



University  
of Glasgow

Miranda, Alexandra de Sousa Montenegro (2006) Characterisation of LVI-1 (WDR76) as a candidate tumour suppressor gene. PhD thesis.

<http://theses.gla.ac.uk/4967/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

**CHARACTERISATION OF *LVI-1* (*WDR76*) AS A CANDIDATE TUMOUR  
SUPPRESSOR GENE**

**Alexandra de Sousa Montenegro Miranda**

**A thesis submitted for the degree of Doctor of Philosophy**



**Molecular Oncology Laboratory  
Institute of Comparative Medicine  
Faculty of Veterinary Medicine  
University of Glasgow  
April, 2006**

# List of contents

<b>Contents</b>	i
<b>List of figures and tables</b>	ix
<b>Abbreviations</b>	xiv
<b>Acknowledgments</b>	xx
<b>Declaration</b>	xxi
<b>Summary</b>	xxii
<b>Chapter I: General Introduction</b>	1
<b>1.1 Retroviruses</b>	1
1.1.1 Classification of retroviruses	1
1.1.2 General features of the retrovirus replication cycle	2
1.1.3 Endogenous retroviruses	9
<b>1.2 Retroviral Oncogenesis</b>	9
1.2.1 Retroviral transduction of cellular genes	10
1.2.2 Retroviral cis-insertional activation of cellular genes	10
<i>Viral promoter insertion</i>	12
<i>Viral enhancer activation</i>	13
<i>Truncation of cellular gene product</i>	13
1.2.3 Retroviral trans-activation of cellular genes	17
1.2.4 Insertional inactivation of cellular genes	20
<b>1.3 Cellular genes involved in cancer</b>	22
1.3.1 Oncogenes	23
1.3.2 Tumour suppressor genes: History and genetic characterisation	26
<i>Functional classification of tumour suppressor genes</i>	27
<i>Retinoblastoma TSG</i>	29

<i>P53 TSG</i>	30
<b>1.4 LVI-1 (WDR76) locus</b>	33
<b>1.5 LVI-1 gene as a member of the WD-40 gene family</b>	37
<b>1.6 Aims of the project</b>	39
<b>Chapter II: Materials and methods</b>	40
<b>2.1 Materials</b>	40
2.1.1 Cell culture materials	40
<i>Sources of cell lines</i>	40
<i>Plasticware</i>	41
<i>Solutions, media and supplements</i>	41
<i>Media</i>	41
Supplements	42
2.1.2 Bacterial strains	42
<i>E.coli One Shot<sup>®</sup> TOP10</i>	42
<i>E.coli BL21</i>	42
2.1.3 Complete kits	43
2.1.4 DNA	43
<i>Plasmid vectors</i>	43
<i>Molecular size standards</i>	44
<i>Oligonucleotide primers</i>	44
2.1.5 Enzymes	46
<i>Restriction enzymes</i>	46
<i>T4 DNA Ligase</i>	46
<i>Alkaline Phosphatase, Calf Intestinal (CIAP)</i>	46
<i>2x ReddyMix<sup>™</sup> PCR Master Mix</i>	46
<i>Extensor Hi-Fidelity PCR Master Mix, Buffer 1, ReddyMix<sup>™</sup></i>	48

<i>Ribonuclease A Type III-A</i>	48
2.1.6 RNA	48
2.1.7 Radiochemicals	48
2.1.8 Protein SDS-PAGE standards	48
2.1.9 General chemicals	49
2.1.10 Equipment	49
<i>Major Equipment</i>	49
<i>Consumables</i>	50
2.1.11 Buffers, solutions and growth media	51
<i>Water</i>	51
<i>Antibiotics</i>	51
<i>Buffers and solutions</i>	51
<i>Bacteriological media</i>	54
<b>2.2 Methods</b>	56
2.2.1 Growth and manipulation of mammalian cells	56
<i>Cryopreservation of cells</i>	56
<i>Cell counting</i>	56
<i>Maintenance of mammalian adherent cell lines</i>	57
<i>Maintenance of mammalian suspension cell lines</i>	57
2.2.2 Recombinant DNA techniques	58
<i>Storage and growth of bacteria</i>	58
<i>Extraction and purification of plasmid DNA</i>	58
<i>Large-scale plasmid preparation</i>	58
<i>Small-scale plasmid preparation</i>	59
<i>Determination of nucleic acid concentration and quality</i>	59
<i>Determination by spectrophotometry</i>	59

<i>Estimation of double stranded DNA concentration and quality by agarose gel electrophoresis</i>	59
<i>Restriction endonuclease digestion</i>	60
<i>Electrophoresis of DNA</i>	60
<i>Agarose gel electrophoresis</i>	60
<i>Polyacrylamide gel electrophoresis</i>	61
<i>Purification of restriction enzyme fragments</i>	62
<i>Cloning of hybrid DNA molecules</i>	62
<i>Ligation of vector and insert DNA</i>	62
<i>Transformation of bacteria with plasmid DNA</i>	63
<i>Transformation of commercially prepared competent bacteria</i>	63
<i>Preparation of freshly competent bacteria and subsequent transformation</i>	63
<i>Screening of transformants for desired recombinant plasmids</i>	64
2.2.3 Preparation of total RNA	65
<i>RNA extraction using RNA-Bee™</i>	65
<i>Assessment of RNA</i>	66
<i>Assessment of RNA using RNA 6000 Nano LabChip Kit</i>	66
<i>Assessment of RNA using agarose gel electrophoresis</i>	66
2.2.4 Amplification of DNA by polymerase chain reaction	66
<b><i>Primer design</i></b>	67
<i>Preparation of PCR reactions</i>	67
<i>Reaction conditions</i>	68
<i>Purification and assessment of PCR products</i>	68
<i>First strand DNA synthesis for reverse transcriptase (RT)-PCR</i>	69
2.2.5 DNA sequence analysis	69
<i>Automated sequencing</i>	69

<i>Sample preparation</i>	69
<i>Sample sequencing and evaluation</i>	70
2.2.6 DNA and RNA hybridisation analysis	72
<i>Southern blot transfer of DNA</i>	72
<i>Northern blot transfer of RNA</i>	72
<i>Preparation of radiolabelled DNA probes</i>	73
<i>Hybridisation of labelled probes to membrane bound nucleic acids</i>	73
2.2.7 Protein analysis	74
<i>Protein extraction</i>	74
<i>Estimation of protein concentration</i>	74
<i>SDS - PAGE of proteins</i>	75
<i>Electroblotting</i>	76
<i>ECL detection</i>	77
<b>Chapter III: Structure and coding potential of the human and murine</b>	
<b><i>LVI-1</i> genes</b>	79
<b>3.1 Introduction</b>	79
<b>3.2 Analysis of <i>LVI-1</i> genomic organisation in human and murine</b>	79
<b>3.3 Coding potential of the <i>LVI-1</i> genes</b>	83
<b>3.4 Predicted <i>LVI-1</i> gene products and their properties</b>	85
<b>3.5 Discussion</b>	92
<b>Chapter IV: Expression of the human and murine <i>LVI-1</i> genes</b>	93
<b>4.1 Introduction</b>	93
<b>4.2 Materials and methods</b>	93
4.2.1 Genomic PCR	93
4.2.2 Generation of a human <i>LVI-1</i> and $\beta$ -actin probes	93

4.2.3 Reverse transcription-polymerase chain reaction (RT-PCR)	96
4.2.4 5'Rapid Amplification of cDNA Ends Assay (5'RACE)	96
<b>4.3 <i>Lvi-1</i> gene expression</b>	<b>97</b>
4.3.1 Generation of a murine specific <i>Lvi-1</i> probe	97
4.3.2 Analysis of murine <i>Lvi-1</i> gene expression	99
4.3.3 Coding potential of the murine <i>Lvi-1</i> gene	103
4.3.4 Identification of the transcription initiation site of the murine <i>Lvi-1</i> gene	106
4.3.5 Identification of the transcription initiation site of the human <i>LVI-1</i> gene	110
4.3.6 Promoter switching and alternative <i>LVI-1</i> exon usage	113
<b>4.4 Discussion</b>	<b>118</b>
<b>Chapter V: Characterisation of LVI-1 protein products</b>	<b>120</b>
<b>5.1 Introduction</b>	<b>120</b>
<b>5.2 Material and methods</b>	<b>120</b>
5.2.1 Production of recombinant expression vector	120
<i>Preparation of DNA encoding the hLVI-1 C-terminus</i>	121
<i>Preparation of DNA encoding the mLvi-1 C-terminus</i>	121
<i>Preparation of the pGEX vector</i>	121
<i>Cloning of the insert DNA into the pGEX vector</i>	123
5.2.2 Expression of LVI-1-GST fusion proteins	123
<i>Transformation of E.coli BL21 cells</i>	123
<i>Small scale expression of LVI-1-GST fusion proteins</i>	124
<i>Small scale purification of LVI-1-GST fusion proteins</i>	124
<i>Analytical scale preparation of insoluble LVI-1-GST fusion proteins</i>	125

<i>Preparative scale preparation of insoluble LVI-1-GST fusion proteins</i>	125
<i>Extraction of the solubilised LVI-1-GST fusion proteins from SDS-PAGE</i>	126
<i>Dialysis and concentration of the eluted LVI-1-GST fusion proteins</i>	127
<b>5.2.3 Identification of LVI-1-GST fusion proteins</b>	127
<i>Identification on Coomassie stained gels</i>	128
<i>Identification by western blot analysis</i>	128
<b>5.2.4 Mass spectrometry analysis</b>	128
<b>5.2.5 Measurement of purity and estimation of concentration</b>	129
<i>Estimation of LVI-1-GST fusion proteins concentration by silver staining</i>	129
<i>Estimation of LVI-1-GST fusion protein by Bio-Rad Protein Assay</i>	129
<b>5.2.6 Production of polyclonal antibodies against GST fusion protein and its immunodetection by western blotting</b>	129
<b>5.2.7 Validation of polyclonal antisera</b>	130
<i>Sera affinity chromatography</i>	130
<i>Detection of specificity</i>	130
<b>5.3 Construction and purification of LVI-1 fusion proteins</b>	131
<b>5.3.1 Peptide selection</b>	131
<b>5.3.2 GST gene fusion system</b>	131
<b>5.3.3 Generation and expression of LVI-1 GST fusion proteins</b>	133
<b>5.3.4 Purification of LVI-1 GST fusion proteins</b>	136
<b>5.3.5 Validation by mass spectrometry</b>	136
<b>5.3.6 Quantification of LVI-1-GST fusion proteins for immunisation</b>	140
<b>5.3.7 Immunisation Schedule</b>	140
<b>5.3.8 Validation of polyclonal antisera</b>	144
<b>5.3.9 Endogenous expression of murine Lvi-1</b>	146
<b>5.4 Discussion</b>	149

<b>Chapter VI: Clues to LVI-1 protein function</b>	151
<b>6.1 Introduction</b>	151
<b>6.2 Material and methods</b>	152
6.2.1 Phylogenetic analysis of LVI-1 amino acid sequences	152
6.2.2 Cellular exposure to UV irradiation	152
<i>UV treatment of adherent cells</i>	152
<i>UV treatment of suspension cells</i>	152
6.2.3 Determination of the <i>LVI-1</i> transcripts levels after UV irradiation by RT-PCR	153
6.2.4 Cell cycle analysis by flow cytometry	153
6.2.5 Immunofluorescent labelling and confocal microscopy	155
6.2.6 Combined immunoprecipitation and western blot analysis	156
<b>6.3 Homologues of LVI-1</b>	157
<b>6.4 Functional analysis</b>	161
6.4.1 Responses of Lvi-1 to UV irradiation	161
<i>Effects on Lvi-1 RNA</i>	161
<i>Effects on Lvi-1 protein levels</i>	163
<i>Effects on LVI-1 protein localisation</i>	167
6.4.2 Lvi-1 protein interactions	173
<b>6.5 Discussion</b>	177
 <b>Chapter VII: General Discussion</b>	 179
 <b>References</b>	 184

## List of figures and tables

### Chapter I

Figure 1-1.	Schematic representation of the retrovirus virion depicting the relationship of the retrovirus genes and the location of the different proteins	3
Figure 1-2.	Genomic organization of the proviral FIV genome	5
Figure 1-3.	Retroviral replication cycle	7
Figure 1-4.	Retroviral transformation by transduction of an oncogene	11
Figure 1-5.	Mechanisms of retroviral insertional activation	14
Figure 1-6.	Retroviral <i>trans</i> -activation of an oncogene	19
Figure 1-7.	Genomic structure of the human chromosome 15q15 surrounding the <i>LVI-1</i> gene, based on the draft human genome ( <a href="http://www.genome.ucsc.edu">http://www.genome.ucsc.edu</a> )	36
Table 1-1.	Genes inactivated by retroviral integrations in cancer	21
Table 1-2.	Some well characterised proto-oncogenes	25
Table 1-3.	Representative TSGs	28

### Chapter II

Figure 2-1.	Organization of pGEX -5X-1 and pGEX-5X-2 vectors (GST gene fusion system)	45
Table 2-1.	Restriction enzymes	47
Table 2-2.	Sequencing primers	71

### Chapter III

Figure 3-1.	Genomic structure of the human and murine homologues of <i>LVI-1</i>	80
Figure 3-2.	Exon structure of the 5' region of <i>LVI-1</i> gene	82

Figure 3-3.	Fine structure of the P1 and P2 promoter regions of the human and murine <i>LVI-1</i> genes	84
Figure 3-4.	Predicted amino acid sequence comparison of human and murine <i>LVI-1</i> cDNAs	86
Figure 3-5.	Potential start sites and predicted protein sizes	88
Figure 3-6.	Structural features of predicted human and murine <i>LVI-1</i> gene products	89
Table 3-1.	Predicted location of predicted <i>LVI-1</i> gene products	91
 <b>Chapter IV</b>		
Figure 4-1.	Amplification of murine <i>Lvi-1</i> exon 13	98
Figure 4-2.	Conservation of the nucleotide sequence and coding potential of exon 13 of the human and mouse <i>LVI-1</i> gene	100
Figure 4-3.	Validation of murine exon 13 fragment as a single copy gene probe	101
Figure 4-4.	Expression of <i>Lvi-1</i> gene transcripts in normal murine tissues	102
Figure 4-5.	Northern blot analysis of <i>LVI-1</i> expression in various human and murine cell lines	104
Figure 4-6.	Analysis of murine <i>Lvi-1</i> transcripts structure	105
Figure 4-7.	Nucleotide and predicted amino acid sequence of the murine <i>Lvi-1</i> cDNA isolated from normal mouse thymus	107
Figure 4-8.	Nucleotide sequence comparison of the 5' start of murine and human <i>LVI-1</i> cDNAs	108
Figure 4-9.	Determination of the murine <i>Lvi-1</i> transcription start site	111
Figure 4-10.	Nucleotide sequence of the murine <i>Lvi-1</i> RACE clones	112
Figure 4-11.	Determination of the human <i>LVI-1</i> transcription start site	114

Figure 4-12.	Nucleotide sequence of human <i>LVI-1</i> RACE clones	115
Figure 4-13.	Expression from the P2 promoter of murine <i>Lvi-1</i> gene	117
Table 4-1.	Primer sequences (by hand)	94
Table 4-2.	Primer combinations, PCR annealing conditions, and applications	95

## Chapter V

Figure 5-1.	Kyle-Doolittle amino acid hydrophilicity plot of AK023035 and m <i>Lvi-1</i> amino acid sequences	132
Figure 5-2.	Strategy for construction of pGEX-5X-1-h <i>LVI-1</i> and pGEX-5X-2-m <i>Lvi-1</i> plasmids	134
Figure 5-3.	Induction of the h <i>LVI-1</i> -GST and m <i>Lvi-1</i> -GST fusion proteins with IPTG	135
Figure 5-4.	Small scale purification of the human and murine pGEX recombinants for fusion protein expression	137
Figure 5-5.	Solubilisation of m <i>Lvi-1</i> -GST and h <i>LVI-1</i> -GST fusion proteins	138
Figure 5-6.	Western blot analysis of <i>LVI-1</i> -GST fusion proteins	139
Figure 5-7.	Identification of <i>LVI-1</i> -GST fusion proteins by mass spectrometry	141
Figure 5-8.	Silver staining detection of m <i>Lvi-1</i> -GST fusion proteins	142
Figure 5-9.	Analysis of reactivities in rabbit sera during immunisation	143
Figure 5-10.	Absorption of GST reactivity from antisera to <i>LVI-1</i> fusion protein	145
Figure 5-11.	Reactivity of antisera to <i>LVI-1</i> -GST fusion proteins after absorption	147
Figure 5-12.	Western blot analysis of <i>Lvi-1</i> expression in various murine cell lines	148

**Chapter VI**

Figure 6-1. Output from BLAST search of UniProt protein database (Swiss-Prot + TrEMBL) using the full length predicted human LVI-1 protein (626 amino acids) 158

Figure 6-2. Phylogenetic tree of homologues of LVI-1 160

Figure 6-3. Determination of human *LVI-1* transcript levels after UV irradiation 162

Figure 6-4. Determination of murine *Lvi-1* transcript levels after UV irradiation 164

Figure 6-5. *Lvi-1* expression in p/m 16i cell line after UV irradiation 165

Figure 6-6. *Lvi-1* expression in H/3T3 cell line after UV irradiation 166

Figure 6-7. Detection of *Lvi-1* transcript after UV irradiation of H/3T3 cells 168

Figure 6-8. P53 protein expression in H/3T3 cell line after UV irradiation 169

Figure 6-9. Analysis of p53 expression in various murine cell lines by western blot 170

Figure 6-10. *Lvi-1* promoter use is independent of p53 status in primary murine fibroblasts (Mefs) 171

Figure 6-11. Localisation of LVI-1 protein by immunofluorescence in H/3T3, p/m 16i, and Jurkat cell lines 172

Figure 6-12. Percentage of cytoplasmic staining of the murine (H/3T3 and p/m 16i) and human (Jurkat) cell lines before and after UV treatment 174

Figure 6-13. Interaction of MSH6 with MSH2 is detected by Co-immunoprecipitation 175

Table 6-1. Primers used for analysis of LVI-1 expression in response to UV irradiation 154



## Abbreviations

°C	Degrees centigrade
6-4PP	6-4 pyrimidine-pyrimidone photoproducts
ALV	Avian leucosis virus
Apaf-1	Apoptotic protease activating factor-1
APC	<i>Adenomatous Polyposis Coli</i>
APS	Ammonium persulfate
ARF	Alternative reading frame
ATM	Ataxia telangiectasia mutated
ATR	ATM and Rad3-related
A <sub>x</sub>	Absorbance <sub>wavelength (nm)</sub>
BIV	Bovine immunodeficiency viruses
BLM	<i>Bloom mutated</i>
BLV	Bovine leukaemia virus
Bp	Base pair
BRCA 1	<i>Breast cancer-associated 1</i>
BRCA 2	<i>Breast cancer-associated 2</i>
BSA	Bovine serum albumin
CA	Capsid
CAEV	Caprine arthritis-encephalitis virus
Cdks	Cyclin dependent kinases
cDNA	Complementary deoxyribonucleic acid
CIAP	Calf intestinal alkaline phosphatase
Cip1	Cdk-interacting protein 1
<i>c-onc</i>	Cellular proto-oncogenes
COP1	Constitutively photomorphogenic 1
CPDs	Cyclobutane pyrimidine dimers
CpG	Cytosine phosphate guanosine
cpm	Counts per minute
C-terminus	Carboxyl-terminus
Cul4A	Cullin
DAPI	4',6-diamidino-2-phenylindole

DDB	Damaged DNA-binding
ddH <sub>2</sub> O	Double distilled water
DEPC	Diethylpyrocarbonate
dH <sub>2</sub> O	Distilled water
DMEM	Dulbecco's Modified Eagles Medium
<i>Dmp1</i>	<i>Cyclin D-binding Myb-like protein</i>
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNTP	3' deoxyribonucleoside 5' triphosphate
Ds	Double stranded
DTT	Dithiethreitol
DU	dUTPase
dUMP	Deoxyuridine monophosphate
dUTPase	Deoxyuridine triphosphatase
dUTP	Deoxyuridine triphosphate
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetra-acetic acid, disodium salt
EGTA	Ethylenediguaninetetra-acetic acid, disodium salt
EIAV	Equine infectious anaemia virus
ENTV	Enzootic nasal tumour virus
<i>Env</i>	<i>Envelope</i>
ERVs	Endogenous retroviruses
ETS	E26-Transformation-Specific
<i>Evi-2</i>	Ecotropic viral integration site 2
FBS	Foetal bovine serum
FeLV	Feline leukaemia virus
FITC	Fluorescein isothiocyanate
FIV	Feline immunodeficiency viruses
F-MuLV	Friend murine leukaemia virus
FV-P/FV-A	Friend virus-polycythemia/ Friend virus-anaemia
g	Gram
G	Gravity
<i>Gag</i>	<i>Group-specific antigen</i>
GAPDH	Glyceraldehyde phosphate dehydrogenase

GGR	Global genomic repair
GH	Glycine-histidine
Gro/TLE	Groucho /transducin-like enhancers
GST	Glutathione S-Transferase
Hepes	N-(2-hydroxyethyl) piperazine-N'-(2' ethanesulphonic acid)
HERV	Human endogenous retrovirus
HIF-1	Hypoxia inducible transcription factor 1
HIV	Human immunodeficiency viruses
HPR	Horseradish peroxidase
HPRT	Hypoxanthine phosphoribosyl-transferase
Hr	Hour
HSF	Heat shock factor
HSP	Heat shock protein
HSTF-2	Heat shock transcription factor-2
HTLV	Human T-lymphotropic virus
HUGO	Human genome organisation
IN	Integrase
IU	International units
IPTG	Isopropyl $\beta$ -D-thiogalactoside
JSRV	Jaagsiekte sheep retrovirus
kb	Kilobase pairs
kDa	Kilodalton
L	Litre
LB	Luria broth
LOH	Loss of heterozygosity
LTRs	Long terminal repeats
<i>LVI-1</i>	<i>Lentivirus integration-1</i>
m	Milli
M	Molar
MA	Matrix protein
MALDI-TOF	Matrix-Associated Laser Desorption Ionization-Time of Flight
MES	2-(N-morpholino) ethane sulfonic acid
<i>MFAP1</i>	<i>Microfibrillar associated protein 1</i>
MLV	Murine leukaemia virus

MLVs	Murine leukaemia viruses
MMTV	Mouse mammary tumour virus
MOPS	3-N-Morpholino-propanesulfonic acid
mRNA	Messenger ribonucleic acid
MVV	Maedi-visna virus
n	Nano
NC	Nucleocapsid
NER	Nucleotide excision repair
<i>NF1</i>	<i>Neurofibromatosis-1</i>
<i>NFATc3</i>	<i>Nuclear factor of activated T cell 3</i>
Nfe2	Nuclear factor erythroid 2
N-Terminus	Amino-terminus
OD	Optical density
Orfs	Open reading frames
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI	Propidium iodide
PMSF	Phenylmethanesulfonyl fluoride
<i>Pol</i>	<i>Polymerase</i>
<i>PTEN</i>	<i>Phosphatase and tensin homologue</i>
PR	Protease
R	Repeat
Rb	Retinoblastoma
RFLPs	Restriction fragment length polymorphisms
RNase	Ribonuclease
RNasin	Ribonuclease inhibitor
rpm	Revolutions per minute
RPMI	Rosewall Park Memorial Institute medium
RSV	Rous sarcoma virus
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase- Polymerase chain reaction
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis
SFFV	Friend spleen focus-forming virus

SFV	Simian foamy virus
SHWFGF	Sir Henry Wellcome Functional Genomics Facility
siRNA	Small interfering RNA
SIV	Simian immunodeficiency viruses
Sp1	Stimulating protein 1
SRV	Simian retrovirus D-type
Ss	Single stranded
SU	Surface
SV40 T	Simian vacuolating virus 40 large tumour
T	Deoxythymidine
TAE	Tris-acetate-EDTA
Taq	<i>Thermus aquaticus</i>
TBE	Tris-borate-EDTA
TBST	Tris-buffered-saline-tween
TCR	Transcription-coupled repair
TE	Tris-EDTA
TEMED	N,N,N',N',-Tetramethylethylenediamine
TM	Transmembrane
Tris	2-amino-2-(hydroxymethyl)-1,3- propanediol
TSG	Tumour suppressor gene
TSS	Transformation and Storage Solution
TTP	Thymidine triphosphate
U	Units
U3	Unique 3' region
U5	Unique 5' region
UTR	Untranslated region
UV	Ultraviolet
V/v	Volume for volume
VHL	<i>Von Hippel-Lindau</i>
v- <i>onc</i>	Viral oncogenes
W/v	Weight for volume
W-D	Tryptophan-aspartic acid
WDSV	Walleye dermal sarcoma virus
Wt	Wild type

<i>WT1</i>	<i>Wilms Tumour</i>
X-Gal	5-bromo-4-chloro-3-indolyl- $\beta$ galactoside
XP	Xeroderma pigmentosum
$\mu$	Micro

**Single letter amino-acid code:**

G	Glycine
A	Alanine
L	Leucine
M	Methionine
F	Phenylalanine
W	Tryptophan
K	Lysine
Q	Glutamine
E	Glutamic acid
S	Serine
P	Proline
V	Valine
I	Isoleucine
C	Cysteine
Y	Tyrosine
H	Histidine
R	Arginine
N	Asparagine
D	Aspartic acid
T	Threonine

Throughout this thesis genes are indicated by italics e.g. *LVI-1*, whilst proteins remain unitalicised e.g. LVI-1.

## Acknowledgements

First, I would like to thank my sponsor, the Portuguese Foundation (FCT, Fundação para a Ciência e a Tecnologia) for funding me throughout this study and also during my first year of the Portuguese PhD programme (GABBA, programa Graduado em Áreas da Biologia Básica e Aplicada), and to the Molecular Oncology Laboratory for accepting me.

Second, I would like to express my sincere gratitude to my supervisors, Prof. James Neil and Dr. Monica Stewart. I fully appreciate their contribution to the project and their supervision while I was writing my thesis, and also for their constant support. In addition, I would like to acknowledge Dr Anna Kilbey for being so patient when teaching new techniques and also for helping me through my writing.

I would also like to express my sincere thanks to the many people who helped me throughout the duration of my PhD. In particular, Dr Linda Hanlon, Ms Sharon Mackay, Dr. Karen Blyth, Dr. Francois Vaillant, Mrs Anne Terry, Ms Alma Jenkins, Dr. Sandy Wotton, Mrs Nancy Mackay, Mrs Fenella Long, Mrs Margaret Bell, and Prof Ewan Cameron. Without them I would certainly not have gained the experience and knowledge I now have.

My appreciation also goes to others at the Veterinary School such as Prof. Massimo Palmarini, Mr Tom MacPherson, Ana Monteiro, Drew McConnell, Linda Hanlon, Sharon Mackay, Karen Blyth, Ana Mateus, Susan Ridha, Anna Garcia, Marco Caporale, Chiara Pinoni, Nuno Costa, Peck Toung Ooi, Mariana Varela, Pablo Murcia, Manuela Mura, Claudio Murgia, Kamil Wolyniec, Ralph Hector, Louise Grant and all of those who made me feel very much a part of the team and contributed to making my time at the Veterinary school a very pleasurable and memorable one. I feel very privileged to have been a part of that for three years. They also made me laugh a lot!

Finally I would like to thank to my parents, Tom, all my family, and friends for their support, patience and encouragement. However, I want to dedicate this thesis to my grandmother (avó) and grandfather (avô).

## **Declaration**

The studies described in this thesis were carried out in the Molecular Oncology Laboratory at the University of Glasgow Veterinary School between September 2002 and September 2005. The author was responsible for all results unless otherwise stated.



Alexandra de Sousa Montenegro Miranda, April 2006

## Summary

The central aim of this study was to characterise the expression of the candidate tumour suppressor gene, *LVI-1* (*lentivirus integration-1*) and its products. The *LVI-1* gene (*WDR76*) was discovered as a target for disruption by proviral insertional mutagenesis in a case of pre-B cell lymphoma in a FIV infected cat {Beatty, 1998 5 /id; Beatty, 2002 7 /id}. As *LVI-1* is highly conserved, my work focused on the human and murine orthologues to take advantage of the superior resources available for these species. My work showed potentially important differences in expression of the human and mouse genes with respect to promoter use and length of 3' untranslated sequences. The murine gene is transcribed mainly from the distal P1 promoter, which appears to be a bi-directional element shared with the adjacent *Mfap1* gene, while the human gene transcripts are derived exclusively from the proximal P2 promoter. Direct analysis by RT-PCR showed that the murine gene could also be expressed from the P2 promoter. These findings have significant implications for Lvi-1 protein expression as the P2 transcripts are predicted to express larger proteins with a distinct N-terminal sequence. To characterise the *LVI-1* gene products, rabbit polyclonal antisera were raised to GST fusion proteins expressed in bacteria. Use of the murine anti-mLvi-1 antiserum in Western blotting identified a protein of the expected size (58 kDa) based on translation of the major mRNA species from the P1 promoter. Immunofluorescence and confocal microscopy suggested that this protein is localised mainly in the cytoplasm. Although the function of LVI-1 is unknown, its closest relative in the human genome is DDB2, a protein involved in repair of UV-induced DNA damage. Regulation of LVI-1 expression was examined after UV irradiation, providing preliminary evidence of responses at transcriptional and post-transcriptional levels. Further leads were followed by analogy with the yeast orthologue of LVI-1, YDL156W, but no evidence of a complex between LVI-1 and MSH6 was found. In conclusion, while function of the LVI-1 gene remains to be established, it provides the basis for future characterisation of this highly conserved and potentially important eukaryotic gene.

# Chapter I

## General Introduction

### 1.1 Retroviruses

The study of retroviruses began with the isolation of the first transmissible cancer-causing agents by Ellermann and Bang in 1908 (1) and by Rous in 1911 (2). Retroviruses are associated with a variety of diseases including malignancies, neurological disorders, wasting diseases and immunodeficiency, however some retroviruses are apathogenic. Their name derives from their capacity for reverse transcription, a process by which the viral RNA genome is converted into DNA by the viral enzyme reverse transcriptase (RT). The DNA copy can incorporate stably into the host cell genome as a provirus and, in some cases after a period of latency, direct the cellular machinery to reproduce the virus. The study of retroviruses led to the discovery and development of the oncogene theory of tumorigenesis: some retroviruses actually contained oncogenes within their genomes, while others interact with oncogenes in either a direct or indirect way to contribute to tumour formation (3, 4).

#### 1.1.1 Classification of retroviruses

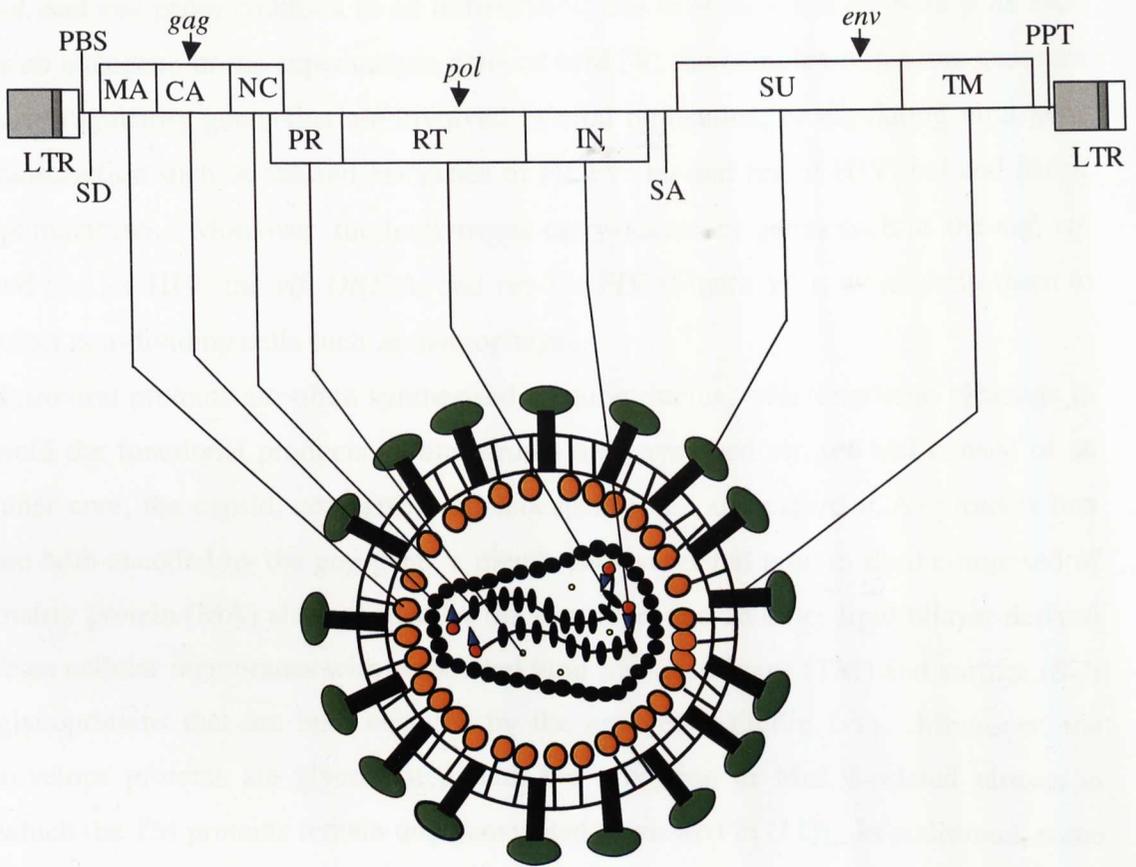
Historically, because of their pattern of pathogenicity, the retroviruses were grouped into three subfamilies. The subfamily Oncovirinae, the acute oncogenic retroviruses; Lentivirinae, associated with 'slow' diseases or those with long latent periods, and the Spumavirinae, the "foamy" viruses, named because of the pathogenic changes observed in infected cells. Retroviruses were classified (Types A-D) according to the electron microscopic appearance of their nucleocapsid structures (5). The most recent classification of the *Retroviridae* family on the basis of sequence relatedness defines seven genera: alpharetrovirus, such as Rous sarcoma virus (RSV), and avian leucosis virus (ALV); betaretrovirus, such as mouse mammary tumour virus (MMTV), simian retrovirus D-type (SRV), Jaagsiekte sheep retrovirus (JSRV), and enzootic nasal tumour virus (ENTV); gammaretrovirus, such as murine leukaemia virus (MuLV) and feline leukaemia virus (FeLV); deltaretrovirus, such as bovine leukaemia virus (BLV) and human T-lymphotropic virus (HTLV); epsilonretrovirus, such as Walleye dermal

sarcoma virus (WDSV); lentivirus, such as maedi-visna virus (MVV) of sheep, caprine arthritis-encephalitis virus (CAEV), equine infectious anaemia virus (EIAV) and the human, simian, feline and bovine immunodeficiency viruses (HIV, SIV, FIV, BIV) and spumavirus such as simian foamy virus (SFV). These genera can be grouped into simple retroviruses (alpharetrovirus, betaretrovirus, gammaretrovirus) and complex retroviruses (deltaretrovirus, lentivirus, spumavirus, and epsilonretrovirus).

### 1.1.2 General features of the retrovirus replication cycle

Retroviral virion particles contain two identical single stranded, genomic RNA molecules of positive polarity that are replicated through a double-stranded DNA intermediate (6, 7). The 5' end of the genomic RNA begins with the "R" (for repeat) and "U5" (for unique 5' region) segments, followed by the viral genes *gag* ('group-specific antigen'), *pol* ('polymerase'), *env* ('envelope'), which are the three major open reading frames (orfs) (Figure 1-1). However, this is not the basic structure because many retroviruses contain additional orfs. The 3' end of the genomic RNA terminates with the U3 (for unique 3' region) and R (identical to the 5' region) regions and a polyA tail. Following reverse transcription of the RNA genome into a double strand (ds) DNA copy, this then integrates into the host cell chromosomal DNA, where it is thereafter referred to as the "provirus". Because of the mechanism used for reading and utilising the viral RNA template during reverse transcription (7-9), the U3 and U5 regions are duplicated such that the 5' and 3' ends of the proviral genome differ in structure from the ends of the RNA genome.

Each end of the proviral genome is made up of regions called long terminal repeats (LTRs) which contain the proviral U3, R, and U5 regions. This arrangement of both termini of the viral genome enables appropriate expression of the viral genes. The U3 region contains the enhancer and promoter sequences that drive viral RNA transcription from the 5' LTR. In many retroviruses the viral *gag* and *pol* genes are expressed from an unspliced *gagpol* transcript while, the *env* gene is invariably expressed from a spliced transcript (the splice donor is located between the 3' end of the 5' LTR and the 5' end of the *gag* gene, with the splice acceptor located at the 5' end of the *env* gene) (10). While the genome of simple retroviruses carries the *gag*,

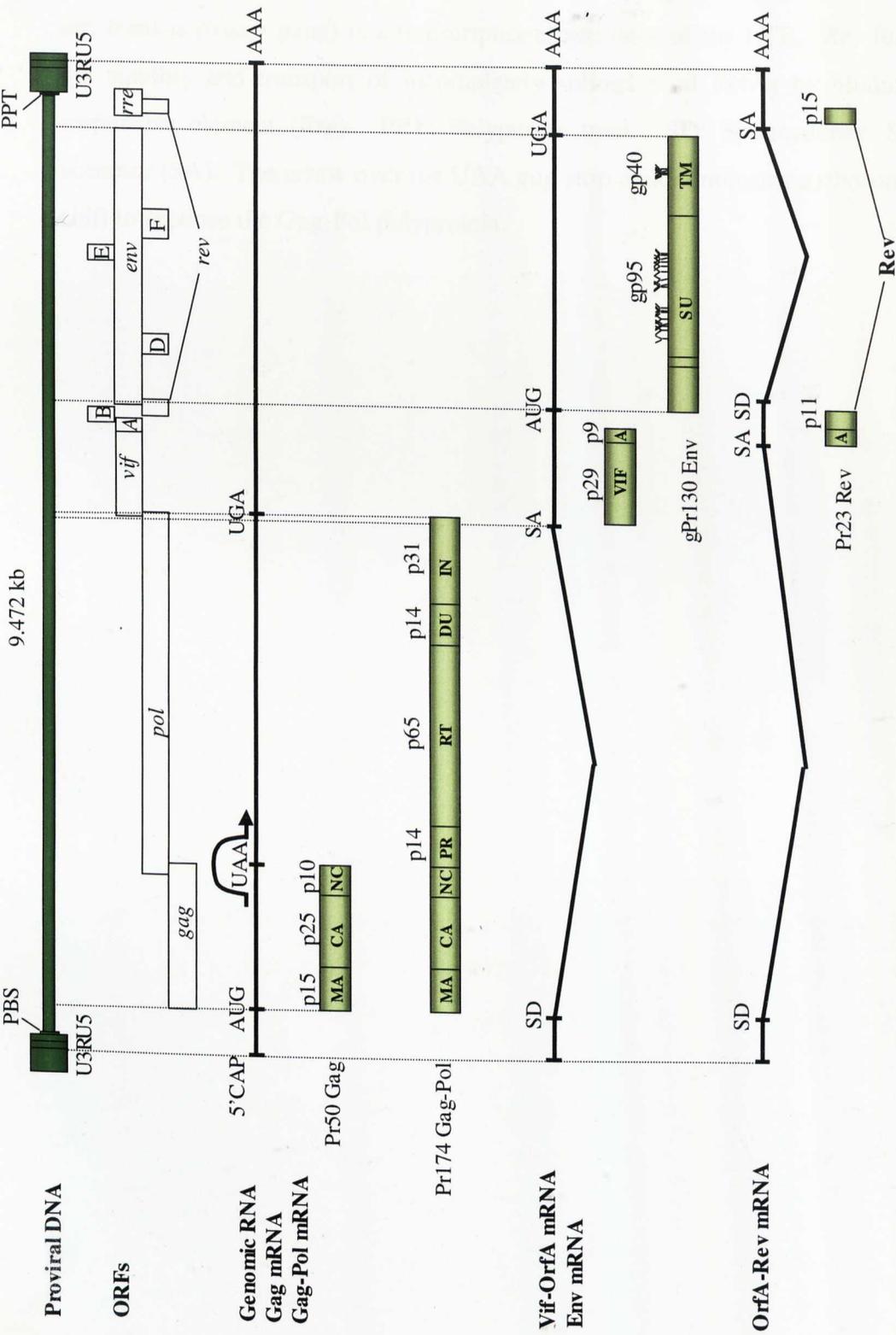


**Figure 1-1. Schematic representation of the retrovirus virion depicting the relationship of the retrovirus genes and the location of the different proteins.** LTR: long terminal repeat, PBS: primer binding site, PPT: polyurine tract. This figure is not to scale and the regions vary in length between different retroviruses. *Gag* encodes structural viral proteins: matrix protein (MA), capsid protein (CA), and nucleocapsid protein (NC). *Pol* encodes the enzymatic functions: protease (PR), reverse transcriptase (RT), and integrase (IN). *Env* encodes viral glycoproteins: surface protein (SU), and transmembrane protein (TM).

*pol*, and *env* genes common to all retroviruses, and in some cases an extra gene such as an oncogene or the superantigen gene of MMTV; the complex retrovirus genomes carry regulatory genes that are involved in viral replication, or regulating viral gene transcription such as *tax* and *rex* genes of HLTV; *tat* and *rev* of HIV; *bel* and *bet* of spumaviruses. Moreover, the lentiviruses carry accessory genes such as the *nef*, *vif*, and *vpr* for HIV; the *vif*, *ORF A*, and *rev* for FIV (Figure 1-2), which help them to infect non-dividing cells such as macrophages.

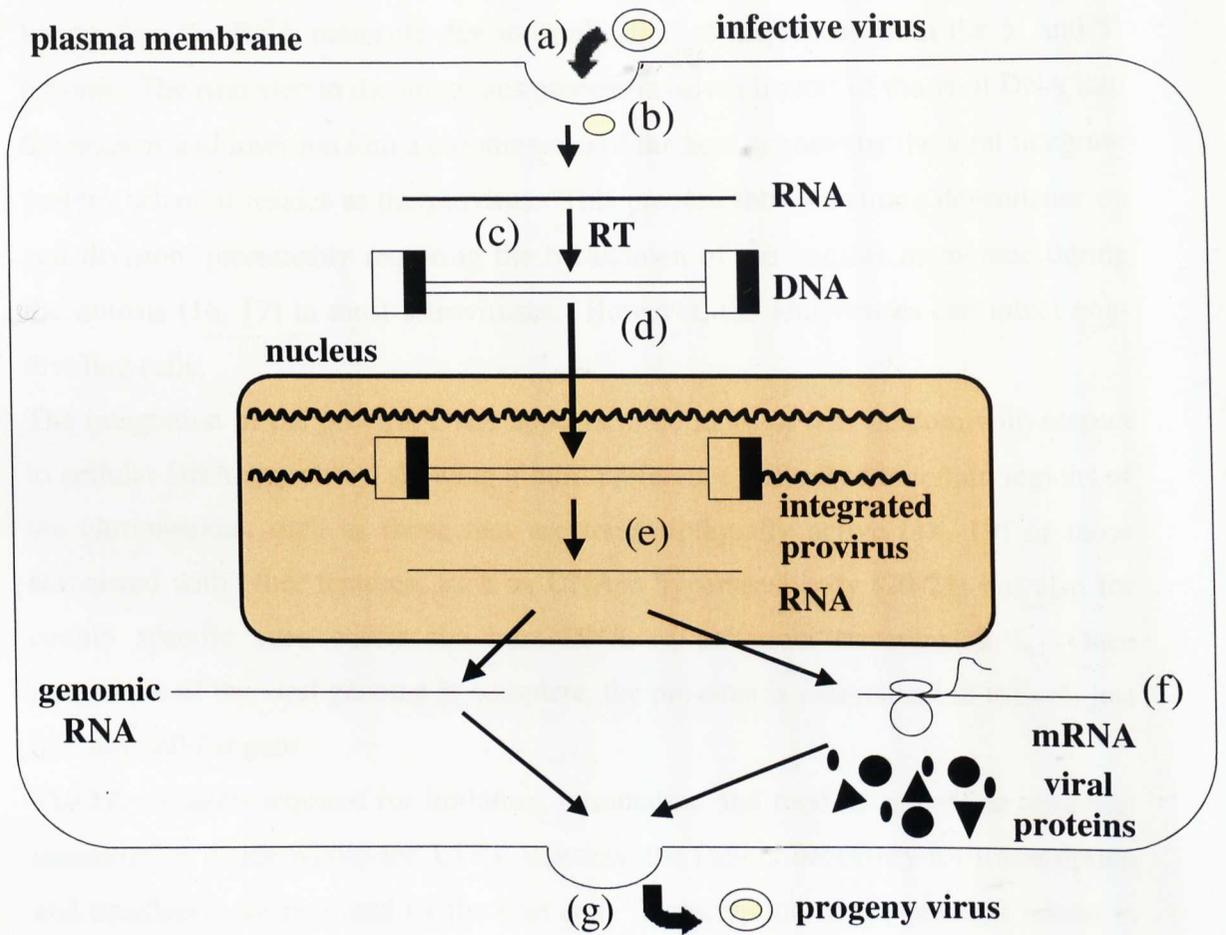
Retroviral proteins are often synthesised as polyproteins, which undergo cleavage to yield the functional products. Retroviruses are enveloped viruses and consist of an inner core, the capsid, comprising nucleocapsid (NC) and capsid (CA) proteins that are both encoded by the *gag* gene, a membrane-associated protein shell composed of matrix protein (MA) also encoded by the *gag* gene, and an outer lipid bilayer derived from cellular membranes with embedded viral transmembrane (TM) and surface (SU) glycoproteins that are both encoded by the *env* gene (Figure 1-1). Moreover, the envelope proteins are glycosylated with the exception of MuLV-related viruses in which the TM proteins remain unglycosylated (reviewed in (11)). In addition, some retroviruses also encode various other small proteins within the *gag* orf. The *pol* gene encodes the viral replication enzymes: reverse transcriptase (RT), an RNA-dependent DNA polymerase necessary for transcription of viral RNA into DNA; integrase (IN), necessary for integrating the virus into the host cell genome; protease (PR) which is encoded within *pol* for most retroviruses, is necessary for proteolytic cleavage of primary translation products; *pol* also encodes an RNase H activity that removes the RNA strand from the hybrid RNA-DNA double strand (reviewed in (12)). *Pol* orf of certain lentiviruses such as FIV, EIAV, CAEV, and visna virus encodes a deoxyuridine triphosphatase (dUTPase, DU) (Figure 1-2), which hydrolyzes dUTP (deoxyuridine triphosphate) to dUMP (deoxyuridine monophosphate). In fact, this reaction not only provides a source of dUMP, an essential precursor of TTP (thymidine triphosphate), but also maintains a low dUTP/TTP ratio that minimizes dUTP incorporation into chromosomal DNA which leads to excision repair, DNA breakage, and loss of viability (13, 14).

Retroviruses generally require actively dividing cells for productive infection (15). The retroviral replication cycle (Figure 1-3) is initiated by binding of the viral envelope glycoprotein to a specific cell surface receptor. This membrane fusion results in release of the viral core into the cytoplasm. The viral RNA is uncoated and



**Figure 1-2. Genomic organization of the proviral FIV genome.** The provirus is flanked by long terminal repeats (LTRs) at both ends, which can be subdivided into: U3, R and U5. U3 regions. PBS: tRNA primer binding site. *Gag* encodes structural viral proteins: matrix protein (MA), capsid protein (CA), and nucleocapsid protein (NC). *Pol* encodes the enzymatic functions: protease (PR), reverse transcriptase (RT) with

RNAse H, dUTPase (DU), and integrase (IN). *Env* encodes viral glycoproteins: surface protein (SU) and transmembrane protein (TM). Y: glycosylation sites. *Vif* (virion infectivity factor) controls virion release from lymphocytes. *ORF A* (putative *tat*, *trans-activator* gene) is a transcriptional activator of the LTR. *Rev* functions in the stability and transport of incompletely spliced viral RNAs by binding to *rev*-responsive element (*Rre*). PPT: Polypurine tract. SD: Splice donor. SA: splice acceptor (SA). The arrow over the UAA *gag* stop codon indicates a ribosomal frame-shift to produce the Gag-Pol polyprotein.



**Figure 1-3. Retroviral replication cycle.** Retrovirus binds to cellular receptors (a) through virally encoded receptor binding proteins. Entry into the cell (b) is followed by reverse transcription of the viral genome (c) to produce a DNA intermediate or provirus. The provirus then integrates into the genome (d) and the random nature of this process typically results in insertions into one allele of diploid genomes. Transcription (e) and translation (f) are followed by packaging of progeny virus and budding (g) from the host cell.

reverse transcribed into a double-stranded, blunt-ended, linear DNA copy. This process is catalyzed by the viral RT, resulting in a DNA molecule which is slightly longer than the RNA molecule due to duplication of sequences from the 5' and 3' termini. The next step in the infectious process involves import of the viral DNA into the nucleus and insertion into a chromosome of the host genome by the viral integrase protein, where it resides as the provirus. This process shows a strong dependence on cell division, presumably requiring the breakdown of the nuclear membrane during the mitosis (16, 17) in most retroviruses. However, the lentiviruses can infect non-dividing cells.

The integration of the proviral DNA appears to be more or less random with respect to cellular DNA sequences, showing a mild preference not only for certain regions of the chromosome, such as those that are transcriptionally active (18, 19) or those associated with other features, such as DNase hypersensitivity (20-22) but also for certain specific sites within the host DNA of unknown structure (23). Once integration of the viral genome is complete, the provirus is maintained in the cell just like any cellular gene.

The *cis*-elements required for initiation, termination and modulation of the retroviral transcription reside within the LTRs, whereas, the factors necessary for transcription and translation are provided by the host cell. Transcription of the provirus results in production of copies of viral genomic RNA, which can either be translated to produce viral proteins or packaged by viral proteins to form new core particles (10). Two types of viral messenger RNAs from the provirus are produced in infected cells. The major spliced RNA is the sub-genomic RNA for the env proteins, ensuring adequate levels to populate the unit membrane that surrounds the viral particle. Splicing also results in removal of the viral packaging signal preventing encapsidation of the sub-genomic RNA into virions. The unspliced genomic RNAs that do carry the packaging signal become assembled into viral cores that bud through the plasma membrane of infected cells.

The ability of a retrovirus to propagate in a given cell type is mainly determined by the tissue specificity of the enhancer sequences in the LTR, and by the viral envelope gene (24).

### **1.1.3 Endogenous retroviruses**

The ability of retroviruses to integrate in the host cell genome has provided an opportunity for these viruses to colonise during the evolution of the genome of virtually all vertebrates. Retroviruses usually infect somatic cells, and consequently most retroviral genes integrated into genomic DNA are not passed on to host progeny. However, some classes of retrovirus occasionally infect germ line cells and in this way, the integrated DNA provirus can be passed on to the next generation without undergoing further viral replication. Offspring that develop from infected germ cells will carry the integrated retrovirus as part of their genomes, and these retroviruses can therefore be subsequently transmitted vertically from one host generation to the next in Mendelian fashion. Retroviruses that enter the germline in this way are referred to as endogenous retroviruses (ERVs) to distinguish them from horizontally transmitted, “exogenous” retroviruses (25). The great majority of these ERVs are truncated and inactivated by point mutations and deletions accumulated in the course of evolution, however some ERVs are transcriptionally active, and both detrimental and beneficial effects to the host have been suggested and more rarely, confirmed (26-30). In general, ERVs are proposed to contribute to shaping the genome of the host and influencing gene expression by promoting chromosomal rearrangements through homologous recombination between distant loci (31) and by directly influencing gene expression (32).

Several exogenous retroviruses such as; feline leukaemia virus in cats, mouse mammary tumour virus (MMTV) and murine leukaemia viruses (MuLVs) in mice, and avian leukaemia viruses in chickens have endogenous counterparts in the germ line of their host species (28). The only retroviral genera for which clear endogenous counterparts have not been identified are lentiviruses and deltaretroviruses (33).

The draft human genome sequence shows that 8% of human DNA sequences represent human endogenous retrovirus (HERV) and related retrotransposons (34).

### **1.2 Retroviral Oncogenesis**

Retroviruses have been valuable tools in the study of oncogenesis. They were first associated with malignant diseases in animals approximately one century ago. In 1908 Ellerman and Bang observed that avian erythroleukaemia could be transmitted by cell-free filtrates and were caused by a virus (1). Years later, Rous reported that

extracts from certain poultry sarcomas were infectious and could induce these tumours when inoculated into healthy poultry (2). These two pioneering events marked the beginning of oncoretrovirus research. Since the 1960s retroviruses have been shown to cause leukaemia, lymphoma, and other forms of cancer in a wide variety of vertebrate animals.

### **1.2.1 Retroviral transduction of cellular genes**

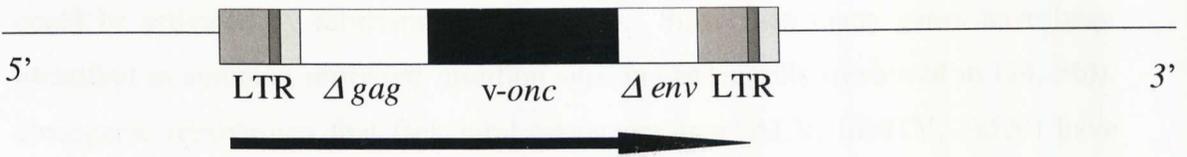
During replication, retroviruses can acquire cellular sequences through aberrant recombination events. Such recombinant retroviruses (Figure 1-4) are usually defective. They have lost parts of the viral genome and require the aid of intact retroviruses via trans-complementation for replication or integration. If the acquired cellular gene is a proto-oncogene (a gene involved in such functions as cellular homeostasis, growth control, and signalling cascades), its overexpression under the control of strong viral promoters can lead to malignant transformation. The incorporation of a proto-oncogene within the retroviral genome can occur when transcription from the 5'LTR reads through the packaging signal of viral RNA at the beginning of *gag* and onwards, often spliced, into cellular sequences.

Such viruses with incorporated viral oncogenes (*v-onc*) derived from cellular proto-oncogenes (*c-onc*) have been called 'acute transforming viruses' because they usually induce polyclonal tumours rapidly after a short incubation period (2-3 weeks) and are capable of transforming cells in culture.

Transduction of an oncogene has not been observed in human cancer, however it has proven to be very useful in experimental settings for studying the action of oncogenes in cancer (12).

### **1.2.2 Retroviral cis-insertional activation of cellular genes**

Insertion of a retrovirus into or nearby a cellular proto-oncogene may lead to activation of this gene by promoters or enhancer elements in the viral LTR. Simple retroviruses typically act oncogenically by *cis*-activation of host genes. In the large majority of B-cell lymphomas in ALV-infected chickens, for example, the cellular proto-oncogene *c-myc* has been found to be affected by this mechanism of provirus integration. Hayward and colleagues demonstrated that the *c-myc* proto-oncogene



**Figure 1-4. Retroviral transformation by transduction of an oncogene.**

Transduction of an oncogene in which the provirus contains a viral oncogene (*v-**onc***) originally derived from the host cell. Parts of the viral genome are deleted or truncated, and the virus relies on help from intact viruses for replication or integration (coinfection with a helper virus).

could be activated by retroviral insertion (35). Since then many genes have been identified as common retroviral insertion sites in cancer cells (reviewed in (24, 36)). Oncogenic retroviruses that lack viral oncogenes (e.g. ALV, MMTV, FeLV) have been called 'slow-transforming retroviruses' because they usually induce mono- or oligo- clonal tumours after a latent period of 3-12 months, and are usually unable to transform cells *in vitro* (24). Instead, the predominant mechanism by which they induce tumours is a process called insertional mutagenesis.

Insertion of a provirus within or adjacent to cellular proto-oncogenes alters the normal organisation of the locus and also introduces strong promoter and enhancer regions into the locus. The major ways by which a proto-oncogene becomes activated by the insertion of a retrovirus are referred to as promoter insertion and enhancer activation.

### ***Viral promoter insertion***

Gene activation by promoter insertion requires the integration of a provirus upstream of the gene in the same transcriptional orientation. Transcription of the downstream endogenous gene is initiated from the viral promoter sequences in either the 5' or the 3'LTR, thereby replacing the function of the normal promoter. Promoter insertions where the transcript is initiated from the 5'LTR usually yield fusion transcripts as a result of transcript splicing using the splice donor (SD) or cryptic SD sites within the viral DNA toward the second exon or, in case a cryptic splice acceptor is present, the first exon of the endogenous gene (37, 38) (Figure 1-5b). Examples of 5'LTR insertion are the AVL integrations in the *c-erbB*/EGF-receptor gene in chicken-erythroleukaemias (39, 40), Mo-MuLV integrations in *c-myb* found in mouse lymphosarcomas (41, 42) and Mo-MuLV integrations activating the *c-lck* gene in some murine T cell lymphomas (37). 3'LTR promoter insertion is similar to 5' promotion except that transcription is initiated at the 3' promoter of the LTR (Figure 1-5c). The most studied example represents AVL proviral insertions occurring predominantly in the first intron of the *c-myc* gene in avian bursal lymphomas (20, 43, 44). In many tumours carrying promoter insertions the 5'LTR is deleted suggesting that the removal of the 5'LTR promotes the transcription of the endogenous gene by the 3'LTR (45).

## ***Viral enhancer activation***

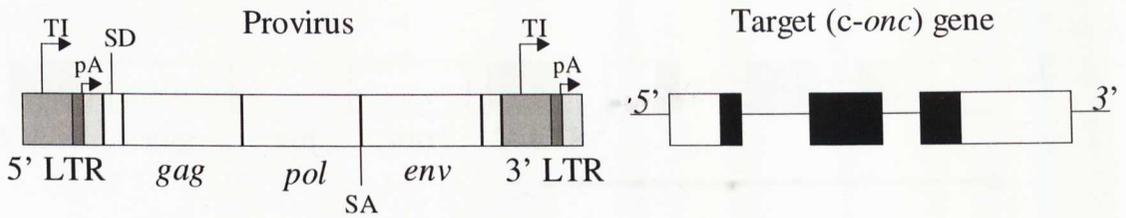
Proviral insertions can also stimulate transcription of genes from their authentic promoters through the activity of the enhancers sequences contained within the LTRs. Enhancer insertion involves integration of a provirus either at the 5' end of a gene in the inverse transcriptional orientation (Figure 1-5d), or at the 3' end in the same transcriptional orientation (Figure 1-5e). Typical examples are the activation of *c-myc* by ALV in some bursal lymphomas (46), the activation of *int* genes by MMTV in murine mammary carcinomas (47, 48), and the activation of *pim-1* and *c-myc* in MuLV-induced T-cell lymphomas (49). The group of 3' enhancer insertions also includes proviral insertions within the 3' untranslated region of the gene, in the same transcriptional orientation, in which the mRNA transcription is terminated at the poly (A), polyadenylation site present in the 5'LTR (Figure 1-5f). The result of these insertions is often the stabilisation of mRNA transcripts by the removal of sequences in the 3' untranslated region, which contain negative regulatory elements, such as mRNA-destabilising motifs. Examples include *pim-1* and *pim-2* in MuLV-induced T-cell lymphomas (50-52), *N-myc* in MuLV-induced T-cell lymphomas (53).

Viral enhancer activation is the most frequent mechanism of gene activation by insertional mutagenesis, due to the fact that activation by enhancement allows more flexibility with respect to both proviral orientation, and the distance between proviral integration and the target gene.

