employed was based on the use of fluorescent dye combined with Giemsa stain to make permanent cytological preparations of differentially labelled chromatids (Perry and Wolff 1974, modifications by Goto *et al.* 1977) (see Figure 5.7.1 and Figure 5.7.2). Immortalised cell lines and primary cultures from mid-gestation embryos were allowed to incorporate the base analogue 5-bromodeoxyuridine (BrdU) for two rounds of DNA replication, whereupon colcemid was added to arrest the cell division in metaphase of mitosis. Cells were fixed and dropped onto microscope slides to spread the chromosomes, before being stained with Hoechst 33258 fluorescent dye. Sister chromatids stained in this way fluoresce differentially under UV illumination, but the dye is not photostable and the image fades rapidly. Finally, to produce a more permanent image, spreads were aged by exposure to 'blacklight' (UV-A) with heating, and thereafter incubated in hot 2X SSC. Thereafter, chromosomes were stained with Giemsa and permanently mounted. Differentially stained sister chromatids could then be visualised by light microscopy, and photographed for scoring of SCEs numbers.

Table 5.7.1 details the results of analysis of the numbers of sister chromatid exchanges in cultures of human and mouse fibroblasts. SCE numbers were expressed in terms of the number of exchanges per chromosome, and also the average number of SCE per metaphase. Because immortalised cell lines displayed variable numbers of chromosomes in each cell, it was assumed that each murine metaphase included the normal diploid complement of 40 chromosomes, and that each human cell had 46.

All wild-type samples whether human (MRC5V1) or mouse (PF20T, *Lig1* +/+) primary fibroblasts) displayed SCE rates consistent with previous studies (~10 SCEs per metaphase) (e.g. McDaniel and Schultz 1992). There also appeared to be no detectable difference between primary cultures and immortalised lines in this respect. Samples from -/- mouse embryos did not show any significant difference in the rate of SCE formation compared to wild-type (ANOVA $F_{4,71}$= 1.42, P>0.05), and there was again no change following immortalisation. It was therefore concluded that the absence of DNA ligase I does not result in the elevation of SCE rates. An increased baseline level of SCEs had been noted previously in 46BR cells (Henderson *et al.* 1985). Examination of 46BR.1G1 samples in this study also revealed a statistically significant increase in SCE formation (~20 SCEs/metaphase on average) compared to the control human cell line (Mann-Whitney U test: U=36.5, P<0.05).

Table 5.7.1: The frequency of SCEs in human and mouse fibroblast samples of different *Lig1* genotypes.

	Chromosomes Scored	Number of SCEs	SCEs per chromosome	SCEs per metaphase
MRC5V1	548	99	0.18 ±0.02	8.2 ±1.1
46BR.1G1	556	241	0.43 ±0.05	19.7 ±2.3
Lig1 +/+	525	120	0.24 ±0.02	9.8 ±1.0
Lig1 -/-	486	131	0.28 ±0.02	11.1 ±0.9
$PF20^T$ (+/+)	660	154	0.23 ±0.03	9.2 ±1.1
$PFL13^T$ (-/-)	694	158	0.22 ±0.03	8.7 ±1.3

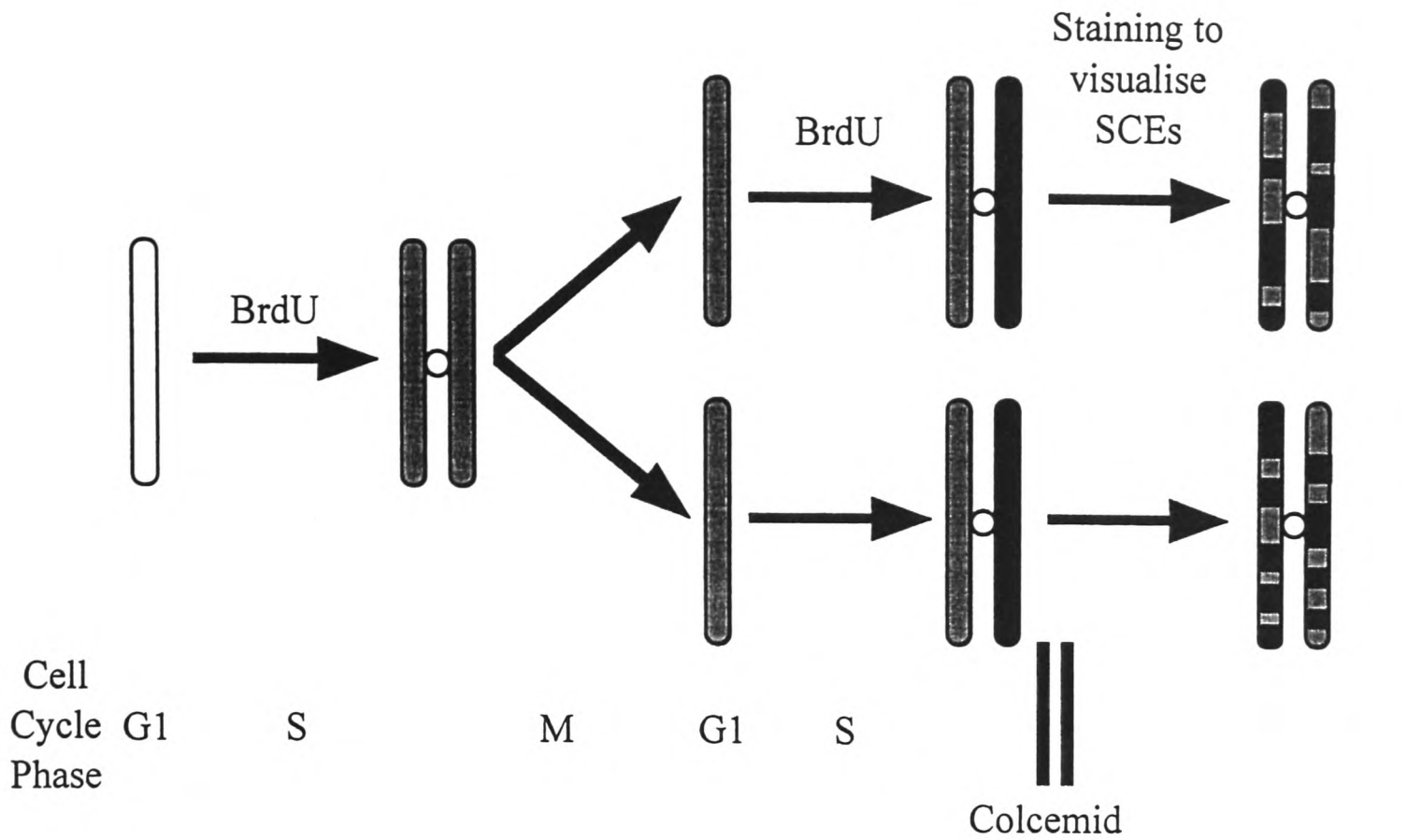
Cultures of immortalised human (MRC5V1, 46BR.1G1) and mouse ($PF20^T$, $PFL13^T$) fibroblasts and primary cultures from *Lig1* +/+ and -/- midgestation embryos were grown in BrdU-containing medium for sufficient time for two rounds of DNA replication to occur. Cells were arrested in metaphase by addition of colcemid, fixed and spread onto microscope slides. Chromosomes were then stained by incubation with Hoechst 33258, followed by exposure to a UV-A light source, incubation in 2X SSC and staining with Giemsa. Differentially stained metaphase chromosomes were examined by light microscopy and photographed using a green filter for maximal contrast between sister chromatids. Photographs were then visually scored for SCEs. The mean number of SCEs per chromosome scored was calculated for each sample (± standard error). Additionally, this value was expressed as the average number of SCEs per metaphase (± standard error), assuming that each cell contained a normal diploid complement of chromosomes (mouse; 2n = 40, human; 2n = 46).

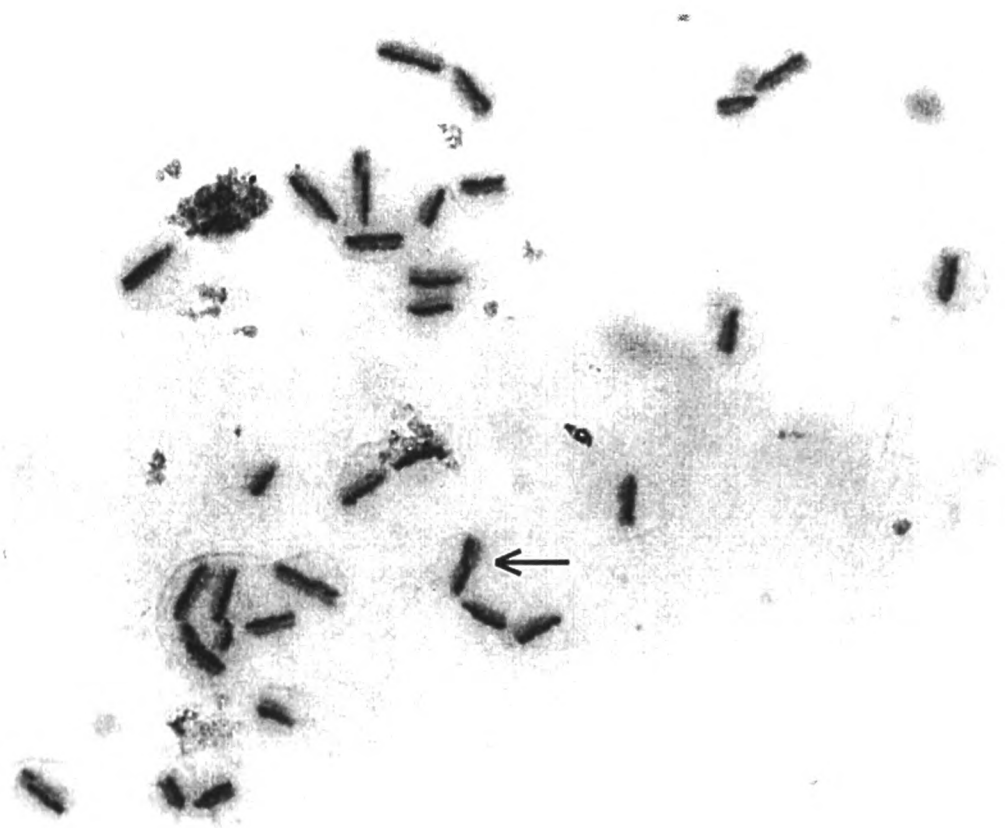
Figure 5.7.1: Schematic representation of sister chromatid differentiation by BrdU-dye staining techniques.

Adapted from Latt *et al.* 1983. Each oblong represents a single chromosome (or chromatid in a pair) composed of one double stranded DNA molecule. Circles represent centromeres. Cells are incubated with BrdU for two rounds of DNA replication. Semi-discontinuous DNA replication means the base analogue is incorporated only into daughter molecules. Hence after the second round of replication one chromatid will be composed entirely of BrdU containing DNA (bifilarly substituted), whilst the other chromatid is a hybrid of one parental strand and one BrdU containing strand (unifilarly substituted). The two sister chromatids will fluoresce differentially when stained with fluorescent dye such as Hoechst 33258 during metaphase. The unifilarly substituted chromatid fluoresces more brightly than the bifilarly substituted sister chromatid. It is then possible to identify locations where reciprocal exchanges have occurred between two sister chromatids by examination of the differential staining pattern.

Figure 5.7.2: A differentially stained mouse metaphase

A photograph of a typical wild-type mouse fibroblast metaphase spread stained by the fluorescence plus Giemsa method. A primary culture from a wild-type mouse embryo was cultured, stained and photographed as detailed previously. One reciprocal exchange between differentially stained homologous chromatids (SCE) is indicated (arrow).


Staining to
visualise
SCEs
BrdU
BrdU
Cell
Cycle
Phase
G1
S
M
G1
S
Colcemid



CHAPTER SIX:

CONDITIONAL DELETION OF *LIG1*

We have shown that DNA ligase I is required for normal development, and its absence results in death *in utero*. Prenatal lethality limits the scope of any investigation into the role of DNA ligase I in a whole animal system. Not only does it prevent study of the effects of DNA ligase I-deficiency on later developmental events, but also could conceal other consequences, such as predisposition to malignant transformation. It would therefore be desirable to be able to bypass the embryonic lethality, and to analyse DNA ligase I-deficient cells in an adult environment. Technical difficulties aside, *in vivo* haematopoietic repopulation of lethally-irradiated recipients had certain associated constraints, not least being study of only a restricted number of cell lineages. We therefore decided to pursue an alternative strategy for continued *in vivo* investigation.

Conditional gene targeting offers the potential to control the spatial and temporal boundaries of gene knockouts. The strategy makes use of the *loxP*/Cre recombinase system from the bacteriophage P1. Cre is a site-specific DNA recombinase which recognises a 34bp target sequence (*loxP* site) (Kilby, Snaith and Murray 1993). By flanking a gene of interest with two *loxP* sites (termed 'floxing'), it is subsequently possible to excise the intervening genomic locus (and hence delete the gene) via Cre-mediated recombination. Control of Cre expression can therefore be used to limit knockout of the 'floxed' gene to a particular tissue, cell lineage, or timepoint, and hence permit analysis of gene knockouts that are otherwise lethal (reviewed by Kühn and Schwenk 1997).

This chapter describes the introduction of two *loxP* sites into the *Lig1* locus by means of a 'double replacement' gene targeting strategy. The resulting mice were crossed to a line carrying a Cre transgene, and analysis of Cre-mediated recombination at the *Lig1* locus *in vivo* is also detailed. Work was carried out in association with Carolanne McEwan, Anne-Marie Ketchen, Niki Redhead, Jim Selfridge, Caroline Holmes, Stefan Selbert, Dominic Rannie, Paula Lourenço, Christine Watson, and Alan Clarke and published in *Transgenic Research* (Selbert *et al.* 1998). A reprint of the paper is included at the rear of this thesis. Construction of the targeting vectors was done by Carolanne McEwan, and targeting ES cell culture performed by Anne-Marie Ketchen. Mice expressing Cre from a mammary-specific transgene were produced by Stefan Selbert, Dominic Rannie, Paula Lourenço, Christine Watson, and Alan Clarke. Genotyping of stock mice was carried out by Carolanne McEwan and Caroline Holmes. Southern blot analysis to quantify the levels of Cre-mediated recombination in mammary tissue samples was performed by Jim Selfridge.

6.1 Double replacement gene targeting to introduce two *loxP* sites into *Lig1*

Double replacement gene targeting is a strategy for introducing subtle alterations into endogenous genes by means of two sequential rounds of gene targeting in ES cells (Reid *et al.* 1990). It is based around the use of *HPRT* selectable markers in the HPRT-deficient ES cell line HM-1, and is dependant on the ability to select both for and against *HPRT* expression (Stacey *et al.* 1994, Moore *et al.* 1995). Although previously only employed to introduce point mutations, the method is equally suitable to engineer *loxP* sites into a genomic locus.

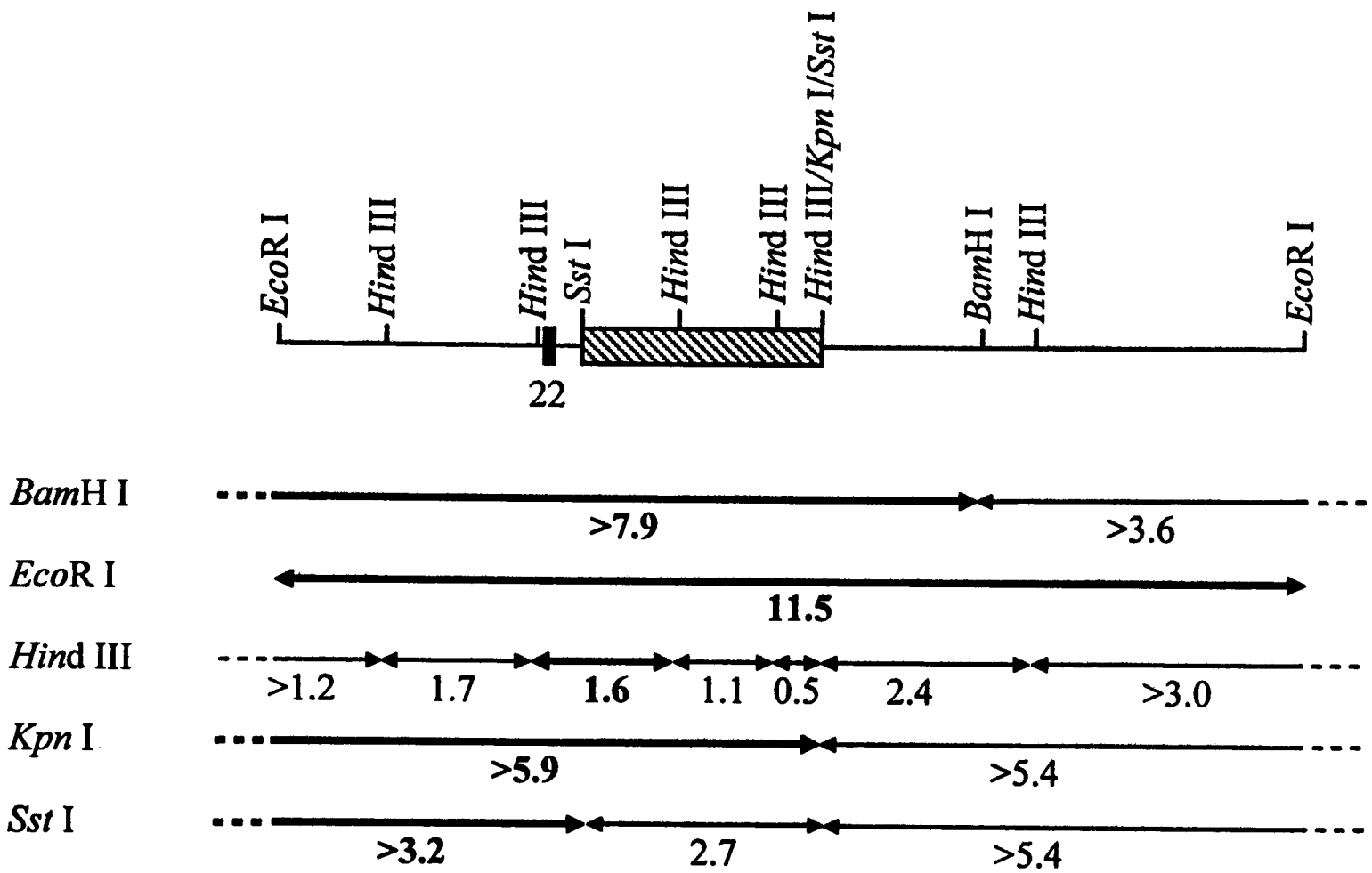
6.1.1 Retargeting the *Lig1* locus

The first step, inactivation of the *Lig1* locus by replacement of exons 23 to 27 with an *HPRT* minigene, was carried out successfully and mice lacking DNA ligase I produced. However, when the second targeting step (replacement of the *HPRT* marker with an altered region of the target gene to reconstruct the locus) was attempted, the targeting frequency was much lower than anticipated (D. W. Melton, unpublished data). Subsequent analysis indicated that *HPRT* minigene expression had been lost, resulting in a high level of spontaneous 6-thioguanine resistance which precluded the isolation of second step targeted clones (Melton, Ketchen and Selfridge 1997). In order to overcome this problem, it proved necessary to retarget the *Lig1* locus with a different *HPRT* minigene.

The original targeting vector (MLTV) included an *HPRT* minigene under the control of the constitutively expressed mouse phosphoglycerate kinase (*PGK*-1) promoter. The structure of this *PGK-HPRT* minigene was described in Magin *et al.* 1992a, whilst construction of MLTV was outlined in Jessop 1995. The targeting strategy and structure of the original targeted *Lig1* locus (designated *Lig1/PGK-HPRT*) are depicted in Figure 3.1.1 (Chapter 3). An almost identical targeting strategy was employed to repeat targeting of the *Lig1* locus, the only difference being the use of minigene DWM110 as the selectable marker (D. W. Melton, unpublished data). DWM110 is an *HPRT* expression cassette under the control of part of the natural mouse *HPRT* promoter, and contains certain key transcription control elements (Magin *et al.* 1992a). The structure of the new targeted allele (designated *Lig1*/DWM110) is shown in Figure 6.1.1.

Figure 6.1.1: Restriction map of the retargeted *Lig1* locus (*Lig1/DWM110*)

Predicted restriction map of the 3' end of the targeted *Lig1* allele *Lig1*/DWM110 based on the structures of the wild-type *Lig1* allele and *Lig1*/DWM110 targeting vector (data not shown). Drawn to scale: 1cm = 2kb. Black boxes indicate exons; lines, introns or flanking DNA, hatched box; *HPRT* minigene. Restriction sites of *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I are indicated, along with the arrangement of resulting restriction fragments for each individual enzyme. Sizes are given in kb. Fragments (and associated sizes) which contain exons, and hence will be detected on Southern blots probed with *Lig1* cDNA, are in bold.


EcoR I
Hind III
Hind III
Sst I
Hind III
Hind III
Hind III/Kpn I/Sst I
BamH I
Hind III
EcoR I
22
BamH I
>7.9
>3.6
EcoR I
11.5
Hind III
>1.2
1.7
1.6
1.1
0.5
2.4
>3.0
Kpn I
>5.9
>5.4
Sst I
>3.2
2.7
>5.4

6.1.2 Second step gene targeting

A second step gene targeting vector was constructed for use in a second homologous recombination step to reconstruct a functional *Lig1* allele (C. McEwan, D. W. Melton, unpublished data). The targeting vector was introduced into an ES cell clone heterozygous for the new targeted *Lig1* allele (HM-1/*Lig1*/DWM110 #12). Selection for the loss of the *HPRT* marker was applied, using the purine analogue 6-thioguanine (6-TG), and those 6-TG resistant colonies arising were screened for the presence of a 'floxed' *Lig1* allele (denominated f) by PCR (A.-M. Ketchen, D. W. Melton, unpublished data. See Figure 6.1.2). 14 6-TG resistant colonies gave positive results from the PCR assay, and of these 8 were subsequently analysed further (designated HM-1/*Lig1*/*loxP* #7, 9, 41, 48, 58, 62, and 68). A schematic representation of the strategy used to produce the floxed *Lig1* allele is shown in Figure 6.1.2.

6.1.3 Confirmation of the structure of the floxed *Lig1* allele

The second step gene targeting vector used to reconstruct the *Lig1* allele was based around the same 13.9kb *Eco*RI genomic fragment that had been used in construction of the first step knockout vectors. The second step vector contained the whole 13.9kb *Eco*RI fragment, but with two *loxP* sites engineered into *Sst*I sites within the genomic fragment. One site was located between exons 22 and 23, whilst the other was downstream of the final exon (27). Insertion of the two Cre recognition sequences during vector construction had resulted not only in the loss of the two SstI sites, but also the introduction of two novel restriction sites. The 5' *loxP* site was tagged with a *Hin*dIII site, while the downstream sequence was marked by the presence of an *Eco*RI site. The presence of these additional restriction enzyme cleavage sites thus permitted the floxed *Lig1* allele to be distinguished from the wild-type allele.

Figure 6.1.3 shows the predicted structure of the floxed *Lig1* allele arising after the second step gene targeting, along with predicted DNA fragments arising from digestion with *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I restriction enzymes. Those fragments highlighted in bold contain exons of the *Lig1* gene and should thus hybridise to a *Lig1* cDNA probe and be detected on Southern blots. When the 8 ES cell clones were analysed by Southern blotting the pattern of fragments observed was consistent with the proposed structure of the floxed *Lig1* allele. It was thererfore concluded that all 8 clones had correctly undergone the desired homologous recombination event, and were heterozygous for the floxed allele (+/f). Analysis of two of the 8 clones is shown in Figure 6.2.4, and discussed further in section 6.2.4.

Figure 6.1.2: Schematic representation of the double replacement gene targeting strategy to introduce *loxP* sites into the *Lig1* locus

Schematic representation of 13.9kb at the 3' end of the wild-type mouse DNA ligase I locus, targeted gene structure, and the structure of the reconstructed 'floxed' locus following second step gene targeting. In the first targeting step, the last 5 exons of the *Lig1* locus were replaced with an *HPRT* minigene by homologous recombination. A second homologous recombination step was subsequently used to reconstruct a functional allele, but with exons 23 to 27 flanked by two *loxP* sites

Drawn to scale: 1cm = 2kb. Solid boxes indicate exons (numbered according to the human convention); hatched box, *HPRT* minigene (DWM110); thin lines, *Lig1* introns or flanking regions; solid triangles, *loxP* sites. Arrows indicate the positions at which PCR primer pairs used for genotyping bind. *Bam*HI restriction sites are marked by B, *Eco*RI sites by E, and *Hin*dIII sites by H.

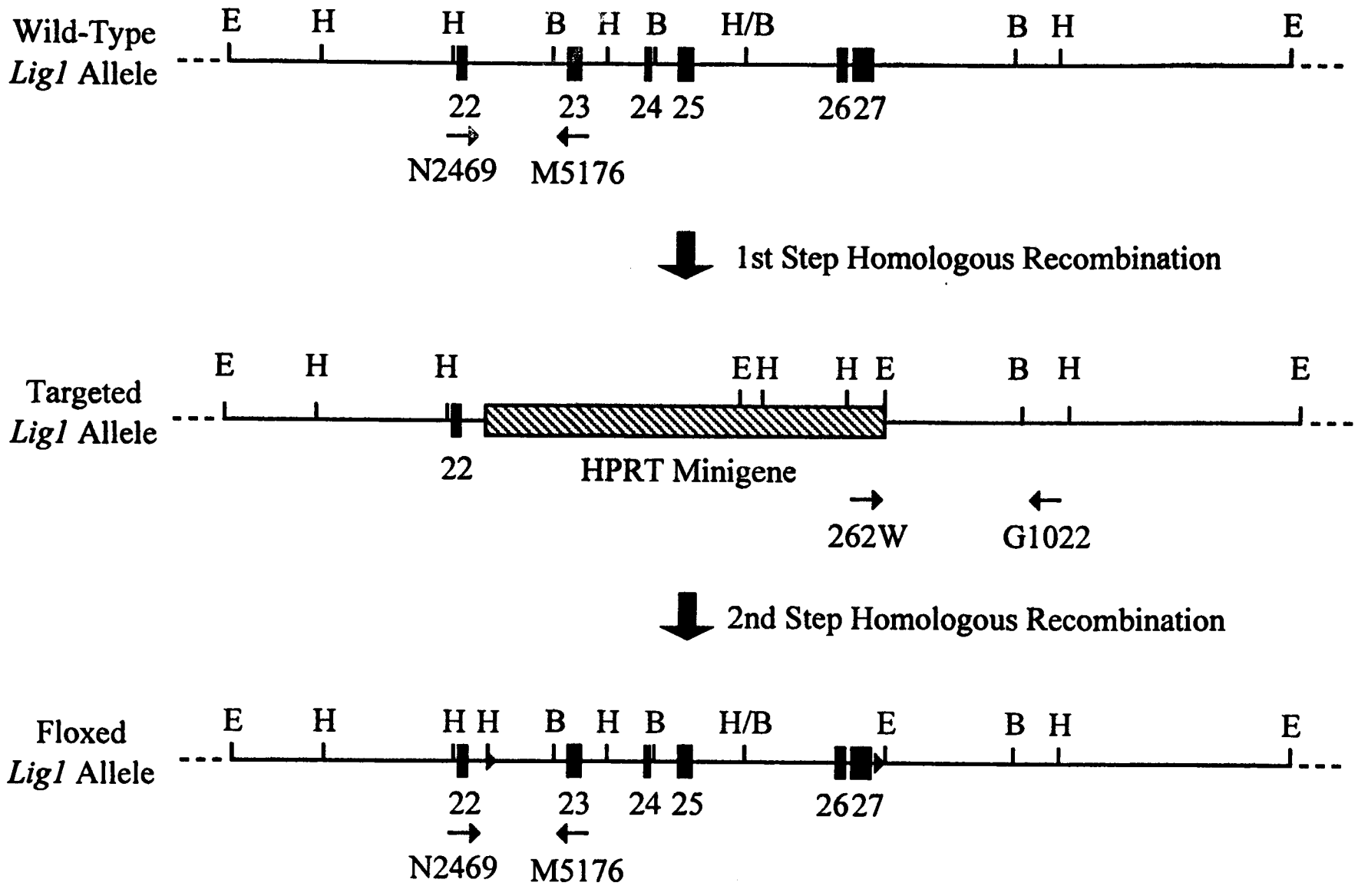
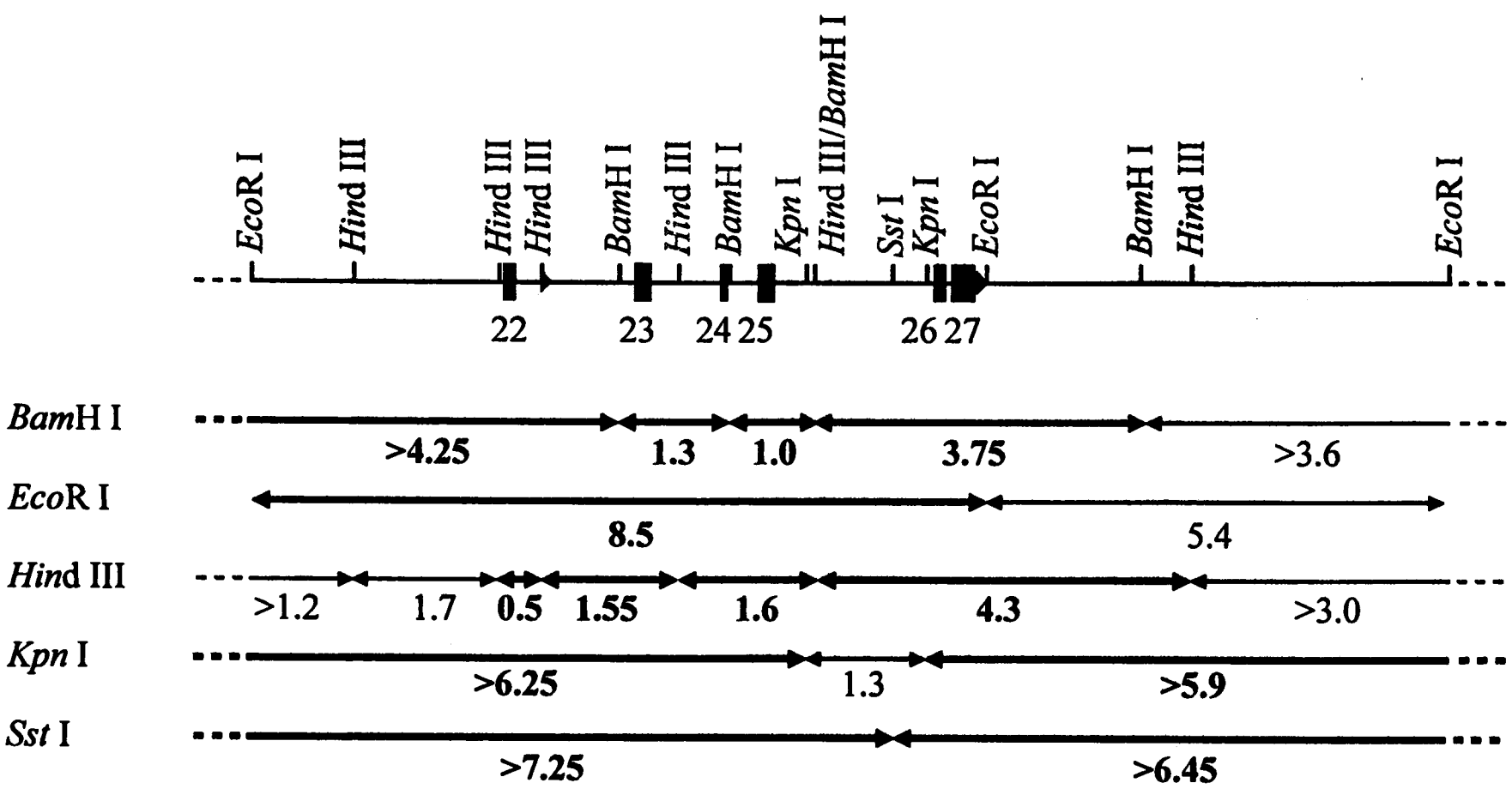

Wild-Type
Lig1 Allele
E H H B H B H/B B H E
22 23 24 25 2627
N2469 M5176
1st Step Homologous Recombination
Targeted
Lig1 Allele
E H H EH H E B H E
22 HPRT Minigene
262W G1022
2nd Step Homologous Recombination
Floxed
Lig1 Allele
E H H H B H B H/B E B H E
22 23 24 25 2627
N2469 M5176

Figure 6.1.3: Restriction map of the floxed *Lig1* locus

Predicted restriction map of the 3' end of the floxed *Lig1* allele based on the structures of the wild-type *Lig1* allele and second step targeting vector (data not shown). Drawn to scale: 1cm = 2kb. Black boxes indicate exons; lines, introns or flanking DNA; solid triangles, *loxP* sites. Restriction sites of *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I are indicated, along with the arrangement of resulting restriction fragments for each individual enzyme. The orientation of the *loxP* sites is indicated by the direction of the triangle, both *loxP* sites being in the same orientation. Sizes are given in kb. Fragments (and associated sizes) which contain exons, and hence will be detected on Southern blots probed with *Lig1* cDNA, are in bold.


EcoR I
Hind III
Hind III
Hind III
BamH I
Hind III
BamH I
Kpn I
Hind III/BamH I
Sst I
Kpn I
EcoR I
BamH I
Hind III
EcoR I
22
23
24 25
26 27
BamH I
>4.25
1.3
1.0
3.75
>3.6
EcoR I
8.5
5.4
Hind III
>1.2
1.7
0.5
1.55
1.6
4.3
>3.0
Kpn I
>6.25
1.3
>5.9
Sst I
>7.25
>6.45

6.1.4 Second step targeting restores expression to the floxed *Lig1* allele

Whilst second step targeting regenerated the 3' end of the *Lig1* locus, it remained possible that the engineered changes might disrupt expression or result in an altered transcript being produced. To determine whether expression from the floxed allele was comparable to the wild-type allele, northern blot analysis was carried out on ES cell clones heterozygous for the floxed *Lig1* allele (Figure 6.3.5). Total RNA was separated by electrophoresis, transferred onto nylon membrane, and probed with a *Lig1* cDNA probe. The wild-type *Lig1* transcript was easily detected in the wild-type ES cell line (HM-1), and present at reduced levels in the ES cell clone heterozygous for the knockout *Lig1* allele (#53). The samples heterozygous for the floxed *Lig1* allele (#58 and #68) displayed a single transcript of normal size, and the signal was of equal intensity compared to wild-type samples. Thus it was concluded that the reconstructed *Lig1* allele was transcriptionally functional, and that expression from the floxed allele was not detectably different from wild-type.

6.2 Conditional deletion of *Lig1* by Cre-mediated recombination

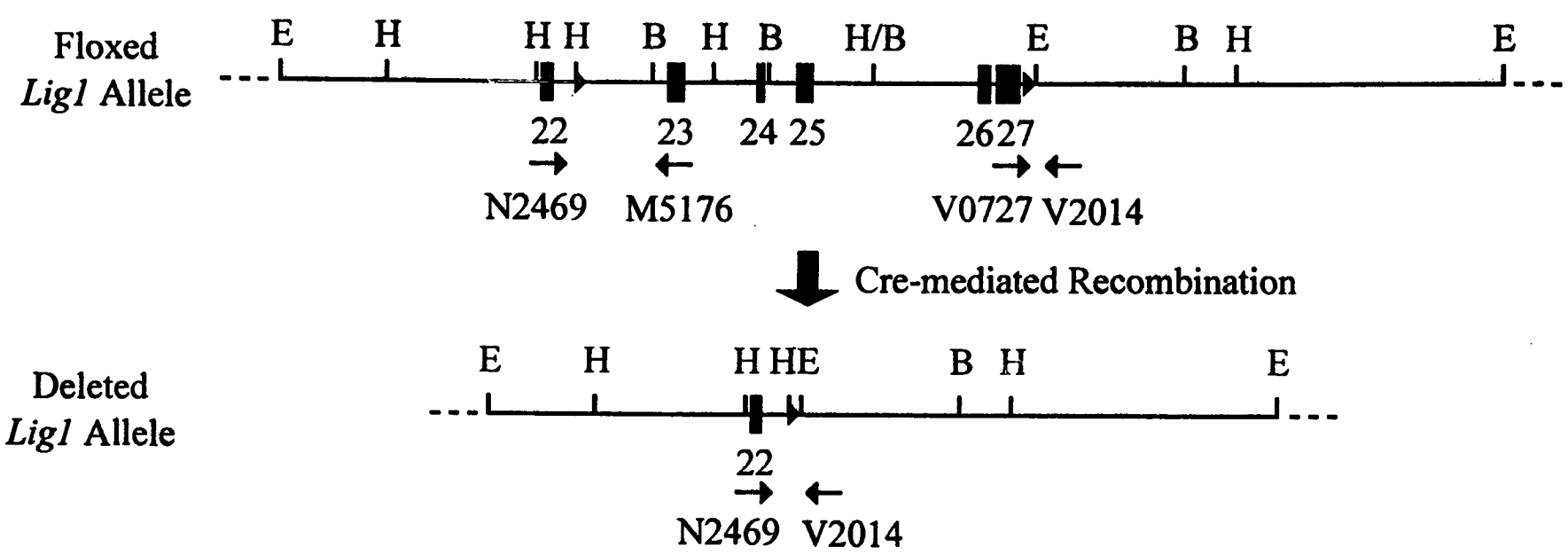
Cre-mediated recombination between the two *loxP* sites within the floxed *Lig1* allele has a number of consequences for the structure of the *Lig1* locus (Figure 6.2.1). Firstly, because the two *loxP* sites are arranged in the same orientation, recombination will result in excision of the intervening genomic region as a small circular DNA molecule. This molecule will lack elements necessary for maintenance and propagation within the cell, will be prone to nucleolytic degradation, and hence will subsequently be lost. Because exons 23 to 27 are contained within the region being excised, they will effectively be deleted by Cre-mediated recombination. Exons 23 to 27 encode the terminal 175 amino acids of the DNA ligase I protein. It has been reported that removal of any more than 15 amino acids from the C-terminus of the protein abolishes enzymatic activity (Kodama *et al.* 1991). Deletion of the final 5 exons should therefore disrupt gene function.