## ***Truncation of cellular gene product***

The integration of a provirus into the coding sequence of a gene can have effects on the resulting protein, inactivating or mutating it in such a way that an aberrant gene product is produced. Transcript truncation by proviral insertions may also produce a functional protein with oncogenic properties by deletion of either NH<sub>2</sub>- or COOH-terminal domains (54). This may happen in combination with enhancer or promoter activation. Protein truncation from proviral insertions often occurs by transcription termination using the viral polyadenylation signal. If the promoter insertion occurs in the first intron of a gene (Figure 1-5g), transcripts arising from the endogenous promoter may be truncated by the viral poly-A, while the provirus promotes transcription of the downstream exons. If the start codon is located in the second exon as is the case for *c-myc* or *c-fms* a full-length protein will be produced (55, 56),

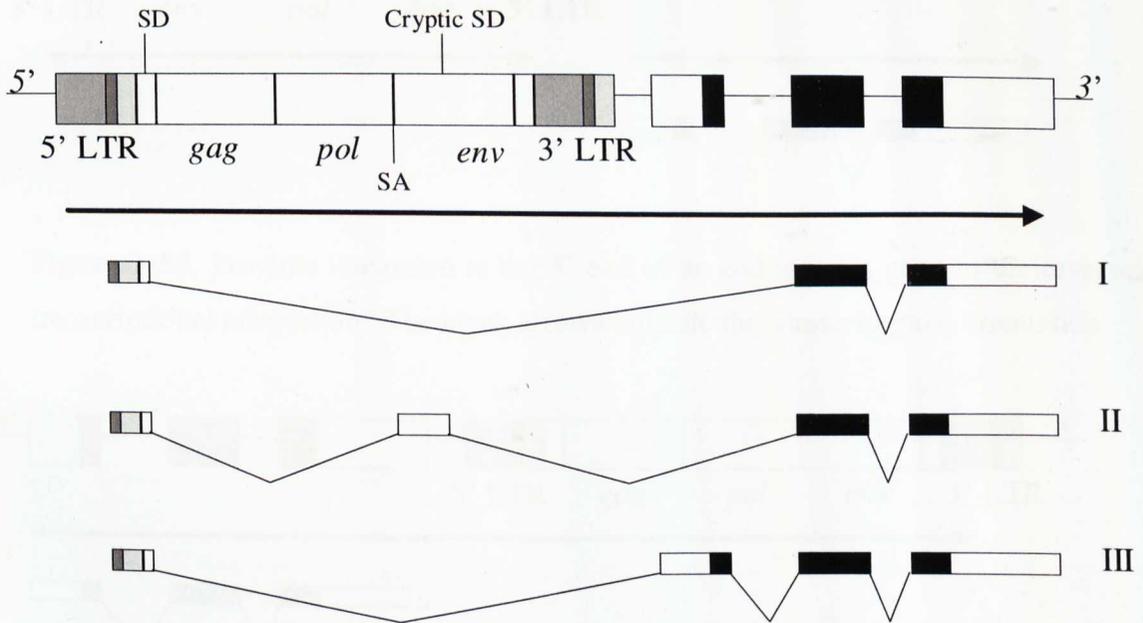
## Figure1-5. Mechanisms of retroviral insertional activation



**Figure 1-5a.** Proviral enhancer and promoter sequences in the U3 region of the LTR, and the poly(A) signal in R are utilised to deregulate endogenous genes. Structural features of a provirus, and an endogenous gene (*c-onc*), which in this case consists of 3 coding exons (black boxes) as well as 5' and 3' untranslated regions (open boxes). Abbreviations: TI, transcriptional initiation; pA, polyadenylation signal; SD, splice donor; SA, splice acceptor.

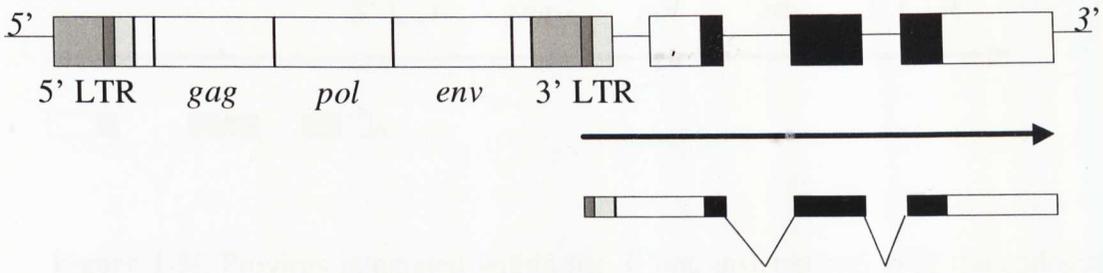
### Activation by promoter insertion

#### 5'LTR promotion



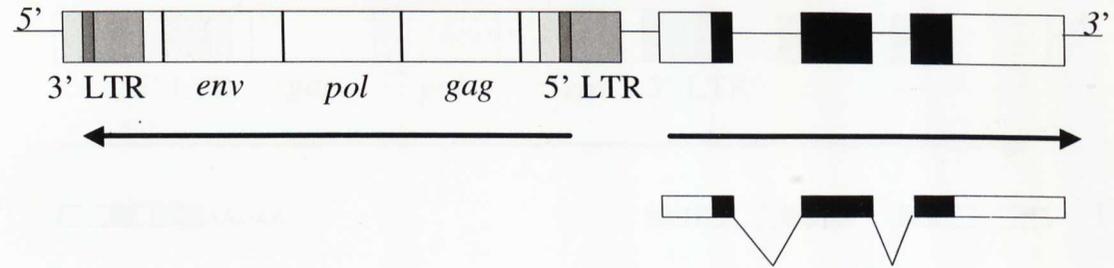
**Figure 1-5b.** Promoter insertion in which the 5'LTR promoter is used to drive transcription of a downstream endogenous gene. In this case, the transcription results in the formation of fusion transcripts containing both viral and cellular sequences, due to a read through at the 3' LTR polyadenylation site and subsequent splicing using the subgenomic mRNA splice donor (I), a cryptic splice donor (II), or a cryptic splice acceptor site (III). The black arrow indicates the transcriptional orientation.

### 3'LTR promotion

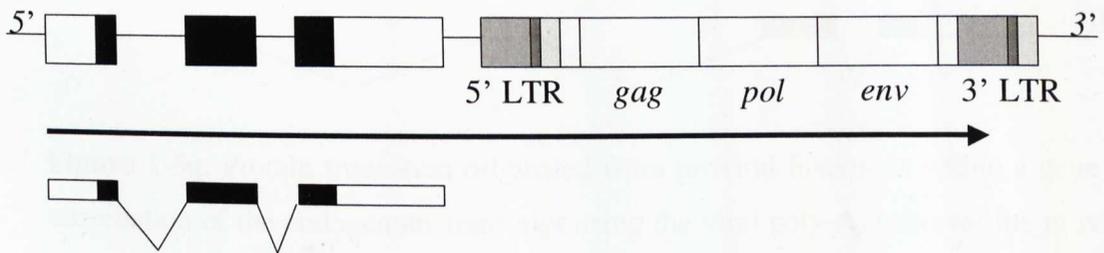


**Figure 1-5c.** Promoter insertion in which the 3' LTR promoter is used to drive transcription of a downstream endogenous gene. The black arrow indicates the transcriptional orientation.

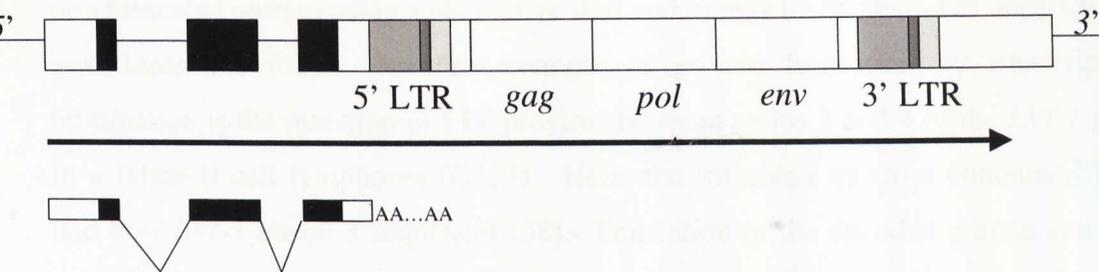
### Activation by enhancer insertion



**Figure 1-5d.** Provirus integrated at the 5' end of an endogenous gene in the reverse transcriptional orientation. The black arrows indicate the transcriptional orientation.

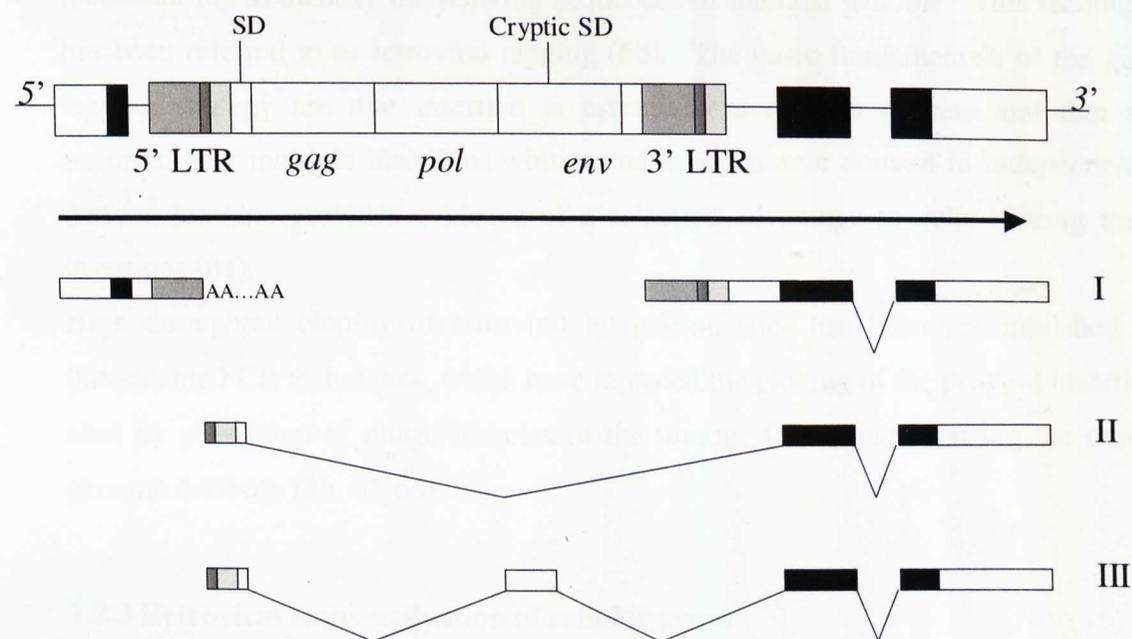


**Figure 1-5e.** Provirus integrated at the 3' end of an endogenous gene in the same transcriptional orientation. The black arrow indicates the transcriptional orientation.



**Figure 1-5f.** Provirus integrated within the 3' untranslated region of the endogenous gene and in the same transcriptional orientation. The transcript is cleaved at the polyadenylation site of the 5' LTR. The black arrow indicates the transcriptional orientation.

### Protein truncation by transcription termination at viral poly-A site



**Figure 1-5g.** Protein truncation originated from proviral insertions within a gene via termination of the endogenous transcript using the viral poly-A, however the provirus promotes transcription of the downstream exons (I). Protein truncation originated by promoter insertion in which the 5' LTR promoter is used and transcription results in the formation of fusion transcripts, due to a read through at the 3' LTR polyadenylation site and subsequent splicing using the subgenomic mRNA splice donor (II), or cryptic splice donor (III) sites. The black arrow indicates the transcriptional orientation.

or a truncated protein using an alternative start codon may be expressed as occurs for *c-myb* insertions (57). Another example of protein truncation by transcription termination is the insertion of FIV provirus between exons 3 and 4 of the *LVI-1* gene in a feline B-cell lymphoma (Q254). Here, the truncated transcript contains 3'LTR and host *LVI-1* intron 3 sequences (58). Truncation of the encoded protein can also occur by proviral insertions within a gene in the opposite transcriptional orientation due to the presence of a cryptic poly-A site in the virus (Figure 1-5h). An example is the insertion of MMTV into *int-6* gene in murine mammary tumours (59).

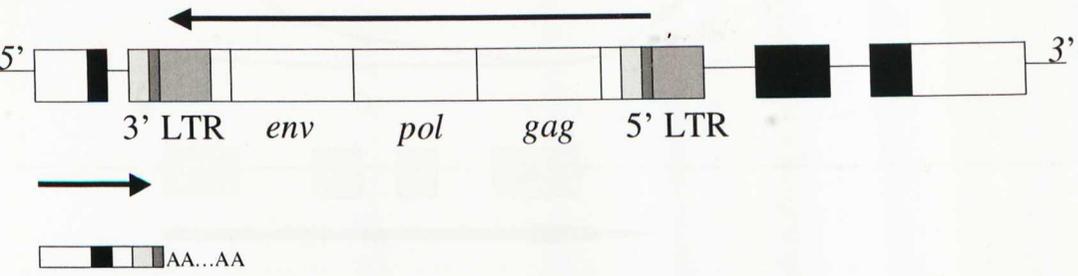
Somatic and clonal integration of retroviruses in the genome can enhance proto-oncogene expression or inactivate tumour suppressor genes. A cell that has such integrations acquires a growth advantage, and is then clonally selected to grow and may result in tumour formation. In these cases, the proviral DNA can be utilized as a molecular tag to identify the flanking sequences of the host genome. This technique has been referred to as retroviral tagging (60). The basic fundamentals of the gene tagging strategy are that insertion is essentially a random process and that the occurrence of multiple insertions within a narrow genomic domain in independently derived tumours provides evidence of a selective advantage to cells bearing these insertions (61).

High throughput cloning of retroviral integration sites has been accomplished by introducing PCR techniques, which have replaced the cloning of the proviral insertion sites by generation of phage libraries of the tumour DNA, and by using the mouse genome database (38, 62-64).

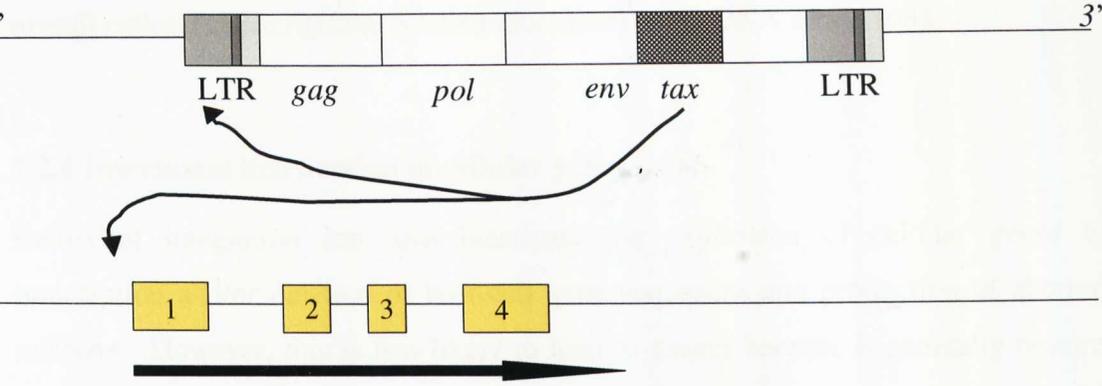
### **1.2.3 Retroviral trans-activation of cellular genes**

Some complex retroviruses encode viral proteins (i.e. not acquired through recombination with cellular sequences) that can act oncogenically. Examples are the viruses of the deltaretrovirus genera and possibly also the simple retroviruses JSRV and ENTV (65). Deltaretroviruses such as BLV and HTLV encode a protein Tax that acts as a transactivator of transcription (66, 67) (Figure 1-6). It primarily activates and regulates the transcription of viral genes by activation of the viral LTR, but also alters the transcription pattern of cellular genes, including proto-oncogenes. This can lead to a loss of growth control and malignant transformation. A number of host genes are up-regulated (and some down-regulated) by Tax, and its effect on the

## Protein truncation by transcription termination at viral cryptic poly-A site



**Figure 1-5h.** Protein truncation originated from proviral insertions within a gene in the opposite transcriptional orientation by the presence of a cryptic poly-A site in the virus.



**Figure 1-6. Retroviral *trans*-activation of an oncogene.** The provirus encodes viral proteins (eg. HTLV-Tax) that are not acquired through recombination with cellular sequences and can activate not only the viral LTR, but also a set of cellular genes, including proto-oncogenes. Exons of the proto-oncogene are indicated by numbered boxes.

overall cellular transcriptome is being elucidated using DNA arrays (68).

#### 1.2.4 Insertional inactivation of cellular genes

Retroviral integration can also inactivate the expression of cellular genes by interruption and/or deletion of host-cell gene sequences and production of aberrant mRNAs. However, this is less likely to lead to cancer because it generally requires the inactivation of both alleles (either by two integrations or by one integration followed by transcriptional silencing or by loss of the second allele) in order to detect a phenotype for most genes. Examples of the oncogenic effects of retroviral insertion through inactivation of the gene activity were seen in leukaemogenesis induced by the Friend virus complex that is composed of the Friend murine leukaemia virus (F-MuLV) and by the Friend spleen focus-forming virus (SFFV) (Table 1-1). Erythroleukaemic cell lines derived from spleens from mice infected with F-MuLV were shown to have integrated proviruses into the *p53* tumour suppressor gene with resultant loss of gene expression (69-71). It initially seems unlikely that gene inactivation would arise by two insertional events into both alleles, but independent insertional mutations that knock out both alleles of the *p53* gene were reported for one tumour cell line (72). In some tumours, proviral integration is associated with deletions of *p53*-coding sequences, while in others, integration into intronic sequences results in the synthesis of aberrant *p53* mRNAs. Therefore, *p53* gene is inactivated in almost all erythroleukaemia cell lines induced by polycythemia- or anaemia-inducing strains of Friend virus (FV-P/FV-A) or F-MuLV strain due to retroviral insertions, deletions, and point mutations (reviewed in (73, 74)). Moreover, in one erythroleukaemia cell line, CB3, induced by both FV-P and F-MuLV strains of Friend virus the *Fli-2* (Friend leukaemia integration 2) locus was identified as a common site for retroviral integration in which the proviral insertion was associated with the loss of the second allele, resulting in the inactivation of the erythroid cell- and megakaryocyte-specific gene *p45<sup>Nfe2</sup>* (75). The transcription factor Nfe2 (nuclear factor erythroid 2) plays a critical role in the regulation of erythroid cell specific gene expression (76). Its large subunit, *p45<sup>Nfe2</sup>* has been found to be tissue specific and with expression restricted to erythroid cells, megakaryocytes, and mast cells (77).

In a similar way, the *Neurofibromatosis-1* (*NFI*) tumour suppressor gene, which encodes the neurofibromin GTP-ase activating protein, has also been shown to be

**Table 1-1. Genes inactivated by retroviral integrations in cancer.**

Gene/Locus	Chromosome		Insertional mutagen <sup>a</sup>	Gene product/Function	Tumour	References
	Mouse	Human				
p53	11	17p13.1	F-MuLV/ SFFV/ CasBrE-MuLV/ M- MuLV	P53/ Transcription factor	Erythroleukaemia/ non-B and non-T leukaemia/ T and B lymphoma	(70, 72, 82)d}d}d}
NF1/Evi-2	11	17q11.2	BXH2/ CasBrE-MuLV	GTPase activating/ GTP/GDP exchange factors	Myeloid leukaemia/ non-B and non-T leukaemia	(78, 80, 83)d}d}d}
Nfe2/Fli-2	15	12q13	F-MuLV	NF-E2/ Transcription factor	Erythroleukaemia	(75, 84)d}d}d}
Nfatc3	8	16q22.1	SL3-3-MuLV	Nuclear factor of activated T cell	T-cell lymphomas	(81)d}d}d}

Key to table: <sup>a</sup> Insertional mutagen: F-MuLV, Friend murine leukaemia virus; SFFV, Friend spleen focus-forming virus; M-MuLV, Moloney murine leukaemia virus; CasBrE-MuLV, Casitas Brain Ectropic -murine leukaemia virus; BXH2 is a recombinant inbred strain derived from a cross between C57BL/6J and C3H/HeJ mice.

inactivated by provirus integration of MuLV and a murine AIDS virus-related endogenous virus at a common viral integration site, *Evi-2* (ecotropic viral integration site 2) in myeloid leukaemia of BXH-2 mice (78, 79). This integration resulted in the production of truncated *Nf-1* transcripts and no stable, full length neurofibromin, which is involved in the development of neurofibrosarcomas and myeloid leukaemia (80). As with *p53*, inactivation of the second allele can occur by integration of a second provirus or by a nonviral mechanism.

A recent study identified *NFATc3* (nuclear factor of activated T cell 3) as tumour suppressor for the development of murine T-cell lymphoma induced by the murine lymphomagenic retrovirus SL3-3. Furthermore, only one allele of the *Nfatc3* gene is targeted by T-cell lymphomagenic retrovirus SL3-3 MuLV so the repression of NFATc3 expression can be associated with additional transcription repression effects at either the provirus targeted or untargeted allele (81).

### **1.3 Cellular genes involved in cancer**

Cancer is a genetic disease due in part to the accumulation of mutations, by which the tumour cells become insensitive to controls by the local cellular environment and the whole organism. An attempt to summarise the various traits of cancerous cells has been made by Hanahan and Weinberg. They suggest that most, if not all cancers acquire the same set of functional capabilities during their development, namely self sufficiency in growth signals, insensitivity to growth inhibitory signals, tissue invasion and metastasis, limitless replicative potential, sustained angiogenesis, and evasion of programmed cell death (apoptosis) (85). This model predicts that each individual tumour will exhibit as many mutation events, in as many genes, as are necessary to acquire these six traits. Tumour growth by clonal expansion of a single somatic cell is generally a slow and multistep process, in which an estimated three to six independent genetic steps are necessary to complete tumorigenesis, and acquire these six characteristics (reviewed in (86)). Therefore, the risk of cancer development depends not only on mutations initiating tumorigenesis, but also on subsequent mutations driving tumour progression (87). Unlike diseases such as cystic fibrosis or muscular dystrophy where mutations in one gene can cause disease, no single gene defect is sufficient to cause cancer. So, an invasive cancer develops only when several genes are defective (88). Mutations occur randomly throughout the

genome; among these are mutations in genes that function in normal cells to maintain the stability of the genome and to guarantee that it is faithfully copied and transmitted to progeny during each cell division (89).

It has been established that environmental factors, such as tobacco, diet, sunlight, radiation, pollution, and stress can play an important role in the etiology of cancer (reviewed in (90)). In addition, oncogenic viruses have been related to human as well as animal cancers. Examples of human cancer viruses include the papilloma viruses in cervical cancers (91) and the Epstein-Barr virus in Burkitt lymphoma, in nasopharyngeal cancer (92) and in B cell lymphomas (reviewed in (93)). Moreover, it has been suggested that multiple, low penetrance genes segregating in the human population confer cancer susceptibility and resistance to environmental carcinogens (94, 95). Therefore, these low penetrance genes determine cancer predisposition by affecting processes such as carcinogen metabolism, DNA repair efficiency, inflammation, and the immune response to provide relative degrees of tumour resistance (96)

Although environmental and other non-genetic factors have roles in many stages of tumorigenesis, it is widely accepted that cancer-susceptibility genes play a significant role. Most of these genes normally function to restrict cancer development and are thereby defined as tumour suppressor genes (TSGs) (97, 98).

### 1.3.1 Oncogenes

Oncogenes are genes that undergo gain-of-function alterations in cancer cells. The study of oncogenes began with experiments on RSV, one of the acutely transforming retroviruses (reviewed in (99)). These viruses were shown to possess, in addition to their normal genetic content, extraneous DNA sequences, which were responsible for acute oncogenesis. In fact, RSV is unique in having all genes necessary for viral replication (*gag*, *pol*, *env*) in addition to the oncogenic *v-src*; all other acute transforming retroviruses lack one or more of the structural genes, rendering them replication incompetent. The homologous cellular genes, known as *c-onc*'s are called proto-oncogenes because they have served as targets for genetic transduction by retroviruses. Proto-oncogenes were identified initially by the homology of their nucleotide sequences to retroviral oncogenes (reviewed in (100)). Some retroviral oncogenes have undergone mutation in their coding regions, altering the biochemical

properties of the gene product. Others, however, harbour no detectable alterations and are oncogenic by virtue of the fact that they are under the transcriptional control of the constitutively active viral LTRs. Furthermore, oncogenes are generally considered as dominant genes because an activating somatic mutation in one allele of an oncogene is sufficient to confer a selective growth advantage on the cell.

The high degree of evolutionary conservation of the proto-oncogenes across all multicellular organisms points to basic functions that these genes might have for the cell. Proto-oncogenes generally encode a class of proteins that are selectively active only when the proper regulatory signals allow them to be activated (Table 1-2). Their products are elements of cell cycle positive control pathways, including growth-factor receptors, signal transduction proteins, and transcriptional regulators. Other proto-oncogene products function to negatively regulate apoptotic pathways. Disregulation of these gene products will therefore result in insensitivity to antigrowth signals and evasion of apoptosis, two of the hallmarks of cancer mentioned previously (85).

There are in fact a number of mechanisms where by proto-oncogenes can be activated aberrantly within normal cells, resulting in transformation. Such include retroviral transduction or integration as discussed previously, genetic modification such as point mutation or insertion mutation, gene amplification, chromosomal rearrangement or epigenetic modification leading to over-expression.

The products of the proto-oncogenes can act in a number of different ways, which include phosphorylation of proteins at serine, threonine, or tyrosine residues and in this case the role of the proto-oncogene product can be induction of phosphorylation (as with some growth factors such as PDGF (*c-sis* gene product) (101)), or self-catalysis (as with the receptors for some growth factors such as *c-erbB* gene product (102)); transmission of signals by GTP-binding proteins, as exemplified by the GTPases (products of the *Ras* genes) (103); and the control of transcription from DNA as in the case of the transcription factors such as *Myc* gene (104). Even with differences in their normal roles, these genes all contribute to unregulated cell division if they are present in a mutant (oncogenic) form. The mutant oncoproteins are uncoupled from the regulatory pathway that should be controlling their activation, leading to continuous unregulated expression or function of the oncoproteins.

The isolation of the majority of proto-oncogenes to date has been done not only by molecular cloning of genes, which have been activated by proviral insertion, but also

**Table 1-2. Some well characterised proto-oncogenes.**

Proto-oncogene nomenclature	Properties of protein		Alteration in cancer	Tumour	Reference
	Function	Location			
<i>c-sis</i>	Secreted growth factor (PDGF B chain)	Extracellular	Overexpression	Astrocytoma, osteosarcoma	(105)d}d}d}
<i>c-erbB</i>	Growth factor receptor (Receptor tyrosine kinase)	Transmembrane	Amplification or overexpression	Epidermal cancer, glioblastoma, erythroblastoma	(39)d}d}d}
<i>c-bcl2</i>	Apoptosis inhibitor (upstream inhibitor of caspase cascade)	Cytoplasm	Chromosomal translocation or mutation	Leukaemia, lymphoma, small lung carcinoma cell lines, carcinomas, colorectal adenomas, neuroblastoma	(106)d}d}d}
<i>c-abl</i>	Intracellular signal transducer (protein tyrosine kinase)	Cytoplasm	Chromosomal translocation	Chronic myeloid leukaemia	(107)d}d}d}
<i>c-ras</i> (c-Ha-ras)	Intracellular signal transducer (GTP/GDP-binding protein: GTPase)	Cytoplasm	Mutation	Bladder cancer, pancreatic carcinoma, NSCLC, colon cancer, acute myeloid leukaemia	(108)d}d}d}
<i>c-myc</i>	Nuclear transcription regulator (Transcription factor)	Nucleus	Chromosomal translocation <sup>a</sup> , proviral insertion or gene amplification <sup>b</sup>	<sup>a</sup> Burkitts lymphoma, <sup>b</sup> SCLC, <sup>b</sup> neuroblastoma, <sup>b</sup> retinoblastomas, <sup>b</sup> astrocytomas, <sup>b</sup> gliomas,	(109)d}d}d)}

Key to table: NSCLC, Non-small cell lung carcinoma; SCLC, small cell lung carcinoma.

by identification of genes which have been transduced by a retrovirus (reviewed in (24)).

### 1.3.2 Tumour suppressor genes: History and genetic characterisation

Failure of growth inhibition is one of the hallmarks of cancer (85). The proteins responsible for arresting cellular proliferation are the products of TSGs. The study of TSGs began with experiments involving somatic cell fusion (110). Harris and colleagues observed that the growth of murine cancer cells could be suppressed by fusion with non-tumourigenic cells. Such somatic cell hybrids tended to revert to the malignant phenotype over time; an event that coincided with the loss of chromosomes from the cell. Subsequent experiments using chromosome segregation showed that the transfer of single chromosome carrying a wild type gene suppresses tumour formation of malignant cells. Together, these studies pointed to the existence of genes that could suppress tumourigenicity (110, 111). Indeed, these observations supported the view that malignancy was a recessive trait, and that a particular chromosome or a small number of genes might be responsible for the tumour suppression.

At the same time, Knudson was conducting epidemiological investigations into the childhood cancer retinoblastoma (112) and proposed that two hits or mutagenic events were necessary for tumour development (98, 113-116). In inherited retinoblastoma, he hypothesised that susceptibility was due to a mutation present in the germ line, but that neoplastic disease required a second mutation which could occur early in life, leading to strong penetrance and the characteristic dominant autosomal inheritance pattern. By contrast, the sporadic form of retinoblastoma, which requires two de novo mutational events to inactivate both alleles, is much rarer.

However, this two-hit paradigm of TSG mutation in which the inactivation of both TSG alleles was required to develop a cancer phenotype may not be universal. There is now considerable evidence suggesting that loss of only one allele of a TSG might contribute to the tumourigenesis, a process known as haploinsufficiency. This has been shown in mice carrying single, inactivated alleles of the genes encoding *p27<sup>Kip1</sup>* (117), *p53* (118), *TGF- $\beta$ 1* (119), and *Dmp1* (cyclin D-binding Myb-like protein) (120). In each case, the genetically deficient mice are predisposed to tumour development, but the resulting tumours frequently retain a functional wild-type allele.

Moreover, the  $p27^{Kip1}$  and *Dmp1* mutant mice show at least slightly more severe tumour susceptibility phenotypes in the homozygous state than in the heterozygous state for most tumours types. These observations have led to the suggestion that haploinsufficiency effects can be strong or weak (121, 122). Whereas a strongly haploinsufficient effect results in tumour formation in the absence of loss of the wild type allele, in a weaker haploinsufficient effect loss of the wild type allele is required to complete tumourigenesis, a process known as loss of heterozygosity (LOH). LOH can occur by loss of the whole chromosome carrying the wild type allele, by mitotic recombination, large or small chromosomal deletions, or point mutation of the wild type allele (123).

In addition to family linkage analysis, as implied by the two hit model of Knudson, a search for consistent LOH, which occurs in particular tumour types (and frequently in familial cancers) is used as a tool for narrowing the region of interest by deletion mapping, and finally for searching the intact homologous chromosomal segment for mutated genes whose functions can be demonstrated to protect against cancer development. However, this approach will exclude identification of genes that either require two functional copies to adequately suppress tumour cell growth and thus will have only one inactivated allele in tumours, or when the second allele of the target gene is transcriptionally silenced through promoter methylation of its cytosine phosphate guanosine rich regions (CpG islands). Indeed, there are many regions of LOH identified in human cancer for which no TSG has been identified. One such site is the long arm of human chromosome 7 where the *DMP1* resides (124). In fact, it is the linkage method and the occasionally observed cytogenetic changes that have been most fruitful in identifying TSGs such as *WT1* (*Wilms Tumour*), *BRCA-1* and *-2*, *APC* (*Adenomatous Polyposis Coli*) *NF1*, *VHL* (*von Hippel-Lindau*), and *BLM* (*Bloom mutated*).

### ***Functional classification of tumour suppressor genes***

According to the functional classification proposed by Kinzler and Vogelstein (97, 98), the TSGs genes belong to one of the three classes: gatekeepers, caretakers and landscapers (Table 1-3). Defects in gatekeeper genes lead to abnormal cellular proliferation, differentiation, and apoptosis. Caretakers, by contrast, suppress tumourigenesis indirectly (125, 126). They may function in maintaining the genomic

**Table 1-3. Representative TSGs.**

TSG	Class of		Chromosome	Function	Cancer syndrome/Human tumour types
	TSG	Mouse			
<i>RB</i> <sup>a</sup>	Gatekeeper	14	13q14.2	Cell cycle block through transcriptional repression (inhibits progression from G1 to S phase)	Retinoblastoma/Retinoblastoma, osteosarcoma, bladder cancer
<i>p53</i> <sup>a</sup>	Gatekeeper	11	17p13.1	Transcription factor (cell cycle block through transcriptional activation), response to DNA damage and stress; apoptosis	Li-Fraumeni Syndrome/Soft tissue sarcoma, osteosarcoma, colorectal cancer, breast cancer, thyroid carcinoma, hepato-cellular carcinoma, nasopharyngeal carcinoma
<i>p16INK4a</i> <sup>a</sup>	Gatekeeper	4	9p21	Cell cycle G1 control, inhibitor of cdk4 kinase, antagonist of D1 cyclin (RB activation), stabiliser of p53	Melanoma, nervous system tumours, pancreatic cancer, orolaryngeal cancer, cutaneous malignant melanoma, bladder cancer
<i>BRCA-1</i> <sup>b</sup>	Caretaker	11	17q21	Involved in DNA repair, transcriptional regulator	HBOC/Breast, ovarian and other cancers
<i>NF1</i> <sup>c</sup>	Gatekeeper	11	17q11.2	Negative regulator of RAS-GAP activity, GTPase activating protein for RAS	Neurofibromatosis type 1/Neurofibrosarcoma, AML, brain tumours, sarcomas, gliomas

**Key to table:** a. TSG involved in apoptosis and in the cell cycle; b. TSG involved in DNA damage repair; c. TSG involved in cell signalling and differentiation; HBOC, hereditary breast and ovarian cancer syndrome; HNPCC, hereditary non-polyposis colorectal cancer; FAP, familial adenomatous polyposis; NBCCS, nevoid basal cell carcinoma syndrome.

integrity of the cell and regulate DNA repair mechanisms or chromosome segregation (126). Caretaker defects lead to genetic instabilities that contribute to the accumulation of mutations in other genes that directly affect cell proliferation and survival (127). An increasing number of TSGs are included in this category. The two highly penetrant breast cancer-associated genes, *BRCA-1* and *BRCA-2* (hereditary breast cancer) (128) have been placed in this class of TSGs. There are fewer examples of landscapers which are defined as genes that regulate the stromal environment (129). The landscaper genes may contribute to the viability and metastatic potential of tumour cells by, for example, mediating paracrine growth stimulation. In fact, the 'caretaker and gatekeeper' model proposed by Kinzler and Vogelstein suggest that TSGs of both types required two inactivating mutation events. However, their model did not predict any increased genetic instability resulting from loss of a single allele of a caretaker gene (130).

Two of the most frequent targets of mutation during tumourigenesis are the gatekeepers: *retinoblastoma (RB)* TSG and *p53* TSG (reviewed in (131, 132)). Loss of RB protein results in deregulated cell proliferation and apoptosis, whereas loss of p53 desensitises cells to checkpoint signals, including apoptosis. Therefore, to recognize their role in the development of most human cancer, it is important to understand the cell cycle biology. The cell cycling process is carefully regulated and responds to the specific needs of a certain tissue or cell type. In normal, physiological conditions there is a balance between cell death (programmed cell death or apoptosis) and proliferation (cell division) producing a steady state (133). Conventionally, the cell cycle has been divided into four phases: S phase, DNA replication or synthesis; M phase, mitosis; G1 and G2, gap phases when the cell is metabolically active and preparing for DNA replication and mitosis, respectively. Hartwell and colleagues identified a number of genes whose products are required for the progression of the cell cycle (reviewed in (134)). These proteins are now known as the cyclins and the cyclin dependent kinases (cdks). Although initially identified and characterised in the yeast *Saccharomyces cerevisia*, these molecules are remarkably well conserved and appear to function similarly in all eukaryote organisms.

### ***Retinoblastoma* TSG**

The Rb protein was the first TSG product to be recognised (135). Although germline mutations in *Rb* are associated with rare cancers such as childhood retinoblastomas

and osteosarcomas (136), somatic mutations in *RB* also contribute to the development of common human tumours, such as lung and bladder carcinomas. *Rb* is a member of a gene family that includes *p107* and *p130*, which collectively co-repress the transcription of genes that regulate cell cycle progression, apoptosis, and differentiation. The Rb family proteins exert much of their growth suppressive control during the first gap phase (G1) of the cell cycle and entry into DNA synthesis (S phase). *In vivo* phosphorylation of Rb is cell cycle dependent. The hypophosphorylated form is present in G0/G1 where it binds and sequesters proteins such as E2F. Phosphorylation in late G1 allows E2F to be released and activate the transcription of genes required for DNA synthesis (S phase) (137-139). Phosphorylation of Rb is mediated by cdks in complex with their cognate cyclins: cyclin D (D1, D2 and D3)-cdk4/6 and cyclin E-cdk2 (132). Inhibitors of these kinase complexes include the INK4 proteins ( $p16^{\text{INK4a}}$ ,  $p15^{\text{INK4b}}$ ,  $p18^{\text{INK4c}}$ ,  $p19^{\text{INK4d}}$ ) which block the ability of cyclin D-cdk4/6 to phosphorylate and regulate Rb's growth suppressive functions (140-143).

Functional inactivation of the RB protein in the absence of an RB mutation can also lead to tumourigenesis indirectly through alterations in upstream regulatory proteins such as cyclin D1 overexpression in breast carcinoma, head and neck tumours, and mantle cell lymphomas; cdk4 overexpression in glioblastomas; and  $p16^{\text{INK4a}}$  loss in melanoma, pancreatic carcinoma and T cell lymphoma (144). Viral oncoproteins from DNA tumour viruses, such as adenovirus E1A, simian vacuolating virus 40 large tumour (SV40 T) antigen and the E7 proteins of human papillomaviruses interact with the hypophosphorylated form of Rb and interrupt its function, thus facilitating viral replication as well as promoting tumorigenesis if replication fails.

### **P53 TSG**

Another well-known TSG is the *p53* gene, that was first isolated in immunoprecipitates of large tumour antigen (large T) from SV40-transformed rodent cells (145, 146). It was originally thought to be an oncogene, due the high levels of expression of the p53 protein (so called because of its molecular weight of 53 kDa) in transformed mouse cells, but not in normal mouse cells (147). In fact, reintroduction of wt p53 into p53 deficient cells results in growth inhibition; with cells exhibiting growth arrest at or near the G1/S phase transition. Therefore, the wt p53 acts negatively to block transformation, (148), so it functions as a classical TSG.

p53 TSG is mutated in more than 50% of human cancers, and mutation in other genes that affect p53 function occur in many tumours that retain a normal p53 gene (144). Mutations of p53 underlie the familial Li-Fraumeni cancer susceptibility syndrome (149), which predisposes to a range of tumour types including breast cancer, sarcomas, brain tumours and leukaemia. Unlike RB, mutation at one allele may be sufficient to give rise to the altered cell phenotype (150). Moreover, p53 mutations may result in genomic instability, which then permits accumulation of the multiple mutations required for tumour development (127).

The p53 phosphoprotein is normally unstable and rapidly degraded. However, cellular stress, including irradiation, hypoxia, drug-induced genotoxic damage, and even oncogene activation are responsible for the activation and stabilization of p53 (151). Then, the activated p53 restricts cellular growth by inducing senescence, cell cycle arrest (at G1 and/or G2 phase) to allow repair to take place or apoptosis if the damage is excessive (152). For these reasons, p53 has been called as 'guardian of the genome' (153). The p53 protein functions as a transcription factor and a number of target genes containing p53-binding sites in their promoters have been identified, one of the most important of which is the WAF1 gene product, p21 (154) also known as Cdk-interacting protein 1 (Cip1) (155). The p21 protein binds to and inhibits G1 cyclin-cdk complex kinase activity, preventing the phosphorylation of cdk substrates and thereby blocking the transition from G1 to S phase in the cell cycle (154, 155). Therefore, the stabilisation of p53 by DNA damage results in P21<sup>WAF1/CIP1</sup> induction and arrest in G1 (156).

A number of negative regulators of p53 have been identified, including the MDM2 (157, 158) which functions as a ubiquitin E3 ligase, targeting p53 for ubiquitination and degradation by the proteasome. Indeed, MDM2 is the key cellular regulator of p53 (159). However, the expression of MDM2 is itself induced by p53, for instance following ultraviolet (UV) irradiation (160), which may serve to autoregulate p53 activity in normal cells (161). Therefore, p53 can transactivate the *Mdm2* gene, thus establishing a negative feedback loop that maintains the p53 protein at low levels in nonstressed cells (159).

Oncogenic signals activate p53 by induction of the ARF (alternative reading frame) tumour suppressor (human p14<sup>ARF</sup> and murine p19<sup>ARF</sup>) encoded by an alternative open reading frame in the p16<sup>INK4a</sup> locus, which inhibits the E3 ligase function of MDM2 (162, 163). However, DNA damage signals activate p53 via ARF-

independent pathways involving the regulatory kinases ATM (ataxia telangiectasia mutated), CHK2, and ATR (ATM and Rad3-related). In fact, DNA damage induces phosphorylation of p53 at serine 15 by ATM and ATR and at serine 20 by CHK2, resulting in disruption of the p53-MDM2 complex (164).

Examples of other genes that function as TSGs include the *phosphatase* and *tensin* *homologue* (*PTEN*), *VHL*, and *cadherins*. Each of these genes is associated with well-defined clinical syndromes in humans; in the case of *PTEN* (mapped to human chromosome 10q23) deletion is associated with a high frequency of endometrial carcinomas and glioblastomas (165). *PTEN* activity causes cell cycle arrest and apoptosis as well as inhibition of cell mobility, so loss of function of this gene alone confers three of the hallmarks of cancer (165, 166).

With *VHL* gene (chromosome 3p), germline mutations are associated with hereditary renal cancers, pheochromocytomas, haemangioblastomas of the central nervous system, retinal angiomas and renal cysts. The main substrate for *VHL* gene product is hypoxia inducible transcription factor 1 (HIF-1), which itself upregulates a number of pathways associated with increased angiogenesis. Lack of *VHL* activity prevents degradation of HIF-1 and is associated with increased levels of angiogenic growth factors (167).

Loss of *cadherin* (particularly *E-cadherin* located on chromosome 16q) is associated with increased ability to invade tissue and metastasise. This is because the cadherin family of glycoproteins act as 'glues' between epithelial cells, so loss of cadherin allows cells to disaggregate and move around more easily. Loss of cadherin function is associated with many types of cancers, including those that arise in the oesophagus, colon, breast, ovary and prostate (168).

The losses of function of the TSGs described above are due to a structural change or germline mutation of the genes involved, however this is not the only mechanism by which TSG function may be lost. In addition, TSGs may be silenced by epigenetic change involving hypermethylation of promoter sequences without a change in DNA base sequence (169). Hypermethylation can be maintained through rounds of cell division, and examples of this phenomenon include *BRCA-1* in breast cancer and *VHL* in renal cell carcinomas (170). Recent data demonstrating that genomic hypermethylation causes chromosomal instability and induces tumours in mice

strengthens the evidence that epigenetic change contributes directly to tumour development (171).

#### 1.4 *LVI-1* (*WDR76*) locus

The *LVI-1* (*lentivirus integration-1*) locus named by the Human Genome Organisation (HUGO) at <http://www.hugo-international.org> as WDR76, was identified as a target for proviral insertional mutagenesis in a case of a pre-B cell lymphoma in a cat infected with FIV (58, 172).

Lymphoma is the most common malignancy in cats and most cases are induced by FeLV, especially in younger cats, but lymphoma can occur independently of FeLV infection (173, 174). Many epidemiological studies have described lymphoproliferative malignancies in association with FIV (175-182), and an increased risk of malignant lymphoma (lymphosarcoma) in FIV-infected cats (183-185).

The feline lentivirus, FIV, is a significant feline pathogen, with an overall infection prevalence of approximately 11% in cats worldwide (186, 187). FIV is a T-lymphotropic virus and studies have been centred on the alteration of T-cell subsets, especially the T helper cells (CD4) (188). However, despite these defects in T-cell function, infection induces a marked activation and expansion of B cells. In fact, with increased time post infection, the major reservoir for proviral DNA appears to be CD21<sup>+</sup> cells (B cells) and CD8<sup>+</sup> T cells (189, 190). In the early stages of the FIV infection there is a generalized lymphadenopathy characterized by hyperplasia of B-cell areas with prominent and irregular expansion of lymphoid follicles (191) and a persistent polyclonal activation of B cells with the production of antibodies to a variety of non-viral antigens including self-antigens (192) similar to that described in HIV-infected patients (193, 194).

The clinical course of FIV infection parallels that seen in HIV infection. At 4-6 weeks after experimental infection with FIV, cats develop an acute clinical syndrome similar to that seen in HIV-1 infection, including mild pyrexia and lymphadenopathy (195) that usually disappears after 2-3 months (186). This is followed by a long asymptomatic period of latent infection lasting at least 3-5 years (196-198) in which the CD4<sup>+</sup>: CD8<sup>+</sup> T cell ratio declines because a gradual decrease in CD4<sup>+</sup> cells (188, 199, 200) and in some cases, an increase in CD8<sup>+</sup> cells as well (188, 199, 201). In naturally infected cats, the asymptomatic period is followed by the development of a

variety of disorders associated with a progressive immune dysfunction which eventually becomes manifest clinically as feline acquired immune deficiency syndrome (FAIDS) (202-204).

The majority of FIV associated lymphomas have been high-grade B cell lineage (178, 205, 206). The characteristics of FIV-associated lymphomas, thus, parallel those of the lymphomas arising in HIV-1 infection of humans (207, 208) and SIV infection of non-human primates (209, 210).

In contrast to the directly oncogenic retroviruses such as murine and feline leukaemia viruses, lentiviruses are generally considered to play an indirect role in tumour development. Most HIV-associated lymphomas are believed to develop secondary to one or a combination of factors, including immunodeficiency, polyclonal B-cell activation, and the involvement of other infectious agents, notably Epstein-Barr virus or HHV-8 (211). Between 40% and 50% of HIV-1 related non-Hodgkin's lymphomas are associated with the  $\gamma$ -herpesvirus Epstein-Barr virus (EBV) (212). It is possible that a similar virus may be involved in the aetiopathogenesis of FIV-associated lymphoma (213).

Tumours associated with HIV and related lentiviruses have generally been found to lack viral sequences, suggesting that the tumour predisposition conferred by this genus of retroviruses is due to an indirect mechanism. The demonstration of integrated sequences in only one of several FIV-associated tumours examined by restriction fragment analysis to date (213, 214) suggests that, consistent with the current knowledge of the other lentiviruses, FIV facilitates tumour development by indirect mechanisms in the majority of the cases. It has been suggested that lymphomas may emerge secondary to the immune dysfunction which accompanies FIV infection (214).

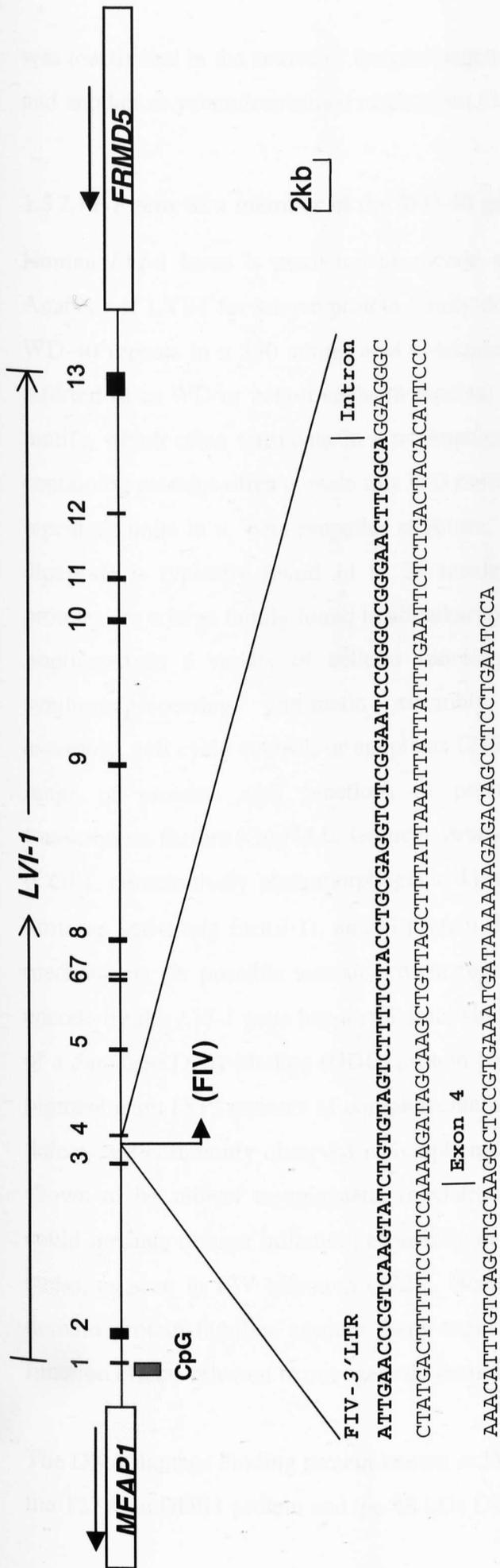
Despite these generally negative observations, a number of studies have suggested that there may be exceptions to this rule. The apparent capacity of HIV to transform non-immortalised B lymphocytes *in vitro* (215) and the detection of traces of HIV nucleic acids in AIDS lymphomas by PCR have led to speculation that HIV plays an initiating role in lymphomagenesis in at least some cases (216). A single copy, clonally integrated HIV-1 provirus was identified in a rare HIV-1 associated T-cell lymphoma. The authors speculated that the virus might have integrated near a cellular oncogene such as *c-MYC*, thereby causing constitutive *c-MYC* expression and cellular transformation (217). Another monoclonal integration of viral sequences in

the tumour DNA was demonstrated for one case of HIV-associated T-cell lymphoma where viral integration was mapped to the *FUR* gene and was associated with over expression of the nearby *FPS/FES* proto-oncogene (208).

A similar observation was found in a case of B-cell lymphoma (Q254), which arose six years and 3 months post infection, in a 10-month-old neutered male specific-pathogen-free cat that had been infected with the Glasgow-8 strain of FIV as part of a long term pathogenesis study (191). The tumour was extranodal, high grade, and of immature B-cell origin as demonstrated by immunocytochemical staining and antigen receptor gene analysis. Moreover, the tumour cells harboured a single integrated proviral copy detected by Southern blot analysis (172). This observation indicated that a single infection event preceded clonal expansion of the transformed cell and raised the possibility of a direct oncogenic role for FIV (172). Further analysis indicates that FIV integration in lymphoma Q254 occurred within a conserved gene on feline chromosome B3. RT-PCR analysis suggested that the feline gene appeared to have been transcriptionally down-regulated in the genesis of the lymphoma. Searches for matches to known genes by BLAST analysis revealed a highly significant match to a human gene (*FLJ12973/Q9H967*) that maps to chromosome 15q15 (58). The human *LVI-1* gene is highly conserved, as shown by its close homology to its murine homologue on chromosome 2 (*Q9CTV4* at band E5).

The structure of the human *LVI-1* gene based on the draft human genome sequence shows a gene comprising 13 exons and covering 40kb of the human genome on chromosome 15q15. The first exon occurs within a CpG island, which also marks the boundary with the next 5' gene (*MFAP1*, *microfibrillar associated protein 1*), which maps in the opposite orientation. The relative position of the FIV insertion in the homologous feline gene is between exons 3 and 4, in the same transcriptional orientation (Figure 1-7). RT-PCR analysis revealed a very low level of expression of a truncated viral-host gene fusion transcript arising from the 3' viral LTR adjacent to exon 4, but no transcripts were detected from the apparently unaffected allele. Moreover, this position of insertion within the gene suggests the presence of a truncated or fused mRNA species that was shown by RT-PCR analysis. However, the presence of transcripts in the control tissue (salivary gland) from the same animal demonstrated that this evident loss of expression was not a constitutive defect.

Therefore, integration resulted in truncation of the target gene, while the unaffected allele appeared to have been transcriptionally down-regulated, suggesting that *LVI-1*



**Figure 1-7. Genomic structure of the human chromosome 15q15 surrounding the *LVI-1* gene, based on the draft human genome (<http://www.genome.ucsc.edu>). *LVI-1* gene comprises 13 exons, which are indicated by numbered vertical lines and black boxes. A green box indicates the location of a CpG island at the 5' end of the gene. The relative position of the FIV insertion in the orthologous feline gene between exon 3 and exon 4 is indicated by an arrow. Adjacent genes *MFAP1* and *FRMD5*, which maps in the opposite orientation are indicated by open boxes. Nucleotide sequence of the fusion transcript derived from the FIV insertion site is shown below. Bold text and underlined letter represent the FIV LTR sequence and the exon 4 splice acceptor site, respectively. This structure suggests that the transcript arises from activity of the 3' LTR promoter element (58).**

was inactivated in the course of lymphomagenesis by a combination of viral insertion and another as yet undetermined mechanism (58).

### **1.5 *LVI-1* gene as a member of the WD-40 gene family**

Human *LVI-1* locus is predicted to encode a protein of 626 amino acid residues. Analysis of *LVI-1* for known protein family domains revealed a series of five tandem WD-40 repeats in a 250 amino acid C-terminal domain (58). WD-40 repeats (also referred to as WD or beta-transducin repeats) are approximately 40 amino acid long motifs, which often terminate in a tryptophan-aspartic acid (W-D) dipeptide. WD-containing proteins often contain this WD motif at the C-terminus, and consist of 4-16 repeating units in a 'beta-propeller structure.' In addition a glycine-histidine (GH) dipeptide is typically found 11 to 24 residues from the N-terminus. WD-repeat proteins are a large family found in all eukaryotes but not in prokaryotes (218) and are implicated in a variety of cellular functions such as signal transduction, RNA synthesis/processing, chromatin assembly, vesicular trafficking, cytoskeletal assembly, cell cycle control, or apoptosis (219). WD-40 repeats are found in a wide range of proteins with functions of potential relevance to cancer, including transcription factors (Gro/TLE, Groucho /transducin-like enhancers), ubiquitin ligases (COPI1, constitutively photomorphogenic 1), apoptotic regulators (Apaf-1, apoptotic protease activating factor-1), and G-protein  $\beta$ -subunits, presenting many alternative mechanisms for possible tumour growth regulation (219). The WD repeat protein encoded by the *LVI-1* gene has a structural similarity to DDB2 (58), the small subunit of a damaged DNA-binding (DDB) protein complex which is defective in xeroderma pigmentosum (XP) patients of complementation group E (220). Indeed, DNA repair defects are commonly observed in lymphomas, and this class of gene has also been shown to be subject to epigenetic regulation (221), suggesting that such an event could mediate tumour initiation, especially in a B-cell population under proliferative stress, as seen in FIV infection (192). However, as mentioned above, the WD-40 domain protein families execute many regulatory functions, and thus other loss-of-function effects relevant to oncogenesis should not be excluded (58).

The DNA damage binding protein known as DDB or UV-DDB (222) is constituted by the 127 kDa DDB1 protein and the 48 kDa DDB2 protein (220, 223, 224). The DDB

subunits are also found in a protein complex that includes components of a Cullin 4A (Cul4A)-based ubiquitin E3 ligase as well as the COP9/signalosome (225).

In contrast to human tissues, in rodent tissues DDB is absent or present at very low levels. Rodent genomes contain the DDB1 and DDB2 genes, but expression of DDB is suppressed in many tissues. In the mouse, DDB2 mRNA levels are relatively low in skin, brain, lung, muscle and heart, but relatively high in testes, liver and kidney (226). It was suggested that rodents may not have as great a need for DDB as humans because their skin is protected by fur and additionally because they are nocturnal animals exposure to UV is kept to a minimum (222).

Despite the ability of the DDB protein complex to bind UV light-irradiated DNA (227), the specific role of DDB in nucleotide excision repair (NER) is uncertain because NER can be reconstituted with purified components and damaged naked DNA in the absence of DDB (228-230). However, DDB has been shown to stimulate binding of the repair factors, XPA and RPA to damaged DNA (231), suggesting that DDB stimulates NER by recruiting core NER factors to damaged DNA via interaction with XPA and RPA.

NER recognises a variety of helix-distorting DNA lesions, including the cyclobutane pyrimidine dimers (CPDs) and the 6-4 pyrimidine-pyrimidone photoproducts (6-4PP) induced by UV irradiation (232). It is the only pathway in humans that removes UV-induced photoproducts (233). NER can operate via two pathways: global genomic repair (GGR) and transcription-coupled repair (TCR). GGR repairs the DNA damage over the entire genome (232), while TCR that preferentially removes DNA lesions from the transcribed strand of active genes (234).

As mentioned above, mutations in the *DDB2* gene, which encodes a WD-40 repeat containing protein, have been found in XP-E cell lines that are lacking the damaged DNA binding activity (235), suggesting that *DDB2* might contribute to NER process. In fact, *in vivo* studies showed that XP-E cells and rodent cells lacking DDB activity are selectively defective in GGR of CPDs, and transfection of the *DDB2* gene into the rodent cells complements the deficiencies (236, 237), suggesting that DDB may be involved in GGR especially for CPDs.

## 1.6 Aims of the project

By virtue of their ability to integrate into host DNA, the immunosuppressive lentiviruses have the potential to act as direct agents in carcinogenesis. Therefore it is important to monitor the sites of the integration of these agents *in vivo* and to examine their capacity to silence as well as activate host cell genes. The aims of this project were to carry out a comprehensive characterization of the *LVI-1* gene. However, the research was restricted to the human and murine homologues because of the limitations of the cat genome sequence information.

### 1. *LVI-1* gene expression studies

The first aim was to carry out an analysis of human and murine *LVI-1* gene expression *in vivo* and in cell lines *in vitro*. In view of limited knowledge of the murine gene, cDNA cloning was undertaken to establish the full sequence and coding potential of the *Lvi-1* transcripts in representative murine tissues.

### 2. LVI-1 gene product(s) studies

The aims were to identify and determine the localization of the gene product(s) as a clue to protein function. Specific antisera were raised to the human and murine gene product(s) and these were tested by immunofluorescence and analysis of cell fractions by western blotting.

### 3. Clues to LVI-1 function

The similarity of LVI-1 to DDB2 indicates a possible role for this gene in DNA repair pathways. In addition, as many UV damage repair genes show a direct response to induced damage by increased transcription, post-translational modification or intracellular translocation, the effects of the UV irradiation were assessed using several different types of human and murine cell lines of known p53 status.

The following chapters will describe the experimental techniques used to achieve these objectives and the results that were obtained.

## Chapter II

### Materials and methods

#### 2.1 Materials

Materials in regular use such as equipment and general solutions are detailed in this Section.

##### 2.1.1 Cell culture materials

###### *Sources of cell lines*

WI-38: The WI-38 human diploid cell line derived from normal embryonic (3 months gestation) lung tissue (238).

CEM: Human T lymphoblastoid cell line derived from peripheral buffy coat of a case of acute lymphoblastic leukaemia (239).

U-937: Human histiocytic lymphoblastoid cell line derived from pleural effusion in a case of diffuse histiocytic lymphoma (240).

KG-1a: Human acute myeloid lymphoblastoid cell line derived from bone marrow of a case of erythroleukemia that developed into acute myeloid leukaemia (241).

K-562: Human chronic myeloid lymphoblastoid cell line derived from pleural effusion in a case of chronic myeloid leukemia (CML) in blast crisis (242).

Jurkat: The Jurkat E61 clone is a human acute T cell leukaemia cell line (243).

Kasumi-1: Human acute myeloid lymphoblastoid cell line derived from a case of with acute myeloid leukaemia (244).

SW480: Human colon adenocarcinoma cell line derived from tumour of a case of colon adenocarcinoma (grade 4, Duke class B) (245).

P/mi (p/m16i, p/m32i and p/m47i): The p/mi mouse cell lines were derived from Moloney Murine Leukemia virus (Mo-MuLV) induced thymic lymphomas in p53-null CD2-MYC mice (246).

BW5147: Mouse T lymphoblastoid cell line derived from a spontaneous AKR/J thymoma (247).

3T3: Mouse fibroblast cell line derived from disaggregated Swiss albino mouse embryos (248).

H/3T3: Mouse fibroblast cell line H-*ras* transformed 3T3 cells (H-3T3 cells were provided by Prof. Alan Balmain previously of Beatson Institute for Cancer Research, University of Glasgow).

### ***Plasticware***

Tissue culture flasks, pipettes (5, 10, and 25 ml), 100 mm cell culture dish and cell scraper were supplied by Costar (Corning Incorporated, Netherlands). Cryotubes 1.8 ml were supplied by Nunc (DK 400, Roskilde, Denmark). Falcon conical centrifuge tubes (15 and 50 ml) were supplied by Becton Dickinson Labware Europe. Chamber Slide System-2 well glass slide was supplied by Lab-Tek (Nalge Nunc International, Rochester, New York).

### ***Solutions, media and supplements***

All solutions and media for cell culture were supplied by Invitrogen Life Technologies, unless otherwise stated.

### ***Media***

Dulbecco's Minimum Essential Medium (DMEM) with 4500 mg/L glucose, L-Glutamine and pyruvate.

RPMI 1640 medium with L-glutamine.

## **Supplements**

Foetal Bovine Serum (FBS), HyClone, Perbio: virus and mycoplasma screened. FBS was heat inactivated at 56 °C for 30 minutes, then stored in 50 ml aliquots at -20 °C until use.

L-glutamine: supplied as a 200 mM (100x) stock solution, and stored at -20 °C in 5 ml aliquots. This was routinely added to culture media prior to use.

Penicillin/streptomycin (P/S): supplied as a 100x stock solution of 10,000 units penicillin and 10,000 units streptomycin per millilitre and stored in 5 ml aliquots at -20°C.

Trypsin-EDTA: supplied as 10x liquid, stored at -20 °C. This was diluted 1:10 in sterile PBS prior to use and stored at 4 °C.

2-Mercaptoethanol (2-ME): supplied as 50 mM solution in Dulbecco's PBS.

N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid (Hepes): supplied as 1 M solution in dH<sub>2</sub>O.

### **2.1.2 Bacterial strains**

#### ***E.coli One Shot<sup>®</sup> TOP10***

*E. coli* TOP10: One Shot Chemically Competent *E.coli* cells (~1 x 10<sup>9</sup> colony forming units/μg) (Invitrogen). Genotype: F<sup>-</sup> *mcrA* Δ(*mrr-hsdRMS-mcrBC*)Φ80 *lacZ*ΔM15 Δ*lacX74 deoR recA1 araD139 Δ(ara-leu)7697 galU galK rpsL* (Str<sup>R</sup>) *endA1 nupG*. The genotype φ80*lacZ*ΔM15 marker enables blue-white screening by α-complementation of β-galactosidase encoded by vector DNA (e.g. pCR<sup>™</sup> II).

#### ***E.coli BL21***

*E. coli* BL21: A protease-deficient *E. coli* host strain for optimal expression of recombinant proteins (Amersham Pharmacia Biotech). Genotype: *E. coli* B F<sup>-</sup>, ompT, *hsdS* (*r<sub>B</sub><sup>-</sup>, m<sub>B</sub><sup>-</sup>*), *gal, dcm*.

### 2.1.3 Complete kits

5'/3' RACE Kit, 2<sup>nd</sup> Generation (Roche Applied Science)  
Big Dye<sup>®</sup> Terminator Version 1.1 Cycle Sequencing Kit (Applied Biosystems)  
Bulk GST Purification Module (Amersham Biosciences)  
ECL<sup>™</sup> Western Blotting Detection Reagents (Amersham Biosciences)  
EndoFree<sup>®</sup> Plasmid Maxi Kit (Qiagen)  
GST Detection Module (Amersham Biosciences)  
High Pure PCR Product Purification Kit (Roche Applied Science)  
High Pure Plasmid Isolation Kit (Roche Applied Science)  
Improm-II<sup>™</sup> Reverse Transcription System (Promega)  
QIAquick<sup>®</sup> PCR Purification Kit (Qiagen)  
QIAquick<sup>®</sup> Gel Extraction kit (Qiagen)  
Rad Prime DNA Labeling System (Invitrogen Life Technologies)  
RNA 6000 Nano LabChip<sup>®</sup> Kit (Agilent Technologies)  
TOPO TA Cloning<sup>®</sup> (Invitrogen Life Technologies)

### 2.1.4 DNA

Plasmid, molecular weight markers and oligonucleotide DNAs were stored at -20 °C.

#### *Plasmid vectors*

pCR<sup>®</sup> 2.1-TOPO vector (Invitrogen): *lacZ*<sup>+</sup>, *amp*<sup>+</sup>, *kan*<sup>+</sup>, T7 promoter and priming site, M13 forward and reverse primer annealing sites. Plasmid designed for direct cloning of PCR products with 3' deoxyadenosine residues (A - overhangs), generated by the non-template dependent activity of *Taq* polymerase. The vector is supplied as linearised DNA with single 3' deoxythymidine (T) residues allowing for efficient ligation of target sequence to vector.

pCR<sup>®</sup>II-TOPO (Invitrogen): As above but also has SP6 promoter and priming sites.

pUC19 (Invitrogen): Plasmid supplied with TA cloning kit for use as positive control for verifying the transformation efficiency of competent bacteria; concentration of 0.1 µg/ml.

Glutathione S-transferase fusion vectors, pGEX-5X-1 and pGEX-5X-2 (Amersham Pharmacia Biotech): The vectors contain the Glutathione S-Transferase (GST) gene (with an ATG and ribosome-binding site, and is under the control of the *tac* promoter), an expanded multiple cloning site (MCS) that contains six restriction sites and provides all three translational reading frames beginning with the *EcoRI* restriction site, and an internal *lac I<sup>q</sup>* gene for use in any *E. coli* host (Figure 2-1).

pGEX-5X-1: Glutathione S-transferase gene region: *tac* promoter: -10: 205-211; -35: 183-188; *lac* operator: 217-237; Ribosome binding site for GST: 244; Start codon (ATG) for GST: 258; Coding region for factor Xa cleavage: 921-932. MCS: 934-969. Beta-lactamase gene region: Promoter: -10: 1333-1338; -35: 1310-1315; Start codon (ATG): 1380; Stop codon (TAA): 2238. *lacI<sub>q</sub>* gene region: Start codon (GTG): 3321; Stop codon (TGA): 4401. Plasmid replication region: Site of replication initiation: 2998; Region necessary for replication: 2305-3001. Sequencing primers: 5' pGEX Sequencing Primer binds nucleotides 869-891; 3' pGEX Sequencing Primer binds nucleotides 1044-1022.

pGEX-5X-2: Glutathione S-transferase gene region: *tac* promoter: -10: 205-211; -35: 183-188; *lac* operator: 217-237; Ribosome binding site for GST: 244; Start codon (ATG) for GST: 258; Coding region for factor Xa cleavage: 921-932. MCS: 934-970. Beta-lactamase gene region: Promoter: -10: 1334-1339; -35: 1311-1316; Start codon (ATG): 1381; Stop codon (TAA): 2239. *lacI<sub>q</sub>* gene region: Start codon (GTG): 3322; Stop codon (TGA): 4402. Plasmid replication region: Site of replication initiation: 2999; Region necessary for replication: 2306-3002. Sequencing primers: 5' pGEX Sequencing Primer binds nucleotides 869-891; 3' pGEX Sequencing Primer binds nucleotides 1045-1023.

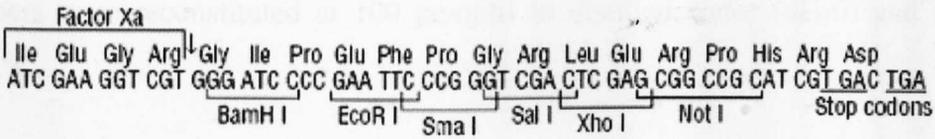
### ***Molecular size standards***

Molecular size standards used include  $\phi$ X174 RF DNA/HaeIII fragments (size range 72 bp to 1,353 bp),  $\lambda$  DNA/HindIII fragments (size range 125 bp to 23,130 bp), and Low DNA Mass™ Ladder (size range: 100-2000 bp) all supplied by Invitrogen Life Technologies.

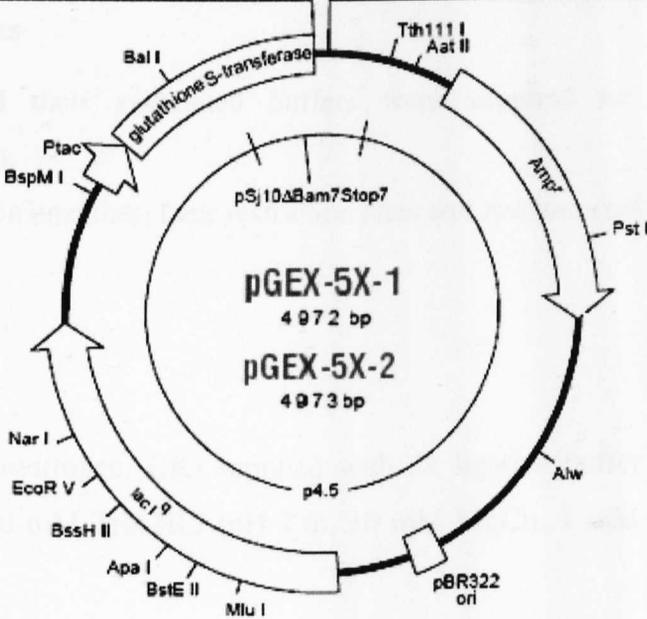
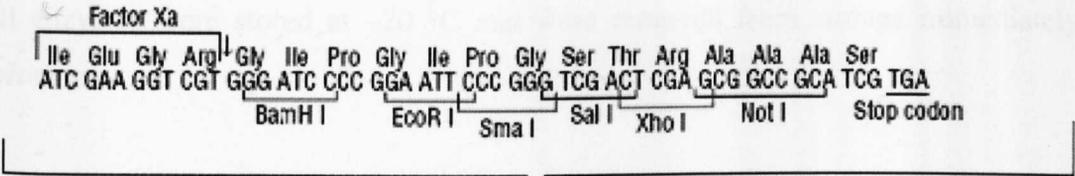
### ***Oligonucleotide primers***

Oligonucleotides for polymerase chain reaction (PCR) and cycle sequencing were

**Multiple cloning site of pGEX-5X-1**



**Multiple cloning site of pGEX-5X-2**



**Figure 2-1. Organization of pGEX-5X-1 and pGEX-5X-2 vectors (GST gene fusion system).** Plasmid map: Glutathione-S-transferase gene region (*tac* promoter, *lac* operator); multiple cloning site (MCS); Amp<sup>r</sup>, ampicillin resistance; ori, pBR322 origin of replication. Coding sequences of the MCS of both vectors are indicated above, in which the stop codon(s) is located at 3' of the *Not* I site.

synthesised by MWG Biotech, and delivered as high purity salt free lyophilised DNA. Primers were reconstituted at 100 pmol/ $\mu$ l in distilled water (dH<sub>2</sub>O) and stored at -20 °C.

### **2.1.5 Enzymes**

All enzymes were stored at -20 °C and were removed from storage immediately before use.

#### ***Restriction enzymes***

All enzymes and their associated buffers were supplied by Invitrogen Life Technologies (UK).

Details of restriction enzymes, their restriction sites and reaction conditions are shown in Table 2-1.

#### ***T4 DNA Ligase***

T4 DNA ligase (Invitrogen, UK) supplied with 5x ligation buffer (used at a final concentration of 50 mM Tris-HCl (pH 7.6), 10 mM MgCl<sub>2</sub>, 1 mM DTT and 1 mM ATP).

#### ***Alkaline Phosphatase, Calf Intestinal (CIAP)***

Calf Intestinal Alkaline Phosphatase (CIAP) (Promega, UK) supplied with 10x reaction buffer (used at a final concentration of 0.05 mM Tris-HCl (pH 9.3 at 25 °C), 1 mM MgCl<sub>2</sub>, 0.1 mM ZnCl<sub>2</sub> and 1 mM Spermidine)

#### ***2x ReddyMix™ PCR Master Mix***

2x ReddyMix™ PCR Master Mix (1.5 mM MgCl<sub>2</sub>) (Abgene, UK). Standard PCR reactions were performed in a final volume of 25  $\mu$ l using 12.5  $\mu$ l of 2x ReddyMix™ PCR Master Mix and 12.5  $\mu$ l of a mixture of template, primers and dH<sub>2</sub>O. Each reaction contains *Taq* DNA Polymerase (1.25 units), 75 mM Tris-HCl (pH 8.8 at 25 °C), 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.5 mM MgCl<sub>2</sub>, 200  $\mu$ M of each dNTP, precipitant, and red dye for electrophoresis.

**Table 2-1. Restriction enzymes.**

Restriction enzyme	Restriction site	Buffer	Temperature of incubation (°C)
<i>EcoRI</i>	[5'-G↓AATTC-3']	REACT <sup>®</sup> 3	37
<i>BamHI</i>	[5'-G↓GATC-3']	REACT <sup>®</sup> 3	37
<i>SstI</i>	[5'-GAGCT↓C-3']	REACT <sup>®</sup> 2	37
<i>EcoRV</i>	[5'-GAT↓ATC-3']	REACT <sup>®</sup> 2	37
<i>XbaI</i>	[5'-T↓CTAGA-3']	REACT <sup>®</sup> 2	37

### ***Extensor Hi-Fidelity PCR Master Mix, Buffer 1, ReddyMix™***

Extensor Hi-Fidelity PCR Master Mix, Buffer 1, ReddyMix™ (Abgene, UK). Standard PCR reactions were performed in a final volume of 25 µl using 12.5 µl of Extensor Hi-Fidelity PCR Master Mix, Buffer 1, ReddyMix™ and 12.5 µl of a mixture of template, primers and dH<sub>2</sub>O. Each reaction contains *Taq* DNA Polymerase (1.25 units), 2.25 mM MgCl<sub>2</sub>, 350 µM of each dNTP, buffer 1 (used for amplifications up to 12 kb), precipitant, and red dye for electrophoresis.

### ***Ribonuclease A Type III-A***

Ribonuclease A Type III-A (Sigma, UK) was used to purify DNA from contaminating RNA.

## **2.1.6 RNA**

### ***Molecular size standards***

Molecular size standards (5 µg/lane) include a 0.24 kb to 9.5 kb RNA ladder and a 0.5 kb to 10 kb RNA ladder all supplied by Invitrogen Life Technologies. RNA ladders were stored at -70 °C.

## **2.1.7 Radiochemicals**

$\alpha$ -<sup>32</sup>P dCTP, specific activity 3000 Ci/mmol, for DNA labelling was supplied by Amersham Life Science.

## **2.1.8 Protein SDS-PAGE standards**

Prestained SDS-PAGE (Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis) standard was the Precision Plus Protein™ Standards, supplied by Biorad. This standard contains ten highly purified recombinant proteins ranging from 10 kDa to 250 kDa. The stock was aliquoted and stored at -20 °C. Each aliquot was heated to 95 °C for 3 minutes prior to use to dissolve any precipitated solids.

### **2.1.9 General chemicals**

Chemicals used were of analytic, ultrapure or molecular grade quality and were supplied by Sigma Chemical Company (Dorset, England), Fisher Scientific UK Limited (Leicestershire, UK), Invitrogen Life Technologies, unless stated otherwise.

Bacterial agar, and tryptone, and phosphate buffered saline (DulbeccoA): Oxoid.

### **2.1.10 Equipment**

#### ***Major Equipment***

Automatic Sequencing Apparatus: ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, USA)

Flow cytometer: Coulter Epics XL MCL (Beckman Coulter Company, Florida, USA)

Gel cast: AE-6210 Slab Gel Cast (Atto Incorporation, Japan); casting stand and casting frames from Mini-Protean 3 Cell and Systems, and gel caster from Sub-Cell Systems (Bio-Rad Hertfordshire, UK)

Gel drier: Easy Breeze Gel Dryer (Hoeter Scientific Instruments, San Francisco, USA)

Gel electrophoresis systems: Biomax QS710 Horizontal Unit (Scientific Imaging Systems – Eastman Kodak Company, New York, USA); Sub-Cell GT Submerged Horizontal Agarose Gel Electrophoresis System, and MiniProteanII™ (Bio-Rad, Hertfordshire, UK); AE-6200 Slab EP Chamber, and AE-6220 Slab EP Chamber, for two gels (Atto Incorporation, Japan)

Hybridisation oven: Mini Oven MKII, Hybaid Limited

PCR Machines: PCR Express Thermal Cycler (Hybaid Limited, Middlesex, UK), DNA Thermal Cycler (Perkin Elmer, Boston Ma, USA), and PTC-220 DNA Engine Dyad® Cycler (MJ Research Incorporated, Massachusetts, USA)

Phosphoimager: Molecular Dynamics Storm 840 equipped with ImageQuant software (Amersham Pharmacia Biotech, Sunnyvale, USA)

Power packs: Electrophoresis Power Supply EPS 500/400 (Pharmacia Fine Chemicals, Uppsala, Sweden), and Power PAC 300 (Bio-Rad, Hertfordshire, UK)

Semi-dry electrophoretic transfer: Trans-blot® SD Semi-Dry Transfer Cell (Bio-Rad Hertfordshire, UK)

Sonicator: Soniprep 150 MSS 150.CX 3.5 (Sanyo Gallenkamp PLC, Leicester, UK)

Spectrophotometer: DU<sup>®</sup> 640 Spectrophotometer (Beckman, California, USA),  
Agilent 2100 Bioanalyzer (Agilent Technologies, Germany)  
Ultraviolet trans-illuminator: High Performance Ultraviolet Transluminator (UVP-  
Ultra Violet Products, San Gabriel, CA), and GeneFlash Syngene Bio Imaging (Fisher  
Scientific- Synoptics Ltd, UK)  
UV Cross Linker: Spectrolinker<sup>™</sup> XL-1500 UV Crosslinker (Spectrotronics  
Corporation, New York, USA)

### *Consumables*

21G sterile syringe needles were supplied by Becton Dickinson S. A. (Fraga, Spain)  
Acrodisc syringe filters (0.2  $\mu\text{m}$  and 0.45  $\mu\text{m}$ ) were supplied by Sartorius (Hannover,  
Germany); used for sterilising of filtering small volumes of solutions  
Centrifugal filters: Microcon and Centricon YM-30 (Millipore Corporation, Bedford  
USA)  
Eppendorf tubes: Flip-top 0.5 ml and 1.5 ml, and screw-top 1.5 ml (Elbkay  
Laboratory Products, Hampshire, UK), and Rnase-free Microfuge Tubes (Ambion,  
Cambridgeshire, UK)  
Film: Hyperfilm MP, and Hyperfilm ECL (Amersham Biosciences,  
Buckingshamshire, UK)  
Filter paper: Qualitative Circles 185 mm  $\varnothing$  (Whatman<sup>®</sup>, Kent, England)  
Filter tip pipette tips: A range of capacities (10, 20, 200, 1000)  $\mu\text{l}$  supplied by Rainin  
Instrument Co (Woburn, MA); for use in setting up PCR reactions  
Fine Tip Pastette Sterile were supplied by Alpha Laboratories (Hampshire, UK)  
Flat ended gel loading tips were supplied by Sorenson Bioscience Inc (Utah, USA)  
Hybridisation buffer: Rapid-hyb buffer (Amersham Pharmacia Biotech,  
Buckingshamshire, UK)  
Gel exclusion column: Nick<sup>™</sup> columns (Amersham Biosciences, Uppsala, Sweden)  
Membranes: Hybond<sup>™</sup>-N nylon membrane and Hybond<sup>™</sup>-ECL nitrocellulose  
membrane (Amersham Biosciences, Buckingshamshire, UK)  
Parafilm was supplied by Pechiney Plastic Packaging (Menasha WI)  
Petri dishes were supplied by Sterilin (Staffordshire, UK)  
Pipette tips were supplied by Sarstedt (Nümbrecht, Germany)

Pipette tips RNase free: 1000 µl and 200 µl supplied by Starlab (Ahrensburg, Germany), and 1 µl to 20 µl supplied by Elkay Laboratory Products Ltd (Hampshire, England)

Scalpel blades: Sterile disposable scalpels (Schwann-Morton, Sheffield, UK)

Spreaders, L shaped were supplied by VWR International (Lutterworth, UK)

Sterile plastic containers: Bijoux and Universals (Greiner Bio-One, Gloucestershire, UK)

Syringes (2, 5, 10, 20, 50) ml were supplied by Becton Dickinson

### **2.1.11 Buffers, solutions and growth media**

#### ***Water***

Diethylpyrocarbonate (DEPC) treated RNase free water was supplied by Ambion, Cambridgeshire, UK. Tissue culture grade distilled water was supplied by Invitrogen Life Technologies. Ultrapure water (for procedures involving recombinant DNA, PCR etc.) was provided by a Milli Q<sup>®</sup> ultrapure water purification system (Millipore Ltd., Watford, UK). An Elix<sup>®</sup> purification system (Millipore Ltd., Watford, UK) was used to supply water for preparation of general solutions and media.

#### ***Antibiotics***

Ampicillin (Sigma): 100 mg/ml in ddH<sub>2</sub>O. Filtered through a 0.22 µm filter, aliquotted and stored at -20 °C until use.

#### ***Buffers and solutions***

10x Dry Blot Buffer: 0.5 M Tris base, 0.4 M Glycine, 12.8 mM SDS. Working stock was prepared prior to use at a 1:10 dilution with 20% (v/v) Methanol.

10x MOPS (3-N-Morpholino-propanesulfonic acid) Buffer: 200 mM MOPS, 50 mM NaOAc (3H<sub>2</sub>O), 10 mM Na<sub>2</sub>EDTA, adjusted to pH 7.0 using either NaOH if free acid MOPS or HCl if Na salt MOPS. Stored at 4 °C in the dark.

10x SDS-PAGE Running Buffer: 0.25 M Tris base, 1.92 M Glycine, 34.7 mM SDS

10x TBE (Tris-borate-EDTA) Buffer Solution: 0.9 M Tris base, 0.9 M Boric acid, 23 mM Na<sub>2</sub>EDTA. pH 8.2/8.3. Stored at room temperature.

10x TBST (Tris-buffered-saline-tween) Buffer Solution: 0.1 M Tris base, 1.5 M NaCl, 0.5% Tween-20, adjusted to pH 8.0.

10x Tris-glycine buffer for electrotransfer of proteins onto membranes (transfer buffer): 0.25 M Tris base, 1.92 M Glycine. Working stock was prepared prior to use at a 1:10 dilution with 15% (v/v) Methanol.

1M Tris HCl: 121 g Tris base, 800 ml dH<sub>2</sub>O. Adjusted to desired pH with concentrated HCl and made up to 1 L.

1x Phosphate buffered saline (PBS): 140 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.3)

2.5x RadPrime Buffer: 125 mM Tris-HCL, pH 6.8; 12.5 mM Magnesium chloride; 25 mM 2-Mercaptoethanol; 150 µg/ml Oligodeoxyribonucleotide primers

20x Propidium iodide (PI) stock solution: made to a working solution of 1 mg/ml with ddH<sub>2</sub>O in a fume cupboard. This was filtered through a 0.22 µm syringe filter, and stored at 4 °C in the dark.

20x SSC: 3 M NaCl, 0.3 M Sodium citrate, pH 7.0 in dH<sub>2</sub>O. Stored at room temperature.

50x TAE (Tris-acetate-EDTA) Buffer Solution: 2 M Tris base, 50 mM Na<sub>2</sub>EDTA, 1 M Glacial acetic acid.

Acrylamide elution buffer: 500 mM Ammonium acetate, 10 mM Magnesium acetate, 0.1 mM Na<sub>2</sub>EDTA, 0.1% (w/v) SDS)

Ammonium persulphate: 10% (w/v) stock solution in ddH<sub>2</sub>O, freshly made.

Denaturation buffer: 1.5 M NaCl, 0.5 M NaOH in dH<sub>2</sub>O. Stored at room temperature.

Destain solution for protein gels: 5% Methanol, 7.5% Acetic acid in ddH<sub>2</sub>O.

Ethidium bromide: made to a working solution of 3 mg/ml with ddH<sub>2</sub>O in a fume cupboard. Stored at room temperature in the dark.

Gel loading buffers:

1. DNA- 30% Glycerol, 0.25% (w/v) Bromophenol blue, 0.25% (w/v) Xylene cyanol, in ddH<sub>2</sub>O. Stored at room temperature and used at a 1:5 dilution.

2. RNA- 50% (v/v) Formamide, 2.2 M Formaldehyde, 1x MOPS, in nuclease-free water. Made fresh. Immediately prior to loading on the gel add RNA running dye: 50% Glycerol, 1x MOPS buffer, 0.5% (w/v) Bromophenol blue, in nuclease-free water (stored at -20 °C).

3. Protein- SDS-PAGE Protein 2x Sample Loading Buffer: 120 mM Tris-HCl, pH 6.8, 20% (v/v) Glycerol, 5% SDS, 7.4% β-mercaptoethanol, 0.5% (w/v) Bromophenol blue in ddH<sub>2</sub>O. Stored in aliquots at -20 °C. Sample diluted at least 1:2 with buffer and heated at 100 °C for 5 minutes prior to loading gel.

Glutathione elution buffer 1: 10 mM glutathione, 50 mM Tris-HCl, pH 8.0

Glutathione elution buffer 2: 75 mM Hepes pH 7.4, 150 mM NaCl, 10 mM reduced glutathione, 5 mM DTT, 2% N-octyl glucoside

Lysis buffer for immunoprecipitation: 50 mM Tris pH 8.0, 150 mM NaCl, 0.5% (v/v) Nonidet P 40, 5 mM Na<sub>2</sub>EDTA, 1 mM Dithiethreitol (DTT). Stored in aliquots at -20 °C. Working Lysis buffer was prepared prior to use by addition of 0.2 mM Na<sub>2</sub>VO<sub>4</sub>, and protease inhibitors as follows: 10 μg/ml Aprotinin, 10 μg/ml Leupeptin, 2.5 μg/ml Pepstatin A, 10 mM β-glycerophosphate, 1 mM Phenylmethanesulfonyl fluoride (PMSF).

Neutralisation buffer: 1.5 M NaCl, 0.5 M Tris-HCl pH 8.0, in dH<sub>2</sub>O. Stored at room temperature.

PI Staining solution: 50 µg/ml 20x Propidium Iodide (PI) stock solution, 100 U/ml Rnase A in sample buffer (1 g/l Glucose in PBS. This was filtered through a 0.22 µm syringe filter, and stored at 4 °C).

SDS-PAGE Protein Gel Fix-Stain Solution: 0.25% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) Methanol, 10% (v/v) Acetic acid in ddH<sub>2</sub>O.

STE buffer: 10 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA

TE Buffer: 10 mM Tris-HCl, 1 mM Na<sub>2</sub>EDTA adjusted to pH 8.0. Autoclaved and stored at room temperature.

Whole Cell Extract Buffer: 20 mM Hepes pH 7.0, 5 mM EDTA, 10 mM EGTA, 5 mM NaF, 1 mM DTT, 0.4 M KCl, 0.4% (v/v) Triton X-100, 10% (v/v) Glycerol. Stored in aliquots at -20 °C. Working Whole Cell Extract buffer was prepared prior to use by addition of 0.1 µg/ml Okadaic acid, and protease inhibitors as follows: 5 µg/ml Aprotinin, 5 µg/ml Leupeptin, 5 µg/ml Pepstatin A, 1 mM Benzamidine, 50 µg/ml PMSF.

X-Gal (5-bromo-4-chloro-3-indolyl-βgalactoside) solution: prepared as 40 mg/ml stock in dimethylformamide; stored in aliquots at -20 °C in the dark.

### ***Bacteriological media***

Media was sterilised by autoclaving at 121 °C for 15 minutes, unless otherwise stated.

2x YTA medium: 1.6% (w/v) Tryptone, 1% (w/v) Yeast extract, 0.5% (w/v) NaCl in dH<sub>2</sub>O. pH then adjusted to 7.0 with NaOH.

LB (Luria Bertani) medium: 1% (w/v) Tryptone, 0.5% (w/v) Yeast extract, 1% (w/v) NaCl in dH<sub>2</sub>O. pH then adjusted to 7.0 with NaOH.

LB-agar: as for LB medium but also containing 1.5% (w/v) agar.

LBG medium: 1% (w/v) Tryptone, 0.5% (w/v) Yeast extract, 0.5% (w/v) NaCl in dH<sub>2</sub>O. After the medium has cooled to (50 to 60) °C, 20 mM Glucose was added.

LBG-agar: as for LBG medium but also containing 1.5% (w/v) agar.

SOC medium: 2% Tryptone, 0.5% Yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, and 20 mM Glucose. (Invitrogen Life Technologies, UK).

TSS (Transformation and Storage Solution): 1% (w/v) Tryptone, 0.5% (w/v) Yeast extract, 0.5% NaCl, 10% (w/v) Polyethylene glycol (M<sub>r</sub> 3350), 5% Dimethylsulfoxide (DMSO), 0.05 mM MgCl<sub>2</sub> in dH<sub>2</sub>O. pH then adjusted to 6.5. Sterilised by filtering through a 0.22 µm syringe filter, and stored at 4 °C. Stable for up to 6 months.

## **2.2 Methods**

Methods used commonly throughout the thesis are detailed in this Section whilst more specific methods are dealt with in the relevant chapter. Many of the methods described here are based on standard protocols, which are detailed in several laboratory manuals (249, 250).

### **2.2.1 Growth and manipulation of mammalian cells**

All procedures involving mammalian cells were carried out in a laminar flow hood using standard aseptic procedures.

#### ***Cryopreservation of cells***

Stocks of cells for long-term storage were preserved over liquid nitrogen. Cells to be frozen were grown to mid-log phase (as described below) and removed into a sterile 50 ml centrifuge tube (using 1x trypsin-EDTA solution where necessary). Cells were centrifuged at 1000 rpm for 5 minutes and the supernatant discarded. The cells were then re-suspended in freeze medium at a concentration of approximately  $3 \times 10^6$  cells/ml. Freeze medium consisted of the appropriate culture medium supplemented with FBS to 20% and 10% dimethylsulphoxide (DMSO, Sigma Aldrich) as a cryoprotectant. The cell suspension was transferred in one millilitre aliquots, to labelled cryovials (Nunc cryotubes vials, Nunc) and brought to  $-70\text{ }^{\circ}\text{C}$  at a controlled rate of  $-1\text{ }^{\circ}\text{C}$  per minute using a NALGENE™ Cryo 1  $^{\circ}\text{C}$  Freezing container (NALGENE, USA). The vials were then transferred to a liquid nitrogen freezer for long-term storage. Cell stocks were revived by rapid thawing in a  $37\text{ }^{\circ}\text{C}$  water bath and maintained following standard techniques (described below).

#### ***Cell counting***

Cells were counted in a haemocytometer (Bright-Line Hemacytometer, Sigma) as follows. Cell pellets were suspended in an appropriate volume of media to allow ease of counting in the haemocytometer chamber. A  $10\text{ }\mu\text{l}$  volume of the cell suspension was then diluted 1:1 in trypan blue stain 0.4% (Invitrogen Life Technologies) and incubated at room temperature for 1 minute. The suspension was then introduced to

the haemocytometer chamber and cell counts made using an inverted microscope with a 4x or 10x objective. Cells lying on the top and right hand perimeter of each large square (1 mm) were included; those on the bottom or left hand perimeters were excluded. Cell concentrations (cells/ml) were calculated by multiplying the mean numbers of cells per large marked square by  $2 \times 10^4$  to account for the volume of the haemocytometer chamber and to correct for the dilution factor. Dead cells were differentiated by uptake of the trypan blue stain and appear blue.

### ***Maintenance of mammalian adherent cell lines***

WI-38, SW480, 3T3, and H/3T3, (Section 2.1.1). All cell lines form an adherent monolayer in culture, and were cultured in 75 cm<sup>2</sup> tissue culture flasks kept at 37 °C in a humidified incubator containing 5% (v/v) CO<sub>2</sub>. Cells were cultured in 20 ml DMEM (Section 2.1.1) supplemented with 10% FBS, 100 international units (IU)/ml penicillin, 100 µg/ml streptomycin, and 2 mM glutamine. Cultures were split typically 1:3, every 2 to 3 days, when sub-confluent and maintained at densities of 10<sup>6</sup> cells/ml. The medium was decanted from the cell monolayer and cells washed with 10 ml of warm PBS. Then, 3 ml to 5 ml of 1x trypsin-EDTA solution was added, cells were incubated at 37 °C for 5 to 15 minutes and observed by microscopy to ensure cell layer dispersal. The detached cells were then washed in fresh medium, pelleted by centrifugation at 1000 rpm for 5 minutes, prior to resuspending in fresh medium and seeding new tissue culture flasks.

### ***Maintenance of mammalian suspension cell lines***

CEM, U-937, KG-1a, K-562, Jurkat, Kasumi-1, p/mi, and BW5147 (Section 2.1.1). Cells were cultured in suspension in RPMI 1640 medium (Section 2.1.1) supplemented with 10% FBS, 100 IU/ml penicillin, 100 µg/ml streptomycin, 2 mM glutamine, and  $5 \times 10^{-5}$  M β-mercaptoethanol with the exception of KG-1a cell line which was maintained in DMEM supplemented with 10% FBS, 100 IU/ml penicillin, 100 µg/ml streptomycin, and 2 mM glutamine. Cultures were maintained at 37 °C in a humidified incubator with 5% CO<sub>2</sub> in upright 75 cm<sup>2</sup> tissue culture flasks. Cells were split every two to three days, typically 1:5 to 1:7, and maintained at densities of between  $5 \times 10^5$  cells/ml and  $5 \times 10^6$  cells/ml. The suspension cells were pelleted by

centrifugation at 1000 rpm for 5 minutes, prior to resuspending in fresh medium and seeding new tissue culture flasks.

## **2.2.2 Recombinant DNA techniques**

### ***Storage and growth of bacteria***

Plasmid DNA was maintained and stored in the *E.coli* strain One Shot Chemically Competent cells<sup>®</sup>. Glycerol stocks were prepared from transformed bacteria and their long-term storage as outlined below.

The desired bacterial culture was streaked onto a 1.5% LB agar plate (Section 2.1.11); as the plasmid conferred ampicillin resistance the medium was supplemented with 100 µg/ml ampicillin. The plate was then incubated overnight at 37 °C and the following day single colonies were picked using a sterile pipette tip into a sterile universal containing 5 ml LB medium supplemented with 100 µg/ml ampicillin. Cultures were incubated at 37 °C overnight in a horizontal orbital incubator at 220 rpm. Confirmation that overnight cultures were derived from bacteria containing the correct plasmid was achieved by DNA isolation and restriction digestion. Glycerol stocks were prepared by addition of 150 µl of glycerol to 850 µl of culture broth to produce a 15% glycerol mixture. Glycerol stocks were then stored at -20 °C and -70 °C. Bacterial stocks were revived for subsequent work by using a sterile platinum wire to scratch the surface of the glycerol stock, and streaking this onto an agar plate as outlined above.

### ***Extraction and purification of plasmid DNA***

Plasmid DNA was isolated using a modification of the alkali lysis technique described by Birnboim and Doly (1979) (251).

### ***Large-scale plasmid preparation***

Large quantities of highly pure, endotoxin free plasmid were prepared using the EndoFree<sup>®</sup> Plasmid Maxi Kit (Qiagen,UK). 50 µl of glycerol stock from the desired transformant was used to seed an overnight multiplier culture in 100 ml of LB medium incubated at 37 °C with shaking at 220 rpm. These exponentially growing

bacteria were harvested by centrifugation of the culture in 50 ml sterile centrifuge tubes at 4000 g for 20 minutes at 4 °C. The remainder of the protocol was performed according to the manufacturer's instructions. DNA was stored at -20 °C.

#### *Small-scale plasmid preparation*

Requirements for small amounts of plasmid DNA, such as sequencing, were met using the High Pure Plasmid Isolation Kit (Roche Applied Science, Germany). This kit allows isolation of plasmid DNA from 2 ml to 3 ml LB broth cultures of exponentially growing bacteria. Bacteria were harvested by centrifugation (13,000 rpm for 1 minute), and DNA isolated following the manufacturer's protocol. The method involved lysis of the bacterial cells to release plasmid DNA, which was then purified by centrifugation, filtration using spin columns, washing and final elution into 100 µl of elution buffer (10 mM Tris-HCl, pH 8.5). DNA storage was at -20 °C.

#### *Determination of nucleic acid concentration and quality*

##### *Determination by spectrophotometry*

Nucleic acid samples were diluted either at 1:20 by addition of 25 µl of resuspended nucleic acid in 475 µl of dH<sub>2</sub>O, or 1:100 by addition of 5 µl of the resuspended nucleic acid in 495 µl of dH<sub>2</sub>O. Optical density readings were taken at 260 nm and 280 nm, using dH<sub>2</sub>O as a blank. An optical density reading of 1.0 at 260 nm corresponds to an approximate nucleic acid concentration of 50 µg/ml for double stranded (ds) DNA and 40 µg/ml for single stranded (ss) RNA. The ratio of the readings taken at 260 nm and 280 nm ( $OD_{260}/OD_{280}$ ) was used to give an estimate of the purity of the nucleic acid. Pure preparations of DNA and RNA have an  $OD_{260}/OD_{280}$  of 1.8 and 2.0 respectively; a lower value suggests contamination, typically with protein or phenol.

##### *Estimation of double stranded DNA concentration and quality by agarose gel electrophoresis*

This method was used when there were insufficient amounts of dsDNA for spectrophotometry, or when purity of a particular DNA fragment needed to be

investigated. The concentration of dsDNA was determined by running the sample on an agarose gel electrophoresis as described below, and the intensity of the fluorescence of the unknown DNA was compared to that of a known quantity of the appropriate size marker ( $\phi$ X174 RF DNA/HaeIII fragments or  $\lambda$  DNA/HindIII fragments) following staining with ethidium bromide and visualisation by UV transillumination. Smearing of a DNA band indicated degradation of the sample and resulted in exclusion of that DNA from further analysis.

### ***Restriction endonuclease digestion***

Typically, 1  $\mu$ g of miniprep DNA was digested in a 20  $\mu$ l reaction mix containing the appropriate buffer for the restriction endonuclease(s), 2 mM spermidine and 10 units of the desired restriction enzyme. The reactions were incubated at 37 °C for a minimum of 1 hour. Where the isolation of restriction fragments was required, larger quantities of DNA, generally 20  $\mu$ g, were digested with 100 units of enzyme, and the reaction volume and other components increased proportionally. In this case, reactions were incubated at 37 °C for 12 to 18 hours.

### ***Electrophoresis of DNA***

#### ***Agarose gel electrophoresis***

DNA fragments of 1.0 kb to 10 kb were separated by agarose gel electrophoresis using a Biomax QS710 Horizontal Unit (Scientific Imaging Systems – Eastman Kodak Company, New York, USA) or Sub-Cell GT Submerged Horizontal Agarose Gel Electrophoresis Systems (Bio-Rad, Hertfordshire, UK). Typically, 0.6 g to 0.9 g or 2 g to 3 g agarose was added to 60 ml or 200 ml of 1x TBE or 1x TAE buffer, respectively; melted in a microwave for 1 minute to 2 minutes and mixed to produce a 1% to 1.5% gel (w/v). Once the gel mix had cooled to 55 °C, the gel was poured into a 100 mm x 65 mm or 150 mm x 200 mm gel support in its casting tray and an appropriate gel comb (eight, twelve or twenty wells) inserted. The gel was allowed to solidify before transferring to an electrophoresis tank; the gel was immersed in 1x TBE or 1x TAE buffer and the comb carefully removed. Gel loading buffer (Section 2.1.11) was added to DNA samples at a final concentration of 1x and samples loaded into the wells using a micropipette. Known concentrations of DNA size markers were

prepared similarly and run alongside the samples in order to gauge both product size and yield. Gels were run at 100 volts for 60 to 120 minutes or overnight at 20 volts, then removed from the gel apparatus and stained in buffer solution containing 0.5 µg/ml ethidium bromide for 30 minutes. Following destaining for 30 minutes in dH<sub>2</sub>O, gels were visualised on a medium wave UV transilluminator (UVP Inc. and Geneflash Syngene Bio Imaging) and photographed using black and white Video Graphic Printer Sony UP-895MD (Fisher Scientific- Synoptics Ltd, UK).

### *Polyacrylamide gel electrophoresis*

In order to separate, visualise and determine the size of DNA fragments under approximately 1.2 kb (including PCR products and small products of restriction digests), non-denaturing polyacrylamide gel electrophoresis was employed. Glass plates of 16 cm x 16 cm size were assembled with a 0.75 mm spacer in a casting stand (Atta, Atto Corporation). 4% gel was prepared with 3.3 ml of 30%: 1.579% acrylamide/bisacrylamide solution (Severn Biotech Ltd, Kidderminster, UK), 2.5 ml of 10x TBE buffer (Section 2.1.11) and ddH<sub>2</sub>O added to 25 ml total volume. Following the addition of 32 µl of TEMED (Sigma) and 200 µl 10% APS (Section 2.1.11), the gel solution was poured between the assembled gel plates and a comb (12 wells) inserted. After polymerisation, the gel plates were removed from the casting apparatus, the spacer removed and the plates transferred to the gel electrophoresis apparatus. The apparatus was filled with 1x TBE buffer, the gel comb removed and the wells flushed with buffer. Samples were prepared as described above and loaded onto the gel using 0.4 mm flat ended gel loading tips; φX174 RF DNA/Hae III fragments and Low DNA Mass Ladder were loaded as molecular size standards. Gels were electrophoresed at 250 volts for 45 minutes then removed and stained in buffer solution containing 0.5 µg/ml ethidium bromide for 30 minutes. Following destaining for 30 minutes in dH<sub>2</sub>O, gels were visualised on a UV transilluminator (UVP Inc. and Geneflash Syngene Bio Imaging) and photographed using black and white Video Graphic Printer Sony UP-895MD (Fisher Scientific- Synoptics Ltd, UK).

### ***Purification of restriction enzyme fragments***

When purification of DNA fragments was required for construction of recombinant plasmids, the DNA was purified from agarose gels using the QIAquick<sup>®</sup> Gel Extraction Kit (Qiagen). DNA fragments of interest were cut from an agarose gel using a clean scalpel blade and the protocol was then performed according to the manufacturer's instructions.

### ***Cloning of hybrid DNA molecules***

#### ***Ligation of vector and insert DNA***

The quantity of vector and insert DNA used in ligations was calculated to produce a molar ratio between 1:1 and 1:5 using the equation:

$$\frac{X \text{ ng of vector} \times Y \text{ kb of insert}}{Z \text{ kb of vector}} \times \text{insert : vector ratio} = \text{ng of insert required}$$

Vector DNA was linearised using appropriate restriction enzyme(s). To prevent recircularisation of the vector DNA following digestion with a single enzyme, both 5'-phosphate groups were hydrolysed twice with 0.01 units of CIAP (Promega) at 37 °C for 30 minutes. Generally, 50 ng of vector DNA and a three fold molar excess of insert DNA were ligated, with an appropriate volume of ligation buffer and 4 units T4 DNA ligase (Invitrogen), in a final volume of 20 µl. A control ligation, omitting insert DNA was generally set up in parallel to the above, in order to check for 'background' when performing subsequent bacterial transformations. Ligations of DNA fragments generated by PCR were carried out according to the manufacturer's instructions with the TOPO TA Cloning Kit (Invitrogen Life Technologies). TOPO TA Cloning provides a highly efficient, 5 minute, one step cloning strategy for the direct insertion of *Taq* polymerase-amplified PCR products into a plasmid vector. *Taq*-polymerase has a nontemplate-dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3' ends of PCR products. The linearized vector supplied in the kit has single, overhanging 3'-deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector. Ligations were carried out overnight at 14 °C except when using the TA-TOPO Cloning kit where ligations

were carried out at room temperature for 5 minutes. Stored thereafter at  $-20\text{ }^{\circ}\text{C}$  if not used immediately.

### *Transformation of bacteria with plasmid DNA*

All transformations included a positive control using supercoiled plasmid DNA (eg pUC19 10 pg/ml) to monitor the efficiency of the transforming bacteria and a ligation control (vector DNA, no insert) to monitor the background rate of transformation due to recircularisation of vector DNA.

### *Transformation of commercially prepared competent bacteria*

2  $\mu\text{l}$  of a 1:5 dilution of ligation mix, 2  $\mu\text{l}$  of neat ligation mix or 2  $\mu\text{l}$  of TA ligation mix were added to 50  $\mu\text{l}$  of One Shot<sup>®</sup> TOP10 Chemically Competent *E. coli* cells (Invitrogen Life Technologies). Cells and DNA were mixed by gentle tapping to avoid damage to the bacterial cells. In all cases the reactions were then incubated on ice for 30 minutes. Cells transformed with ligation mix and TA ligation mix were subsequently heat shocked at  $42\text{ }^{\circ}\text{C}$  for exactly 40 seconds and 30 seconds respectively, and then returned to ice. 250  $\mu\text{l}$  of SOC medium (Invitrogen) was then added and incubated at  $37\text{ }^{\circ}\text{C}$  for 1 hour with shaking at 220 rpm. Cells (50  $\mu\text{l}$  and 200  $\mu\text{l}$ ) were plated onto LB plates containing 100  $\mu\text{g}/\text{ml}$  ampicillin and 1.6 mg X-Gal (Section 2.1.11). Plates were incubated overnight at  $37\text{ }^{\circ}\text{C}$ .

### *Preparation of freshly competent bacteria and subsequent transformation*

*E. coli* host strain: BL21 cells (Amersham Pharmacia Biotech), used for procedures involving the pGEX plasmid were made competent using a modification of the procedure described by Chung *et al.* (252). A single colony was picked from an LB agar plate after overnight growth and used to inoculate 50 ml of LB broth. Bacteria were grown at  $37\text{ }^{\circ}\text{C}$ , with shaking at 250 rpm, to an  $\text{OD}_{600}$  of 0.4 to 0.5. The cells were then sedimented at 2500 g for 15 minutes at  $4\text{ }^{\circ}\text{C}$ , the supernatant discarded, and then resuspended in 5 ml of ice-cold TSS buffer. Cells were kept on ice until used for transformations, within 2 to 3 hours of preparation. Then, 1 ng or 5 ng of pGEX plasmid DNA was added to 1 ml of freshly competent *E. coli* host strain: BL21 cells. The mixture was incubated on ice for 45 minutes followed by heat shock at  $42\text{ }^{\circ}\text{C}$  for

exactly 2 minutes and returned to ice. 100  $\mu$ l of the transformed cells were added to 900  $\mu$ l of LBG medium (Section 2.1.11) and incubated with shaking at 250 rpm for 60 minutes. Cells (50  $\mu$ l and 200  $\mu$ l) were plated onto LBAG plates (LB plates containing 100  $\mu$ g/ml ampicillin; 20 mM glucose) and were incubated overnight at 37°C.

#### *Screening of transformants for desired recombinant plasmids*

All plasmid strains used in this project conferred ampicillin resistance upon host bacteria, allowing selection and maintenance of transformed bacteria with ampicillin supplemented media.

The pCR<sup>®</sup> II plasmid contains genes encoding the *lacZ $\alpha$*  fragment of  $\beta$ -galactosidase and the *lac* promoter and is thus capable of complementation with the  $\phi$  fragment encoded by the *E.coli* host strains, giving active  $\beta$ -galactosidase. The incorporation of X-gal into LB agar plates allows the selection of transformants based on blue-white screening. Disruption of *lacZ $\alpha$*  expression occurs with the cloning of fragments into the multiple cloning site of this vector, hence recombinants with plasmid containing insert DNA appear white whilst non-recombinants, expressing a functional  $\beta$ -galactosidase, appear blue. White colonies were picked from plates containing X-Gal and used to inoculate 5 ml of LB medium containing ampicillin (100  $\mu$ g/ml) and grown overnight at 37 °C with shaking. Alternatively, when large number of colonies need to be screened to identify recombinants, colony hybridisation analysis was performed using the method first described by Grunstein and Hogness (253). Briefly, white or light blue colonies were picked, spotted not only on nitrocellulose filters over LB agar plates containing 100  $\mu$ g/ml ampicillin but also onto reference plates, and then were incubated overnight at 37 °C. After that, the reference plate was placed at 4 °C and the filter upon which single bacterial clones had grown was treated as for Southern blot (see 2.2.6) in order to allow identification of recombinants clones by hybridisation with an appropriate probe (see below). Following identification of recombinant clones, the relevant colony was picked from the reference plate and expanded for further analysis.

Recombinant plasmid DNA, isolated by small-scale plasmid purification as described above, was subjected to restriction digest with the appropriate enzyme(s), and the

resulting products of digestion run on a polyacrylamide or agarose gel, stained with ethidium bromide and visualized by UV illumination. Bacteria with plasmids containing inserts of the desired size were stored as glycerol stocks as detailed above.

### **2.2.3 Preparation of total RNA**

The unstable nature of isolated RNA and the necessity for very pure isolates for downstream applications required that all materials used in RNA preparation be totally free from ribonuclease (RNase) activity. RNase is a very stable, ubiquitous enzyme that degrades RNA requiring no cofactors for function. All plasticware used were RNase free and all solutions were prepared using DEPC treated dH<sub>2</sub>O. Gloves were worn and changed frequently.

#### ***RNA extraction using RNA-Bee™***

Various methods have been described for the isolation of undegraded RNA, and progress in the field have led to the development of single step methods for the procedure (254, 255). RNA-Bee™ (Biogenesis) is a monophasic solution containing phenol and guanidine thiocyanate. This RNA isolation reagent is a complete and ready-to-use reagent for isolation of total RNA from samples of human, animal, plant, bacterial and viral origin. The samples used for RNA isolation in this project were cell pellets harvested from tissue culture experiments. These were homogenised directly in RNA-Bee™ by simply pipetting vigorously several times. 2 ml of RNA-Bee™ was used per 10<sup>7</sup> cells. The homogenate was then incubated on ice for 5 minutes to allow dissociation of nucleoproteins from the nucleic acids. Chloroform was then added (0.2x starting volume), and thorough mixing was carried out by shaking the tube vigorously for approximately 30 seconds. The mixture was then centrifuged at 12000 g for 15 minutes at 4 °C to allow separation of the mixture into three distinct phases, an upper aqueous phase containing the RNA, a semi-solid interphase containing most of the DNA, and a lower organic phase. The upper aqueous phase was removed by careful pipetting and transferred to a clean 1.5 ml RNase-free microfuge tube. This was followed by addition of an equal volume of isopropanol, mixing, and incubation of the sample on ice for 1 hour. The RNA was then pelleted by centrifugation at 12000 g for 15 minutes at 4 °C. The supernatant was

decanted, and a wash step using 75% ethanol was carried out, followed by centrifugation at 7500 g for 5 minutes. The supernatant was decanted and the pellets air dried for no longer than 10 minutes (complete drying of the pellet makes resuspension difficult) and resuspended in 50  $\mu$ l of DEPC treated dH<sub>2</sub>O. RNA was stored at -70 °C and its concentration determined either by spectrophotometry (Section 2.2.2) or using the RNA 6000 Nano LabChip Kit (see below).

### ***Assessment of RNA***

#### *Assessment of RNA using RNA 6000 Nano LabChip Kit*

RNA 6000 Nano LabChip Kit (Agilent Technologies) allows the characterisation of total RNA samples and their concentration, with only nanograms of sample. Each RNA LabChip contains an interconnected set of microchannels that is used for separation of nucleic acids fragments based on their size as they are driven through it electrophoretically. For determination of RNA concentration, total RNA in sample must be between 25 ng/ $\mu$ l to 250 ng/ $\mu$ l. The manufacturer's instructions were followed using 1  $\mu$ l of RNA sample either neat or 1:10 dilution with DEPC treated dH<sub>2</sub>O. All procedures were performed using RNase-free microcentrifuge tubes and pipette tips.

#### *Assessment of RNA using agarose gel electrophoresis*

1 $\mu$ g of each RNA sample was run on a 1% agarose TBE gel. Assessment of RNA quality was carried out by checking the integrity of the 18S and 28S ribosomal subunit bands, and examining their rate of migration in comparison to a  $\lambda$  DNA/HindIII fragments molecular weight standard (Invitrogen Life Technologies).

### **2.2.4 Amplification of DNA by polymerase chain reaction**

The polymerase chain reaction (PCR) is a technique that allows the amplification of DNA sequences between two specific oligonucleotides using very little starting material.

### ***Primer design***

Primers were designed manually according to a protocol defined by QIAGEN (256). Briefly, the guidelines for designing PCR primers followed are in terms of primer sequence, to avoid running 3 or more G or C bases at the 3'-terminal sequence since it may stabilize non-specific annealing of the primer. A thymidine at the 3' end should also be avoided, since it is more prone to mispriming than other nucleotides (257). Mismatches between template and primer DNA at the 3' end should be avoided, as DNA polymerase will be acting in a 5' to 3' direction, thus it is best if these are located close to the 5' end of the primer. Moreover the formation of primer-dimers can be avoided by not using primers with complementary sequences in the 3' region; design of primers in this respect was aided by the Primer express software program (Version 2.0 – Applied Biosystems).

Primer length should be between 18 and 30 nucleotides; any increase in size beyond this is unlikely to improve primer specificity significantly and this size of primer is sufficient for sequence as complex as the human genome. Other potential features of the primer sequence, such as restriction enzyme sites or promoter elements should be engineered onto the 5' end of the primer, because bases located at the 5' end of the primer are less critical for primer annealing. Such additions can have a detrimental effect on primer specificity at low temperatures, and so are best used when amplifying from a single template vector.

Primer GC content of 40 - 60% is recommended.

The optimal annealing temperature was obtained from the estimated melting temperature ( $T_m = 2 \times (A+T) + 4 \times (C+G)$ ).

### ***Preparation of PCR reactions***

The sensitive nature of PCR means that very stringent steps must be taken if contamination is to be avoided. First, physical separation of the PCR area from bench space used for other work is recommended. In the case of PCRs carried out in this project, a dedicated PCR suite isolated from the main laboratory was used (PCR 6 Vertical Laminar Airflow Cabinet with UV Sterilisation). In addition, a set of instruments used to aliquot reagents (micropipettes and their tips) were kept within the PCR suite and restricted to PCR use. Filter tip pipette tips were used to reduce the

chance of carry over of reaction components from one tube to the next and master mixes of reagents were used whenever possible to reduce the number of pipetting steps required per reaction. Reaction components including primers were aliquoted prior to use and stored at  $-20\text{ }^{\circ}\text{C}$ .

### ***Reaction conditions***

Standard PCRs were carried out using 2x ReddyMix<sup>TM</sup> PCR Master Mix (ABgene), which contains the *Taq* DNA polymerase isolated from *Thermus aquaticus*. Standard reaction mixes were set up in either 50  $\mu\text{l}$  or 25  $\mu\text{l}$  in 0.2 ml reaction tubes containing 0.2 mM each dNTP, 75 mM Tris-HCl, pH 8.8, 20 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1.5 mM  $\text{MgCl}_2$ , 0.01% (v/v) Tween 20, 1.25 units *Taq* DNA polymerase, 2.0  $\mu\text{M}$  each primer and 10 ng to 1  $\mu\text{g}$  of template (genomic DNA or cDNA). Thermal cycling were carried out in two programmable thermal cyclers; PCR Express Thermal Cycler (Hybaid Limited, Middlesex, UK) and PTC-220 DNA Engine Dyad<sup>®</sup> Cycler (MJ Research Incorporated, Massachusetts, USA). The exact parameters of the reactions varied, but generally consisted of an initial denaturation at  $94\text{ }^{\circ}\text{C}$  for 5 minutes (genomic DNA template), followed by 20 to 35 cycles of: denaturation at  $94\text{ }^{\circ}\text{C}$ , for 1 minute; annealing at  $50\text{ }^{\circ}\text{C}$  to  $60\text{ }^{\circ}\text{C}$  for 1 minute; and extension at  $72\text{ }^{\circ}\text{C}$  for 1 minute. Reaction products were separated by polyacrylamide or agarose gel electrophoresis as detailed in 2.2.2, using 5  $\mu\text{l}$  or 10  $\mu\text{l}$  of reaction product per well.

### ***Purification and assessment of PCR products***

Single PCR products were excised either from polyacrylamide or from agarose gels. Small PCR products were excised from 4% polyacrylamide gel and eluted using acrylamide elution buffer (Section 2.1.11) at  $37\text{ }^{\circ}\text{C}$  overnight. Eluate solution (13000 rpm for 1 minute) was precipitated using 2.5x volume of 100% ethanol at  $-70\text{ }^{\circ}\text{C}$  for 30 minutes. Pelleted DNA (13000 rpm for 20 minutes) was dried by centrifugation under vacuum. The pellet was then resuspended in 25  $\mu\text{l}$  of ddH<sub>2</sub>O and highly purified DNA was obtained using the QIAquick<sup>®</sup> PCR purification Kit (Qiagen) according to the manufacturer's protocol. Long PCR products were excised from agarose gel and highly purified DNA was obtained using the QIAquick<sup>®</sup> Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. Then, 1/10 of

eluted DNA from polyacrylamide and agarose gels were assessed by 4% polyacrylamide and 1% TBE agarose gels electrophoresis against a  $\phi$ X174 RF DNA/HaeIII fragments and a  $\lambda$  DNA/HindIII fragments molecular weight standard (Invitrogen Life Technologies), respectively and an estimation of quantity obtained.

### ***First strand DNA synthesis for reverse transcriptase (RT)-PCR***

Analysis of gene expression requires accurate determination of mRNA levels, but as PCR is based on amplification of DNA, the process of amplifying RNA sequence requires an initial step of conversion of the RNA to cDNA by reverse transcription. In order to maximise the likelihood of obtaining full-length cDNA copies of mRNA, a commercial cDNA synthesis kit was employed (ImProm-II<sup>TM</sup> Reverse Transcription System, Promega).

Typically, 1  $\mu$ g of total RNA was placed in a RNase free microcentrifuge tube with 0.5  $\mu$ g Oligo(dT)<sub>15</sub> primer and brought to 5  $\mu$ l with nuclease-free water. The mixture was denatured at 70°C for 5 minutes then chilled on ice. A reverse transcription reaction mix was assembled on ice containing 1x RT buffer, 5 mM MgCl<sub>2</sub>, 0.5 mM dNTP mix, 1 U/ $\mu$ l Recombinant Rnasin<sup>®</sup> Ribonuclease Inhibitor, and 1  $\mu$ l Improm-II<sup>TM</sup> RT. The template-primer combination was added to the reaction mix on ice. Following an initial annealing at 25 °C for 5 minutes, the reaction was incubated at 42 °C for up to 1 hour. The reverse transcriptase was thermally inactivated at 70 °C for 15 minutes prior to amplification by PCR using primers specific for the gene of interest. Freshly synthesised cDNA was kept at -20 °C for long-term storage.

## **2.2.5 DNA sequence analysis**

### ***Automated sequencing***

#### ***Sample preparation***

DNA sequencing was carried out using the Big Dye<sup>®</sup> Terminator Version 1.1 Cycle Sequencing Kit (Applied Biosystems). Generally, 200 ng to 500 ng of highly purified DNA template was added into the cycle sequencing reaction containing 3.2 pmol of sequencing primers (Table 2-2), 0.5x sequencing buffer (80 mM Tris-HCl, 2 mM MgCl<sub>2</sub>) and 4  $\mu$ l of cycle sequencing mix (Big Dye<sup>®</sup> Terminator Version 1.1 Cycle

Sequencing Kit, Applied Biosystems). Samples were prepared in the PCR Express Thermal Cycler (Hybaid Limited) incorporating 25 cycles of amplification, each cycle consisting of a denaturing step at 96 °C for 30 seconds followed by an annealing temperature of 50 °C for 15 seconds, and an elongation step of 60 °C for 4 minutes. Sequencing products were purified by precipitation using ethanol (80%) at room temperature for 15 minutes. Pelleted DNA (14000 rpm for 20 minutes) was washed in ethanol (70%) and repelleted (14000 rpm for 10 minutes) before all ethanol was removed and the pellet dried at 90 °C for 1 minute. The pellet was then resuspended in 25 µl of Hi-Di Formamide (Applied Biosystems).

#### *Sample sequencing and evaluation*

Samples were loaded into 96 well plates and run on the ABI PRISM<sup>®</sup> 3100 Genetic Analyzer (PE Applied Biosystems, UK) under standard sequencing conditions for generation of automated sequence data. Electrophoresis on the ABI PRISM<sup>®</sup> 3100 Genetic Analyzer was performed using 36 cm and 80 cm capillaries which reads up to 500 nucleotides and 950 nucleotides respectively, with 98.5% base calling accuracy with less than 2% ambiguity. A series of different computational software were utilised for sequence analysis including the 'Blast' search engine within the NCBI database (<http://www.ncbi.nlm.nih.gov/>), and the 'Blat' search engine within the UCSC Genome Browser Home (<http://www.genome.ucsc.edu/>).

**Table 2-2. Sequencing primers.**

Primer name	Primer sequence	Description
Universal M13 (-20) Forward	5'd[GTAAAACGACGGCCAG]3'	Forward cycle sequencing primer from pCR <sup>®</sup> II-TOPO vector
Universal M13 Reverse	5'd[CAGGAAACAGCTATGAC]3'	Reverse cycle sequencing primer from pCR <sup>®</sup> II-TOPO vector
pGEX 5'	5'd[GGGCTGGCAAGCCACGTTTGGTG]3'	Forward sequencing primer from pGEX series
pGEX 3'	5'd[CCGGGAGCTGCATGTGTCAGAGG]3'	Reverse sequencing primer from pGEX series

## 2.2.6 DNA and RNA hybridisation analysis

### *Southern blot transfer of DNA*

This was carried out essentially as described by Southern (258). Generally, 20 µg of high molecular weight genomic DNA was digested with 100 units of the desired restriction enzyme, in the appropriate restriction enzyme buffer and 2 mM spermidine. The reactions were incubated overnight at 37 °C. Digestion products were then precipitated with half volume of 7.5 M ammonium acetate and two volumes of 100% ethanol. Pelleted DNA (14000 rpm for 20 minutes) was washed with 70% ethanol and repelleted (14000 rpm for 20 minutes) before all ethanol was removed and the pellet dried by centrifugation under vacuum for 10 minutes and resuspended in TE pH 8.0 for several hours at 37 °C. DNA samples were electrophoresed as described in the Section 2.2.2 on a 0.8% agarose gel in 1x TAE buffer overnight at 23 volts. 5 x 10<sup>3</sup> counts per minute of <sup>32</sup>P-end labelled marker (1 µg of λ DNA/HindIII fragments) together with 1 µg of unlabeled marker were also loaded on the gel. Following electrophoresis, the gels were stained in buffer solution containing 0.5 µg/ml ethidium bromide for 30 minutes, viewed by UV transillumination and photographed as described above; submerged in denaturation buffer for 30 minutes, followed by submersion in neutralisation buffer for 30 minutes. The gel was equilibrated in 20x SSC and the DNA was transferred onto Hybond<sup>TM</sup>-N nylon membrane (Amersham Biosciences) by capillary blotting overnight in 20x SSC as described by Sambrook *et al.*, 1989 (249). The membrane was removed and the DNA crosslinked to the membrane (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation).