A second consequence of the recombination event is that previously unassociated regions will be brought into juxtaposition. Removal of the intervening sequences will bring exon 22 and the 3' flanking region into close proximity, in a manner unique to the deleted *Lig1* allele. We exploited this novel association to design a PCR assay specific for the deleted allele (denominated Δ).

Figure 6.2.1: Schematic representation of the effect of Cre-mediated recombination on the floxed *Lig1* allele

Schematic representation of 13.9kb at the 3' end of the floxed mouse DNA ligase I locus, and the structure of the locus following Cre-mediated recombination between the two *loxP* sites. Cre-mediated recombination between the two *loxP* sites flanking exons 23 to 27 will result in excision of the intervening genomic region as a circular molecule (not shown), which will subsequently be lost from the cell. Consequently the last five exons of *Lig1* will effectively be deleted, and previously unassociated regions will be brought into juxtaposition.

Drawn to scale: 1cm = 2kb. Solid boxes indicate exons (numbered according to the human convention); thin lines, *Lig1* introns or flanking regions; solid triangles, *loxP* sites. Arrows indicate the positions at which PCR primer pairs used for genotyping bind. *Bam*HI restriction sites are marked by B, *Eco*RI sites by E, and *Hin*dIII sites by H.


Floxed
Lig1 Allele
E H H H B H B H/B E B H E
22 23 24 25 2627
N2469 M5176 V0727 V2014
Cre-mediated Recombination
Deleted
Lig1 Allele
E H H HE B H E
22
N2469 V2014

6.2.1 PCR assay to detect the deleted *Lig1* allele

Oligonucleotide primers which bound within exon 22 (N2469) and the 3' flanking region of the *Lig1* gene downstream of exon 27 (V2014) were employed in a PCR reaction specific for the deleted allele. Prior to recombination, the distances between the two primer binding sites were too great for efficient amplification. Thus the wild-type and floxed alleles could not give a product from this reaction. Only after Cre-mediated recombination had excised the intervening region were the two primer binding sites close enough for a product to be amplified. Hence, the 0.5kb product spanning the Cre-induced join was diagnostic of the deleted *Lig1* allele. Figure 6.2.2 shows the results of a screen for Cre-mediated recombination using the PCR assay to identify ES cell colonies containing the deleted *Lig1* allele.

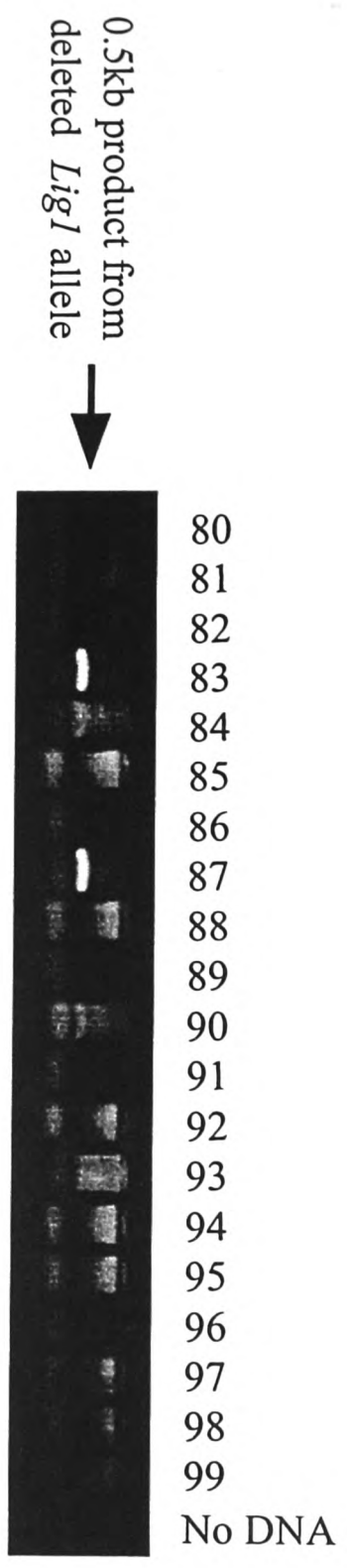
6.2.2 Conditional deletion of *Lig1 in vitro* by transient expression of Cre recombinase

To test whether Cre-mediated recombination could be induced between the two *loxP* sites in a genomic context, Cre recombinase was transiently expressed in ES cell clones carrying the floxed *Lig1* allele. The Cre expression vector pMC/CreN was electroporated into two ES clones heterozygous for the floxed allele (#58 and #68). In each case approximately 200μg of circular plasmid DNA was introduced into $20X10^6$ cells, which were then plated out as a single pool of cells in a large flask (denoted #58 1° and #68 1°). Two days later, samples of each pool were taken for genomic DNA preparation, after which a second round of electroporation was carried out on the remaining cells, again using 200μg of pMC/CreN. The cells were again plated out as pools (denoted #58 2° and #68 2°), allowed to proliferate for 2 days, and then samples taken for genomic DNA preparation.

Introduction of the circular pMC/CreN plasmid into ES cells should be sufficient to permit expression of Cre recombinase in a transient manner. The reported efficiency of the Cre-mediated recombination between *loxP* sites is very high. We therefore hypothesised that transient expression of Cre in ES cells carrying the floxed *Lig1* allele should be sufficient to induce Cre-mediated deletion of the allele, albeit at a low frequency. In order to detect such rare events, a PCR-based screening strategy was designed whereby generation of a PCR product specific for the deleted *Lig1* allele was employed as a marker of a recombination event (see section 6.2.1). DNA samples from pools of ES cells were screened for the presence of a Cre-mediated recombination event using this PCR reaction. Positive results were

Figure 6.2.2 Screening for the deleted *Lig1* allele by PCR

The results of a screen of ES cell colonies for Cre-mediated recombination using a PCR reaction specific for the deleted *Lig1* allele. A Cre expression vector was introduced into ES cell clones carrying the floxed *Lig1* allele, and the surviving cells plated out at single cell density. ES cell colonies arising were picked and cells digested with proteinase K, before the lysates were used as templates for PCR. A PCR reaction specific for the deleted *Lig1* allele was carried out using the primer pair N2469+V2014. The products were separated by agarose gel electrophoresis, and the 0.5kb product specific for the deleted allele is indicated by an arrow. Figures are the numerical identifiers of the colonies being screened, numbers 83 and 87 proving to be positive for the deleted allele.


0.5kb product from
deleted *Lig1* allele
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
No DNA

obtained from all four of the ES cell pools transfected with the Cre expression plasmid (data not shown).

6.2.3 Isolation of ES cell clones heterozygous for the deleted *Lig1* allele

The PCR assay for Cre-mediated recombination of the floxed *Lig1* allele produced a 520bp product when tested on pools of *Lig1* +/f ES cells expressing Cre recombinase. In order to confirm that the positive result was due to the presence of a deleted *Lig1* allele within these cell populations, it was important to isolate an ES cell clone heterozygous for the deleted allele (+/Δ) from the pools.

Pools of cells from the second sequential round of electroporation (#58 2° and #68 2°) were plated out at single cell density to allow colony formation. When colonies were sufficiently large, 96 were picked at random and screened for the presence of the deleted *Lig1* allele by PCR. Individual colonies were harvested by scraping and mechanically disaggregated in a small volume of medium. Thereupon one half of each cell suspension was deposited into a well of a multiwell plate for continued expansion of the colony, whilst the other half was used for the rapid colony screening procedure. Cells were pelleted by brief centrifugation and resuspended in 1X PCR buffer plus proteinase K. After 2 hours digestion, the proteinase was inactivated by heating, whereupon a small aliquot was used in the PCR assay specific for the deleted *Lig1* allele. Out of 96 colonies screened in this manner, 3 produced positive signals (#69, 83, 87) (data not shown). Of these, the first was isolated from ES cell pool derived from clone HM-1/*Lig1*/*loxP* #58, and hence was designated HM-1/*Lig1*/*loxP* #58/Δ69, while the latter two were isolated from HM-1/*Lig1*/*loxP* #68, and hence were designated HM-1/*Lig1*/*loxP* #68/Δ83 and Δ87.

6.2.4 Confirmation of the structure of the deleted *Lig1* allele

Cre-mediated recombination between the two *loxP* sites within the floxed *Lig1* allele results in excision of the central *loxP*-flanked DNA segment as a circular molecule. The circular episome will not be replicated and is susceptible to degradation, and thus will ultimately be lost from the cell. Therefore, Cre-mediated recombination will effectively delete exons 23 to 27 from the endogenous floxed *Lig1* locus. Figure 6.2.3 shows the predicted structure of the deleted *Lig1* allele arising from Cre-mediated recombination of the floxed allele, along with the predicted DNA fragments arising from digestion with *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I restriction enzymes. Those fragments highlighted in bold contain exons of the *Lig1* gene, and should thus hybridise to a *Lig1* cDNA probe and be detected by Southern blotting. The deleted allele contains a single 0.5kb *Hin*dIII fragment that includes an exon, whereas the floxed allele is composed of four exon containing

fragments (0.5, 1.5, 1.6 and 4.3kb). Thus Cre-mediated deletion of the *Lig1* allele should result in the disappearance of the 1.5, 1.6, and 4.3kb fragments.

DNA was prepared from the original ES cell line (HM-1), a clone heterozygous for the targeted knockout allele (#53), two clones heterozygous for the floxed allele (#58 and #68), and the three clones which had given a positive result from the PCR reaction specific for the deleted allele (Δ69, Δ83 and Δ87). Following digestion with *Hin*dIII, the DNA was subjected to Southern blot analysis (Figure 6.2.4), results of which are summarised in Table 6.2.1. Using a *Lig1* cDNA fragment as a probe, 6 bands were detected in wild-type and +/- samples (10.7, 4.3, 3.9, 2.0, 1.8, 1.6kb). Previous work had shown that 10.7, 3.9 and 1.8kb bands were derived from the central region of the *Lig1* gene and should thus be unaffected by targeting. The sample heterozygous for the first step targeted allele displayed an identical banding pattern, except that 4.3, 2.0 and 1.6kb bands were reduced in intensity, consistent with the structure of the *Lig1/PGK-HPRT* allele. Both lanes from HM-1/*Lig1/loxP* ES cell clones (#58 and #68) exhibited identical patterns of 7 bands (10.7, 4.3, 3.9, 2.0, 1.8, 1.6, and 1.5kb). 6 bands were shared with +/+ and +/- samples, although the 2.0kb band appeared less intense when compared to the 1.8kb reference band. The seventh band was a novel 1.5kb fragment.

The 13.9kb *Eco*RI genomic region corresponding to the 3' end of the wild-type allele contains 3 *Hin*dIII fragments that include exons (2.0, 1.6 and 4.3kb respectively). The reconstructed allele has an additional *Hin*dIII site within the 2.0kb fragment, hence dividing this fragment into two. Thus second step gene targeting should result in the disappearance of the 2.0kb fragment and appearance of two novel bands of 0.5 and 1.5kb, but leave both other wild-type bands (1.6 and 4.3kb) unchanged. Whilst an additional novel 0.5kb band had been predicted to arise from second step targeting, this fragment contained only a single exon (22) which had previously been observed to be difficult to detect by Southern blotting (Jessop 1995). Failure to detect this small fragment was hence regarded as a technical problem rather than a conflict with the predicted restriction map. Therefore, the pattern of fragments observed was not inconsistent with the proposed structure of the floxed *Lig1* allele.

The 1.5kb band detected in +/f samples was not present in lanes corresponding to clones Δ69 and Δ83. In these samples, six bands were observed in a wild-type pattern (10.7, 4.3, 3.9, 2.0, 1.8, 1.6kb), although the 4.3, 2.0 and 1.6kb fragments were reduced in intensity compared to wild-type samples. As previously noted, the predicted 0.5kb band arising as a result of reconstruction of the floxed allele, which should also be present in +/Δ samples, could not be detected. The pattern of fragments observed in samples Δ69 and Δ83 was therefore not inconsistent with the

Table 6.2.1: Summary of the predicted *Hin*dIII restriction fragments from wild-type, floxed and deleted *Lig1* alleles, and bands detected by Southern blotting

Predicted *Hin*dIII Fragments			Observed Banding Pattern		
Wild-Type Allele	Floxed Allele	Deleted Allele	+/+ Samples	+/f Samples	+/Δ Samples
4.3	4.3	0.5	(10.7)	(10.7)	(10.7)
2.0	1.6		4.3	4.3	4.3*
1.6	1.5		(3.9)	(3.9)	(3.9)
	0.5		2.0	2.0*	2.0*
			(1.8)	(1.8)	(1.8)
			1.6	1.6	1.6*
				1.5	

Predicted and actual *Hin*dIII restriction fragments arising from the 3' end of the mouse DNA ligase I locus when probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA. Fragment sizes in brackets are those arising from parts of the gene outwith the 13.9kb *Eco*RI fragment used in the targeting experiments. Fragments marked with an asterisk are those which were present at a reduced intensity compared to +/+ samples.

proposed structure of the deleted allele. It was concluded that these two ES cell clones had correctly undergone Cre-mediated deletion of the floxed *Lig1* allele and were heterozygous for the deleted *Lig1* allele.

However, while the ES cell clone Δ87 had given a positive signal from the *Lig1* deleted PCR reaction, the pattern of fragments observed by Southern blotting was not consistent with this sample being heterozygous for the deleted *Lig1* allele. Seven fragments were present in the lane corresponding to Δ87 in Figure 6.2.4 (10.7, 4.3, 3.9, 2.0, 1.8, 1.6, and 1.5kb), and in this respect closely resembled samples heterozygous for the floxed *Lig1* allele. However, the 1.5kb *Hin*dIII fragment specific for the floxed allele did appear slightly reduced in intensity when compared to reference bands. The nature of the Δ87 sample is discussed further in chapter 7, but Southern blotting did not eliminate the possibility that a fraction of the cells from which the sample was derived had undergone Cre-mediated deletion of the floxed *Lig1* allele.

Figure 6.2.3: Restriction map of the deleted *Lig1* locus

Predicted restriction map of the 3' end of the deleted *Lig1* allele based on the structure of the floxed *Lig1* allele. Drawn to scale: 1cm = 2kb. Black boxes indicate exons; lines, introns or flanking DNA; solid triangles, *loxP* sites. Restriction sites of *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I and *Sst*I are indicated, along with the arrangement of resulting restriction fragments for each individual enzyme. Sizes are given in kb. Fragments (and associated sizes) which contain exons, and hence will be detected on Southern blots probed with *Lig1* cDNA, are in bold.

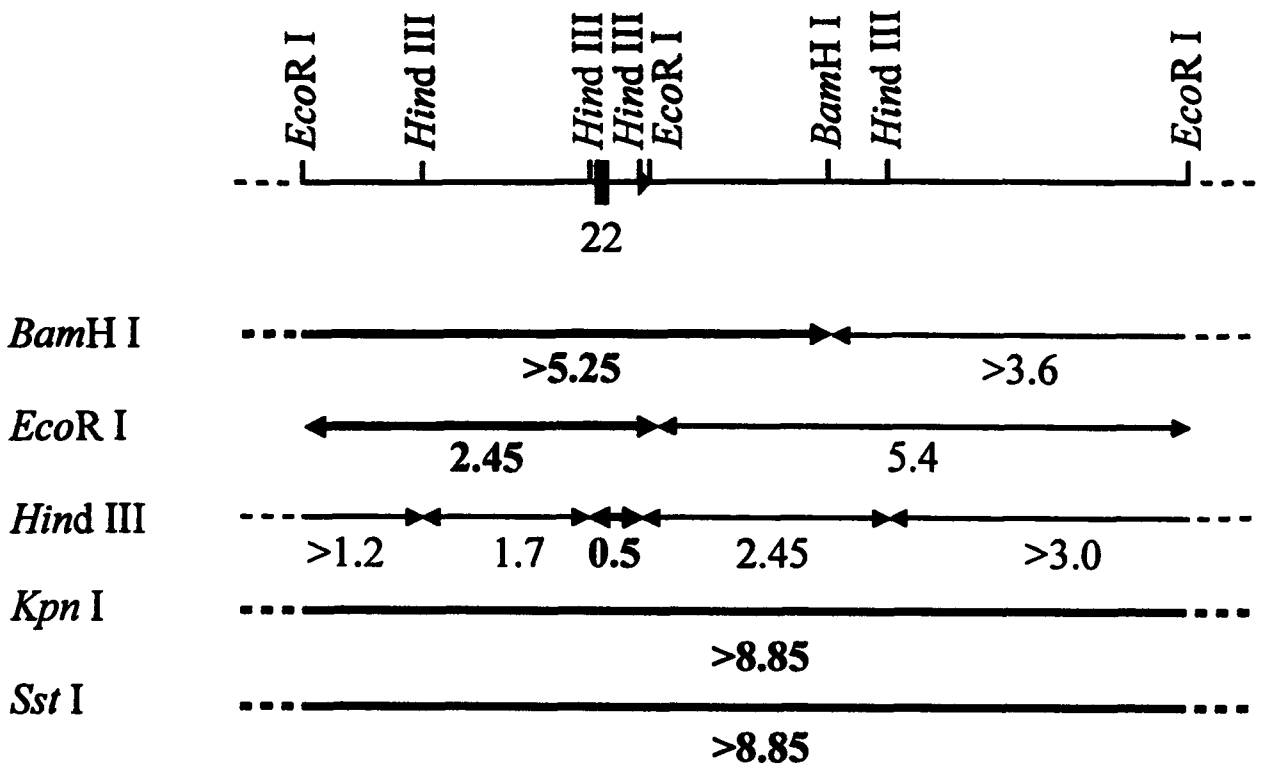
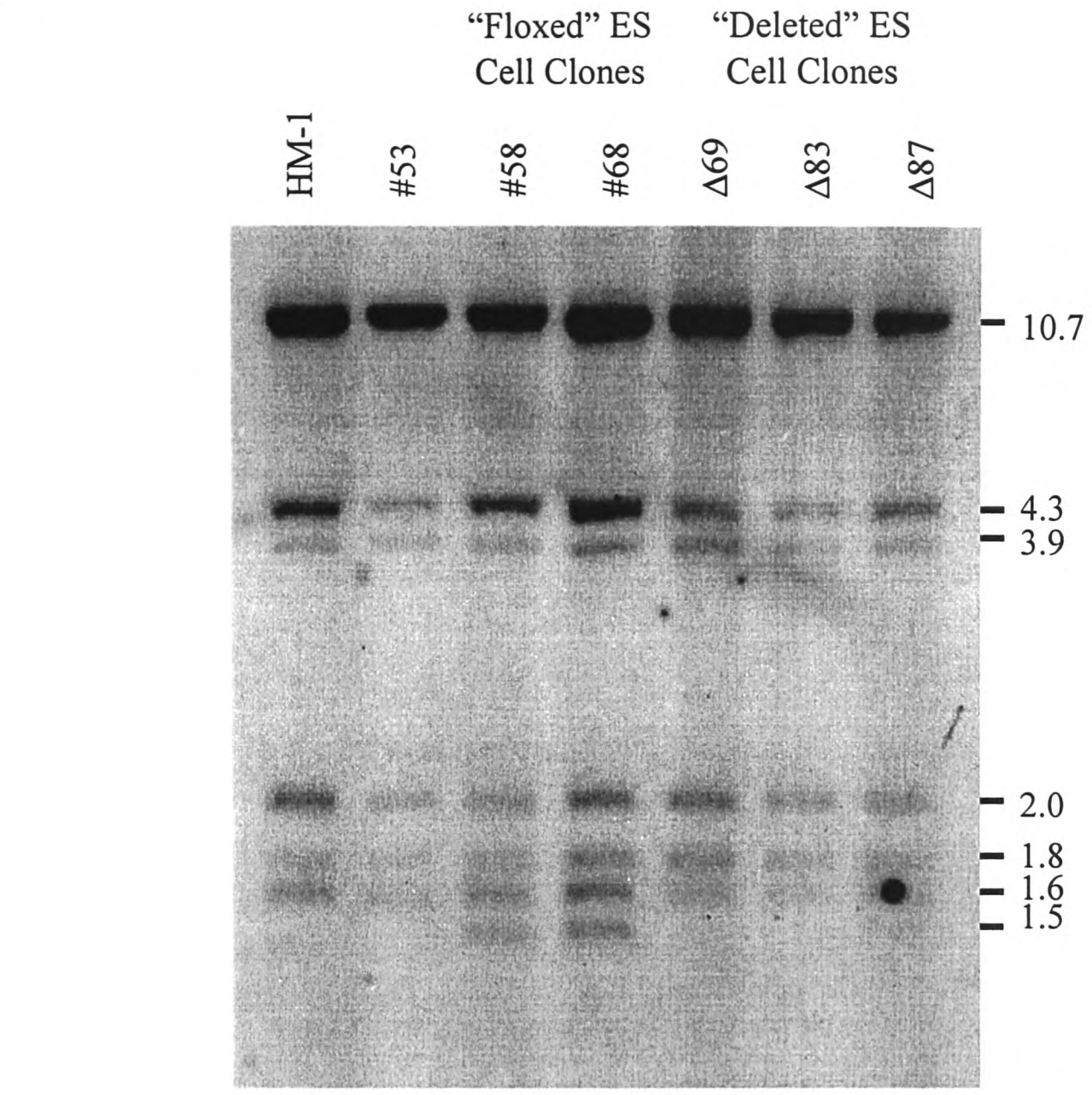

EcoR I
Hind III
Hind III
Hind III
EcoR I
BamH I
Hind III
EcoR I
22
BamH I
>5.25
>3.6
EcoR I
2.45
5.4
Hind III
>1.2
1.7
0.5
2.45
>3.0
Kpn I
>8.85
Sst I
>8.85

Figure 6.2.4: Southern blot analysis of ES cell clones positive for the *Lig1* deleted allele PCR assay

Confirmation of the structure of the deleted *Lig1* allele by Southern analysis. Genomic DNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the targeted *Lig1/PGK-HPRT* allele (#53), 2 ES cell clones heterozygous for the floxed *Lig1* allele (HM-1/*Lig1/loxP* #58 and #68), and the 3 ES cell clones which had given positive results from PCR screening for the deleted *Lig1* allele (HM-1/*Lig1/loxP* #58/Δ69, HM-1/*Lig1/loxP* #68/Δ83 and Δ87). DNA was then digested with *Hin*dIII, electrophoresed through an agarose gel, blotted onto nylon membrane and probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA, corresponding to exons 8 to 27. *Lig1* genotypes of the different samples, derived by PCR, are shown, along with approximate sizes of the restriction fragments detected.


"Floxed" ES Cell Clones
"Deleted" ES Cell Clones
HM-1
#53
#58
#68
Δ69
Δ83
Δ87
10.7
4.3
3.9
2.0
1.8
1.6
1.5
Lig1 genotype
+/+
+/-
+/f
+/f
+/Δ
+/Δ
+/Δ

6.2.5 Expression from the deleted *Lig1* allele is abolished

Cre-mediated recombination of the floxed *Lig1* locus deletes the same five exons as are replaced by an *HPRT* minigene in the targeted *Lig1/PGK-HPRT* and *Lig1*/DWM110 alleles. There was no detectable expression from the targeted allele, but in order to determine the nature of any transcript produced by the deleted *Lig1* allele, northern blot analysis was carried out on samples heterozygous for the deleted allele (Figure 6.2.5). Total RNA was isolated from the same panel of ES cell samples that had been employed during Southern analysis of the structure of the deleted allele. After electrophoresis and transfer onto nylon membrane, the RNA was probed with a *Lig1* cDNA probe.

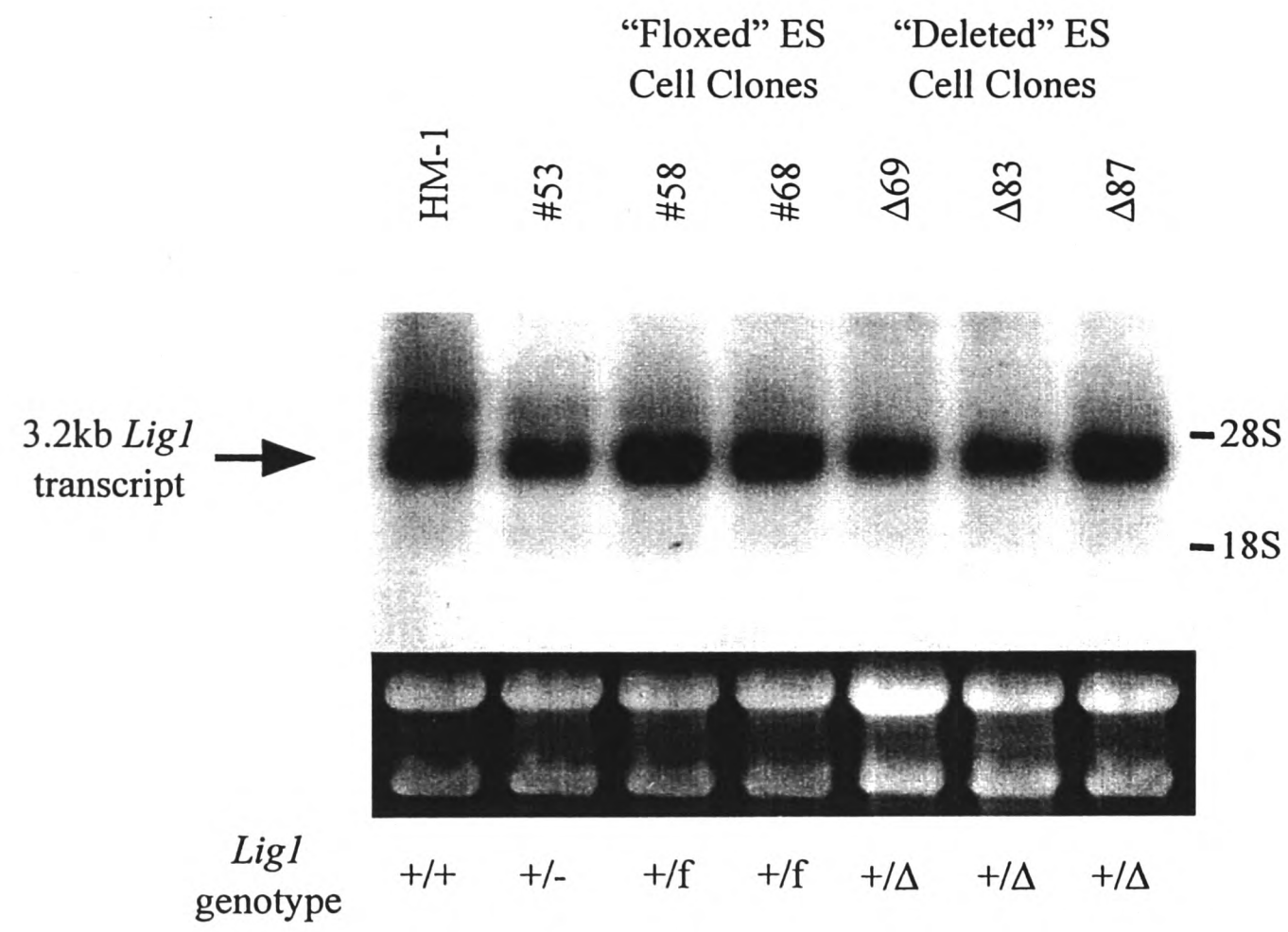
As discussed previously, second step gene targeting restores *Lig1* expression to wild-type levels in samples heterozygous for the floxed allele (section 6.3). A single transcript of wild-type size is present in samples heterozygous for the deleted *Lig1* allele (Δ69 and Δ83). However, the intensity of this band is reduced by approximately 50% compared to the wild-type sample, suggesting that no stable transcript is produced by the deleted allele. The absence of expression would thus indicate the deleted *Lig1* allele to be functionally equivalent to the targeted *Lig1* allele. The band in the lane corresponding to Δ87 appeared to be of equal intensity compared to +/+ and +/f samples.

6.3 Production of mice carrying the floxed *Lig1* allele and a Cre recombinase transgene

Following the confirmation that the floxed *Lig1* allele had the correct structure, was apparently functional, and that Cre-mediated recombination could be induced between the two *loxP* sites *in vitro*, +/f ES cells were used to produce mice carrying the floxed *Lig1* allele. A number of the +/f ES cell clones were injected into host blastocysts, and the resulting chimaeric males were mated with Balb/c females to produce mice heterozygous for the floxed allele. These animals were interbred to derive mice homozygous for the floxed *Lig1* allele (f/f). f/f mice developed normally to adulthood, were fertile, and to date no abnormalities have been observed in these mice. Furthermore, mice which were compound heterozygotes for the floxed and targeted *Lig1* alleles also developed apparently normally (J. Selfridge and D. W. Melton, unpublished data) These data confirmed that the floxed *Lig1* allele is functionally indistinguishable from the wild-type allele.

Figure 6.2.5: Expression from the floxed *Lig1* allele is indistinguishable from wild-type transcript, and expression from the deleted *Lig1* allele is not detectable

Northern blot analysis of ES cell samples carrying the deleted *Lig1* allele. Total RNA was prepared from the original ES cell line (HM-1), one ES cell clone heterozygous for the targeted *Lig1/PGK-HPRT* allele (#53), 2 ES cell clones heterozygous for the floxed *Lig1* allele (#58 and #68), and the 3 ES cell clones which had given positive results from PCR screening for the deleted *Lig1* allele (Δ69, Δ83 and Δ87). Upper, RNA was electrophoresed through an agarose-formaldehyde gel, blotted onto nylon membrane and probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA, corresponding to exons 8 to 27. Positions of the 28S and 18S ribosomal RNA transcripts are indicated flanking the 3.2kb wild-type *Lig1* transcript. *Lig1* genotypes of the different samples are shown. Lower, ethidium bromide staining to verify equal RNA loading.


"Floxed" ES Cell Clones
"Deleted" ES Cell Clones
HM-1
#53
#58
#68
Δ69
Δ83
Δ87
3.2kb *Lig1* transcript
28S
18S
Lig1 genotype
+/+
+/-
+/f
+/f
+/Δ
+/Δ
+/Δ

To test the ability of the floxed *Lig1* allele to undergo Cre-mediated recombination *in vivo,* we crossed f/f mice to transgenic mice expressing Cre recombinase under the control of the ovine β-lactoglobulin (BLG) promoter (Selbert *et al.* 1998). BLG is the major milk protein in ruminants, but is not present in rodent milk. Previously it has been shown that the expression of transgenes under the control of the BLG promoter is largely position independent, and can be restricted exclusively to mammary secretory epithelial cells (Whitelaw *et al.* 1992). A BLG-Cre expression construct was used to produce a number of lines of transgenic mice. Two lines of mice (74 and 93), each carrying multiple copies of the transgene, had high levels Cre expression specific to the mammary gland of lactating females. Hemizygotes from both lines of BLG-Cre transgenic mice were mated to mice homozygous for the floxed *Lig1* allele, and female progeny positive for the Cre transgene and the floxed *Lig1* allele selected for further study.

6.3.1 Conditional deletion of *Lig1 in vivo* was predominantly limited to lactating mammary tissue

Female mice heterozygous for the floxed *Lig1* allele and carrying the BLG-Cre transgene were mated to wild-type male mice. The females were subsequently killed at various time points during mammary gland development, and genomic DNA prepared from the mammary gland and a variety of other tissues. To assay for Cre-mediated deletion of the floxed *Lig1* allele *in vivo*, the PCR reaction specific for the deleted *Lig1* allele was carried out on genomic DNA samples from line 74 mice. The correct product was amplified in mammary and tail tissue samples from lactating females, but not any other tissue (data not shown). We therefore concluded that deletion of the floxed *Lig1* was occurring *in vivo*, and that this occurred in a spatially restricted manner.

6.3.1.1 PCR quantification

In order to estimate the level of Cre-mediated recombination occurring in the different tissues we performed quantitative PCR on genomic DNA samples. Quantitative PCR was carried out according to the same basic protocol outlined in chapter 4, but employing the PCR reaction specific for the deleted *Lig1* allele. The *Lig1* deleted product was amplified as part of a duplex reaction with a pair of primers specific for the wild-type mouse *Ercc-1* gene. 20 cycles of PCR were carried out prior to agarose gel electrophoresis, transfer onto nylon membrane, and hybridisation to α-^{32}P-labelled probes specific for the two PCR products. Levels of radiolabel bound to each PCR product were analysed by phosphorimagery, and the ratio of the two signals calculated.

A standard curve was produced by spiking varying amounts of genomic DNA from one of the ES cell clones heterozygous for the deleted *Lig1* allele (+/Δ) into a wild-type DNA sample, thereby creating a series of samples containing known percentages of the deleted allele. These standards were subjected to PCR quantitation in parallel to the experimental samples, and the ratio between the two signals plotted against the known percentage of the deleted *Lig1* allele to produce a standard curve. In order to estimate the percentage of cells in each tissue sample which had undergone Cre-mediated recombination, experimental signal ratios were compared against the standard curve. Figure 6.3.1 shows a typical standard curve produced.

We estimated the percentage of cells in each tissue sample which had undergone Cre-mediated recombination by means of quantitative PCR for the deleted *Lig1* allele (Table 6.3.1). Line 74 displayed very low levels of Cre-mediated recombination in all non-mammary tissues, the only exception being the sample taken from the tail tip. This sample had previously given a positive signal in a non-quantitative PCR assay for the deleted *Lig1* allele, and quantification revealed that approximately 1% of cells in this sample contained the deleted allele.

Levels of recombination in non-mammary tissues were slightly higher in line 93. Although estimated figures were still below 1% in most tissues, they were over 1% in samples from the salivary gland, skeletal muscle and skin. This suggested that ectopic expression of Cre was marginally higher in line 93 compared to line 74. However, data from both lines were in agreement with published data suggesting that BLG-driven transgene expression is largely position independent (Whitelaw *et al.* 1992), and confirmed data indicating that expression of Cre recombinase was limited to the mammary gland (Selbert *et al.* 1998).

Using quantitative PCR on samples of mammary tissue taken from line 74 mice at various time points during mammary gland development, we were able to determine the time course along which recombination occurred (Table 6.3.2). Samples from virgin mice displayed low levels of recombination in the mammary gland, with approximately 7% of cells containing the deleted *Lig1* allele. This level did not change significantly with age (data not shown), suggesting that the recombination events detected had arisen from a transient burst of Cre expression occurring during early development. Levels of the deleted allele in the mammary gland rose during pregnancy and birth, peaking about 10 days after birth. Cre-expression was thus tightly linked to the lactation status of the mammary gland, and this profile closely paralleled the published kinetics of BLG expression (Harris *et al.* 1991). Between days 10 and 14 of lactation around 70 to 80% of cells in the mammary gland had undergone recombination. The proportion which had not

Figure 6.3.1: A typical standard curve for PCR quantification of the deleted *Lig1* allele

Varying amounts of genomic DNA from the ES cell clone HM-1/*Lig1loxP* #58/Δ69 was spiked into DNA from the wild-type ES cell line HM-1, thereby producing a series of standards containing known percentages of the deleted allele. These standards were used as templates for PCR reactions specific for the *Lig1* deleted allele using primers N2469+V2014. The product from the deleted allele was amplified as part of a duplex reaction with primers 033M+035M, which amplify the wild-type *Ercc-1* gene. This second reaction was used as an independent control for the amount of input DNA. Products were separated by electrophoresis, blotted and hybridised to probes specific for the two PCR products. The relative levels of the two signals were quantified by phosphorimagery and the calculated ratios plotted against the known percentages of the deleted allele. Using this graph it was possible to derive estimates of the fraction of cells in tissue samples which had undergone Cre-mediated recombination.