### *Northern blot transfer of RNA*

20 µg of total cellular RNA was lyophilised (VR-1 Hetovac, Heto Laboratory Equipment), then resuspended in 20 µl of RNA loading buffer and denatured for 15 minutes at 65°C. Subsequently, 5 µl of RNA running dye was added and RNAs were separated on 1% agarose gel containing 2.2 M formaldehyde. Electrophoresis was carried out for 3 to 5 hours at 100 volts in 1x MOPS, with buffer recirculation (302S, Watson Marlow Pumps). Moreover, 5 µg of RNA ladder (Invitrogen Life Technologies), treated in the same way as the RNA samples, were used as molecular

weight markers. Following electrophoresis, the gel was stained in ddH<sub>2</sub>O containing 3 µg/ ml ethidium bromide for 10 minutes, then destained twice in ddH<sub>2</sub>O for 30 minutes, viewed by UV transillumination and photographed to confirm the presence and integrity of RNA. The marker lanes were removed using a clean scalpel, destained overnight in ddH<sub>2</sub>O and photographed. The rest of the gel was equilibrated in 10x SSC for 30 minutes, and then transferred overnight onto Hybond™-N nylon membrane (Amersham Biosciences) by capillary blotting in 10x SSC as described by Sambrook *et al.*, 1989 (249). The membrane was removed and the RNA crosslinked to the membrane (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation).

### ***Preparation of radiolabelled DNA probes***

PCR or restriction fragment DNA probes were gel purified from non-specific or vector sequence (Section 2.2.4) and radioactively labelled using the RadPrime DNA Labeling System (Invitrogen Life Technologies), and  $\alpha$ -<sup>32</sup>P dCTP, specific activity 3000 Ci/mmol (Amersham Life Science). Briefly, 25 ng of DNA fragment was made up to 21 µl with ddH<sub>2</sub>O, boiled for 5 minutes and chilled on ice. Then, the heat denatured DNA was radiolabeled using 20 µl of 2.5x RadPrime Buffer (Section 2.1.11), 10 pmol/µl of each three nucleoside triphosphates (dATP, dGTP, and dTTP), 1 µl of Klenow fragment (Large Fragment of DNA Polymerase I) and 1.85 Mbq of  $\alpha$ -<sup>32</sup>P dCTP in a final volume of 50 µl and incubated at 37°C for 10 minutes. Unincorporated <sup>32</sup>P- labelled nucleotides were removed by gel-filtration through Sephadex-G50 beads (Nick™Columns, Pharmacia Biotech) and labelled fragments were eluted in 400 µl of TE buffer and specific activity was determined by liquid scintillation analysis (MicroBeta Jet Liquid Scintillation and Luminescence Counter, Perkin Elmer Life Sciences). Probes were labelled to specific activities of 10<sup>8</sup> cpm/µg or greater. ~

### ***Hybridisation of labelled probes to membrane bound nucleic acids***

Membranes which had been pre-wetted in 2x SSC and rolled into Hybaid hybridisation bottles were pre-hybridised in 15 ml of pre-warmed hybridisation buffer (Rapid-hyb buffer, Amersham Pharmacia Biotech) at 65 °C for 1 hour in a hybridisation oven (Mini Oven MKII, Hybaid Limited) with continuous rotation.

Freshly, radiolabelled dsDNA probe was boiled for 5 minutes and chilled on ice prior to addition to the pre-hybridisation solution:  $10^6$  cpm/ml was used for genomic Southern and northern blots and  $10^5$  cpm/ml was used for plasmid DNA Southern blots. Filters were hybridised at 65 °C for 2 hours in a hybridisation oven with continuous rotation. After rinsing briefly with 2x SSC, the membrane was washed for 20 minutes with three changes of 0.1x SSC, 0.5% SDS at 60 °C, unless otherwise stated. Membranes were then sealed in polythene and exposed either to X-ray film (Hyperfilm MP, Amersham Biosciences) at -70 °C or the Molecular Dynamics Storage phosphor screen (Molecular Dynamics, Amersham Pharmacia Biotech) at room temperature. In order to reprobe with an alternative probe, filters were stripped of existing probe by continuous shaking in boiling ddH<sub>2</sub>O containing 0.1% SDS until the solution has cooled to room temperature.

## **2.2.7 Protein analysis**

### ***Protein extraction***

Whole cell extracts, from cells grown in tissue culture were prepared for western blot analysis essentially according to the method of Marais *et al.*, (259). All procedures were performed at 4 °C. Typically, cells were washed twice with ice-cold PBS, drained thoroughly, lysed in 100 µl working whole cell lysis buffer (Section 2.1.11) per  $10^6$  cells or 60 mm dish and rotated for 10 minutes. The crude extract was cleared in a bench top centrifuge at 14000 rpm for 30 minutes and the supernatant transferred to a fresh pre-cooled eppendorf. Extracts were quantitated as described below, and stored at -70 °C.

### ***Estimation of protein concentration***

In order to estimate protein yields, a Bio-Rad Protein Assay (Bio-Rad), based on the method of Bradford was used, in which a differential colour change of a dye occurs in response to various concentrations of protein (260).

A protein standard (bovine serum albumin, BSA) was diluted with ddH<sub>2</sub>O to concentrations ranging from 0.025 to 0.8 mg/ml to allow the production of a standard curve each time the assay was performed. Aliquots of protein for quantification were diluted in order to give a concentration in the above range. The dye reagent was

prepared by addition of 4 volumes of ddH<sub>2</sub>O to 1 volume of Dye Reagent Concentrate (Bio-Rad) and filtration through a Whatman #1 to remove particulates. Samples were diluted 1/500 in diluted dye reagent, and incubated in sterile polystyrene cuvettes at room temperature for 5 minutes to 1 hour. OD<sub>595</sub> was determined against the reagent blank. The absorbance of the protein standards was used to construct a standard curve from which an approximate concentration of the unknown samples could be read.

### ***SDS - PAGE of proteins***

The separation and analysis of proteins was facilitated by one dimensional denaturing discontinuous gel electrophoresis, as originally described by Laemmli (1970) (261).

Mini SDS-PAGE gels (8.0 cm x 7.3 cm) were formed and run using the Mini-PROTEAN II electrophoresis system (Bio-Rad) as recommended by the manufacturer. Glass plates were assembled with 0.75 mm spacers in a casting stand. The separating gel was poured to a depth of approximately 5 cm; consisting of 4 ml 30%: 0.8% w/v acrylamide/bisacrylamide (giving a 12% gel), 2.05 ml ddH<sub>2</sub>O, 3.75 ml 1 M Tris-HCl (pH 8.8), 100 µl 10% SDS, 100 µl 10% APS and 4 µl TEMED. This was overlain with ddH<sub>2</sub>O-saturated butanol and allowed to polymerise. The butanol was then poured off, the surface of the separating gel rinsed with ddH<sub>2</sub>O and the stacking gel poured; consisting of 830 µl 30%: 0.8% w/v acrylamide/bisacrylamide, 3.4 ml dH<sub>2</sub>O, 630 µl 1 M Tris-HCl (pH 6.8), 50 µl 10% SDS, 50 µl 10% APS and 5 µl TEMED. A 10 well comb was inserted and the gel allowed polymerising. Gels were run in 1x running buffer. Equal volumes of protein sample and SDS-PAGE Protein 2x Sample Loading Buffer (Section 2.1.11) were vortexed, heated to 95 °C for 5 minutes and chilled on ice prior to loading on gel. Wells were flushed and samples loaded using 0.2 mm flat-ended gel loading tips (Sorenson Bioscience Inc). Moreover, 7 µl of a protein molecular weight standard (Precision Plus Protein™ Standards, Bio-Rad) was loaded in one or both outer wells to allow size estimation of sample proteins.

Standard SDS-PAGE gels (13.8 cm x 13 cm) were formed and run using the AE-6200 Slab EP Chamber, and AE-6220 Slab EP Chamber, for two gels (Atto Incorporation). The separating gel was poured to a depth of approximately 10 cm; consisting of 16.7

ml 30%: 0.8% w/v acrylamide/bisacrylamide (giving a 10% gel), 13.4 ml dH<sub>2</sub>O, 18.75 ml 1 M Tris-HCl (pH 8.8), 500 µl 10% SDS, 375 µl 10% APS and 20 µl TEMED. This was overlain with ddH<sub>2</sub>O-saturated butanol and allowed to polymerise. The butanol was then poured off and the surface of the separating gel rinsed with ddH<sub>2</sub>O. Excess water was removed with a sheet of 3mm Whatman filter paper, and the stacking gel poured; consisting of 3.4 ml 30%: 0.8% w/v acrylamide/bisacrylamide, 13.6 ml ddH<sub>2</sub>O, 2.5 ml 1 M Tris-HCl (pH 6.8), 200 µl 10% SDS, 200 µl 10% APS and 20 µl TEMED. A 14 well comb was inserted and the gel allowed polymerising. Gels were run in 1x running buffer, and samples were treated as for mini SDS-PAGE gels (see above). Moreover, 14 µl of a protein molecular weight standard (Precision Plus Protein™ Standards, Bio-Rad) was loaded in one or both outer wells to allow estimation of the size of sample proteins.

Mini SDS-PAGE gels were electrophoresed at 110 volts to 150 volts for 55 to 90 minutes, and standard SDS-PAGE gels at 230 volts to 250 volts for 2 hours until the bromophenol blue dye reached the bottom of the separating gel. The gel was then removed from the glass plates, the stacking gel discarded, and the protein bands detected either by staining with SDS-PAGE Protein Gel Fix-Stain Solution (Section 2.1.11) or by immunodetection as described below. Visualisation of protein bands with Coomassie Brilliant Blue involved staining for one hour in four to five gel volumes of protein fix-stain solution followed by destaining for approximately 4 to 24 hours in destain solution (Section 2.1.11), with four to five changes of solution. Gels were then removed and preserved by drying for two to six hours, sandwiched between prewetted cellulose film (Amersham Bioscience), in a gel drying apparatus (Easy Breeze Air Gel Dryer - Hoefer Scientific Instruments).

### ***Electroblotting***

Electroblotting was necessary for the transfer of protein after polyacrylamide resolving gel electrophoresis to a nitrocellulose membrane.

Following electrophoresis:

Mini SDS-PAGE gels were removed from the glass plates, and the resolving gel equilibrated in 1x transfer buffer with 15% methanol (Section 2.1.11) for 5 minutes at

room temperature. Two Scotch-Brite pads and a further 4 pieces of 3MM Whatmann filter paper of the same size as the gel were presoaked for 10 minutes in 1x transfer buffer, carefully removing air pockets from the Scotch-Brite pads by repeated squeezing and agitation. These pieces were necessary for building up a sandwich around the gel for electroblotting. One piece of Hybond ECL nitrocellulose membrane (Amersham Biosciences) was presoaked for 1 minute in 1x transfer buffer. Electroblotting experiments were carried out using the Mini-PROTEAN® II cell (Bio-Rad). The electroblot apparatus contains a gel holder consisting of hinged white and black clamps on which the components of the sandwich are built up on the black clamp as follows; Scotch Brite pad, 2x filter papers, gel, nitrocellulose membrane, followed by the remaining 2 pieces of filter paper and the second Scotch Brite pad. At each stage any air bubbles are carefully removed from between the layers. The gel holder is then closed and placed in the transfer cell with the black clamp facing the cathode. An ice block is included to prevent overheating of the transfer buffer. The cell is then filled with 1x transfer buffer and electroblotted for 45 minutes at 80 volts in the cold room (4 °C). Molecular weight marker positions were highlighted and the membrane stored at 4 °C.

Standard SDS-PAGE gels were removed from the glass plates, and the resolving gel equilibrated in 1x dry blot buffer with 20% methanol (Section 2.1.11) for 15 minutes at room temperature. The gel was covered in pre-wetted nitrocellulose membrane (1x dry blot buffer) and sandwiched between 12 pieces of Whatman 3MM rinsed in 1x dry blot buffer. The sandwich was turned over and placed on the semi dry blotter (Trans-blot® SD Semi-Dry Transfer Cell, Bio-Rad). Air bubbles were removed and transfer effected at 20 volts for 2 hours. Molecular weight marker positions were highlighted and the membrane stored at 4 °C.

### ***ECL detection***

The detection of proteins by immunoblotting (western blotting) is a rapid and sensitive technique that exploits the inherent specificity of antigen recognition by antibodies (262).

Hybond-ECL nitrocellulose membrane (Amersham Biosciences) was immersed in 1x TBST buffer (Section 2.1.11) containing 5% low fat dried milk (Marvel - Premier

Beverages, Stafford, UK) overnight at 4 °C, to block non-specific binding sites, rinsed briefly with 1x TBST buffer, washed once for 15 minutes, and then twice for 10 minutes, with shaking at room temperature in 1x TBST buffer. The membrane was then incubated with the primary antibody, at a pre-determined dilution in blocking buffer (either 5% low fat dried milk, TBST or, 5% BSA, TBST), with shaking. This was followed by washing three times in 1x TBST buffer for 10 minutes, with shaking at room temperature, prior to incubating with the secondary antibody (HRP labelled), appropriately diluted in 5% low fat dried milk, TBST, for 1 hour, at room temperature, with shaking. The membrane was washed as above before detection using the ECL™ Western Blotting Detection Reagents (Amersham Biosciences). Enhanced Chemiluminescence (ECL) is a light-emitting-non-radioactive method for detecting immobilized specific primary antibodies conjugated with horseradish peroxidase (HRP)-linked secondary antibodies. Light emission is generated by the HRP/hydrogen peroxide catalysis of luminal to an excited oxidated state, and it is dramatically enhanced in the ECL system by chemical enhancers such as phenol (263). Briefly, an equal volume of detection reagent A was mixed with reagent B (typically 4 ml each). Excess buffer was drained from the membrane and the detection solution was pipetted onto the surface of the membrane carrying the protein. After incubation for one minute at room temperature, excess reagent was drained from the membrane, which was then wrapped in plastic film. The membrane was exposed to blue-light sensitive autoradiography film (Hyperfilm ECL, Amersham Biosciences) for seconds or minutes, before developing in an automated processor (SRX-101A Compact X4, Xograph Imaging Systems).

Unless otherwise stated, membranes were stripped for subsequent analysis in 200 ml of 0.2 M Glycine, pH 2.5, 1% SDS, for 1 hour at room temperature. The membranes were washed three times at room temperature in 1x TBST buffer, blocked overnight and then probed with 1:1000 dilution of goat anti  $\beta$ -actin (I-19) sc-1616 polyclonal antibody (Santa Cruz Biotechnology) for 4 hours at room temperature. As a secondary antibody, HRP conjugated rabbit anti-goat IgG polyclonal antibody (DakoCytomation) was used at dilution of 1:2000 for 1 hour at room temperature.

## Chapter III

### Structure and coding potential of the human and murine *LVI-1* genes

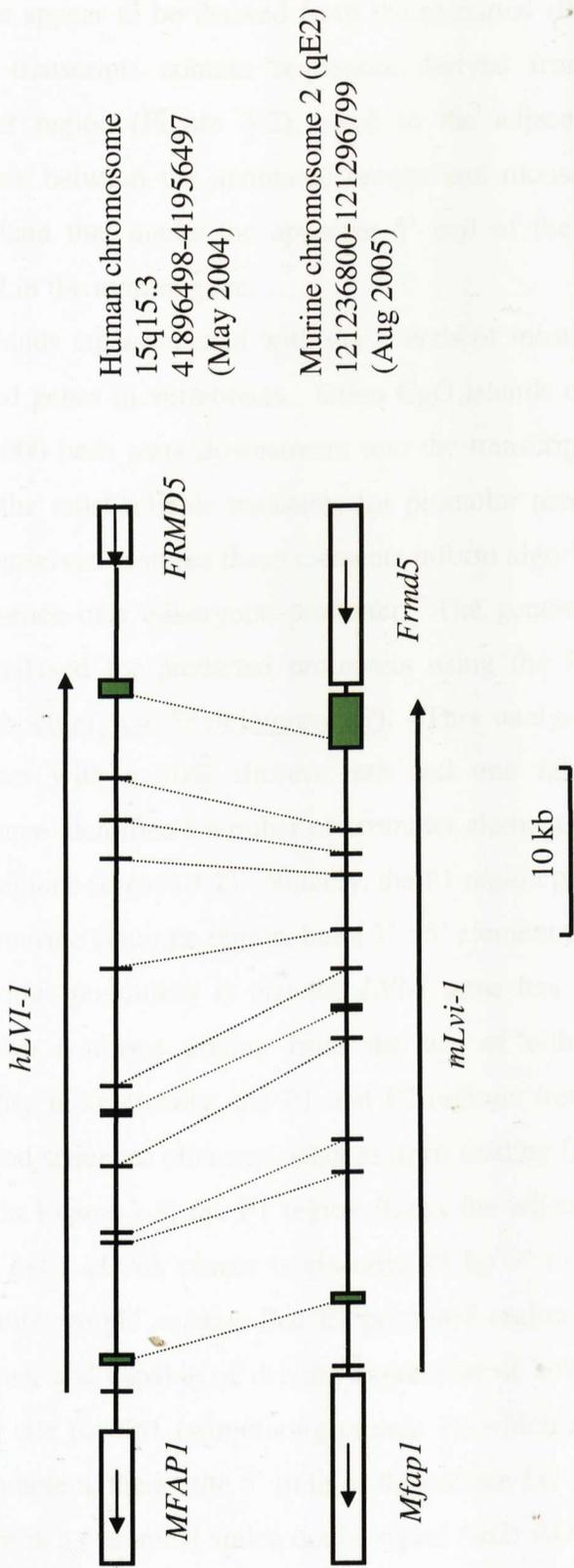
#### 3.1 Introduction

Due to the limiting resources of feline genome and cDNA sequence information, it was decided to focus on the human and murine orthologues of the *LVI-1* gene, which could ultimately be used to model the effects of truncation or transcriptional silencing of the gene. Previous work indicated that the transcript sizes for *LVI-1* were different for the human, murine and feline genes (58). It was therefore considered important to establish the basic transcription pattern and the coding potential of the *LVI-1* genes. At the outset of my work, very little was known about the murine gene, and no full-length cDNA had been reported. However, the available database information has grown steadily, allowing comparison and validation of my experimentally derived information with public domain sequence information. The chapter describes an analysis of the structure of the human and murine *LVI-1* genes based on the currently available information in 2006.

#### 3.2 Analysis of *LVI-1* genomic organisation in human and murine

The mammalian *LVI-1* gene is highly conserved and comprises 13 exons covering 40 kb of the human genome on chromosome 15q15 and 37 kb of the murine genome on chromosome 2 band E5. The overall genomic organisation of the human and murine *LVI-1* genes is shown in Figure 3-1. This and other figures are based on the latest public domain releases of the human and murine genome sequences, ESTs and cDNA clones (<http://www.genome.ucsc.edu>). Homology is evident over exons 2 – 13 of the murine and human genes. The murine gene has a longer 3'UTR (2.3 vs 0.6 kb), which is sufficient to account for the differences in the major transcripts detected by northern blot analysis (4.5 vs 2.8 kb) in previous studies (58). There is an obvious difference, however, at the 5' end of the genes and the 5' exons display no obvious homology.

Comparison of the predicted amino acid sequence of the human and murine cDNAs shows a high degree of homology from exon 2-13, but 5' start regions show a difference between species as a result of the utilisation of different 5' exons and

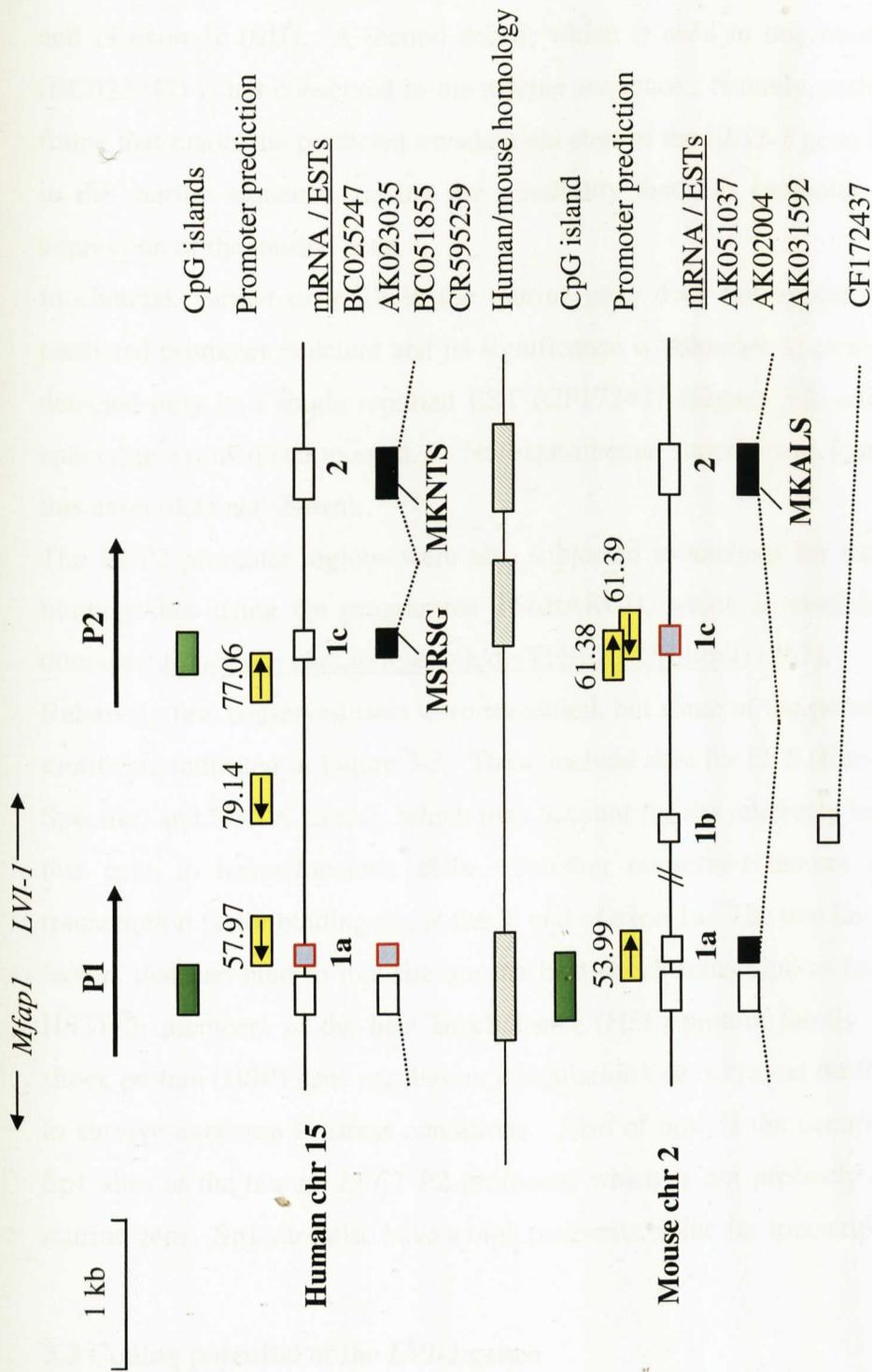


**Figure 3-1. Genomic structure of the human and murine homologues of *LVI-1*.** The open boxes represent the conserved adjacent genes. Black arrows represent the transcriptional orientation. The coding exons of the *LVI-1* genes are indicated by vertical bars and green boxes.

promoters, which I have designated P1 and P2. Human *LVI-1* transcripts in the database appear to be derived from the proximal (P2) promoter region whereas the murine transcripts contain sequences derived from further upstream in the P1 promoter region (Figure 3-2), close to the adjacent 5' gene (*Mfap1*). Another difference between the annotated human and mouse genome sequences lies in the CpG island that marks the apparent 5' end of the human *LVI-1* gene and is not detected in the murine gene.

CpG islands are associated with the 5' ends of most house-keeping genes and many regulated genes in vertebrates. Often CpG islands overlap the promoter and extend about 1000 base pairs downstream into the transcription unit. In fact, they provide one of the most reliable measures for promoter prediction (264, 265). Along with other conserved features these elements inform algorithms that can be used to predict the presence of a eukaryotic promoter. The genomic regions spanning P1 and P2 were analysed for predicted promoters using the Proscan programme version 1.7 (<http://thr.cit.nih.gov/molbio/proscan/>). This analysis is reported to predict pol II promoters with a 70% success rate and one false positive/ 14kb (266). The programme identifies a number of promoter elements in the human and murine *LVI-1* P1/P2 regions (Figure 3-2). Notably, the P1 region predicts a 5'→3' promoter activity for the murine genomic region, but a 3'→5' element for the equivalent human.

An obvious possibility is that the *LVI-1* gene has a dual promoter structure, with alternative isoforms arising from the use of either element. To examine this possibility more closely, the P1 and P2 regions were examined for the presence of conserved sequence elements, such as open reading frames and splice donor sites. As shown in Figure 3-3, the P1 region flanks the adjacent *Mfap1* gene and the longest murine *Lvi-1* cDNA clones begin only 98 bp 3' to *Mfap1* transcription unit. This observation would suggest that P1 promoter region of the murine *Lvi-1* gene is bi-directional and capable of driving expression of both the *Lvi-1* and *Mfap1* gene. A binding site for Sp1 (stimulating protein 1), which is frequently found within CpG-rich promoters, marks the 5' limit of the murine *Lvi-1* cDNAs. At the 3' end of exon 1a, there is a canonical splice donor signal (sd2: AGGT) and a variant (sd1: TGGT), which has been utilized in one of the m*Lvi-1* cDNA clones (AK031592). The most commonly used splice donor (sd2) is conserved in the human genomic sequence. At the P2 region, it is clear that the mouse sequence is also rich in CpGs, although the



**Figure 3-2. Exon structure of the 5' region of *LVI-1* gene.** P1 and P2 represent the P1 and P2 promoter regions, respectively. Black boxes represents the exons utilised by the major *LVI-1* cDNAs of human and mouse with corresponding accession codes. Striped boxes indicate regions of significant interspecies homology. Green boxes show CpG islands, while yellow boxes indicate predicted polIII promoters and their associated probability scores. The black arrows indicate the transcriptional orientation. The locations of alternative translational initiation sequences are also indicated.

reduced stretch of these linkages presumably explains the failure of standard algorithms to score a CpG island in the mouse sequence. However, in contrast to P1, the essential elements appear conserved here, including the major splice donor at the end of exon 1c (sd1). A second donor, which is used in one human *LVI-1* cDNA (BC025247) is not conserved in the murine sequence. Notably, a short open reading frame that marks the predicted translational start of the h*LVI-1* gene is also conserved in the murine sequence, raising the possibility that this promoter may be used in expression of the murine gene.

In contrast, variant exon 1b in the murine gene does not appear to correspond to predicted promoter structure and its significance is unknown at present as it has been detected only in a single reported EST (CF172437 (Figure 3-2, exon 1b)) that was spliced to exons distal to exon 2. No experimental support was found for the use of this exon (data not shown).

The P1/P2 promoter regions were also subjected to analysis for transcription factor binding sites using the programme TFSEARCH, which is available in the public domain (<http://www.cbrc.jp/research/db/TFSEARCH.html>) (267).

Relatively few conserved sites were identified, but some of the potentially significant motifs are indicated in Figure 3-3. These include sites for ETS (E26-Transformation-Specific) and GATA factors, which may account for the relatively high expression of this gene in haematopoietic cells. Another conserved feature is a heat shock transcription factor binding site at the 3' end of exon 1a. The two known transcription factors that can bind to this site are the heat shock transcription factor HSTF-1 and HSTF-2; members of the heat shock factor (HSF) protein family involved in heat shock protein (HSP) gene regulation; a regulation that is critical for the ability of cells to survive exposure to stress conditions. Also of note is the occurrence of adjacent Sp1 sites at the human *LVI-1* P2 promoter, which is not precisely conserved in the murine gene. Sp1 sites also have a high predictive value for transcriptional initiation.

### **3.3 Coding potential of the *LVI-1* genes**

At the outset of this project there was limited sequence information about the mammalian *LVI-1* genes and no full-length cDNA had been cloned for the murine gene. However, a number of cDNA sequences have subsequently been submitted to



the databases. An alignment of these sequences is presented in Figure 3-4. The largest predicted gene product of the human cDNAs (AK023035, BC051855) encode a product of 626 amino acids with a predicted molecular weight of 69.75 kDa, while the longest murine open reading encodes a protein of 57.96 kDa. However, as will be seen from the sequence alignment, the principal difference between the human and murine products lies in the N-terminal domain, which is derived from exon 1c in most of the predicted human gene products, but within exon 2 in the mouse products. One of the human cDNAs, which has an alternative splicing structure that places the exon 1c start out of frame has a homologous start site to the murine gene products. A more complete list of the potential start sites in exons 1c and 2 is given in Figure 3-5. These potential start sites were subjected to analysis using the programme TIS Miner (<http://research.i2r.a-star.edu.sg/DNAFSMiner>), which predicts the probability of utilisation of each AUG in mammalian mRNAs based on match to consensus sequences and an algorithm based on a large dataset of several thousand eukaryotic mRNAs (268). This analysis suggested that the first ATG of the murine cDNA sequence (MKALS) is the most likely translational initiation site in this mRNA. However, a similar analysis of the human cDNA sequences suggested that the second ATG (MKVNE) was more likely to be utilised than the first (MSRSG), while the use of the next in-frame start, which matches to the murine start site (MKNTS) was only slightly less likely to be utilised. The second site (MKVNE) has no equivalent in the hypothetical murine exon 1c cDNA. However, analysis of this theoretical structure strongly suggests that the first initiator will be used (MSGSK, 0.955) although the next in-frame start is also quite strongly predicted (MKALS, 0.815). These results do not provide a strong prediction with regard to the most likely initiation site, and highlight the need for experimental test of these predictions.

### **3.4 Predicted *LVI-1* gene products and their properties**

The most striking feature of the *LVI-1* gene products is the C-terminal 250 amino acids which consists of a series of WD-40 repeats (see Chapter I). No other recognised protein motifs are found, although domain architecture programmes such as SMART (Simple Modular Architecture Research Tool) available online at (<http://smart.embl-heidelberg.de/>) define low complexity regions in the variable N-terminal domains. This outline structure is depicted in Figure 3-6.

exon 1c / exon 2

BC025247	-----	} Human
AK023035	MSRSGAAAEKADSRQRPQMKVNEYKENQNIAYVSLRPAQTTVLIKTAKVYLAPFSLSNYQ	
BC051855	MSRSGAAAEKADSRQRPQMKVNEYKENQNIAYVSLRPAQTTVLIKTAKVYLAPFSLSNYQ	
CR595259	MSRSGAAAEKADSRQRPQMKVNEYKENQNIAYVSLRPAQTTVLIKTAKVYLAPFSLSNYQ	
<b>Human consensus</b>	<b>MSRSGAAAEKADSRQRPQMKVNEYKENQNIAYVSLRPAQTTVLIKTAKVYLAPFSLSNYQ</b>	
BC094676	-----	} Murine
AK031592	-----	
AK162181	-----	
AK051037	-----	
AK020004	-----	
<b>Mouse consensus</b>	-----	

BC025247	-----	-MKNTSSKAESTLQNSSSAVHTES
AK023035	LDQLMCPKSLSEKNSNNEVACKKTKIKKTCRRIIPPKMKNNTSSKAESTLQNSSSAVHTES	
BC051855	LDQLMCPKSLSEKNSNNEVACKKTKIKKTCRRIIPPKMKNNTSSKAESTLQNSSSAVHTES	
CR595259	LDQLMCPKSLSEKNSNNEVACKKTKIKKTCRRIIPPKMKNNTSSKAESTLQNSSSAVHTES	
<b>Human consensus</b>	-----	<b>-MKNTSSKAESTLQNSSSAVHTES</b>
BC094676	-----	-MKALSSKADSLKSSVDAYTES
AK031592	-----	-MKALSSKADSLKSSVDAYTES
AK162181	-----	-MKALSSKADSLKSSVDAYTES
AK051037	-----	-MKALSSKADSLKSSVDAYTES
AK020004	-----	-MKALSSKADSLKSSVDAYTES
<b>Mouse consensus</b>	-----	<b>-MKALSSKADSLKSSVDAYTES</b>

/ exon 3

BC025247	NKLQPKRTADAMNLSVDVLESSQDGDSDDEDTTPSLD-FSGLSPYERKRLKNIENADFFAS
AK023035	NKLQPKRTADAMNLSVDVLESSQDGDSDDEDTTPALD-FSGLSPYERKRLKNIENADFFAS
BC051855	NKLQPKRTADAMNLSVDVLESSQDGDSDDEDTTPALD-FSGLSPYERKRLKNIENADFFAS
CR595259	NKLQPKRTADAMNLSVDVLESSQDGDSDDEDTTPSLD-FSGLSPYERKRLKNIENADFFAS
<b>Human consensus</b>	<b>NKLQPKRTADAMNLSVDVLESSQDGDSDDEDTTPSLD-FSGLSPYERKRLKNIENADFFAS</b>
BC094676	TRLGPKRTSDSATLSVDAES-----SDEDSAPGLDDFSGLSPYERKRLRNIRENANFFAS
AK031592	TRLGPKRTSDSATLSVDAES-----SDEDSAPGLDDFSGLSPYERKRLRNIRENANFFAS
AK162181	TRLGPKRTSDSATLSVDAES-----SDEDSAPGLDDFSGLSPYERKRLRNIRENAXFFAS
AK051037	TRLGPKRTSDSATLSVDAES-----SDEDSAPGLDDFSGLSPYERKRLRNIRENANFFAS
AK020004	TRLGPKRTSDSATLSVDAES-----SDEDSAPGLDDFSGLSPYERKRLRNIRENANFFAS
<b>Mouse consensus</b>	<b>TRLGPKRTSDSATLSVDAES-----SDEDSAPGLDDFSGLSPYERKRLRNIRENANFFAS</b>

(exon 3a) / exon 4 / exon 5

BC025247	LQLSE-----	SAARLREMIEKRQPPKSKRKKPKR-ENGIGCRRSMRLK
AK023035	LQLSE-----	SAARLREMIEKRQPPKSKRKKPKR-ENGIGCRRSMRLK
BC051855	LQLSE-----	SAARLREMIEKRQPPKSKRKKPKR-ENGIGCRRSMRLK
CR595259	LQLSE-----	SAARLREMIEKRQPPKSKRKKPKR-ENGIGCRRSMRLK
<b>Human consensus</b>	<b>LQLSE-----</b>	<b>SAARLREMIEKRQPPKSKRKKPKR-ENGIGCRRSMRLK</b>
BC094676	LQLAE-----	SAARLRGMIKKRESPESKRKRPKKKENEIGCRRSMRLK
AK031592	LQLAE-----	SAARLRGMIKKRESPESKRKRPKKKENEIGCRRSMRLK
AK162181	LQLAE-----	SAARLRGMIKKRESPESKRKRPKKKENEIGCRRSMRLVK
AK051037	LQLAE-----	SAARLRGMIKKRESPESKRKRPKKKENEIGCRRSMRLK
AK020004	LQLAETLVSLLVGVACCPSLNSAARLRGMIKKRESPESKRKRPKKKENEIGCRRSMRLK	
<b>Mouse consensus</b>	<b>LQLAE-----</b>	<b>SAARLRGMIKKRESPESKRKRPKKKENEIGCRRSMRL-K</b>

/ exon 6 /

BC025247	VDPSGVSLPAAPTPTTLVADE-TPLLPPGPLEMTSENQEDNNERFKGFLHTWAGMSKPSS
AK023035	VDPSGVSLPAAPTPTTLVADE-TPLLPPGPLEMTSENQEDNNERFKGFLHTWAGMSKPSS
BC051855	VDPSGVSLPAAPTPTTLVADE-TPLLPPGPLEMTSENQEDNNERFKGFLHTWAGMSKPSS
CR595259	VDPSGVSLPAAPTPTTLVADE-TPLLPPGPLEMTSENQEDNNERFKGFLHTWAGMSKPSS
<b>Human consensus</b>	<b>VDPSGVSLPAAPTPTTLVADE-TPLLPPGPLEMTSENQEDNNERFKGFLHTWAGMSKPSS</b>
BC094676	VDPLGVSLPASPTQPTLVVEEENPLPPGPLEMIPENQDDSSSELLKASLKTWAEMSQTSN
AK031592	VDPLGVSLPASPTQPTLVVEEENPLPPGPLEMIPENQDDSSSELLKASLKTWAEMSQTSN
AK162181	VDPLGVSLPASPTQPTLVVEEENPLPPGPLEMIPGNQDDSSSELLKASLKTWAEMSQTSN
AK051037	VDPLGVSLPASPTQPTLVVEEENPLPPGPLEMIPENQDDSSSELLKASLKTWAEMSQTSN
AK020004	VDPLGVSLPASPTQPTLVVEEENPLPPGPLEMIPENQDDSSSELLKASLKTWAEMSQTSN
<b>Mouse consensus</b>	<b>VDPLGVSLPASPTQPTLVVEEENPLPPGPLEMIPENQDDSSSELLKASLKTWAEMSQTSN</b>

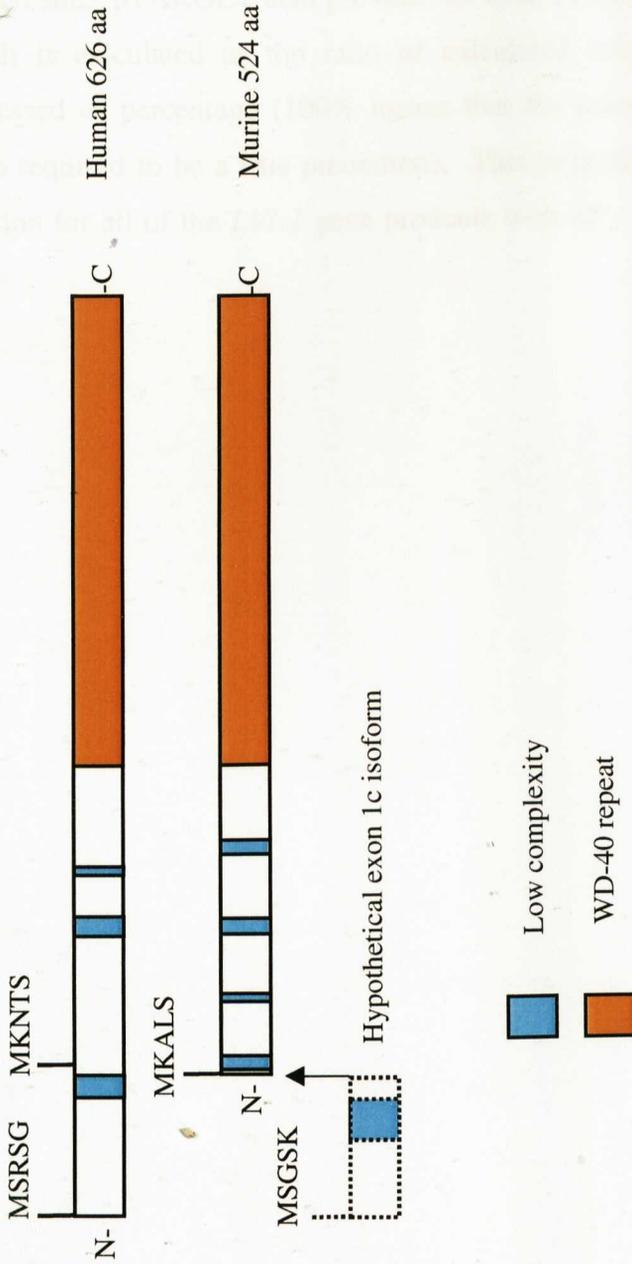
exon 7 / exon 8      WD-40 repeat

BC025247	<u>KNTEKGLSSI</u> IKSYKANLNGMVIS <u>EDTVYKVT</u> TGPIFSMALHP <u>SETRTLVAVGAKFGQVGL</u>
AK023035	<u>KNTEKGLSSI</u> IKSYKANLNGMVIS <u>EDTVYKVT</u> TGPIFSMALHP <u>SETRTLVAVGAKFGQVGL</u>
BC051855	<u>KNTEKGLSSI</u> IKSYKANLNGMVIS <u>EDTVYKVT</u> TGPIFSMALHP <u>SETRTLVAVGAKFGQVGL</u>
CR595259	<u>KNTEKGLSSI</u> IKSYKANLNGMVIS <u>EDTVYKVT</u> TGPIFSMALHP <u>SETRTLVAVGAKFGQVGL</u>
<b>Human consensus</b>	<b><u>KNTEKGLSSI</u>IKSYKANLNGMVIS<u>EDTVYKVT</u>TGPIFSMALHP<u>SETRTLVAVGAKFGQVGL</u></b>
BC094676	<u>EKTKKGLSSI</u> IKSYKANLSGMVIS <u>EATVRKV</u> TGAISSVALHP <u>SEVRTLVAAGAKSGQIGL</u>
AK031592	<u>EKTKKGLSSI</u> IKSYKANLSGMVIS <u>EATVRKV</u> TGAISSVALHP <u>SEVRTLVAAGAKSGQIGL</u>
AK162181	<u>EKTKKGLSSI</u> IKSYKANLSGMVIS <u>EATVRKV</u> TGAISSVALHP <u>SEVRTLVAAGAKSGQIGL</u>
AK051037	<u>EKTKKGLSSI</u> IKSYKANLSGMVIS <u>EATVRKV</u> TGAISSVALHP <u>SEVRTLVAAGAKSGQIGL</u>
AK020004	<u>EKTKKGLSSI</u> IKSYKANLSGMVIS <u>EATVRKV</u> TGAISSVALHP <u>SEVRTLVAAGAKYQIGL</u>
<b>Mouse consensus</b>	<b><u>EKTKKGLSSI</u>IKSYKANLSGMVIS<u>EATVRKV</u>TGAISSVALHP<u>SEVRTL</u>---G---GQ---</b>

	/ exon 9	WD-40 repeat	/	
BC025247	<u>CDLTQQPKEDGVVVFHPHSQPVSCLYFSPANPAHILSLSYDGTLCRGDFSR</u>			} Human
AK023035	<u>CDLTQQPKEDGVVVFHPHSQPVSCLYFSPANPAHILSLSYDGTLCRGDFSR</u>			
BC051855	<u>CDLTQQPKEDGVVVFHPHSQPVSCLYFSPANPAHILSLSYDGTLCRGDFSR</u>			
CR595259	<u>CDLTQQPKEDGVVVFHPHSQPVSCLYFSPANPAHILSLSYDGTLCRGDFSR</u>			
<b>Human consensus</b>	<b><u>CDLTQQPKEDGVVVFHPHSQPVSCLYFSPANPAHILSLSYDGTLCRGDFSR</u></b>			
BC094676	WDLTQQ- <u>SEDAMVVFYAHSRVVSCLSFSP</u>			} Murine
AK031592	WDLTQQ- <u>SEDAMVVFYAHSRVVSCLSFSP</u>			
AK162181	-----			
AK051037	WDLTQQ- <u>SEDAMVVFYAHSRVVSCLSFSP</u>			
AK020004	WDLTQQ- <u>SEDAMVVFYAHSRVVSCLSFSP</u>			
<b>Mouse consensus</b>	<b><u>WDLTQQ-SEDAMVVFYAHSRVVSCLSFSP</u></b>			
	exon 10	WD-40 repeat		
BC025247	<u>ERSSFSFDFLAEDASTLIVGHWDGNMSLVDRRTPGTSYEKLTSSSMGKIR</u>			
AK023035	<u>ERSSFSFDFLAEDASTLIVGHWDGNMSLVDRRTPGTSYEKLTSSSMGKIR</u>			
BC051855	<u>ERSSFSFDFLAEDASTLIVGHWDGNMSLVDRRTPGTSYEKLTSSSMGKIR</u>			
CR595259	<u>ERSSFSFDFLAEDASTLIVGHWDGNMSLVDRRTPGTSYEKLTSSSMGKIR</u>			
<b>Human consensus</b>	<b><u>ERSSFSFDFLAEDASTLIVGHWDGNMSLVDRRTPGTSYEKLTSSSMGKIR</u></b>			
BC094676	EGNSPSSFDL- <u>NDSSLLVGHWDGHL</u>			
AK031592	EGNSPSSFDL- <u>NDSSLLVGHWDGHL</u>			
AK162181	-----			
AK051037	MYLRPDIR-----			
AK020004	QEA-----			
<b>Mouse consensus</b>	<b><u>EGNSPSSFDL-NDSSLLVGHWDGHL</u></b>			
	/ exon 11	WD-40 repeat	/	
BC025247	<u>QYFITAGLRDTHIYDARRLNSRRSQPLISL</u>			
AK023035	<u>QYFITAGLRDTHIYDARRLNSRRSQPLISL</u>			
BC051855	<u>QYFITAGLRDTHIYDARRLNSRRSQPLISL</u>			
CR595259	<u>QYFITAGLRDTHIYDARRLNSRRSQPLISL</u>			
<b>Human consensus</b>	<b><u>QYFITAGLRDTHIYDARRLNSRRSQPLISL</u></b>			
BC094676	<u>QYFVTAGLRDVHVDARFLKSRG</u>			
AK031592	<u>QYFVTAGLRDVHVDARFLKSRG</u>			
AK162181	-----			
AK051037	-----			
AK020004	-----			
<b>Mouse consensus</b>	<b><u>QYFVTAGLRDVHVDARFLKSRG</u></b>			
	exon 12	/ exon 13		
BC025247	<u>IFDSSCISSKIPLLT</u>	<u>IRHNTFTGRWLR</u>	<u>TRFQAMWDPKQEDCV</u>	<u>IVGSM</u>
AK023035	<u>IFDSSCISSKIPLLT</u>	<u>IRHNTFTGRWLR</u>	<u>TRFQAMWDPKQEDCV</u>	<u>IVGSM</u>
BC051855	<u>IFDSSCISSKIPLLT</u>	<u>IRHNTFTGRWLR</u>	<u>TRFQAMWDPKQEDCV</u>	<u>IVGSM</u>
CR595259	<u>IFDSSCISSKIPLLT</u>	-----	-----	-----
<b>Human consensus</b>	<b><u>IFDSSCISSKIPLLT</u></b>	<b><u>IRHNTFTGRWLR</u></b>	<b><u>TRFQAMWDPKQEDCV</u></b>	<b><u>IVGSM</u></b>
BC094676	<u>VFDSSSISSQLPLL</u>	<u>STIRHNTVTGRWLR</u>	<u>TRFQAVWDPKQEDCF</u>	<u>IVGSM</u>
AK031592	<u>VFDSSSISSQLPLL</u>	<u>STIRHNTVTGRWLR</u>	<u>TRFQAVWDPKQEDCF</u>	<u>IVGSM</u>
AK162181	-----	-----	-----	-----
AK051037	-----	-----	-----	-----
AK020004	-----	-----	-----	-----
<b>Mouse consensus</b>	<b><u>VFDSSSISSQLPLL</u></b>	<b><u>STIRHNTVTGRWLR</u></b>	<b><u>TRFQAVWDPKQEDCF</u></b>	<b><u>IVGSM</u></b>
		WD-40 repeat		
BC025247		<u>KRVHSGGEYLVSVCS</u>	<u>INAMHPTRYILAGGNSSGKI</u>	<u>HVFMNEKSC</u>
AK023035		<u>KRVHSGGEYLVSVCS</u>	<u>INAMHPTRYILAGGNSSGKI</u>	<u>HVFMNEKSC</u>
BC051855		<u>KRVHSGGEYLVSVCS</u>	<u>INAMHPTRYILAGGNSSGKI</u>	<u>HVFMNEKSC</u>
CR595259		-----	-----	-----
<b>Human consensus</b>		<b><u>KRVHSGGEYLVSVCS</u></b>	<b><u>INAMHPTRYILAGGNSSGKI</u></b>	<b><u>HVFMNEKSC</u></b>
BC094676		<u>KNVHSLWGECLVSVCS</u>	<u>LSAVHPTRYILAGGNSSGKL</u>	<u>HVFMHQET-</u>
AK031592		<u>KNVHSLWGECLVSVCS</u>	<u>LSAVHPTRYILAGGNSSGKL</u>	<u>HVFMHQET-</u>
AK162181		-----	-----	-----
AK051037		-----	-----	-----
AK020004		-----	-----	-----
<b>Mouse consensus</b>		<b><u>KNVHSLWGECLVSVCS</u></b>	<b><u>LSAVHPTRYILAGGNSSGKL</u></b>	<b><u>HVFMHQET-</u></b>

**Figure 3-4. Predicted amino acid sequence comparison of human and murine *LVI-1* cDNAs.** Published human (BC025247, AK023035, BC051855, and CR595259) and murine (BC094676, AK031592, AK162181, AK051037, AK020004) cDNA is shown. Bold text indicates the human and mouse consensus sequences. Underlined text represents the WD-40 repeats.





**Figure 3-6. Structural features of predicted human and murine *LVI-1* gene products.** This figure is based on analysis of both gene products for conserved protein motifs using the SMART (Simple Modular Architecture Research Tool) programme at <http://smart.embl-heidelberg.de/>. The light blue and the orange boxes represent the low complexity regions and the WD-40 repeats, respectively.

The predicted proteins were also examined for physicochemical properties (Chapter VI) and likely subcellular location using the online programme pTARGET available on the web page (<http://bioinformatics.albany.edu>) (269), which lists nine potential subcellular locations and assigns proteins to these on the basis of known motifs and N- and C-terminal amino acid composition which are characteristic of proteins found at each site. pTARGET also provides us with a confidence value for each prediction which is calculated as the ratio of calculated score to the total score required, expressed as percentage (100% means that the calculated score is the same as the score required to be a true prediction). This programme strongly predicts a nuclear location for all of the *LVI-1* gene products with 88 – 100% confidence (see Table 3-1).

**Table 3-1. Predicted location of predicted LVI-1 gene products.** Analysed by pTARGET programme available online at: <http://bioinformatics.albany.edu/~ptarget/> (269).

cDNA	Localisation	Confidence
Murine BC094676	Nuclear	87.6%
Murine AK031592	Nuclear	87.6%
Murine AK162181	Nuclear	93.9%
Murine AK00204	Nuclear	93.9%
Hypothetical exon 1c mRNA	Nuclear	100%
Human BC025247	Nuclear	100%
Human AK023035	Nuclear	100%
Human BC051855	Nuclear	100%
Human CR595259	Nuclear	100%

### 3.5 Discussion

Analysis of the human and murine *LVI-1* genes and known transcripts reveals a difference in 3'UTR length, which presumably accounts for the differences in mRNA size detected on northern blot analysis from a previous study (58). The most significant difference, however, is seen at the 5' end where the human and murine genes appear to use alternative promoters. Based on the evidence of the available cDNAs and numerous ESTs, the murine gene is transcribed from the distal P1 promoter, which appears to be a bi-directional element shared with the adjacent *Mfap1* gene. In contrast, the human gene transcripts are derived from the proximal P2 promoter. Analysis of gene structure in the P1 and P2 regions provides some theoretical support for this species differences.

Promoter prediction algorithms score the forward P1 promoter only for the murine gene. The P2 promoter adjacent to exon 1c is more strongly predicted in the human genome and it is conceivable that the use of the distal P1 promoter by the murine gene would suppress transcription here by the phenomenon of promoter interference (270). However, it is notable that the basic architecture of exon 1c is conserved in the murine including with its open reading frame and splice donor sites. The possibility remains, therefore, that the P2 promoter might be used by the murine gene in some circumstances.

The issue of *LVI-1* gene promoter use is important, as there are implications for the encoded gene products. Use of P2 and exon 1c would result in mRNA encoding a protein with close to 100 additional amino acids, compared to the use of the next common ATG. However, in view of the relative uncertainty of translational initiation predictions, it is clear that direct experimental analysis will be necessary to determine which proteins are actually generated. The significance of variant proteins encoded by rare alternatively spliced mRNAs are unclear and may be artefactual.

Structural analysis of the *LVI-1* proteins reveals only a single specific feature, which is the presence of multiple WD-40 repeats at the C-terminus. However, it is notable that a composition-based prediction programme strongly suggests a nuclear location for the *LVI-1* gene products. A caveat with this programme is that it is poor at predicting the location of proteins that shuttle between compartments e.g. nucleus – cytoplasm (269).

## Chapter IV

### Expression of the human and murine *LVI-1* genes

#### 4.1 Introduction

At the outset of this study, there was little information on transcripts from the murine *Lvi-1* gene and no full-length cDNAs were evident from the database. One of the first priorities, therefore, was to generate a probe to allow detection of *Lvi-1* transcripts containing the 3' exons. Also, it has become clear that most of the human *LVI-1* transcripts are derived from the use of the P2 promoter, with exon 1c as the 5' exon, while the murine *Lvi-1* transcripts are derived from the P1 promoter, with exon 1a as the 5' exon. However, as will be seen from the comparative sequence analysis in Chapter III, it is conceivable that the gene has a dual promoter structure and that these promoters are differentially regulated. The potential for use of these alternative promoters was assessed by 5'RACE and RT-PCR analyses. These studies are important as they have implications for the structure of the *LVI-1* gene products as well as regulation of expression of the gene.

#### 4.2 Materials and methods

Variations from previous chapter only are included.

##### 4.2.1 Genomic PCR

1  $\mu$ g of murine genomic DNA (p/m 23i cell line) was used to amplify exon 13 of the murine *Lvi-1* gene using 100 pmol of each of the gene-specific primers. Primer sequences and PCR conditions are detailed in Tables 4-1, 4-2 and Section 2.2.4. The 463 bp amplified product was gel purified (Section 2.2.4), cloned (Section 2.2.2), and its identity verified by sequencing (Section 2.2.5).

##### 4.2.2 Generation of a human *LVI-1* and $\beta$ -actin probes

Miniprep DNA from the human pGEX-LVI-1 fusion plasmid (Section 5.3.14) was used to amplify the exon 13 of the human *LVI-1* gene with 50 pmol of the specific primers. For primer sequence and PCR conditions see Table 4-1, Table 4-2, and also

**Table 4-1. Primer sequences (by hand)**

Number of primer	Primer	Strand	Sequence (5'→3')
1	mExon13F	Forward	ggt ggc taa cga ggt tcc a
2	mExon13R	Reverse	cat gaa aac acc cgt ttc tgc
3	mFLJexon1Fa	Forward	gcg gct gct tca tga gtg c
4	mFLJexon2F	Forward	gag acc gat aca gat tac aac
5	mFLJexon5R	Reverse	gca tgg acc ttc gac atc c
6	mFLJexon5F	Forward	tgg atg tgc aag gtc cat g
7	mFLJexon11R	Reverse	ctg agc tta cag tca gca c
8	mouseFLJexon2RSP1	Reverse	tgg tgc act gtc ttc gtc
9	mouseFLJexon2RSP2	Reverse	cac act gag agt cgc tga
10	mouseFLJexon2RSP3	Reverse	caa ccg ggg att tgg aca
11	mFLJexon1aF	Forward	gct taa agg gct agg gtg ag
12	mFLJexon1cF	Forward	atg tct ggg tcc aag gct g
13	mFLJexon2R	Reverse	cag gcc tgg tgc act gtc
14	mHPRT5	Forward	ggg ggc tat aag ttc ttt gc
15	mHPRT3	Reverse	tcc aac act tcg aga ggt cc
16	huHPRT5	Forward	ggg ggc tat aaa ttc ttt gc
17	huHPRT3	Reverse	tcc aac act tcg aga ggt cc
18	humanFLJexon2RSP1	Reverse	ggc tgg tgt ggt atc ttc
19	humanFLJexon2RSP2	Reverse	gat tca tcg cat ctg ccg
20	humanFLJexon2RSP3	Reverse	agc gtg gat tct gcc ttg
21	humanFLJexon1aF	Forward	ggg ctt tgc tgt tag gac tc
22	humanFLJexon1cF	Forward	gaa gcc gaa aga tgt cca g
23	HumanFLJEXON13F	Forward	cac ttt cac tgg gcg atg g
24	HumanFLJEXON13R	Reverse	aca tgt atc ttc ccg ctg ga

**Table 4-2. Primer combinations, PCR annealing conditions, and applications.**

Primer combination	Product size	Annealing temperature (°C)	Application
1 + 2	463 bp	60	Genomic PCR
3 + 2	2383 bp	62	Analysis of transcript(s)
2 + 4	1985 bp	62	Analysis of transcript(s)
4 + 5	590 bp	55	Analysis of transcript(s)
6 + 7	924 bp	60	Analysis of transcript(s)
11 + 13	513 bp	55	Analysis of P1 promoter
12 + 9	461 bp	55	Analysis of P2 promoter
14 + 15	294 bp	55	HPRT control
16 + 17	313 bp	55	HPRT control
21 + 19	373 bp	55	Analysis of P1 promoter
22 + 19	411 bp	55	Analysis of P2 promoter
23 + 24	234 bp	60	Plasmid PCR

Section 2.2.4. The PCR product was then purified as described in Section 2.2.4 and radioactively labelled with  $\alpha$ -<sup>32</sup>P dCTP as detailed in Section 2.2.6.

Human  $\beta$ -actin cDNA control probe (Clontech Laboratories, Palo Alto, USA) was radioactively labelled with  $\alpha$ -<sup>32</sup>P dCTP as detailed in Section 2.2.6.

#### **4.2.3 Reverse transcription-polymerase chain reaction (RT-PCR)**

1 $\mu$ g of total mouse RNA (thymus (Ambion, Inc), brain, lung, testis, liver, kidney, muscle, spleen and heart), and human RNA (Jurkat, CEM, KG-1a, and K-562 cell lines) were reverse transcribed using the ImProm-II Reverse Transcription System (Promega) (Section 2.2.4). PCR was carried out using 2.5  $\mu$ l of freshly synthesised cDNA and 50 pmol of each forward and reverse specific primers (see Tables 4-1, 4-2, and Section 2.2.4 for primer sequence, primer combination, and PCR conditions). 5  $\mu$ l aliquots of each reaction were separated on 4% polyacrylamide gels, stained with ethidium bromide and visualized by UV illumination. Amplification of the long fragments (murine *Lvi-1* exons 1-13, 2-13, and 5-11) was performed using Extensor PCR Master Mix, Buffer 1, ReddyMix<sup>TM</sup> version (ABgene) (Section 2.1.5) and 2.5  $\mu$ l aliquots of mouse thymus cDNA with 5 pmol of the specific primers (see Table 4-1 and Table 4-2 for primer sequence, primer combination and annealing temperatures). Cycling conditions were 94 °C 2 minutes, followed by 94 °C 10 seconds, 60-62 °C 30 seconds, 68 °C 3 minutes for 10 cycles, then 94 °C 20 seconds, 60-62 °C 30 seconds, 68 °C 9 minutes for 20 cycles and finally 68 °C for 7 minutes. 15 $\mu$ l aliquots of each reaction were separated on 1% agarose gels, stained with ethidium bromide and visualized by UV illumination.

#### **4.2.4 5'Rapid Amplification of cDNA Ends Assay (5'RACE)**

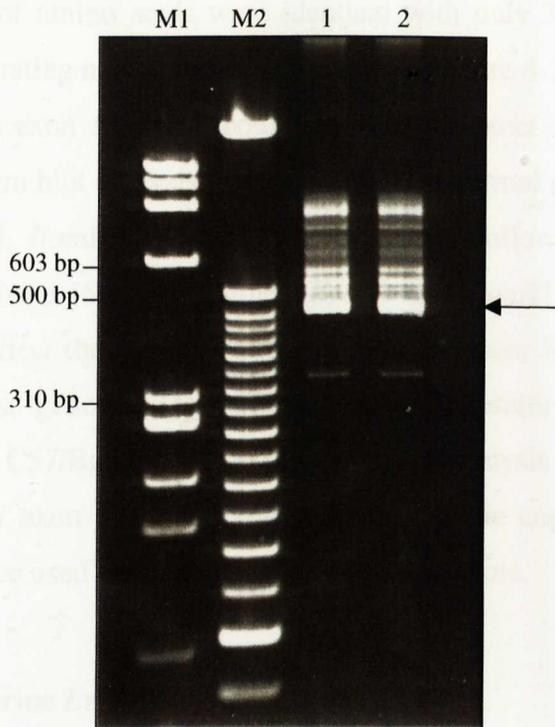
5'RACE was performed on 2  $\mu$ g of total RNA (human and murine) using the 5'RACE Kit 2<sup>nd</sup> Generation (Roche) according to the manufacturer's instructions. First-strand cDNA was synthesized using a gene-specific primer SP1 (12.5  $\mu$ M) (Table 4-1, primers 8 and 18). The first-strand cDNA was purified from unincorporated nucleotides and primers using the QIAquick PCR Purification Kit (Qiagen) according to the manufacture's protocol. Then, terminal transferase was used to add a homopolymeric A-tail to the 3'end of the cDNA. The tailed cDNA was amplified by

PCR using a gene-specific primer SP2 (12.5  $\mu$ M) (Table 4-1, primers 9 and 19) and the oligo dT-anchor primer. The oligo dT-anchor primer is a mixture of primers with one non T (A, C or G) nucleotide at the 3'end, so that the oligo dT-anchor primer is forced to bind to the inner end of the poly(A)-tail. Therefore, the actual length of the added poly(A)-tail is unimportant. The obtained cDNA was further amplified by a second PCR using a nested specific primer SP3 (12.5  $\mu$ l) (Table 4-1, primers 10 and 20), and the PCR anchor primer (supplied in 5'RACE Kit 2<sup>nd</sup> Generation, Roche). The PCR conditions were 94 °C for 2 minutes, followed by 10 cycles of 30 seconds denaturation at 94 °C, 30 seconds annealing at 55 °C, 40 seconds elongation at 72 °C and then 25 cycles of 30 seconds denaturation at 94 °C, 30 seconds annealing at 55 °C, 40 seconds (plus 20 seconds for each cycle) elongation at 72 °C with a final elongation at 72 °C for 7 minutes. The resulting RACE products were cloned into the pCR<sup>®</sup> 2.1-TOPO vector (Section 2.2.2), and the identity of the amplified fragments verified by sequencing (Section 2.2.5).

### **4.3 *Lvi-1* gene expression**

#### **4.3.1 Generation of a murine specific *Lvi-1* probe**

BLAST analysis of murine cDNA and EST databases using the full-length human *LVI-1* cDNA sequence (AK023035, derived from the teratocarcinoma cell line NT2 and comprising 13 exons) identified a murine cDNA AK020004 (derived from adult male thymus) that showed a high degree of homology across exons 2 to 9. This murine cDNA appeared to lack exons 10 to 13. However, further BLAST searching using the human *LVI-1* sequence coding for exons 10 through to 13 identified homologous murine sequences on murine chromosome 2. Murine *Lvi-1* exon 13 specific primers were designed (mExon13F, mExon13R, see Table 4-1) and used to amplify an exon 13 specific fragment from normal murine genomic DNA. In addition to the expected 463 bp PCR product, a number of other products were observed, most probably due to non-specific amplification of genomic sequences (Figure 4-1). The putative exon 13 specific fragment was gel purified, cloned, and several clones isolated, and sequenced. Sequence analysis revealed 82% homology with the equivalent human sequence confirming that the fragment amplified indeed represents exon 13 of the murine *Lvi-1* gene (Figure 4-2). Comparison of the predicted amino acid sequence of both human and murine exon 13 also revealed a high degree of



**Figure 4-1. Amplification of murine *Lvi-1* exon 13.** Murine *Lvi-1* exon 13 amplified by DNA PCR using the primers mExon13F and mExon13R (Table 4-1). PCR products were separated on a 4% polyacrylamide gel (lanes 1 and 2), stained with ethidium bromide and visualized by UV illumination. M1:  $\Phi$ X174RF DNA size marker, M2: 25 bp DNA ladder size marker. The black arrow indicates the 463 bp, amplified fragment of murine *Lvi-1* exon 13.

conservation. 78% of amino acids were identical with only 3 of the remaining 19 amino acids demonstrating non-conservative changes (Figure 4-2).

To confirm that this exon 13 probe could be used to detect a unique sequence in murine DNA, Southern blot analysis was carried out on normal murine genomic DNA digested with *EcoRI*, *BamHI* and *SstI* (Figure 4-3). Following washing at high stringency, a single band (*SstI*) or doublet (*EcoRI* and *BamHI*) was observed. The doublets probably reflect the presence of restriction fragment length polymorphisms (RFLPs) in the mouse genomic DNA (purified from laboratory mice with a mixed genetic background, C57/Balbc, NIH, 129SV). This analysis clearly demonstrates that the murine *Lvi-1* exon 13 DNA probe acts as a single copy probe and since it spans an exon it can be used to detect murine *Lvi-1* transcripts.

#### 4.3.2 Analysis of murine *Lvi-1* gene expression

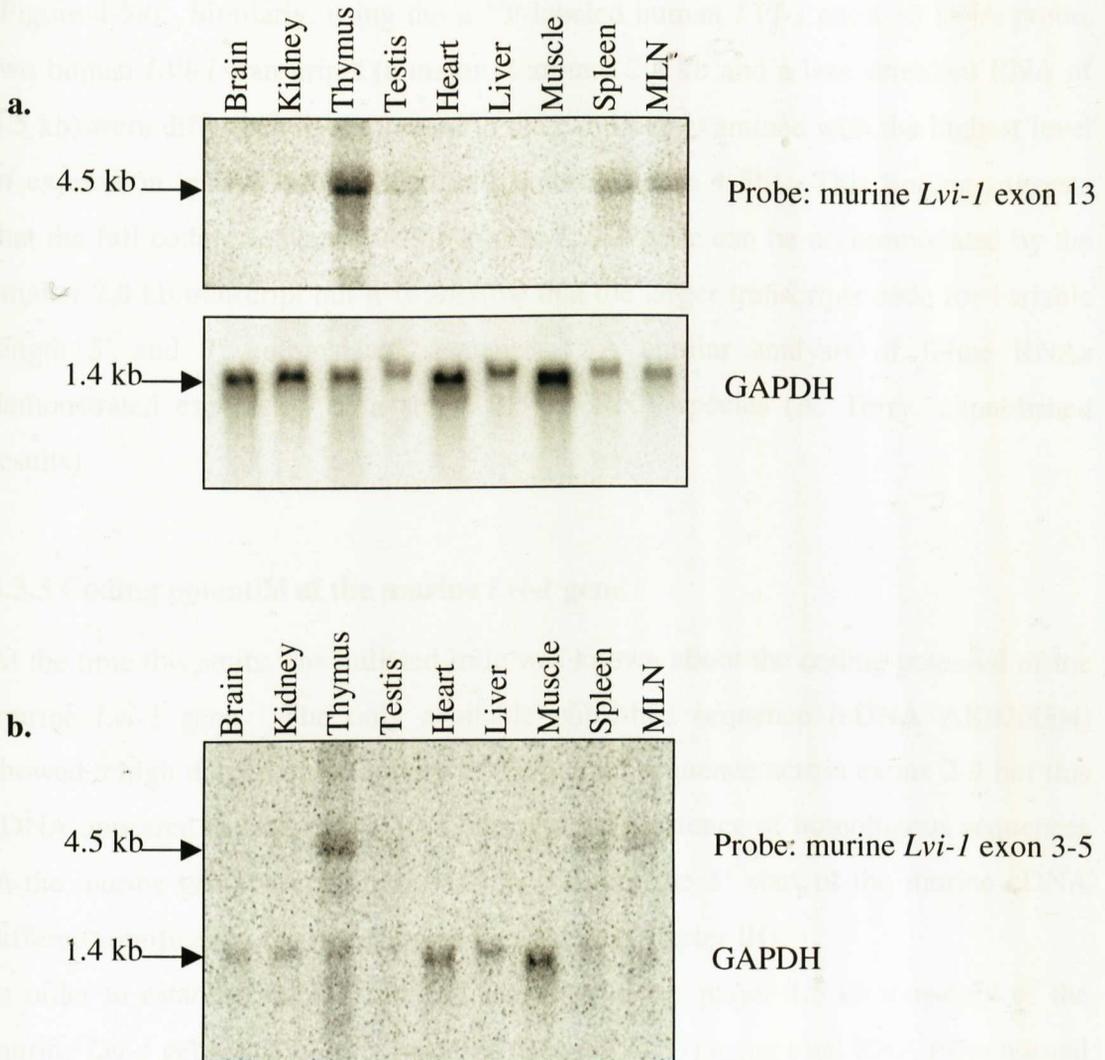
The expression pattern of the murine *Lvi-1* gene was explored by northern blot analysis of total RNA isolated from a range of normal murine tissues including brain, kidney, thymus, spleen, mesenteric lymph node, testis, heart, liver, and muscle (Figure 4-4). Following hybridisation and washing at high stringency, a single RNA species of 4.5 kb was identified with the  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe (Figure 4-4a) and with an alternative *Lvi-1* probe which spans the coding region from exon 3 through to exon 5 (Figure 4-4b). This probe fragment was generated in the same way as the exon 13 specific probe described in Section 4.3.1 (A. Terry, unpublished results). The highest levels of expression were detected in lymphoid-derived tissues especially murine thymus. However, low level expression was observed across the range of tissues examined. A similar ubiquitous expression pattern was observed across a range of normal human tissues with the highest expression observed in human thymus RNA. However, in contrast to the single murine *Lvi-1* transcript, the human *LVI-1* gene appears to be expressed in the form of 2 transcripts: a major species of 2.8 kb and a less abundant RNA of 4.5 kb (146). These findings were also confirmed in a panel of 6 murine cell lines (4 T lymphoblastoid cell lines and 2 fibroblast cell lines) and 6 human lymphoblastoid cell lines. Thus, using the  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe, a single murine *Lvi-1* transcript of 4.5 kb was also differentially expressed in the cell lines examined with the highest level of expression in p/m16i, then p/m47i, and BW5147

		H N T <b>F</b> T G R W L T R F Q A <b>M</b> W D
Human	CCATCAGGCACAACACTTTC	ACTGGGCGATGGCTGACCAGGTTCCAAGCCATGTGGGATC
	*****	***** * ***** ** ***** *
Murine	CCATCAGGCACAACACTGT	CACTGGGAGGTGGCTAACGAGGTTCCAGGCTGTGTGGGACC
		H N T <b>V</b> T G R W L T R F Q A <b>V</b> W D
		P K Q E D C <b>V</b> I V G S M <b>A</b> H P R R V E <b>I</b>
Human	CTAAACAAGAAGACTGTG	TGCATAGTTGGCAGCATGGCCATCCACGACGGGTAGAAATCT
	*****	** ***** ***** ** ***** ** *
Murine	CTAAACAAGAAGATTGCT	TTCATAGTTGGTAGCATGGACCACCCACGACGGGTGGAAGTCT
		P K Q E D C <b>F</b> I V G S M <b>D</b> H P R R V E <b>V</b>
		F H E <b>T</b> G K <b>R</b> V H S <b>F</b> <b>G</b> G E <b>Y</b> L V S V C
Human	TCCATGAGACAGGAAAG	AGGGTGCATTCGTTTGGTGGAGAATACCTTGTCTCTGTGTGTT
	****	*** * ***** ***** ** * ***** ***** *****
Murine	TCCACGAGTCGGGAAAG	AATGTGCATTCCTTTGGGAGAATGTCTTGATCTGTGTGTT
		F H E <b>S</b> G K <b>N</b> V H S <b>L</b> <b>W</b> G E <b>C</b> L V S V C
		S <b>I</b> N A <b>M</b> H P T R Y I L A G G N S S G K
Human	CCATCAATGCCATGCAC	CCAACCTCGGTATATTTGGCTGGAGGTAATCCAGCGGGAAGA
	**	*** ** * ***** ***** ***** ***** ** *
Murine	CCCTCAGTGCTGTGC	ATCCTACGCGATATATCCTGGCTGGAGGTAATCCAGTGGCAAGT
		S <b>L</b> <b>S</b> A <b>V</b> H P T R Y I L A G G N S S G K
		<b>I</b> H V F M <b>N</b> E K S <b>C</b> -
Human	TACATGTTTTTATGA	ATGAAAAAAGCTGCTGA
	*****	** ** * ** *
Murine	TACATGTTTTTATGC	ATCAAGAAACCTGAGTC
		<b>L</b> H V F M <b>H</b> Q E T - V

**Figure 4-2. Conservation of the nucleotide sequence and coding potential of exon 13 of the human and murine *LVI-1* gene.** Identical nucleotides are indicated by \*. Conservative amino acid changes are indicated in blue bold and non-conservative changes in red bold.



**Figure 4-3. Validation of murine exon 13 fragment as a single copy gene probe.** Murine genomic DNA was digested with *EcoRI* (lane 1), *BamHI* (lane 2) and *SstI* (lane 3), separated on a 0.8% agarose TAE gel and transferred to Hybond N membrane. The blot was hybridized with  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe. Doublets identified in lanes 1 and 2 reflect the RFLPs due the genomic parental DNA (three breeders: C57/Balbc, NIH and 129SV). M:  $\lambda$  DNA/*HindIII* fragments labeled with <sup>32</sup>P-d CTP.



**Figure 4-4. Expression of *Lvi-1* gene transcripts in normal murine tissues.** **a.** 20  $\mu$ g samples of total mouse RNA from various tissues sources were separated on 1% agarose-formaldehyde gels and transferred to Hybond-N membrane blot. The blot was probed with radiolabeled  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe and then reprobed with a radiolabeled  $\alpha$ -<sup>32</sup>P-GAPDH fragment, to control for RNA loading and integrity. **b.** The same blot was stripped and then reprobed with a radiolabeled  $\alpha$ -<sup>32</sup>P-murine *Lvi-1* exon 3-5 DNA probe. MLN, mesenteric lymph nodes.

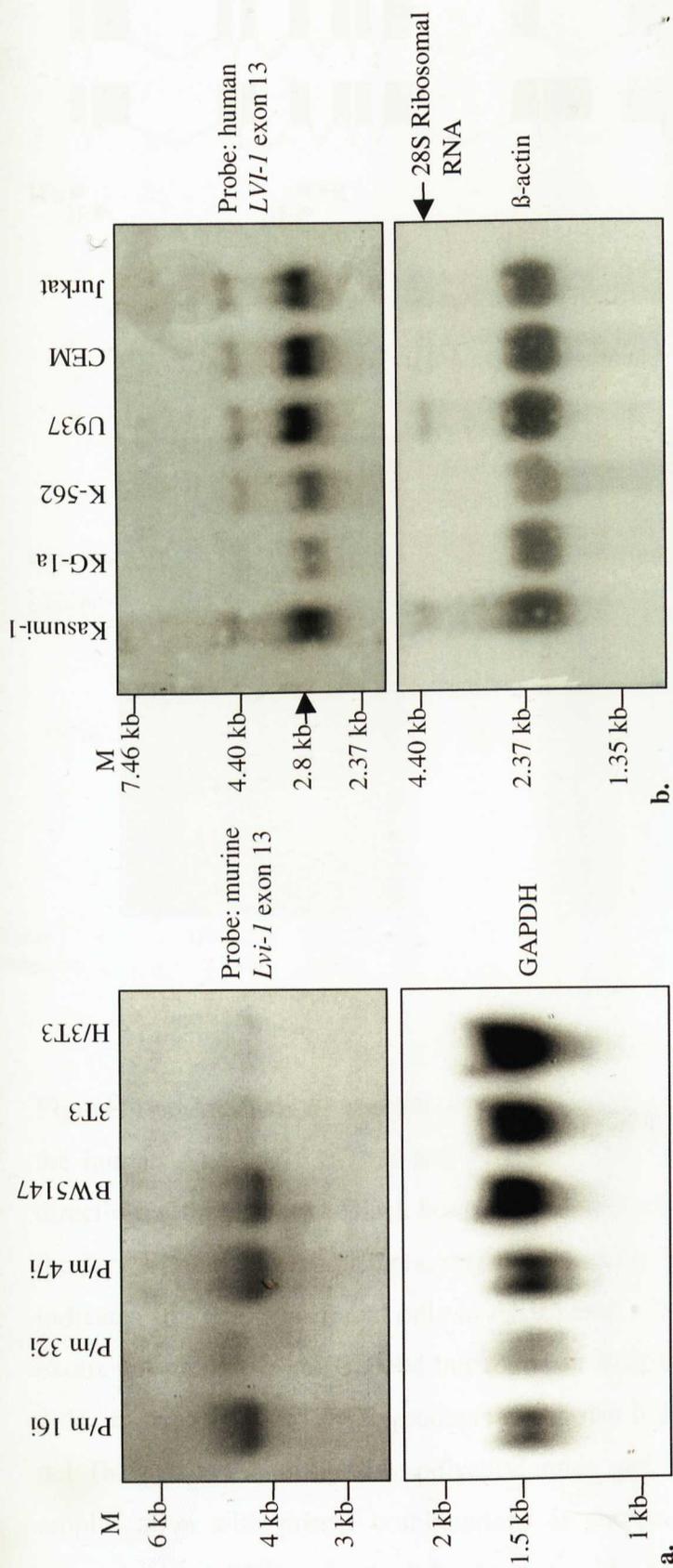
(Figure 4-5a). Similarly, using the  $\alpha$ -<sup>32</sup>P-labeled human *LVI-1* exon 13 DNA probe, two human *LVI-1* transcripts (a major at around 2.8 kb and a less abundant RNA of 4.5 kb) were differentially expressed in the cell lines examined with the highest level of expression in U937, then CEM and Jurkat, (Figure 4-5b). This finding suggests that the full coding sequence of the human *LVI-1* gene can be accommodated by the smaller 2.8 kb transcript but it is possible that the larger transcripts code for variable length 5' and 3' untranslated sequences. A similar analysis of feline RNAs demonstrated expression of a single 3.5 kb RNA species (A. Terry, unpublished results).

#### 4.3.3 Coding potential of the murine *Lvi-1* gene

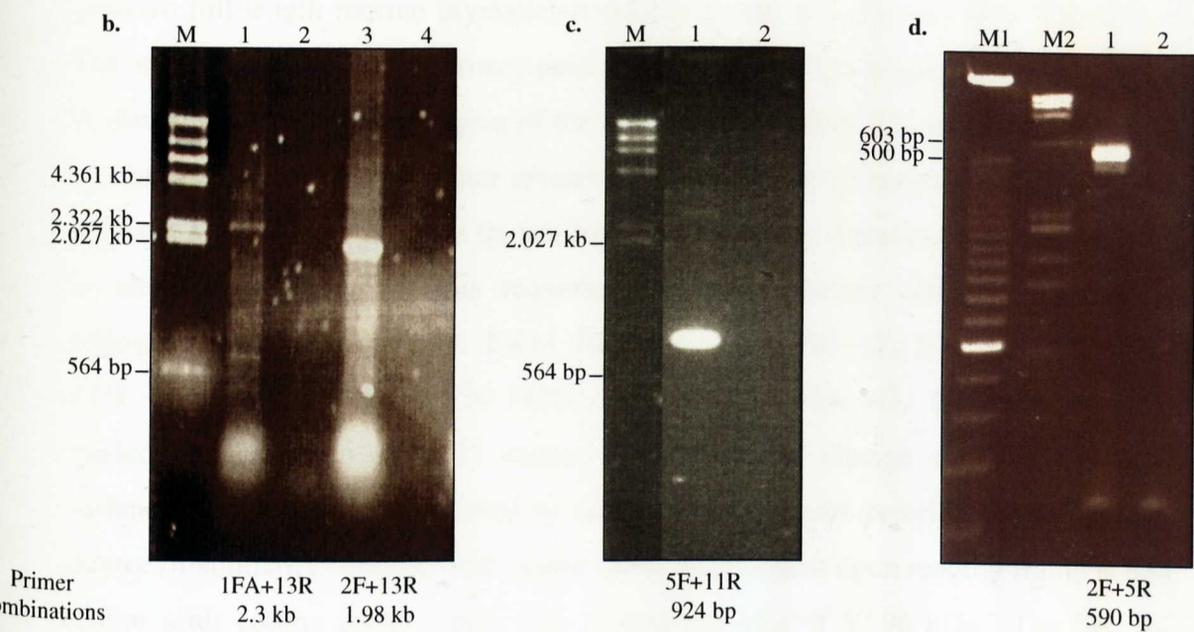
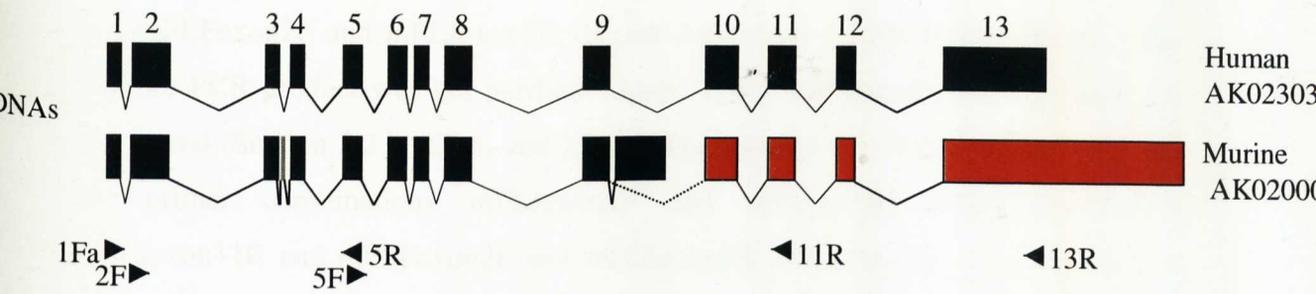
At the time this study was initiated little was known about the coding potential of the murine *Lvi-1* gene. The only available published sequence (cDNA AK020004) showed a high degree of homology to the human sequence across exons 2-9 but this cDNA appeared to lack exons 10-13 despite the existence of homologous sequences in the murine genome (Figure 4-6a). In addition the 5' start of the murine cDNA differed significantly from the human equivalent (Chapter III).

In order to establish the 5' start and sequence of the major 4.5 kb transcript of the murine *Lvi-1* gene, cDNA was prepared (Section 4.2.3) using total RNA from normal murine thymus (the highest expresser of the 4.5 kb *Lvi-1* transcript). Subsequent PCR was performed on aliquots of thymus cDNA using a variety of combinations of murine *Lvi-1* primers (Table 4-1).

Careful analysis of the murine genomic sequence on chromosome 2 spanning the potential start site of the *Lvi-1* gene was carried out and a primer designed mFLJexon1Fa (Table 4-1). This forward primer lies upstream of the predicted promoter region of the murine *Lvi-1* gene (see Chapter III) and is upstream of the sequences contained in the published murine cDNA clones (AK020004, AK051037 and AK031592). It was hoped that amplification in combination with an exon 13 specific reverse primer would generate full-length P1-derived cDNA. PCR was carried out using mFLJexon1Fa in combination with mExon13R (Table 4-1) and a 2.383 kb fragment was generated (Figure 4-6b, lane 1). Additional PCRs included combination of mFLJexon2F and mExon13R (Figure 4-6b, lane 3; 1.985 kb amplicon), mFLJexon5F and mFLJexon11R (Figure 4-6c, lane 1; 898 bp amplicon),



**Figure 4-5. Northern blot analysis of *LVI-1* expression in various murine and human cell lines. a.** Murine RNA blot was hybridized sequentially with  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe, and GAPDH probe (to control for equal loading of RNA). Cell lines analysed were p/m 16i, p/m 32i, p/m 47i, BW5147, 3T3, and H/3T3. Size markers (0.5-10 kb RNA ladder, Invitrogen Life Technologies) are indicated. **b.** Human RNA blot was hybridized sequentially with  $\alpha$ -<sup>32</sup>P-labeled human *LVI-1* exon 13 DNA probe and a  $\beta$ -actin probe (to control for equal loading of RNA). Cell lines analysed were Kasumi-1, KG-1a, K-562, U-937, CEM, and Jurkat. The relative migration of RNA size markers (0.24-9.5 kb RNA ladder, Invitrogen Life Technologies) is indicated. The blot exposure time for the human and murine *LVI-1* transcripts and controls ( $\beta$ -actin and GAPDH) was 7 days and 3 hours, respectively.



**Figure 4-6. Analysis of murine *Lvi-1* transcripts structure.** **a.** Exon structure of the human AK023035 cDNA and murine AK020004 cDNA showing position and direction of the primers. Black boxes and arrowheads represent the coding exons and the forward and reverse primers, respectively. The grey box between exons 3 and 4 indicates an extra exon found only in AK020004. The red boxes represent the coding exons not present in AK020004 but found in BC094676 and AK031592 (see Figure 3-4). **b., c., and d.** RT-PCR products are shown following separation by 1% agarose gel (**b.** and **c.**) and by 4% polyacrylamide gel (**d.**). Lanes 1 and 3 indicate amplification with primer combinations as shown. Lanes 2 and 4 represent the corresponding PCR primer only controls. M, M1, and M2 are the marker  $\lambda$  DNAHindIII, the marker 25bp DNA Ladder, and the marker  $\Phi$ X174RF DNA, respectively.

and mFLFexon2F and mFLJexon5R (Figure 4-6d, lane 1; 590 bp amplicon). The 2.38 kb PCR product was gel purified, cloned and a number of clones isolated and sequenced (Section 2.2.2, 2.2.4, and 2.2.5). The sizes of the PCR products obtained for primer combinations mFLJexon2F and mExon13R, mFLJexon5F and mFLJexon11R and mFLFexon2F and mFLJexon5R were exactly as predicted from the available *Lvi-1* cDNA sequences. Products were cloned and sequenced (as above), and DNA sequence analysis (Section 2.2.5) confirmed the specificity of the amplified products. The consensus sequence and predicted open reading frame of the putative full-length murine thymus-derived *Lvi-1* cDNA is illustrated in Figure 4-7. The mFLJexon1Fa forward primer used to generate the full-length cDNA lies in the bi-directional P1 promoter region of the *Lvi-1* gene (Section 3.2) and examination of the murine *Lvi-1* cDNA sequence reveals that the first 116 bp represent the start of the adjacent *Mfap-1* gene which is transcribed in the opposite direction to the *Lvi-1* gene. In addition, alignment of this sequence with other murine cDNAs revealed the following nucleotide changes: EX1a 107T→C, EX8 76C→T, EX8 88C→T, EX9 111C→T, EX9 129T→C, EX10 192G→A, and EX12 50A→G. However, only the nucleotide change in exon 12 causes an amino acid change (I437V). Indeed, isoleucine is frequently substituted by valine in homologous proteins due to the high degree of similarity between both amino acids. The longest open reading frame is 524 amino acids coding for a protein with a predicted size of 57.96 kDa. The first in-frame ATG is located in exon 2. This methionine is the first start codon conserved across species (Figure 4-8 and 3-5). Moreover, murine cDNAs share homology in the 5' start region, but no in-frame start codon is present in the upstream exon 1a. Although the human cDNAs contain two additional in-frame ATG codons in exon 1c and exon 2 (Figure 3-2), these are not conserved in the mouse gene and it is unclear if these are utilised. The possibility exists that the human exon 1c represents a 5' untranslated sequence. This analysis demonstrates that at least some murine *Lvi-1* cDNAs are initiated upstream of the P1 promoter region of the *Lvi-1* gene and matches the other murine and human cDNAs downstream of exon 2, but the precise transcription initiation site remains to be established.

#### 4.3.4 Identification of the transcription initiation site of the murine *Lvi-1* gene

In order to identify the transcriptional start site of the murine *Lvi-1* gene, 5'RACE



Exon1a

AK020004	-----TGCAGCAAGTTT-----CCC	CGCTTCGACATTTTCCAGA	ACTGG	}	Murine	
AK051037	GGAGATTTGCGCGCCTTTCGAGCAAGTTC	-----CCC	CGCTTCGACATTTTCCAGA			ACTGG
AK031592	-----GCCC	CGCTTCGACATTTTCCAGA	ACTGG			
mLvi-1	TGAGATTTGCGCGCCTTTCGAGCAAGTTT	-----CCC	CGCTTCGACATTTTCCAGA			ACTGG
Mouse consensus	-----CCC	CGCTTCGACATTTTCCAGA	ACTGG			
AK023035	-----A--	GCCGCGC-CTA-AGAAGCC-GAA--AG		}	Human	
BC051855	-----	CCGCGC-CTA-AGAAGCC-GAA--AG				
BC025247	-----	GCGCCGCGC-CTA-AGAAGCC-GAA--AG				
Human consensus	-----	CCGCGC-CTA-AGAAGCC-GAA--AG				

Exon1c

AK020004	AAGGTCCT--CCTGTC-TGACAGTTGTTTCTTTTGGTTTTTGGTAAG-GTCC----	AAGA
AK051037	AAGGTCCT--CCTGTC-TGACAGTTGTTTCTTTTGGTTTTTGGTAAG-GTCC----	AAGA
AK031592	AAGGTCCT--CCTGTC-TGACAGTTGTTTCTTTTGGTTTTTGGT-----	
mLvi-1	AAGGTCCT--CCTGTC-TGACAGTTGTTTCTTTTGGTTTTTGGTAAG-GTCC----	AAGA
Mouse consensus	AAGGTCCT--CCTGTC-TGACAGTTGTTTCTTTTGGTTTTTGG-----	
AK023035	ATG-TCCAGGTCGGGCGCGCGGCTGAGA-AGGCGGACTCCAGACAGCGACCCAGATGA	
BC051855	ATG-TCCAGGTCGGGCGCGCGGCTGAGA-AGGCGGACTCCAGACAGCGACCCAGATGA	
BC025247	ATG-TCCAGGTCGGGCGCGCGGCTGAGA-AGGCGGACTCCAGACAGCGACCCAGATGA	
Human consensus	ATG-TCCAGGTCGGGCGCGCGGCTGAGA-AGGCGGACTCCAGACAGCGACCCAGATGA	

↓Exon 2

AK020004	AGGTGA-----	ATGAATATAAAGAAAATGAACATATTGCT
AK051037	AGGTGA-----	ATGAATATAAAGAAAATGAACATATTGCT
AK031592	AGGTGA-----	ATGAATATAAAGAAAATGAACATATTGCT
mLvi-1	AGGTGA-----	ATGAATATAAAGAAAATGAACATATTGCT
Mouse consensus	AGGTGA-----	ATGAATATAAAGAAAATGAACATATTGCT
AK023035	AGGTAA-----	ATGAATATAAAGAAAATCAAAACATCGCT
BC051855	AGGTAA-----	ATGAATATAAAGAAAATCAAAACATCGCT
BC025247	AGGTGAGAAACGGACTGGTGCTCCAGGTAATGAATATAAAGAAAATCAAAACATCGCT	
Human consensus	AGGT-A-----	ATGAATATAAAGAAAATCAAAACATCGCT

↑Exon2

AK020004	TATACTTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
AK051037	TATACTTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
AK031592	TATACTTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
mLvi-1	TATACTTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
Mouse consensus	TATACTTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
AK023035	TATGTGTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
BC051855	TATGTGTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
BC025247	TATGTGTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT
Human consensus	TATGTGTCTCTGAGACCGATACAGATTACAACCTTTAAGGAAAACAGCTAAGGTTTATCTT

AK020004	TACCCGTTTTCACTCAGCAATTC	CAAGCTAGGCTACTTAAAGTTGTCCAAATCCCCGGTT
AK051037	TACCCGTTTTCACTCAGCAATTC	CAAGCTAGGCTACTTAAAGTTGTCCAAATCCCCGGTT
AK031592	TACCCGTTTTCACTCAGCAATTC	CAAGCTAGGCTACTTAAAGTTGTCCAAATCCCCGGTT
mLvi-1	TACCCGTTTTCACTCAGCAATTC	CAAGCTAGGCTACTTAAAGTTGTCCAAATCCCCGGTT
Mouse consensus	TACCCGTTTTCACTCAGCAATTC	CAAGCTAGGCTACTTAAAGTTGTCCAAATCCCCGGTT
AK023035	GCCCCTTTTCACTCAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAAATCCCTATCA	
BC051855	GCCCCTTTTCACTCAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAAATCCCTATCA	
BC025247	GCCCCTTTTCACTCAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAAATCCCTATCA	
Human consensus	GCCCCTTTTCACTCAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAAATCCCTATCA	

AK020004	GTAATAAATCTAGCAAAATCAGTGGTCCATAAGAAGAAAGATCGGAAGAAGACTCGTAGA
AK051037	GTAATAAATCTAGCAAAATCAGTGGTCCATAAGAAGAAAGATCGGAAGAAGACTCGTAGA
AK031592	GTAATAAATCTAGCAAAATCAGTGGTCCATAAGAAGAAAGATCGGAAGAAGACTCGTAGA
mLvi-1	GTAATAAATCTAGCAAAATCAGTGGTCCATAAGAAGAAAGATCGGAAGAAGACTCGTAGA
Mouse consensus	GTAATAAATCTAGCAAAATCAGTGGTCCATAAGAAGAAAGATCGGAAGAAGACTCGTAGA
AK023035	GAAAAGAATTCTAACAAATGAAGTGGCGTGAAGAAGACTAAAATAAAGAAAACCTGCAGA
BC051855	GAAAAGAATTCTAACAAATGAAGTGGCGTGAAGAAGACTAAAATAAAGAAAACCTGCAGA
BC025247	GAAAAGAATTCTAACAAATGAAGTGGCGTGAAGAAGACTAAAATAAAGAAAACCTGCAGA
Human consensus	GAAAAGAATTCTAACAAATGAAGTGGCGTGAAGAAGACTAAAATAAAGAAAACCTGCAGA

M

AK020004	AAGGTTTTAACTTCAAAAATGAAGGCCTTATCTTCCAAGGCAGATTCCTTCTGCTGAAA
AK051037	AAGGTTTTAACTTCAAAAATGAAGGCCTTATCTTCCAAGGCAGATTCCTTCTGCTGAAA
AK031592	AAGGTTTTAACTTCAAAAATGAAGGCCTTATCTTCCAAGGCAGATTCCTTCTGCTGAAA
mLvi-1	AAGGTTTTAACTTCAAAAATGAAGGCCTTATCTTCCAAGGCAGATTCCTTCTGCTGAAA
Mouse consensus	AAGGTTTTAACTTCAAAAATGAAGGCCTTATCTTCCAAGGCAGATTCCTTCTGCTGAAA
AK023035	AGGATTATACCTCCAAAAGATGAAAAACACATCTTCCAAGGCAGATTCACCGCTGCAAAAAT
BC051855	AGGATTATACCTCCAAAAGATGAAAAACACATCTTCCAAGGCAGATTCACCGCTGCAAAAAT
BC025247	AGGATTATACCTCCAAAAGATGAAAAACACATCTTCCAAGGCAGATTCACCGCTGCAAAAAT
Human consensus	AGGATTATACCTCCAAAAGATGAAAAACACATCTTCCAAGGCAGATTCACCGCTGCAAAAAT

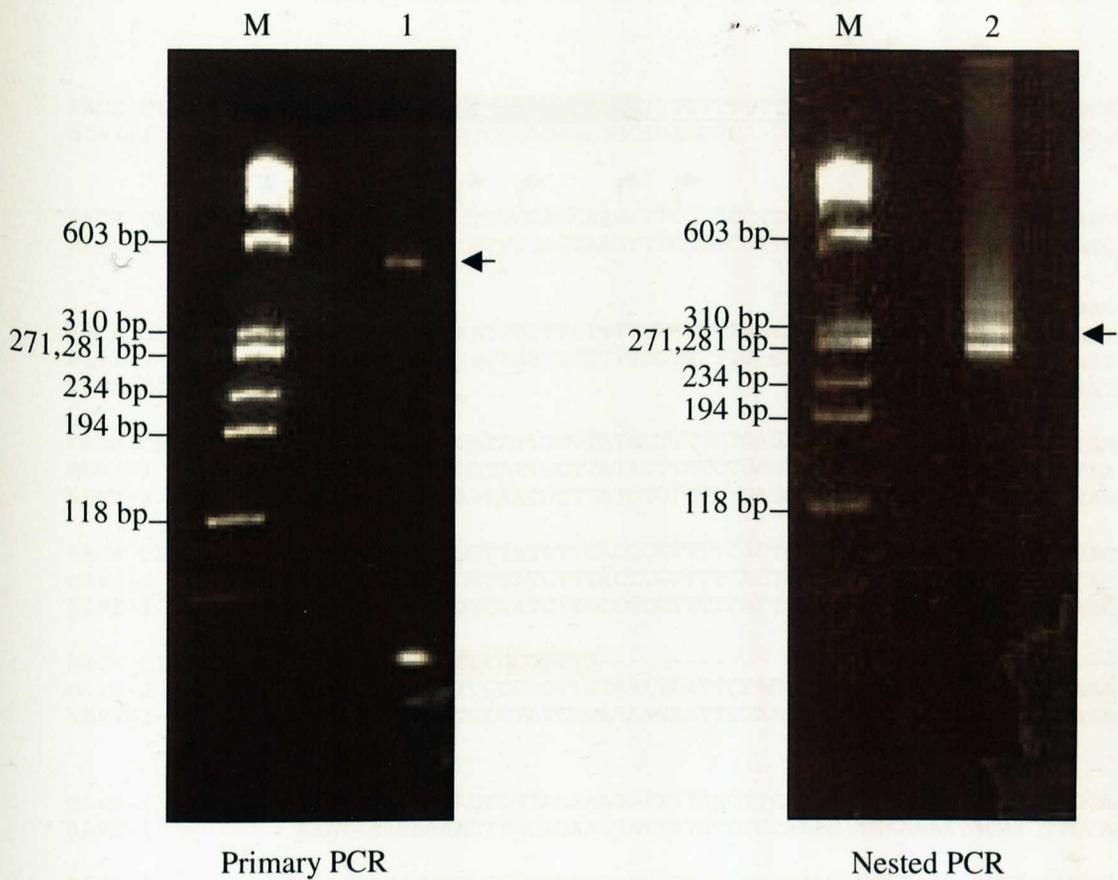
**Figure 4-8. Nucleotide sequence comparison of the 5'start of murine and human *LVI-1* cDNAs.** Published murine (AK020004, AK051037, AK031592), human (AK023035, BC05185, BC025247) and consensus sequence for murine thymus *Lvi-1* (*mLvi-1*) cDNAs is shown. Bold text indicates the mouse and human consensus sequences. Red text: first conserved in frame ATG start codon. Blue text: exon1c ATG present in all human cDNAs.

was performed using reverse primers complementary to murine *Lvi-1* exon 2. First-strand cDNA was synthesised from total mouse thymus RNA (Ambion) using the mouseFLJexon2RSP1 (Table 4-1). The amplification of the dA-tailed cDNA was carried out using the mouseFLJexon2RSP2 (Table 4-1) and the oligo dT anchor primer (Figure 4-9, lane 1); followed by a nested PCR with the specific primer 3 (SP3) mouseFLJexon2RSP3 (Table 4-1) (Figure 4-9, lane 2). Amplified fragments following the nested PCR were recovered by ligation into pCR2.1-TOPO vector, and 15 clones sequenced. The sequences were compared with the consensus sequence for murine thymus *Lvi-1* (*mLvi-1*, see Figure 4-7) and seven presumptive transcriptional initiation sites were identified. One out of the fifteen 5'RACE clones (clone 23), revealed sequence which extended back to -358 nucleotides from the conserved translation initiation codon ATG in exon 2 (Figure 4-10). Seven of the fifteen sequenced 5'RACE clones revealed an alternative transcriptional start site at -323 nucleotides upstream of the translation start site ATG. Of the remaining 7 clones 4 showed an alternative transcriptional start site located at position -327 or -316; the remaining 3 showed an alternative transcriptional start site located at position -352, -312, and -295. This analysis would indicate that the major transcriptional start site is located at position -358, however 47% of the clones sequenced revealed a transcriptional start site located at position -323.

No RACE clones were found to contain sequences derived from either exon1b or exon1c (Figure 3-2) demonstrating that expression of the murine *Lvi-1* gene is driven predominantly by the P1 promoter.

#### **4.3.5 Identification of the transcription initiation site of the human *LVI-1* gene**

The presumptive 5'exon of human *LVI-1* is conserved in mammalian genomes (human, chimp, mouse, rat) and contains the N-terminus favoured by gene prediction algorithms. However, a 5'RACE was performed using reverse primers complementary to human *LVI-1* exon 2. First-strand cDNA was synthesised from total RNA isolated from the Jurkat cell line using the humanFLJexon2RSP1 (Table 4-1). Amplification of the dA-tailed cDNA was carried out using the humanFLJexon2RSP2 (Table 4-1) and the oligo dT anchor primer (Figure 4-11, lane 1); followed by a nested PCR with the gene-specific primer 3 (SP3)



**Figure 4-9. Determination of the murine *Lvi-1* transcription start site.** cDNA was prepared using total RNA from normal mouse thymus. Primary and nested PCR products are shown following separation by 4% polyacrylamide gel. Lane 1: primary PCR product using the oligo dT-anchor primer and mouseFLJexon2RSP2 primer. Lane 2: nested PCR using PCR anchor primer and mouseFLJexon2RSP3. M is the  $\Phi$ X174RF DNA size marker.

*Mfap1* ↓

m*Lvi-1* TGACAGAAACAACCTCCGGTAGGGGGTAGAGGAACCAGAAGTACATAACACATAACAGAG

RACE C123 GACCACGCGTATC--GATGTCGACTTTT TTTT TTTT TTTT TTTT TTTT CGCTTAAAGGGCTAGGGTGA  
m*Lvi-1* GATTGTGGGTAATCGGAGGGCGGCTGAGCGC-----GGATCGCTTAAAGGGCTAGGGTGA

RACE C123 GATTTGCGCGCCTTTTCGAGCAAGTTTCCCGCCTTCGACATTTTCCAGAACTGGAAGGTCC  
m*Lvi-1* GATTTGCGCGCCTTTTCGAGCAAGTTTCCCGCCTTCGACATTTTCCAGAACTGGAAGGTCC

RACE C123 TCCTGTCTGACAGTTGTTTCTTTTGGTTTTTGGCAAGGTCCAAGAAGGTGAATGAATATA  
m*Lvi-1* TCCTGTCTGACAGTTGTTTCTTTTGGTTTTTGGCAAGGTCCAAGAAGGTGAATGAATATA  
h*LVI-1* GAAGGTAATGAATATA

RACE C123 AAGAAAATGAACATATTGCTTATACTTCTCTGAGACCGATACAGATTACAACTTTAAGGA  
m*Lvi-1* AAGAAAATGAACATATTGCTTATACTTCTCTGAGACCGATACAGATTACAACTTTAAGGA  
h*LVI-1* AAGAAAATCAAAACATCGCTTATGTGTCTCTGAGACCAGCACAGACTACAGTTTAAATAA

RACE C123 AAACAGCTAAGGTTTATCTTTACCCGTTTTCACTCAGCAATTCCAAGCTAGGCCTACTTA  
m*Lvi-1* AAACAGCTAAGGTTTATCTTTACCCGTTTTCACTCAGCAATTCCAAGCTAGGCCTACTTA  
h*LVI-1* AAACAGCTAAGGTTTATCTTTACCCGTTTTCACTCAGCAATTCCAAGCTAGGCCTACTTA

RACE C123 AGTTGTCCAATCCCGGTTG-----  
m*Lvi-1* AGTTGTCCAATCCCGGTTGTAATAATTCTAGCAAATCAGTGGTCCATAAGAAGAAAG  
h*LVI-1* TGTGCCCAAATCCCTATCAGAAAAGAATTCTAACAAATGAAGTGGCGTGTAAGAAGACTA

m*Lvi-1* ATCGGAAGAAGACTCGTAGAAAGGTTTTAACTTCAAAAATGAAGGCCTTATCTTCCAAGG  
h*LVI-1* AAATAAAGAAAACCTGCAGAAGGATTATACCTCCAAGATGAAGAAACACATCTTCCAAGG

m*Lvi-1* CAGATTCCCTTCTGCTGAAATCATCCGTGGATGCTTATACTGAAAGTACCAGGCTGGGAC  
h*LVI-1* CAGAATCCACGCTGCAAAATTCATCCTCAGCTGTTCATACTGAAAGTAAACAGCTACAAC

m*Lvi-1* CCAAGAGGACCTCGGATTAGCGACTTCTCAGTGTGGACGCGGAAAGCAGC---GACGAAG  
h*LVI-1* CCAAGAGAACGGCAGATGCGATGAATCTCAGTGTGATGTGGAAAGTAGTCAGGATGGAG

m*lvi-1* ACAGTG CACCAGGCCTG  
h*LVI-1* ACAGTGATGAAGATACCACACCAGCCCTG

-358

↓exon2

+1 →

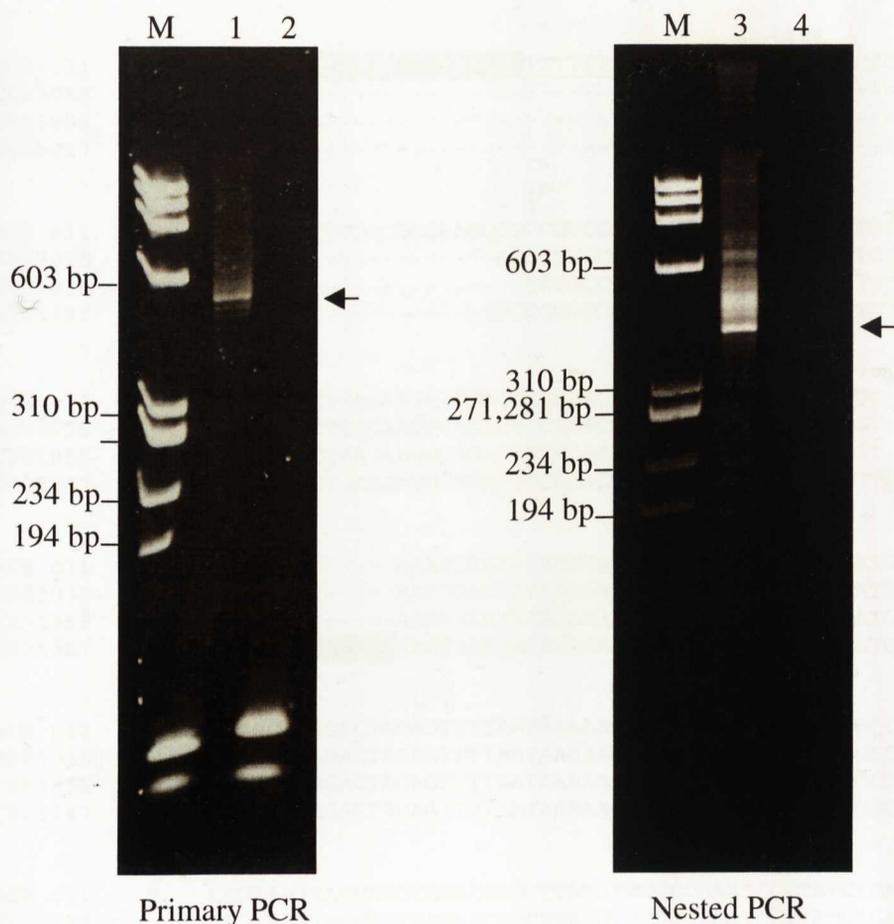
**Figure 4-10. Nucleotide sequence of the murine *Lvi-1* RACE clones.** The most upstream and downstream transcription start sites are indicated by red and black arrows, respectively. Underline letter: first nucleotide of *Mfap1* gene. Blue text: first nucleotide of each transcriptional start site. Red text: first conserved in frame ATG start codon. +1 and -128 indicate the coding start and transcription initiation site, respectively. The grey and yellow highlighted regions are the PCR anchor primer and mouseFLJexon2RSP3 primer. The underlined region represents the Oligo dT-anchor primer. The human sequence h*LVI-1* is the human cDNA AK023035.

humanFLJexon2RSP3 (Table 3-1) (Figure 4-11, lane 3). Amplified fragments following the nested PCR were recovered by ligation into pCR2.1-TOPO vector, and 4 clones sequenced. The sequences were compared with the human cDNA AK023035 and 4 presumptive transcriptional initiation sites were identified. One out of these four 5'RACE clone revealed sequence which extended back to -356 nucleotides from the conserved translation initiation codon ATG in exon 2 (Figure 4-12). The remaining 3 clones showed an alternative transcriptional start site located at positions -331, -313, or -305. This analysis would indicate that the major transcriptional start site is located at position -356. Interestingly, one of the clones showed the same transcriptional initiation site found in the published human BC051855 cDNA. Moreover, like the human cDNAs AK023035 and BC051855, all RACE clones lack the additional sequence observed only in the published human BC025247 cDNA. This sequence appears to be derived from the use of an alternative splice donor (see Figure 3-3).

In conclusion, 5'RACE analysis of the *LVI-1* gene has shown that the major transcriptional start sites differ between human and murine genes. The major transcription start site of the murine *Lvi-1* gene is located in exon 1a, under the control of the distal P1 promoter. In contrast, the proximal P2 promoter generally drives transcription of the human *LVI-1* gene with transcripts initiating in exon 1c.

#### **4.3.6 Promoter switching and alternative *LVI-1* exon usage**

As mentioned above, there is the possibility of two promoter regions in which, the human *LVI-1* transcript(s) are mainly transcribed from P2 promoter whereas in mouse the transcription appears to be driven by P1 promoter. Moreover, the murine homologue of the human P2 promoter appears to be a putative start site with the predicted amino acid sequence start MSGSK (Figure 3-3, exon 1c). Thus, the potential expression of the murine *Lvi-1* gene from the P2 promoter was explored by RT-PCR. Amplification was carried out on aliquots of 2.5  $\mu$ l of each murine cDNA (thymus, brain, lung, testis, liver, kidney, muscle, spleen, and heart) with 50 pmol of the primers: mFLJexon1cF and mouseFLJexon2RSP2 (Table 4-1 and Table 4-2). As a control for relative quantification of each mRNA transcripts, and in order to check the quality of the cDNA samples used, PCR was performed using the housekeeping



**Figure 4-11. Determination of the human *LVI-1* transcription start site.** cDNA was prepared using total RNA from the Jurkat cell line. Primary and nested PCR products are shown following separation by 4% polyacrylamide gel. Lane 1: primary PCR product using the oligo dT-anchor primer and human FLJexon 2RSP2 primer. Lane 3: nested PCR using PCR anchor primer and humanFLJexon2RSP3. Lanes 2 and 4: negative controls (PCR primers only). M:  $\Phi$ X174RF DNA size marker.

-356

RACE c11 GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTTCCGCCCCCTCCGGCCTCCCCG  
 AK023035 -----  
 BC051855 -----  
 BC025247 -----



RACE c11 CCGCAATCTTGGCGGGAAGGCGCCGGCCGCTAAGAAGCCGAAAGATGTCCAGGTCGGGC  
 AK023035 -----AGCCGGCCGCTAAGAAGCCGAAAGATGTCCAGGTCGGGC  
 BC051855 -----CCGGCCGCTAAGAAGCCGAAAGATGTCCAGGTCGGGC  
 BC025247 -----GCGCCGGCCGCTAAGAAGCCGAAAGATGTCCAGGTCGGGC

↓ exon2

RACE c11 GCGGCGGCTGAGAAGGCGGACTCCAGACAGCGACCCAGATGAAGGT-----  
 AK023035 GCGGCGGCTGAGAAGGCGGACTCCAGACAGCGACCCAGATGAAGGT-----  
 BC051855 GCGGCGGCTGAGAAGGCGGACTCCAGACAGCGACCCAGATGAAGGT-----  
 BC025247 GCGGCGGCTGAGAAGGCGGACTCCAGACAGCGACCCAGATGAAGGTGAGAAACGGACT

RACE c11 -----AAATGAATATAAAGAAAATCAAAACATCGCTTATGTGTCTCTGAGA  
 AK023035 -----AAATGAATATAAAGAAAATCAAAACATCGCTTATGTGTCTCTGAGA  
 BC051855 -----AAATGAATATAAAGAAAATCAAAACATCGCTTATGTGTCTCTGAGA  
 BC025247 GGTGCTTCCAGTAAATGAATATAAAGAAAATCAAAACATCGCTTATGTGTCTCTGAGA

RACE c11 CCAGCACAGACTACAGTTTTAATAAAAAACAGCTAAGGTCTATCTTGCCCCCTTTTCACT  
 AK023035 CCAGCACAGACTACAGTTTTAATAAAAAACAGCTAAGGTCTATCTTGCCCCCTTTTCACT  
 BC051855 CCAGCACAGACTACAGTTTTAATAAAAAACAGCTAAGGTCTATCTTGCCCCCTTTTCACT  
 BC025247 CCAGCACAGACTACAGTTTTAATAAAAAACAGCTAAGGTCTATCTTGCCCCCTTTTCACT

RACE c11 CAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAATCCCTATCAGAAAAGAATTCTA  
 AK023035 CAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAATCCCTATCAGAAAAGAATTCTA  
 BC051855 CAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAATCCCTATCAGAAAAGAATTCTA  
 BC025247 CAGTAATTACCAGCTAGACCAGCTTATGTGCCCAAATCCCTATCAGAAAAGAATTCTA

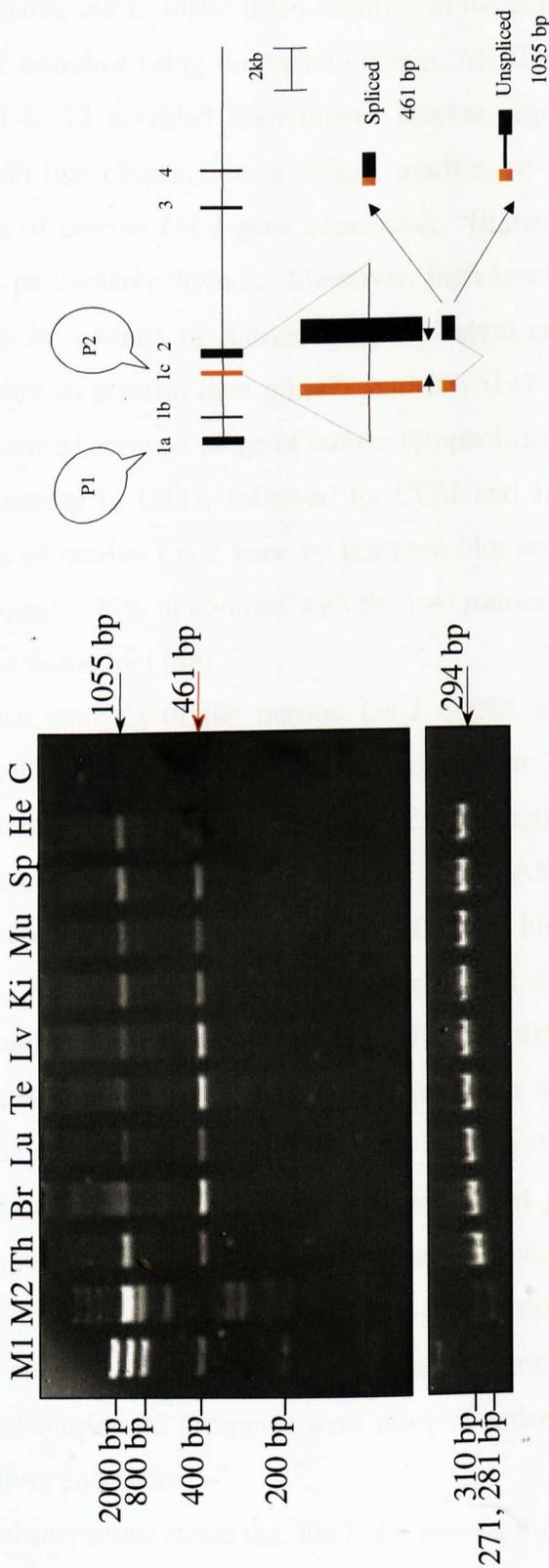
RACE c11 ACAATGAAGTGGCGTGTAAAGAAGACTAAAAATAAGAAAACCTTGAGAAGGATTATACCT  
 AK023035 ACAATGAAGTGGCGTGTAAAGAAGACTAAAAATAAGAAAACCTTGAGAAGGATTATACCT  
 BC051855 ACAATGAAGTGGCGTGTAAAGAAGACTAAAAATAAGAAAACCTTGAGAAGGATTATACCT  
 BC025247 ACAATGAAGTGGCGTGTAAAGAAGACTAAAAATAAGAAAACCTTGAGAAGGATTATACCT

+1→

RACE c11 CCAAAGATGAAAAACACATCTTCCAAGGCAGAATCCACGCT-----  
 AK023035 CCAAAGATGAAAAACACATCTTCCAAGGCAGAATCCACGCTGCAAAATTCATCCTCAGC  
 BC051855 CCAAAGATGAAAAACACATCTTCCAAGGCAGAATCCACGCTGCAAAATTCATCCTCAGC  
 BC025247 CCAAAGATGAAAAACACATCTTCCAAGGCAGAATCCACGCTGCAAAATTCATCCTCAGC

**Figure 4-12. Nucleotide sequence of human *LVI-1* RACE clones.** The most upstream and downstream transcription start sites are indicated by red and black arrows, respectively. Blue text: first nucleotide of each transcriptional start site. Red text: first conserved in frame ATG start codon. +1 and -356 indicate the coding start and transcription initiation site, respectively. The grey, green and yellow highlighted regions are the PCR anchor primer, extra nucleotide sequence from the human BC025247 cDNA and humanFLJexon2RSP3 primer, respectively. The underlined region represents the Oligo dT-anchor primer.

gene hypoxanthine-guanine phosphoribosyl transferase (*HPRT*), that it is generally expressed in high levels in most tissue types. *HPRT* was amplified using 1  $\mu$ l of each murine cDNA with 30 pmol of the primers HPRT5 and HPRT3 (Table 4-1 and Table 4-2). All tissues examined were positive for the presence of the alternative P2 promoter form; in addition an unspliced RT-PCR product is present and is probably due to genomic DNA contamination (Figure 4-13). Intriguingly, the relative expression levels did not correspond to northern blot analysis, which showed higher levels in lymphoid tissues, particularly thymus. Instead, brain, testis and liver showed the highest level of expression of the spliced exon 1c transcript. These results suggest that the murine *Lvi-1* gene may be expressed from either P1 or P2 promoters in a tissue-specific manner. In contrast, no expression was detected from the human P1 promoter by RT-PCR using the primer combination humanFLJexon1aF and humanFLJexon2RSP2 (Table 4-1 and Table 4-2), suggesting that the human gene is expressed only from P2.



**Figure 4-13. Expression from the P2 promoter of the murine *Lvi-1* gene.** cDNA was prepared using murine total RNA. RT-PCR products are shown following separation by 4% polyacrylamide gel. Lanes corresponding to the following tissues: Th, thymus; Br, brain; Lu, Lung; Te, testis; Lv, liver; Ki, Kidney; Mu, muscle; Sp, spleen, and He, heart cDNAs were amplified using the mFLJexon1cF primer and mouseFLJexon2RSP2 primer (upper panel), or using the mHPRT5 primer and mHPRT3 primer (lower panel). M1, M2, and C are the marker Low Mass DNA ladder, the marker  $\Phi$ X174RF DNA, and the negative control (PCR primers only). The black arrow indicates the unspliced RT-PCR product. The explanatory diagram in the panel on the right shows the original and splicing pattern of the P2 transcripts.

#### 4.4 Discussion

The aim of this chapter was to characterise the expression of the human and murine *LVI-1* genes and to relate these findings to the gene structure analysis in Chapter III. BLAST searches using the human cDNA AK023035 sequence coding for exons 10 through to 13 revealed homologous murine sequences on murine chromosome 2. Based on this observation, a unique murine *Lvi-1* probe was generated for further analysis of murine *Lvi-1* gene expression. Higher levels were detected in lymphoid tissues, particularly thymus. Moreover, high levels of expression of *Lvi-1* were also detected in a range of murine lymphoblastoid cell lines with the highest level of expression in p/m16i, then p/m47i, and BW5147. A similar expression pattern was also observed across a range of human lymphoblastoid cell lines with the highest level of expression in U937, followed by CEM and Jurkat. In addition, the expression patterns of murine *Lvi-1* gene by northern blot analysis detected the presence of one transcript (4.5 kb), in contrast with the two transcripts (2.8 kb and 4.5 kb) detected in a human tissue blot (58).

Sequence analysis of the murine *Lvi-1* cDNA confirmed that the murine gene is expressed in a similar manner to its human orthologue, a finding which has subsequently been confirmed by further full-length sequences lodged in the database (human cDNA AK023035 and murine cDNA AK031592). Furthermore, this study shows that the genomic organization of *LVI-1* is highly conserved between human and mouse. 5'RACE confirmed the transcriptional start sites of the murine and human genes and supported evidence from EST and cDNA clones that the murine gene is expressed predominantly from the P1 promoter while the human gene is expressed from the P2 promoter. While RT-PCR cloning showed that some murine transcripts may be initiated from within the adjacent *Mfap1* gene, 5'RACE suggested that most transcripts are initiated 3' to this gene. In addition, direct analysis by RT-PCR showed that the murine gene could also be expressed from the P2 promoter, while no transcripts were detected from the human P1 promoter region. Moreover, transcripts from the murine P2 promoter were more abundant in non-lymphoid tissues, notably brain, liver and testis.

These observations argue that the *Lvi-1* gene is differentially expressed and regulated, with expression of the murine gene mainly coming from the highly lymphoid specific P1 promoter, while the human gene is expressed from the more widely utilised P2 promoter. These findings also have significant implications for *Lvi-1* protein

expression as the P2 transcripts are predicted to express larger proteins with a distinct N-terminal sequence (Figure 3-6). "

All the results in this chapter are centred on the expression of human and murine *LVI-1* gene. It was considered important to extend this analysis to study protein expression, as the nature and abundance of the protein products could not be predicted accurately from these RNA-based studies.

## Chapter V

### Characterisation of LVI-1 protein products

#### 5.1 Introduction

At the outset of this study, no information was available on the protein products of the *LVI-1* gene. Analysis of human and murine *LVI-1* gene cDNA clones showed that the major human *LVI-1* transcripts have an open reading frame of 626 amino acids encoding a protein of 69.75 kDa, while the murine *LVI-1* transcript has an open reading frame of 524 amino acids encoding a protein of 57.96 kDa (Figure 3-5, 3-6). However, I have shown that the P2 promoter of the murine gene is also active, with the potential to encode a larger protein of 69 kDa, while the use of an internal AUG in the human mRNA would generate a product of 58.78 kDa. To characterise the products directly rabbit polyclonal antisera were generated using the GST gene fusion system. The specificities of these antisera were then tested by comparison of western blot signals before and after blocking with the homologous GST fusion protein immobilised on glutathione sepharose beads. Further studies were carried out to compare protein and mRNA expression levels of the *LVI-1* gene in a panel of human and murine cell lines, in which the polyclonal antisera were used to examine the LVI-1 protein expression by western blotting. It was important to generate and validate these immunological reagents before subsequent use in studies of protein localisation and function.

#### 5.2 Material and methods

##### 5.2.1 Production of recombinant expression vector

The recombinant<sup>™</sup> plasmids: mLvi-1-pGEX (pGEX-5X-2 plasmid containing the coding sequence for the predicted immunogenic region of the murine *Lvi-1* gene) and hLVI-1-pGEX (pGEX-5X-1 plasmid containing the coding sequence for the predicted immunogenic region of the human *LVI-1* gene), were produced as described below. Peptides were selected from the predicted immunogenic regions of the LVI-1 proteins that are generally hydrophilic (regions in contact with an aqueous environment). The Protein Hydrophilicity/Hydrophobicity Search is available online at the Weizmann Institute's Bioinformatics & Biological Computing web page

(<http://www.weizmann.ac.il/>) and was the computer program used to predict the hydrophilic regions of the human and murine LVI-1 proteins. Although other domains were more hydrophilic, the C-terminal domains were chosen based on the laboratory's experience of the GST system and previous successes in generating antisera to C-terminal fusion fragments (271, 272).