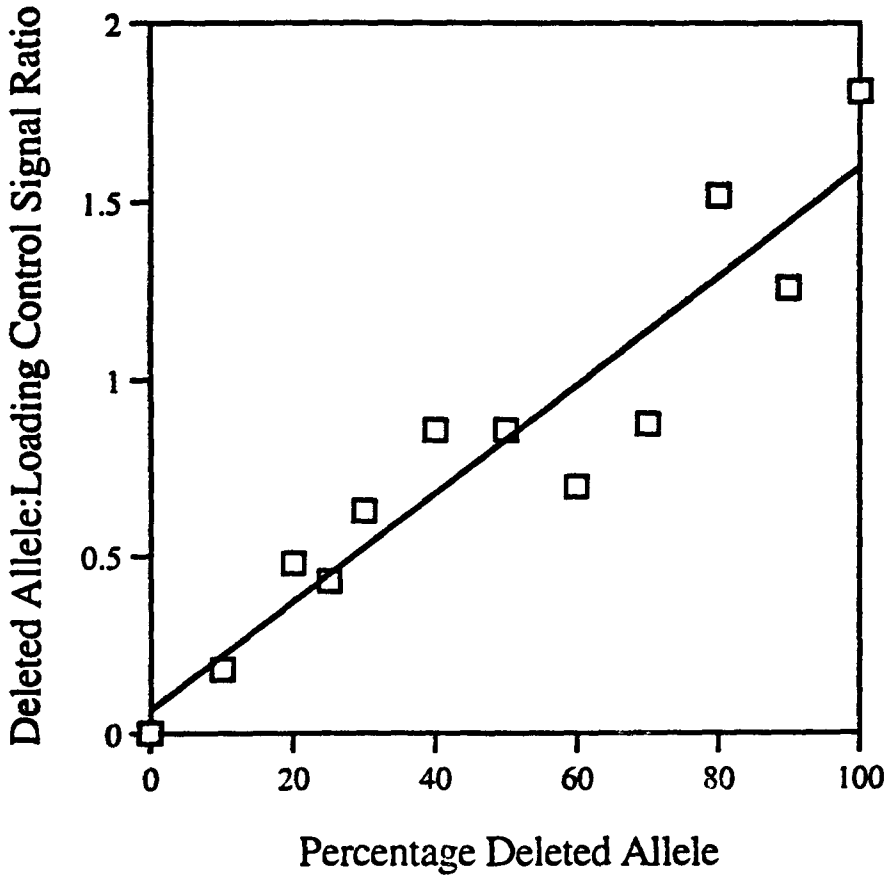

Deleted Allele:Loading Control Signal Ratio
2
1.5
1
0.5
0
0
20
40
60
80
100
Percentage Deleted Allele

Table 6.3.1: Tissue distribution of Cre-mediated recombination in two BLG-Cre transgenic lines

Tissue	Percentage Deleted Allele (%)	
	74 line	93 line
Brain	0.1	0.6
Heart	0.0	0.1
Intestine (Large)	0.1	0.1
Intestine (Small)	0.0	0.1
Kidney	0.0	0.0
Liver	0.1	0.0
Lung	0.0	0.2
Ovary	0.0	0.0
Pancreas	0.1	0.2
Salivary Gland	0.1	2.4
Skeletal Muscle	0.1	1.4
Skin	0.1	1.5
Spleen	N.D.	0.1
Tail	1.1	0.5
Thymus	0.0	0.1

Levels of Cre-mediated recombination in a panel of tissues from two different BLG-Cre transgenic lines (74 and 93) were estimated by quantitative PCR for the deleted *Lig1* allele. The percentage deleted allele is an approximation of the proportion of cells in each tissue which contain the deleted allele.

Table 6.3.2: Levels of Cre-mediated recombination are maximal in the lactating mammary gland

	Percentage Deleted Allele (%)		
	Quantitative PCR		Southern Blotting
Developmental Timepoint	74 line	93 line	74 line
Virgin	7.0 ±5.2	N.D.	15
Gestation (E15)	30.1 ±4.7	N.D.	N.D.
Birth	20.7 ±14.5	N.D.	N.D.
Lactating (10 days)	69.1 ±32.3	38.6 ±28.6	73

Levels of Cre-mediated recombination in mammary tissue of two different BLG-Cre transgenic lines (74 and 93) at different stages of development were estimated by quantitative PCR and Southern blotting. The percentage deleted allele is an approximation of the proportion of cells in the mammary gland that contain the deleted *Lig1* allele. Figures for quantitative PCR represent the average of at least three independent experiments for each timepoint (± standard deviation). N.D., not determined.

undergone recombination (20-30%) most likely reflected non-epithelial cells present in the tissue samples. The BLG promoter will not be activated in fat cells, fibroblasts, lymphocytes etc.. These will therefore lack Cre expression, and hence not undergo recombination.

In addition, quantitative PCR was employed to estimate the frequency of cells containing the deleted *Lig1* allele in ES cell samples. Genomic DNA from the four pools of ES cells into which the Cre-expression plasmid pMC/CreN had been introduced (#58 1°, #58 2°, #68 1°and #68 2°) (see section 6.2) was analysed, along with the ES cell clone which had given a positive signal from the *Lig1* deleted PCR reaction but an incompatible banding pattern on Southern blots (HM-1/*Lig1*/*loxP* #68/Δ87). The calculated frequencies of cells containing the deleted *Lig1* allele are shown in Table 6.3.3. Estimates for pools #58 2° and #68 2° were close to frequencies calculated from colony screening (section 6.2.3, 2.1% and 4.2% respectively). Results for the HM-1/*Lig1*/*loxP* #68/Δ87 indicate that only 1/4 of cells contained the deleted allele.

Table 6.3.3: PCR quantification of the frequency of the deleted *Lig1* allele in ES cell samples

ES Cell Sample	Percentage Deleted Allele (%)
58 1° Pool	2.1 ±0.9
58 2° Pool	3.7 ±0.3
68 1° Pool	1.3 ±0.7
68 2° Pool	3.4 ±1.6
HM-1/*Lig1*/*loxP* #68/Δ87	26.4 ±4.4

Levels of Cre-mediated recombination in ES cell samples were estimated by quantitative PCR. The percentage deleted allele is an approximation of the proportion of cells in each sample that contain the deleted *Lig1* allele. Figures for quantitative PCR represent the average of at least three independent experiments for each timepoint (± standard deviation).

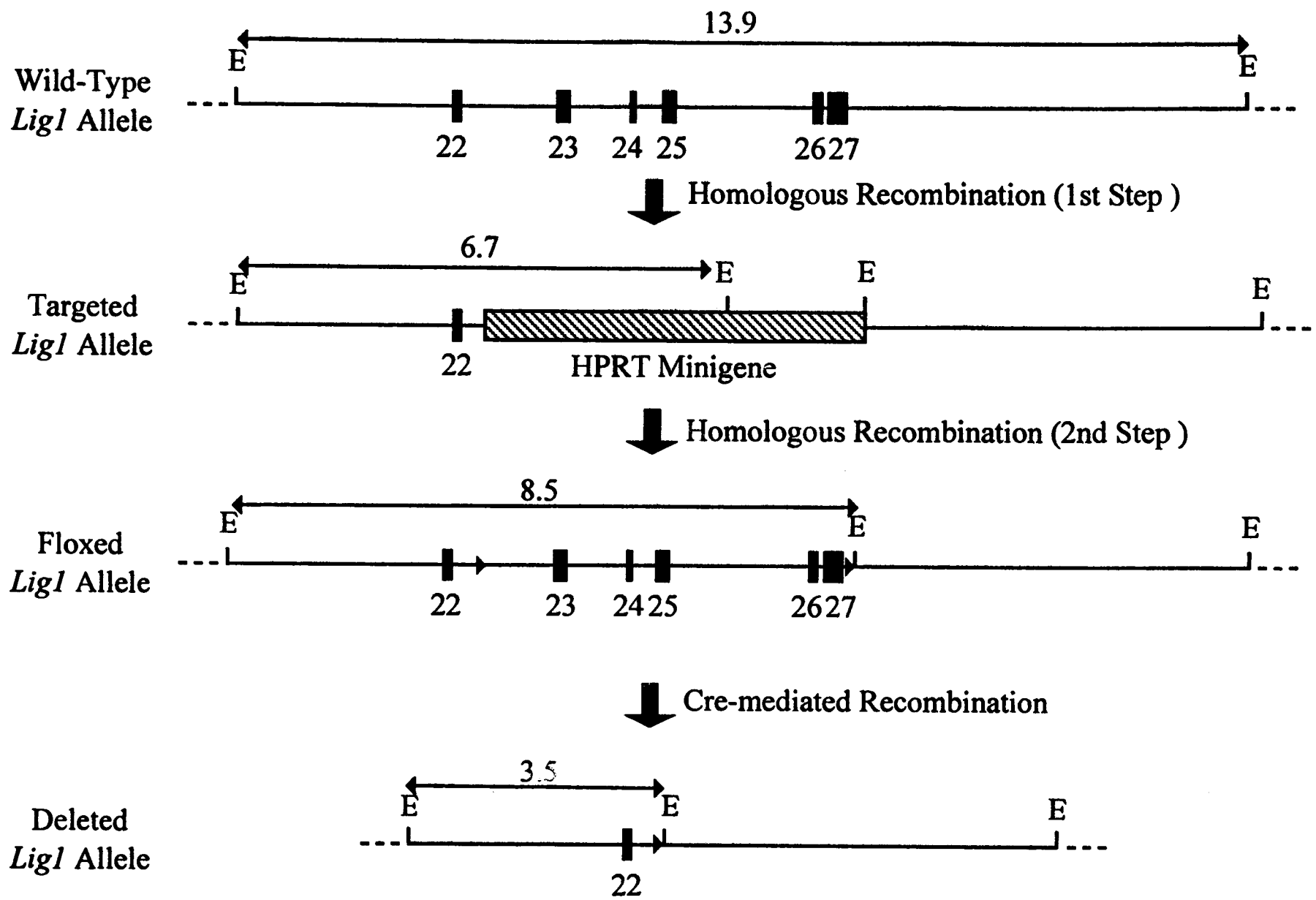
6.3.1.2 Southern blotting

The reliability of quantitative PCR analysis to estimate levels of Cre-mediated recombination was confirmed independently by Southern analysis. Figure 6.3.2 summarises the double replacement gene targeting strategy used to engineer two *loxP* sites into the 3' end of the endogenous *Lig1* locus, and the subsequent effect of Cre-mediated recombination on the floxed locus. The final 6 exons of the wild-type *Lig1* gene are located within a 13.9kb *Eco*RI fragment. The first round of homologous recombination, which replaced exons 23 to 27 with an *HPRT* minigene, also resulted in production of a novel 6.7kb fragment diagnostic of the *Lig1*/DWM110 targeted allele. A second gene targeting event was subsequently used to reconstruct a functional *Lig1* allele and introduce two *loxP* sites into locus. The presence of additional *Eco*RI restriction enzyme cleavage site tagged to the 3' *loxP* sequence permitted the floxed *Lig1* allele to be discriminated from the wild-type allele. Following *Eco*RI digestion, the floxed allele was marked by the presence of an 8.5kb restriction fragment on Southern blots. Subsequent Cre-mediated excision of exons 23 to 27 from the genomic locus reduced the size of this *Eco*RI fragment to 3.5kb. Thus, the deleted *Lig1* allele could be identified by the presence of a unique 3.5kb *Eco*RI fragment on Southern blots. Hence, increasing levels of Cre-mediated deletion of the floxed *Lig1* allele should be reflected by disappearance of the 8.5kb band and the simultaneous appearance of the 3.5kb *Eco*RI fragment.

Figure 6.3.2: Schematic representation of the strategies used to produce the deleted *Lig1* allele

Schematic representation of the 3' end of the wild-type mouse DNA ligase I locus, first step targeted locus, reconstructed 'floxed' locus, and the structure of the locus following Cre-mediated recombination between the two *loxP* sites. In the first targeting step, the last 5 exons of the *Lig1* locus were replaced with an *HPRT* minigene by homologous recombination. A second homologous recombination step was subsequently used to reconstruct a functional allele, but with exons 23 to 27 flanked by two *loxP* sites. Cre-mediated recombination between the two *loxP* sites flanking exons 23 to 27 results in excision of the intervening genomic region, effectively deleting the last five exons of gene.

Drawn to scale. Solid boxes indicate exons (numbered according to the human convention); hatched box, *HPRT* minigene (DWM110); thin lines, *Lig1* introns or flanking regions; solid triangles, *loxP* sites. *Eco*RI restriction sites are marked by E, and the fragments diagnostic for each allele indicated.

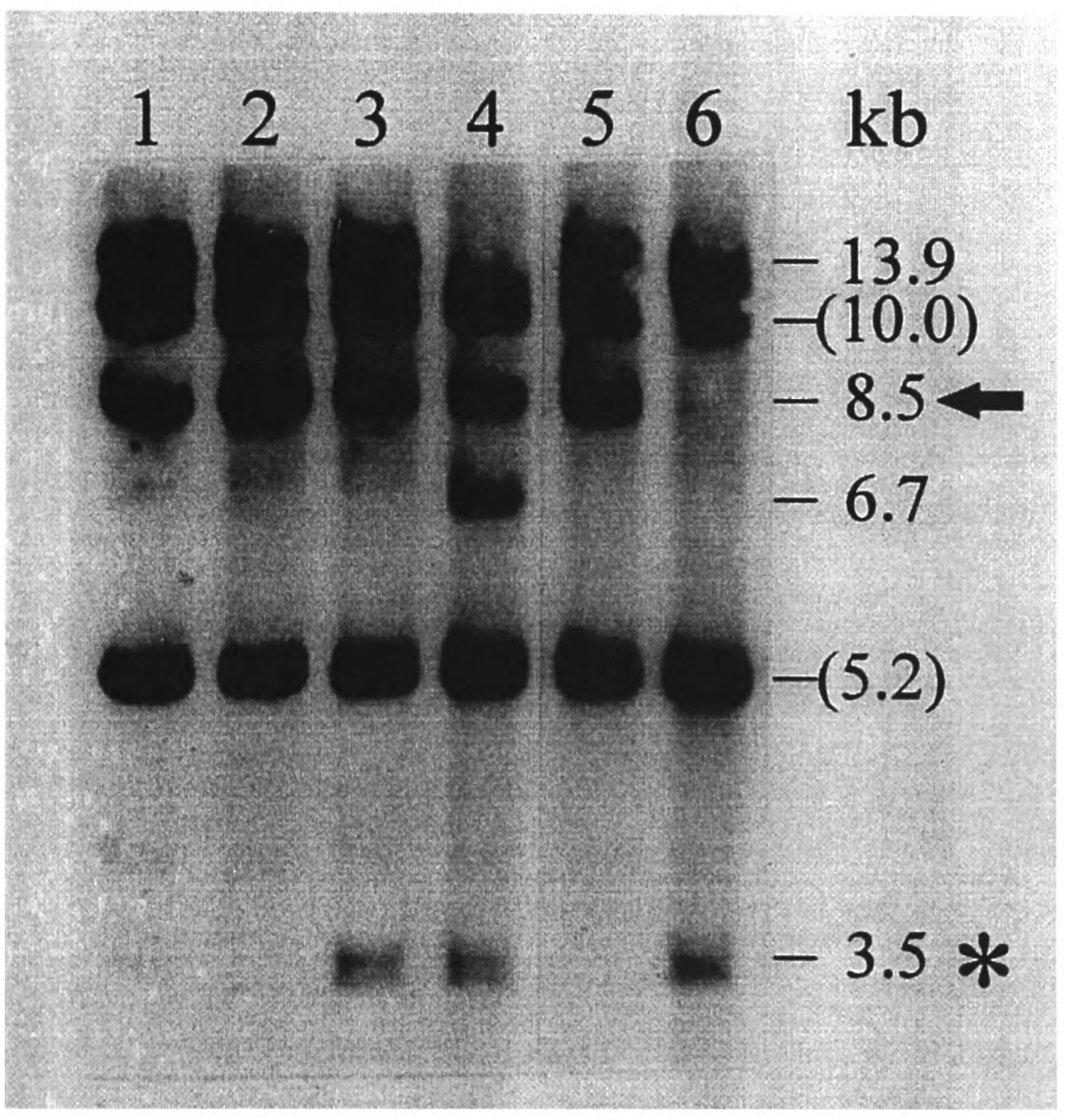

13.9
E
E
Wild-Type
Lig1 Allele
22
23
24 25
2627
Homologous Recombination (1st Step)
6.7
E
E
E
E
Targeted
Lig1 Allele
22
HPRT Minigene
Homologous Recombination (2nd Step)
8.5
E
E
E
Floxed
Lig1 Allele
22
23
2425
2627
Cre-mediated Recombination
3.5
E
E
E
Deleted
Lig1 Allele
22

DNA was prepared from an ES cell clone heterozygous for the floxed allele (#58), an ES cell clone heterozygous for the deleted allele (Δ69) and mammary tissue samples from line 74 mice at various time points during mammary gland development. Following digestion with *Eco*RI, the DNA was subjected to Southern blot analysis (Figure 6.3.3). Using a *Lig1* cDNA fragment as a probe, 4 bands were detected in the +/f ES cell sample (13.9, 10.0, 8.5 and 5.2kb), and 4 bands in the +/Δ ES cell sample (13.9, 10.0, 5.2, and 3.5kb), as predicted from the known structures of the wild-type, floxed and deleted *Lig1* alleles. Samples from the mammary gland two virgin females displayed a pattern of bands almost identical to the +/f sample. A weak signal corresponding to a 3.5kb fragment could be detected one of the virgin samples, but not in the other. However, this fragment specific for the deleted *Lig1* allele was present at high levels in both samples from lactating females, and was accompanied by a marked reduction in the intensity of the signal from the 8.5kb fragment.

Phosphorimagery was then used to quantify the level of the 3.5kb fragment relative to the 5.2kb band. The 5.2kb fragment was common to all *Lig1* alleles, and hence acted as an independent control for the amount of DNA loaded in each sample. The calculated ratio of the two signals in each lane were compared against the ratio of signals in the +/Δ ES cell sample (known to have undergone recombination in 100% of cells) to derive estimates of the fraction of cells in tissue samples which had undergone Cre-mediated recombination. The results of this analysis are shown in Table 6.3.2, and are in close agreement with those obtained by quantitative PCR. Thus Southern blotting provided independent confirmation that BLG-driven Cre expression results in conditional deletion of the floxed *Lig1* allele in a spatially and temporally controlled manner, being predominantly limited to the lactating mammary gland.

Figure 6.3.3: Detection the deleted *Lig1* allele by Southern blotting

Southern blot analysis of Cre-mediated recombination in mammary tissues from BLG-Cre transgenic mice (line 74). Genomic DNA was prepared from one ES cell clone heterozygous for the floxed *Lig1* allele (HM-1/*Lig1/loxP* #58), one ES cell clone heterozygous for the deleted *Lig1* allele (HM-1/*Lig1/loxP* #58/Δ69), and mammary tissues from line 74 mice heterozygous for the floxed *Lig1* allele. The developmental status of each of the mammary samples was: lane 1; virgin 16 weeks old, lane 2; virgin 6 weeks old, lane 3; mother 14 days lactation, lane 4; mother 10 days lactation. Lane 5 contains DNA from the +/f ES cell clone #58, while lane 6 represents the +/Δ ES cell clone Δ69. DNA was then digested with *Eco*RI, electrophoresed through an agarose gel, blotted onto nylon membrane and probed with a 2.1kb *Bam*HI/*Eco*RI fragment of the *Lig1* cDNA, corresponding to exons 8 to 27. The sizes of each hybridising fragment (in kb) are shown. Sizes in parentheses indicate fragments derived from the 5' and central region of the *Lig1* locus which are common to all four *Lig1* alleles. The 8.5 kb fragment specific for the floxed allele is indicated by an arrow, while the novel 3.5kb fragment generated by Cre-mediated recombination is marked by an asterisk. The sample in lane 4 was taken from a mouse that was a compound heterozygote for the targeted *Lig1* allele and the floxed *Lig1* allele. The 6.7kb fragment evident in this lane is derived from the *Lig1/PGK-HPRT* targeted allele.


1 2 3 4 5 6
kb
— 13.9
—(10.0)
— 8.5
— 6.7
—(5.2)
— 3.5 *

CHAPTER SEVEN:

DISCUSSION

In order for the genome to be successfully maintained and propagated, cells must be able to join DNA molecules. Enzyme-catalysed DNA joining (ligation) is therefore one of the basic activities required for cell viability. At least four biochemically distinct DNA ligase activities have been identified within mammalian cells, and the genes encoding three have been cloned (Barnes *et al.* 1990, Wei *et al.* 1994). The three proteins are designated DNA ligase I, III and IV, and of these, DNA ligase I is the most extensively studied.

As reviewed in chapter 1, published evidence implicates DNA ligase I in both DNA replication and DNA repair. A wealth of data supports the premise that DNA ligase I is central to lagging strand synthesis. The data can be summarised as follows:

- Gene expression and protein activity are both upregulated in proliferating cells
- The protein can be isolated from replication complexes, and has been proven to interact with PCNA
- The protein localises to replication foci in the nuclei of replicating cells
- Human cells with reduced DNA ligase I activity display an abnormal profile of replication intermediates
- The purified protein has been used to reconstitute DNA replication *in vitro*, and complements the replication disturbances of ligase-deficient humans, *S. cerevisiae*, and *E. coli* mutants

DNA ligase I is also thought to function in a number of DNA repair pathways:

- Gene expression and protein activity are upregulated upon UV irradiation
- The protein has been isolated from a complex of proteins involved in NER, and the purified protein has been used in reconstitution of NER *in vitro*
- The protein has been isolated from a complex of proteins involved in BER, and the purified protein has been used in reconstitution of BER *in vitro*
- Activity of the protein is stimulated by addition of Ku, and the purified protein has been used in the reconstitution of V(D)J recombination *in vitro*
- DNA ligase I-deficient human cells are sensitive to a wide variety of DNA damaging agents

Taken together, these data indicate that DNA ligase I functions in a number of cellular processes. It was consequently proposed that DNA ligase I is essential for mammalian cell viability:

> *"It is difficult to envisage that a cell can survive without the predominant DNA ligase."*
>
> *Hoeijmakers and Bootsma 1992*

Apparent proof of this hypothesis came in 1995 from work by Weaver and co-workers. They were unable to isolate cells lacking DNA ligase I *in vitro*, and from this concluded that DNA ligase carries out essential functions within mammalian cells (Petrini, Xiao and Weaver 1995).

Loss of function of a single *LIG1* allele has no effect on phenotype (i.e. loss of function mutations are recessive to wild-type). A single functional DNA ligase I allele must be sufficient for all cellular processes. However, not all mutations will result in complete loss of function. The product of certain alleles may retain enough residual ligase activity to fulfil the cellular requirements, albeit in a reduced manner. Providing the protein also retains the ability to interact normally with other cellular proteins, compound heterozygotes of a loss-of-function allele and a partially functional allele could be viable. Nevertheless, the high requirement for DNA ligase activity in dividing cells places stringent requirements on the minimum level of activity sufficient for survival. Most mutants will fall below such a threshold, and will in effect resemble non-functional alleles. Only a limited subset of mutant alleles will thus be compatible with cellular viability. It has been hypothesised that the 46BR patient represents such an occurrence.

Given the presumed inviability of cells homozygous for complete loss-of-function alleles, it was rationalised that the only way to produce a viable experimental model of DNA ligase I-deficiency was to reproduce the 46BR mutations. Double replacement gene targeting is an ideal method by which to introduce analogous point mutations into the mouse genome. A two step scheme was devised to introduce the Arg771→Trp change into exon 24 of the endogenous locus. The first step, introduction of an *HPRT* marker gene into the *Lig1* locus, had the additional advantage of deleting the last 5 exons of the gene, thereby creating a 'knockout' allele.

7.1 Embryos lacking DNA ligase I are viable

Using conventional techniques we were able to generate mice heterozygous for the knockout *Lig1* allele. These mice were phenotypically normal, but when interbred, no pups homozygous for the targeted allele were found. It was therefore concluded that DNA ligase I is essential for normal development. Contrary to expectation, when embryos were collected from heterozygote matings and genotyped, live embryos homozygous for the targeted mutation were found.

The PCR reactions used to follow the wild-type and targeted *Lig1* alleles, and to genotype embryos, were sensitive and robust. False negatives were not a problem under optimal reaction conditions (data not shown). Conversely, false positives were not observed from the targeted reaction. Providing care was taken to avoid contamination of yolk sac samples with maternally-derived tissue, false positives were

also not observed from the wild-type reaction (data not shown). These assertions are corroborated by the ratios of embryo genotypes observed from heterozygote matings (e.g. E10.5, Table 3.3.1). No significant deviations from the expected mendelian ratios were observed in samples from early embryos. Similarly, Southern blotting of random embryos confirmed the genotypes determined by PCR. We therefore concluded that embryos genotyped as -/- by PCR were indeed homozygous for the targeted allele. Moreover, Southern blots of -/- samples were all consistent with the predicted pattern of restriction fragments produced from the targeted allele. We thus inferred that the targeted *Lig1* locus had indeed undergone the correct homologous recombination event, and exons 23 to 27 had been deleted.

In order to ascertain the nature of any mRNA produced from the targeted allele, northern analysis of *Lig1* expression was carried out. Transcript levels were reduced by approximately 50% in +/- samples compared to wild-type, and no signal could be detected in -/- embryos or cell lines. The lack of any kind of detectable transcript from the targeted allele was surprising given the targeting strategy employed. Replacement of exons 23 to 27 by a selectable marker cassette had been expected to result in the production of some sort of truncated or *Lig1/HPRT* fusion transcript (e.g. Porter *et al.* 1995). The complete absence of detectable expression could have two possible explanations; rapid transcript degradation, or no transcription from the targeted allele.

In addition to replacing five exons, homologous recombination removed the polyadenylation signal. Addition of a poly-A tail is known to be essential for transcript stability (Jackson and Standardt 1990), and it had proved difficult to express *Lig1* cDNA clones truncated at the 3' end (Jessop 1995). It is therefore possible that *Lig1* transcripts lacking a poly-A tail might be unstable, and hence be rapidly degraded. Alternatively, transcription from the targeted *Lig1* locus may not occur. It has been shown that expression from the *HPRT* minigene within the targeted *Lig1* locus is highly unstable (Melton, Ketchen and Selfridge 1997). It was hypothesised that this might reflect an alteration to the locus associated with the targeting event itself, for example a change in chromatin structure. Deletion of a structural control element found within the 3' end of the *Lig1* locus could lead to transcriptional silencing of the region. Although such a switch might be limited to the *Lig1* locus only, alteration from an 'open' to a 'closed' chromatin structure would be likely to have longer range effects. It should thus be possible to test this alternate hypothesis by investigating whether transcription from other genes is affected by targeting. Other investigators have observed that insertion of a marker gene into a locus can alter expression of surrounding genes. These results were attributed to the nature of the selectable marker employed. The fact that expression from the deleted *Lig1* allele (Δ)

was also abolished suggests that deletion of sequences is the crucial factor in this situation.

Whatever the cause, no transcription could be detected from the targeted *Lig1* allele by northern blot analysis. Similarly we were unable to detect any kind of *Lig1*-derived polypeptide in -/- samples by western blot analysis. Western analysis was performed using antibodies raised against purified bovine DNA ligase I, and also using antibodies raised against the C-terminal conserved peptide. Both produced identical results. Because were able to detect levels of DNA ligase I down to as little as 1% of normal, we were confident that any *Lig1*-derived proteins expressed in -/- samples would have been detected. No intact, truncated, or fusion protein of any size could be detected, and thus we concluded that expression from the targeted *Lig1* allele is abrogated, and that DNA ligase I is absent from -/- embryos.

Confirmation of this result was achieved using an assay for DNA ligase activity. All DNA ligases share the ability to form DNA ligase-AMP intermediates, and by using radioactively labelled ATP it is possible to visualise the adenylated products. When we carried out this assay on whole cell extracts from -/- embryos, no adenylation activity correlating to DNA ligase I was observed. No additional adenylated products were detected compared to the wild-type controls (data not shown). Whilst acknowledging the limited sensitivity of this technique, we concluded that there was no DNA ligase activity attributable to the targeted *Lig1* allele.

Lindahl and co-workers have suggested that survival of -/- embryos may arise as a consequence of low levels of a truncated *Lig1*-derived protein. They cite work by Subramanya *et al.* reporting that a single domain of bacteriophage T7 DNA ligase retains some enzymatic activity (Subramanya *et al.* 1996).

> *"The resulting truncated mRNA, which still encodes the active site of the enzyme was reported to be unstable. However, low levels of this mRNA might generate a large fragment of DNA ligase I, that by analogy with a homologous fragment of bacteriophage T7 DNA ligase, could have enough residual activity to allow survival of null embryos, at least until very rapid cellular proliferation is required during early erythropoiesis."*
>
> *Mackenney, Barnes and Lindahl 1997*

Extrapolating this result to the theoretical product of our targeted *Lig1* allele, Mackenney, Barnes and Lindahl suggested that low levels of a truncated DNA ligase I polypeptide may retain enough activity to permit cell survival during early development. The implication is that the targeted *Lig1* allele we have generated is not a true knockout allele, and it is this which allows -/- cells to survive. We would

counter this argument by citing previous work carried out in the laboratory of Dr Lindahl. Kodama, Barnes and Lindahl expressed human DNA ligase I in *E. coli*, and used systematic mutagenesis to define the regions of the protein important for catalytic function. Engineered constructs were assayed for function by the ability to complement *E. coli* DNA ligase mutants, and the altered proteins were also assayed for DNA ligase activity by formation of adenylated intermediates *in vitro*. It was found that the majority of the catalytic domain was absolutely required for enzyme activity. Removal of more than 16aa from the C-terminus of the human DNA ligase I rendered the protein enzymatically inactive (Kodama, Barnes and Lindahl 1991). Our targeting strategy replaced exons 23 to 27, in effect deleting the final 175aa of the protein. Any protein produced from the targeted *Lig1* allele should therefore be catalytically inactive. Consequently, we believe that the model proposed by Mackenney, Barnes and Lindahl is invalid. Given the high levels of DNA ligase I activity normally found in proliferating cells, and the conspicuous absence of any detectable *Lig1*-derived protein in -/- cells, we find it inconceivable that an undetectable amount of a theoretical protein that should have no activity could meet the catalytic requirements of most proliferating cells.

Levin and Tomkinson have suggested that the viability of -/- embryos may alternatively be explained by developmental changes in DNA ligase expression (Levin and Tomkinson 1997). They propose that another DNA ligase complements for the absence of DNA ligase I during early development. The presence of DNA ligase I only becomes essential later in development, and it is at this point which -/- embryos die.

We have not addressed this hypothesis directly, but circumstantial evidence would suggest that this is not true. As with other cell types, ES cells proliferating *in vitro* contain high levels of *Lig1* expression, alluding to the presence of high levels of the protein in very early development. DNA ligases III and IV are expressed at much lower levels (data not shown). Studies of DNA ligase I expression have shown that DNA ligase I expression is switched on very soon after fertilisation of the egg (reviewed by Signoret and David 1986). Later in development, as cells differentiate, a switch between the forms of DNA ligase does occur. However, this leads to the replacement of *Lig1* by another DNA ligase rather than vice versa. Thus the available data indicates that DNA ligase I represents the predominant ligase activity in early eukaryotic development. Moreover, cells homozygous for the targeted allele are viable in non-embryonic environments. Fibroblast cell lines have been isolated from -/- embryos (chapter 5), and -/- cells were able to rescue lethally-irradiated mice and to repopulate the adult haematopoietic system (chapter 4). Recent data using Cre-mediated deletion of *Lig1* has also indicated that -/Δ cells lacking a functional *Lig1*

allele can be generated *in vivo*. Taken together these data suggest that the viability of early -/- embryos cannot be explained simply in terms of a developmental changes in DNA ligase expression. Whether *Lig1* is required specifically for certain developmental events will be discussed further in section 7.4, but the fact that -/- embryos and cells can be isolated indicates that, contrary to expectations, DNA ligase I-deficient cells are viable *in vivo*. A previously unsuspected redundancy must therefore exist among the mammalian DNA ligases.

7.2 DNA ligase I is not essential for mammalian cell viability

Our data indicates that DNA ligase I-deficient cells are viable. This conflicts with conclusions drawn by Weaver and co-workers from studies of ES cells. Having carried out gene targeting of the *Lig1* locus in ES cells, Petrini, Xiao and Weaver were unable to isolate cells lacking DNA ligase I *in vitro* (Petrini, Xiao and Weaver 1995). This led to the conclusion that DNA ligase I is essential for cell viability. We believe that the contrast between this conclusion and our data arises mainly from an inherent difficulty in isolating ES cells homozygous for a disabling mutation *in vitro*. In order to enrich for cells which have undergone the desired recombination events, selection must be applied. Selection typically works by killing undesired cells via a toxic metabolite. Given the relative infrequency of targeting or gene conversion events, high levels of enrichment are required, and hence high selective pressures have to be applied. Whilst the desired cells will theoretically be resistant to killing, the drug may still affect the viability of these cells relative to non-selective conditions. It is thus not unfeasible that drug resistant cells with a pre-existing survival handicap will not be able to form colonies in culture when selection is applied. The handicap itself may not be lethal, and hence cells will be viable *in vivo*, but in the context of selection in low cell density conditions, cells may not proliferate. As a result, it would be impossible to isolate such cells *in vitro*. In support of this hypothesis, we cite data generated from targeting of the *Ercc-1* locus. Production of ES cells homozygous for a knockout mutation at the *Ercc-1* locus has not been achieved to date, even though *Ercc-1*-deficient mice were described in 1993 (McWhir *et al.* 1993, J. Selfridge and D.W. Melton, unpublished data).

The failure of Petrini, Xiao and Weaver to isolate cells homozygous for a targeted *Lig1* allele may also have been influenced by the targeting strategy they employed. They used an replacement gene targeting strategy to introduce a *neo* selectable marker cassette into the middle of exon 17. The cassette replaces half of exon 17 and all of exons 18 and 19, in effect deleting sequences encoding the active site of the protein. This region has previously been demonstrated to be essential for enzyme activity (Kodama, Barnes and Lindahl 1991), and thus any protein produced

should be enzymatically inactive. However, while the product of the targeted allele would have no catalytic activity, we believe that an aberrant polypeptide is produced from the targeted allele, and that this affects normal cellular processes in a deleterious manner.

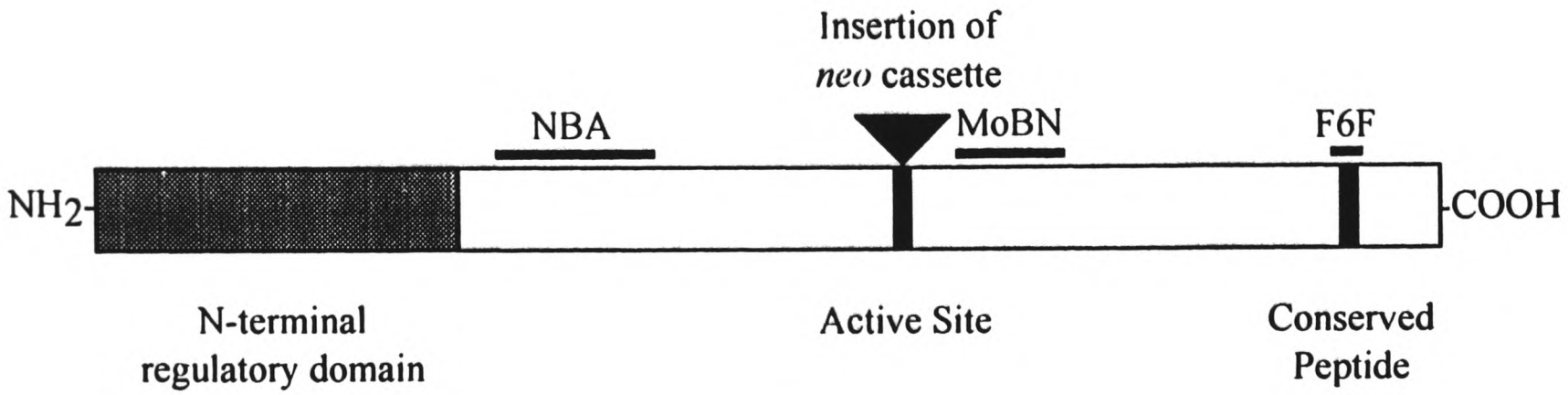
Given that the *neo* cassette was inserted directly into an exon, we would expect that some sort of truncated or fusion peptide would be produced by the targeted allele. Investigation of ES cells heterozygous for the targeted allele detected only the product of the wild-type allele. However, the "MoBN" *Lig1* cDNA fragment used in northern blot analysis corresponds only to exons 18 to 21 (depicted in Figure 7.2.1). The exons are downstream of the point of insertion of the *neo* module, and will therefore not detect a truncated or fusion transcript generated from the *Lig1* region 5' to the marker gene. Similarly, the antibody used for western analysis of cell extracts (F6F) was raised against a peptide close to the C-terminus of the wild-type protein (see Figure 7.2.1). The antiserum will therefore not recognise *Lig1*-derived peptides lacking the C-terminus of the wild-type protein. Investigation carried out using these molecular probes would therefore not detect if a polypeptide was generated from the targeted allele. Thus the published data does not rule out the presence of a truncated, enzymatically inactive polypeptide derived from the targeted allele.

While Petrini, Xiao and Weaver state that ES cells heterozygous for the targeted allele were indistinguishable from wild-type controls, examination of the data may indicate otherwise. Heterozygous cells show a marginal increase in the percentage of newly replicated DNA present as small fragments, but also show increased sensitivity to killing by 3AB. One interpretation of these results might be that a polypeptide produced from the targeted *Lig1* allele interferes with cell function. Though catalytically inert, the peptide would still retain the N-terminal sequences known to be involved in protein-protein interactions. It is feasible that the polypeptide could interact with components of the replication and repair machinery, and thereby impair the normal functioning of these complexes. Disruption would be represented by decrease in the measurable replication and repair parameters. It is thus possible that the targeted allele may function as a dominant-negative mutation, and put cells at a survival disadvantage. Failure to isolate cells homozygous for the targeted mutation may hence occur because the targeting strategy significantly compromises cell viability.

Petrini, Xiao and Weaver reported that they were able to isolate ES cells homozygous for the targeted *Lig1* allele if they introduced a *LIG1* cDNA construct. Cell lines rescued using a cDNA lacking 60 nucleotides from exon 2 were sensitive to

Figure 7.2.1: Schematic representation of the *Lig1* cDNA

The relative locations of certain features within the *Lig1* cDNA are shown schematically. Numbers indicate locations within the coding sequence, which is numbered according to the published sequence (Jessop and Melton 1995). Information about the binding sites of cDNA probes NBA and MoBN are taken from Petrini, Xiao and Weaver 1995, as is the relative location of the peptide recognised by the polyclonal antiserum F6F.


Insertion of
neo cassette
NBA
MoBN
F6F
NH2-
-COOH
N-terminal
regulatory domain
Active Site
Conserved
Peptide

3AB, but only displayed minor perturbations in DNA replication. These data were interpreted as evidence that the missing region was important for interaction of DNA ligase I with repair proteins. Whilst this is a plausible explanation, if a truncated protein is being produced from the targeted allele, an alternative interpretation can be made. The apparent difference between the efficacy of replication and repair processes in these cells may actually reflect the effects of the inactive protein. Parallels can be drawn with 46BR cells. Although 46BR cells do display an altered profile of replication intermediates, they are hypersensitive to the effects of 3AB treatment. It is thus possible that an inactive or partially active form of DNA ligase I will have differential effects on the different cellular processes. Employing such cells as a system in which to test the effects of mutations on DNA ligase I protein function may thus not give unequivocal results. It would clearly be advantageous to express altered forms of DNA ligase I in cells devoid of any other form of the protein. In this respect, cells isolated from our -/- embryos, which can be proven not to contain any *Lig1*-derived polypeptide, would be a better model system in which to test the effects of mutations on protein function.

7.3 DNA ligase I is required for foetal liver erythropoiesis

The finding that embryos homozygous for the targeted *Lig1* allele are viable was unexpected. We demonstrated that expression from the targeted allele was abrogated, and that DNA ligase I was absent from -/- embryos. Furthermore, two completely independent strains of *Lig1* targeted mice displayed exactly identical phenotypes (chapter 3 and unpublished data). When second step gene targeting was used to reconstruct a functional *Lig1* allele, the resulting homozygous mice were phenotypically normal (J. Selfridge and D.W. Melton, unpublished data). We were thus able to rule out the possibility that the phenotype and molecular characteristics of -/- samples had arisen as a result of artifactual changes to the ES cell genome. We were certain that the phenotype observed is the result of targeted alteration of the *Lig1* locus.

Embryos lacking DNA ligase I developed normally until mid-term, and were indistinguishable from their wild-type counterparts until day 11.5 of development. Thereupon acute anaemia developed. At E11.5, all -/- embryos were noted to be paler and smaller than normal. As development proceeded, the differences became exacerbated, so that by E15.5 the two genotypes were strikingly different. Appearance of the anaemia phenotype correlated very closely with developmental switch in the location of blood cell production (haematopoiesis).

The mature haematopoietic system comprises of eight mature cell lineages. All eight lineages are ultimately derived from a common pluripotent precursor or stem

cell (HSC). These cells have two main characteristics; the capacity for apparently unlimited self-renewal, and the ability to differentiate into all haematopoietic lineages. HSCs are separated from the mature cells of the peripheral blood by a series of intermediates of increasingly limited developmental potential, the most restricted being committed to a single lineage and having limited capacity for self-renewal. Figure 7.3.1 shows a schematic representation of the genealogy of the different haematopoietic lineages.

While the bone marrow is the predominant focus of blood cell production in the adult, haematopoiesis is of a migratory nature during ontogeny (Figure 7.3.2). In the mouse embryo, blood forming islands develop in the yolk sac at E7. These arise from mesodermal cells (hemangioblasts) whose precursors originate from the primitive streak region of the blastoderm and migrate to a position in the yolk sac in contact with extraembryonic endoderm. By the time circulation is established between the yolk sac and embryo, the blood islands are connected by a network of capillaries, which evolve into 'erythropoietic sinusoids'. Yolk sac haematopoiesis is thus intravascular, and is predominantly limited to the production of primitive nucleated erythrocytes (10μm diameter).

Recent work by Medvinsky and Dzierzak has identified a region within the body of the embryo also involved in early embryonic haematopoiesis (reviewed by Dzierzak and Medvinsky 1995). The aorta-genital ridge-mesonephros (AGM) region has been shown to contain haematopoietic progenitor cells between E8 and E10. The precise relationship between the AGM region, yolk sac, and liver remains to be delineated, but a limited number of HSCs migrate from either the AGM or yolk sac to colonise the foetal liver around day 9 of development.

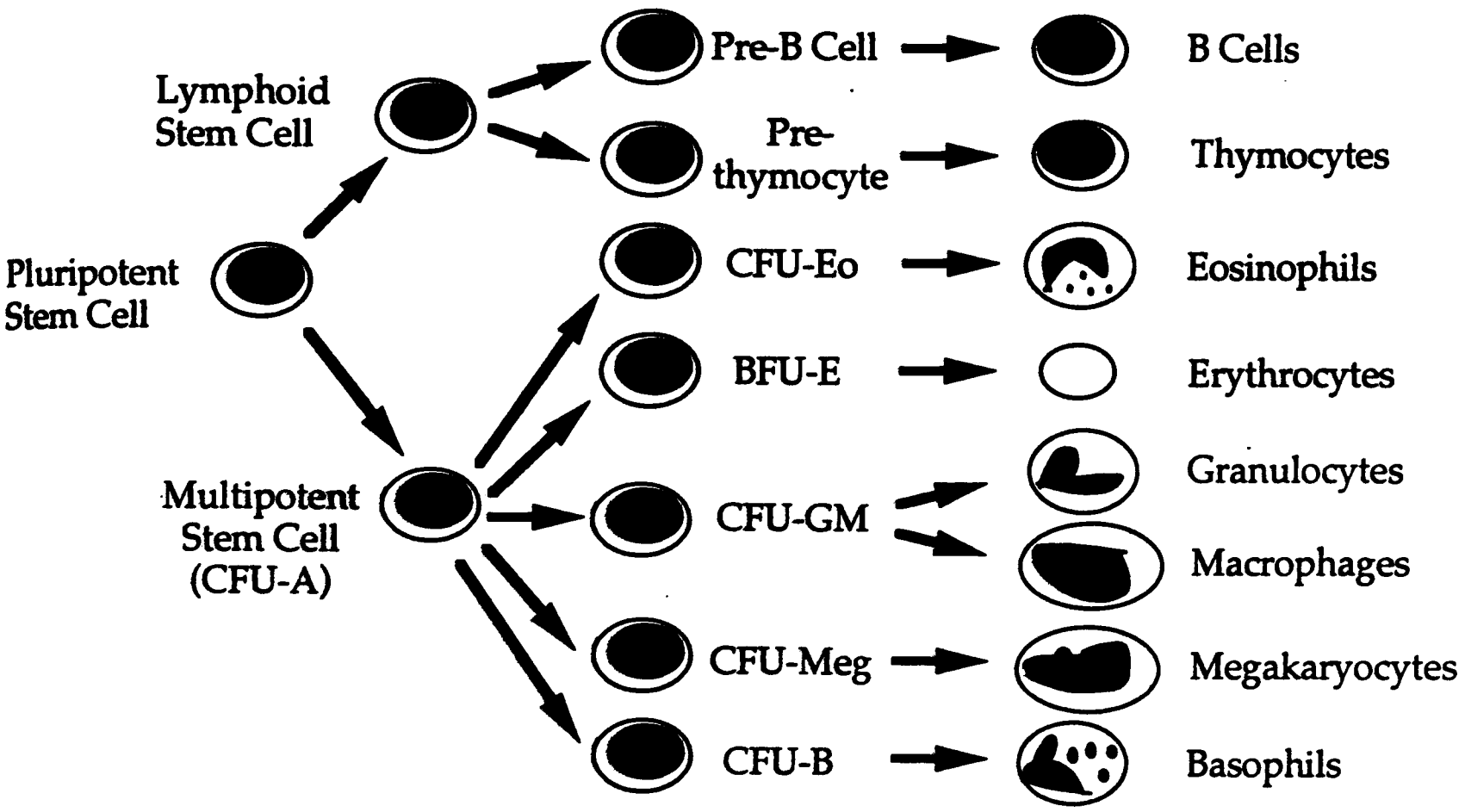
At E9, the liver rudiment begins to form, and on the same day the first erythroid progenitor cells become visible within this structure. From E11.5 onwards the liver is the major site of blood production in the developing embryo, and it remains so until shortly after birth. During this period, erythroid cells can comprise up to 70% of the total liver mass (Metcalf and Moore 1971). Haematopoiesis in the foetal liver is distinct from that occurring in the yolk sac in several ways: blood cell production is not limited to erythropoiesis, the erythrocytes do not mature in a synchronous manner, and production is responsive to stimulation by erythropoietin (Epo). However, the most visible difference is the production of mature enucleated erythrocytes (8μm diameter). Thus, as development proceeds, and haematopoiesis switches from the yolk sac to the foetal liver, the proportion of nucleated cells in the peripheral circulation diminishes.

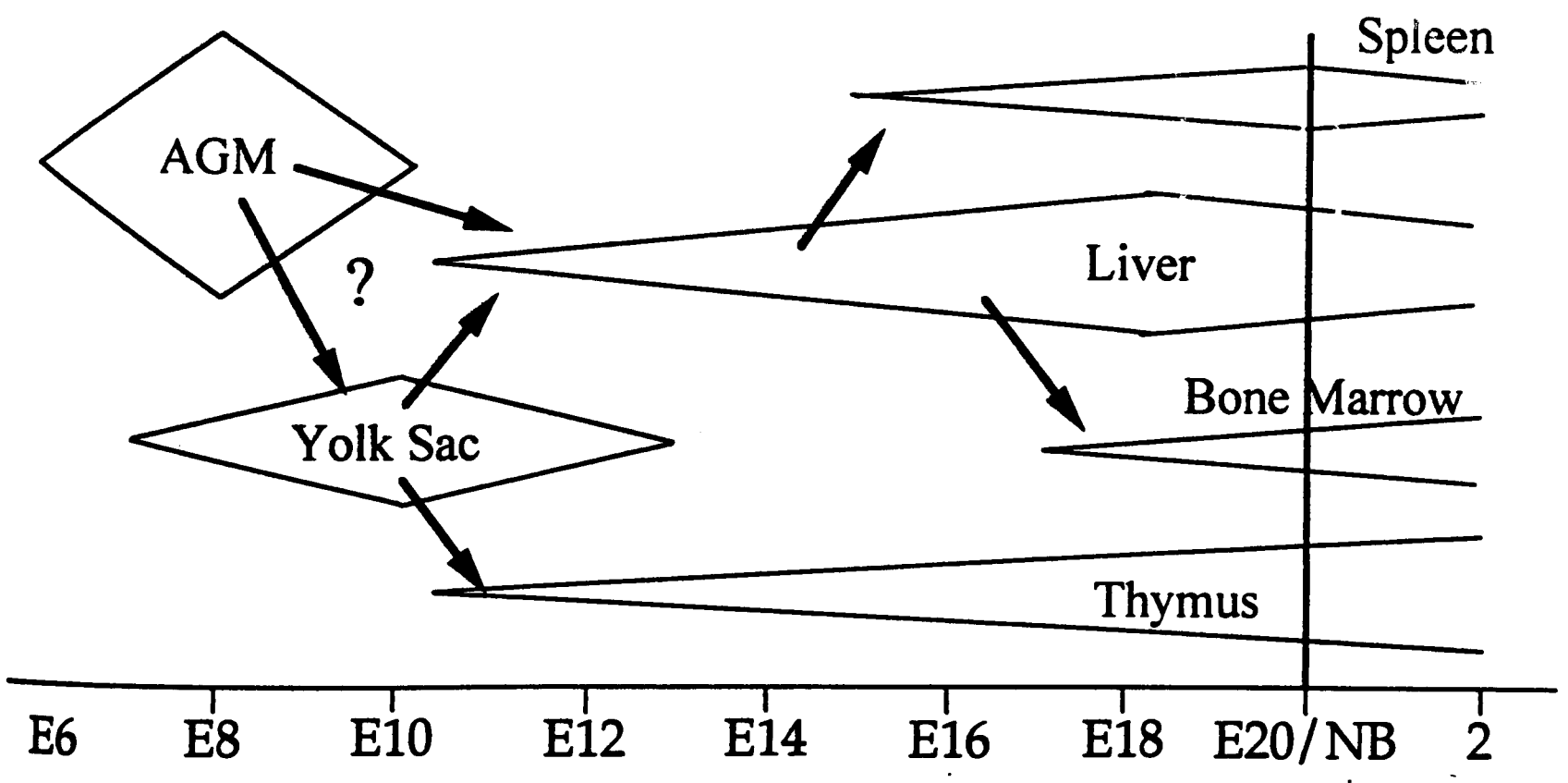
Figure 7.3.1: Schematic representation of the haematopoietic system

A depiction of the relationship of the eight mature haematopoietic cell lineages to progenitor cells. The lineages are identified on the basis of the terminally differentiated cell type they produce. For non-lymphoid lineages, the in vitro assays for committed precursor cells are indicated. The exact correlation between cells which produce colonies in the CFU-A assay and the multipotent progenitor cell remains unclear.

Figure 7.3.2: Schematic representation of the loci of haematopoiesis during mouse embryonic development

A depiction of the migration of haematopoiesis in the developing mouse embryo. Arrows indicate the possible migration pathways of HSCs from one organ to another, while numbers on the x-axis indicated the approximate timing of events during gestation. Question marks indicated the uncertain relationship between the AGM region, the yolk sac and the foetal liver.


Lymphoid Stem Cell
Pluripotent Stem Cell
Multipotent Stem Cell (CFU-A)
Pre-B Cell
Pre-thymocyte
CFU-Eo
BFU-E
CFU-GM
CFU-Meg
CFU-B
B Cells
Thymocytes
Eosinophils
Erythrocytes
Granulocytes
Macrophages
Megakaryocytes
Basophils


AGM
?
Yolk Sac
Spleen
Liver
Bone Marrow
Thymus
E6
E8
E10
E12
E14
E16
E18
E20/NB
2

Examination of embryonic blood samples revealed a deficiency of mature enucleated erythrocytes in -/- embryos. The percentage of nucleated erythrocytes in the peripheral circulation did not show the normal decrease characteristic of the shift from the yolk sac to the foetal liver. Blood from wild-type embryos consisted predominantly of enucleated liver-derived erythrocytes. In contrast samples from -/- embryos displayed a high number of nucleated cells, and those enucleated cells present tended to be abnormal. It was hence concluded that foetal liver erythropoiesis was disrupted. Yolk sac haematopoiesis was, however, apparently normal. -/- embryos displayed no visible phenotype up until the switch to foetal liver haematopoiesis whereas mice deficient in yolk sac erythropoiesis die before E11 (e.g. *GATA-2* knockout mice, Tsai *et al.* 1994). Inspection of the peripheral blood and yolk sac histology of E10.5 embryos revealed no significant difference from wild-type.

In contrast, examination of livers from -/- embryos after E11.5 revealed striking differences. Total cell numbers were reduced at E11.5, and the reduction became more pronounced with time. E15.5 -/- livers lacked extensive cellularization and contained no readily identifiable erythroid islands. An *in vitro* BFU-E (blast forming unit-erythroid) colony forming assay was used to estimate numbers of erythroid-committed progenitors. Fewer colonies were formed from -/- livers compared to wild-type. Qualitative scoring also revealed BFU-E colonies derived from -/- livers to be smaller (i.e. they contained fewer cells), although cell morphology appeared normal. The presence of cells capable performing (almost) normally in this assay indicated that the anaemia phenotype was not the consequence of a complete block in erythroid differentiation. Instead the depleted numbers of BFU-E progenitors, and the reduced colony size were interpreted to be the result of restricted cell proliferation.

Hepatocyte and megakaryocyte lineages appeared normal in liver histology sections. Antibody staining for the cell surface marker T200 indicated that the total number of non-erythroid haematopoietic cells was normal in -/- livers. Similarly, numbers of B220-positive cells were not reduced. These data suggested that members of non-erythroid lineages were relatively unaffected. Thus we concluded that disruption caused by the lack of DNA ligase I was limited to the erythroid lineage.

The observed disruption to foetal liver erythropoiesis in -/- embryos could have had a number of causes. These included failure of HSCs to migrate to the liver, inability of the hepatic microenvironment to support erythropoiesis, or an intrinsic inability of DNA ligase I-deficient cells to complete erythroid differentiation. To quantify the numbers of haematopoietic progenitors within foetal livers, and to examine the capacity of these cells to differentiate, an *in vitro* assay was performed. Single cell suspensions from embryonic livers were plated out under conditions optimised for the proliferation and differentiation of primitive progenitors (CFU-A

assay). Colonies were formed from -/- samples, and, at E11.5, the total numbers per liver were not significantly different from wild-type. However, as development proceeded, numbers of CFU-A progenitors in -/- livers failed to increase by the same amount as wild-type controls. Furthermore, at each developmental stage investigated, qualitative scoring revealed colonies derived from -/- embryos to be smaller. Again cell morphology was apparently normal, from which it was concluded that the differences observed between -/- and wild-type samples could be explained by a quantitative defect in cell proliferation.

Considering all of these data, we were able to propose a model to explain the phenotype observed. Central to this model was the hypothesis that -/- cells were unable to proliferate as rapidly as wild-type cells. Measurements of the rate of cell division were made using primary fibroblasts isolated from mid-gestation embryos. It was possible to demonstrate that cells from -/- embryos divided slightly more slowly than wild-type cells (A.-M. Ketchen and D.W. Melton unpublished data). We believe that -/- cells are able to divide quickly enough to meet the requirements of early developmental events. However, -/- cells are apparently unable to meet the replicative demands of foetal liver erythropoiesis.

During initiation of haematopoiesis in the foetal liver, a limited number of HSCs migrated from the yolk sac. Upon reaching the foetal liver, they not only have to proliferate to produce a pool of stem cells, but also to generate large numbers of terminally differentiated erythrocytes. It is possible to give some indication of the high replicative requirements placed upon the stem cells using numerical data from the *in vitro* colony forming assays. At E11.5 approximately 250 CFU-A progenitors can be detected in the wild-type foetal liver. At E15.5 approximately $6X10^7$ total cells are present within the foetal liver. Taking the number of CFU-A progenitors to represent the number of HSCs present, and assuming that all $6X10^7$ cells are members of haematopoietic lineages and derived from the original 250 stem cells, it is possible to calculate that 18 cell divisions are required to generate $6X10^7$ cells from 250. Given that this proliferation takes place over 4 days, this equates to 4.5 cell divisions per day, or an average of one cell division every 5 hours 20 minutes. Even this crude calculation indicates the phenomenal rate of division to which cells are subject during foetal liver erythropoiesis. Any failure of -/- cells to match the required rate of cell division would preclude exponential proliferation. The normal geometric expansion of erythroid cells would thus be abbreviated and ultimately lead to fewer terminally erythrocytes. In terms of the calculation carried out above, extension of the length of the cell cycle from 5 hours 20 minutes to 6 hours would mean that 45,000,000 fewer cells would be generated after 5 days.

The number of HSCs colonising the foetal liver rudiment is apparently normal in -/- embryos, as indicated by CFU-A numbers at E10.5). However, if the cells are not able to divide rapidly enough to meet the requirement for erythroid cells, the embryo will quickly become anaemic. There is a well defined homeostatic mechanism for maintaining numbers of red cells in the peripheral circulation. Anaemia stimulates the release of the growth factor erythropoietin. Epo stimulates erythroid precursors to differentiate and divide more rapidly, thereby increasing the numbers of erythrocytes released into the peripheral circulation. The most important parameter controlled by Epo regulation *in vivo* is the number of cell divisions in the CFU-E and erythroid cell stages (Pantel *et al.* 1990). However, if -/- cells are unable to divide more rapidly, then increased Epo production will have no effect on the numbers of circulating red blood cells. In the context of haematopoiesis in -/- embryos, the feedback system may even have a counterproductive effect. By forcing stem cells to differentiate to meet the demand for terminally differentiated erythrocytes, Epo may deplete the pool of stem cells. This would lead to fewer cells being subsequently available to proliferate and differentiate, thereby exacerbating the anaemia. Such a positive feedback loop could explain why the anaemia phenotype appears very rapidly between E10.5 and E11.5. Because the embryonic tissues keep growing, the volume of the peripheral circulation keeps increasing, and so does the demand for large numbers of erythrocytes. Once the anaemia has become established, and there is no 'spare' proliferative capacity, the haematopoietic system will never be able to meet demand. As a result, the anaemia will become progressively worse.

Because the -/- embryos grew apparently normally until the onset of anaemia, we believe that growth retardation was secondary to the anaemia. The primary function of erythrocytes is to transport oxygen around the body. Chronic anaemia results in a drop in the oxygen concentration in the tissues (hypoxia), and because all cells require oxygen for respiration, under hypoxic conditions the embryonic tissues will not develop at a normal rate. Prolonged hypoxia will cause cell death, and ultimately will result in the death of the embryo.

Although erythroid cells predominate, haematopoiesis in the foetal liver is not limited to erythropoiesis. Data indicated the other haematopoietic cell lineages to be relatively unaffected by the absence of DNA ligase I. Reduced numbers of pluripotent progenitors were present, but it has been observed in other knockout mice strains that normal numbers of terminally differentiated haematopoietic cells can be present even when numbers of progenitors are reduced (e.g. Ohbo *et al.* 1996). Moreover, we believe that the results can be explained on the basis of known differences between erythroid differentiation and other haematopoietic lineages. It has been estimated that up to 15 cell divisions are required to generate a mature erythrocyte from a

haematopoietic stem cell (Pantel *et al.* 1990). Non-erythroid lineages on the other hand only require 5-6 cell divisions. With the relatively small number of non-erythroid cells present in the foetal liver, the absence of direct regulation of non-erythroid cell numbers, and the dispensability of these cells during development, it is plausible that a limit on the maximum rate of cell division would not significantly disrupt prenatal development of non-erythroid haematopoietic lineages.

7.4 Lethally-irradiated mice could be rescued with -/- haematopoietic stem cells

Having devised a model to explain the specific disruption to foetal liver erythropoiesis caused by the absence of DNA ligase I, we are able to make a number of predictions. Central to the model is the conjecture that although the absence of DNA ligase I results in prenatal lethality, this arose as a secondary consequence of the anaemia. Furthermore, despite the observation that foetal liver erythropoiesis was disrupted, there is no specific block to erythroid differentiation. In other words we propose a quantitative rather that qualitative defect. Anaemia arises simply because cells are unable to proliferate rapidly enough to meet the requirements of the developing embryo. Because rapid erythrocyte production is essential for survival, inefficient foetal liver erythropoiesis results in lethality.

If there is no block to erythropoiesis *per se*, we hypothesised that -/- cells would be able to undergo normal erythropoiesis in a non-developmental context. This appeared to be true *in vitro*, but in order to test the idea *in vivo*, we employed radiation chimaeras. Adult mice whose haematopoietic system has been ablated by a controlled dose of γ radiation can be rescued by the introduction of HSCs to repopulate the haematopoietic system. One consequence of this procedure is that the haematopoietic system of rescued mice will be composed of cells derived from the donor HSCs. Using foetal livers as a source of HSCs, we attempted to rescue lethally-irradiated mice with cells lacking DNA ligase I.

Despite being able to prove that primitive pluripotent progenitors were present, suspensions from single -/- livers were unable to effect rescue. We therefore concluded that the observed disruption to foetal liver erythropoiesis was not simply a consequence of an inappropriate hepatic microenvironment. However, it was possible to rescue mice by injecting cell suspensions from multiple -/- livers into a single recipient, and long-term repopulation of the haematopoietic system was achieved.

Initial experiments using single foetal livers had achieved rescue with wild-type samples. However, subsequent procedures using multiple foetal livers carried out using Balb/c X 129/Ola F1 recipients failed to work even with wild-type samples. Experiments were carried out under identical conditions, apart from the strain of recipients used. Following irradiation, the body-weight profile of these animals

dropped rapidly and did not recover. The majority of animals died (data not shown). We were unable to explain the failure of this experiment in terms other than incompatibility of donor and recipient. F1 recipients will be obligate heterozygotes for all loci which are not isogenic between the Balb/c and 129/Ola strains. Because donor embryos were an ill defined mix of 129/Ola and Balb/c strains, all should be immunologically compatible with F1 recipients. The failure of even wild-type samples to effect rescue was therefore puzzling. The only explanation we had for this (beyond technical problems) involved the hybrid histocompatibility locus (*hh*). The genetics are still poorly understood, but this locus is known to be outwith the MHC complex, and is expressed mainly on natural killer (NK) cells. The locus has the peculiar property of not being expressed in animals which are heterozygous at the locus, although homozygotes do express. Because F1 mice will be obligate heterozygotes but donor embryos could potentially be homozygotes, it is possible that the recipient haematopoietic cells remaining after irradiation could recognise the donor cells as foreign. An immune reaction would thus prevent repopulation and rescue. However, in view of the low survival of inbred recipients in the subsequent experiment, technical problems seem a more likely cause. The failure of wild-type liver suspensions to effect rescue even when donors and recipients were matched for major MHC compatibility could have been due to incompatibility of minor histocompatibility loci, but the body-weight profile suggested otherwise. After an initial drop, the body mass of mice which were rescued returned to normal, and remained stable. Those experimental animals which died, did so in parallel with irradiation controls. This suggested that donor cells were unable to effect rescue rather than rejection of the host tissue by injected donor cells (graft versus host disease, gvhd). We therefore concluded that the experiment was not successful because of a lack of donor cell viability, although the precise cause remained elusive.

Those recipients which were rescued, displayed donor markers in all haematopoietic tissues. FACScan analysis revealed the cellular profile of tissues from animals rescued with -/- cells to be indistinguishable from wild-type controls in most situations. However, whilst there was no extensive failure to contribute to any particular lineage or lineages, the numbers of cells in each haematopoietic tissue were reduced. Quantitative PCR data also indicated that the percentage of donor cells in each haematopoietic tissue were reduced compared to wild-type. When total differences in cell numbers were included, the calculated number of cells of donor origin were greatly reduced in animals rescued with -/- cells. Although the small sample size precluded meaningful statistical analyses of the data, the general trend towards reduced cell numbers was striking.

In order to determine the origins of those mature cells present in the haematopoietic tissues of rescued animals, we employed a lymphocyte proliferation assay. Following mitogen stimulation of B and T lymphocytes in cell suspensions from haematopoietic tissues, PCR was again used to quantify the percentage of donor-derived cells. The reasoning behind this was that only mature lymphocytes proliferate in response to mitogen stimulation, and if a high percentage of these proliferating cells were of donor origin, it must follow that donor cells in haematopoietic tissues were terminally differentiated and capable of responding to stimulation. Whilst this proved to be true for wild-type samples, no significant donor-marker signal was observed in samples from mice rescued with -/- cells. Whether the failure of -/- cells to proliferate in these assays was due to the initial absence of mature cells, or a failure to respond to stimulation could not be determined from these data. It was thus not possible to determine if any of the cells which stained positive for mature cell surface markers by antibody staining were derived from -/- donor material.

Theoretically recipient haematopoietic cells should be mitotically inactive after irradiation, and thus incapable of dividing. However, it is acknowledged that a certain percentage of the haematopoietic system of rescued animals will be of recipient origin. It is possible for a small number of stem cells to effect repopulation of the haematopoietic system in a clonal manner (Micklem *et al.* 1989). Survival of even a few recipient HSCs could thus result in a significant proportion of the haematopoietic system of rescued mice being composed of host-derived cells. Therefore it is entirely feasible that the mature cells in the haematopoietic tissues of rescued animals observed by FACScan analysis were of recipient origin. These cells would be be wild-type with regards to DNA ligase I. Consequently, it was only possible to draw inferences about the properties of -/- cells from the cell populations which were absent from rescued recipients.

Data from mice rescued with -/- cells displayed a general trend towards reduced cell numbers in all haematopoietic lineages. This was consistent with the hypothesis that -/- cells are unable to proliferate as rapidly as wild-type cells. However, it appeared that the B cell lineage was underrepresented in comparison to T cells, and that in turn T cells were depleted compared to granulocytes. These results could be explained by a block to B cell (and T cell) differentiation, for example caused by a failure of -/- cells to undergo functional V(D)J recombination. Alternatively, -/- stem cells might preferentially differentiate rather than proliferate, and may even be biased towards differentiation into granulocytes. Both of these mechanisms would, if proven to be correct, implicate DNA ligase I in cellular processes beyond DNA replication.

The published data does lend tentative support to this hypothesis. It has been shown that single strand breaks present in resting peripheral blood lymphocytes are

rapidly rejoined after PHA stimulation (Boulikas 1991 and references therein). During differentiation of HL-60 along the myeloid lineage, the number of single strand breaks increases, as also happens during induction of differentiation of primary avian myoblasts, murine erythroleukaemia cells, CFU-GM and during ageing. Inhibitors of PARP interfere with mitogen induced activation of human peripheral blood lymphocytes (Colon-Otero *et al.* 1987), and in addition, myeloid cell differentiation inducing protein (MGI-2) causes formation of single strand breaks in closed circular SV40. Exactly why DNA strand breaks affect differentiation is not clear, but one possibility involves the alteration of gene expression by relaxation of supercoiled domains in the genome. In DNA ligase I-deficient cells, the accumulation of strand breaks may result in a predisposition to differentiate rather than be maintained as stem cells, an effect that is only observed in the haematopoietic system where cell division rates are maximal. Alternatively, DNA structural changes may be involved more directly with the differentiation pathway taken by HSCs. From a study of PARP-mediated ligation, it was concluded that strand breakage is involved in the control of differentiation. During granulocyte-macrophage switching, monocytic differentiation is associated with the closure of existing breaks, while granulocytic differentiation involves creation of new disruptions (Kahn and Francis 1987). Analogously, UV treatment of lymphocytes has been shown to induce immunoglobulin class switching (Rosen and Klein 1983). A DNA ligase-deficiency might therefore bias differentiation towards or away from a particular pathway. Mice lacking the Rb protein display a phenotype similar to our DNA ligase I-deficient mice. It has been demonstrated that this is the result of differentiation of haematopoietic progenitor cells into certain non-erythroid cell types (Jacks *et al.* 1992). A similar mechanism could be acting in -/- embryos.

There is conflicting evidence as to which DNA ligase is involved in V(D)J recombination. Gellert and co-workers used DNA ligase I in reconstitution of V(D)J recombination *in vitro*, but 46BR cells exhibit normal V(D)J recombination. Additionally, DNA ligase IV has been shown to interact with, and be stimulated by, XRCC4, a protein required for both V(D)J recombination and double strand break repair. Failure of -/- cells to undergo V(D)J recombination would be strong evidence for the requirement of DNA ligase I in this process. Unfortunately, the data from rescued animals is inconclusive because it was not possible to determine the origins of those B cells observed in mice rescued with -/- cells. It should, however, be possible to address the question directly by differentiating stem cells from -/- livers *in vitro*. *In vitro* B cell differentiation would enable close examination of V(D)J recombination processes in cells lacking DNA ligase I.

Whilst it is possible that the disparity between observed cell numbers in the different haematopoietic lineages of rescued mice was the result of specific disturbance to haematopoietic differentiation, a simpler explanation can be proposed. This suggestion involves the known kinetic demands of the repopulation process. Published data has indicated that all cells in the haematopoietic system of rescued mice may be derived from as few as 3 progenitors (Micklem *et al.* 1987). The stem cells involved are therefore under high proliferative pressure. Donor cells injected into the peripheral circulation of recipients initially home to the spleen, whereupon they proliferate to form colonies composed of erythroid and myeloid lineages. Thereafter other haematopoietic tissues become colonised, probably by migration of progenitor cells from the spleen. Because the development of lymphoid cells takes place only in certain locations (e.g. thymus and bone marrow), lymphoid cells are only generated once haematopoiesis has begun in these tissues. If secondary migration of progenitor cells from the spleen to other haematopoietic tissues does not occur, then mature lymphoid cells would not be generated, and would thus not be present in the peripheral circulation. If -/- cells do not proliferate sufficiently rapidly to completely fulfil the requirements of erythropoiesis, secondary migration of progenitors may not occur to the normal extent, and hence subsequent generation of mature lymphocytes may be abbreviated. Furthermore, Epo-mediated processes may cause bone marrow progenitors to preferentially differentiate into erythrocytes rather than B cells.

7.5 Types of anaemia

There are 3 basic cause of anaemia in humans; impaired red cell formation, lysis of red blood cells (haemolysis), and blood loss.

Impaired red cell formation can be caused by a number of mechanisms. These include metabolic disorders, toxic dyshaematopoiesis, bone marrow infiltration by malignant cells, impaired haemoglobin synthesis, and impaired DNA synthesis. In this last case, abnormal DNA synthesis typically arises from vitamin B12 or folate deficiency, or from drugs which inhibit DNA replication. In these situations, megaloblastic anaemia occurs. The features of megaloblastic anaemia include the presence of macrocytic erythrocytes in the peripheral circulation with profound variability of size and shape (anisocytosis and poikilocytosis). Inclusions are often observed within the erythrocytes (e.g. Howel-Jolley bodies, nuclear remnants), but few reticulocytes are seen.

Aplastic anaemia also arises from impaired red cell production. Fanconi's anaemia is a congenital aplastic anaemia, but similar symptoms can arise from bone marrow failure in adults. In each case, anaemia can be accompanied by macrocytosis with occasional nucleated cells, but there are reduced numbers of all haematopoietic

cell types (pancytopenia) and bone marrow hypoplasia. The reticulocyte count is always low in aplastic anaemia.

Another form of aplasia is pure red cell aplasia. In this ill-defined group of diseases, there is isolated depletion of the erythroid lineage (rather than pancytopenia). Pure red cell aplasia can exist either as an acute condition or a chronic complaint. One example of a chronic red cell aplasia is Diamond-Blackfan anaemia (DBA). Work to date has suggested that DBA is caused by selective deficiency of erythroid progenitors, resulting in a failure to produce the normal number of erythrocytes after birth. Although splenomegaly is observed, reticulocyte numbers are again always low.

Haemolytic anaemias are caused by reduced red cell survival. There are a number of different types of haemolytic anaemias, but all have certain common features. Erythrocyte life-span is reduced, and to compensate there is an increased rate of red cell production. This is manifested as erythroid hyperplasia with high numbers of reticuloocytes in the peripheral circulation. In severe cases, nucleated erythrocytes may be released into the circulation, and red cells may display morphological abnormalities.

7.6 The anaemia phenotype of mice rescued with -/- cells displays characteristics of increased red cell turnover

The phenotype of -/- embryos matched the characteristics of megaloblastic anaemia, which was consistent with the hypothesis that the anaemia arose as a result of impaired DNA replication. The phenotype displayed by mice rescued with -/- cells, however, did not match any of the documented types of anaemia found in humans. Animals rescued with -/- cells were anaemic. The severity of this anaemia was variable, but where severe, was accompanied by pronounced reticulocytosis and macrocytosis. Whilst there was a shortage of circulating erythrocytes, the spleen contained large numbers of erythroid cells, of which a large proportion appeared to be DNA ligase I-deficient. The percentage of reticulocytes is always low in anaemia arising from impaired red cell formation, and hence the anaemia observed in mice rescued with -/- cells does not correspond to megaloblastic anaemia.

The phenotype also did not match exactly with red cell aplasia syndromes such as DBA. However, our interest in DBA was stimulated by publication of data mapping the DBA gene to a 1.8Mb region on human chromosome 19q13.3 (Gustavsson *et al.* 1997). *LIG1* maps to 19q13.2-13.3, and given the anaemia phenotype of DNA ligase I-deficient mice, was thus a potential DBA candidate gene. However, we were unable to identify any changes to the *LIG1* locus in DBA samples, and subsequent mapping of the disease locus proved that DBA is not congruent with *LIG1* (S.E. Ball, personal communication).

Reticulocytosis and erythroid hyperplasia are two characteristics of haemolytic anaemias arising as a consequence of increased red cell turnover. The fact that mice rescued with -/- cells displayed reticulocytosis and splenomegaly would suggest that red cell turnover was increased. The logical conclusion to be drawn from these data is that anaemia in -/- embryos and mice rescued with -/- cells arises because one of the final steps of erythroid differentiation is blocked. This would contradict the hypothesis that anaemia (and pancytopenia) was the result of a reduced rate of cell proliferation. Resolution of this conflict requires additional experiments to be carried out. The most obvious way to investigate this issue would be to carry out more detailed studies of erythroid differentiation of -/- cells *in vitro* (e.g. CFU-E assays carried out on foetal liver cell suspensions).

A specific disruption to erythroid differentiation would, however, not be easily reconciled with the pancytopenia observed, nor the variability in the degree of reticulocytosis. It remains possible that all the features observed in mice rescued with -/- cells may result from a reduced proliferative capacity. As outlined earlier, repopulation of the haematopoietic system is a dynamic process that involves rapid proliferation of stem cells. All sampling points we employed were after long term repopulation had been achieved, and thus only represented the final 'steady state'. It is possible that reticulocytosis may represent a secondary event arising from perturbation to the repopulation process.