#### ***Preparation of DNA encoding the hLVI-1 C-terminus***

1 µg of human genomic DNA from CEM cell line was used to amplify DNA encoding amino acids 540 to 617 of the human LVI-1 protein (hLVI-1) with 100 pmol of specific primers. Both primers contain an *EcoRI* restriction site at the 5' end to allow cloning of the amplified PCR product into the pGEX 5X-1 vector as detailed below. For primer sequence and PCR conditions see Tables 5-1a, 5-1b, and also Section 2.2.4. The amplified fragment was then purified using the QIAquick<sup>®</sup> PCR purification Kit (Qiagen) according to the manufacturer's protocol. The *EcoRI* fragment was obtained by digestion of the purified DNA with the restriction enzyme (as detailed in Section 2.2.2.) and then purified as described above. The insert DNA was then quantified by 4% polyacrylamide gel electrophoresis. Following staining with ethidium bromide and inspection under UV illumination DNA was quantified by comparison with the molecular weight standard DNA (ϕX174 RF DNA/HaeIII fragments).

#### ***Preparation of DNA encoding the mLvi-1 C-terminus***

DNA encoding amino acids 446 to 524 of the murine Lvi-1 protein (mLvi-1) was generated by digestion of the plasmid DNA containing exon 13 of the murine *Lvi-1* gene (Section 4.2.1) with the restriction enzyme *EcoRI* (as detailed in Section 2.2.2) followed by purification of the *EcoRI* fragment (as detailed in Section 2.2.2). The purified fragment was then quantified by 1% TBE agarose gel electrophoresis as for the hLVI-1 protein.

#### ***Preparation of the pGEX vector***

In order to prepare the vectors pGEX-5X-1 and pGEX-5X-2 for cloning the human and murine *LVI-1* DNA inserts respectively, 2 µg of each vector was digested with

**Table 5-1. Primers used for generation of hLVI-1-GST fusion constructs.**

**Table 5-1a. Primer sequences (by hand).**

Number of primer	Primer	Strand	Sequence (5'→3')
1	HumanFLJ13F	Forward	gct gaa ttc cac aac act ttc ac
2	HumanFLJ13R	Reverse	agc gaa ttc tat ctt ccc gct gg

**Table 5-1b. Primer combinations, PCR conditions, and applications.**

Primer combination	Product size	Annealing temperature (°C)	Application
1 + 2	252 bp	50	Genomic PCR

the restriction enzyme *EcoRI* (Section 2.2.2) followed by gel purification (Section 2.2.2). The purified and linearised vectors were then dephosphorylated by digestion with CIAP (Section 2.2.2) followed by purification using the QIAquick® PCR purification kit. Vector DNA was quantified by gel analysis of an aliquot as described above for the insert fragments.

### ***Cloning of the insert DNA into the pGEX vector***

The human and murine *LVI-1* DNA inserts were ligated into the pGEX-5X-1 and pGEX-5X-2 vectors, respectively (as detailed in Section 2.2.2). The ligation mix was then used to transform One Shot®TOP10 Chemically Competent *E. coli* cells (as detailed in Section 2.2.2). The mLvi-1-pGEX and hLVI-1-pGEX recombinant plasmids were identified by Grunstein-Hogness screening (as detailed in Section 2.2.2) using the  $\alpha$ -<sup>32</sup>P-labeled human *LVI-1* exon 13 DNA probe (Section 4.2.2 and 2.2.6) and  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe (Section 4.3.1 and 2.2.6), respectively. The presence of inserts in the correct orientation was verified by sequencing the clones containing the recombinant plasmids by using the primers pGEX5' and pGEX3' as detailed in Section 2.2.5. Moreover, mLvi-1-pGEX recombinant plasmids were also verified by dual restriction enzyme digestion with *XbaI* and *EcoRV* enzymes (Section 2.2.2) followed by electrophoresis through a 1% TAE agarose gel (Section 2.2.2).

## **5.2.2 Expression of LVI-1-GST fusion proteins**

### ***Transformation of E.coli BL21 cells***

Competent *E.coli* BL21 cells were freshly prepared and transformed with the mLvi-1-pGEX and hLVI-1-pGEX recombinant plasmids as described in Section 2.2.2. The cells were also transformed with the parental pGEX plasmid for use as a control in expression studies. Single colonies with the mLvi-1-pGEX and hLVI-1-pGEX recombinant plasmids were picked after overnight growth at 37 °C on LBAG agar plates and used for production of the hLVI-1-GST and mLvi-1-GST fusion proteins.

### ***Small scale expression of LVI-1-GST fusion proteins***

A small scale screening procedure was carried out to analyse protein expression of transformants of BL21 *E. coli* in combination with pGEX vectors. A single colony expressing the pGEX-LVI-1 fusion plasmids was inoculated into 5 ml of 2xYTA medium and incubated overnight at 37 °C with vigorous shaking. The overnight culture was diluted 1:100 into fresh pre-warmed 2xYTA medium (20 ml), and incubated at 37 °C with shaking until an A<sub>600</sub> of 0.8. Isopropyl β-D-thiogalactoside (IPTG) was added to a final concentration of 0.1 mM and cultures incubated at 37 °C for a further 2 hours with vigorous shaking. The bacterial culture was transferred to an appropriate centrifuge container and centrifuged at 4000 rpm for 10 minutes at 4 °C to sediment the cells. The supernatant was discarded and the bacterial pellet resuspended in 1 ml of ice-cold 1xPBS. Lysozyme was added to a final concentration of 1 mg/ml and the cells lysed by 10 cycles of alternative freezing in a dry ice /isopropanol bath and 37 °C warming to thaw. The supernatant was then removed to a clean tube prior to purification of the fusion protein by affinity chromatography (see below).

### ***Small scale purification of LVI-1-GST fusion proteins***

50% Glutathione Sepharose 4B slurry (Amersham Biosciences) was prepared according to the manufacture's instructions. 20 µl of 50% glutathione sepharose 4B slurry was added to the cleared lysate derived from 20 ml of bacterial culture and incubated at room temperature for 5 minutes with gentle agitation. The matrix was then sedimented by centrifugation at 500 g for 5 minutes, the supernatant discarded and the matrix resuspended in 100 µl of 1x PBS and sedimented by centrifugation at 500 g for 5 minutes. This was repeated a further two times for a total of three washes. Elution of the fusion protein was carried out by the addition of 10 µl of glutathione elution buffer 1 (Section 2.1.11). The sepharose beads were resuspended by gentle agitation at room temperature for 5 minutes and then sedimented by centrifugation at 500 g for 5 minutes. The supernatant was collected to a fresh tube. The elution and collection steps were repeated twice more for a total of three eluates.

### ***Analytical scale preparation of insoluble LVI-1-GST fusion proteins***

A single colony expressing the pGEX-LVI-1 fusion plasmid was inoculated into 1 ml of LB medium containing 100 µg/ml ampicillin and incubated overnight at 37 °C with vigorous shaking. 30 µl of overnight culture was added to 1.5 ml of LB (1:50 dilution) containing 100 µg/ml ampicillin and incubated at 37 °C with shaking for 90 minutes. IPTG was added to a final concentration of 0.1 mM and the culture incubated for a further 4 hours. The bacterial culture was transferred to an appropriate centrifuge container and centrifuged at 4000 rpm for 10 minutes at 4 °C to sediment the cells. The supernatant was discarded and the bacterial pellet washed once with 200 µl of ice-cold STE (Section 2.1.11), resuspended by repeated pipetting in 135 µl of STE containing 100 µg/ml of lysozyme, and incubated on ice for 15 minutes. Dithiothreitol (DTT) was then added to a final concentration of 5 mM. N-laurylsarcosine (sarkosyl) sodium salt to a final concentration of 1.5% from a 10% stock in STE was added and the solution vortexed for 5 seconds. The cell suspension was lysed by the freeze/thaw method as described above. The supernatant was transferred to a clean eppendorf tube, Triton X-100 added to a final concentration of 2% from a 10% stock in STE and the lysate vortexed for 5 seconds. 30 µl of washed and swollen glutathione agarose beads suspension (50% v/v in PBS) was added, and the lysate incubated at 4 °C on an Orbitron Rotator for 15 minutes. The beads were washed 6 times with ice-cold PBS by repeated centrifugation at 500 g for 5 minutes. The supernatant was aspirated, and ½ volume of ice-cold glutathione elution buffer 2 (Section 2.1.11) was added. The tip of a 24-gauge needle was used to make a hole in the bottom of the eppendorf. Applying positive pressure with air hose to the top of the eppendorf tube and collecting drops from the bottom captured eluted protein, in the absence of beads. Another ½ volume of elution buffer was added to the beads, the procedure was repeated, and the eluates pooled.

### ***Preparative scale preparation of insoluble LVI-1-GST fusion proteins***

A single colony expressing the pGEX-LVI-1 fusion plasmid was inoculated into 40 ml of LB (Luria broth) medium containing 100 µg/ml ampicillin and incubated overnight at 37 °C with vigorous shaking. 40 ml of overnight culture was added to 2000 ml of LB (1:50 dilution) containing 100 µg/ml ampicillin and incubated at 37 °C

with shaking for 90 minutes. IPTG was added to a final concentration of 0.1 mM and the culture incubated for a further 4 hours. The bacterial culture was transferred to an appropriate centrifuge container and centrifuged at 7000 g for 7 minutes at 4 °C to sediment the cells. The supernatant was discarded and the bacterial pellet washed once with 120 ml of ice-cold STE (Section 2.1.11), resuspended by repeated pipetting in 120 ml of STE containing 100 µg/ml of lysozyme, and incubated on ice for 15 minutes. DTT was then added to a final concentration of 5 mM. Sarkosyl was then added to a final concentration of 1.5% from a 10% stock in STE and the solution vortexed for 5 seconds. Cells were lysed by sonication as follows. The cell suspension was sonicated for approximately 30 seconds to disrupt the bacteria. A model Sonicator Ultrasonic Processor (Sonicator XL 2220 Labcaire Systems Ltd) equipped with a standard probe was used for sonication; the sonicator was tuned prior to use, following the manufacturer's instructions. The power level was generally set to between five and six (where level 10 were maximal) which avoided frothing, which may denature fusion proteins. A decrease in viscosity and slight darkening in colour of the cell suspension indicated the degree of sonication. Cell debris was pelleted by centrifugation at 10000 g for 10 minutes at 4 °C in a Beckman JA20 rotor. The supernatant was then removed to a clean tube, Triton X-100 added to a final concentration of 2% from a 10% stock in STE and the lysate vortexed for 5 seconds. 2.4 ml of washed and swollen glutathione agarose bead suspension (50% v/v in PBS) was added, and the lysate incubated at 4 °C overnight on an Orbitron Rotator. The beads were washed 6 times with 10 ml of ice-cold PBS by repeated centrifugation at 500 g for 5 minutes. The supernatant was aspirated, and ½ volume of ice-cold glutathione elution buffer 2 (Section 2.1.11) was added. The beads were resuspended by gentle agitation at room temperature for 10 minutes and then sedimented by centrifugation at 500 g for 5 minutes. The supernatant was collected to a fresh tube. The elution and collection steps were repeated twice more for a total of three eluates.

### ***Extraction of the solubilised LVI-1-GST fusion proteins from SDS-PAGE***

Equal volumes of sample and SDS-PAGE Protein 2x Sample Loading Buffer (Section 2.1.11) were vortexed, heated to 95 °C for 5 minutes and chilled on ice prior to loading on gel. All samples were separated on 10% Bis-Tris SDS-PAGE gels, stained with Electro-Blue™ Staining Solution from Qbiogene™ for 1 hour and then destained

with several ddH<sub>2</sub>O changes until a clear background was obtained. The gel slice containing the LVI-1-GST fusion proteins was excised, transferred to a ProteoPLUS™ tube from Qbiogene (that was previously hydrated, according to the manufacturer's instructions) and then the tube was filled with 800 µl of protein running buffer (the same buffer used to run the gel). The supporting tray containing the ProteoPLUS tube was then placed in a horizontal electrophoresis tank (Biomax QS710 Horizontal Unit) containing protein-running buffer and the protein electroeluted from the gel slice at 100 volts for 60 minutes. To increase the amount of protein recovered, the polarity of the electric current was reversed for 60 seconds. The solution was transferred to a clean tube and centrifuged at 14000 rpm for 1 minute to remove gel residues. The protein-containing solution was transferred to a clean tube.

#### ***Dialysis and concentration of the eluted LVI-1-GST fusion proteins***

The dialysis of the eluted LVI-1-GST fusion proteins was carried out using a ProteoPLUS™ tube from Qbiogene at 4 °C, overnight with two changes of the dialysis buffer (PBS). The sample was loaded into the ProteoPLUS™ tube (that was previously hydrated, according to the manufacturer's instructions) and placed in a stirred beaker containing sterile PBS (100× the volume of the protein sample).

The concentration of the GST fusion proteins was carried out using a Centricon Centrifugal Filter Devices from Millipore. The sample was added to the sample reservoir of the Centricon Centrifugal Filter Devices and centrifuged at 5000 g for 25 minutes at 4 °C. Then the filtrate vial was separated from the membrane support base, and the retentate vial placed over sample reservoir and then centrifuged at 1000 g for 2 minutes to transfer the concentrate into the retentate vial.

### **5.2.3 Identification of LVI-1-GST fusion proteins**

#### ***Identification on Coomassie stained gels***

Recombinant proteins were identified by SDS-PAGE gels followed by staining with Coomassie Brilliant Blue, as outlined in Section 2.2.7. Typically 10 µl of each sample was loaded into each well of a mini 12% SDS-PAGE gel. This was used to estimate the molecular size, yield and purity of the recombinant protein.

### ***Identification by western blot analysis***

In order to confirm the identity of the protein bands seen on Coomassie stained gels as an hLVI-1-GST and mLvi-1-GST fusion proteins, western blotting was performed by the protocol of ECL detection system (Section 2.2.7). Briefly, 1 µl aliquots of the gel purified sarkosyl solubilised mLvi-1-GST and hLVI-1-GST fusion proteins were separated on a standard 10% SDS-PAGE gel and transferred to a nitrocellulose ECL membrane. The membrane was blocked overnight, incubated with 1:1000 goat anti-GST antibody (Amersham Pharmacia Biotech) for 1 hour at room temperature, washed 3-4 times in 1x TBST then incubated with 1:2000 HRP conjugated rabbit anti-goat IgG polyclonal antibody (DakoCytomation) for 1 hour at room temperature, prior to immunodetection.

### **5.2.4 Mass spectrometry analysis**

LVI-1-GST fusion proteins aliquots equivalent to 185 ng were resolved on NuPAGE® Novex 4-12% Bis-Tris gel (Invitrogen Life Technologies) as described in Section 5.2.5, and stained with the SimplyBlue Safe Stain (Invitrogen Life technologies) according to the manufacturer's protocol. Bands corresponding to the LVI-1-GST fusion proteins were excised from the gel, placed into a 1.5 ml eppendorf tube and sent to Sir Henry Wellcome Functional Genomics Facility (SHWFGF) in the University of Glasgow for mass spectrometry by MALDI-TOF (matrix assisted laser desorption ionisation-time of flight) using the mass spectrometer, Voyager-De PRO Workstation from Applied Biosystems. Briefly, protein identification by mass spectrometry requires the digestion of the protein of interest with trypsin, a serine protease that specifically cleaves at the carboxylic side of lysine and arginine residues. The resultant mixture of peptides displays a mixture of distinct molecular weights that after elution can be accurately measured by MALDI-TOF. In fact, the MALDI produces a peak list in which each peak corresponds to the exact mass of a peptide ion. The resulting peak molecular weights from the MALDI spectrum were then used to search for corresponding proteins using the Mascot program available on the Matrix Sciences web page (<http://www.matrixscience.com>).

## **5.2.5 Measurement of purity and estimation of concentration**

### ***Estimation of LVI-1-GST fusion proteins concentration by silver staining***

Equal volumes of 1:40 diluted LVI-1-GST fusion proteins and different dilutions (3.9-780 ng) of protein standard BSA were resolved on a 1.0 mm NuPAGE® Novex 4-12% Bis-Tris Gel (Invitrogen Life Technologies) using 1x NuPAGE® MES (2-(N-morpholino) ethane sulfonic acid)-SDS Running Buffer (Invitrogen Life Technologies) as detailed in the NuPAGE® Technical Guide from Invitrogen Life Technologies. After gel electrophoresis, the proteins were silver stained using the SilverQuest Silver Staining Kit (Invitrogen Life Technologies) according to the manufacturer's protocol.

### ***Estimation of LVI-1-GST fusion protein by Bio-Rad Protein Assay***

Pooled preparations of LVI-1-GST fusion proteins were measured using the Bio-Rad Protein Assay as detailed in Section 2.2.7. Typically, BSA was diluted with ddH<sub>2</sub>O to concentrations ranging from 0.025 to 1.4 mg/ml to enable the production of a standard curve each time the assay was performed.

## **5.2.6 Production of polyclonal antibodies against GST fusion protein and its immunodetection by western blotting**

Immunizations of 2 New Zealand White-Barrier Reared rabbits per fusion protein proteins, (hLVI-1 and mLvi-1) were performed by Harlan Sera-Lab (Leicester, UK) in accordance with the Animals (Scientific Procedures) Act. The rabbits were immunized subcutaneously with 95.46 µg of hLVI-1-GST protein and 105 µg of mLvi-1-GST protein in Freund's Complete adjuvant and boosted a further 5 times, at 2 weekly intervals, with the same amount of antigen in Freund's incomplete adjuvant and the animals were sacrificed at 11 weeks. Bleeds were taken 7 days after each immunisation and the serum divided into aliquots for long term storage at -70 °C. The pre-immune and test bleeds were submitted to western blot analysis to confirm the anti GST fusion protein antibody status. Fusion protein samples and GST protein alone were separated on 10% SDS-PAGE gels and subsequently electroblotted onto a nitrocellulose ECL membrane (Section 2.2.7). Western blotting was performed as described previously (Section 2.2.7). The pre immune and test sera were diluted

1:1000 and 1:2000 and incubated on strip blots for 3 hours at room temperature. Protein A linked to HRP (Amersham Biosciences), was used as a secondary antibody according to the manufacturer's instructions.

## **5.2.7 Validation of polyclonal antisera**

### ***Sera affinity chromatography***

The intrinsic specificity of the test sera for their cognate proteins was assessed by the sequential removal of GST and then LVI-1-GST reactive moieties by affinity chromatography. Essentially Glutathione Sepharose beads (Amersham Biosciences) were prepared as 50% slurry according to the manufacturer's instructions and 4 ml coupled with 4 mg of GST protein (Sigma) by incubation with gentle agitation at 4 °C overnight. After coupling of GST, the beads were pelleted by centrifugation at 500 g for 5 minutes, washed with 10 bed volumes of PBS (1 bed volume is 0.5 x the volume of 50% glutathione sepharose slurry used) and repelleted by centrifugation at 500 g for 5 minutes. The wash was repeated twice for a total of three washes. 2 ml ice-cold PBS was added to the pellet to make 50% glutathione sepharose slurry coupled with GST. Then, 250 µl of neat or diluted (1:10, 1:100, and 1:1000) anti-mLvi-1-GST sera (7328 rabbit) or anti-hLVI-1-GST sera (7330 rabbit) were added to 1 ml of GST coupled slurry and the mixtures were incubated with gentle agitation at 4 °C overnight. The beads were collected by centrifugation 500 g for 5 minutes and the GST reactive depleted supernatant tested for reactivity against GST by western blotting of GST protein (1 µg and 0.1 µg,) at 1:3000 dilution for 4 hours at room temperature (Section 2.2.7). A second round of affinity chromatography was then performed using the respective LVI-1-GST fusion proteins as the immobilised substrate. 50% glutathione sepharose slurry was coupled with 76.5 µg mLvi-1-GST or 9.1 µg hLVI-1-GST fusion proteins as described above. The LVI-1-GST depleted supernatants were collected as before and compared with their original and GST-absorbed sera for reactivity against their respective LVI-1-GST fusion proteins.

### ***Detection of specificity***

To analyse the specificity of each antisera for their native proteins whole cell lysates were prepared from H/3T3 and Jurkat cells as described in Section 2.2.7. 30 µg of

lysate was loaded into multiple wells and resolved on a 10% SDS-PAGE gel. Parallel strip blots were probed against an appropriate dilution of neat, GST absorbed, and LVI-1-GST absorbed fractions and the reactive bands visualised by ECL detection (Section 2.2.7).

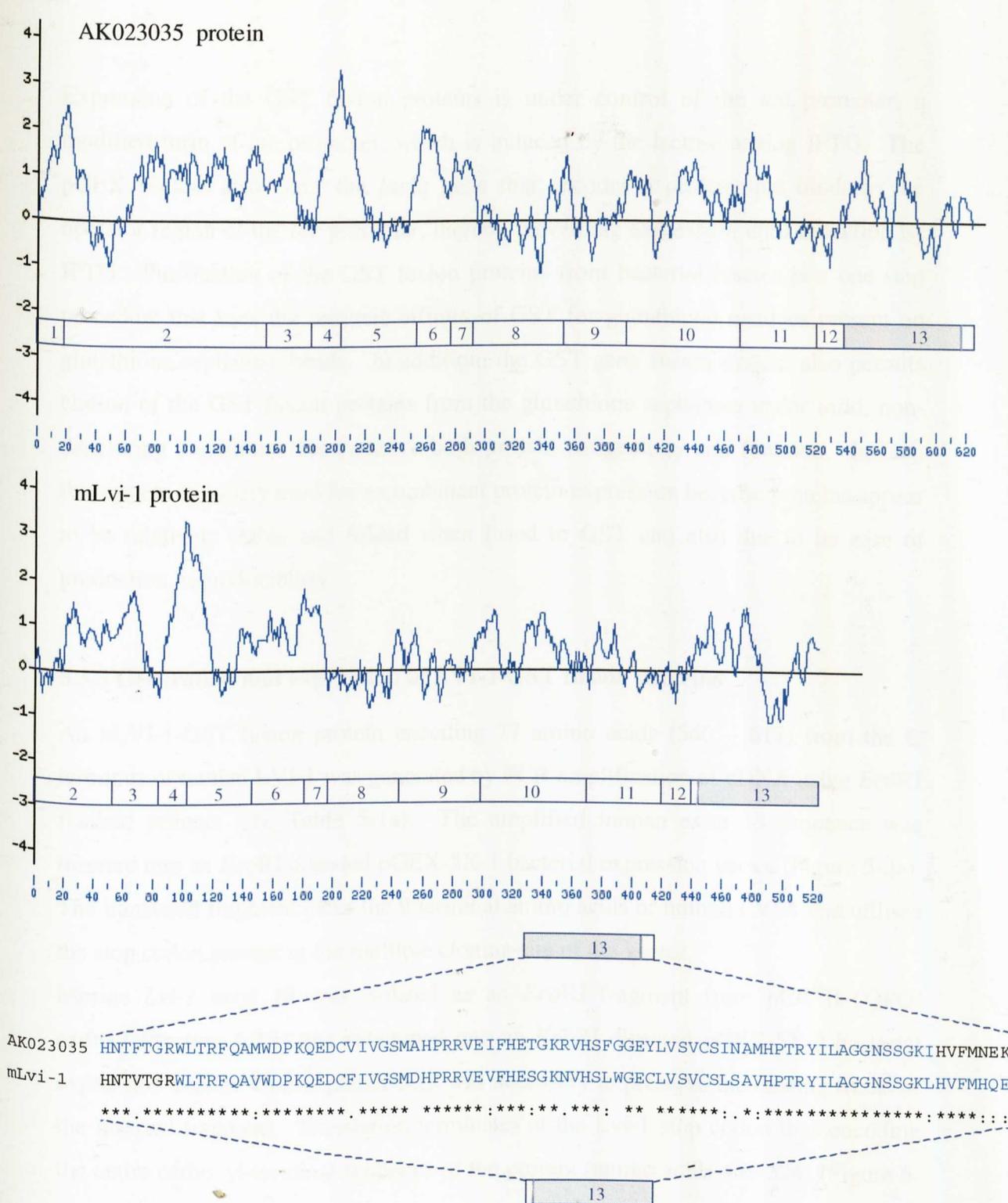
### **5.3 Construction and purification of LVI-1 fusion proteins**

#### **5.3.1 Peptide selection**

Immunogenicity or the capacity of a molecule to induce an immune response is determined by the intrinsic chemical structure of the injected molecule and its recognition by the host animal. In choosing a region of LVI-1 to generate a polyclonal antibody a number of independent factors were taken into account. Firstly a hydrophilicity plot of the human and murine amino acids sequences was generated to locate hydrophilic peptides (Figure 5-1). Previous studies (273-275) suggest that hydrophilic peptides are more likely to be exposed on the surface of native proteins and therefore be more antigenic to the host animal. Care must be taken with these plots because many hydrophilic amino acid sequences can be buried in water pockets or form intra or inter-molecular bonds thus excluding other potential interactions with anti-native antibodies. To minimise the risk of this we focussed on the extreme carboxyl (C) and amino (N) termini as these regions are commonly exposed in native proteins (276) and therefore make good antigenic targets. The murine and human amino acid sequences of LVI-1 differ most markedly at their amino termini so the carboxyl region encoded by the nucleotide sequence of exon 13 was identified as the main sequence of interest. It was predominantly hydrophilic and was sufficiently homologous (78% homology) between human and murine sequences to potentially generate an antibody with cross species reactivity.

#### **5.3.2 GST gene fusion system**

To produce a sufficient quantity of murine and human LVI-1 proteins, within a limited time scale, we used the GST gene fusion system for the expression, purification, and detection of fusion proteins (Figure 2-1). The system utilises pGEX vectors for inducible, high-level intracellular expression in *E. coli* of genes or gene fragments fused to the C-terminus of *Schistosoma japonicum* GST (277).



**Figure 5-1. Kyle-Doolittle amino acid hydrophilicity plot of AK023035 and mLvi-1 amino acid sequences.** The positive values indicate the hydrophilic regions. The numbered boxes indicate the translated exons of *LVI-1* gene. The blue stippled portions of exon 13 and blue lettered amino acid sequences letters indicate the peptide regions included in the GST fusion proteins. An alignment of human and murine amino acid sequences encoded by exon 13 is shown below.

Expression of the GST fusion proteins is under control of the *tac* promoter, a modified form of *lac* promoter, which is induced by the lactose analog IPTG. The pGEX vectors also carry the *lacIq* gene that encodes a protein that binds to the operator region of the *tac* promoter, thereby preventing expression until induction by IPTG. Purification of the GST fusion proteins from bacterial lysates is a one step procedure that uses the intrinsic affinity of GST for glutathione residues present on glutathione sepharose beads. In addition, the GST gene fusion system also permits elution of the GST fusion proteins from the glutathione sepharose under mild, non-denaturing conditions that preserve both protein antigenicity and function. In fact, this system is widely used for recombinant protein expression because proteins appear to be relatively stable and folded when fused to GST and also due to its ease of production, reproducibility.

### 5.3.3 Generation and expression of LVI-1-GST fusion proteins

An hLVI-1-GST fusion protein encoding 77 amino acids (540 – 617) from the C terminus of human LVI-1 was generated by PCR amplification of cDNA using *EcoRI* flanked primers (see Table 5-1a). The amplified human exon 13 sequence was inserted into an *EcoRI* digested pGEX-5X-1 bacterial expression vector (Figure 5-2a). The translated fragment lacks the 9 terminal amino acids of human LVI-1 and utilises the stop codon present in the multiple cloning site of the vector.

Murine *Lvi-1* exon 13 was isolated as an *EcoRI* fragment from pCR<sup>®</sup>II-TOPO/exon13 (Section 4.2.1) and subcloned into an *EcoRI* digested pGEX-5X-2 bacterial expression vector. Use of pGEX-5X2 was necessary to preserve the reading frame of the inserted fragment. Translation terminates at the *Lvi-1* stop codon thus encoding the entire carboxyl-terminal sequence of the protein (amino acids 446-524) (Figure 5-2b).

Expression of the fusion proteins was confirmed by transformation of *E. coli* BL21 cells with the recombinant vectors. Cell lysates were extracted pre and post IPTG induction and analysed for LVI-1 expression using Coomassie Brilliant Blue to visualise the proteins (Figure 5-3). The predicted molecular weights of human and murine LVI-1-GST fusion proteins are 37.03 kDa and 36.09 kDa respectively, of which 26 kDa is acquired from the GST moiety (<http://www.expasy.ch>). As can be seen in Figure 5-3 a predominant protein species was observed in each track that

a. Human *LVI-1* selected nucleotide sequence

```

F H N T F T G R W L T R F Q A M W D P K Q E D C V I
aattccacaacactttcactgggcatggctgaccaggttccaagccatgtgggatcctaacaagaagactgtgtcata
V G S M A H P R R V E I F H E T G K R V H S F G G E Y
gttggcagcatggccatccacgacgggtagaaatcttccatgagacaggaagagggtgcattcgtttggaggagaatac
L V S V C S I N A M H P T R Y I L A G G N S S G K I
cttgtctctgtgttccatcaatgccatgcaccaactcggatattttggctggaggttaattccagcgggaagatag

```

pGEX-5X-1 Ile Glu Gly Arg Gly Ile Pro Gly Phe Pro Gly Arg Leu Glu Arg Pro His Arg Asp  
Cloning site ATC GAA GGT CGT GGG ATC CCC GAA TTC CCG GGT CGA CTC GAG CGG CCG CAT CGT GAC TGA  
EcoRI Stop codon

b. Murine *Lvi-1* selected nucleotide sequence

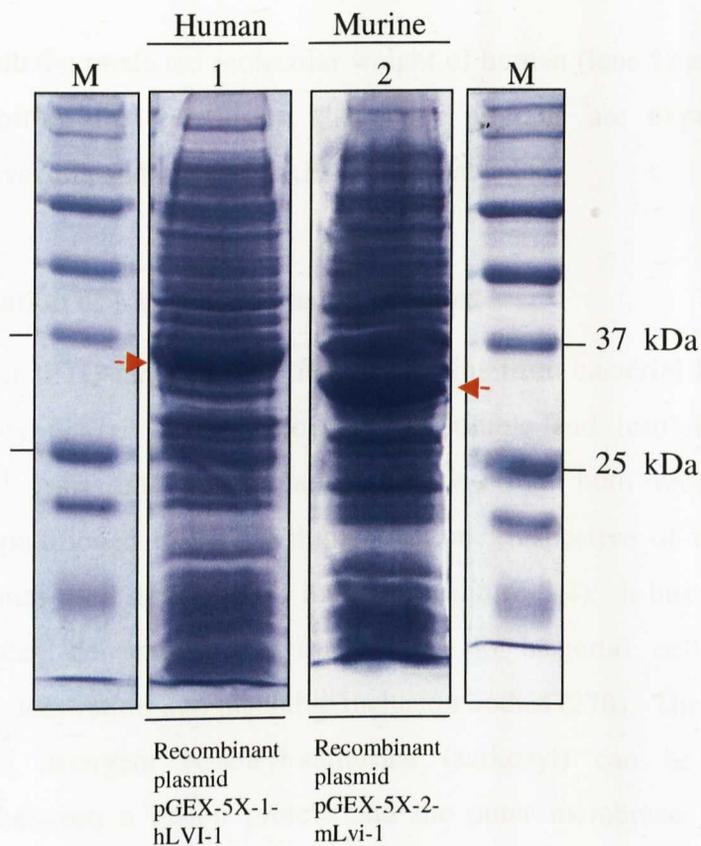
```

I R P W L T R F Q A V W D P K Q E D C F I V G S M D H
aattcgcccttggtcaacgaggttccaggctgtgtggaccctaacaagaagattgcttcatagttggtagcatggaccac
P R R V E V F H E S G K N V H S L W G E C L V S V C S
ccacgacgggtggaagtcttccacgagtcgggaaagaatgtgcattctcttggggagaatgtcttgcattctgtgttcc
L S A V H P T R Y I L A G G N S S G K L H V F M H Q E T
ctcagtgctgtgcattccacgcgatatactcgtggaggttaattccagtgcaagttacatgttttatgcatcaagaacc
tgagtcattggataggttcaacaaatgttcaactgttctgcagctagggaaacctagagactcagcttagtgcattctgttgg
tcacgtgttatatttagtaataataaactgcttctcagtgagacctccacacagaatgtaataattagcaagggaggaacttg
aagccatttctagaggatgagggagagacctggagcagaacgggtgttttcatgaagggcg

```

pGEX-5X-2 Ile Glu Gly Arg Gly Ile Pro Gly **Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser**  
Cloning site ATC GAA GGT CGT GGG ATC CCC GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG TGA  
EcoRI

**Figure 5-2. Strategy for construction of pGEX-5X-1-hLVI-1 and pGEX-5X-2-mLvi-1 plasmids.** a. Human *LVI-1* nucleotide sequence encoding the peptide sequence starting from 540 to 617 was digested with *EcoRI* restriction enzyme and inserted into *EcoRI* cloning site on pGEX-5X-1. b. Murine *Lvi-1* nucleotide of sequence encoding the peptide sequences starting from 446 to 524 from the pCR<sup>®</sup>-II-TOPO vector was digested with *EcoRI* restriction enzyme and inserted into *EcoRI* cloning site pGEX-5X-2. The blue, green, underlined, and red letters represent the *EcoRI* site, nucleotides encoded by the pCR<sup>®</sup>-II-TOPO vector, the stop codon, and the amino acids encoded by the vector pGEX-5X-2 that are not translated in the GST fusion protein, respectively.



**Figure 5-3. Induction of the hLVI-1-GST and mLvi-1-GST fusion proteins with IPTG.** Bacterial cells carrying the recombinant pGEX-5X-1 (lane 1) and pGEX-5X-2 (lane 2) plasmids were induced with IPTG, analysed by SDS gel electrophoresis on a 12% polyacrylamide gel and stained with Coomassie Brilliant Blue. Red arrows indicate the hLVI-1-GST and mLvi-1-GST fusion proteins. M denotes protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

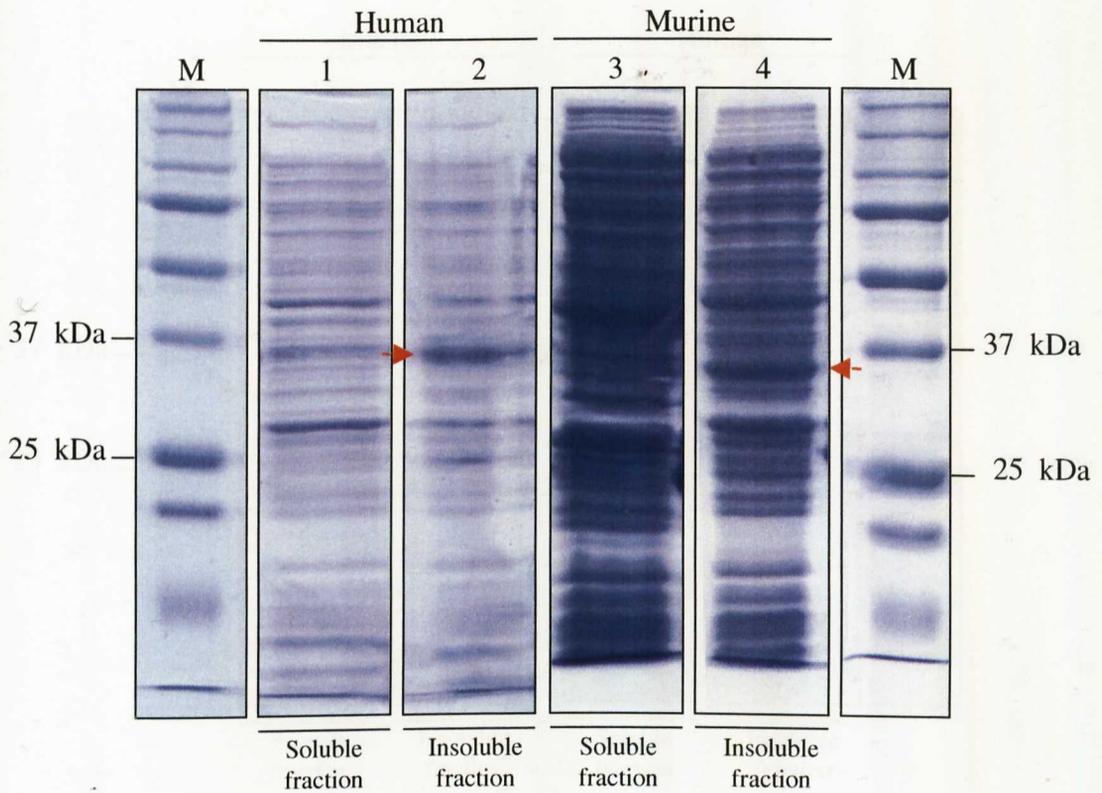
correlated with the predicted molecular weight of human (lane 1) and murine (lane 2) LVI-1 supporting the conclusion that both proteins are expressed from their recombinant vectors and are non-toxic to bacterial cells.

### **5.3.4 Purification of LVI-1-GST fusion proteins**

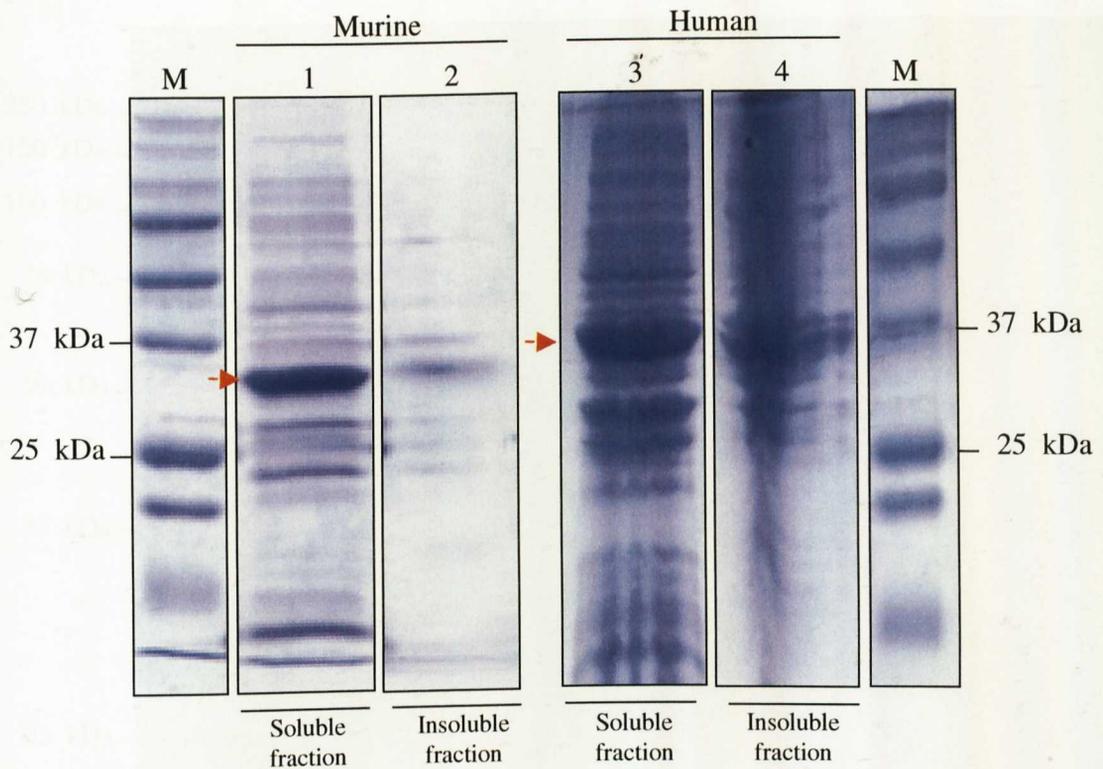
Purification of IPTG-induced GST fusion proteins from bacterial lysates depends on their solubility (277). Examination of the soluble and insoluble fractions after bacterial cell lysis and centrifugation revealed that both recombinant proteins consistently partitioned to the insoluble fraction irrespective of the temperature or incubation times used during IPTG induction (Figure 5-4). It has been reported that insolubility can be associated with binding to bacterial cell membranes and subsequent sequestration into insoluble inclusion bodies (278). The sodium salt of the alkyl anionic detergent N-lauryl-sarcosine (sarkosyl) can be used to disrupt interactions between a fusion protein and the outer membrane components or to prevent such an interaction from occurring if present during bacterial cell lysis (279). Using the recommended concentration of sarkosyl (1.5%) both human and murine LVI-1-GST partitioned to the soluble fraction after bacterial cell lysis (Figure 5-5). However, the presence of sarkosyl in the cell lysate severely reduced the affinity of the GST fusion proteins for glutathione sepharose beads (Figure 5-5). The non-ionic detergent Triton X-100 was used in an attempt to sequester sarkosyl into micelles and improve glutathione affinity but with little success. Consequently sarkosyl solubilised LVI-1 fusion proteins were typically resolved on PAGE SDS gels, visualised with "Simply Blue Safe Stain" (Invitrogen Life Technologies), electroeluted, dialysed and concentrated prior to use. The extracted proteins were verified by western blotting using an anti-GST antibody. Proteins corresponding to the predicted molecular weight for human and murine LVI-1-GST fusion proteins were detected indicating that the combination of sarkosyl solubilisation and gel extraction was a valid method for purification of these fusion proteins (Figure 5-6).

### **5.3.5 Validation by mass spectrometry**

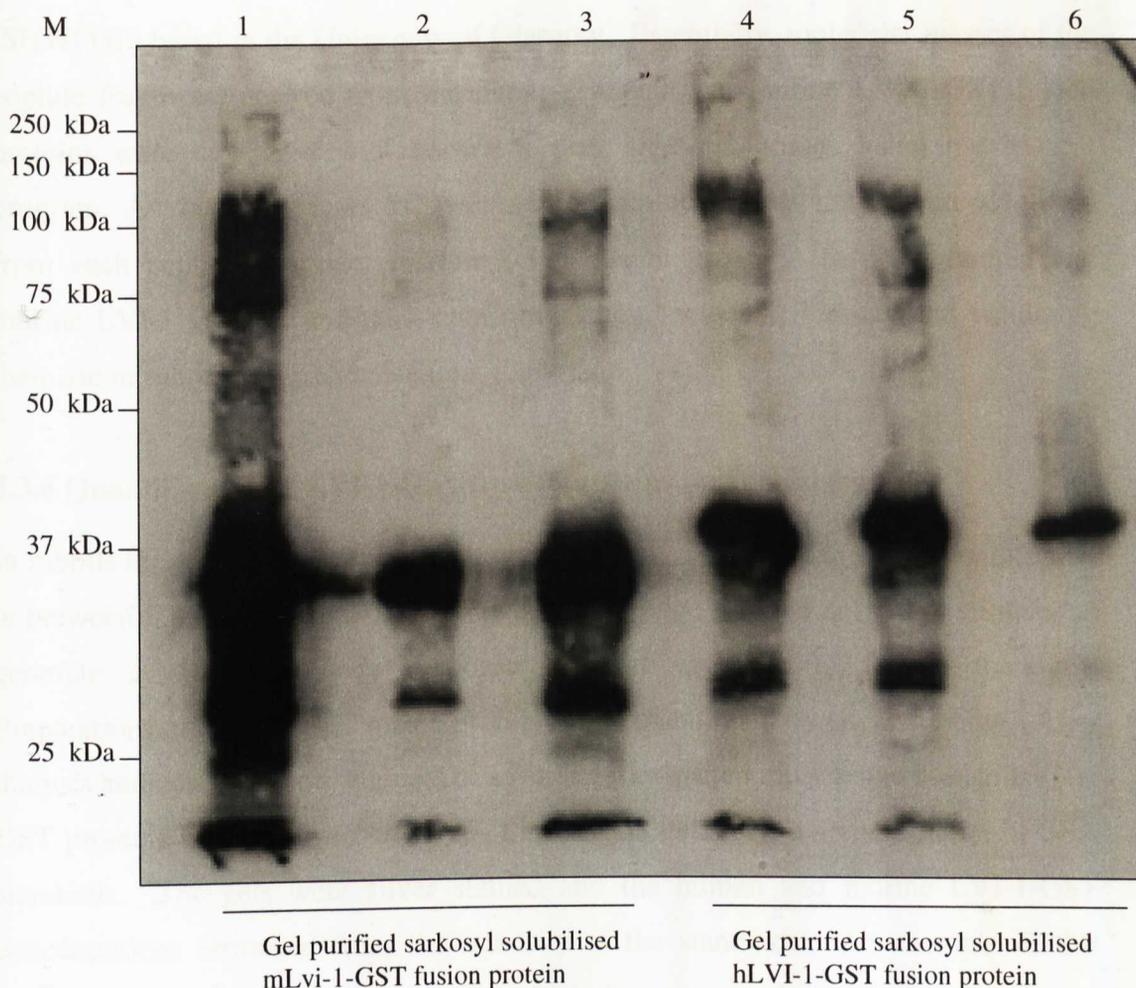
To confirm the molecular identity of both GST reactive proteins MALDI-TOF mass spectrometry was performed at the Sir Henry Wellcome Functional Genomics Facility



**Figure 5-4. Small scale purification of the human and murine pGEX recombinants for fusion protein expression.** Bacterial cells carrying the recombinant pGEX vectors expressing hLVI-1 (lanes 1, 2) and mLvi-1 (lanes 3, 4) fusion proteins were induced with IPTG and whole cell lysates analysed by 12% SDS-PAGE. The gels were stained with Coomassie Brilliant Blue. Lanes 1 and 3 are the cleared lysates (soluble fractions) from the bacterial cultures with hLVI-1 and mLvi-1, respectively. Lane 2 and 4 are the corresponding pellets (insoluble fractions). Red arrows indicate the hLVI-1-GST and mLvi-1-GST fusion proteins. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.



**Figure 5-5. Solubilisation of mLvi-1-GST and hLVI-1-GST fusion proteins.** Bacterial cell lysates expressing murine and human LVI-1 pGEX fusion proteins were solubilised with 1.5% sarkosyl. Lanes 1 and 2 show soluble and insoluble fractions from an analytical scale preparation of cells expressing the murine fusion protein. Lanes 3 and 4 show soluble and insoluble fractions from a preparative scale preparation of cells expressing the human fusion protein. Red arrows indicate the hLVI-1-GST and mLvi-1-GST fusion proteins. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.



**Figure 5-6. Western blot analysis of LVI-1-GST fusion proteins.** Gel purified sarkosyl solubilised mLvi-1-GST and hLVI-1-GST fusion proteins were separated on 10% SDS-PAGE and transferred to nitrocellulose membrane. The membrane was incubated with the anti-GST antibody followed by an anti-goat secondary antibody and processed for immunodetection. Lanes 1 to 3 and 4 to 6 represent the mLvi-1-GST fusion protein (36.09 kDa) and hLVI-1-GST fusion protein (37.03 kDa). The blot exposure time was 10 seconds. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

(SHWFGF) based in the University of Glasgow. Essentially, molecular masses of the peptide fragments derived from the putative human and murine LVI-1-GST fusion proteins were compared with known trypsin digest databases using the Mascot program. As shown in Figure 5-7 two predicted amino acid sequences were identified from each peptide fragment mixture (GST fusion proteins) for both human and murine LVI-1 proteins and thus, confirming their molecular identity and validating their use in subsequent immunisations.

### **5.3.6 Quantification of LVI-1-GST fusion proteins for immunisation**

In rabbits the recommended dose of protein antigen in adjuvant at each immunisation is between 0.05-1 mg (276). Doses within this range provide sufficient stimulus to generate a strong antibody response and are unlikely to induce tolerance. Preparations of LVI-1-GST fusion proteins were quantitated by silver staining. After dialysis and concentration, aliquots of several gel extracted murine and human LVI-1-GST proteins were resolved on a 12% PAGE/SDS gels with a serial dilution of BSA standards. The gels were silver stained and the human and murine LVI-1-GST concentrations estimated from the intensity of the standards. An example of this quantification is shown in Figure 5-8 in which the concentration of one of the samples of murine Lvi-1-GST protein was estimated at 0.78  $\mu\text{g}/\mu\text{l}$ . Parallel samples were pooled and the approximate concentrations confirmed using a modification of the Bradford method (260). In this way I estimated the concentration of the human and murine LVI-1-GST preparations to be approximately 0.74  $\mu\text{g}/\mu\text{l}$  and 3  $\mu\text{g}/\mu\text{l}$ , respectively.

### **5.3.7 Immunisation Schedule**

The production of polyclonal antibodies against human and murine LVI-1-GST fusion proteins was carried out by the Harlan Sera-Laboratory in Leicester (Section 5.2.6). The sera were tested for reactivity against the appropriate LVI-1 fusion and GST protein alone by western blotting. As expected the pre-immune sera gave no reactivity against either GST or the LVI-1 fusion protein (Figure 5-9). Reactivity was observed using the first test bleed of all 4 animals. Protein species that did not correspond to the predicted molecular weight of LVI-1-GST or GST alone were also

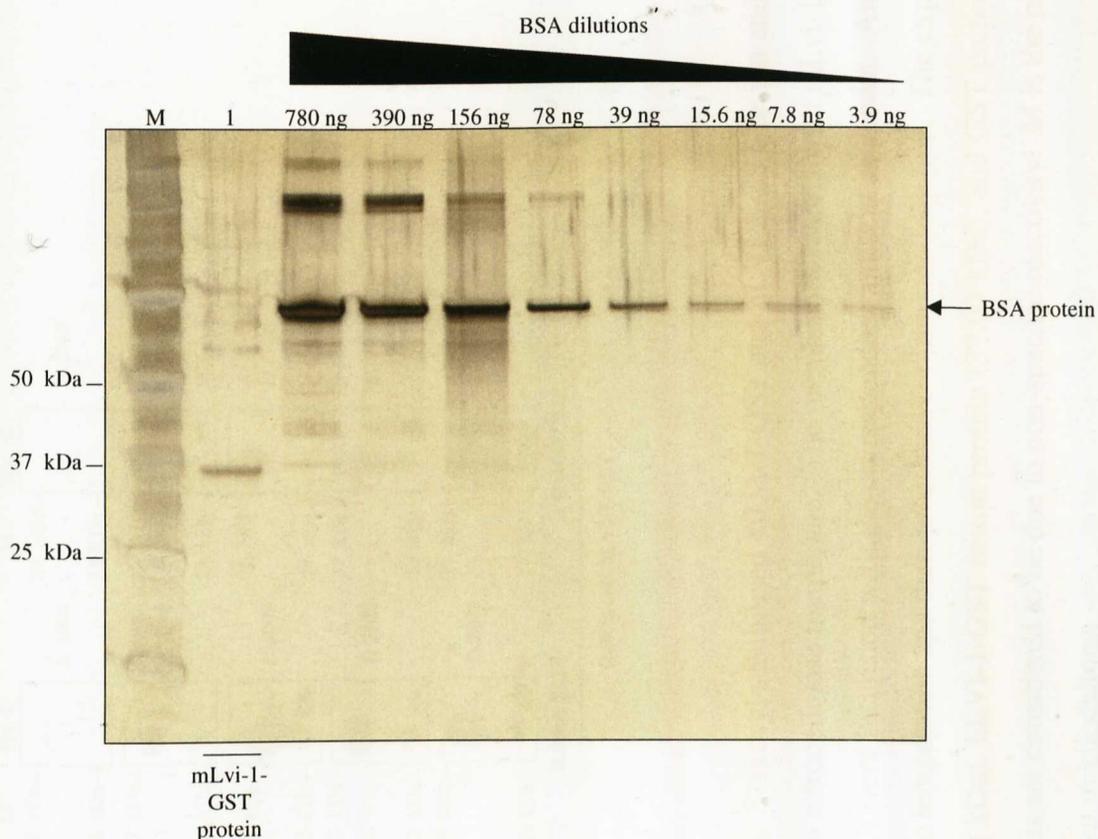
### Selected human LVI-1 peptide region

HNTFTGRWLTTR FQAMWDPK QEDCVIVGSMAPR R VEIFHETGK R VHSFGGEYLLVSVCSINAMHPTR YIILAGNSSGK I

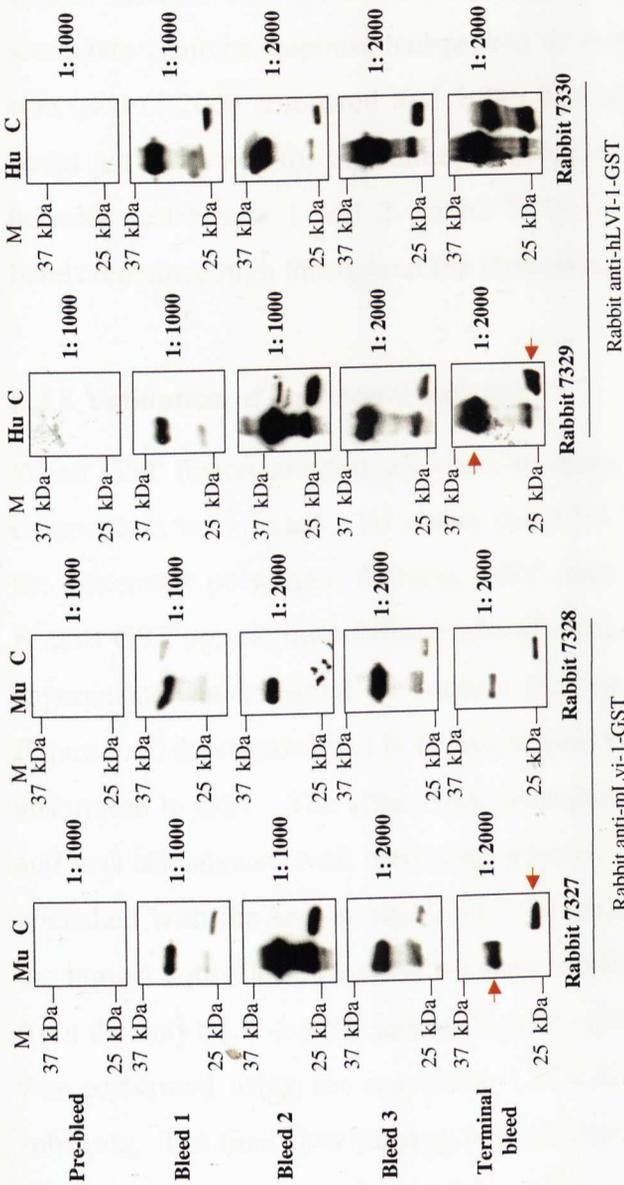
### Selected murine Lvi-1 peptide region

WLTTR FQAVWDPK QEDCFIVGSMDDPR R VEVFHEGK NVHSLWGECLVSVCSLSAVHPTR YIILAGNSSGK LHVFMHQET

**Figure 5-7. Identification of LVI-1-GST fusion proteins by mass spectrometry.** LVI-1-GST fusion proteins were resolved by gel electrophoresis, the band corresponding to their predicted molecular weights excised from the gel and digested with trypsin. The digestion products were purified and then analysed by mass spectrometry to determine their molecular weights. The Mascot program was then used as a search engine to identify the proteins resolved on the gel by comparison of their masses with known trypsin digest databases. Blue letters are the digested peptides identified by the Mascot program as specific peptides of the human and murine LVI-1 protein. Black arrows indicate the trypsin digestion site at the carboxyl side of lysine (K) and arginine (R) residues.



**Figure 5-8. Silver staining detection of mLvi-1-GST fusion proteins.** Equal volumes of aliquots of 1:50 diluted gel purified sarkosyl solubilised mLvi-1-GST fusion protein and different dilutions (3.9-780 ng) of BSA protein standard were resolved by NuPAGE<sup>®</sup> Novex 4-12% Bis-Tris gel and silver stained. Lane 1 contains the mLvi-1-GST fusion protein. The next lanes show dilutions of BSA, 1:5 (780 ng), 1:10 (390 ng), 1:25 (156 ng), 1:50 (78 ng), 1:100 (39 ng), 1:250 (15.6 ng), 1:500 (7.8 ng), and 1:1000 (3.9 ng). M is the protein molecular weight marker (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

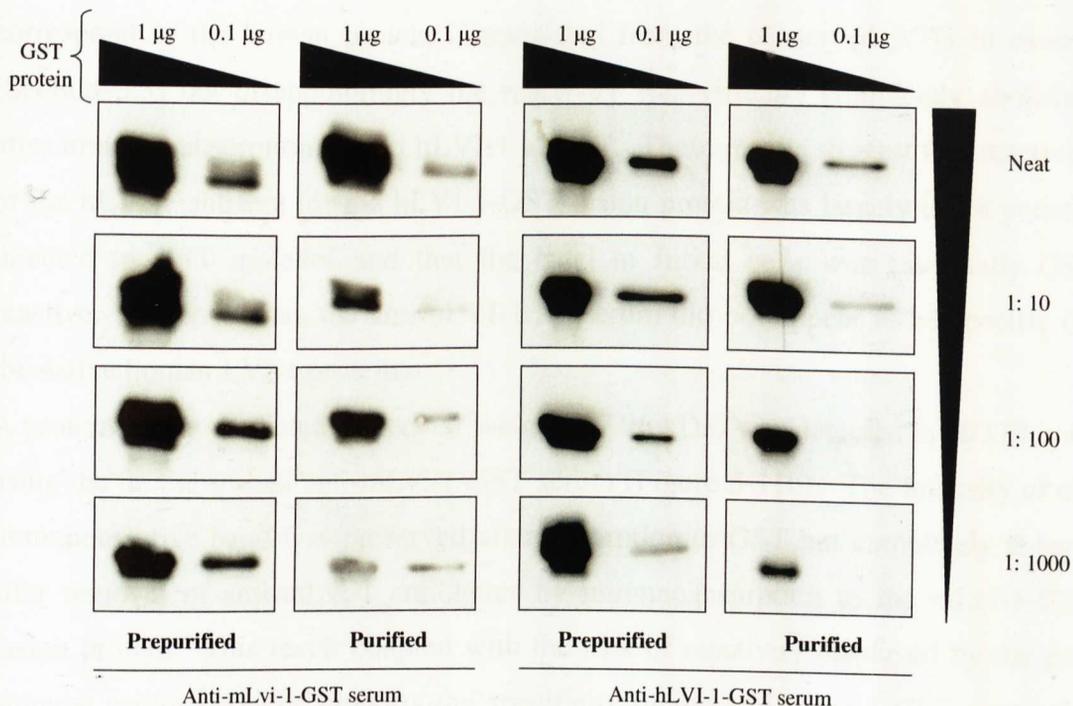


**Figure 5-9. Analysis of reactivities in rabbit sera during immunisation.** Sarkosyl solubilised mLvi-1-GST and hLVI-1-GST fusion and GST protein were gel purified, separated on 10% SDS-PAGE, and transferred to nitrocellulose membrane. The membranes loaded with mLvi-1-GST fusion protein (Mu), hLVI-1-GST fusion protein (Hu) and GST protein (C) were incubated with the corresponding antisera as shown. Antisera were used at the dilution indicated (1:1000 and 1:2000) for 3 hours at room temperature and then processed for immunodetection. The exposure time of the blots was 30 seconds. The mLvi-1-GST fusion protein (36.09 kDa), hLVI-1-GST fusion protein (37.03 kDa), and GST protein (26 kDa) are indicated by a red arrow only in the lower panels. The other bands are considered to be due to non-specific reactivity. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

detected. These persisted throughout the analysis of anti human LVI-1-GST sera (compare bleed 2 with terminal bleed for Rabbits 7329 and 7330) but were abolished when the terminal bleeds of rabbits injected with mLvi-1-GST (7327 and 7328) were used in the detection. In conclusion, it appeared that the specificity of the murine antisera increased during the immune response. This conclusion is supported by the increased signal intensity observed between test bleed 1 and 2 (rabbit 7327). A similar increase was not observed at later time points. This could indicate that the secondary immune response had peaked or more likely that the increased dilution of sera used (1:2000 compared to 1:1000) had ablated the effect. Similar conclusions could not be drawn for the human antisera, as despite an increased signal intensity between test bleeds 1 and 2 (rabbit 7329) the intensity of additional non-specific bands remained high throughout the immunisation regime.