The processes of repopulation and initiation of foetal liver erythropoiesis are similar in that both involve the proliferation and differentiation of a small number of HSCs to form a large population of terminally differentiated cells on a relatively short time-scale. The model devised to explain the disruption to foetal liver erythropoiesis in -/- embryos may thus be adapted to the repopulation of lethally-irradiated recipients with -/- cells. If -/- HSCs struggle to meet the initial proliferative demands, and fail to produce the normal number of differentiated cells, hypoplasia and anaemia would occur. Anaemia would result in Epo production, which in turn stimulates erythroid differentiation. As noted previously, this would have little effect on a system that was already operating to its full proliferative potential. In the absence of the required numbers of mature erythrocytes, immature erythroid cells could be prematurely released into the peripheral circulation, and reticulocytosis would occur. Reticulocytes are rapidly cleared from the peripheral circulation by the liver, and have a life-span of 1-2 days, compared to 28 days for a mature erythrocyte. As a result, rapid clearance of these cells from the circulation would effectively mimic haemolytic events. A positive feedback loop would thus become established, anaemia forcing the release of reticulocytes, which are cleared rapidly, thereby exacerbating the anaemia. Providing sufficient numbers of stem cells are present to enable recipients to survive (and by

definition rescued animals have to survive) then over the course of time, Epo stimulation of erythroid differentiation will expand numbers of erythroid lineage cells in the spleen, resulting in hyperplasia. Again the Epo response may diminish the pool of stem cells, precluding subsequent correction of the normal balance of haematopoietic lineages. Because the demand for erythroid cells in the adult mouse is static, given time, the anaemia should theoretically resolve. We did not observe resolution, and in fact the opposite occurred. However, given the repeated sampling which removed large volumes of blood, and the relatively short period prior to final sacrifice, this would not be inconsistent with the model presented.

7.7 DNA ligase I-deficient cells display an abnormal profile of replication intermediates

Whatever the precise nature of the disturbance to erythroid differentiation, there remains the fact that -/- cells are capable of cell division. It therefore follows that DNA replication must be occurring in the absence of DNA ligase I. This was highly unexpected given the wealth of data on the involvement of DNA ligase I in lagging strand DNA replication. In order to investigate the process of DNA replication in -/- cells, pulse-chase experiments were carried out. Radioactively labelled thymidine was introduced into cells for a short period of time. By separating DNA molecules on the basis of size, and using scintillation to detect the radiolabelled fragments, it proved possible to identify the newly replicated DNA.

Using pulse-chase labelling in human cells containing an altered form of DNA ligase I (46BR.1G1), we were able to demonstrate diminution of the average size of newly replicated DNA fragments. The 46BR.1G1 cells contained a raised proportion of small labelled fragments after the pulse procedure, though these small fragments were apparently incorporated into larger molecules with time. In this respect, these results were broadly similar to other studies of DNA replication carried out using 46BR and 46BR.1G1 fibroblasts.

An immortalised fibroblast line derived from a -/- embryo exhibited a broadly similar replication defect to the 46BR.1G1 cells. There was a raised proportion of small labelled fragments after the pulse period and these were "chased away" with time. However, comparison between -/- cells and 46BR.1G1 revealed that the average fragment size was greater in -/- cells. Put another way, there were more small fragments in 46BR.1G1 cells. This would suggest that the initial joining of DNA fragments was more efficient in -/- cells. However, following the initial pulse period, fragments were chased away at broadly similar rates in both cell types. Whether any firm conclusions can be drawn from these results would be reliant on a more detailed study of the kinetics of the systems, but a number of general inferences can be made.

The observation that DNA replication processes are disturbed in cells lacking DNA ligase I would be consistent with the hypothesis that -/- cells have a reduced proliferative capacity. Although the cell cycle was apparently unaffected by the absence of DNA ligase I (A.-M. Ketchen and D.W. Melton, unpublished data), the analysis would not detect subtle changes to timings of the cell cycle phases. Retarded joining of Okazaki fragments in -/- cells could therefore extend the time taken to complete S phase without the effect being noticeable by FACscan analysis of whole cell populations.

Satoh and Lindahl have presented a model to explain the hypersensitivity of 46BR cells to 3AB in terms of the interaction of polyADPribose polymerase with DNA strand breaks (Lindahl *et al.* 1993). PARP binds tightly to breaks, and apparently prevents access of excision repair enzymes. Release occurs following automodification of PARP, since heavily autoribosylated PARP no longer binds effectively to DNA. 3AB acts on PARP by inhibiting release of the molecule from DNA. Normally, PARP does not gain access to DNA strand interruptions during lagging strand synthesis, due to the protection of DNA termini by the replisome. Hence the effects of PARP on DNA joining is usually limited to BER of alkylation damage (PARP being known to associate with the BER complex of proteins). 46BR cells on the other hand, exhibit delayed joining of the Okazaki fragments, which may result in their escape from the replisome, and thereby become targets for PARP binding. In turn, this would delay conversion of the small intermediates into higher molecular weight forms, an effect greatly exacerbated by the presence of 3AB, which would thus result in reduced cell viability. However, cells from -/- embryos are not sensitive to the effects of killing by 3AB (A.-M. Ketchen and D.W. Melton, unpublished data). This implies that either the model proposed by Satoh and Lindahl is inappropriate, or that newly synthesised DNA fragments are processed in different ways in the two cell types.

It has already been noted that the initial numbers of small fragments are greater in 46BR.1G1 cells, but that subsequent joining occurs at broadly similar rates in 46BR.1G1 and -/- cells. The assumption implicit in this interpretation is that the mechanisms by which the labelled fragments are generated, and subsequently processed, are identical in the two cell types, and that it is only the rates of reaction that are different.

It has been suggested that DNA ligase I is part of a 'core' multiprotein replication complex via interactions with PCNA. Nevertheless, data from Loor *et al.* indicated that DNA ligase I may not be essential for the formation of replication complexes. Analysis of protein complexes which bound to PCNA failed to identify DNA ligase I as a component of complexes containing other replication proteins

(Loor *et al.* 1997). Our suggestion would be that replication complexes form and DNA polymerisation is carried out as normal in 46BR and -/- cells. Conflicting interpretations have been made from analyses of DNA replication in 46BR cells. Lehmann *et al.* found that incorporation of Okazaki fragments into high molecular weight forms occurred at a reduced rate (Lehmann *et al.* 1988). Prigent *et al.* determined that the initial rate of joining was normal, but that a percentage of small fragments remained unligated after prolonged chase periods (Prigent *et al.* 1994). Our data on 46BR cells was consistent with both models. In each case the assumption is that the defective DNA ligase I protein interacts with the replication complex in a normal manner. Joining of Okazaki fragments at the replication fork is carried out by DNA ligase I, albeit at a reduced rate. Subsequently unligated fragments are joined by the ligation activity of a repair pathway.

In cells lacking DNA ligase I, ligation during lagging strand synthesis cannot be carried out by DNA ligase I. The simplest explanation would be that another DNA ligase assumes the role of DNA ligase I within the replication complex. However, to date no other DNA ligase has been observed to participate in replication, and, at least *in vitro*, DNA ligases III and IV have been shown not to function (Waga, Bauer and Stillman 1994, Mackenney, Barnes and Lindahl 1997). Whether another DNA ligase acts to replace DNA ligase I directly in -/- cells remains to be determined. One potential way of doing this would be to isolate replication complexes from -/- cells. It would thus be possible to identify which, if any, other DNA ligase was present within the complex.

Alternatively, it is possible that all Okazaki fragments formed during lagging strand DNA synthesis in -/- cells are processed by a 'repair' mechanism not directly associated with the replication complex. It has been noted that single strand breaks generated during DNA replication are recombinogenic. The *S. cerevisiae cdc9* DNA ligase mutant displays an increased mitotic recombination rate (Game, Johnston and von Borstel 1979). Recently Zou and Rothstein demonstrated that Holliday junction replication intermediates are increased in certain yeast replication mutants, including *cdc9* (DNA ligase deficient) cells. They proposed that replication-related lesions are repaired by recombination via a Rad52-dependent pathway (Zou and Rothstein 1997). It is thus potentially possible that lesions arising from lagging strand DNA replication could be repaired by this, or other pathways, in the absence of DNA ligase I. Although speculative, different relative activities of 'replication' and 'repair'-mediated ligation between 46BR and -/- cells could explain the different profiles of replication intermediates observed.

Increased levels of mitotic recombination have already been observed in human cells containing aberrant DNA ligase I. Bloom's syndrome cells display high numbers

of exchanges between sister chromatids (SCEs) and between homologous chromosomes. Elevated SCE levels have also been recorded in 46BR.1G1 cells following treatment with DNA damaging agents. We were therefore very interested to determine whether the frequency of SCEs was elevated in -/- cells. BS is now known to be caused by mutations within a helicase protein, and SCEs are thought to be a mechanism by which BS cells process abnormal structures formed during replication. 46BR cells on the other hand are sensitive to a wide variety of DNA damaging agents, which has been presumed to indicate deficiency in a number of DNA repair pathways. It is therefore believed that induction of SCEs by treatment with DNA damaging agents represents a failure to repair lesions before entering S phase. Potentially therefore SCEs could be a marker for either perturbation of DNA replication, or inefficient repair of DNA damage.

Cells derived from -/- embryos did not display an elevated SCE frequency. No significant difference was observed between -/- samples and wild-type controls. These data indicate that perturbations in replication arising from the absence of DNA ligase I were not resolved by a mechanism that induced sister chromatid exchanges. 46BR cells had been observed to have a marginally raised baseline level of SCEs (Henderson *et al.* 1985), but this was not observed in -/- cells. When we carried out the assay on 46BR.1G1 cells, the numbers of SCEs observed were higher than reported elsewhere. Nevertheless we do not believe that this result was significant given the differences between the experiments. The disparity most likely stems from the use of primary fibroblasts, whereas we employed the SV40 immortalised 46BR.1G1 line. Either the immortalisation process or the prolonged period of time spent in culture could have caused molecular changes which elevated the level of SCEs in 46BR.1G1 cells.

46BR cells were reported to be sensitive to the induction of SCEs by DNA damaging agents. Whether this is also true for -/- cells remains to be determined. The response of 46BR cells has been linked to the sensitivity of these cells to killing by DNA damaging agents. Studies by A.-M. Ketchen and D.W. Melton have indicated that -/- cell lines are not sensitive to a variety of DNA damaging agents. As such, it would be predicted that -/- cells would not show induction of the rate of SCE formation in response to DNA damage.

7.8 Molecular changes arising from the absence of DNA ligase I

DNA ligase I is the predominant DNA ligase activity in proliferating cells, and is thought to be central to lagging strand DNA replication. Because cells lacking DNA ligase I are nonetheless viable, some other gene product or pathway must assume the role of DNA ligase I in joining small replication intermediates into high molecular

weight forms. The obvious candidate for this would be one of the other known DNA ligases. Because these other enzymes are poorly expressed in proliferating cells in comparison to DNA ligase I, it would not be unreasonable to expect expression of the complementing DNA ligase to be upregulated. To test whether DNA ligases III or IV were upregulated in the absence of DNA ligase I, northern and western blot analysis was carried out.

In order to perform northern analysis, fragments of the mouse homologues of *LIG3* and *LIG4* were cloned by RT-PCR using primers based on the published human sequence. The fragment of the putative *Lig3* cDNA cloned had a predicted amino acid sequence with high homology to the human protein. Furthermore, the fragment detected a transcript of the correct size (~4kb) which was highly expressed in testes. However, a ~6kb transcript was also detected by this probe, which also showed maximal expression in testes. Tomkinson and co-workers observed an identical phenomenon using a different region of the cDNA (Husain *et al.* 1996). It is therefore probable that the additional transcript was not an artefact, but instead represents a splicing intermediate.

The predicted amino acid sequence of the putative *Lig4* cDNA fragment displayed homology with the published sequence of the human protein, and detected a fragment of the predicted size. The tissue distribution of this transcript correlated with published data on expression of *LIG4*, being most highly expressed in testes, thymus and bone marrow. The high levels observed in lymphoid tissues would lend credence to the supposition that DNA ligase IV is involved in V(D)J recombination.

In short, the properties of the cDNA fragments showed close agreement with data available for *LIG3* and *LIG4*, and we were confident that we had cloned sequences from the mouse *Lig3* and *Lig4* homologues. The cloned fragments were thus used to probe RNA samples from embryos representative of the three genotypes and cell lines derived from mid-gestation embryos. There was no detectable change in levels of *Lig3* or *Lig4* transcripts in samples lacking DNA ligase I. Similarly, western blotting using antibodies specific for DNA ligases III and IV revealed no alterations to the respective protein levels in -/- samples. Furthermore, an *in vitro* assay for adenylation activity showed no detectable differences between wild-type and -/- samples (data not shown). We therefore concluded that targeted deletion of *Lig1* does not affect expression of either DNA ligase III or DNA ligase IV. Whilst it is still theoretically possible that one (or both) of the two enzymes acts to complement for the absence of DNA ligase I, these data would suggest that the role is filled by some other enzyme.

Published evidence has suggested that there may be up to six distinct DNA ligase activities in mammalian cells. Biochemical strategies have been used to purify

proteins designated as DNA ligase II, DNA ligase V, NHR ligase, and a mitochondrial DNA ligase. Little is known about any of these proteins, and in the case of DNA ligase V, NHR ligase and mitochondrial DNA ligase, references are limited to a single publication apiece. All putative DNA ligases have been shown to display DNA ligation activity. Whether any are actually derived from novel genes awaits cloning of the loci responsible. Without cloned cDNAs or reliable antisera against the proteins, it was impossible for us to investigate whether the absence of DNA ligase I affected expression of these other putative DNA ligases. The use of a functional assay, e.g. the ability to form adenylated intermediates, would have enabled us to investigate this possibility. However, when we carried out this assay on crude cell extracts, the sensitivity of the assay proved to be low. Without purification of the active proteins, it proved impossible to identify any differences in the levels of the known DNA ligases or to identify any novel adenylated products. This approach was therefore discounted as an option.

DNA ligases III and IV were both cloned on the basis of homology to DNA ligase I. Comparison of the protein sequences demonstrated that all known DNA ligases share certain common motifs, notably the active site and conserved C-terminal polypeptide. We attempted to exploit possible homology between DNA ligases in order to clone novel DNA ligases. The RT-PCR fingerprinting technique developed by Boehm had previously been used to clone novel members of the tyrosine phosphatase gene superfamily (Boehm 1993). Degenerate PCR primers are used to amplify the region between two conserved motifs. Thereupon the products are digested with frequently cutting restriction enzymes, and separated on a high resolution gel system. Divergence between homologous genes within the intervening region results in different family members generating different restriction patterns, from which it is possible to identify and clone previously unidentified members of the gene family.

Using PCR primers based on the sequences of DNA ligases I, III and IV under low stringency conditions, we found that large numbers of bands were produced even prior to restriction digestion. Without the appropriate cDNA controls it proved impossible to identify even the bands corresponding to the known DNA ligases. In this respect, the attempt to employ RNA fingerprinting was a failure. However, it was noted that patterns differed between wild-type and -/- samples. Cloning of those bands increased in intensity in -/- samples enabled us to identify two known genes and one previously unidentified gene whose expression was upregulated in mid-gestation -/- embryos.

One of the genes we were able to identify encoded transferrin, a protein known to be upregulated in conditions of anaemia. This was not surprising given the anaemia

phenotype of -/- embryos. The second identifiable transcript was a fragment of the MRP14 gene. MRP14 is known to be expressed during granulopoiesis, but why it should apparently be upregulated in -/- embryos remains unclear. It is possible that the increased level of transcript in -/- RNA samples compared to wild-type controls is merely an artefact. With far fewer erythroid cells present in the livers of -/- embryos, cell lineages present in normal numbers will represent a higher proportion of cells. It would thus follow that any lineage specific transcript not expressed in erythroid cells would appear to be 'upregulated' in -/- RNA samples. However, FACScan analysis of mice rescued with -/- cells does suggest that there is a disparity between numbers of cells in the different haematopoietic lineages. It is thus possible to speculate that the raised levels of MRP14 transcript observed actually represents an increase in the numbers of granulocytes in the livers of -/- embryos compared to wild-type controls. Given the consequent implication that non-erythroid haematopoietic cell lineages are differentially affected by the absence of DNA ligase I, the detection of MRP14 as a gene that is 'upregulated' may prove to be highly significant.

The third cloned fragment (#10) did not have homology to any sequence in the public databases. It therefore may represent a 'novel' gene of undetermined function. Although the transcript was not expressed in one -/- cell line, and we subsequently failed to detect expression in primary fibroblasts (data not shown), the fragment #10 transcript was detected at high levels in one of the DNA ligase I-deficient cell lines. This raised the intriguing possibility that the gene may be involved in cellular processes directly affected by the lack of DNA ligase I. Further investigation is, however, required to clarify the matter.

Although not what was originally intended, the results achieved from this attempt to employ RNA fingerprinting do demonstrate the potential of the approach. It is possible that RNA fingerprinting could be made to work using specifically designed primers. However, this technique can only be employed to investigate the members of a defined gene family. The related technique of 'differential display' is far more versatile in that it permits analysis of the expression levels of all genes. It is highly sensitive, and can be used to identify quantitative differences in gene expression levels as well as simple qualitative changes. Furthermore, the technique is relatively cheap, and easy to perform, and thus represents an ideal method to screen for, and clone, genes whose expression is altered following gene targeting. The major limitation of this technique is that it only detects changes in gene transcription. It will not detect translational or post-translational control mechanisms, and identification of genes whose products might complement for the absence of DNA ligase I is therefore totally reliant on the assumption that transcriptional upregulation will occur.

7.9 Cells lacking DNA ligase I are repair competent

In order to address the question of whether -/- cells are also repair deficient, -/- cell lines were challenged with a variety of different DNA damaging agents (A.-M. Ketchen and D.W. Melton, unpublished data). Results to date have indicated that -/- cells are not sensitive to the killing effects of any damaging agent. There are a number of potential explanations for these results.

DNA ligase I may not function in the pathway employed to repair the damage. For example, accumulating evidence points to DNA ligase IV, and not DNA ligase I, being involved in non-homologous double strand break repair. Alternatively, there may be two redundant repair pathways capable of repairing the damage lesion. In the case of BER, there are known to be two pathways which repair alkylated bases; 'short patch' (PCNA-independent) and 'long patch' (PCNA-dependent). While DNA ligase III is involved in the former pathway, DNA ligase I is only thought to function in the latter, and hence the absence of the protein would not abolish repair of alkylation damage. It is possible that another DNA ligase may be able to fill the role of DNA ligase I. For example, it has been demonstrated that Ku can stimulate the activity of all three known DNA ligases, and so all may be able to function in Ku-mediated end joining. Redundancy is especially plausible if the final ligation step is not tightly associated with the initial lesion processing. Although NER and BER are thought to be carried out by multiprotein complexes, no such protein associations have been identified for the mismatch repair pathway. If free DNA ligase I protein carries out the final ligation step, then it would be feasible for another free DNA ligase to perform the same role equally well.

Although -/- cells are not sensitive to any exogenous DNA damaging agents, it has been shown that they do display a high level of genomic instability (D.W. Melton, unpublished data). Studies of metaphase cells demonstrated that there was a high variability in chromosome number, and the rate of micronucleus formation *in vivo* was also increased. As -/- cells are repair competent for all types of damaging agents, the implication is that either instability arises as a result of the replication defects observed, or that possibly DNA ligase I is required for a more general genome surveillance mechanism. Whichever is the case, *Lig1* -/- cells represent the first example of separation of genomic instability from sensitivity to DNA damaging agents.

7.10 Deletion of DNA ligase I *in vivo*

Whilst the ability to generate DNA ligase I-deficient mice demonstrated that DNA ligase I was not required for cell viability, the associated developmental lethality restricted the scope of investigation of the consequences of the absence of the protein

in vivo. Studies of cell lines *in vitro* have contributed towards understanding the molecular changes involved, but a viable whole animal model would have far greater potential. In order to develop such a model, we chose to use the Cre/*loxP* recombination system. Cre recombinase mediates site-specific recombination between 34bp sequences termed *loxP* sites. By flanking a gene of interest with two *loxP* sites, recombination will delete that gene when the Cre protein is expressed. Because Cre-mediated recombination is a highly efficient process, it is possible to carry out recombination *in vivo* without selection, and by limiting Cre expression to a defined cell lineage or timepoint, it is possible to perform gene knockout in a conditional manner.

We employed a double replacement gene targeting strategy to introduce *loxP* sites into the endogenous *Lig1* locus. Using ES cells heterozygous for the targeted *Lig1* knockout allele, we reconstructed a functional *Lig1* allele but with *loxP* sites flanking exon 23 and 27. Southern blotting confirmed that 8 ES cell clones identified by PCR had undergone the correct homologous recombination event. These clones were then used to produce mice heterozygous for the 'floxed' *Lig1* allele. Concurrently, we also used the ES cell clones to test Cre-mediated recombination of the floxed *Lig1* allele.

A PCR-based assay was devised to detect Cre-mediated recombination at the *Lig1* locus. This assay was used to screen ES cells into which a Cre-expression plasmid had been introduced. Three clones gave positive signals from this experiment, and Southern analysis confirmed that two had undergone Cre-mediated deletion of the last 5 exons of the floxed *Lig1* allele. The structure of the deleted *Lig1* allele determined by Southern blotting matched the predictions made from the known structure of the floxed allele. Northern blot analysis was also carried out on ES cell clones heterozygous for the floxed and deleted *Lig1* alleles. Samples heterozygous for the floxed allele were indistinguishable from wild-type samples. We thus concluded that the floxed allele was analogous to the wild-type allele, despite the engineered changes. Following Cre-mediated recombination, the level of *Lig1* transcript dropped by approximately 50%, suggesting that there was no expression from the deleted allele. The absence of expression is significant because it meant that the deleted allele is functionally equivalent to the targeted *Lig1* knockout allele. It should thus be possible to compare the properties of Δ/Δ, Δ/-, and -/- cells in the knowledge that no protein is produced from either allele.

The third ES cell clone which had given positive results from the PCR screen did not show an appreciably lower level of *Lig1* transcript signal compared to wild-type controls. Similarly, restriction patterns produced by Southern blotting did not show disappearance of the fragment specific for the floxed *Lig1* allele. When PCR

quantification of the proportion of cells carrying the deleted allele was carried out, it was calculated that 75% of cells did not contain the deleted allele. The simplest explanation for this result was that during plating of cells at single-cell density a cell which had undergone Cre-mediated recombination remained attached to three cells which had not undergone recombination. The resulting ES cell colony would thus not be derived from a single cell, but would be a mixture of cells of different genotypes.

The ES cell clones heterozygous for the deleted allele were used to develop a quantitative PCR-based assay for the deleted *Lig1* allele. By spiking known amounts of DNA from samples heterozygous for the deleted allele into wild-type DNA, and then subjecting to PCR, a linear standard curve was produced. In this way it proved possible to quantify the level of Cre-mediated recombination which had occurred in other samples. Although the values calculated from this technique displayed a high variance, the validity of the data was confirmed independently by Southern blotting. Both PCR and Southern blotting gave results which were in close agreement.

The data from these *in vitro* studies gave us confidence that not only could Cre-mediated recombination occur between the two *loxP* sites at the *Lig1* locus, but that we were able to detect this event even at low levels. To investigate recombination *in vivo*, ES cells heterozygous for the floxed allele were used to produce mice heterozygous for the floxed allele. These mice were subsequently interbred to produce mice homozygous for the floxed *Lig1* allele (f/f). The f/f mice were phenotypically normal, indicating that the f allele was functionally equivalent to the wild-type allele. f/f mice were crossed to two lines of transgenic mice (74 and 93) which expressed Cre recombinase under the control of the beta lactoglobulin promoter (BLG). BLG is the major whey protein in ruminants, and the BLG promoter had previously been shown to direct expression of transgenes exclusively to the secretory epithelial cells of mammary tissue, where expression peaks at day 10 of lactation. Quantitative PCR and Southern analysis demonstrated that in one of the transgenic lines (74), Cre-mediated recombination occurred in 70-80% of the cells in the lactating mammary gland, while background recombination levels in non-mammary tissues were minimal. Very low recombination levels were observed in virgin mammary tissue, increasing throughout gestation and birth, to peak around day 10 of lactation. This pattern of Cre activity closely paralleled the published kinetics of BLG expression. We were thus able to conclude that the BLG-Cre construct specifies mammary-specific gene deletion, and that deletion is tightly controlled both spatially and temporally.

These results confirm other independent data on Cre expression from the BLG transgene in line 74 transgenic mice. However, they also provided valuable confirmation that high levels of Cre-mediated deletion of the *Lig1* allele can occur *in vivo*. This will hence facilitate investigation of the effects of the absence of DNA

ligase I in an adult environment. By crossing mice carrying both the BLG-Cre transgene and the *Lig1* floxed allele to mice heterozygous for the targeted *Lig1* allele, it will be possible to produce f/- mice which undergo Cre-mediated deletion of *Lig1* in the lactating mammary gland. This should, in effect, produce Δ/- secretory epithelial cells lacking DNA ligase I. However, the availability of mice carrying a floxed *Lig1* allele offers wider potential than simple investigation of the effect of DNA ligase I-deficiency in the mammary gland. A wide range of conditional or inducible *Lig1* knockouts could be generated. By crossing the floxed *Lig1* allele with mice expressing Cre in different tissues, it will be possible to study the consequences of loss of DNA ligase I activity in other cell lineages. Given the phenotype of -/- embryos and mice rescued with -/- cells, the obvious starting point would be deletion of *Lig1* from specific haematopoietic cell lineages. This offers the immediate opportunity to resolve the current debate about the requirements of DNA ligase I for V(D)J recombination, and to determine whether the course of haematopoietic differentiation can be affected by the absence of DNA ligase I.

In short, the generation of mice carrying a floxed *Lig1* allele will allow the production of a number of different experimental mouse models of DNA ligase I-deficiency in the future. With these mice it should be possible to answer long-standing questions over the importance of DNA ligase I to the immune system and the processes of neoplasia.

CHAPTER SEVEN: DISCUSSION

CHAPTER EIGHT:

REFERENCES

Aboussekhra, A., Biggerstaff, M., Shivji, M.K.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.M., and Wood, R.D. (1995) Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80(6):** 859-868

Applegren, N., Hickey, R.J., Kleinschmidt, A.M., Zhou, Q., Coll, J., Wills, P., Swaby, R., Wei, Y., Quan, J.Y., Lee, M.Y.W.T., and Malkas, L.H. (1995) Further characterization of the human cell multiprotein DNA replication complex. *J. Cell. Biochem.* **59:** 91-107

Arrand, J.E., Willis, A.E., Goldsmith, I. and Lindahl, T. (1986) Different substrate specificities of the two DNA ligases of mammalian cells. *J. Biol. Chem.* **261:** 9079-9082

Barnes, D.E., Johnston, L.H., Kodama, K., Tomkinson, A.E., Lasko, D.D., and Lindahl, T. (1990) Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. (USA)* **87:** 6679-6683

Barnes, D.E., Kodama, K.-I., Tynan, K., Trask, B.J., Christensen, M., De Jong, P., Spurr, N.K., Lindahl, T., and Mohrenweiser, H.W. (1992a) Assignment of the gene encoding DNA ligase I to human chromosome 19q13.2-13.3. *Genomics* **12:** 164-166

Barnes, D.E., Tomkinson, L.H., Lehmann, A.R., Webster, A.D.B., and Lindahl, T. (1992b) Mutations in the DNA ligase I gene in an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell* **69:** 495-504

Bentley D.J., Selfridge J., Millar J.K., Samuel K., Hole N., Ansell J.D. and Melton, D.W. (1996) DNA ligase I is required for foetal liver erythropoiesis but is not essential for mammalian cell viability. *Nature Genetics* **13:** 489-491

Bloom, D. (1954) Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs. Probably a syndrome entity. *Am. J. Dis. Child* **88:** 754-758

Bloom, D. The syndrome of congenital telangiectatic erythema and stunted growth: Observations and studies. *J. Pediatr.* (1966) **68:** 103-113.

Boehm, T. (1993) Analysis of multigene families by DNA fingerprinting of conserved domains: directed cloning of tissue-specific protein tyrosine phosphatases. *Oncogene* **8:** 1385-1390

Bogue, M., and Roth, D.B. (1996) Mechanism of V(D)J recombination. *Curr. Opin. Immunol.* **8:** 175-180

Caldecott, K.W., Aoufouchi, S., Johnson, P., and Shall, S. (1996) XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly(ADP-ribose) polymerase, and DNA ligase III is a novel molecular nick-sensor *in vitro Nucleic Acids Res.* **24(22):** 4387-4394

Caldecott, K.W., McKeown, C.K., Tucker, J.D., Ljungquist, S., and Thompson, L.H. (1994) An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Mol. Cell Biol.* **14:**68-76

Caldecott, K.W., Tucker, J.D., Stanker, L.H., and Thompson, L.H. (1995) Characterization of the XRCC1-DNA ligase III complex in vitro and its absence from mutant hamster cells. *Nucleic Acids Res.* **23(23):** 4836-4843

Callebaut, I., and Mornon, J.P. (1997) From BRCA1 to RAP1: A widespread BRCT module closely associated with DNA repair. *FEBS Lets.* **400(1):** 25-30

Capecchi, M.R. (1989) The new mouse genetics: Altering the genome by gene targeting. *Trends Genet.* **5(3):** 70-76

Cappelli, E., Taylor, R., Cevasco, M., Abbondandolo, A., Caldecott, K., and Frosina, G. (1997) Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. *J. Biol. Chem.* **272(38):** 23970-23975

Cardoso, M.C., Joseph, C., Rahn, H.P., Reusch, R., NadalGinard, B., and Leonhardt, H. (1997) Mapping and use of a sequence that targets DNA ligase I to sites of DNA replication *in vivo*. *J. Cell Biol.* **139(3):** 579-587

Chaganti, R.S.K., Schonberg, S., and German, J. (1974) A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc. Natl. Acad. Sci.(USA)* **71:** 4508-4512

Chan, J.Y.-H., and Becker, F..F (1985) DNA ligase activities during hepatocarcinogenesis induced by N-2-acetylaminofluorene. *Carcinogenesis* **6:** 1275-1277

Chan, J.Y.H., and Becker, F,.F (1988) Defective DNA ligase I in Bloom's syndrome cells. *J. Biol. Chem.* **263:** 18231-18235

Chen, J.W., Tomkinson, A.E., Ramos, W., Mackey, Z.B., Danehower, S., Walter, C.A., Schultz, R.A., Besterman, J.M., and Husain, I. (1995) Mammalian DNA ligase III. Molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination. *Mol. Cell. Biol.* **15(10):** 5412-5422

Colon-Otero, G., Sando, J.J., Sims, J.L., McGrath, E., Jensen, D.E., and Queensberry, P.J. (1987) Inhibition of haematopoietic growth-factor induced proliferation by adenosine diphosphate-ribosylation inhibitors. *Blood* **70(3):** 686-693

Critchlow, S.E., Bowater, R.P., Jackson, S.P. (1997) Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Curr. Biol.* **7(8):** 588-598

Dagert, M. and Ehrlich, S.D. (1974) Prolonged incubation in calcium phosphate improves competence of *Escherichia coli* cells. *Gene* **6:** 23-28

Derbyshire, M.K., Epstein, L.H., Young, C.S.H., Munz, P.L., Fishel, R. (1994) Nonhomologous recombination in human cells. *Mol. Cell. Biol.* **14(1):** 156-159

Digweed, M., and Sperling, K. (1996) Molecular analysis of Fanconi anaemia. *Bioessays* **18(7):** 579-585

Dzierzak, E., and Medvinsky, A. (1995) Mouse embryonic hematopoiesis. *Trends Genet.* **11(9):** 359-366

Elder, R.H., and Rossignol, J.-M. (1990) DNA ligases from rat liver. Purification and partial characterisation of two molecular forms. *Biochem.* **29:** 6009-6017

Ellis, N.A., Groden, J., Ye, T.-E., Straughen, J., Lennon, D.J., Ciocci, S., Proytcheva, M., and German, J. (1995a) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* **83:** 655-666

Ellis, N.A., Lennon, D.J., Proytcheva, M., Alhadeff, B., Henderson, E.E., and German, J. (1995b) Somatic intragenic recombination within the mutated locus *BLM* can correct the high SCE phenotype of Bloom syndrome cells. *Am. J. Hum. Genet.* **57:** 1019-1027

Essers, J., Hendriks, R.W., Swagemakers, S.M.A., Troelstra, C., de Wit, J., Bootsma, D., Hoeijmakers, J.H.J., and Kanaar, R. (1997) Disruption of mouse RAD54 reduces ionizing radiation resistance and homologous recombination. *Cell* **89:** 195-204

Evans, M., and Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292:** 154-155

Fairman, M.P., Johnson, A.P., and Thacker, J. (1992) Multiple components are involved in the efficient joining of double strand breaks in human cell extracts. *Nucleic Acids Res.* **20:** 4115-4132

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13

Frankenberg, D., Frankenberg-Schwager, M., Blöcher, D., and Harbich, R. (1981) Evidence for double-strand breaks as the critical lesions in yeast cells irradiated with sparcely or densely ionizing radiation under oxic and anioxic conditions. *Radiat. Res.* **88:** 524-532

Freshney, M.G., in *Culture of Haematopoietic Cells* (eds Freshney, R.I., Pragnell, I.B. and Freshney, M.G.) 265-268 (Wiley-Liss, New York, 1994)

Friedberg, E.C., Walker, G.C., and Siede, W. *DNA repair and mutagenesis.* (American Society of Microbiology, Washington DC, 1995)

Game, J.C., Johnston, L.H., and von Borstel, R.C. (1979) Enhanced mitotic recombination in a ligase-defective mutant of the yeast *Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. (USA)* **76(9):** 4589-4592

Genetics Computer Group (1996) Program manual of the GCG package, version 9. Genetics Computer Group, Madison, Wisconsin.

German, J. (1969) Bloom's syndrome. I. Genetical and clinical observations in the first twenty-seven patients. *Am. J. Hum. Genet.* **21:** 196-227.