### **5.3.8 Validation of polyclonal antisera**

When GST fusion proteins are used as immunogens an antibody response to both components is expected. To ensure that LVI-1 reactive species were present within the generated polyclonal antisera, GST reactive antibodies were immunoabsorbed against GST protein immobilised onto glutathione sepharose beads. The non reactive supernatant was screened by western blotting against GST protein. As shown in Figure 5-10 the reactivity of both sera against GST was severely compromised by pre-absorption to GST. The effect was more dramatic using the anti human LVI-1 sera and was accentuated with increasing dilution. GST reactivity was never completely abolished with the anti-murine LVI-1 antisera indicating a higher titre compared to the human equivalent. Because we were unable to remove all GST reactive moieties from the anti mLvi-1-GST serum (Figure 5-10) a second round of immunoabsorption was performed using the respective LVI-1-GST fusion protein as the immobilised substrate. The final flow through was compared with the original and GST absorbed fractions for their reactivity against LVI-1 using whole cell lysates prepared from Jurkat (human) and H/3T3 (murine) cells. These cell lines were chosen because they had been shown to express the *LVI-1* transcript in northern analysis of total RNA (Figure 4-5). The predicted molecular weights of murine and human LVI-1 proteins were estimated at 57.96 kDa and 69.75 kDa, respectively (<http://www.expasy.ch>). As



**Figure 5-10. Absorption of GST reactivity from antisera to LVI-1 fusion protein.**

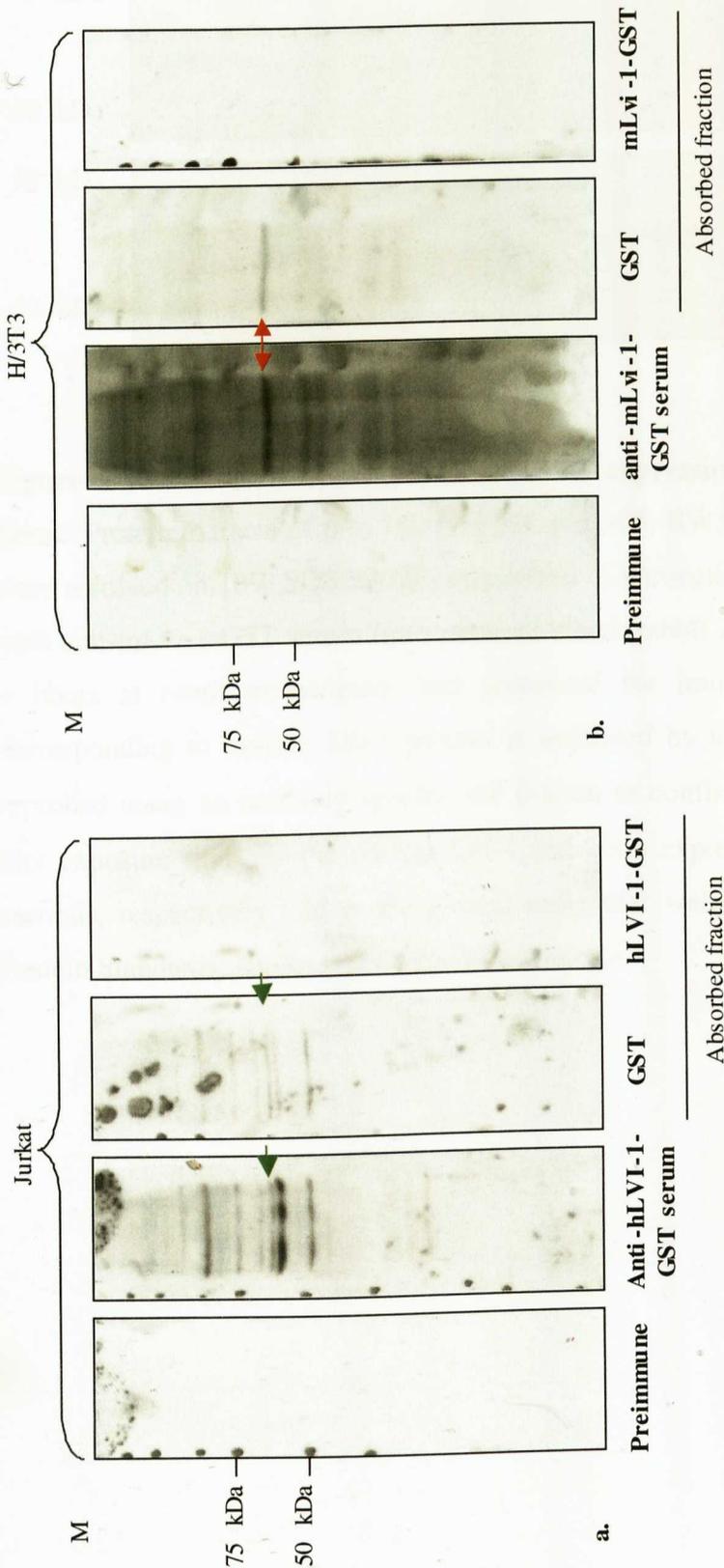
1 and 0.1  $\mu\text{g}$  of GST protein samples were separated by 10% SDS-PAGE, and immobilized onto a nitrocellulose membranes. The membranes were incubated with dilutions (neat, 1:10, 1:100 and 1:1000) of rabbit sera as shown. Prepurified and purified denote antiserum before and after absorption with GST. The blot exposure time was 15 seconds.

shown in Figure 5-11a a smaller protein of approximately 59.3 kDa was detected in Jurkat cells using the neat anti-hLVI-1-GST serum." The immunoreactive band would correspond to the human protein if translated from the conserved ATG in exon 2 (Section 3.3) but disappointingly the reactivity was virtually completely abolished after immuno-absorption of anti hLVI-1 to GST. These results suggest that reactivity of the hLVI-1 antisera for the hLVI-1-GST fusion protein was largely if not entirely directed to GST moieties and that the band in Jurkat cells was essentially GST reactive. In conclusion, the anti-hLVI-1 antiserum did not appear to be specific for the native human LVI-1 protein.

A protein of the predicted molecular weight (57.96 kDa) was detected in H/3T3 cells using the non-absorbed anti-mLvi-1-GST serum (Figure 5-11b). The intensity of the immunoreactive band was preserved after absorption to GST but completely ablated after removal of anti-mLvi-1 antibodies by immunoabsorption to the mLvi-1-GST fusion protein. This result coupled with the lack of reactivity displayed by the pre-immune serum strongly supports the specificity of the anti mLvi-1-GST serum for endogenous murine Lvi-1 protein and validates its use in subsequent expression analyses.

### **5.3.9 Endogenous expression of murine Lvi-1**

The expression of Lvi-1 protein was then examined using anti-mLvi-1-GST serum on a panel of murine cell lines previously shown to express the *Lvi-1* transcript (Figure 4-5b). As shown in Figure 5-12 a protein of the predicted molecular weight (57.96 kDa) was detected in each lysate including that of H/3T3 cells that had been used in the validation exercise. A strict correlation was not observed between the levels of Lvi-1 protein and mRNA (compare BW5147 with 3T3 and H/3T3) suggesting additional levels of post transcriptional control. The cross reactivity of anti-mLvi-1 was then tested against two human cell lines (WI-38 and Jurkat cells); however, no protein of the expected molecular weight (69.75 kDa) was detected (data not shown).



**Figure 5-11. Reactivity of antisera to LVI-1-GST fusion proteins after absorption.** Western blots of lysates from Jurkat (a.) or H/3T3 (b.) cells were used to test the effects of absorption with GST and GST fusion proteins, as indicated. Anti-hLVI-1-GST serum and anti-mLvi-1-GST serum refer to the terminal bleeds from rabbit 7330 and 7328, respectively. The mLvi-1 protein specific band around 58 kDa is highlighted by a red double arrowhead. A candidate human LVI-1 product of approximately 60 kDa which is not absorbed by GST, is indicated by a green arrow. The blot exposure time was 15 minutes. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.



**Figure 5-12. Western blot analysis of Lvi-1 expression in various murine cell lines.** Protein extracts of p/m 16i, p/m 32i, p/m 47i, BW5147, 3T3, and H/3T3 cells were resolved on 10% SDS-PAGE, transferred to nitrocellulose membrane, incubated with anti-mLvi-1-GST serum from terminal bleed (rabbit 7328) at 1:2000 dilution for 4 hours at room temperature, and processed for immunodetection. The band corresponding to murine Lvi-1 protein is indicated by a red arrow. The blot was reprobed using an antibody specific for  $\beta$ -actin to confirm the loading levels. The blot exposure time for the murine Lvi-1 and actin expression was 1 minute and 5 seconds, respectively. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

## 5.4 Discussion

The generation and subsequent validation of polyclonal antibodies for the characterisation of the *LVI-1* gene products revealed that only the anti-mLvi-1 serum displayed strong specificity for the cognate protein. One explanation for this could be the absence of the 9 carboxyl-terminal amino acids in the selected human fragment. The murine fragment included the extreme carboxyl terminus of the protein and looked more promising both in terms of titre and the specificity generated (presumably due to its greater antigenicity, as predicted). An alternative explanation could be the absolute quantities of LVI-1-GST fusion proteins used for the immunisation. The poor solubility of both fusion proteins required further solubilisation by sarkosyl and an alternative purification method by electroelution that inevitably decreased the overall yield of protein. As a result, the injected amounts used in the immunisation were at the lower end of the optimal range.

Although the validation exercise suggests that the anti-mLvi-1-GST serum is specific, failure to remove all the GST reactive moieties in the first round of immunoabsorption means that the possibility cannot be excluded that the detected band in H/3T3 cells was GST reactive and that the only difference between the human and murine antisera was their respective titre. To resolve this issue multiple rounds of immunoabsorption against GST protein would be necessary for a complete removal of GST reactive species before the final immunoabsorption against the mLvi-1-GST fusion protein. If successful the cross species reactivity of anti-mLvi-1-GST serum in hLVI-1-GST fusion protein could be tested. In addition, validation of the specificity of the antisera for the human and murine LVI-1 proteins could then be confirmed by testing each serum against *in vitro* translated versions of both LVI-1 proteins using a basic *in vitro* transcription kit. Although this approach could provide solid evidence of LVI-1 reactive moieties if reactivity is obtained and would help to resolve the cross reactivity of both antisera, it would not exclude the possibility of false negative results as protein folding and modifications of the *in vitro* translated proteins may not represent the *in vivo* situation.

In conclusion, definitive confirmation of LVI-1 antiserum specificity would then assist in characterisation of the native proteins. Determination of the sizes of the products by SDS-PAGE would also help in resolving uncertainties over the transcripts structures and coding potential (Chapter III and IV). Additionally, MALDI-TOF

could be used to confirm the identity of LVI-1 reactive proteins using immunoprecipitation to isolate LVI-1 reactive species from native cell extracts.

## Chapter VI

### Clues to LVI-1 protein function

#### 6.1 Introduction

Clues to mammalian gene function can often be found by comparison with homologous genes in model eukaryotes such as *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Saccharomyces cerevisiae*. Homologous genes in different species that have evolved from a common ancestor are called orthologues and these are generally expected to share the same function (reviewed in (280)). In addition, clues to the LVI-1 protein function could also be provided by comparison with homologues in the mammalian genome, the most closely related of which is DDB2, the small subunit of the DNA damage binding complex which is required for repair of UV-induced DNA damage and is defective in Xeroderma pigmentosum patient of complementation group E (XP-E) (58). However, identification of orthologues of LVI-1 is complicated by the fact that many functionally diverse proteins contain the WD repeat domain (219). It is therefore difficult to determine by sequence comparison alone which of the homologous genes in other species is a true orthologue. Nevertheless, computer-aided analysis strongly suggested that the yeast protein, YDL156W is an orthologue of LVI-1. It was therefore decided to test whether LVI-1 shares functional properties with this yeast protein as well as with its closest mammalian homologue, DDB2. Moreover, human *DDB2* mRNA is induced in a p53 dependent manner in response to DNA damage. The role of murine DDB2 is less well characterised. It is poorly expressed in the skin and lacks the functional p53 binding site found in the human gene (226). It has been suggested that this reflects species-specific differences in capacity for DNA repair in response to UV damage and that rodents may have a reduced requirement for DDB2 because they are nocturnal animals with skin protected by fur (222). It was therefore interesting to assess the effects of UV irradiation on expression of the human and murine *LVI-1* genes.

## 6.2 Material and methods

### 6.2.1 Phylogenetic analysis of LVI-1 amino acid sequences

LVI-1 amino acid sequences were obtained from the public domain databases: Ensembl database (<http://www.ensembl.org>) and UCSC human gene sorter at <http://www.genome.ucsc.edu> for *Homo sapiens*, *Pan troglodytes*, *Macaca mulatta*, *Bos taurus*, *Canis familiaris*, *Monodelphis domestica*, *Rattus norvegicus*, *Mus musculus*, *Fugu rubripes*, *Danio rerio*, *Xenopus tropicalis*, *Saccharomyces cerevisiae*, *Ciona intestinalis*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Oryza sativa*, and *Arabidopsis thaliana*. Phylogenetic analysis of these peptide data was conducted using MEGA 3.1 (Molecular Evolutionary Genetics Analysis version 3.1) software (<http://www.meagsoftware.net>) in which ClustalW was used for alignment of the amino acids sequences and the phylogenetic tree was generated using the NJ (Neighbor-Joining) method with 500 bootstrapped replications and treating gaps as missing data by complete deletion.

### 6.2.2 Cellular exposure to UV irradiation

#### *UV treatment of adherent cells*

$2 \times 10^6$  adherent cells (Section 2.2.1) were plated onto a 100 mm diameter dish with 10 ml of medium and were grown for 24 hours before treatment. Cells were rinsed with 10 ml of warmed PBS and were then exposed to  $10 \text{ J/m}^2$  UVC irradiation (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation). Following treatment, 10 ml of fresh medium was added to the plates and cells were harvested at different time points post irradiation.

#### *UV treatment of suspension cells*

$6 \times 10^6$  suspension cells (Section 2.2.1) were washed once with 1 ml of warmed PBS, pelleted by centrifugation at 1000 rpm for 5 minutes, prior to resuspending in 1 ml of warmed PBS and dispensing into the centre of a 100 mm diameter dish. Cells were exposed to  $10 \text{ J/m}^2$  UVC irradiation (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation), returned to T25  $\text{cm}^2$  culture flasks, and then harvested at different time points post irradiation.

### **6.2.3 Determination of the *LVI-1* transcripts levels after UV irradiation by RT-PCR**

Each of the total cellular RNA samples were prepared from p/m16i and Jurkat cells grown in tissue culture and UV treated as described in Section 6.2.2. Wt MEFs (mouse embryonic fibroblasts) cells, and RNA from p53-null MEFs cells were generous gifts from Dr Linda Hanlon (Faculty of Veterinary Medicine, University of Glasgow). The RNA was isolated and quantified as detailed in Section 2.2.3. Aliquots of 104 ng to 1 µg of total RNA were reverse transcribed as detailed in Section 2.2.4 and analysis of cDNAs were carried out using 2.5 µl of reverse transcription reaction with 50 pmol of the specific primers (see Table 6-1 and Section 2.2.4 for primer sequence, primer combination, and PCR conditions). 5 µl or 10 µl aliquots of each reaction were separated on 4% polyacrylamide gels, stained with ethidium bromide and visualized by UV illumination.

### **6.2.4 Cell cycle analysis by flow cytometry**

$5 \times 10^5$  H/3T3 cells (Section 2.2.1) were plated onto a 100 mm diameter dish with 10 ml of medium and were grown for 18 hours before treatment. Cells were rinsed with 10 ml of warmed PBS and were exposed to (5, 10, 50, and 100)  $J/m^2$  UVC irradiation (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation). Following treatment 10 ml of fresh medium was added to the plates and cells were harvested at different time points post irradiation for flow cytometry analysis. Briefly, cells were harvested by trypsinisation as described in Section 2.2.1 and then added to the medium. The cells were then removed by centrifugation at 1000 rpm for 10 minutes, resuspended in cold sample buffer (1g/L glucose in PBS), and repelleted (1000 rpm for 10 minutes). The wash in cold sample buffer was repeated twice more for a total of three washes. The cell pellet was vigorously vortexed in residual fluid (~0.2 ml) for 10 seconds before adding 1 ml of ice-cold 70% ethanol in a drop-wise manner. Cells were fixed for a minimum of 18 hours at 4 °C. For staining, cells were centrifuged at 3000 rpm for 5 minutes at 4 °C, and cell pellet gently vortexed in residual ethanol (~0.2 ml). 1 ml PI staining solution (Section 2.1.11) was added to each sample, and the mixture transferred to FACS tubes. Samples were incubated at

**Table 6-1. Primers used for analysis of LVI-1 expression in response to UV irradiation.**

**Table 6-1a. Primer sequences (by hand).**

Number of primer	Primer	Strand	Sequence (5'→3')
1	mFLJexon1aF	Forward	gct taa agg gct agg gtg ag
2	mFLJexon1cF	Forward	atg tct ggg tcc aag gct g
3	mFLJexon2R	Reverse	cag gcc tgg tgc act gtc
4	mHPRT5	Forward	ggg ggc tat aag ttc ttt gc
5	mHPRT3	Reverse	tcc aac act tcg aga ggt cc
6	huHPRT5	Forward	ggg ggc tat aaa ttc ttt gc
7	huHPRT3	Reverse	tcc aac act tcg aga ggt cc
8	humanFLJexon1cF	Forward	gaa gcc gaa aga tgt cca g
9	HumanFLJexon2RSP2	Reverse	gat tca tcg cat ctg cgg

**Table 6-1b. Primer combinations, PCR conditions, and applications.**

Primer combination	Product size	Annealing temperature (°C)	Application
1 + 3	513 bp	55	Analysis of P1 promoter
2 + 3	451 bp	58.8	Analysis of P2 promoter
4 + 5	294 bp	55	HPRT control
6 + 7	313 bp	55	HPRT control
8 + 9	411 bp	55	Analysis of P2 promoter

room temperature on a rocking platform for at least 30 minutes and analysed on a Beckman-Coulter Epics XL within 6 hours. Analysis was performed using Expo32 software package.

### **6.2.5 Immunofluorescent labelling and confocal microscopy**

Glass chamber slides (Nalge Nunc International) and microscope slides (BDH, VWR International) were routinely pretreated with poly-L-lysine (Sigma) at 37 °C overnight according to the manufacture's instructions. H/3T3 cell line was plated on two-well glass chamber slides at  $1 \times 10^4$  cells/well and also  $2 \times 10^4$  cells/well two days before fixation. In the case of suspension cells (Jurkat and p/m 16i cell lines), the microscope slides were mounted with the filter card (Shandon) and the cuvette (sample chamber) in the metal holder and  $5 \times 10^4$  cells in 250  $\mu$ l were loaded in the cuvette, cytocentrifuged onto microscope slides at 500 rpm for 5 minutes, air dried, and then the area around the cytocentrifuged cells marked using a wax pen (DAKO Pen S2002) before fixation. H/3T3 cells were washed once in ice-cold PBS and then fixed in 3.3% paraformaldehyde for 15 minutes at room temperature. Jurkat and p/m 16i cells were fixed in 3.3% paraformaldehyde for 15 minutes at room temperature. All further steps were performed on adherent and suspension cells in parallel. The cells were washed three times in ice-cold PBS, immersed in 20 mM glycine (Section 2.1.11) for 5 minutes at room temperature to remove free aldehyde groups to give a clear background, and then permeabilised with 0.1% Triton X-100/PBS (Section 2.1.11) three times for 10 minutes at room temperature. The slides were blocked in 0.1% Triton X-100/10%FBS/PBS (Section 2.1.11) for 1 hour at room temperature and then incubate overnight at 4 °C in a humid environment with the primary antibody (rabbit polyclonal antibody against mLvi-1-GST protein from terminal bleed of rabbit n.7328) diluted at 1:100 in block buffer (Section 2.1.11). The cells were then washed twice in 0.1% Triton X-100/PBS for 5 minutes and once in block buffer for 5 minutes at room temperature before incubation with 1:100 fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit (Jackson Immunoresearch Laboratories) in block buffer for 45 minutes at room temperature in a humid environment. After three washes in 0.1% Triton X-100/PBS for 5 minutes and one wash in ice-cold PBS at room temperature; the slides were mounted using VECTASHIELD® Hard + Set™

mounting medium with DAPI (4',6-diamidino-2-phenylindole) from Vector Laboratories and fluorescent images captured using a confocal microscope (Leica DMIRE 2).

### **6.2.6 Combined immunoprecipitation and western blot analysis**

MSH6 and the LVI-1 protein expression levels of H/3T3 and SW480 cells (Section 2.1.1) were examined by immunoprecipitation and western blot analysis. Typically, cells grown in tissue culture were washed twice with ice-cold PBS, drained thoroughly, lysed in 100  $\mu$ l working IP lysis buffer (Section 2.1.11) per  $10^6$  cells and rotated for 10 minutes at 4 °C. The crude extract was cleared in a bench top centrifuge at 14000 rpm for 30 minutes at 4 °C and the supernatant transferred to a fresh pre-cooled Eppendorf tube. Extracts were quantitated (Section 2.2.7) and 1 mg of each extract in a total volume of 500  $\mu$ l in IP lysis buffer was pre-cleared by incubation (1 hour at 4 °C with rocking) with 50  $\mu$ l of 50% protein G sepharose 4B slurry. The preparation of the 50% protein G sepharose 4B slurry was as follows: Protein G sepharose 4B Fast Flow (Sigma) beads were washed with ice-cold PBS and pelleted by centrifugation at 1000 rpm for 1 minute, for a total of three washes. The pellet was resuspended in equal volume of PBS, washed with IP lysis buffer and pelleted (14000 rpm for 1 minute) for a total of three washes. The pellet was then resuspended in equal volume of IP lysis buffer.

The pre-cleared lysate was separated from the beads by centrifugation at 14000 rpm for 1 minute and then rotated overnight at 4 °C with 1  $\mu$ g of anti-MSH6 monoclonal antibody (BD Biosciences, 610918). The beads were kept at - 70 °C for further examination by SDS-PAGE. The immune complexes were collected by incubation (1 hour at 4 °C with rocking) with 50  $\mu$ l of 50% protein G sepharose 4B slurry, pelleted (14000 rpm for 1 minute), washed five times in ice-cold IP lysis buffer, repelleted, resuspended in equal volume of 2x SDS Sample Loading Buffer (Section 2.1.11), and boiled for 10 minutes. The immunoprecipitates were loaded on 7.5% and 10% SDS-PAGE gels (Section 2.2.7). After electrophoresis, the proteins were transferred onto nitrocellulose membranes (Section 2.2.7) for western blot analysis (Section 2.2.7). The membranes were blotted with rabbit anti-MSH2 polyclonal antibody (Calbiochem, PC57) at dilution of 1:100 for 4 hours at room temperature, and with rabbit serum anti mLvi-1-GST at dilution 1:2000 for 4 hours at room temperature. As

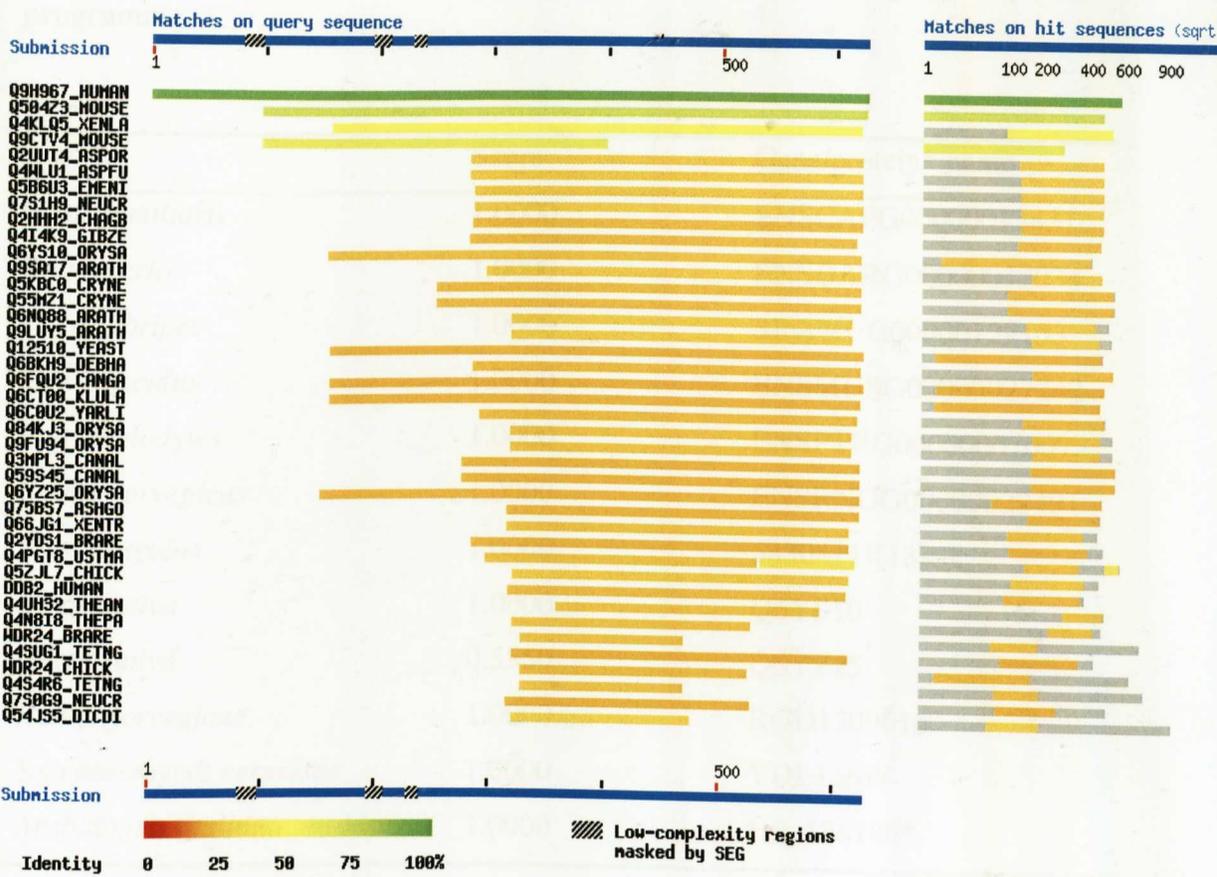
a secondary antibody, HRP conjugated swine anti-rabbit IgG polyclonal antibody (DakoCytomation) at 1:3000 dilution, for 1 hour at room temperature.

### 6.3 Homologues of LVI-1

The LVI-1 gene is a unique gene that is highly conserved in mammalian species, with closely related genes detected in human, chimpanzee, murine, rat, canine and bovine genomes. Other gene products related to LVI-1 were identified by BLAST searching of the Swissprot protein database (<http://www.expasy.ch/tools/blast/>). This analysis shows that many proteins from a wide range of eukaryotes show homology to LVI-1, but most of the relatedness is seen in the C-terminal WD repeat region (Figure 6-1) and multiple gene products from the same species appear in the list. This analysis illustrates the problems in identifying true orthologues of LVI-1 in more distantly related organisms.

An attempt to overcome this problem was made with the use of the Inparanoid programme (<http://www.inparanoid.cgb.ki.se>), which makes pairwise comparisons between whole genomes to assign the most likely orthologues. Other sequences are added to this group if they are closely related to one of the main orthologues and are called In-Paralogs. A confidence value is provided for each in-paralog that shows how closely related it is to the main orthologue. It is interesting that this programme fails to assign orthologues in *Drosophila* or *Caenorhabditis*, but confidently predicts the yeast gene YDL156W as an orthologue. Selected output from this analysis is shown in Table 6-2.

To show the evolutionary relationship in graphic form, a tree was generated using the orthologues detected by the Inparanoid and other gene products from the list provided by the Ensembl database, which automatically assigns putative orthologues by reciprocal BLAST analysis. The human DDB2 protein was added for comparison. The phylogenetic tree shown in Figure 6-2 was generated with MEGA 3.1 software using the Neighbour Joining method. This tree shows the relatively close homology between the mammalian proteins and the increasingly distant homology to lower species. However, the fact that DDB2 is more distantly related than the yeast YDL156W protein is consistent with the assignment of the latter as an orthologue.

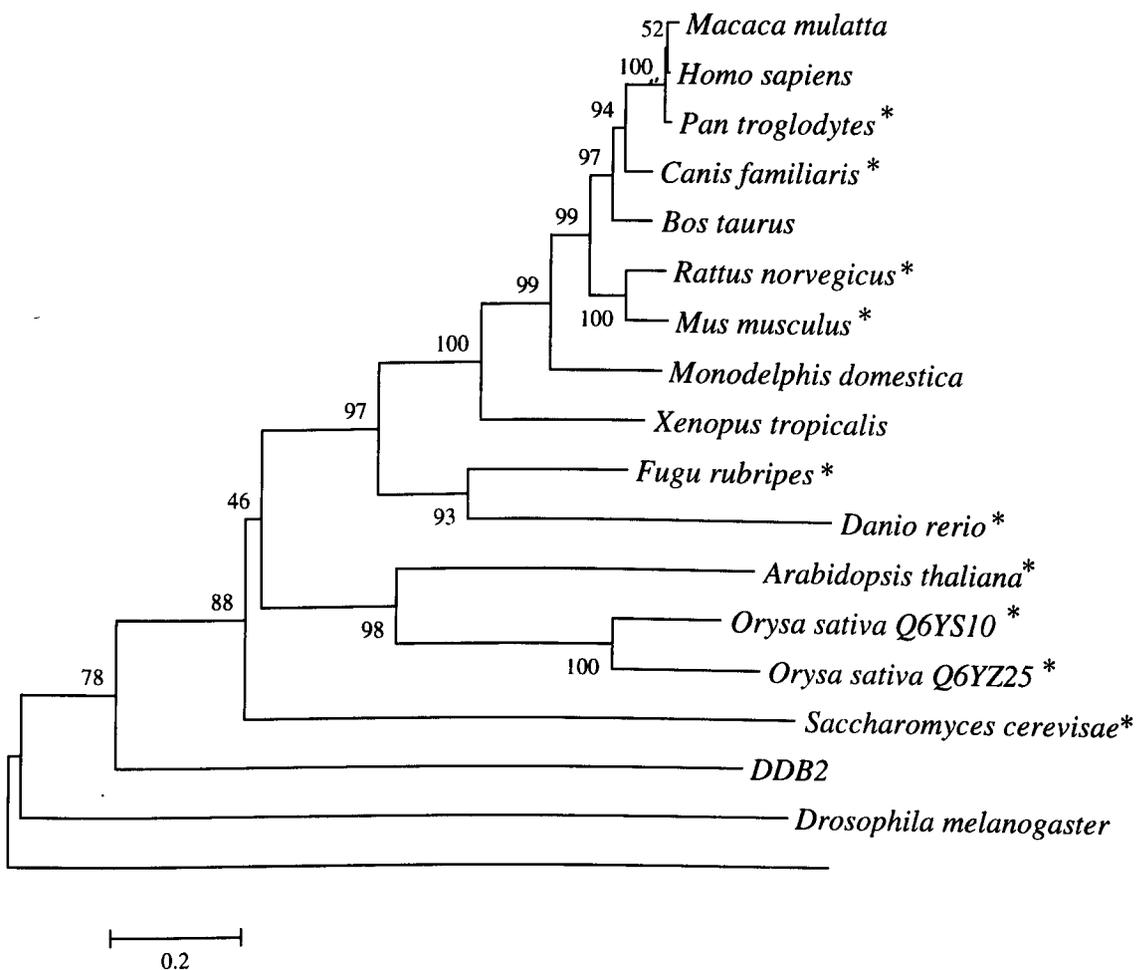


**Figure 6-1. Output from BLAST search of UniProt protein database (Swiss-Prot + TrEMBL) using the full length predicted human LVI-1 protein (626 aa). Only the first 40 hits are depicted. UniProt codes are shown in the left column. The coloured bars illustrated the homologous regions, coloured according to the degree of identity as shown in the key at the bottom left corner of the figure. Grey denotes regions of non-homology in the target sequences.**

**Table 6-2. Orthologues and in-paralogs of LVI-1 detected by the Inparanoid programme.**

Species	Score	Gene/protein* name
<i>Canis familiaris</i>	1.0000	ENSCAFG00000013431
<i>Danio rerio</i>	1.0000	ENSDARG00000012034
<i>Fugu rubripes</i>	1.0000	SINFRUG000000129153
<i>Mus musculus</i>	1.0000	ENSMUSG00000027242
<i>Pan troglodytes</i>	1.0000	ENSPTRG00000007007
<i>Rattus norvegicus</i>	1.0000	ENSRNOG00000022491
<i>Mus Musculus</i>	1.0000	5830411K18Rik
<i>Oryza sativa</i>	1.0000	Q6YS10
<i>Oryza sativa</i>	0.5550	Q6YZ25
<i>Rattus norvegicus</i>	1.0000	RGD1309616
<i>Saccharomyces cerevisiae</i>	1.0000	YDL156W
<i>Arabidopsis thaliana</i>	1.0000	NP_178186*

Note: In *Mus musculus*, ENSMUSG00000027242 and 5830411K18Rik (RIKEN cDNA 5830411K18 gene) represent the same gene from the ENSEMBL and NCBI databases respectively. In *Rattus norvegicus*, ENSRNOG00000022491 and RGD1309616 represent the same gene from the ENSEMBL and RGD databases respectively. In *Oryza sativa*, Q6YS10 and Q6YZ25 represent independent genes



**Figure 6-2. Phylogenetic tree of homologues of LVI-1.** A Neighbor-Joining tree was constructed using MEGA 3.1 software. Numbers (neighbor-joining bootstrap values) indicate the percentage that each branch was retained in 500 bootstrap samplings of the data. The asterisc (\*) indicates the LVI-1 orthologues identified by the Inparanoid programme (<http://www.inparanoid.cgb.ki.se>).

## 6.4 Functional analysis

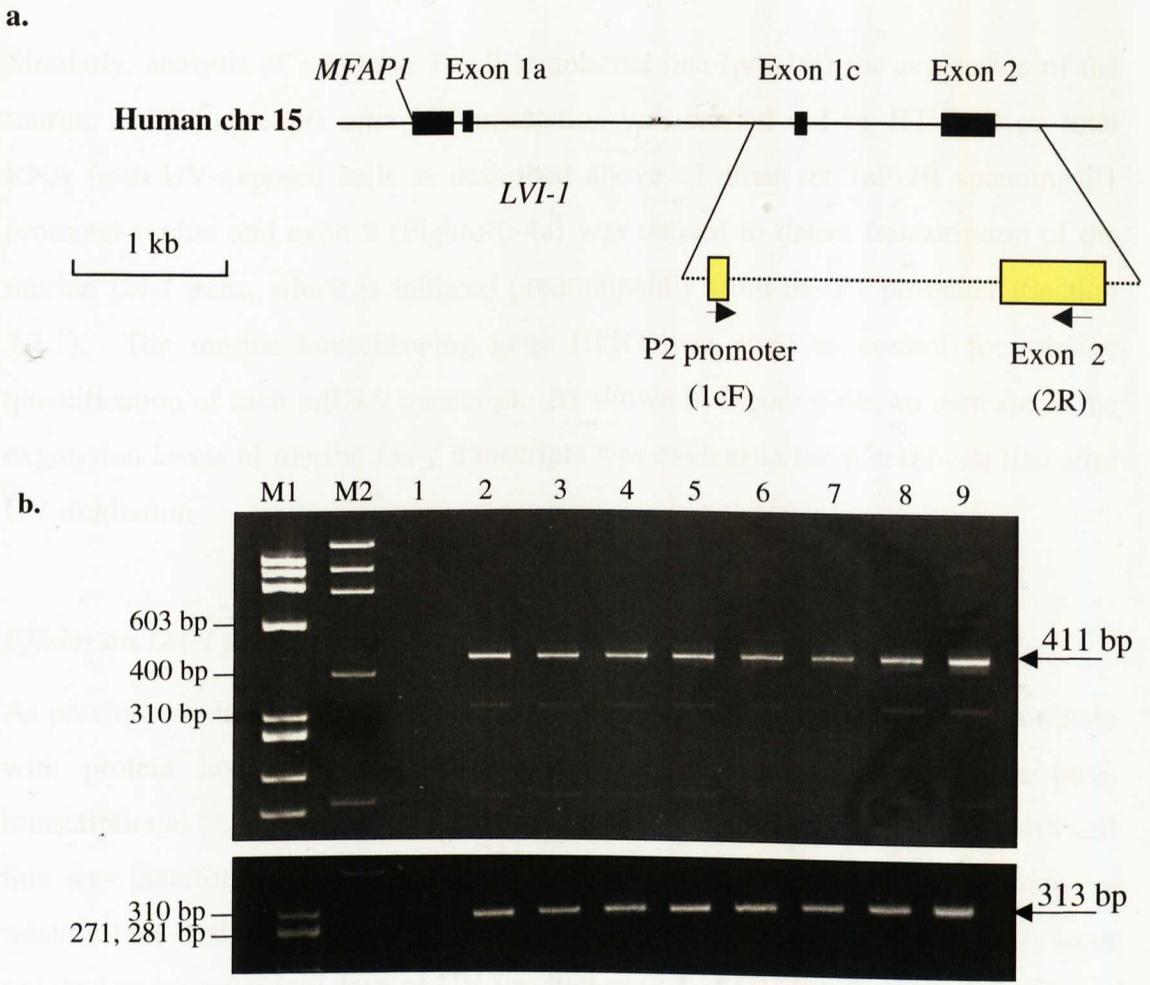
### 6.4.1 Responses of Lvi-1 to UV irradiation

The close structural relationship of LVI-1 to DDB2 indicates a possible role for *LVI-1* gene in DNA repair pathways, particularly in response to UV damage. It was therefore decided to test the effects of UV irradiation on the expression of LVI-1 mRNA and proteins. In human cells induction of DDB2 in response to UV damage is dependent on p53 (236) and the possibility of p53-dependent effects on Lvi-1 was also of interest. Also, the dose of irradiation is critical, as low doses generally induce cell cycle arrest and DNA repair, while high doses lead to apoptosis (281). Many established cell lines have mutations in p53 or a defective response, so it was decided to test primary fibroblasts, although these cells generally have lower expression of Lvi-1 than haematopoietic cells. Protein analysis was focused on the murine system, as I could not be confident that the antiserum to the human LVI-1 protein was specific. There were also practical problems in irradiating non-adherent cells that had to be irradiated in low volumes of buffer before replating.

In preliminary experiments, doses of UV irradiation were tested on H/3T3 murine fibroblasts (5, 10, 50 and 100 J/m<sup>2</sup>). The effects on cell viability were measured by propidium iodide staining and flow cytometry at various time intervals (0-48h) (not shown). On the basis of these results, a dose of 10 J/m<sup>2</sup> was chosen for further studies as a low dose that did not induce significant levels of apoptosis.

#### *Effects on Lvi-1 RNA*

Human and murine cell lines that had previously been shown to express the *LVI-1* gene (Figure 4-5) were used in this analysis. The human T-cell line Jurkat was exposed to 10 J of UV irradiation/m<sup>2</sup> and the cells analysed at different times afterward (0h, 2h, 4h, 8h, and 24h) by RT-PCR (Figure 6-3). Primer set 1cF-2R spanning P2 promoter region and exon 2 (Figure 6-3a) was chosen because transcription of the human *LVI-1* gene is mainly initiated from the P2 promoter (Section 3.3.1). The human housekeeping gene HPRT was used as control for relative quantification of each mRNA transcripts. During the time course after UV exposure, the levels of the human *LVI-1* transcripts remained unchanged, indicating that the human *LVI-1* is not induced in response to DNA damage by UV irradiation (Figure 6-3b).



**Figure 6-3. Determination of human *LVI-1* transcript levels after UV irradiation.**

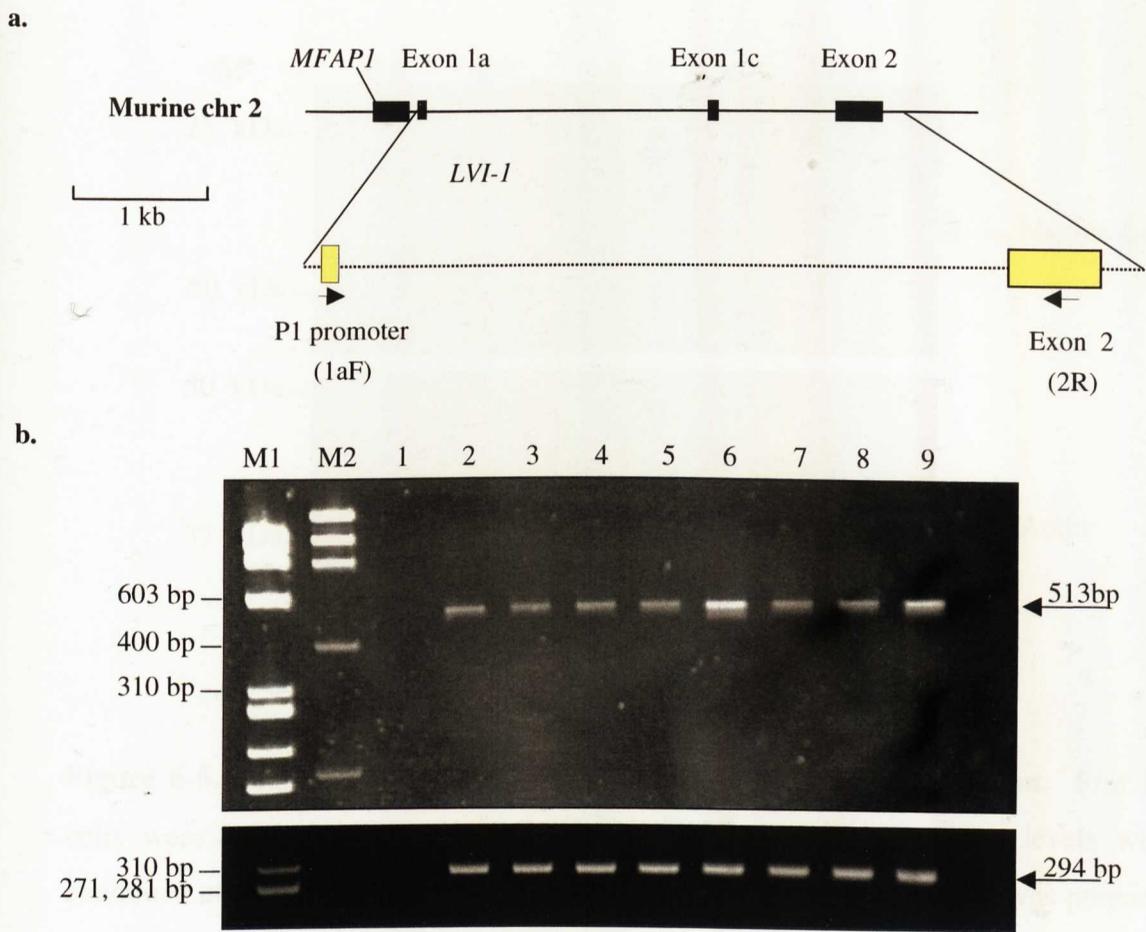
**a.** Structure of the 5' region of the human *LVI-1* gene. The coding exons are represented by black boxes. The yellow boxes represent the exons amplified by RT-PCR using the primer set 1cF-2R (see sequences in Table 6-1a). **b.** RT-PCR analysis of the human *LVI-1* expression after UV irradiation. Jurkat cells were treated with 10 J of UV irradiation/ m<sup>2</sup> and total RNA harvested after different time intervals. cDNA was prepared from RNA samples from Jurkat cells non-UV treated (0h, lane 2, and 24h, lane 9), and UV treated (0h, lane 3; 2h, lane 4; 4h, lane 5; 8h, lane 6; 12h, lane 7; and 24h, lane 8), as described in Section 2.2.4. PCR was carried out using the primer set 1cF-2R (411 bp) and the control housekeeping gene (HPRT, 313 bp) (see sequences in Table 6-1a). Lane 1 is a negative control (PCR primers only). RT-PCR products are shown following separation by 4% polyacrylamide gels. M1 and M2 are the markers  $\Phi$ X174RF DNA/*Hae* III fragments and Low DNA Mass Ladder, respectively.

Similarly, analysis of a murine T-cell lymphoma line (p/m16i) for expression of the murine *Lvi-1* transcripts after UV irradiation was carried out by RT-PCR on total RNA from UV-exposed cells as described above. Primer set 1aF-2R spanning P1 promoter region and exon 2 (Figure 6-4a) was chosen to detect transcription of the murine *Lvi-1* gene, which is initiated predominantly from the P1 promoter (Section 3.3.1). The murine housekeeping gene HPRT was used as control for relative quantification of each mRNA transcript. As shown in Figure 6-4b, no increase in the expression levels of murine *Lvi-1* transcripts was evident in the p/m16i cell line after UV irradiation.

### ***Effects on Lvi-1 protein levels***

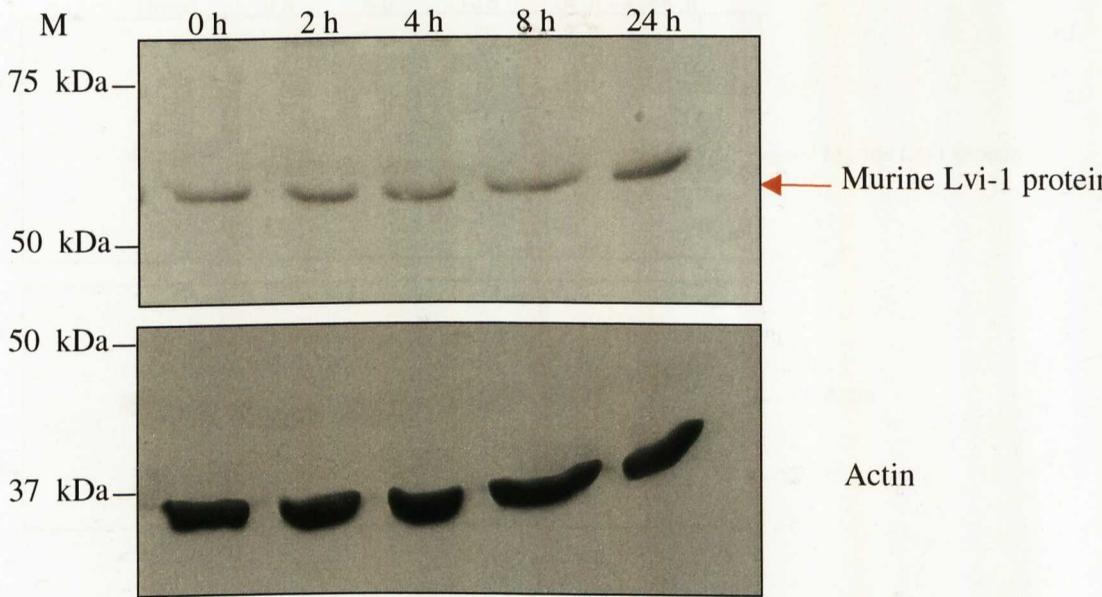
As previous studies showed that the *Lvi-1* mRNA levels did not necessarily correlate with protein levels (Section 4.3.2 and 5.3.9) it was conceivable that post-transcriptional regulation would be affected by UV-induced damage. The p/m16i cell line was therefore examined for *Lvi-1* protein expression after UV irradiation by western blot analysis with the 7328 rabbit serum anti-mLvi-1-GST. The cells were exposed to a similar low dose of UV irradiation (8 J of UV irradiation/m<sup>2</sup>) and lysed at the indicated times thereafter. Apart from a modest increase relative to the actin control observed at 24 hours, *Lvi-1* protein levels were not significantly affected during the time course after UV exposure (Figure 6-5), however, a 48 hour time point might have shown changes similar to those observed in H/3T3 cells as described below. A similar result was observed in another lymphoid cell line (p/m 47i) treated with the same dose of UV irradiation (data not shown).

A concern with regard to these experiments was that the murine lymphoma cell lines p/m16i and p/m47i lack functional p53 (246). In view of the p53-dependence of the human DDB2 gene, it was considered important to examine the expression of *Lvi-1* in UV irradiated p53 wt cells. For this purpose H/3T3 murine fibroblasts were treated with 10 J of UV irradiation/m<sup>2</sup>, lysed, and then analysed at different times afterward (0h, 8h, 12h, 24h, and 48h) by western blotting using the 7328 rabbit serum anti-mLvi-1-GST. Interestingly, consistent induction of LVI-1 protein was observed in UV irradiated H/3T3 cells (Figure 6-6), reaching maximum levels by 48 hours post-UV. To examine whether these results were due to increased transcript levels, RNA

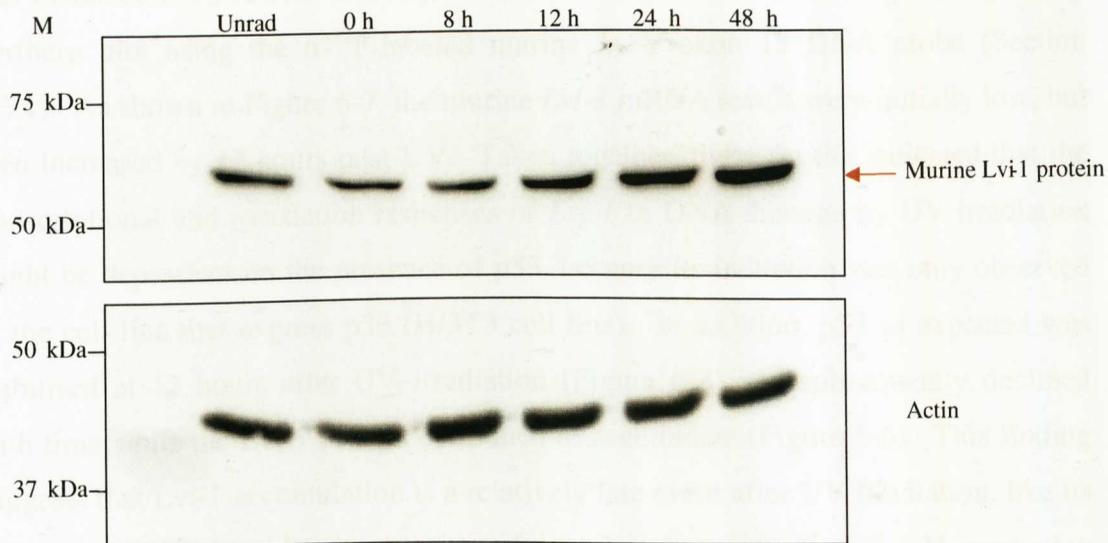


**Figure 6-4. Determination of murine *Lvi-1* transcript levels after UV irradiation.**

**a.** Structure of the 5' region of the murine *Lvi-1* gene. The coding exons are represented by black boxes. The yellow boxes represent the exons amplified in RT-PCR using the primer set 1aF-2R (see sequences in Table 6-1a). **b.** RT-PCR analysis of the murine *Lvi-1* expression after UV irradiation. P/m 16i cells were treated with 10 J of UV irradiation/ m<sup>2</sup> and total RNA harvested after different time intervals. cDNA was prepared from RNA samples from p/m 16i cells non-UV treated (0h, lane 2, and 24h, lane 9), and UV treated (0h, lane 3; 2h, lane 4; 4h, lane 5; 8h, lane 6; 12h, lane 7; and 24h, lane 8), as described in Section 2.2.4. PCR was carried out using the primer set 1aF-2R (513 bp) and the control housekeeping gene (HPRT, 294 bp) (see sequences in Table 6-1a). Lane 1 is a negative control (PCR primers only). RT-PCR products are shown following separation by 4% polyacrylamide gels. M1 and M2 are the markers  $\Phi$ X174RF DNA/*Hae* III fragments and Low DNA Mass Ladder, respectively.



**Figure 6-5. Lvi-1 expression in p/m 16i cell line after UV irradiation.** P/m 16i cells were irradiated with UV (8 J/m<sup>2</sup>). Lvi-1 protein expression levels were examined at various times after UV treatment as shown. Western blot was prepared from whole cell lysates (70 µg) and probed with anti-Lvi-1-GST serum (rabbit 7328) at 1:2000 dilution. The band corresponding to the murine Lvi-1 protein is highlighted with a red arrow. The blot was reprobbed with an antiserum specific for β-actin as a loading control. The blot exposure time for Lvi-1 and actin expression was 30 seconds and 15 seconds, respectively. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

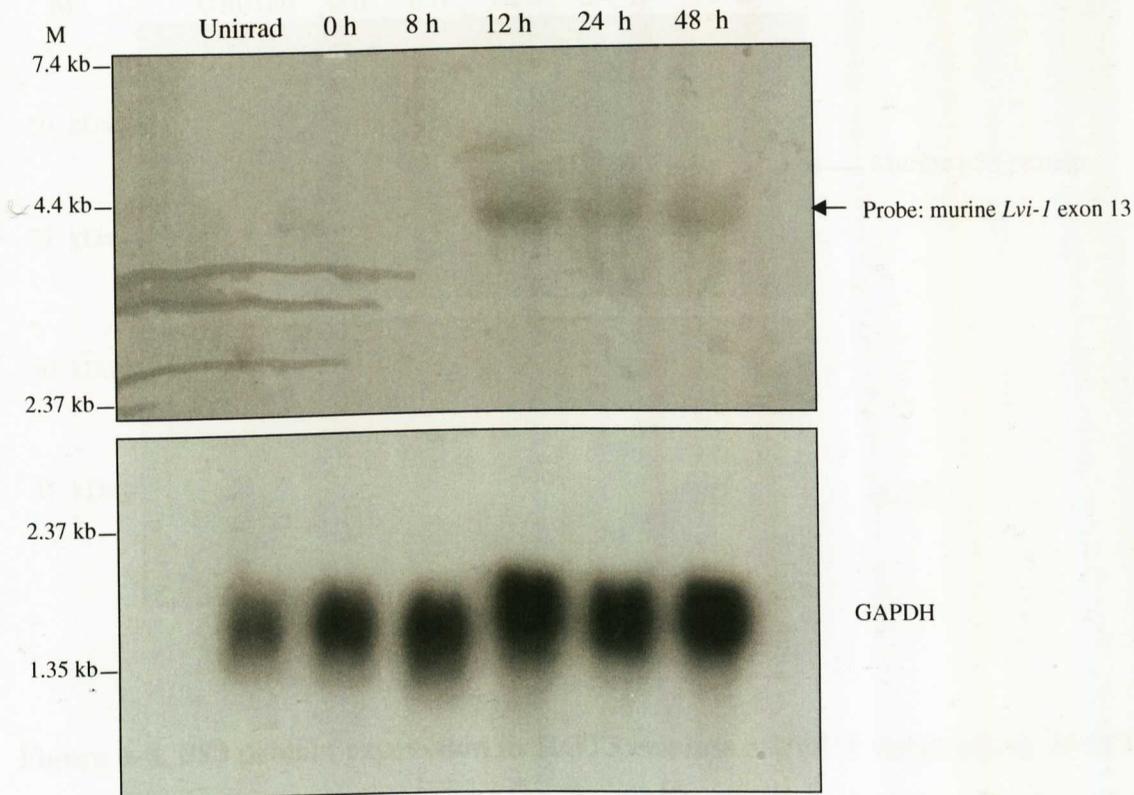


**Figure 6-6. Lvi-1 expression in H/3T3 cell line after UV irradiation.** H/3T3 cells were irradiated with UV (10 J/m<sup>2</sup>). Lvi-1 protein expression levels were examined at various times after UV treatment as shown. Western blot was prepared from whole cell lysates (100 µg) and probed with anti-Lvi-1-GST serum (rabbit 7328) at 1:2000 dilution. The band corresponding to the murine Lvi-1 protein is highlighted with a red arrow. The blot was reprobed with an antiserum specific for β-actin as a loading control. The blot exposure time for Lvi-1 and actin expression was 30 seconds and 1 minute, respectively. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

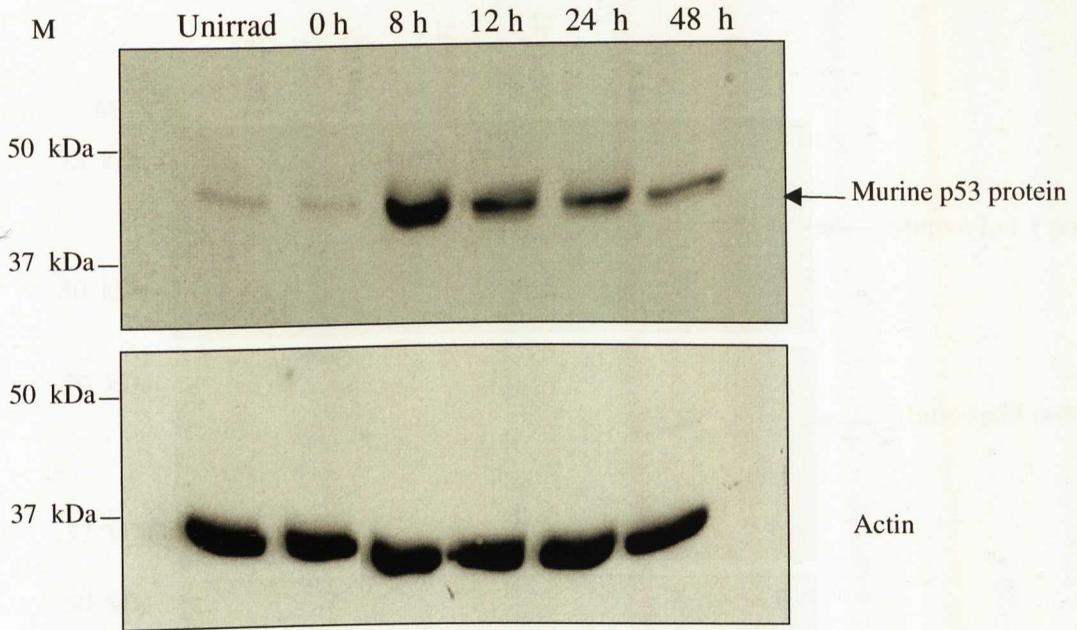
was extracted from H/3T3 cells exposed to 10 J of UV irradiation/m<sup>2</sup> and analysed by northern blot using the  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe (Section 3.3.1). As shown in Figure 6-7, the murine *Lvi-1* mRNA levels were initially low, but then increased by 12 hours post-UV. Taken together, these results indicated that the transcriptional and translation responses of *Lvi-1* to DNA damage by UV irradiation might be dependent on the presence of p53, because its induction was only observed in the cell line that express p53 (H/3T3 cell line). In addition, p53 as expected was stabilised at 12 hours after UV irradiation (Figure 6-8) and subsequently declined with time while the Lvi-1 protein continued to accumulate (Figure 6-6). This finding suggests that Lvi-1 accumulation is a relatively late event after UV irradiation, like its homologue DDB2, which is upregulated as a late event 48 h after UV irradiation (282, 283). However, basal levels of murine Lvi-1 protein appeared to be independent of p53 status in a panel of cell lines (Figure 6-9). In addition, analysis of *Lvi-1* transcription using primers derived from the P1 and P2 promoter regions revealed its presence in primary murine fibroblasts (MEFs) with wt or inactive p53, showing again that the basal levels of *Lvi-1* transcripts from the P1 or from P2 promoter regions are independent of p53 status (Figure 6-10).