German, J. (1993) Bloom Syndrome: A mendelian prototype of somatic mutational disease. *Medicine* **72(6):** 393-406

German, J., Roe, A.M., Leppart, M.F., and Ellis, N.A. (1994) Bloom's syndrome: An analysis of consanguineous families assigns the locus mutated to chromosome band 15q26.1. *Proc. Nat. Acad. Sci. (USA)* **91:** 6669-6673

Goto, K., Maeda, S., Kano, Y., and Sugiyama, T. (1978) Factors involved in differential Giemsa-staining of sister chromatids. *Chromosoma* **66:** 351-359

Grawunder, U., Wilm, M., Wu, X., Kulesza, P., Wilson, T.E., Mann, M., and Lieber, M.R. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells *Nature* **388:** 492-495

Gu, H., Zou, Y.R., and Rajewsky, K. (1993) Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-*loxP*-mediated gene targeting. *Cell* **73(6):** 1155-1164

Gustavsson, P., Willig, T.-N., van Haeringen, A., Tchernia, G., Dianzani, I., Donner, M., Elinder, G., Henter, J.-I., Nilsson, P.-G., Gordon, L., Skeppner, G., van't Veer-Korthof, L., Kreuger, A., and Dahl, N. (1997) Diamond-Blackfan anaemia: genetic homogeneity for a gene on chromosome 19q13 restricted to 1.8Mb. *Nature Genet.* **16:** 368-371

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166:** 557-589

Hand, R., and German, J. (1975) A retarded rate of DNA chain growth in Bloom's syndrome. *Proc. Nat. Acad. Sci. (USA)* **72(2):** 758-762

Harris, S., McClenaghan, M., Simons, J.P., Ali, S., and Clark, J. (1991) Developmental regulation of the sheep β-lactoglobulin gene in the mammary gland of transgenic mice. *Dev. Gen.* **12:** 299-307

He, Z., and Ingles C.J. (1997) Isolation of human complexes proficient in nucleotide excision repair. *Nucleic Acids Res.* **25(6):** 1136-1141

Henderson, L.M., Arlett, C.F., Harcourt, S.A., Lehmann, A.R. and Broughton, B.C. (1985) Cells from an immunodeficient patient (46BR) with a defect in DNA ligation are hypomutable but hypersensitive to the induction of sister chromatid exchanges. *Proc. Natl. Acad. Sci. (USA)* **82:** 2044-2048

Hoeijmakers, J.H.J., and Bootsma, D. (1992) DNA repair: two pieces of the puzzle. *Nature Genet.* **1(8):** 313-314

Hosino, A., Hisayasu, S., and Shimada, T. (1996) Complete sequence analysis of rat transferrin and expression of transferrin but not lactoferrin in the digestive glands. *Comp. Biochem. Physiol.* **113:** 491-497

Hsieh, C.-L., Arlett, C.F., and Lieber, M.R. (1993) V(D)J recombination in ataxia telangiectasia, Bloom's syndrome, and a DNA ligase I-associated immunodeficiency disorder. *J. Biol. Chem.* **268:** 20105-20109

Husain, I., Tomkinson, A.E., Burkhart, W.A., Moyer, M.B., Ramos, W., Mackey, Z.B., Besterman, J.M., Chen, J.W. (1995) Purification and characterization of DNA ligase III from bovine testes. Homology with DNA ligase II and vaccinia DNA ligase. *J. Biol. Chem* **270(16):** 9683-9690

Huschtscha, L.I. and Holliday, R. (1983) Limited and unlimited growth of SV40-transformed cells from human diploid MRC-5 fibroblasts. *J. Cell Sci.* **63**: 77-99

Ish-Horowicz, D., and Burke, J.F. (1981) Rapid and efficient cosmid cloning. *Nucleic Acids Res.* **9:** 2989-2998

Jacks, T., Fazeli, A., Schmitt, E.M., Bronson, R.T., Goodell, M.A., and Weinberg, R.A. (1992) The effects of an Rb mutation in the mouse. *Nature* **359:** 295-300

Jackson, R.J., and Standardt, N. (1990) Do the poly (A) tail and 3' UTR control mRNA translation? *Cell* **62:** 15-24

Jackson, S.P., and Jeggo, P.A. (1995) DNA double-strand break repair and V(D)J recombination: involvement of DNA-PK *Trends Biochem. Sci.* **20:** 412-415

Jessberger, R., Podust, V., Hübscher, U., and Berg, P. (1993) A mammalian protein complex that repairs double-strand breaks and deletions by recombination. *J. Biol. Chem.* **268:** 15070-15079

Jessop, J.K. Structure and expression of the mouse DNA ligase I gene. Ph.D. dissertation, University of Edinburgh, 1995

Jessop, J.K., and Melton, D.W. (1995) Comparison between cDNA clones encoding murine DNA ligase I. *Gene* **160:** 307-308

Johnson, A.P., and Fairman, M.P. (1997) The identification and purification of a novel mammalian DNA ligase. *Mutat. Res. (DNA Repair)* **383:** 205-212

Jonsson, Z.O., Hindges, R., and Hubscher, U. (1998) Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *EMBO* **17(8):** 2412-2425

Kahn, Z., and Francis, G.E. (1987) Contrasting patterns of DNA strand breakage and ADP-ribosylation-dependent DNA ligation during granulocytic and monocyte differentiation. *Blood* **69(4):** 1114-1119

Kenne, K., and Ljunquist, S. (1988) Expression of a DNA ligase-stimulatory factor in Blooms syndrome cell line GM1492. *Eur. J. Biochem.* **174:** 465-470

Kilby, N.J., Snaith, M.R., and Murray, J.A., (1993) Site-specific recombinases: tools for genome engineering. *Trends in Genetics* **9:** 413-421

Kitamoto, T., Nakamura, K., Nakao, K., Shibuya, S., Shin, R.W., Gondo, Y., Katsuki, M., and Tateishi, J. (1996) Humanized prion protein knock-in by Cre-induced site-specific recombination in the mouse. *Biochem. Biophys. Res. Commun.* **222:** 742-747

Kodama, K., Barnes, D.E., Lindahl, T. (1991) *In vitro* mutagenesis and functional expression in *Escherichia coli* of a cDNA-encoding the catalytic domain of human DNA ligase I. *Nucl. Acids Res.* **19(22):** 6093-6099

Kühn, R., and Schwenk, F. (1997) Advances in gene targeting methods. *Curr. Opin. Immunol.* **9:** 183-188

Kunieda, T., Xian, M.W., Kobayashi, E., Imamichi, T., Moriwaki, K., and Toyoda, Y. (1992) Sexing of mouse preimplantation embryos by detection of Y-chromosome-specific sequences using polymerase chain reaction. *Biology of Reproduction* **46:**692-697

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685

Lagasse, E., and Weissman, I.L. (1992) Mouse MRP8 and MRP14, two intracellular calcium-binding proteins associated with the development of the myeloid lineage. *Blood* **79:** 1907-1915

Langlois, R.G, Bigbee, W.L., Jensen, R.H. and German, J. (1989) Evidence for increased *in vivo* mutation and somatic recombination in Bloom's syndrome. *Proc. Natl. Acad. Sci (USA)* **86:** 670-674

Lasko, D.D., Tomkinson, A.E., and Lindahl, T. (1990) Mammalian DNA ligases. Biosynthesis and intracellular localisation of DNA ligase I. *J. Biol. Chem.* **265:** 12618-12622

Latt, S.A., Schreck, R.R., Dougherty, C.P., Gustashaw, K.M., Juergens, L.A., and Kaiser, T.N. Sister-chromatid exchange - The phenomenon and its relationship to chromosome-fragility diseases. In *Chromosome Breakage and Neoplasia* (Alan R. Liss Inc., New York, 1983), pp. 169-178

Lehman, I.R. (1974) DNA ligase: Structure, mechanism and function. *Science* **186:** 790-797

Lehmann, A.R. (1995) Nucleotide excision repair and the link with transcription. *Trends Biochem. Sci.* **20:** 402-405

Lehmann, A.R., Willis, A.E., Broughton, B.C., James, M,R., Steingrimsdottir, H., Harcourt, S.A., Arlett, C.F. and Lindahl, T. (1988) Relation between the human fibroblast strain 46BR and cell lines representative of Bloom's syndrome. *Cancer Res.* **48:** 6343-6347

Lesch, M., and Nyhan, W.L. (1964) A familial disorder of uric acid metabolism and central nervous system function. *Am. J. Med.* **36:** 561-570

Levin, C.J., and Zimmerman, S.B. (1976) A DNA ligase from mitochondria of rat liver. *Biochem. Biophys. Res. Commun.* **69:** 514-520

Levin, D.S., Bai, W., Yao, N., ODonnell, M., Tomkinson, A.E. (1997) An interaction between DNA ligase I and proliferating cell nuclear antigen: Implications for Okazaki fragment synthesis and joining. *Proc. Natl. Acad. Sci. (USA)* **94(24):** 12863-12868

Li, C., Cao, L.G., Wang, Y.-L., and Baril, E.F. (1993) Further purification and characterisation of a multienzyme complex for DNA synthesis in human cells. *J. Cell. Biochem.* **53:** 405-419

Li, C., Goodchild, J., and Baril, E.F. (1994) DNA ligase I is associated with the 21S complex of enzymes for DNA synthesis in HeLa cells. *Nucl. Acids Res.* **22:** 632-638

Li, Z.-W., Stark, G., Gotz, J., Rulicke, T., Müller, U., and Weissmann, C. (1996) Generation of mice with a 200-kb amyloid precursor protein gene deletion by Cre recombinase-mediated site-specific recombination in embryonic stem cells. *Proc. Natl. Acad. Sci. (USA)* **93:** 6158-6162

Liang, P., and Pardee, A.B. (1992) differential display of eukaryotic messenger-RNA by means of the polymerase chain reaction. *Science* **257:** 967-971

Lindahl, T. (1994) DNA surveillance defect in cancer cells. *Curr. Biol.* **4(3):** 249-251

Lindahl, T. (1995) Recognition and processing of damaged DNA *J. Cell Sci.* **S19:** 73-77

Lindahl, T., and Barnes, D.E. (1992) Mammalian DNA ligases. *Annu. Rev. Biochem.* **61:** 251-281

Lindahl, T., and Edelman, G.M. (1968) Polynucleotide ligase from myeloid and lymphoid tissues *Proc. Natl. Acad. Sci. (USA)* **61:** 680-687

Lindahl, T., Karran, P., and Wood, R.D. (1997) DNA excision repair pathways. *Curr. Opin. Genet. Dev.* **7:** 158-169

Lindahl, T., Prigent, C., Barnes, D.E., Lehmann, A.R., Satoh, M.S., Roberts, E., Nash, R.A., Robins, P., and Daly, G. DNA joining in mammalian cells. In *Cold Spring Harbour symposia on quantitative biology* (Cold Spring Harbour Laboratory Press, Cold Spring Harbour, 1993) Volume LVIII, pp. 619-624

Ljungquist, S., Kenne, K., Olsson, L., and Sandström, M. (1994) Altered DNA ligase III activity in the CHO EM9 mutant. *Mutat. Res. (DNA Repair)* **314:** 177-186

Lönn, U., Lönn, S., Nylen, U., and Winblad, G. (1989) Altered formation of DNA replication intermediates in human 46BR fibroblast cells hypersensitive to 3-aminobenzamide. *Carcinogenesis* **10(6):** 981-985

Lönn, U., Lönn, S., Nylen, U., Winblad, G., and German, J. (1990) An abnormal profile of DNA replication intermediates in Bloom's syndrome. *Cancer Res.* **50:** 3141-3145

Loor, G., Zhang, S.J., Zhang, P., Toomey, N.L., and Lee, M.Y.W.T. (1997) Identification of DNA replication and cell cycle proteins that interact with PCNA. *Nucleic Acids Res.* **25(24):** 5041-5046

Lorimore, S.A., Pragnell, I.B., Eckman, L. and Wright, E.G. (1990) Synergistic interactions allow colony formation in vitro by murine haemopoietic stem cells. *Leuk. Res.* **14:** 481-489

MacDonald, R.J., Swift, G.H., Przybyla, A.E. and Chirgwin, J.M. (1987) Isolation of RNA using guanidinium salts. *Methods Enzymol.* **152:** 219-227

Mackenney, V.J., Barnes, D.E., Lindahl, T. (1997) Specific function of DNA ligase I in simian virus 40 DNA replication by human cell-free extracts is mediated by the amino-terminal non-catalytic domain. *J. Biol. Chem.* **272(17):** 11550-11556

Mackey, Z.B., Ramos, W., Levin, D.S., Walter, C.A., McCarrey, J.R., and Tomkinson, A.E. (1997) An alternative splicing event which occurs in mouse pachytene spermatocytes generates a form of DNA ligase III with distinct biochemical properties that may function in meiotic recombination. *Mol. Cell. Biol.* **17(2):** 989-998

Magin, T.M., McEwan, C., Milne, M., Pow, A.M., Selfridge, J., and Melton, D.W. (1992a) A position- and orientation-dependent element in the first intron is required for expression of the mouse *hprt* gene in embryonic stem cells. *Gene* **122:** 289-296

Magin, T.M., McWhir, J., and Melton, D.W. (1992b) A new mouse embryonic stem cell line with good germ line contribution and gene targeting frequency. *Nucleic Acids Res.* **20(14):** 3795-3796

Malkas, L.H., Hickey, R.J., Li, C., Pedersen, N., and Baril, E.F. (1990) A 21S enzyme complex from HeLa cells that functions in simian virus 40 DNA replication *in vitro*. *Biochem.* **29:** 6362-6374

Mandel, M. and Higa, A. (1970) Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* **53:** 159-162

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

Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J., deMurcia, G. (1998) XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Mol. Cell. Biol.* **18(6):** 3563-3571

McDaniel, L.D., and Schultz, R.A. (1992) Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. *Proc. Natl. Acad. Sci. (USA)* **89:** 7968-7972

McMahon, A.P. and Bradley, A. (1990) The *Wnt*-1 (*int*-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62:** 1073-1085

McPherson, I. and Stoker, M. (1962) Polyoma transformation of hamster cell clones: an investigation of genetic factors affecting cell competence.*Virology* **16:** 147-151

McWhir, J., Selfridge, J., Harrison, D.J., Squires, S., and Melton, D.W. (1993) Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nature Genet.* **5:** 217-224

Medvinsky, A.L., Samoylina, N.L., Müller, A.M., and Dzierzak, E.A. (1993) An early pre-liver intra-embryonic source of CFU-S in the developing mouse. *Nature* **364:** 64-67

Melton, D.W. (1994) Gene targeting in the mouse. *BioEssays* **16(9):** 633-638

Melton, D.W. Double replacement gene targeting in embryonic stem cells for the introduction of subtle alterations into endogenous mouse genes. In *Gene cloning and analysis: current innovations* (Horizon Scientific Press, Wymondham, 1997) pp. 147-163

Melton, D.W., Ketchen, A.-M., and Selfridge, J. (1997) Stability of *HPRT* marker gene expression at different gene-targeted loci: observing and overcoming a position effect. *Nucl. Acids. Res.* **25(19)**: 3937-3943

Metcalf, D. and Moore, M.A.S. *Haemopoietic Cells* (North-Holland, Amsterdam, 1971).

Micklem, H.S., Lennon, J.E., Ansell, J.D., and Gray, R.A. (1987) Numbers and dispersion of repopulating haematopoietic cell clones in radiation chimeras as functions of injected cell dose. *Exp. Hematol.* **15:** 251-257

Minty, A.J., Caravatti, M., Robert, B., Cohen, A., Daubas, P., Weydert, A., Gros, F., and Buckingham, M.E. (1981) Mouse actin messenger RNAs. Construction and characterization of a recombinant plasmid molecule containing a complementary DNA transcript of mouse α-actin mRNA. *J. Biol. Chem.* **256(2):** 1008-1014

Montecucco, A., Biamonti, G., Savini, E., Focher, F., Spadari, S., and Ciarrochi, G. (1992) DNA ligase I gene expression during differentiation and cell proliferation. *Nucleic Acids Res.* **20:** 6209-6214

Montecucco, A., Rossi, R., Levin, D.S., Gary, R., Park, M.S., Motycka, T.A., Ciarrocchi, G., Villa, A., Biamonti, G., and Tomkinson A.E. (1998) DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: indentification of a common targeting mechanism for the assembly of replication factories. *EMBO* **17(13):** 3786-3795

Montecucco, A., Savini, E., Biamonti, G., Stefanini, M., Focher, F., and Ciarrocchi, G. (1995a) Late induction of human DNA ligase I after UV-C irradiation. *Nucl. Acids Res.* **23(6):** 962-966

Montecucco, A., Savini, E., Weighardt, F., Rossi, R., Ciarrocchi, G., Villa, A., and Biamonti, G. (1995b) The N-terminal domain of human DNA ligase I contains the nuclear localization signal and directs the enzyme to sites of DNA replication. *EMBO* **14(21):** 5379-5386

Moore, R.C. (1997) Gene targeting studies at the mouse prion protein locus. Ph.D. dissertation, University of Edinburgh, 1997

Moore, R.C. and Melton, D.W. Gene Targeting. In *The molecular and genetic basis of neurological disease.* 2nd edition pp33-48 Eds. Rosenberg, R.N, Prusiner, S.B., DiMauro, S. and Barchi, R.L. (Butterworth-Heinmann, 1997)

Moore, R.C., Redhead, N.J., Selfridge, J., Hope, J., Manson, J.C., and Melton D.W. (1995) Double replacement gene targeting for the production of a series of mouse strains with different prion gene alterations. *Bio/Technology* **13:** 999-1004

Mortensen, R.M., Conner, D.A., Chao, S., Geisterfer-Lowrance, A.A.T., and Seidman, J.G. (1992) Production of homozygous mutant ES cells with a single targeting construct. *Mol. Cell. Biol.* **12:** 2391-2395

Mossi, R., Ferrari, E., and Hubscher, U. (1998) DNA ligase I selectively affects DNA synthesis by DNA polymerases delta and epsilon suggesting differential functions in DNA replication and repair. *J. Biol. Chem.* **273(23):** 14322-14330

Mouse Genome Database (1997)..Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine (http://www.informatics.jax.org/)

Murao, S. (1994) Two calcium-binding proteins, MRP8 and MRP14 - a protein complex-associated with neutrophil and monocyte activation. *Acta Histochemica et Cytochemica* **27(2):** 107-115

Murray, B., Irwin, J., Creissan, D., Tavassoli, M., Durkacz, and Shall, S. (1986) A mammalian cell variant in which 3-aminobenzamide does not potentiate the cytotoxicity of dimethyl sulphate. *Mutat. Res. (DNA Repair)* **165:** 191-198.

Nash, R.A., Caldecott, K.W., Barnes, D.E., Lindahl, T. (1997) XRCC1 protein interacts with one of two distinct forms of DNA ligase III. *Biochem.* 36(17): 5207-5211

Nicholl, I.D., Nealon, K., and Kenney, M.K. (1997) Reconstitution of human base excision repair with purified proteins. *Biochemistry* **36(24):** 7557-5566

Nocentini, S. (1995) Comet assay analysis of DNA strand breaks in normal and deficient human cells exposed to radiations and chemicals. Evidence for a repair pathway specificity of DNA ligation. *Radiation Res.* **144:** 170-180

Noguiez, P., Barnes, D.E., Mohrenweiser, H.W. and Lindahl, T. (1992) Structure of the human DNA ligase I gene. *Nucleic Acids Res.* **20:** 3845-3850

Ohbo, K., Suda, T., Hashiyama, M., Mantani, A., Ikebe, M., Miyakawa, K., Moriyama, M., Nakamura, M., Katsuki, M., Takahashi, K., Yamamura, K.-I., and Sagamura, K. (1996) Modulation of hematopoiesis in mice with a truncated mutant of the interleukin-2 receptor γ chain. *Blood* **87(3):** 956-967

Online Mendelian Inheritance in Man (OMIM), Centre for Medical Genetics, Johns Hopkins University (Baltimore, MD), and National Centre for Biotechnology Information, National Library of Medicine (Bethesda, MA) 1997. http://www.ncbi.nlm.nih.gov/omim/

Ozawa, T., Kondo, N., Motoyoshi, F., Kato, Y., and Orii, T. (1993) Preferential damage to IgM production by ultraviolet B in the cells of patients with Bloom's syndrome. *Scand. J. Immunol.* **38:** 225-232

Pantel, K., Loeffler, M., Bungart, B., and Wichmann, H.E. (1990) A mathematical model of erythropoiesis in mice and rats. Part 4: Differences between bone marrow and spleen. *Cell Tissue Kinet.* **23:** 283-297

Passarge, E. "Bloom's syndrome" in *Chromosome Breakage and Neoplasia* (Alan R. Liss Inc., New York, 1983), pp. 135-167

Perry, P., and Wolff, S. (1974) New Giemsa method for the differential staining of sister chromatids. *Nature* **251:** 156-158

Petrini, J.H.J, Donovan, J.W., Dimare, C., and Weaver, D.T. (1994) Normal V(D)J coding junction formation in DNA ligase I-deficiency syndromes. *J. Immunol.* **152:** 176-183

Petrini, J.H.J., Huwiler, K.G., and Weaver, D.T. (1991) A wild-type DNA ligase I gene is expressed in Bloom's syndrome cells. *Proc. Natl. Acad. Sci. (USA)* **88:** 7615-7619

Petrini, J.H.J., Xiao, Y., and Weaver, D.T. (1995) DNA ligase I mediates essential functions in mammalian cells. Mol. Cell. Biol. **15:** 4303-4308

Poole, T. (ed.), *The UFAW Handbook on the Care and Management of Laboratory Animals* 6th edn (Bath, Longman Scientific and Technical, 1989).

Prasad, R., Singhal, R.K., Srivastava, D.K., Molina, J.T., Tomkinson, A.E., and Wilson, S.H. (1996) Specific interaction of DNA polymerase beta and DNA ligase I in a multiprotein base excision repair complex from bovine testis. *J. Biol. Chem.* **271(27):** 16000-16007

Prigent, C., Lasko, D.D., Kodama, K.-I., Woodgett, J.R., and Lindahl, T. (1992) Activation of mammalian DNA ligase I through phosphorylation by casein kinase II. *EMBO* **11:** 2925-2933

Prigent, C., Satoh, M.S., Daly, G., Barnes, D.E. and Lindahl, T. (1994) Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. *Mol. Cell Biol.* **14:** 310-317

Ramírez-Solis, R., Liu, P., and Bradley, A. (1995) Chromosome engineering in mice. *Nature* **378:** 720-724

Ramos, W., Tappe, N., Talamantez, J., Friedberg, E.C., and Tomkinson A.E. Two distinct DNA ligase activities in mitotic extracts of yeast Saccharomyces cerevisiae. *Nucleic Acids Res.* **25(8):** 1485-1492

Ramsden, D.A., and Gellert, M. (1998) Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role fro Ku in repair of DNA double-strand breaks. *EMBO J.* **17(2):** 609-614

Ramsden, D.A., Paull, T.T., and Gellert, M. (1997) Cell-free V(D)J recombination. *Nature* **388:** 488-491

Ray, J.H., and German, J. The cytogenetics of the "chromosome breakage syndromes". In *Chromosome Breakage and Neoplasia* (Alan R. Liss Inc., New York, 1983), pp. 135-167

Reed, K.C. and Mann, D.A. (1985) Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13:** 7207-7221

Reid, L.H., Gregg, R.G., Smithies, O., and Koller, B.H. (1990) Regulatory elements in the introns of the human *HPRT* gene are necessary for its expression in embryonic stem cells. *Proc. Natl. Acad. Sci. USA.* **87:** 4299-4303

Roberts, E., Nash, R.A, and Lindahl, T. (1994) Different active sites of mammalian DNA ligases I and II. *J. Biol. Chem.* **269:** 3789-3792

Robins, P., and Lindahl, T. (1996) DNA ligase IV from HELA cell nuclei. *J. Biol. Chem.* **271(39):** 24257-24261

Rosen, A., and Klein, G. (1983) UV light-induced immunoglobulin heavy-chain class switch in a human lymphoblastoid cell line. *Nature* **306:** 189-190

Rumbaugh, J.A., Murante, R.S., Shi, S., and Bambara, R.A (1997) Creation and removal of embedded ribonucleotides in chromosomal DNA during mammalian Okazaki fragment processing. *J. Biol. Chem* 272**(36):** 22591-22599

Rünger, T.M., and Kraemer, K.H. (1989) Joining of plasmid DNA is reduced and error-prone in Bloom's syndrome cells. *EMBO* **8:** 1419-1425

Rusquet, R.M., Feon, S.A., and David, J.C. (1988) Association of a possible DNA ligase deficiency with T-cell acute leukemia. *Cancer Res.* **48:** 4038-4044

Sambrook, J., Fritsch, E.F., and Maniatis, T. *Molecular Cloning. A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1989)

Savini, E., Biamonti, G., Ciarrocchi, G., and Montecucco, A. (1994) Cloning and sequence analysis of a cDNA coding for the murine DNA ligase I enzyme. *Gene* **144:** 253-257

Schär, P., Herrmann, G., Daly, G., and Lindahl, T. (1997) A newly identified DNA ligase of *Saccharomyces cerevisiae* involved in *RAD52*-independent repair of DNA double-strand breaks. *Genes. Dev.* **11:** 1912-1924

Scriver, C.R. *The metabolic and molecular bases of inherited disease.* 7th edition. (McGraw-Hill, London, 1995)

Selbert, S., Bentley, D.J., Melton, D.W., Rannie, D., Lourenço, P., Watson, C., and Clarke, A.R. (1998) Efficient BLG-Cre mediated gene deletion in the mammary gland. *Transgenic Research* **7:** 387-396

Selfridge, J., Pow, A.M., McWhir, J., Magin, T.M. and Melton, D.W. (1992) Gene targeting using a mouse HPRT minigene/HPRT-deficient embryonic stem cell system: inactivation of the mouse *ERCC-1* gene. *Somat. Cell Mol. Genet.* **18:** 325-336

Signoret, J., and David, J.-C. (1986) Control of the expression of genes for DNA ligase in eukaryotes. *Int. Rev. Cytol.* **103:** 249-279

Simpson, E.M., Linder, C.C., Sargent, E.E., Davisson, M.T., Mobraaten, L.E., and Sharp, J.J. (1997) Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nat. Genet.* **16:** 19-27

Smith, A.J.H., De Sousa, M.A., Kwabi-Addo, B., Heppell-Parton, A., Impey, H., and Rabbitts, P. (1995) A site-directed chromosomal translocation induced in embryonic stem cells by Cre-*loxP* recombination. *Nature Genet.* **9:** 376-385

Smith, G.E. and Summers, M.D. (1980) The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. *Anal. Biochem* **190:** 123-129

Smith, G.L., Chan, Y.S., and Kerr, S.M. (1989) Transcriptional mapping and nucleotide sequence of a vaccinia virus gene encoding a polypeptide with extensive homology to DNA ligases. *Nucleic Acids Res.* **17:** 9051-9062

Söderhäll, S. (1976) DNA ligases during rat liver regeneration.*Nature* **260:** 640-642

Söderhäll, S., and Lindahl, T. (1973) Mammalian deoxyribonucleic acid ligase. Isolation of an active enzyme-adenylate complex. *J. Biol. Chem.* **248:** 672-675

Somia, N.V., Jessop, J.K., and Melton, D.W. (1993) Phenotypic correction of a human cell line (46BR) with aberrant DNA ligase activity *Mutat. Res. (DNA Repair)* **294:** 51-58

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by agarose gel electrophoresis. *J. Mol. Biol.* **98:** 503-517

Squires, S., and Johnson, R.T. (1983) U.V. induces long-lived DNA breaks in Cockayne's syndrome and cells from an immunodeficient individual (46BR): defects and disturbance in post incision steps of excision repair. *Carcinogenesis* **4:** 565-572

Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A., and Melton, D.W. (1994) Use of double-replacement gene targeting to replace the murine α-lactalbumin gene with its human counterpart in embryonic stem cells and mice. *Mol. Cell. Biol.* **14(2):** 1009-1016

Steihm, E.R. (1993) New and old immunodeficiencies *Pediatr. Res.* **33(1)SS:** 2-8

Strohman, R.C., Moss, P.S., MicouEastwood, J., Spector, D., Przybyla, A. and Paterson, B. (1977) Messenger RNA for myosin polypeptides: isolation from single myogenic cell culture. *Cell* **10**: 265-273

Subramanya, H.S., Doherty, A.J., Ashford, S.R., and Wigley, D.B. (1996) Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. *Cell* **85(4):** .607-615

Tanaka, K., and Wood, R.D. (1994) Xeroderma pigmentosum and nucleotide excision repair of DNA. *Trends Biochem. Sci.* **19:** 83-86

Teo, I.A., and Arlett, C.F. (1982) The response of a variety of human fibroblast cell strains to the lethal effects of alkylating agents. *Carcinogenesis* **3:** 33-37

Teo, I.A., Arlett, C.F., Harcourt, S.A., Priestley, A., and Broughton, B.C. (1983) Multiple hypersensitivity to mutagens in a cell strain (46BR) derived from a patient with immuno-deficiencies. *Mutat. Res.* **107:** 371-386

Teo, S.-H., and Jackson, S.P. (1997) Identification of *Saccharomyces cerevisiae* DNA ligase IV: involvement in DNA double-strand break repair *EMBO J.* **16:** 101-108

Teraoka, H., and Tsukada, K. (1982) Eukaryotic DNA ligase. Purification and properties of the enzyme from bovine thymus, and immunochemical studies of the enzyme from animal tissues. *J. Biol. Chem.* **257:** 4758-4763

Thomas, K.R., and Capecchi, M.R. (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51:** 503-512

Thompson, L.H., Brookman, K.W., Jones, N.J., Allen, S.A., and Carrano, A.V. (1990) Molecular cloning of the human XRCC1 gene which corrects defective DNA strand break repair and sister chromatid exchange. *Mol. Cell. Biol.* **10:** 6160-6171

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., and Melton D.W. (1989) Germ line transmission and expression of a corrected *HPRT* gene produced by gene targeting in embryonic stem cells. *Cell* **56:** 313-321

Thummel, C.S., Boulet, A.M. and Lipshitz, H.D. (1988) Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74:** 445-456

Tomkinson, A.E., and Levin, D.S. (1997) Mammalian DNA ligases *Bioessays* **19(10):** 893-901

Tomkinson, A.E., Lasko, D.D., Daly, G., and Lindahl, T. (1990) Mammalian DNA ligases. Catalytic domain and size of DNA ligase I *J. Biol. Chem.* **265:** 12611-12617

Tomkinson, A.E., Roberts, E., Daly, G., Totty, N. and Lindahl, T. (1991) Three distinct DNA ligases in mammalian cells. *J. Biol. Chem.* **266:**21728-21735

Tomkinson, A.E., Starr, R., and Schultz, R.A. (1993) DNA ligase III is the major high molecular weight DNA joining activity in SV40-transformed human fibroblasts: normal levels of DNA ligase III activity in Bloom syndrome cells. *Nucleic Acids Res.* **21:** 5425-5430

Tomkinson, A.E., Totty, N.F., Ginsburg, M., and Lindahl, T. (1991) Location of the active site for enzyme-adanylate formation in DNA ligases *Proc. Nat. Acad. Sci.* (USA) **88:** 400-404

Torres, R.M., Flaswinkel, H., Reth, M., and Rajewsky, K. (1996) Aberrant B cell development and immune response in mice with a compromised BCR complex. *Science* **272:** 1804-1808

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994) An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* **371:** 221-226

Turchi, J.J, and Bambara, R.A. (1993) Completion of mammalian lagging strand DNA replication using purified proteins *J. Biol. Chem.* **268(20):** 15136-15141

Turchi, J.J., Huang, L., Murante, R.S., Kim, Y., and Bambara, R.A. (1994) Enzymatic completion of mammalian lagging-strand DNA replication. *Proc. Nat. Acad. Sci. (USA)* **91:** 9803-9807

Vijayalaxmi, Evans, H.J., Ray, J.H., German, J. (1983) Bloom's syndrome: evidence for an increased mutation frequency in vivo. *Science* **221:**851-853

Waga, S., Bauer, G., and Stillman, B. (1994) Reconstitution of complete SV40 DNA replication with purified replication factors *J. Biol. Chem.* **269(14):** 10923-10934

Wang, Y.-C.J., Burkhart, W.A., Mackey, Z.B., Moyer, M.B., Ramos, W., Husain, I., Chen, J., Besterman, J.M., and Tomkinson, A.E. (1994) Mammalian DNA ligase II is homologous with Vaccinia DNA ligase *J. Biol. Chem.* **269(50):** 31923-31928

Warren, S.T., Schultz, R.A., Chang, C.C., Wade, M.H., and Trosko, J.E. (1981) Elevated spontaneous mutation rate in Bloom's syndrome fibroblasts *Proc. Nat. Acad. Sci (USA).* **78:** 3133-3137

Weaver, D.T. (1995) What to do at an end: DNA double-strand-break repair. *Trends Genet.* **11(10):** 388-392

Webster, A.D.B., Barnes, D.E., Arlett, C.F., Lehmann, A.R. and Lindahl, T. (1992) Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene. *Lancet* **339:** 1508-1509

Wei, Y.-F., Robins, P., Carter, K., Caldecott, K., Pappin, D.J.C., Yu, G.-L., Wang, R.-P., Shell, B.K., Nash, R.A., Schar, P., Barnes, D.E., Haseltine, W.A., and Lindahl, T. (1995) Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. *Molec. Cell. Biol.* **15(6):** 3206-3216

Weksberg, R., Smith, C., Anson-Cartwright, L., and Maloney, K. (1988) Bloom syndrome: A single complementation group defines patients of diverse ethnic origin. *Am. J. Hum. Genet.* **42:** 816-824

Whitelaw, C.B.A., Harris, S., McClenaghan, M., Simons, J.P., and Clarke, A.J. (1992) Position-independent expression of the ovine β-lactoglobulin gene in transgenic mice. *Biochem. J.* **286:** 31-39

Willis, A.E., and Lindahl, T. (1987) DNA ligase I deficiency in Bloom's syndrome. *Nature* **325:** 355-357

Willis, A.E., Weksberg, R., Tomlinson, S. and Lindahl, T. (1987) Structural alterations of DNA ligase I in Bloom syndrome. *Proc. Natl. Acad. Sci. (USA)* **84:** 8016-8020

Wilson, T.E., Grawunder, U., and Lieber, M.R. (1997) Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* **388:** 495-498

Wu, W., Hickey, R., Lawlor, K., Wills, P., Yu, F., Ozer, H., Starr, R., Quan, J.Y., Lee, M., and Malkas, L. (1994) A 17S multiprotein form of murine cell DNA polymerase mediates polyomavirus DNA replication in vitro. *J. Cell. Biochem.* 54: 32-46

Yang, S., Becker, F.F., and Chan, J.Y.-H. (1993) Biochemical characterisation of a protein inhibitor for DNA ligase I from human cells. *Biochem. Biophys. Res. Comm.* **191:** 1004-1013

Zou, H., and Rothstein, R. (1997) Holliday junctions accumulate in replication mutants via a RecA homologue-independent mechanism. *Cell* **90:** 87-96

DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammaliam cell viability

Darren J. Bentley, Jim Selfridge, J. Kirsty Millar, Kay Samuel,
Nicholas Hole, John D. Ansell & David W. Melton

Reprinted from Nature Genetics, Vol. 13, August 1996

DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability

Darren J. Bentley[1], Jim Selfridge[1], J. Kirsty Millar[1,3], Kay Samuel[2], Nicholas Hole[2], John D. Ansell[2] & David W. Melton[1]

Four distinct DNA ligase activities (I-IV) have been identified within mammalian cells[1–3]. Evidence has indicated that DNA ligase I is central to DNA replication[4–7], as well as being involved in DNA repair processes[8,9]. A patient with altered DNA ligase I displayed a phenotype similar to Bloom's syndrome, being immunodeficient, growth retarded and predisposed to cancer[10]. Fibroblasts isolated from this patient (46BR) exhibited abnormal lagging strand synthesis[11,12] and repair deficiency[13–15]. It has been reported that DNA ligase I is essential for cell viability[16], but here we show that cells lacking DNA ligase I are in fact viable. Using gene targeting in embryonic stem (ES) cells, we have produced DNA ligase I-deficient mice. Embryos develop normally to mid-term, when haematopoiesis usually switches to the fetal liver. Thereupon acute anaemia develops, despite the presence of erythroid-committed progenitor cells in the liver. Thus DNA ligase I is required for normal development, but is not essential for replication. Hence a previously unsuspected redundancy must exist between mammalian DNA ligases.

One patient has been identified with reduced DNA ligase I activity correlating to missense mutations within the protein coding sequence[12,17]. Gene inactivation is the first step of a 'double replacement' strategy[18] to introduce analogous mutations into the mouse DNA ligase I gene (*Lig1*). Utilizing positive-negative selection in the hypoxanthine phosphoribosyltransferase (*HPRT*)-deficient ES cell line HM-1[19], we replaced exons 23 to 27 of *Lig1* with an *HPRT* minigene (Fig. 1*a*). Targeted ES cells were injected into host blastocysts, and resulting germ line chimaeras outbred to produce mice heterozygous for the targeted allele (+/–). Heterozygotes were phenotypically normal, but when interbred, no pups homozygous for the altered allele (–/–) were found. From this it was concluded that the absence of DNA ligase I is developmentally lethal.

To define the developmental defect, embryos ranging from day 10.5 to day 16.5 postcoitum (E10.5–E16.5) were collected from heterozygous matings and genotyped. Live embryos of the three genotypes were obtained in the expected ratios at E10.5 [$n = 77$, $\chi^2_2 = 0.74$ (NS)], and there was no significant loss of any genotype between E11.5 and E15.5 [$n = 167$, $\chi^2_2 = 2.50$ (NS)]. However, while the majority of –/– embryos remained alive at E15.5, all 6 isolated at E16.5 were dead.

To ascertain the nature of any mRNA or protein produced from the targeted *Lig1* allele, northern and western analysis was performed. No intact, truncated or fusion transcripts from the *Lig1* allele could be detected in –/– embryos using cDNA probes corresponding to either exons 8 to 27 (Fig. 1*b*) or exons –1 to 7 (data not shown). Similarly, no peptide of any kind could be detected in protei extracts from –/– embryos when probed with antibodie raised against bovine DNA ligase I (Fig. 1*c* and data nc shown). We conclude that expression from the targete allele is abrogated and DNA ligase I is completely abser from –/– embryos.

At E10.5, –/– embryos were visually indistinguishabl from their wild-type and heterozygous siblings. Howeve at E11.5 all –/– embryos were paler and smaller than nor mal (Fig. 2*a*). By E15.5 this anaemia had become exacer bated (Fig. 2*c*). Although smaller, –/– embryos appeare grossly normal (Fig. 2*e*). Histological examination con firmed that, with the exception of the developing live organogenesis was not perceptibly affected (data no shown).

From E11.5 onwards, the liver is the major site of bloo

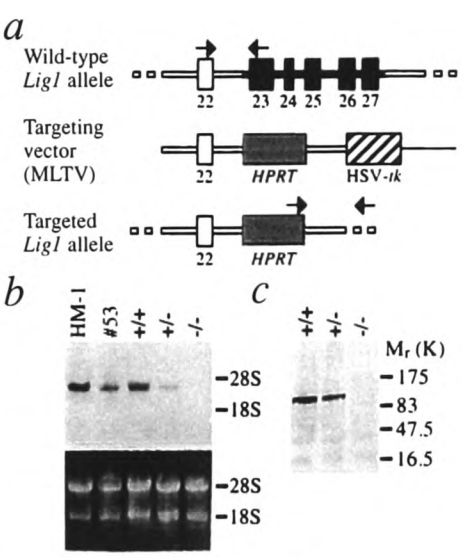


Fig. 1 Targeted disruption of *Lig1* in mouse embryonic stem cells, and expression of the wild-type allele. *a*, Schematic representation of 13.5 kb at the 3' end of the wild-type mouse DNA ligase I locus, targeting vector and gene structure following homologous recombination. Boxes, exons (numbered according to the human convention[31]); thick lines, *Lig1* introns or flanking regions; thin lines, plasmid vector DNA. Solid boxes and lines, regions of *Lig1* replaced by targeting; stippled box, *HPRT* minigene; hatched box, Herpes simplex virus thymidine kinase (HSV-*tk*) cassette. Arrows indicate the positions at which PCR primer pairs used for genotyping bind. *b*, Absence of detectable expression from the targeted *Lig1* allele. Upper, northern analysis: original ES cell line (HM-1), one targeted ES cell clone (#53), and E13.5 embryos of the three genotypes. Positions of the 28S and 18S ribosomal RNA transcripts are indicated flanking the 3.2-kb wild-type *Lig1* transcript. The probe used in the figure was a 2.1-kb *Bam*HI/*Eco*RI fragment of mouse DNA ligase I cDNA, corresponding to exons 8 to 27. Lower, ethidium bromide staining to verify equal RNA loading. *c*, Absence of detectable DNA ligase I in –/– embryos. Western blot on total protein extracts from E15.5 embryos. Positions of molecular weight markers are indicated. *d*, Embryonic expression from the wild-type *Lig1* locus. Upper, northern analysis on a wild-type panel of E10.5 embryo, E10.5 yolk sac, E13.5 embryo carcass minus liver, E13.5 embryonic liver and E15.5 yolk sac. Lower, ethidium bromide staining to verify equal RNA loading. *e*, Expression from the wild-type *Lig1* locus. Upper, northern analysis on a panel of tissues from wild-type adult mice. Lower, ethidium bromide staining to verify equal RNA loading.

[1]*Institute of Cell and Molecular Biology, University of Edinburgh, Darwin Building, King's Buildings, Mayfield Road, Edinburgh, EH9 3JR, UK*
[2]*Institute of Cell, Animal and Population Biology, University of Edinburgh, Ashworth Building, King's Buildings, Mayfield Road, Edinburgh, EH9 3JT, UK*
[3]*Present address: MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK*

Correspondence should be addressed to D.W.M.

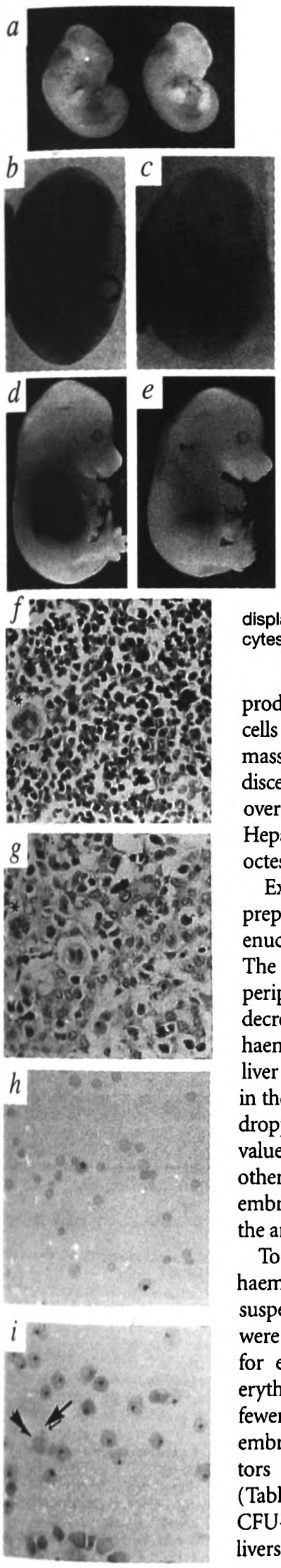


Fig. 2 Phenotype of –/– embryos. *a*, Morphology of wild-type (left) and –/– (right) siblings at E11.5. The –/– embryo is paler and smaller than its sibling. *b–e*, Morphology of wild-type (*b* and *d*) and –/– (*c* and *e*) embryos at E15.5 before (*b*, *c*) and after (*d*, *e*) dissection from the yolk sac. The –/– embryos are severely anaemic, smaller than their wild-type siblings, and the liver is disproportionately reduced in size, but otherwise developmentally normal. *f*, *g*, Histology of livers from wild-type (*f*) and –/– (*g*) E15.5 embryos. The wild-type liver contains numerous islands of erythropoietic cells, and megakaryocytes (asterisked). The –/– liver lacks extensive cellularization and contains no readily identifiable erythropoietic islands, although there is no evidence of increased apoptosis. Hepatocytes appear normal and megakaryocytes (asterisked) are present in normal numbers. *h*, *i*, Cytospin preparations of peripheral blood from wild-type (*h*) and –/– (*i*) E15.5 embryos. Blood from wild-type embryos consists predominantly of mature, enucleated liver-derived erythrocytes, although a few residual nucleated yolk sac-derived erythrocytes are present. In contrast, samples from –/– embryos display an elevated ratio of nucleated to enucleated erythrocytes. Those enucleated cells present display poikilocytosis and anisocytosis, with microcytes (arrow) and macrocytes (arrowhead) evident.

production in the developing embryo, erythroid cells comprising up to 70% of the total liver mass[20]. Livers from –/– embryos contained no discernible erythropoietic islands, giving an overall appearance of acellularity (Fig. 2*g*). Hepatocytes appeared normal, as did megakaryocytes.

Examination of embryonic blood cytospin preparations revealed a deficiency of mature enucleated erythrocytes in –/– embryos (Fig. 2*i*). The percentage of nucleated erythrocytes in the peripheral circulation did not show the normal decrease characteristic of the shift of haematopoiesis from the yolk sac to the fetal liver (Fig. 3). By E15.5 the total number of cells in the peripheral circulation of –/– embryos had dropped to approximately 5% of the wild-type value (data not shown). In the absence of any other visible abnormalities, we postulate that –/– embryos die as a result of hypoxia secondary to the anaemia.

To examine the differentiation capacity of haematopoietic precursors *in vitro*, single-cell suspensions from –/– and control fetal livers were plated under growth conditions optimized for either primitive multipotent (CFU-A) or erythroid (BFU-E) progenitors. Consistently fewer total cells were present in livers of –/– embryos, and numbers of both types of progenitors were also reduced compared to controls (Table 1). Qualitative scoring also revealed both CFU-A and BFU-E colonies derived from –/– livers to be smaller, although cell morphology appeared normal (data not shown). Antibody staining for the cell surface marker T200 indicated that the total number of non-erythroid haematopoietic cells in –/– livers was normal (Table 1). Similarly, the numbers of B220-positive B lineage cells detected in –/– livers were not reduced at this stage (data not shown). These data suggest that, in contrast to the erythroid lineage, members of non-erythroid lineages are relatively unaffected by the lack of DNA ligase I. However, liver suspensions from –/– embryos were unable to rescue lethally irradiated mice (Table 1), indicating the defect is not simply the consequence of an inappropriate hepatic microenvironment.

We have shown that the absence of DNA ligase I has no discernible consequence for most tissues in the developing mouse embryo, the growth retardation observed after E10.5 probably arising as a consequence of the anaemia. No significant abnormality was observed in the yolk sac of –/– embryos at E10.5, and yolk sac erythropoiesis was apparently normal. The specific disruption of liver erythropoiesis alone may reflect the extremely high replicative pressure on this system, embryonic expression of DNA ligase I being highest in E13.5 liver (Fig. 1*d*). Failure of another DNA ligase to completely complement the absence of DNA ligase I would be likely to preclude exponential proliferation. The normal geometric expansion of erythroid cells would thus be abbreviated, resulting in fewer terminally differentiated erythrocytes and anaemia.

DNA ligase I was thought fundamental to lagging strand DNA synthesis. Expression is induced upon cell proliferation[21–23], being highest in embryos and rapidly dividing adult tissues such as spleen, testes and thymus (Fig. 1*e*). The protein has been isolated from replication complexes[4,5], and can function in replication *in vitro*[6,7]. As 46BR cells display abnormal replication intermediate profiles[11,12,14], a role in ligation of Okazaki fragments was proposed. Inability to isolate ES cells lacking DNA ligase I *in vitro*, led Petrini *et al.* to conclude that the enzyme mediates essential functions in mammalian cells[16]. However, our data indicate that DNA ligase I-deficient cells are viable *in vivo*. We believe that this discrepancy stems mainly from an inherent difficulty in isolating ES cells homozygous for a disabling mutation *in vitro*. For example, we have described the production of *Ercc1* deficient mice[24], but have been unable to produce ES cells homozygous for the *Ercc1* mutation. These cells clearly exist *in vivo*, but cannot be isolated *in vitro*, presumably because they cannot survive the demanding selection conditions at low cell density.

Abnormal DNA ligase I activity has been documented in cells from Bloom's syndrome patients, and although *LIG1* is not the Bloom's syndrome gene (*BLM*)[25], interaction between the two gene products seems probable. Studies of DNA ligase I-deficient cells thus may help define the protein interactions occurring within the replication complex. Exactly which other ligase activity replaces DNA ligase I is unclear at present, but a previously unsuspected redundancy must exist among the mammalian DNA ligases.

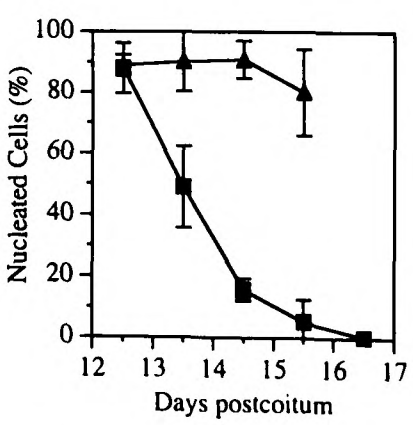


Fig. 3 Incidence of nucleated erythrocytes in the embryonic circulation. The percentage of nucleated cells were determined from cytospin preparations of embryonic peripheral blood between E11.5 and E16.5. A minimum of 500 cells from at least two separate embryos of each genotype were scored. Error bars indicate the standard deviation at each sample point. ■, wild-type; ▲, –/–.

Table 1 Haematopoietic cell numbers per fetal liver and *in vivo* repopulation analysis

Genotype	Total Cells $\times10^{-6}$ at E13.5	CFU-A progenitors $\times10^{-4}$	BFU-E progenitors $\times10^{-4}$	Non-erythroid haematopoietic cells $\times10^{-6}$	Irradiated mice surviving
Wild-type	8.4 ± 0.8	3.1 ± 0.5	1.2 ± 0.2	1.0 ± 0.4	11/13
-/-	3.6 ± 1.0	1.6 ± 0.2	0.4 ± 0.1	1.3 ± 0.6	0/7

All cell numbers are expressed per fetal liver. Cell suspensions were produced by successive passage through 19G and 25G needles, and the number of single cells counted. Numbers of haematopoietic progenitors per isolated E13.5 fetal liver (mean ± s.d.) were determined by standard assays[29,30]. Equal numbers of single cells (cell equivalents) were plated out in conditions optimized for the proliferation and differentiation of primitive progenitors (CFU-A assay), or erythroid differentiation (BFU-E assay), and colony numbers scored after 11 and 8 days respectively. Numbers of non-erythroid haematopoietic cells (mean ± s.d.) were determined by antibody staining. Cell suspensions from E15.5 fetal livers were stained with a PE-conjugated, rat monoclonal antibody raised against mouse CD45 (T200) (Clone 30F 11.1, Pharmingen). For each sample, 25,000 to 50,000 cells were analysed by FACScan, and numbers of T200 positive cells estimated. For *in vivo* repopulation analysis, suspensions from individual E13.5 livers (liver equivalents) were injected into the tail vein of strain 129 or 129/CBA F1 mice exposed to 10.5 Gy of γ radiation. Numbers of mice surviving after 4 weeks were recorded. All irradiated but uninjected control mice died within 3 weeks.

Methods

Gene Targeting. A library of *Eco*RI-digested genomic DNA from the ES cell line E14TG2a was constructed in λ EMBL4. Screening using full-length mouse DNA ligase 1 cDNA as a probe[26], enabled a 13.5-kb fragment, encompassing exons 22 to 27, to be isolated and subcloned into the *Eco*RI site of plasmid pUC8. Replacement of a 4.9-kb *Sst*I and a 3.7-kb *Bam*HI fragment produced a targeting vector (MLTV) in which an *HPRT* minigene is flanked by 3.3-kb (5') and 1.8-kb (3') regions homologous to *Lig1*, plus an HSV-*tk* cassette outwith. MLTV was linearized by digestion with *Eco*RI, and introduced into HM-1 ES cells by electroporation[19]. Transfected cells were enriched for targeting events by growth in HAT medium supplemented with gancyclovir. PCR screening and subsequent Southern blot analysis (data not shown) identified 2 out of 115 *hprt+ tk-* clones as having undergone the predicted homologous recombination event. Targeted ES cells were injected into host blastocysts, and one of the resulting chimaeric males was outbred with BALB/c, to produce mice heterozygous for the targeted allele.

Northern and western analysis. Total RNA was extracted from ES cell lines and tissues and subjected to northern analysis as described[27]. Protein lysates were made by brief homogenisation of tissue in lysis buffer (adapted from ref. 12), and 50 μg loaded onto an 8% SDS-polyacrylamide gel. Proteins were blotted onto a nitrocellulose membrane, blocked in TBST plus 5% non-fat milk powder, and incubated with rabbit polyclonal antibody raised against purified bovine DNA ligase I[28] (kindly provided by T. Lindahl and D. Barnes), which predominantly recognizes epitopes in the N terminus of the protein. Detection of bound antibody was achieved using alkaline phosphatase-conjugated anti-rabbit IgG antibodies.

Embryo analysis. Conceptuses were isolated from timed matings between +/– animals, the day on which a vaginal plug was found being designated as E0.5. Yolk sacs were used to prepare DNA for genotyping by PCR, while embryos were fixed in 4% phosphate buffered formaldehyde, embedded in paraffin and sectioned at 5 μm. Sections were stained with haematoxylin and eosin according to Harris' method. Embryonic blood obtained from the umbilical vessels was diluted in phosphate buffered saline plus 1% fetal calf serum and cytospun. Cells were fixed with methanol, and stained with Giemsa.

Acknowledgments

We thank T. Lindahl and D. Barnes (ICRF, Clare Hall) for anti-DNA ligase I antibodies; P. Robins (ICRF, Clare Hall) for advice on protein purification; C. McEwan for preparation of the targeting vector; A. Pow and A.-M. Ketchen for culture of ES cells; T. White for maintenance of animal stocks; J. Lauder for histological sectioning of embryos; and A. Clarke and D. Harrison for useful discussions. D.J.B. is a BBSRC postgraduate student. This work was supported by the Cancer Research Campaign (Grant no. SP2095/0103).

Received 2 February; accepted 17 May 1996.

1. Barnes, D.E. *et al.* Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **87**, 6679–6683 (1990).
2. Tomkinson, A.E., Roberts, E., Daly, G., Totty, N.F. & Lindahl, T. Three distinct DNA ligases in mammalian cells. *J. Biol. Chem.* **266**, 21728–21735 (1991).
3. Wei, Y.-F. *et al.* Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. *Mol. Cell. Biol.* **15**, 3206–3216 (1995).
4. Wu, W. *et al.* A 17S multiprotein form of murine cell DNA polymerase mediates polyomavirus DNA replication in vitro. *J. Biol. Chem.* **54**, 32–46 (1994).
5. Applegren, N. *et al.* Further characterization of the human cell multiprotein DNA replication complex. *J. Biol. Chem.* **59**, 91–107 (1995).
6. Turchi, J.J., Huang, L., Murante, R.S., Kim, Y. & Bambara, R.A. Enzymatic completion of mammalian lagging-strand DNA replication. *Proc. Natl. Acad. Sci. USA* **91**, 9803–9807 (1994).
7. Waga, S., Bauer, G. & Stillman, B. Reconstitution of complete SV40 DNA replication with purified replication factors. *J. Biol. Chem.* **269**, 10923–10934 (1994).
8. Montecucco, A. *et al.* Late induction of human DNA ligase I after UV-C irradiation. *Nucl. Acids Res.* **23**, 962–966 (1995).
9. Aboussekhra, A. *et al.* Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80**, 859–868 (1995).
10. Webster, A.D.B., Barnes, D.E., Arlett, C.F., Lehmann, A.R. & Lindahl, T. Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene. *Lancet* **339**, 1508–1509 (1992).
11. Lönn, U., Lönn, S., Nylen, U. and Winblad, G. Altered formation of DNA replication intermediates in human 46BR fibroblast cells hypersensitive to 3-aminobenzamide. *Carcinogenesis* **10**, 981–985 (1989).
12. Prigent, C., Satoh, M.S., Daly, G., Barnes, D.E. & Lindahl, T. Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I. *Mol. Cell. Biol.* **14**, 310–317 (1994).
13. Teo, I.A., Arlett, C.F., Harcourt, S.A., Priestley, A. & Broughton, B.C. Multiple hypersensitivity to mutagens in a cell strain (46BR) derived from a patient with immuno-deficiencies. *Mut. Res.* **107**, 371–386 (1983).
14. Henderson, L.M., Arlett, C.F., Harcourt, S.A., Lehmann, A.R. & Broughton, B.C. Cells from an immunodeficient patient (46BR) with a defect in DNA ligation are hypomutable but hypersensitive to the induction of sister chromatid exchanges. *Proc. Natl. Acad. Sci. USA* **82**, 2044–2048 (1985).
15. Lehmann, A.R. *et al.* Relation between the human fibroblast strain 46BR and cell lines representative of Bloom's syndrome. *Cancer Res.* **48**, 6343–6347 (1988).
16. Petrini, J.H.J., Xiao, Y. & Weaver, D.T. DNA ligase I mediates essential functions in mammalian cells. *Mol. Cell. Biol.* **15**, 4303–4308 (1995).
17. Barnes, D.E., Tomkinson, A.E., Lehmann, A.R., Webster, A.D.B. & Lindahl, T. Mutations in the DNA ligase I gene in an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell* **69**, 495–504 (1992).
18. Stacey, A. *et al.* Use of double-replacement targeting to replace the murine α-lactalbumin gene with its human counterpart in embryonic stem cells and mice. *Mol. Cell. Biol.* **14**, 1009–1015 (1994).
19. Selfridge, J., Pow, A.M., McWhir, J., Magin, T. & Melton, D.W. Gene targeting using a mouse HPRT minigene/HPRT-deficient embryonic stem cell system: Inactivation of the mouse *ERCC-1* gene. *Som. Cell molec. Genet.* **18**, 325–336 (1992).
20. Metcalf, D. & Moore, M.A.S. *Haemopoietic Cells* (North-Holland, Amsterdam, 1971).
21. Montecucco, A. *et al.* DNA ligase I gene expression during differentiation and cell proliferation. *Nucl. Acids Res.* **20**, 6209–6214 (1992).
22. Chan, J.Y.-H. & Becker, F.F. DNA ligase activities during hepatocarcinogenesis induced by N-2-acetylaminofluorene. *Carcinogenesis* **6**, 1275–1277 (1985).
23. Gariboldi, M. *et al.* Genetic mapping and expression analysis of the murine DNA ligase I gene. *Mol. Carcinogenesis* **14**, 71–74 (1995).
24. McWhir, J., Selfridge, J., Harrison, D.J., Squires, S. & Melton, D.W. Mice with DNA repair gene (*ERCC-1*) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nature Genet.* **5**, 217–224 (1993).
25. Ellis, N.A. *et al.* The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* **83**, 655–666 (1995).
26. Jessop, J.K. & Melton, D.W. Comparison between cDNA clones encoding murine DNA ligase I. *Gene.* **160**, 307–308 (1995).
27. Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L. & Melton, D.W. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313–321 (1989).
28. Tomkinson, A.E., Lasko, D.D., Daly, G. & Lindahl, T. Mammalian DNA ligases. Catalytic domain and size of DNA ligase I. *J. Biol. Chem.* **265**, 12611–12617 (1990).
29. Lorimore, S.A., Pragnell, I.B., Eckman, L. & Wright, E.G. Synergistic interactions allow colony formation *in vitro* by murine haemopoietic stem cells. *Leuk. Res.* **14**, 481–489 (1990).
30. Freshney, M.G. in *Culture of Haematopoietic Cells* (eds Freshney, R.I., Pragnell, I.B. and Freshney, M.G.) 265–268 (Wiley-Liss, New York, 1994).
31. Noguiez, P., Barnes, D.E., Mohrenweiser, H.W. & Lindahl, T. Structure of the human DNA ligase I gene. *Nucl. Acids Res.* **20**, 3845–3850 (1992).

Transgenic Research 7, 387–396 (1998)

Efficient BLG-Cre mediated gene deletion in the mammary gland

STEFAN SELBERT[1], DARREN J. BENTLEY[2†], DAVID W. MELTON[2], DOMINIC RANNIE[1], PAULA LOURENÇO[1], CHRISTINE J. WATSON[3] and ALAN R. CLARKE[1*]

[1]*Department of Pathology, University Medical School, Teviot Place, Edinburgh, EH8 9AG, UK*
[2]*Institute of Cell and Molecular Biology, University of Edinburgh, Darwin Building, Kings Buildings, Mayfield Road, Edinburgh EH9 3JR, UK*
[3]*Sir Alastair Currie CRC Laboratories, Molecular Medicine Center, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK*

Received 9 April 1998; revised 18 June 1998; accepted 24 June 1998

Using the phage P1-derived Cre/*lox*P recombination system, we have developed a strategy for efficient mammary tissue specific inactivation of floxed genes. Transgenic mice were generated which express Cre DNA-recombinase under the control of the mammary gland specific promoter of the ovine beta-lactoglobulin (BLG) gene. To test the specificity of Cre mediated recombination, we crossed these mice to animals harbouring a floxed DNA ligase I allele. We show that the BLG-Cre construct specifies mammary specific gene deletion, and furthermore that it is temporally regulated, predominantly occurring during lactation. We fully characterised the extent of gene deletion in one line (line 74). In this strain the virgin gland is characterised by low levels (7%) of Cre mediated deletion, whereas 70–80% of cells within the lactating mammary gland have undergone recombination. Immunohistochemistry and indirect *in situ* PCR were used respectively to demonstrate that both Cre protein and Cre activity were evenly distributed throughout the population of secretory epithelial cells. The level of background recombination in non-mammary tissues was found to be ≤1.1%, irrespective of mammary gland developmental status. Crossing the transgenic BLG-Cre strain described here to mice harbouring other floxed alleles will facilitate the functional analysis of those genes during differentiation and development of the mammary gland.

Keywords: beta lactoglobulin promoter; Cre recombinase; DNA ligase I; transgenics; mouse mammary gland

Introduction

The Cre-*lox*P recombination system has recently been shown to allow regulated gene deletion within mammalian cells (Kilby *et al.*, 1993). The availability of such technology has greatly extended the scope of *in vivo* genetic manipulation, particularly where it has been used in conjunction with established gene targeting protocols. Using these approaches it is now possible to perform tissue-specific knock-outs (Gu *et al.*, 1994), inducible knock-outs (Kühn *et al.*, 1995), long targeted deletions (Li *et al.*, 1996), site-directed mutagenesis (Torres *et al.*, 1996), gene-insertions (Kitamoto *et al.*, 1996), site-directed chromosomal translocations (Smith *et al.*, 1995) as well as inversions and duplications (Ramírez-Solis *et al.*, 1995). The ability to generate temporally and/or spatially restricted gene alterations largely resolves two of the main problems associated with conventional gene 'knock-outs', namely embryonic lethality and secondary effects of the targeting event, such as developmental compensation (Kilby *et al.*, 1993).

The Cre-*lox*P system is based on the P1 bacteriophage Cre recombinase, which mediates site-specific recombination between 34 bp sequences referred to as *lox*P sites. When two *lox*P sites are placed in the same orientation on the same DNA molecule, Cre will mediate excision of

*To whom correspondence should be addressed at: University of Edinburgh, Medical School, Dept. of Pathology, Teviot Place, Edinburgh EH8 9AG (Tel: +44 0131 650 2909; Fax: +44 0131 650 6528; Email: aclarke@ed.ac.uk)

†Present address: Wellcome/CRC Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK.

0962–8819

the intervening DNA sequence. Conversely, inversion of the *lox*P flanked ('floxed') DNA occurs if the two *lox*P sites are in a head-to-head orientation (Chambers, 1994). Cre mediated recombination has been successfully employed in mammalian cells (Sauer *et al.*, 1988), mice (Lakso *et al.*, 1992), yeast (Sauer, 1987), bacteria (Maruyama *et al.*, 1992) and plants (Dale *et al.*, 1990). The system itself is not dependent on any cofactors but the outcome of conditional gene targeting is very heavily determined by the characteristics of the promotor used to drive Cre expression. Until now only a limited number of promotors have been used to successfully deliver tight expression of Cre. These include use of the *lck* or *CD19* promotors to generate T cell or B cell specific knock-outs (Gu *et al.*, 1994; Rickert *et al.*, 1997), the *CaMKII* promotor to activate Cre specifically in the forebrain (Tsien *et al.*, 1996) or the *aP2* promotor to deliver adipose tissue expression (Barlow *et al.*, 1997).

We have created transgenic mice which express Cre recombinase under the control of the ovine beta lactoglobulin promotor (BLG). In the past the BLG promotor has been successfully used to target transgenes efficiently and reliable to secretory epithelial cells of the mammary gland (Whitelaw *et al.*, 1992; Farini *et al.*, 1995; Archibald *et al.*, 1990; Clark *et al.*, 1992). To test the specificity of Cre mediated recombination we crossed BLG-Cre expressing mice to animals harbouring a floxed DNA ligase I (*Lig*I) allele. Quantitative PCR and quantitative Southern analysis allowed us to fully assess the extent and specificity of the achieved recombination.

Materials and methods

Cre expression vector and generation of transgenic mice

The BLG-Cre construct is shown in Figure 1 and was developed as follows. A 2.5 kb *Xho*I fragment from plasmid pSP-PGK-Cre containing the Cre gene with a modified translation initiation codon and the mouse metallothionein I polyadenylation signal (Sauer *et al.*, 1990), was blunt-ended and ligated to the *Eco*RV site of 4.2 kb-BLG/SK+ plasmid (Whitelaw *et al.*, 1992). The plasmid was linearized by *Not*I and was purified using a Qiagen purification kit prior to injection into pronuclei of fertilized eggs (CBA × C57Black6). Cre-transgenic mice were generated according to standard procedures and founder mice were detected by PCR using Cre primers (5′-CCGGATCCCTAATCGCCATCTTCCAGCAGGC-3′ and 5′-AAACGTTGATGCCGGTGAACG-3′). The resulting PCR product was 640 bp in size; its localization is indicated in Fig. 1.

To determine the extent of recombination at the DNA ligase 1 floxed allele, doubly transgenic animals were generated by mating female mice carrying one floxed allele of the DNA ligase I gene to male mice positive for Cre recombinase. Both the transgene and the floxed allele segregated according to Mendelian ratios, and thus one mouse out of 16 littermates was Cre positive, carried one floxed *Lig*I allele and was female. The phenotype of these mice was normal at the gross level, with apparently normal mammary gland development and function as assessed by the ability to successfully complete a breeding cycle.

Immunoblotting

Mammary gland tissue (day 10 lactation) was homogenized in TEDABP (10 mM Tris/HCl, pH 7.8, 1 mM EGTA, 1 mM DTT, 0.02% NaN_3, 1 mM benzamidine, 0.5 mM PMSF). Protein was determined according to the Bradford method using bovine serum albumin as a standard. SDS-PAGE was performed according to Laemmli *et al.* (1970), and Western blotting using the method of Towbin *et al.* (1979). Polyclonal anti Cre antibodies (Novagen, Madison, WI) were diluted 1:10,000 in TBS. For detection of antibody binding the Amersham ECL detection system was used.

RNA preparation and RT-PCR

Total RNA was isolated from a range of tissues from lactating Cre mice using TRIZOL (GibcoBRL). One microgram of RNA was reverse transcribed using SUPER-SCRIPT–RT from Gibco BRL and oligo$(dT)_{15}$-primers (Boehringer Mannheim). An aliquot of 2 µl of the RT-reaction was used in a 50 µl PCR reaction. 35 cycles were performed in the presence of three different pairs of primers. The pair of Cre primers mentioned above was used to amplify Cre mRNA and result in a PCR product of 640 bp. β-actin primers were used as a positive control; the product appearing at a molecular size of 440 bp (5′-CGTAGATGGGCACAGTG-3′ and 5′-CTCCGGCATG-TGCCAAAG-3′). As a control for DNA contamination BLG primers were used which amplify the BLG-promoter and give a PCR product of 250 bp (5′-CTTCTGGGGT-CTACCAGGA-3′ and 5′-TCGTGCTTCTGAGCTCTG-CA-3′). The localization of the BLG primers is indicated in Fig. 1.

Quantitative PCR

Oligonucleotide primers N2469 (5′-CCTGTGAGGGCCT-GATGGTGAAGACCTTGG-3′) and V2014 (5′-AGTGTT-CCCATGGCACAAGTGGCTGGAAGC-3′) amplify a 520 bp PCR product specific for the deleted *Lig*I allele. This was amplified as part of a duplex reaction with a pair of primers specific for the wild-type mouse *ERCC1* gene (McWhir *et al.*, 1993) acting as an independent control for the initial amount of DNA. PCR was performed on tissue DNA samples under standard conditions. Cycle parameters: denaturation, 94 °C for 1 min; annealing, 70 °C for 1 min; elongation, 72 °C for 1 min; 20 cycles. PCR products were separated by agarose gel electrophoresis,

transferred onto nylon membrane, and hybridized to [α-^{32}P]-labeled probes specific for the two PCR products. Transfer and hybridization conditions are described elsewhere (Thompson *et al.*, 1989). Levels of radiolabel bound to each PCR product were analyzed by phosphorimager, and the ratio of the two signals calculated. In order to estimate the percentage of the deleted *Lig* I allele present in each tissue sample, calculated ratios were compared against a standard curve (see Fig. 4D). Primers V2014 and V0727 (5′-GTCAGATTCAGAACCAACA-AAG-3′) were used in a control reaction on the wild type *Lig* I sequence to give a 310 bp product.

Southern blotting

Genomic DNA was prepared and subjected to Southern analysis as described (Melton *et al.*, 1997). The source of the mouse DNA ligase I cDNA fragment used as a probe is described in Bentley *et al.* (1996).

Immunohistochemistry

Formalin fixed sections were deparaffinized in xylene, rehydrated in descending grades of alcohol to water and trypsinised for 10 min in 0.01% trypsin/0.1% $CaCl_2$. After washing the sections in Tris buffered saline (TBS) non-specific binding sites were blocked for 20 min with diluted rabbit serum (1:5 with TBS) and for 15 min with an avidin/biotin blocking Kit from Vector Lab. (no. SP-2001). Cre was detected by incubating the sections overnight in a 1:1000 dilution of polyclonal rabbit anti-Cre antibody (Novagen, Madison, WI). Sections were washed twice in TBS and incubated for 30 min with a biotinylated mouse anti-rabbit secondary antibody (Dako), washed twice in TBS and incubated with freshly prepared avidin-biotin alkaline phosphatase complex (Dako). The color reaction was allowed to progress for 20 min in complete darkness (Vector red, Alkaline Phosphatase Substrate Kit I). The sections were counterstained in haematoxylin and mounted in crystal mount.

Indirect in situ PCR

The *Lig* I probe (520 bp) detailed above was labeled using DIG-11-dUTP in a standard PCR reaction (primers N2014 & V2469). The protocol comprised 35 cycles: 94 °C for 1 min, 70 °C for 1 min and 72 °C for 1 min. Before using the labeled probe it was gel-purified using QiatexII extraction kit (Qiagen). The formalin fixed sections were deparaffinized in xylene and rehydrated in descending grades of alcohol to water. The following procedures were performed in a Hybaid PCR thermocycler. Tissue sections were permeabilized by incubation for 15 min at 37 °C in 100 µl of 10 µg/ml Proteinase K (Boehringer Mannheim). The sections were washed four times in PBS and the *in situ* PCR was performed using the standard Cre PCR (see above) (GibcoBRL). Each section was incubated with 100 µl of the following reaction mix: 7 µl 10 × buffer, 2.1 µl 50 mM $MgCl_2$, 1.4 µl 10 mM dNTPs, 0.7 µl primers (N2469 & V2014), 4.9 µl Taq/TaqStart antibodies (Clontech) and 53.9 µl water. The negative control did not include any primers. The *in situ* PCR was performed using 25 cycles with the conditions detailed above. The sections were washed twice 5 min each in 2 × SSC at room temperature and twice in 2 × SSC at 37 °C. Slides were fixed for 15 min in 0.4% paraformaldehyde/PBS on ice. The slides were then washed for 5 min in PBS at room temperature, rinsed in water and incubated overnight in 10 µl hybridization buffer: 20 µl 50 × Denhardt's, 100 µl dextran sulfate, 20 µl denatured salmon sperm DNA, 200 µl 20 × SSC, 100 µl TE, 1 µl 50 ng/ml DIG labeled DNA ligase I probe and 59 µl water. The hybridizing sections were incubated 6 min at 95 °C, 1 min on ice and then left at 42 °C overnight. Sections were washed three times in 2 × SSC at 20 °C and once in 0.1% SSC at 42 °C. For detection the slides were dipped briefly in buffer 1 (100 mM Tris (pH 7.5), 150 mM NaCl, 2 mM $MgCl_2$), incubated in anti-DIG alkaline phosphatase conjugated antibody (Boehringer Mannheim), 1:500 diluted in buffer 2 (buffer 1, 3% BSA) for 1 hour at room temperature. Non-bound antibodies were removed by two washes in buffer 1. For the color reaction the sections were equilibrated 5 min in buffer 3 (0.1 M Tris (pH 9.5), then 20 µl NBT/BCIP color solution (Boehringer Mannheim) were added. The reaction took place for a few hours or overnight in a humid and dark chamber. The sections were rinsed in water and mounted in water soluble Crystal Mount (Biomeda, Foster City, CA, USA).

Results

Generation and characterization of BLG-Cre transgenic animals

We have generated several transgenic lines of mice which carry the bacteriophage P1 Cre-DNA recombinase under the control of the sheep BLG promoter (BLG-Cre, Fig. 1). Transgenic mice were identified by PCR and Southern blot. Of 15 lines that scored positive for Cre DNA by Southern analysis, 5 showed Cre-protein expression in Western blot using polyclonal rabbit anti-Cre-antibodies (Novagen, Madison, WI). The copy number of the Cre transgene varied between 1 and 10 (data not shown) and in those lines with a single integration site (lines 27, 74 and 93) approximately parallels the amount of detectable Cre protein as assessed by Western analysis (Fig. 2). Line 74 showed the highest protein expression and an estimated copy number of 8, line 93 had an estimated copy number of 4 and line 27 had 3 copies. Lines 7 and 16 had multiple integration sites of the transgene. Tissue-specific expression of the Cre-transgene was analyzed by performing RT-PCR with RNA extracted from a variety of tissues (Fig. 3). This analysis revealed mammary gland expression of

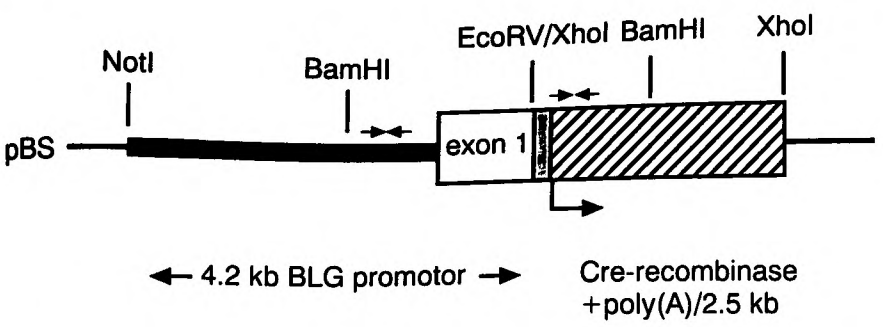


Fig. 1. Structure of the Cre-transgene. Cre recombinase (striated box & spotted box) was cloned as an *Xho* I fragment of pSP-PGK-Cre into the *Eco* RV site of 4.2 kb-BLG/SK+ plasmid. The start of translation is indicated by an arrow. The *Eco* RV site is located in the first untranslated exon of the BLG gene (thick bar). The Cre recombinase fragment containing the polyadenylation signal contains 52 non-coding basepairs upstream of the ATG (spotted box). The small arrows indicate the location of two pairs of primers used to amplify BLG and Cre sequences by PCR.

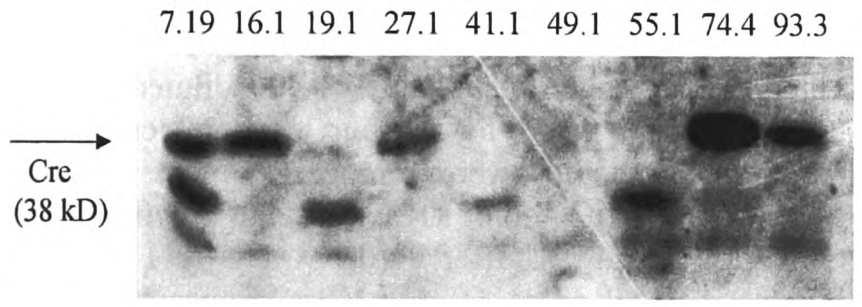


Fig. 2. Western blot analysis of Cre expressing mice. 10 μg of mammary gland protein (day 10 lactation) was separated by SDS-PAGE, transferred onto nylon membrane and incubated with polyclonal anti Cre antibodies (Novagen, Madison, WI). Five of the lines showed Cre expression (7.19, 16.1, 27.1, 74.4 and 93.3) which is indicated by a positive immunoreaction at 38 kDa. Some of the probes also detected a smaller protein which may result from Cre degradation products.

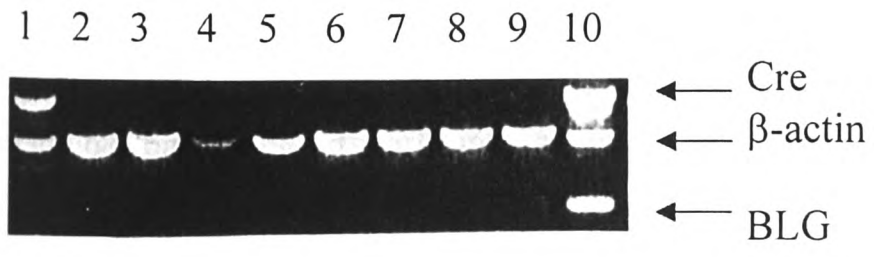


Fig. 3. RT-PCR analysis with RNA extracted from a variety of tissues from Cre transgenic mouse line 74.4. Primers within the Cre gene were used to amplify Cre mRNA which result in a PCR product of 640 bp. Cre mRNA can only be detected in mammary tissue (lane 1) and the non DNAse treated tail DNA (lane 10), which served as a positive control. Salivary gland (lane 2), liver (lane 3), heart (lane 4), kidney (lane 5), spleen (lane 6), small intestine (lane 7), pancreas (lane 8) and skeletal muscle (lane 9) were negative for Cre mRNA. PCR using β-actin primers has been used as a positive control and appears at a molecular size of 440 bp. As a control for DNA contamination BLG primers were used which amplify the BLG-promoter. Only the positive control shows the 250 bp product.

Cre in all the tested lines, and further that no ectopic expression was detected in the majority of these lines. Of the subset (3 lines) that did show expression outside the mammary gland, this was restricted to the salivary gland.

Cre activity is restricted to mammary tissue

To assess Cre activity, Cre expressing mice were mated to mice carrying a floxed target allele. In the study presented here we used the BLG-Cre strains showing the highest Cre-expression as assessed by Western blot and no ectopic Cre-mRNA as confirmed by RT-PCR (lines 74 and 93). As a target gene we used the DNA ligase I gene (*Lig* I) (Bentley *et al.*, 1996), modified by homologous recombination so that it contains two *lox*P sites. Mice null for *Lig* I die during embryogenesis, as a result of disruption of foetal liver erythropoiesis (Bentley *et al.*, 1996). A 'double replacement' strategy was used to introduce two *lox*P sites into the endogenous locus (Fig. 4A). Female progeny positive for Cre (Cre+) with one floxed *Lig* I allele (flox/+) were selected for further study. Double transgenic females (Cre+/flox/+) were mated to wild type mice and the females were sacrificed at various time points during mammary gland development. DNA was extracted from a range of tissues and analyzed. A PCR-based approach was used to detect the Cre-mediated recombination event, and to quantify the level of recombination in each tissue (Fig. 4B–D). Cre mediated deletion of the floxed DNA ligase I allele was scored by amplification of a 520 bp PCR product using primers N2469 and V2014 (see Fig. 4A). This combination of primers gives a product only after recombination between the two *lox*P sites has excised the intervening genomic region. The distance between the two sites in the non-recombined allele is too great to allow amplification. As shown in Fig. 4B the 520 bp fragment specific for the deleted allele is amplified only in mammary tissue (10 day lactation) and to a low extent in tail. A control reaction to assess the quality of the genomic DNA amplifies the non-recombined wildtype allele and gave a positive result for all samples (Fig. 4C).

Quantification of BLG-Cre mediated recombination

Performing quantitative PCR with the primers described above we estimate the amount of Cre mediated recombination in lactating mammary gland (line 74) to be approximately 70% whereas in all other tissues it was found to be ≤1.1% (Table 1). The highest level of ectopic Cre mediated recombination was found in tail. In line 93, we observed 39% recombination in the lactating mammary gland and some background recombination in salivary gland (2%), skeletal muscle and skin (1.5%). A typical calibration curve, used to determine the level of Cre-mediated recombination, is shown in Fig. 4D. The results obtained from this analysis showed relatively large standard deviations, particularly at high levels of deletion. We

therefore chose to use Southern analysis to independently measure the extent of Cre-mediated recombination (Fig. 4E). The 3′ end of the floxed *Lig* I allele is located on an 8.5 kb *Eco* RI restriction fragment (see Fig. 4A). After Cre-mediated recombination this fragment is lost and replaced by a novel 3.5 kb fragment. This is illustrated for control ES cell lines in Fig. 4E before (lane 5) and after recombination (lane 6). For recombination ES cells have transiently been transfected with the Cre expression vector pSP-PGK-Cre. A mammary sample from a 6 week old virgin (lane 2) showed no evidence of the 3.5 kb fragment, while a low level of this fragment was present in a sample from a 16 week old virgin (lane 1). Mammary samples from two lactating females (lane 3, 14 days lactation; lane 4, 10 days lactation) showed a marked reduction in the intensity of the 8.5 kb fragment and a high level of the 3.5 kb fragment. By using a phosphorimager to compare the intensity of the signal from the 3.5 kb fragment relative to a control (5.2 kb) fragment common to all *Lig* I alleles, the level of Cre-mediated recombination in the mammary samples could be compared to the 100% level in the recombined ES cell clone. This analysis gave the following levels of recombination: virgin 16 weeks, (*Lig* I+/flox) (lane 1), 15%; 14 days lactation, (*Lig* I+/flox) (lane 3), 73%; 10 days lactation, (*Lig* I−/flox) (lane 4), 80%. The extent of reduction of the 8.5 kb band is not readily obvious in Fig. 4E as this photograph has purposely been over-exposed to demonstrate the increase in the more weakly hybridizing 3.5 kb fragment. Phosphorimager analysis of the 8.5 kb band however confirms the reduction in intensity of this band. The additional 6.7 kb band in lane 4 results from the introduced *Eco* RI site present in the HPRT selectable marker of the targeted allele (see Fig. 4). Although these values are only based on a single blot, they are in good agreement with those obtained by quantitative PCR (see Table 1) and provide an important independent confirmation of the high level of Cre-mediated recombination in the lactating mammary gland, with a considerably lower level in the glands from virgin animals.

BLG-Cre mediated recombination during mammary gland development

We next used quantitative PCR analysis to establish the time course of recombination throughout mammary gland development. These studies were performed using line 74 (Table 1). We found that in the virgin gland 7% of cells had recombined their floxed allele. At day 15 of gestation and at the time point of birth the level of recombination had risen to between 20 and 30%. The highest percentage was found at day 10 of lactation, where 69% of the floxed allele had been deleted.

Localization of Cre-protein and Cre-activity

To assess the localization of Cre within mammary tissue we performed immunohistochemistry using polyclonal anti Cre antibodies (Novagen). Figure 5A shows that Cre protein was expressed in the vast majority of epithelial cells of the transgenic BLG-Cre strain (line 74) whilst corresponding cells of wild type mice were negative for Cre protein (Fig. 5B). Because some variability was observed in the staining pattern using the anti-Cre antibody, we also analyzed Cre activity by indirect *in situ* PCR using the primer combination N2469/V2014. This identified cells which had undergone Cre mediated recombination. Figure 5C shows that Cre positive nuclei were distributed throughout Cre positive mammary glands, but are absent from Cre negative control mice (Fig. 5D).

Discussion

The Cre-*lox* based recombination system has the potential to become a powerful genetic tool, allowing the precise introduction of genetic alterations within a given time period and within a specific tissue of the animal. It therefore avoids many of the problems regularly arising from conventional knockout studies, in particular those associated with embryonic lethality. In this article we describe transgenic mice which utilize the ovine β-lactoglobulin (BLG) promoter in combination with Cre recombinase. We show that this construct can be used to efficiently deliver gene deletion within the lactating mammary gland.

BLG is the major whey protein in ruminants. It is not present in rodent milk (Ali *et al.*, 1988). Various sizes of the BLG promoter, the smallest of which being 406 bp, have been shown to direct expression of transgenes exclusively to secretory epithelial cells of mammary tissue (Whitelaw *et al.*, 1992), where expression peaks at about day 10 of lactation (Harris *et al.*, 1991).

We utilised the BLG promoter sequence to generate transgenic mice carrying the bacteriophage P1 Cre DNA recombinase driven by the sheep BLG promoter. As expected, Cre mRNA was restricted to the mammary gland (day 10 lactation) in the majority of the transgene positive lines. This finding is in agreement with published data suggesting that BLG expression is largely position independent (Whitelaw *et al.*, 1992). Three lines showed ectopic expression in the salivary gland. Ectopic expression in the salivary gland has been previously reported in transgenic lines using the BLG promoter (Farini *et al.*, 1995), and it has been speculated that this arises because these two tissues are developmentally related.

Cre activity was evaluated by crossing BLG-Cre positive lines (74 and 93) to mice carrying a floxed DNA ligase I gene. Quantitative PCR and Southern blot analysis show that in one of the lines (line 74) Cre mediated recombination of the floxed *Lig* I allele occurs in 70–80% of lactating mammary gland with all other tissues exhibiting recombination levels ≤1.1%. Both

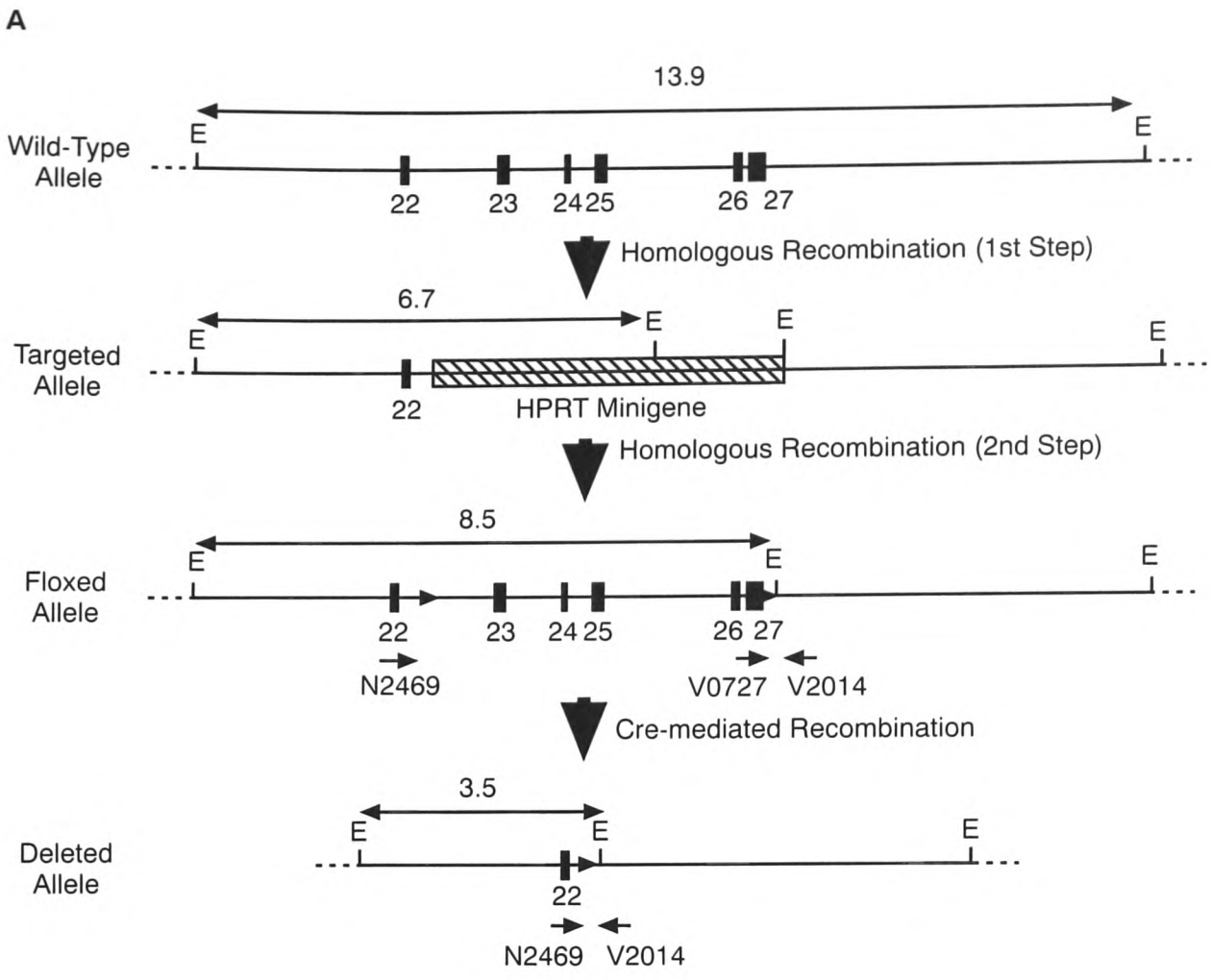
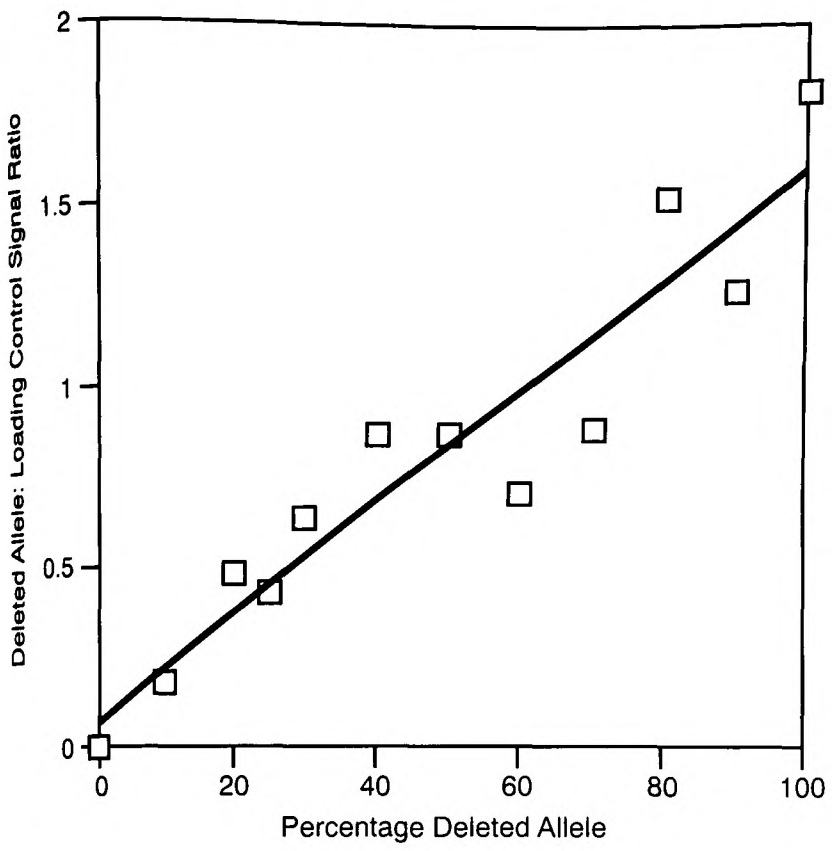


Fig. 4. (A) Schematic representation of the strategies used to produce the floxed DNA ligase I allele and detect Cre-mediated recombination. The last 5 exons of the *Lig* I locus were replaced with an *HPRT* minigene by homologous recombination in HPRT-deficient ES cells (Bentley *et al.*, 1996; McWhir *et al.*, 1993). A second homologous recombination step was subsequently used to reconstruct a functional allele, but with exons 23 to 27 flanked by two *lox*P sites (unpublished data). Cre-mediated recombination between the two *lox*P sites was detected by a PCR reaction specific for the deleted allele (Fig. 4B). Primer N2469 binds within exon 22, while primer V2014 binds downstream of exon 27 and the adjacent *lox*P site. The wild-type and non-recombined floxed alleles do not give a product because the distances between the two primer binding sites are too great for efficient amplification. Only when Cre-mediated recombination has juxtaposed the primer binding sites can a 520 bp product be amplified. Solid boxes, exons (numbered according to convention); lines, *Lig* I introns or flanking regions; hatched box, *HPRT* minigene; solid triangles, *lox*P sites. Arrows indicate the positions at which PCR primers bind. *Eco* RI restriction sites are marked by E and the sizes (in kb) of the *Eco* RI restriction fragments diagnostic for the different *Lig* I allele structures are shown. (B) PCR-Assay for Cre-mediated DNA recombination in various tissues. DNA from thymus (lane 1), ovary (lane 2), mammary gland, day 10 lact. (lane 3), heart (lane 4), tail (lane 5), lung (lane 6), liver (lane 7), pancreas (lane 8), skin (lane 9), spleen (lane 10), small intestine (lane 11), large intestine (lane 12), kidney (lane 13), brain (lane 14) and salivary gland (lane 15) was used as a substrate for the target gene PCR assay. The primer combination V2014 and N2469 will only generate a PCR product (520 bp) if the Cre mediated recombination event brings the two binding sites into close proximity. (C) Primers V0727 and V2014 were used to amplify the wildtype *Lig* I sequence producing a 310 bp product. The tissue samples are loaded in the same order as in Fig. 4B.



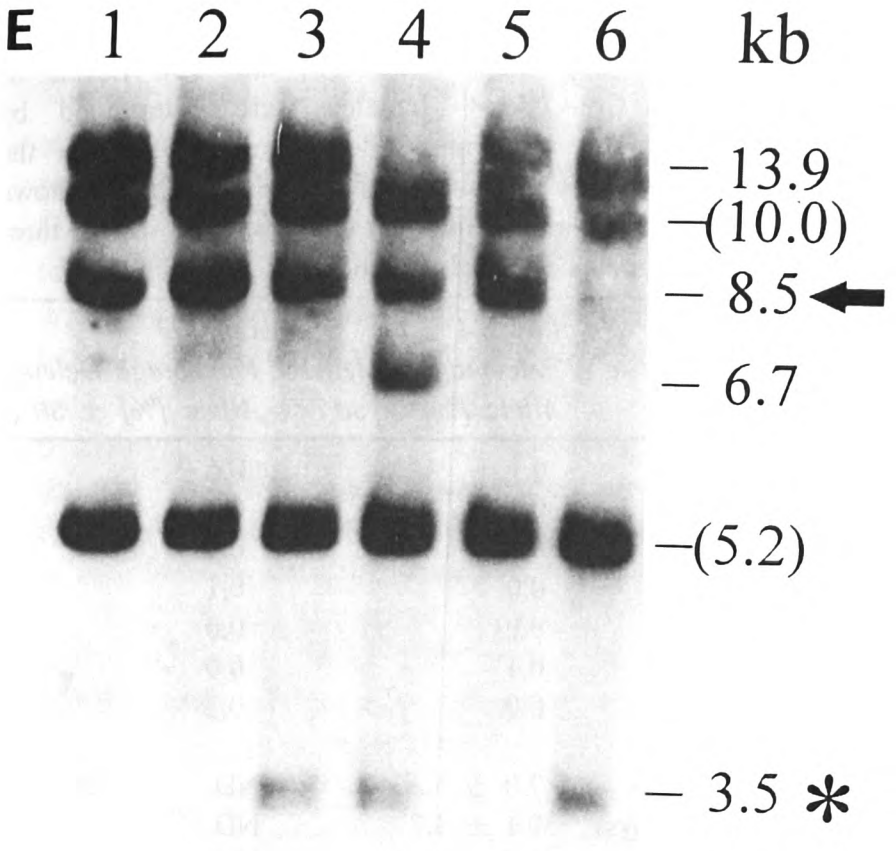


Fig. 4. (D) A typical standard curve for PCR quantitation of the levels of Cre-mediated recombination of the floxed *Lig* I allele. Varying amounts of genomic DNA from an ES cell clone heterozygous for the deleted *Lig* I allele were spiked into wild-type DNA, thereby producing a series of samples containing known percentages of the deleted allele. These standards were then used as templates for PCR reactions specific for the deleted allele, as part of a duplex with an independent reaction acting as a control for the amount of input DNA. Products were separated by electrophoresis, blotted and hybridized to probes specific for the two PCR products. The relative levels of the two signals were quantified by phosphorimagery and the calculated ratios plotted against the known percentages of the deleted allele. Using this graph it was possible to derive estimates of the levels of Cre-mediated recombination in mammary DNA samples. (E) Southern blot analysis of the *Lig* I locus before and after Cre-mediated recombination. Genomic DNA (10 μg), prepared from mouse mammary tissue and mouse ES cells, was digested with *Eco* RI and subjected to Southern analysis using a 2.1 kb *Bam* HI/*Eco* RI fragment of mouse DNA ligase I cDNA as a probe. The cDNA probe used hybridizes to exons 8–27 of the *Lig* I gene. Lanes 1–4: mammary tissue from line 74 Cre transgenic mice heterozygous for the floxed *Lig* I allele (mice in lanes 1–3 are heterozygous floxed/wild-type; the mouse in lane 4 is heterozygous floxed/targeted *Lig* I allele). Lane 1, virgin 16 weeks old; lane 2, virgin 6 weeks old; lane 3, mother 14 days lactation, lane 4, mother 10 days lactation. Lane 5, ES cell clone heterozygous for the floxed *Lig* I allele (heterozygous floxed/wild-type). Lane 6, ES cell clone heterozygous for the deleted *Lig* I allele (heterozygous deleted/wild-type). The sizes (in kb) of all the *Lig* I gene fragments hybridizing to the probe are indicated. Sizes in parentheses indicate fragments derived from the 5′ and central region of the *Lig* I gene which are common to all four *Lig* I alleles and are not shown in panel a). The asterisk indicates the novel 3.5 kb fragment generated by Cre-mediated recombination. The arrow indicates the 8.5 kb fragment of the floxed allele which is lost after Cre-mediated recombination. The 3.5 kb fragment only contains exon 22 and so hybridizes more weakly to the cDNA probe than the 8.5 kb fragment (containing exons 22–27) which it replaces.

PCR and Southern blot analysis gave results which were in good agreement with each other, although the quantitative PCR proved to be quite sensitive to DNA quality, resulting in high standard deviations. The pattern of Cre activity closely reflects the published kinetics of BLG expression (Harris *et al.*, 1991). Very low recombination levels were observed in the virgin mammary gland (7%), increasing throughout gestation and birth (20–30%) and highest at about day 10 of lactation (70–80%). The amount of tissue which still contains the non-recombined floxed allele (~20%) very likely reflects the population of non-epithelial stromal cells (fat cells, fibroblasts and lymphocytes) which do not upregulate the BLG promoter and therefore lack Cre expression. However we cannot formally exclude the possibility that a small percentage of epithelial cells failed to undergo recombination. Such restriction of Cre expression to mammary epithelial cells is supported by the finding that anti-Cre antibodies only stained epithelial cells. Cre expression could be detected in the majority of epithelial cells, although the level of staining varied, which might reflect either staining artifacts or different expression levels of Cre within different cells. The occasional staining inside the lumen

Table 1. Tissue-distribution of Cre-mediated recombination in two different Cre expressing mouse lines (74 and 93). Levels of recombination in Cre+/flox+ females were determined by quantitative PCR. The 'percentage deleted allele' reflects the proportion of cells which contain the deleted allele. Figures shown for the mammary gland samples represent at least three independent experiments for each timepoint

Tissue	*Line 74 Percentage Deleted Allele [%] ± SD*	*Line 93 Percentage Deleted Allele [%] ± SD*
Brain	0.1	0.6
Heart	0.0	0.1
Intestine (Large)	0.1	0.1
Intestine (Small)	0.0	0.1
Kidney	0.0	0.0
Liver	0.1	0.0
Lung	0.0	0.2
Mammary Gland		
Virgin	7.0 ± 5.2	ND
Gestation (15 days)	30.1 ± 4.7	ND
Birth	20.7 ± 14.5	ND
Lactating (10 days)	69.1 ± 32.3	38.6 ± 28.6
Ovary	0.0	0.0
Pancreas	0.1	0.2
Salivary Gland	0.1	2.4
Skeletal Muscle	0.1	1.4
Skin	0.1	1.5
Spleen	ND	0.1
Tail	1.1	0.5
Thymus	0.0	0.1

of the alveoli probably results from the use of avidin to amplify the Cre-signal and represents cross reaction to milk protein; it is also apparent in Cre negative sections. To independently confirm Cre activity in the mammary gland, we used indirect *in situ* PCR to identify cells which had undergone recombination. This showed that Cre activity, as assessed by the presence of the recombined allele, was detectable in virtually every epithelial cell. Occasional negative nuclei were observed, which we assume arise either because of failed recombination or because of technical failure of the PCR reaction (e.g. because the recombined allele is not present in that section of the nucleus). We have recently generated mice which are Cre positive, bear the floxed *Lig* I allele and also the null *Lig* I allele and we have initiated a study of mammary function in these mice. Preliminary data from these animals suggest that DNA ligase I deficiency retards mammary gland development.

Taken together, the data show that BLG driven Cre expression results in Cre mediated excision, which is tightly controlled both spatially and temporally. This is somewhat in contrast to a recently published study reporting the phenotypes of mice transgenic for Cre recombinase that are driven by the mammary gland promotors MMTV or WAP (Wagner *et al.*, 1997). In this study, the MMTV-Cre construct did not show tissue specificity, and furthermore Cre was found to be expressed in the germline. The use of the WAP promoter did result in Cre expression in the mammary gland, however expression was also observed within the brain which is not without precedent as Günzburg *et al.* (1991) have demonstrated high expression levels of WAP driven transgenes specifically in Bergman glia cells. Although WAP-Cre mice are clearly useful for the analysis of gene inactivation within the mammary gland, it remains possible that any phenotypic analysis will be complicated by the observed ectopic expression. Furthermore, no definitive data is given relating to the level of Cre activity in the WAP-Cre mice, and it therefore remains possible that the level and pattern of recombination is insufficient to result in detectable phenotypes in all 'floxed' strains of mice. This is a particularly relevant point for mammary specific transgenesis, as use of MMTV, WAP and even the BLG promotor has previously been shown to lead to mosaic expression within the mammary gland (Rollini *et al.*, 1992; Robinson *et al.*, 1995; Dobie *et al.*, 1996).

Based on the results presented here, we believe that this murine strain (line 74) will prove a valuable addition to the currently available tools for studying the genetic control of mammary biology, particularly for those genes where constitutive inactivation results in embryonic lethality.

Acknowledgements

The authors would like to thank John Clark, Roberta Wallace, John Verth and staff for their assistance in the generation and maintenance of the Cre-transgenics. We also thank Ann-Marie Ketchen for culture of ES cells, Carolanne McEwan for mouse genotyping, Niki Redhead and Jim Selfridge for carrying out the Southern blot shown in Fig. 4E and Robert Belfour for his help in performing the RT-PCR analysis. This work was supported by a DFG grant to SS (Se883/1-1) and by a BBSRC grant to ARC and CJW. ARC is a Royal Society Research Fellow. DJB is a BBSRC postgraduate student, and the DNA ligase I gene targeting work was supported by the Cancer Research Campaign (Grant no. SP2095/0103).

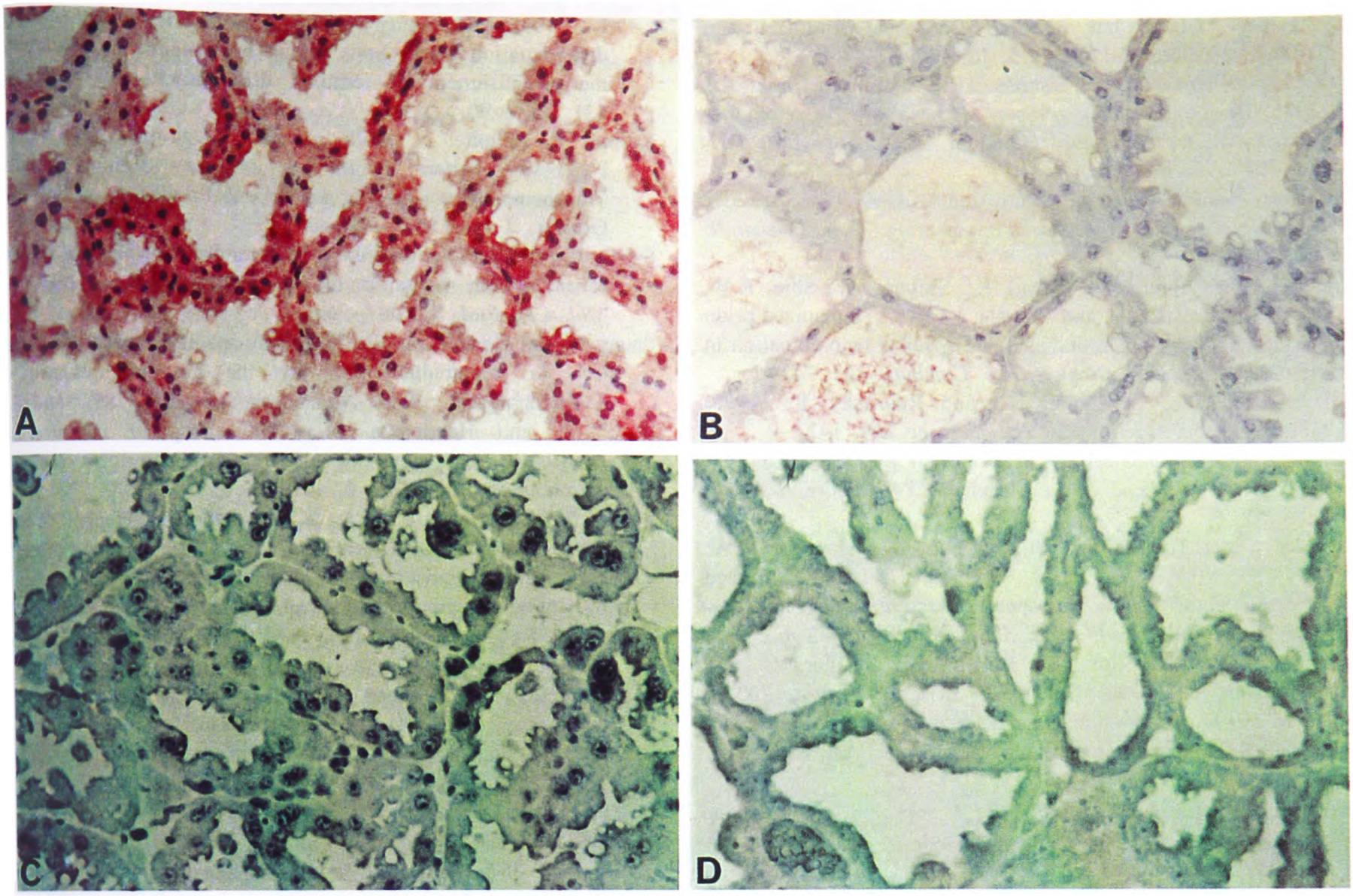


Fig. 5. Localization of Cre protein and Cre activity. Fig. 5A and C show formalin fixed mammary gland sections (day 10 lact.) of Cre expressing line 74, in B and D wild type mammary gland sections are shown. In A and B sections were incubated with 1:1000 dilution of polyclonal anti-Cre. In C and D indirect *in situ* PCR was used to amplify recombined DNA ligase I gene fragments. Sections were incubated with primer combination V2014 and N2469 to amplify the recombined *Lig* I allele and then probed with a DIG-labeled probe. After incubating the sections with alkaline phosphatase labeled anti-DIG antibodies, Cre mice show a homogeneous distribution of Cre (A), Cre activity respectively (C) in mammary tissue, while wildtype mice are Cre negative (B, D). 400 × magnification.

References

Ali, S. and Clark, J. (1988) Characterisation of the gene encoding ovine β-lactoglobulin: similarity to the genes for retinol binding protein and other secretory proteins. *J. Mol. Biol.* **199**, 415–26.

Archibald, A.L., McClenaghan, M., Hornsey, V., Simons, J.P. and Clark, A.J. (1990) High-level expression of biologically active human α1-antitrypsin in the milk of transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5178–82.

Barlow, C., Schroeder, M., Lekstrom-Himes, J., Kylefjord, H., Deng, C.-X., Wynshaw-Boris, A., Spiegelman, B.M. and Xanthopoulos, K.G. (1997) Targeted expression of Cre recombinase to adipose tissue of transgenic mice directs adipose-specific excision of *lox*P-flanked gene segments. *Nucl. Acid. Res.* **25**, 2543–5.

Bentley, D.J., Selfridge, J., Millar, J.K., Samuel, K., Hole, N., Ansell, J.D. and Melton, D.W. (1996) DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability. *Nature Gen.* **13**, 489–91.

Chambers, C.A. (1994) TKO'ed Lox, stock and barrel. *BioEssays* **16**, 865–68.

Clark, A.J., Cowper, A., Wallace, R., Wright, G. and Simons, J.P. (1992) Rescuing transgene expression by co-integration. *Bio/Technology* **10**, 1450–4.

Dale, E.C. and Ow, D.W. (1990) Intra- and intermolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. *Gene* **91**, 79–85.

Dobie, K.W., Lee, M., Fantes, J.A., Graham, E., Clark, A.J., Springbett, A., Lathe, R. and McClenaghan, M. (1996) Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6659–64.

Farini, E. and Whitelaw, C.B.A. (1995) Ectopic expression of β-lactoglobulin transgenes. *Mol. Gen. Genet.* **246**, 734–8.

Gu, H., Marth, J.D., Orban, P.C., Mossmann, H. and Rajewsky, K (1994) Deletion of a DNA polymerase β gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103–6.

Günzburg, W.H., Salmons, B., Zimmermann, B., Müller, M., Erfle, V. and Brem, G. (1991) A Mammary-Specific Promotor Directs Expression of Growth Hormone not only to the

Mammary Gland, but also to Bergman Glia Cells in Transgenic Mice. *Mol. Endo.* **5**, 123–33.

Harris, S., McClenaghan, M., Simons, J.P., Ali, S. and Clark, A.J. (1991) Developmental Regulation of the Sheep β-Lactoglobulin Gene in the Mammary Gland of Transgenic Mice. *Dev. Gen.* **12**, 299–307.

Kilby, N.J., Snaith, M.R. and Murray, J.A. (1993) Site-specific recombinases: tools for genome engineering. *Trends Genet.* **9**, 413–21.

Kitamoto, T., Nakamura, K., Nakao, K., Shibuya, S., Shin, R.W., Gondo, Y., Katsuki, M. and Tateishi, J. (1996) Humanized prion protein knock-in by Cre-induced site-specific recombination in the mouse. *Biochem. Biophys. Res. Commun.* **222**, 742–7.

Kühn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995) Inducible gene targeting in mice. *Science* **269**, 1427–9.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 722–5.

Lakso, M., Sauer, B., Mosinger Jr., B., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L. and Westphal, H. (1992) Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6232–6.

Li, Z.-W., Stark, G., Gotz, J., Rulicke, T., Muller, U. and Weissmann, C. (1996) Generation of mice with a 200-kb amyloid precursor protein gene deletion by Cre recombinase-mediated site-specific recombination in embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6158–62.

Maruyama, I.N. and Brenner, S. (1992) A selective lambda phage cloning vector with automatic excision of the insert in a plasmid. *Gene* **120**, 135–41.

McWhir, J., Selfridge, J., Harrison, D.J., Squires, S and Melton, D.W. (1993) Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nature Gen.* **5**, 217–24.

Melton, D.W., Ketchen, A.-M., and Selfridge, J. (1997) Stability of *HPRT* marker gene expression at different gene-targeted loci: observing and overcoming a position effect. *Nucl. Acid. Res.* **25**, 3937–43.

Ramírez-Solis, R., Liu, P. and Bradley, A. (1995) Chromosome engineering in mice. *Nature (London)* **378**, 720–4.

Rickert, R.C., Roes, J. and Rajewsky, K. (1997) B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucl. Acid. Res.* **26**, 1317–8.

Robinson, G.W., McKnight, R.A., Smith, G.H. and Hennighausen L. (1995) Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. *Development* **121**, 2079–90.

Rollini, P., Billotte, J., Kolb, E. and Diggelmann, H. (1992) Expression Pattern of Mouse Mammary Tumor Virus in Transgenic Mice Carrying Exogenous Proviruses of Different Origins. *J. Virol.* **66**, 4580–6.

Sauer, B. (1987) Functional expression of the Cre-*lox* site-specific recombination system in the yeast Saccharomyces cerevisae. *Mol. Cell Biol.* **7**, 2087–96.

Sauer, B. and Henderson, N. (1988) Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5166–70.

Sauer, B. and Henderson, N. (1990) Targeted insertion of exogenous DNA into the eucaryotic genome by the Cre recombinase. *New Biol.* **2**, 441–9.

Smith, A.J.H., De Sousa, M.A., Kwabi-Addo, B., Heppell-Parton, A., Impey, H. and Rabbitts, P. (1995) A site-directed chromosomal translocation induced in embryonic stem cells by Cre-*lox*P recombination. *Nature Genet.* **9**, 376–85.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L. and Melton, D.W. (1989) Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313–21.

Torres, R.M., Flaswinkel, H., Reth, M. and Rajewsky, K. (1996) Aberrant B cell development and immune response in mice with a compromised BCR complex. *Science* **272**, 1804–8.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350–4.

Tsien, J.Z., Chen, D.F., Gerber, D., Tom, C., Mercer, E.H., Anderson, D.J., Mayford, M., Kandel, E.R. and Tonegawa, S. (1996) Subregion- and Cell Type-Restricted Gene Knockout in Mouse Brain. *Cell* **87**, 1317–26.

Wagner, K.-U., Wall, R.J., St-Onge, L., Gruss, P., Wynshaw-Boris, A., Garrett, L., Minglin, L., Furth, P.A. and Henninghausen, L. (1997) Cre-mediated gene deletion in the mammary gland. *Nucl. Acid. Res.* **25**, 4323–30.

Whitelaw, C.B.A., Harris, S., McClenaghan, M., Simons, J.P. and Clark, A.J. (1992) Position-independent expression of the ovine β-lactoglobulin gene in transgenic mice. *Biochem. J.* **286**, 31–9.