### ***Effects on LVI-1 protein localisation***

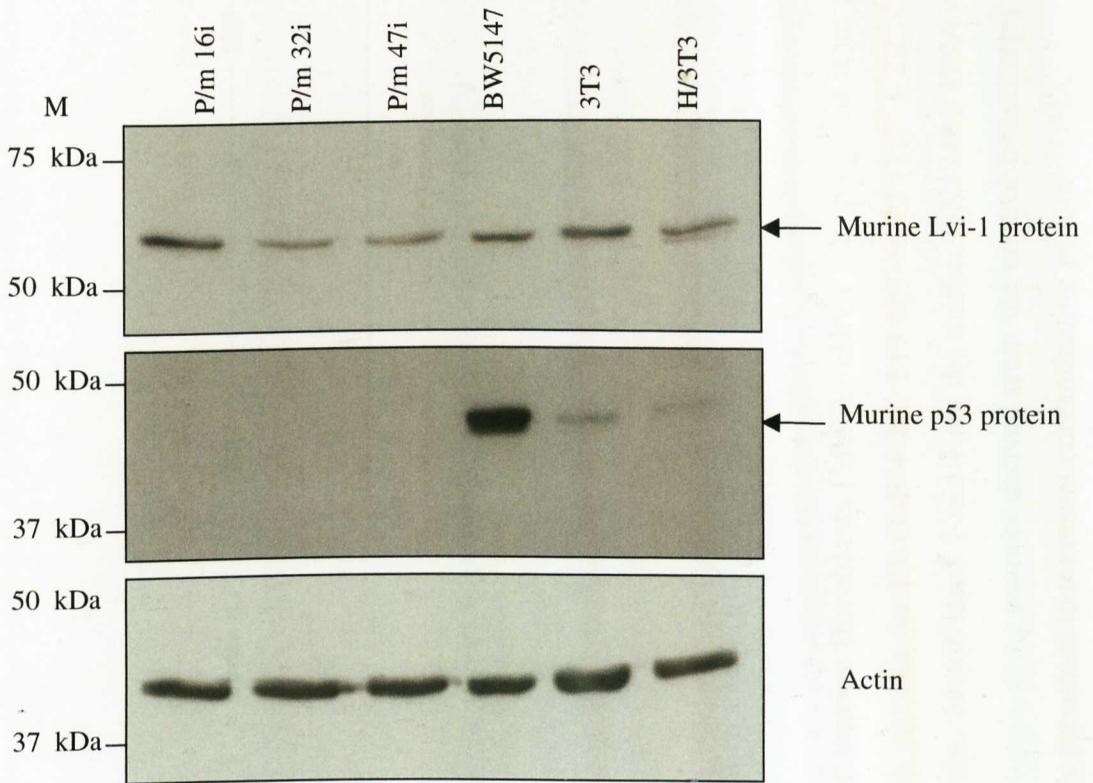
In addition to effects on protein levels, the subcellular location of DDB is regulated in response to DNA damage. Thus, while in undamaged wild type cells DDB2 localizes to the nucleus and DDB1 is found in both cytoplasm and the nucleus (222) in UV damaged cells DDB2 plays a role in transporting DDB1 to the nucleus. To examine whether Lvi-1 might behave in a similar fashion, expression of the human and murine *LVI-1* gene products was analysed in intact cells by confocal immunofluorescence microscopy. An adherent cell line (H/3T3) and the suspension cell lines (p/m 16i and Jurkat) were exposed to 10 J of UV irradiation/m<sup>2</sup>, fixed with paraformaldehyde 48 hours and 24 hours after UV treatment, respectively; and then probed with the prepurified rabbit serum anti-mLvi-1 protein followed by incubation with a FITC-labelled secondary antibody. A similar cytoplasmic staining pattern was detected for all the cell lines before and after UV treatment (Figure 6-11), indicating that Lvi-1 protein is predominantly located in the cytoplasm and that UV treatment did not



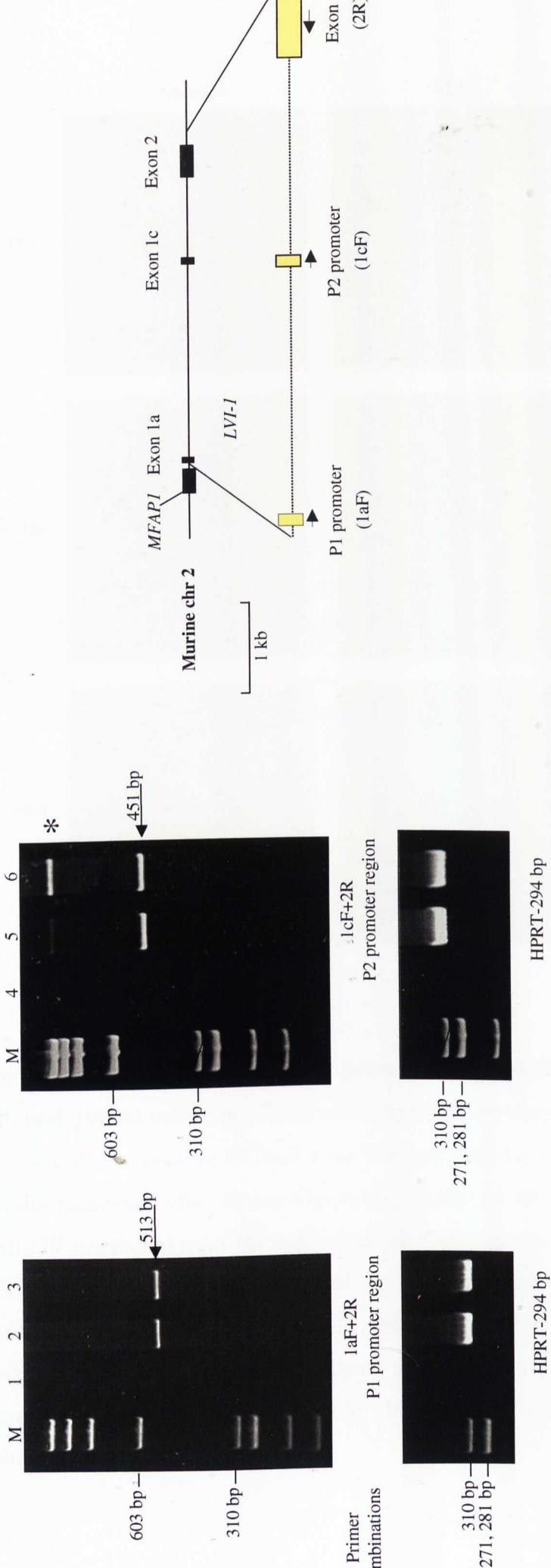
**Figure 6-7. Detection of *Lvi-1* transcript after UV irradiation of H/3T3 cells.** H/3T3 cells were irradiated with UV (10 J/m<sup>2</sup>) and *Lvi-1* transcript levels were examined at various times after UV treatment as shown. A northern blot was prepared from total RNA (20  $\mu$ g samples), probed with  $\alpha$ -<sup>32</sup>P-labeled murine *Lvi-1* exon 13 DNA probe (Section 4.3.1) and the products detected by autoradiography. The band corresponding to the murine *Lvi-1* transcript (4.5 kb) is indicated with a black arrow. The blot was reprobed with GAPDH (1.4 kb) probe to control for loading of mRNA. Unirrad, unirradiated. The relative migration of RNA size markers (0.24-9.5 kb RNA ladder, Invitrogen Life Technologies) is indicated. The blot exposure time for murine *Lvi-1* transcripts and control (GAPDH) was 7 days and 9 hours, respectively.



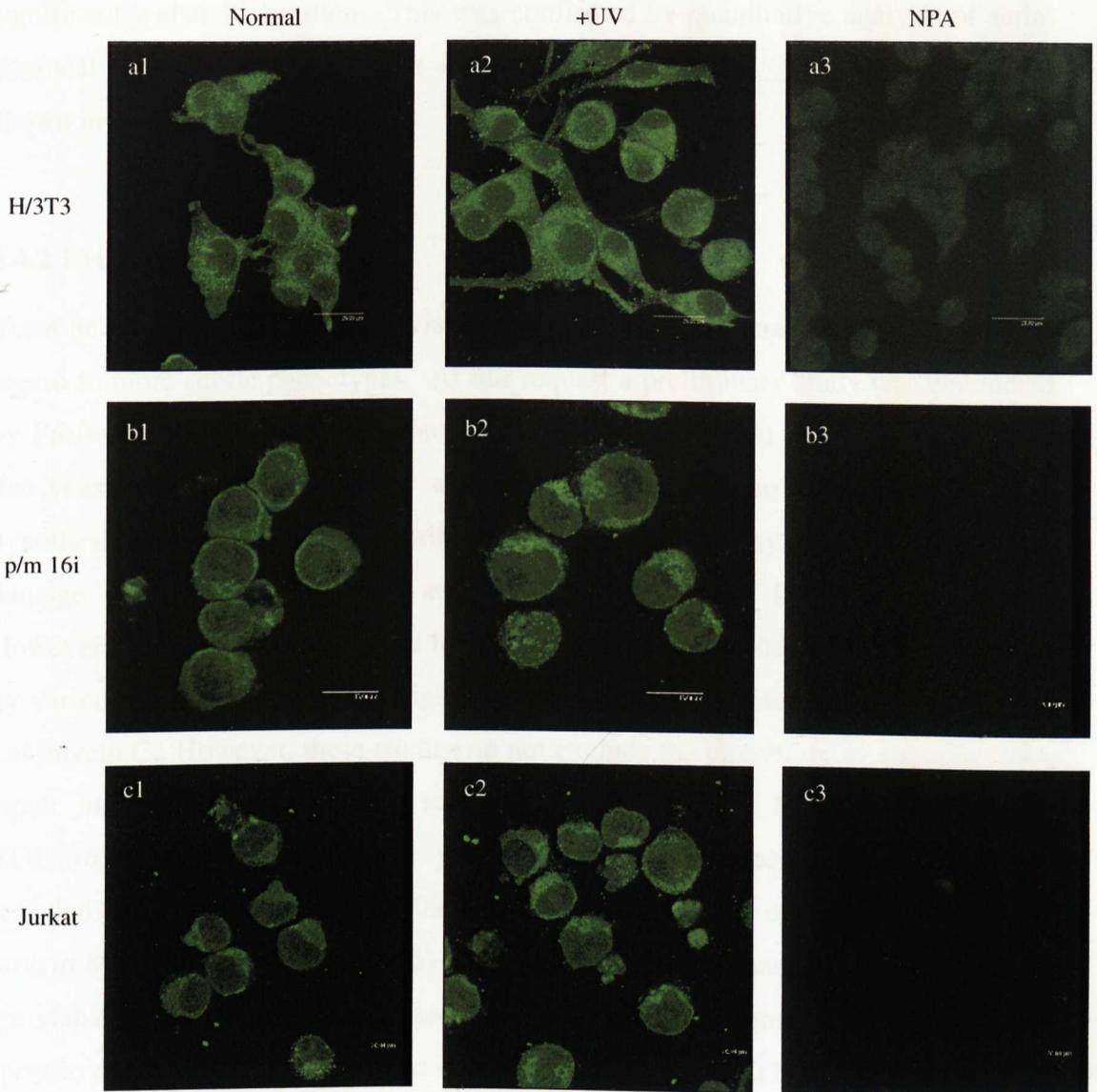
**Figure 6-8. P53 protein expression in H/3T3 cell line after UV irradiation.** H/3T3 cells were irradiated with UV (10 J/m<sup>2</sup>). p53 protein expression levels were examined at various times after UV treatment as shown. Western blot was prepared from whole cell lysates (100 µg) and probed with mouse anti-p53 (1C12) monoclonal antibody, at 1:2000 dilution. The band corresponding to the murine p53 protein is highlighted with a black arrow. The blot was reprobed with an antiserum specific for β-actin as a loading control. The blot exposure time for p53 and actin expression was 10 seconds and 2 seconds, respectively. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.



**Figure 6-9. Analysis of p53 expression in various murine cell lines by western blot.** A western blot was prepared from whole cell lysates (100  $\mu$ g) from the cell lines as shown and probed with anti-Lvi-1-GST serum (rabbit 7328) at 1:2000 dilution and with mouse anti-p53 (1C12) monoclonal antibody at 1:2000 dilution. The band corresponding to murine Lvi-1 and p53 proteins are indicated with a black arrow. The blot was reprobbed with an antiserum specific for  $\beta$ -actin as a loading control. The blot exposure time for Lvi-1, p53 and actin expression was 1 minute, 15 minutes and 5 seconds, respectively. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.



**Figure 6-10. *Lvi-1* promoter use is independent of p53 status in primary murine fibroblasts (Mefs).** cDNA was prepared from p53 wt MEFs (lanes 2, 5) or p53 null MEFs (lanes 3, 6) and amplified with primers specific for P1 (1aF + 2R, 513 bp) or P2 (1cF + 2R, 451 bp) transcripts as shown using total RNA from p53 wt and from p53 null Mefs cells. M and lanes 1 and 4 are the marker  $\Phi$ X174RF DNA and the negative controls (PCR primers only), respectively. RT-PCR of the same samples using primers derived from the control housekeeping gene (HPRT, 294 bp) are shown below. \* The upper PCR product in lanes 5 and 6 is presumed to represent contaminating genomic DNA.



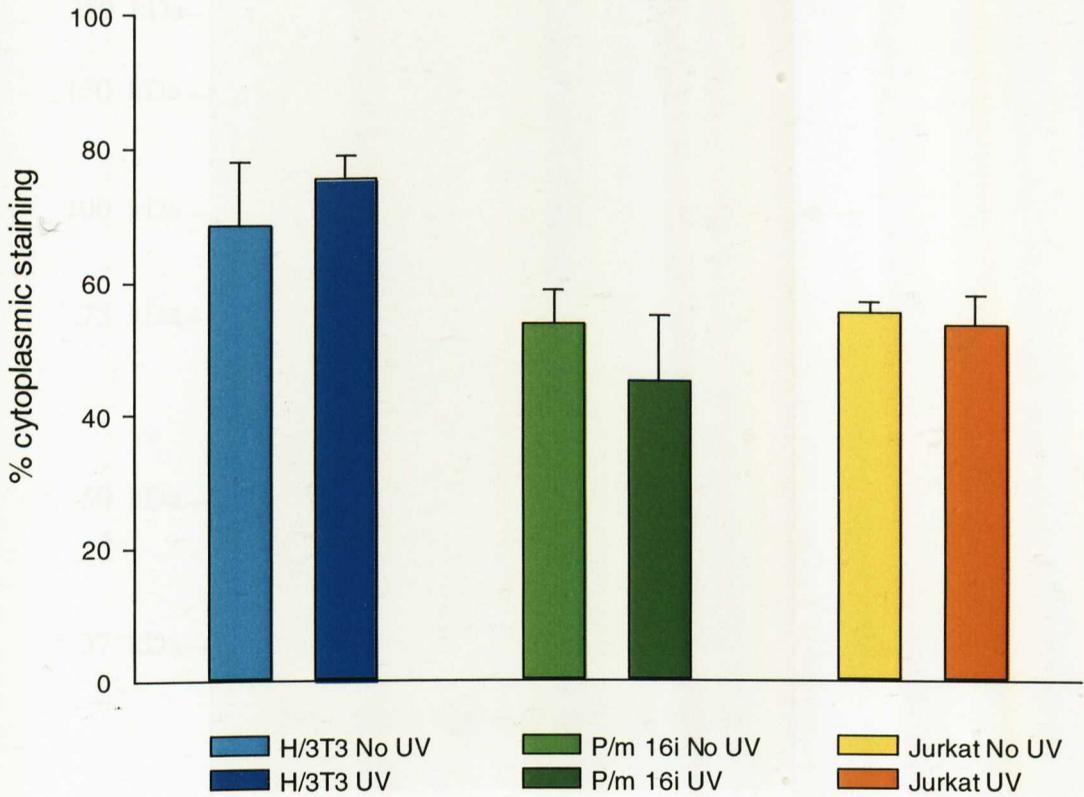
**Figure 6-11. Localisation of LVI-1 protein by immunofluorescence in H/3T3, p/m 16i, and Jurkat cell lines.** Cells which had been untreated (normal), or treated with UV (+ UV) at a dose of 10 J/m<sup>2</sup> were fixed on poly-L-lysine coated glass slides with paraformaldehyde after 48 hours post-UV irradiation (H/3T3 cell line) and 24 hours post-UV irradiation (p/m 16i and Jurkat cell lines) and then stained with anti-mLvi-1-GST serum (rabbit 7328) followed by incubation with FITC-labeled secondary antibody (a1, a2, b1, b2, c1, and c2). NPA (no primary antibody) was included as control, showing no fluorescent patterns when imaged under the same conditions as experimental samples (a3, b3, and c3). Bars, 20  $\mu$ m (H/3T3 and Jurkat) and 16  $\mu$ m (p/m 16i).

significantly alter its location. This was confirmed by quantitative analysis of serial confocal sections using the Image J software package (<http://rsb.info.nih.gov/ij/>) as shown in Figure 6-12.

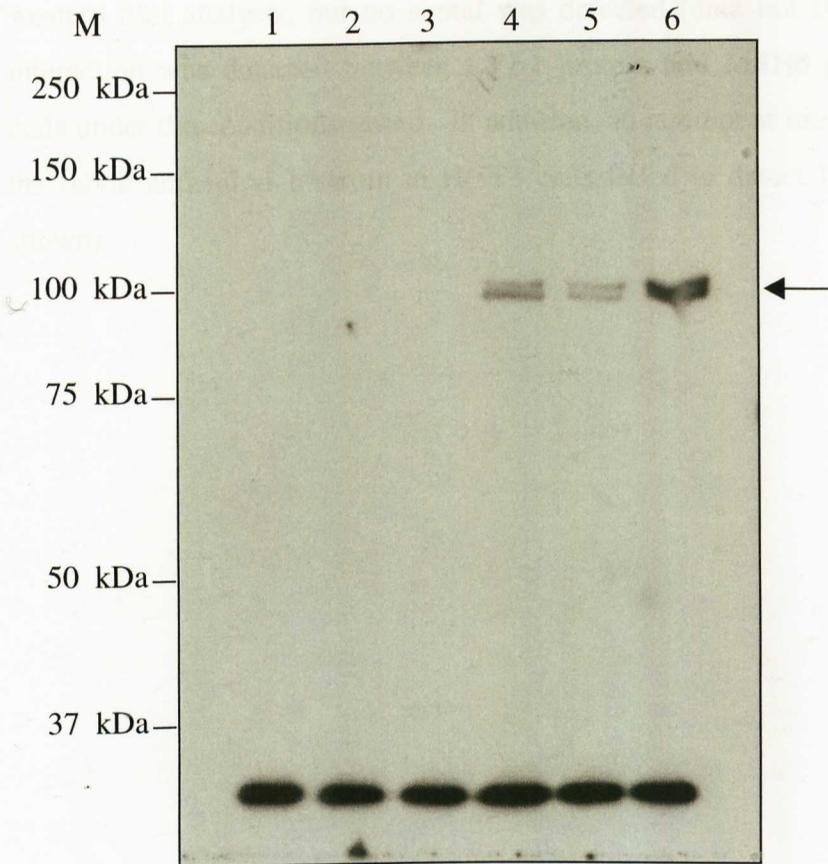
#### 6.4.2 Lvi-1 protein interactions

Yeast deficient in YDL156W are viable but have not been characterised in depth with regard to more subtle phenotypes. At our request a preliminary study was conducted by Professor Ian Hickson's laboratory (University of Oxford) to test the possibility that yeast deficient in YDL156W would be more sensitive to DNA damage. This hypothesis was based on the possibility that the protein may play a role in DNA damage detection or repair by analogy with the human DDB2 gene product. However, a series of studies failed to reveal any unusual sensitivity to DNA damage by various agents including UV light, hydroxyurea, methyl methane sulphonate and mitomycin C. However, these studies do not exclude the possibility of a role in DNA repair as not all functions are tested by these agents. Moreover, studies of YDL156W-interacting proteins by yeast 2-hybrid and co-precipitation analyses has revealed an interaction with YDR097C, the yeast orthologue of the mismatch repair protein MSH6 (MutS homologue 6) (284). It is interesting that *Msh6* knockout mice are viable but are highly predisposed to B- and T-cell lymphomas (285) and have a specific deficit in repair of somatic hypermutation in B-cells (286). It was therefore decided to test the possibility of a similar interaction between the *LVI-1* gene product and MSH6 in a mammalian system.

To address this issue, a combined immunoprecipitation/western blot analysis was carried out in H/3T3 cells and also in SW480 cells that were used as a positive control that is known to express MSH2. In case complex formation might be dependent on induction by DNA damage, H/3T3 cells were exposed to 10 J of UV irradiation/m<sup>2</sup>, and total cell extract was harvested by 48 hours after treatment. MSH2 in complex with MSH6 was detected by immunoprecipitation with MSH6-specific antibody followed by western blotting with the MSH2-specific antibody. This analysis demonstrated the interaction between MSH6 and MSH2 not only in the positive control, SW480 cells (Figure 6-13, lane 6) as expected, but also in the H/3T3 cells. No significant effect was seen in H/3T3 cells following UV treatment (Figure 6-13, lanes 4 and 5). A similar blot was then probed with rabbit anti-mLvi-1 serum in the



**Figure 6-12. Percentage of cytoplasmic staining of the murine (H/3T3 and p/m 16i) and human (Jurkat) cell lines before and after UV treatment.** Serial confocal sections of the murine and human cell lines were quantitated using the Image J software package with multiple colour analysis macro (<http://rsb.info.nih.gov/ij/>).



**Figure 6-13. Interaction of MSH6 with MSH2 is detected by Co-immunoprecipitation.** 1 mg aliquots of total protein from H/3T3 cells, which had been untreated (lanes 1 and 4) or treated with UV dose at 10 J/m<sup>2</sup> (harvested 48 hours post-UV irradiation, lanes 2 and 5), and unirradiated SW480 cells (lanes 3 and 6) were immunoprecipitated with anti-MSH6 antibody (Section 6.2.6). The immunoprecipitated proteins (lanes 4, 5 and 6) and the protein non-specifically bound to the protein G sepharose beads (lanes 1, 2, and 3) were resolved by a 7.5% SDS-PAGE and transferred to a nitrocellulose membrane. The membrane was then immunoblotted with anti-MSH2 antibody. The band corresponding to MSH2 protein (100 kDa) is highlighted with a black arrow. The blot exposure time was 10 minutes. M is the protein molecular weight markers (Precision Plus Protein Standards, BioRad) shown in kilodaltons.

western blot analysis, but no signal was detected (data not shown). Therefore, no interaction was detected between LVI-1 protein and MSH6 protein in mammalian cells under the conditions tested. In addition, an attempt at immunoprecipitation with the rabbit anti-mLvi-1 serum in H/3T3 cells failed to detect Lvi-1 protein (data not shown).

## 6.5 Discussion

Evolutionary analysis shows that *LVI-1* is a unique and highly conserved gene. However, due to the widespread occurrence of WD repeat in functionally diverse proteins, it is difficult to identify orthologues in distantly related eukaryotes. The Inparanoid programme was employed to try to circumvent this problem and this analysis strongly suggested that yeast *YDL156W* is a true orthologue. Some preliminary functional analyses were therefore attempted based on this observation. Also, despite the large number of WD repeat proteins in the mammalian, the organisation and sequence of *LVI-1* is closest to *DDB2*, a gene whose product is known to play a role in responses to UV irradiation. It therefore seemed logical to test the effect of UV irradiation on *Lvi-1* expression and protein localisation.

With hindsight, the design of the UV irradiation experiments could have been improved by inclusion of more controls and testing of a wider range of UV doses. Moreover, better quantitation and sensitivity might have been achieved by real-time PCR, but there was insufficient time available to generate and optimise such assays. The cell lines available were in many cases tumour cells that have mutations in p53 and other pathways that control responses to irradiation and other stresses, which may explain the rather heterogeneous results obtained. Also, the human and murine *LVI-1* genes are driven by distinct promoters (P1, P2) that may well be regulated differently. Furthermore, I was unable to analyse effects on human *LVI-1* protein expression with confidence due to concerns over the specificity of the antiserum raised to this protein. Some evidence of UV responses were seen on murine *Lvi-1* mRNA and protein levels. *Lvi-1* mRNA and protein accumulated following low-dose UV irradiation of H/3T3 fibroblasts. Responses were less pronounced in lymphoid tumour cell lines that lack functional p53. Additional analysis of *Lvi-1* transcription in wt MEFs and p53-null MEFs revealed that the basal levels of *Lvi-1* transcripts from the P1 or from P2 promoter regions are independent of p53 status. However, the basal *Lvi-1* protein levels in the wt MEFs and p53-null MEFs remains to be established.

Experiments on yeast deficient in *YDL156W* have so far failed to show any effect on responses to a number of DNA damaging agents, but these experiments are not exhaustive or conclusive (Prof. I. Hickson, personal communication). However, the interaction between this protein and the yeast orthologue of *MSH6* is suggestive of a role in DNA repair. My preliminary analysis showed no evidence of such a complex

in H/3T3 cells, but studies on cells with more abundant protein expression and under different conditions would be required before drawing clear conclusions.

Initial studies of Lvi-1 protein localisation by confocal microscopy suggested a predominantly cytoplasmic location. This finding is rather surprising from computer-based predictions of a nuclear localisation (Chapter III). However, it was noted that this approach was poor at predicting proteins that shuttle between cellular compartments. It is also interesting that the yeast YDL156W is detected in both nuclear and cytoplasmic fractions (287).

## Chapter VII

### General Discussion

The investigations carried out during this project were principally concerned with the characterisation of *LVI-1* gene and its patterns of expression. The *LVI-1* gene was discovered as a candidate tumour suppressor gene in a case of feline lymphoma, where it appeared to be inactivated by a combination of retroviral insertion and transcriptional down-regulation of the unoccupied allele (58). It was decided at the outset of this project that attempts to define the function of the *LVI-1* gene would be focused initially on murine and human genes. This decision was taken due to the strong conservation of the gene in mammals and the superior genomic and proteomic resources available for human and laboratory mouse systems.

Analysis of the pattern of transcription of the *LVI-1* in murine and human provided clear evidence of differential expression between species. The difference in major transcript size between human and murine (2.8 kb vs 4.5 kb) was found to be due mainly to the differing length of the 3'UTR. Non-coding regions often contain important signals that determine transcript stability or affect nuclear export and subsequent translation (288), suggesting that this difference might be significant.

However, an even more remarkable difference was observed at the 5' end of the mammalian *LVI-1* genes. In the murine, the gene appears to be driven mainly by a bidirectional promoter (P1), which is shared with the adjacent *Mfap1* gene. However, the human gene is expressed from a second promoter (P2) located further downstream at a prominent CpG island, which is only partially conserved in the murine. While the human gene appears to be expressed exclusively from the P2 promoter, the murine gene appears to be capable of utilising P2, particularly in tissues where expression is relatively low. These findings were not only based on the evidence of the available human and murine cDNAs and ESTs, but also from promoter prediction analyses, which scored the forward P1 promoter only for the murine gene, whereas the P2 promoter adjacent to exon 1c is more strongly predicted in the human genome. It is possible that the use of the distal P1 promoter by the murine gene suppresses transcription here by the phenomenon of promoter interference (270).

These results suggest that there may be significant differences in regulation of the human and murine *LVI-1* genes at the transcriptional level, although northern blot analyses revealed that both human and murine *LVI-1* expression are highest in

lymphoid-derived tissues, especially thymus. Preliminary experiments on UV irradiation of cells provided some evidence of differential regulation, but as the results of these studies were inconsistent and in some cases difficult to interpret, it would be useful to extend these analyses using a wider range of UV dose levels. Also, the potential role of p53 in coordinating these responses is worthy of further investigation. The possibility that the gene responds to other non-genotoxic stress signals is also worth considering, particularly in view of the conserved heat shock transcription factor binding site in the P1 promoter region.

The observed differences at the 5' ends of the human and mouse genes also have major implications for the proteins encoded by the respective genes. The major human mRNA has the capacity to encode a larger 70 kDa protein (MSRSG), while the most common murine mRNA species encodes a 58 kDa protein (MKALS). These predictions are not clear-cut, however, as the use of alternative P2 promoter of the murine gene would lead to an mRNA with the capacity to encode the larger 69 kDa protein (MSGSK), while the human mRNA could conceivably be translated from an internal AUG corresponding to the murine initiator, leading to expression of a 59 kDa protein (MKNTS). It is interesting that computer-based predictions based on a consensus of initiation sites favoured the second AUG in the human mRNA suggesting an intermediate product of 68 kDa (MKVNE). These results highlight the uncertainty of computer-based predictions and the importance of testing these by direct experimental analysis.

To investigate the expression of the *LVI-1* genes at the protein level, I attempted to raise antisera to the human and murine gene products using the GST fusion system where C-terminal fragments were expressed in bacterial cells. Some difficulties with protein solubility were overcome and these fusion proteins were used to immunise rabbits and derive polyclonal antisera. The antiserum to the mouse protein appeared to be the better reagent and it was possible to demonstrate its specific reactivity in western blot analyses. The specificity of the antiserum to the human protein could not be demonstrated clearly, however. This difference may have been due to the fact that the human LVI-1-GST fusion protein did not contain the authentic C-terminus and consisted of a tripartite fusion with GST, LVI-1 and additional read-through sequences from the plasmid vector. Based on these observations, it would be helpful to attempt immunisation with a GST fusion expressing the authentic C-terminus of

human LVI-1 protein. Also, it would be useful to use these fusion proteins to raise monoclonal antibodies where specificity could be assessed with greater confidence.

Use of the murine antiserum in western blotting identified a protein of the expected size based on translation of the major mRNA species (58 kDa). The use of the antiserum in immunofluorescence and confocal microscopy suggested that this protein is predominantly cytoplasmic. This was surprising in view of the prediction of a nuclear location on the basis of sequence content (Chapter III). One possibility is that the antiserum is cross-reactive with some cytoplasmic antigen(s), and that the observed pattern does not genuinely reflect the distribution of the protein in the cell. Again, the availability of monoclonal antibodies of more defined specificity would be useful to resolve this question. Also, it would be possible to address the question of subcellular localisation by an independent technique such as cell fractionation followed by western blotting. Another possibility is that the Lvi-1 protein shuttles between the cytoplasm and the nucleus, and this possibility could be addressed by testing the effects of inhibitors of nuclear export such as leptomycin B (289).

WD-40 repeats proteins comprise a large family that encompass many different functions (Chapter I). This complexity makes prediction of the function of a novel family member such as LVI-1 difficult. The leads that were followed in this thesis were based mainly on analogy with the most closely related mammalian homologue, *DDB2*, and its role in repair of UV-induced damage. Another clue that was followed was based on the presumptive yeast orthologue, *YDL156W*. Preliminary analysis of *YDL156W* deficient yeast cells for sensitivity to DNA damaging agents did not reveal any unusual sensitivity (Prof Ian Hickson, personal communication). However, the *YDL156W* protein has been shown to interact with *YDR097C*, the yeast orthologue of the mismatch repair protein *MSH6* (284). Interestingly, mice deficient in *MSH2* or *MSH6* seem destined to develop hematologic malignancy, predominantly B-cell lymphoblastic lymphomas, even in the absence of a thymus and a greatly reduced T-cell component (290).

My studies failed to show evidence that LVI-1 protein interacts with *MSH6*, but it would be interesting to test this hypothetical interaction under a wider range of conditions, particularly genotoxic stresses, where DNA repair complex formation might be induced. Also, as a further clue to function, it would be interesting to use the Lvi-1 protein as target in yeast two-hybrid screens for interacting proteins (291) in

which the full length *Lvi-1* gene product will be used as bait to detect interacting proteins from various expression libraries. "

As reviewed in Chapter I, the LVI-1 homologue DDB2 is an essential component of a complex that binds damaged DNA and suppresses UV-induced mutagenesis (237). Moreover, rodents and humans regulate *DDB2* transcription differently. Thus, p53 binds and activates transcription of the human *DDB2* gene after DNA damage, while murine *DDB2* gene does not contain a functional p53 response element (226). This suggests that p53 protein regulates *DDB2* only in human cells. Therefore, based on these findings it was decided to determine the relationship between p53 induction and *LVI-1* expression.

My data showed that the basal expression level of the murine *Lvi-1* gene and protein in primary fibroblasts and T-cell lymphoma lines is independent of their p53 status. However, protein levels appeared to be induced in the H3T3 fibroblast cell line, but not in the p53-deficient lymphoma line p/m 16i. Unfortunately, the lack of a validated antiserum for the human LVI-1 protein meant that it was not possible to test for post-transcriptional changes in levels of the human LVI-1 protein.

In conclusion, while the *LVI-1* gene shows evidence of responses to UV irradiation at both transcriptional and post-transcriptional levels, there was insufficient time available to test these responses systematically to determine the effects of radiation dose and p53 status as well as possible differences in the responses of the human and murine promoters. As radiation induces distinct gene sets at low and high doses, corresponding to DNA repair and apoptotic pathways respectively (292), it would be very interesting to discover whether *LVI-1* falls into one or other of these gene sets. Moreover, it will be interesting to study the effects of inhibition of *LVI-1* by ablating gene expression *in vitro* using small interfering RNA molecules (siRNA) (293).

As discussed in Chapter I, the *LVI-1* gene was discovered as a target for inactivation in a case of pre-B cell lymphoma. It will be intriguing to examine other B-cell lymphomas from FIV infected cats for evidence of similar genetic events. Furthermore, high expression of *Lvi-1* in lymphoid cells suggests that its loss of expression may predispose to lymphoid neoplasia and or other forms of cancer in other species and it will be equally important to investigate this possibility. Genetic evidence for loci involved in tumour predisposition or somatic tumour suppression at human chromosome 15q15 has been incompletely explored (294), although one study ruled out *LVI-1* point mutations in a series of colorectal cancers (295).

The most direct approach to the putative tumour suppressor function of *Lvi-1* would be to attempt to knock out the function of *Lvi-1* in the mouse and assess the effects of this deletion on phenotype and tumour susceptibility. Recent analysis of public domain resources identified mouse embryonic stem (ES) cell lines carrying gene trap vector insertions within the murine *Lvi-1* gene. If it is established that these insertions result in complete functional inactivation of *Lvi-1*, this resource may aid the generation of *Lvi-1*<sup>-/-</sup> mouse lines after blastocyst injection. If mice homozygous for the inactive gene are non-viable, the causal pathology may be examined, which may be informative with respect to gene function. The pattern of *Lvi-1* expression could also be examined in homozygous and heterozygous animals making use of the  $\beta$ -geo (fusion of  $\beta$ -galactosidase and neomycin transferase) gene expressed by the gene trap vector. The intrinsic effects of *Lvi-1* knockout or heterozygosity could be examined by following groups of mice and assessing general health, background tumour rates and the integrity of the haematopoietic system. Another useful approach would be to cross the knockout mice with strains expressing dominant oncogenes (*Myc*, *Runx*, *Pim*, or *Myb*) and to challenge mice or cell lines derived from these with mutagenic agents to reveal more subtle roles in responses to genotoxic damage (296). Unfortunately, these approaches to *Lvi-1* function were beyond the scope of this study. However, it is hoped that the work described here will provide the groundwork for future characterisation of this highly conserved and potentially important eukaryotic gene.

## References

1. **Ellerman, V. and O. Bang.** 1908. Experimentell leukämie bei Hühnern. *Zentralbl Bakteriol Parasitenkd Infectionskr Hyg Abt Orig.* **46**:595-609.
2. **Rous, P.** 1911. A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *J.Exp.Med.* **13**:397-411.
3. **Poeschla E.M., Jr. Buchschacher GL, and F. Wong-Staal.** 2000. Etiology of cancer: RNA viruses, p. 149-158. *In* Jr. DeVita VT, S. Hellman, and S. A. Rosenberg (eds.), *Cancer: Principles and Practice of Oncology*. Philadelphia: Lippincott, Williams & Wilkins.
4. **Fearon, E. R.** 2000. Oncogenes and tumor suppressor genes, p. 77-118. *In* M. D. Abeloff, J. O. Armitage, A. S. Lichter, and J. E. Niederhuber (eds.), *Clinical Oncology*. Philadelphia: Churchill Livingstone.
5. **Nermut, M. V. and D. J. Hockley.** 1996. Comparative morphology and structural classification of retroviruses. *Curr.Top.Microbiol.Immunol.* **214**:1-24.
6. **Coffin J.M.** 1996. Retroviridae: The viruses and their replication, p. 763-843. *In* B. N. Fields, D. M. Knipe, P. M. Howley, and e. et.al. (eds.), *Fundamental Virology*. Lippincott-Raven Publishers, Philadelphia.
7. **Flint, S. J., L. W. Enquist, and R. M. Krug.** 2000. Reverse transcription and integration: Hallmarks of the retroid viruses, p. 199-234. *In* *Principles of Virology-Molecular Biology, Pathogenesis, and Control*. ASM Press, Washington DC.
8. **Panganiban, A. T. and D. Fiore.** 1988. Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* **241**:1064-1069.
9. **Hu, W. S. and H. M. Temin.** 1990. Retroviral recombination and reverse transcription. *Science* **250**:1227-1233.

10. **Buchschacher, G. L., Jr.** 2001. Introduction to retroviruses and retroviral vectors. *Somat.Cell Mol.Genet.* **26**:1-11. "
11. **Vogt, V. M.** 1997. Retroviral Virions and Genomes, p. 27-70. *In* Coffin J.M., S. H. Hughes, and H. E. Varmus (eds.), *Retroviruses*. Cold Spring Harbor Laboratory Press, New York.
12. **Burmeister, T.** 2001. Oncogenic retroviruses in animals and humans. *Rev.Med.Virol.* **11**:369-380.
13. **Baldo, A. M. and M. A. McClure.** 1999. Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. *J.Virol.* **73**:7710-7721.
14. **Miller, R. J., J. S. Cairns, S. Bridges, and N. Sarver.** 2000. Human immunodeficiency virus and AIDS: insights from animal lentiviruses. *J.Virol.* **74**:7187-7195.
15. **Varmus, H. and R. Swanstrom.** 1982. Replication of retroviruses. *In* RNA Tumor Viruses, New York: Cold Spring Harbour.
16. **Roe, T., T. C. Reynolds, G. Yu, and P. O. Brown.** 1993. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* **12**:2099-2108.
17. **Lewis, P. F. and M. Emerman.** 1994. Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J.Virol.* **68**:510-516.
18. **Mooslehner, K., U. Karls, and K. Harbers.** 1990. Retroviral integration sites in transgenic Mov mice frequently map in the vicinity of transcribed DNA regions. *J.Virol.* **64**:3056-3058.
19. **Scherdin, U., K. Rhodes, and M. Breindl.** 1990. Transcriptionally active genome regions are preferred targets for retrovirus integration. *J.Virol.* **64**:907-912.
20. **Robinson, H. L. and G. C. Gagnon.** 1986. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. *J.Virol.* **57**:28-36.

21. **Vijaya, S., D. L. Steffen, and H. L. Robinson.** 1986. Acceptor sites for retroviral integrations map near DNase-I-hypersensitive sites in chromatin. *J.Virol.* **60**:683-692.
22. **Goodenow, M. M. and W. S. Hayward.** 1987. 5' long terminal repeats of myc-associated proviruses appear structurally intact but are functionally impaired in tumors induced by avian leukosis viruses. *J.Virol.* **61**:2489-2498.
23. **Shih, C. C., J. P. Stoye, and J. M. Coffin.** 1988. Highly preferred targets for retrovirus integration. *Cell* **53**:531-537.
24. **Jonkers, J. and A. Berns.** 1996. Retroviral insertional mutagenesis as a strategy to identify cancer genes. *Biochim.Biophys.Acta* **1287**:29-57.
25. **Vogt, P. K.** 1997. Historical introduction to the general properties of retroviruses, p. 1-25. *In* Coffin J.M., S. H. Hughes, and H. E. Varmus (eds.), *Retroviruses*. Cold Spring Harbor Laboratory Press, New York.
26. **Best, S., P. R. Le Tissier, and J. P. Stoye.** 1997. Endogenous retroviruses and the evolution of resistance to retroviral infection. *Trends Microbiol.* **5**:313-318.
27. **Bock, M. and J. P. Stoye.** 2000. Endogenous retroviruses and the human germline. *Curr.Opin.Genet.Dev.* **10**:651-655.
28. **Boeke, J. D. and J. P. Stoye.** 1997. Retrotransposons, endogenous retroviruses and the evolution of retroelements, p. 343-436. *In* Coffin J.M., S. H. Hughes, and H. E. Varmus (eds.), *Retroviruses*. Cold Spring Harbor Laboratory Press, New York.
29. **Kolson, D. L. and F. Gonzalez-Scarano.** 2001. Endogenous retroviruses and multiple sclerosis. *Ann.Neurol.* **50**:429-430.
30. **Lower, R.** 1999. The pathogenic potential of endogenous retroviruses: facts and fantasies. *Trends Microbiol.* **7**:350-356.

31. **Hughes, J. F. and J. M. Coffin.** 2001. Evidence for genomic rearrangements mediated by human endogenous retroviruses during primate evolution. *Nat.Genet.* **29**:487-489.
32. **Ting, C. N., M. P. Rosenberg, C. M. Snow, L. C. Samuelson, and M. H. Meisler.** 1992. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev.* **6**:1457-1465.
33. **Gifford, R. and M. Tristem.** 2003. The evolution, distribution and diversity of endogenous retroviruses. *Virus Genes* **26**:291-315.
34. **Blaise, S., N. de Parseval, and T. Heidmann.** 2005. Functional characterization of two newly identified Human Endogenous Retrovirus coding envelope genes. *Retrovirology.* **2**:19.
35. **Hayward, W. S., B. G. Neel, and S. M. Astrin.** 1981. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukemia. *Nature* **290**:475-480.
36. **Mikkers, H. and A. Berns.** 2003. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv.Cancer Res.* **88**:53-99.
37. **Voronova, A. F., H. T. Adler, and B. M. Sefton.** 1987. Two lck transcripts containing different 5' untranslated regions are present in T cells. *Mol.Cell Biol.* **7**:4407-4413.
38. **Mikkers, H., J. Allen, P. Knipscheer, L. Romeijn, A. Hart, E. Vink, and A. Berns.** 2002. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nat.Genet.* **32**:153-159.
39. **Fung, Y. K., W. G. Lewis, L. B. Crittenden, and H. J. Kung.** 1983. Activation of the cellular oncogene c-erbB by LTR insertion: molecular basis for induction of erythroblastosis by avian leukemia virus. *Cell* **33**:357-368.
40. **Nilsen, T. W., P. A. Maroney, R. G. Goodwin, F. M. Rottman, L. B. Crittenden, M. A. Raines, and H. J. Kung.** 1985. c-erbB activation in ALV-

induced erythroblastosis: novel RNA processing and promoter insertion result in expression of an amino-truncated EGF<sup>r</sup> receptor. *Cell* **41**:719-726.

41. **Shen-Ong, G. L., M. Potter, J. F. Mushinski, S. Lavu, and E. P. Reddy.** 1984. Activation of the c-myc locus by viral insertional mutagenesis in plasmacytoid lymphosarcomas. *Science* **226**:1077-1080.
42. **Shen-Ong, G. L., H. C. Morse, III, M. Potter, and J. F. Mushinski.** 1986. Two modes of c-myc activation in virus-induced mouse myeloid tumors. *Mol.Cell Biol.* **6**:380-392.
43. **Payne, G. S., S. A. Courtneidge, L. B. Crittenden, A. M. Fadly, J. M. Bishop, and H. E. Varmus.** 1981. Analysis of avian leukosis virus DNA and RNA in bursal tumours: viral gene expression is not required for maintenance of the tumor state. *Cell* **23**:311-322.
44. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. *Cell* **23**:323-334.
45. **Westaway, D., G. Payne, and H. E. Varmus.** 1984. Proviral deletions and oncogene base-substitutions in insertionally mutagenized c-myc alleles may contribute to the progression of avian bursal tumors. *Proc.Natl.Acad.Sci.U.S.A.* **81**:843-847.
46. **Payne, G. S., J. M. Bishop, and H. E. Varmus.** 1982. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. *Nature* **295**:209-214.
47. **Nusse, R. and H. E. Varmus.** 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **31**:99-109.
48. **Peters, G., S. Brookes, R. Smith, and C. Dickson.** 1983. Tumorigenesis by mouse mammary tumor virus: evidence for a common region for provirus integration in mammary tumors. *Cell* **33**:369-377.

49. **Cuypers, H. T., G. Selten, W. Quint, M. Zijlstra, E. R. Maandag, W. Boelens, P. van Wezenbeek, C. Melief, and A. Berns.** 1984. Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. *Cell* **37**:141-150.
50. **Selten, G., H. T. Cuypers, W. Boelens, E. Robanus-Maandag, J. Verbeek, J. Domen, C. Van Beveren, and A. Berns.** 1986. The primary structure of the putative oncogene pim-1 shows extensive homology with protein kinases. *Cell* **46**:603-611.
51. **Domen, J., M. Von Lindern, A. Hermans, M. Breuer, G. Grosveld, and A. Berns.** 1987. Comparison of the human and mouse PIM-1 cDNAs: nucleotide sequence and immunological identification of the in vitro synthesized PIM-1 protein. *Oncogene Res.* **1**:103-112.
52. **van der Lugt, N. M., J. Domen, E. Verhoeven, K. Linders, G. H. van der, J. Allen, and A. Berns.** 1995. Proviral tagging in E mu-myc transgenic mice lacking the Pim-1 proto-oncogene leads to compensatory activation of Pim-2. *EMBO J.* **14**:2536-2544.
53. **van Lohuizen, M., M. Breuer, and A. Berns.** 1989. N-myc is frequently activated by proviral insertion in MuLV-induced T cell lymphomas. *EMBO J.* **8**:133-136.
54. **Uren, A. G., J. Kool, A. Berns, and M. van Lohuizen.** 2005. Retroviral insertional mutagenesis: past, present and future. *Oncogene* **24**:7656-7672.
55. **Fung, Y. K., L. B. Crittenden, and H. J. Kung.** 1982. Orientation and position of avian leukosis virus DNA relative to the cellular oncogene c-myc in B-lymphoma tumors of highly susceptible 1515 X 7(2) chickens. *J.Virol.* **44**:742-746.
56. **Gisselbrecht, S., S. Fichelson, B. Sola, D. Bordereaux, A. Hampe, C. Andre, F. Galibert, and P. Tambourin.** 1987. Frequent c-fms activation by proviral insertion in mouse myeloblastic leukaemias. *Nature* **329**:259-261.

57. **Mukhopadhyaya, R. and L. Wolff.** 1992. New sites of proviral integration associated with murine promonocytic leukemias and evidence for alternate modes of c-myc activation. *J.Virol.* **66**:6035-6044.
58. **Beatty, J., A. Terry, J. MacDonald, E. Gault, S. Cevario, S. J. O'Brien, E. Cameron, and J. C. Neil.** 2002. Feline immunodeficiency virus integration in B-cell lymphoma identifies a candidate tumor suppressor gene on human chromosome 15q15. *Cancer Res.* **62**:7175-7180.
59. **Marchetti, A., F. Buttitta, S. Miyazaki, D. Gallahan, G. H. Smith, and R. Callahan.** 1995. Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. *J.Virol.* **69**:1932-1938.
60. **Goff, S. P.** 1987. Gene isolation by retroviral tagging. *Methods Enzymol.* **152**:469-481.
61. **Neil, J. C. and E. R. Cameron.** 2002. Retroviral insertion sites and cancer: fountain of all knowledge? *Cancer Cell* **2**:253-255.
62. **Devon, R. S., D. J. Porteous, and A. J. Brookes.** 1995. Splinkerettes--improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res.* **23**:1644-1645.
63. **Li, J., H. Shen, K. L. Himmel, A. J. Dupuy, D. A. Largaespada, T. Nakamura, J. D. Shaughnessy, Jr., N. A. Jenkins, and N. G. Copeland.** 1999. Leukaemia disease genes: large-scale cloning and pathway predictions. *Nat.Genet.* **23**:348-353.
64. **Suzuki, T., H. Shen, K. Akagi, H. C. Morse, J. D. Malley, D. Q. Naiman, N. A. Jenkins, and N. G. Copeland.** 2002. New genes involved in cancer identified by retroviral tagging. *Nat.Genet.* **32**:166-174.
65. **Alberti, A., C. Murgia, S. L. Liu, M. Mura, C. Cousens, M. Sharp, A. D. Miller, and M. Palmarini.** 2002. Envelope-induced cell transformation by ovine betaretroviruses. *J.Virol.* **76**:5387-5394.

66. **Sodroski, J. G., C. A. Rosen, and W. A. Haseltine.** 1984. Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. *Science* **225**:381-385.
67. **Sodroski, J., R. Patarca, C. Rosen, F. Wong-Staal, and W. Haseltine.** 1985. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. *Science* **229**:74-77.
68. **Weiss, R. A.** 2001. Retroviruses and cancer. *Current Science* **81**:528-534.
69. **Mowat, M., A. Cheng, N. Kimura, A. Bernstein, and S. Benchimol.** 1985. Rearrangements of the cellular p53 gene in erythroleukaemic cells transformed by Friend virus. *Nature* **314**:633-636.
70. **Ben David, Y., V. R. Prideaux, V. Chow, S. Benchimol, and A. Bernstein.** 1988. Inactivation of the p53 oncogene by internal deletion or retroviral integration in erythroleukemic cell lines induced by Friend leukemia virus. *Oncogene* **3**:179-185.
71. **Ben David, Y., A. Lavigueur, G. Y. Cheong, and A. Bernstein.** 1990. Insertional inactivation of the p53 gene during friend leukemia: a new strategy for identifying tumor suppressor genes. *New Biol.* **2**:1015-1023.
72. **Hicks, G. G. and M. Mowat.** 1988. Integration of Friend murine leukemia virus into both alleles of the p53 oncogene in an erythroleukemic cell line. *J.Virol.* **62**:4752-4755.
73. **Ben David, Y. and A. Bernstein.** 1991. Friend virus-induced erythroleukemia and the multistage nature of cancer. *Cell* **66**:831-834.
74. **Lane, D. P. and S. Benchimol.** 1990. p53: oncogene or anti-oncogene? *Genes Dev.* **4**:1-8.
75. **Lu, S. J., S. Rowan, M. R. Bani, and Y. Ben David.** 1994. Retroviral integration within the Fli-2 locus results in inactivation of the erythroid transcription factor NF-E2 in Friend erythroleukemias: evidence that NF-E2 is essential for globin expression. *Proc.Natl.Acad.Sci.U.S.A.* **91**:8398-8402.

76. **Orkin, S. H.** 1990. Globin gene regulation and switching: circa 1990. *Cell* **63**:665-672.
77. **Andrews, N. C., H. Erdjument-Bromage, M. B. Davidson, P. Tempst, and S. H. Orkin.** 1993. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* **362**:722-728.
78. **Buchberg, A. M., H. G. Bedigian, N. A. Jenkins, and N. G. Copeland.** 1990. Evi-2, a common integration site involved in murine myeloid leukemogenesis. *Mol.Cell Biol.* **10**:4658-4666.
79. **Cho, B. C., J. D. Shaughnessy, Jr., D. A. Largaespada, H. G. Bedigian, A. M. Buchberg, N. A. Jenkins, and N. G. Copeland.** 1995. Frequent disruption of the Nf1 gene by a novel murine AIDS virus-related provirus in BXH-2 murine myeloid lymphomas. *J.Virol.* **69**:7138-7146.
80. **Largaespada, D. A., J. D. Shaughnessy, Jr., N. A. Jenkins, and N. G. Copeland.** 1995. Retroviral integration at the Evi-2 locus in BXH-2 myeloid leukemia cell lines disrupts Nf1 expression without changes in steady-state Ras-GTP levels. *J.Virol.* **69**:5095-5102.
81. **Glud, S. Z., A. B. Sorensen, M. Andrulis, B. Wang, E. Kondo, R. Jessen, L. Krenacs, E. Stelkovics, M. Wabl, E. Serfling, A. Palmethofer, and F. S. Pedersen.** 2005. A tumor-suppressor function for NFATc3 in T-cell lymphomagenesis by murine leukemia virus. *Blood* **106**:3546-3552.
82. **Dreyfus, F., B. Sola, S. Fichelson, P. Varlet, M. Charon, P. Tambourin, F. Wendling, and S. Gisselbrecht.** 1990. Rearrangements of the Pim-1, c-myc, and p53 genes in Friend helper virus-induced mouse erythroleukemias. *Leukemia* **4**:590-594.
83. **Bergeron, D., J. Houde, L. Poliquin, B. Barbeau, and E. Rassart.** 1993. Expression and DNA rearrangement of proto-oncogenes in Cas-Br-E-induced non-T-, non-B-cell leukemias. *Leukemia* **7**:954-962.
84. **Ben David, Y., M. R. Bani, B. Chabot, A. De Koven, and A. Bernstein.** 1992. Retroviral insertions downstream of the heterogeneous nuclear

ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. *Mol.Cell Biol.* **12**:4449-4455.

85. **Hanahan, D. and R. A. Weinberg.** 2000. The hallmarks of cancer. *Cell* **100**:57-70.
86. **Vogelstein, B. and K. W. Kinzler.** 1993. The multistep nature of cancer. *Trends Genet.* **9**:138-141.
87. **Michor, F., Y. Iwasa, and M. A. Nowak.** 2004. Dynamics of cancer progression. *Nat.Rev.Cancer* **4**:197-205.
88. **Vogelstein, B. and K. W. Kinzler.** 2004. Cancer genes and the pathways they control. *Nat.Med.* **10**:789-799.
89. **Loeb, L. A.** 2001. A mutator phenotype in cancer. *Cancer Res.* **61**:3230-3239.
90. **Tominaga, S.** 1999. Major avoidable risk factors of cancer. *Cancer Lett.* **143 Suppl 1**:S19-S23.
91. **Walboomers, J. M., M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. Snijders, J. Peto, C. J. Meijer, and N. Munoz.** 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J.Pathol.* **189**:12-19.
92. **Raab-Traub, N.** 1992. Epstein-Barr virus and nasopharyngeal carcinoma. *Semin.Cancer Biol.* **3**:297-307.
93. **Oudejans, J. J., N. M. Jiwa, A. J. van den Brule, and C. J. Meijer.** 1997. Epstein-Barr virus and its possible role in the pathogenesis of B-cell lymphomas. *Crit Rev.Oncol.Hematol.* **25**:127-138.
94. **Balmain, A., J. Gray, and B. Ponder.** 2003. The genetics and genomics of cancer. *Nat.Genet.* **33 Suppl**:238-244.
95. **Loeb, L. A., K. R. Loeb, and J. P. Anderson.** 2003. Multiple mutations and cancer. *Proc.Natl.Acad.Sci.U.S.A.* **100**:776-781.
96. **Sherr, C. J.** 2004. Principles of tumor suppression. *Cell* **116**:235-246.

97. **Kinzler, K. W. and B. Vogelstein.** 1997. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* **386**:761, 763.
98. **Vogelstein, B. and K. W. Kinzler.** 1998. *The Genetic Basis of Human Cancer.* McGraw-Hill, Toronto.
99. **Weiss, R.** 1984. Tissue-specific transformation by human T-cell leukaemia virus. *Nature* **310**:273-274.
100. **Varmus, H. E.** 1984. The molecular genetics of cellular oncogenes. *Annu.Rev.Genet.* **18**:553-612.
101. **Frackelton, A. R., Jr., P. M. Tremble, and L. T. Williams.** 1984. Evidence for the platelet-derived growth factor-stimulated tyrosine phosphorylation of the platelet-derived growth factor receptor in vivo. Immunopurification using a monoclonal antibody to phosphotyrosine. *J.Biol.Chem.* **259**:7909-7915.
102. **Nair, N., R. J. Davis, and H. L. Robinson.** 1992. Protein tyrosine kinase activities of the epidermal growth factor receptor and ErbB proteins: correlation of oncogenic activation with altered kinetics. *Mol.Cell Biol.* **12**:2010-2016.
103. **Barbacid, M.** 1987. ras genes. *Annu.Rev.Biochem.* **56**:779-827.
104. **Dang, C. V.** 1991. c-myc oncoprotein function. *Biochim.Biophys.Acta* **1072**:103-113.
105. **Ratner, L., B. Thielan, and T. Collins.** 1987. Sequences of the 5' portion of the human c-sis gene: characterization of the transcriptional promoter and regulation of expression of the protein product by 5' untranslated mRNA sequences. *Nucleic Acids Res.* **15**:6017-6036.
106. **Vaux, D. L., S. Cory, and J. M. Adams.** 1988. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* **335**:440-442.

107. **Canaani, E., R. P. Gale, D. Steiner-Saltz, A. Berrebi, E. Aghai, and E. Januszewicz.** 1984. Altered transcription of an oncogene in chronic myeloid leukaemia. *Lancet* **1**:593-595.
108. **Ellis, R. W., D. Defeo, T. Y. Shih, M. A. Gonda, H. A. Young, N. Tsuchida, D. R. Lowy, and E. M. Scolnick.** 1981. The p21 src genes of Harvey and Kirsten sarcoma viruses originate from divergent members of a family of normal vertebrate genes. *Nature* **292**:506-511.
109. **Watson, D. K., E. P. Reddy, P. H. Duesberg, and T. S. Papas.** 1983. Nucleotide sequence analysis of the chicken c-myc gene reveals homologous and unique coding regions by comparison with the transforming gene of avian myelocytomatosis virus MC29, delta gag-myc. *Proc.Natl.Acad.Sci.U.S.A.* **80**:2146-2150.
110. **Harris, H., O. J. Miller, G. Klein, P. Worst, and T. Tachibana.** 1969. Suppression of malignancy by cell fusion. *Nature* **223**:363-368.
111. **Stanbridge, E. J.** 1976. Suppression of malignancy in human cells. *Nature* **260**:17-20.
112. **Knudson, A. G., Jr.** 1971. Mutation and cancer: statistical study of retinoblastoma. *Proc.Natl.Acad.Sci.U.S.A.* **68**:820-823.
113. **Moolgavkar, S. H. and A. G. Knudson, Jr.** 1981. Mutation and cancer: a model for human carcinogenesis. *J.Natl.Cancer Inst.* **66**:1037-1052.
114. **Cavenee, W. K., T. P. Dryja, R. A. Phillips, W. F. Benedict, R. Godbout, B. L. Gallie, A. L. Murphree, L. C. Strong, and R. L. White.** 1983. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* **305**:779-784.
115. **Friend, S. H., R. Bernards, S. Rogelj, R. A. Weinberg, J. M. Rapaport, D. M. Albert, and T. P. Dryja.** 1986. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **323**:643-646.

116. **Knudson, A. G.** 2001. Two genetic hits (more or less) to cancer. *Nat.Rev.Cancer* **1**:157-162.
117. **Fero, M. L., E. Randel, K. E. Gurley, J. M. Roberts, and C. J. Kemp.** 1998. The murine gene p27Kip1 is haplo-insufficient for tumour suppression. *Nature* **396**:177-180.
118. **Venkatachalam, S., Y. P. Shi, S. N. Jones, H. Vogel, A. Bradley, D. Pinkel, and L. A. Donehower.** 1998. Retention of wild-type p53 in tumors from p53 heterozygous mice: reduction of p53 dosage can promote cancer formation. *EMBO J.* **17**:4657-4667.
119. **Tang, B., E. P. Bottinger, S. B. Jakowlew, K. M. Bagnall, J. Mariano, M. R. Anver, J. J. Letterio, and L. M. Wakefield.** 1998. Transforming growth factor-beta1 is a new form of tumor suppressor with true haploid insufficiency. *Nat.Med.* **4**:802-807.
120. **Inoue, K., F. Zindy, D. H. Randle, J. E. Rehg, and C. J. Sherr.** 2001. Dmp1 is haplo-insufficient for tumor suppression and modifies the frequencies of Arf and p53 mutations in Myc-induced lymphomas. *Genes Dev.* **15**:2934-2939.
121. **Cook, W. D. and B. J. McCaw.** 2000. Accommodating haploinsufficient tumor suppressor genes in Knudson's model. *Oncogene* **19**:3434-3438.
122. **Quon, K. C. and A. Berns.** 2001. Haplo-insufficiency? Let me count the ways. *Genes Dev.* **15**:2917-2921.
123. **Largaespada, D. A.** 2001. Haploinsufficiency for tumor suppression: the hazards of being single and living a long time. *J.Exp.Med.* **193**:F15-F18.
124. **Bodner, S. M., C. W. Naeve, K. M. Rakestraw, B. G. Jones, V. A. Valentine, M. B. Valentine, F. W. Luthardt, C. L. Willman, S. C. Raimondi, J. R. Downing, M. F. Roussel, C. J. Sherr, and A. T. Look.** 1999. Cloning and chromosomal localization of the gene encoding human cyclin D-binding Myb-like protein (hDMP1). *Gene* **229**:223-228.

125. **Sieber, O. M., K. Heinemann, and I. P. Tomlinson.** 2003. Genomic instability--the engine of tumorigenesis? *Nat.Rev.Cancer* **3**:701-708.
126. **Rajagopalan, H., M. A. Nowak, B. Vogelstein, and C. Lengauer.** 2003. The significance of unstable chromosomes in colorectal cancer. *Nat.Rev.Cancer* **3**:695-701.
127. **Lengauer, C., K. W. Kinzler, and B. Vogelstein.** 1998. Genetic instabilities in human cancers. *Nature* **396**:643-649.
128. **Nathanson, K. L., R. Wooster, and B. L. Weber.** 2001. Breast cancer genetics: what we know and what we need. *Nat.Med.* **7**:552-556.
129. **Bissell, M. J. and D. Radisky.** 2001. Putting tumours in context. *Nat.Rev.Cancer* **1**:46-54.
130. **Paige, A. J.** 2003. Redefining tumour suppressor genes: exceptions to the two-hit hypothesis. *Cell Mol.Life Sci.* **60**:2147-2163.
131. **Greenblatt, M. S., W. P. Bennett, M. Hollstein, and C. C. Harris.** 1994. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res.* **54**:4855-4878.
132. **Weinberg, R. A.** 1995. The retinoblastoma protein and cell cycle control. *Cell* **81**:323-330.
133. **Sandal, T.** 2002. Molecular aspects of the mammalian cell cycle and cancer. *Oncologist.* **7**:73-81.
134. **Hartwell, L. H.** 1991. Twenty-five years of cell cycle genetics. *Genetics* **129**:975-980.
135. **Lee, W. H., R. Bookstein, F. Hong, L. J. Young, J. Y. Shew, and E. Y. Lee.** 1987. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science* **235**:1394-1399.

136. **Gallie, B. L., C. Campbell, H. Devlin, A. Duckett, and J. A. Squire.** 1999. Developmental basis of retinal-specific induction of cancer by RB mutation. *Cancer Res.* **59**:1731s-1735s.
137. **Nevins, J. R.** 2001. The Rb/E2F pathway and cancer. *Hum.Mol.Genet.* **10**:699-703.
138. **Lee, E. Y., H. Cam, U. Ziebold, J. B. Rayman, J. A. Lees, and B. D. Dynlacht.** 2002. E2F4 loss suppresses tumorigenesis in Rb mutant mice. *Cancer Cell* **2**:463-472.
139. **Trimarchi, J. M. and J. A. Lees.** 2002. Sibling rivalry in the E2F family. *Nat.Rev.Mol.Cell Biol.* **3**:11-20.
140. **Ruas, M. and G. Peters.** 1998. The p16INK4a/CDKN2A tumor suppressor and its relatives. *Biochim.Biophys.Acta* **1378**:F115-F177.
141. **Roussel, M. F.** 1999. The INK4 family of cell cycle inhibitors in cancer. *Oncogene* **18**:5311-5317.
142. **Sherr, C. J. and J. M. Roberts.** 1999. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev.* **13**:1501-1512.
143. **Ortega, S., M. Malumbres, and M. Barbacid.** 2002. Cyclin D-dependent kinases, INK4 inhibitors and cancer. *Biochim.Biophys.Acta* **1602**:73-87.
144. **Sherr, C. J. and F. McCormick.** 2002. The RB and p53 pathways in cancer. *Cancer Cell* **2**:103-112.
145. **Lane, D. P. and L. V. Crawford.** 1979. T antigen is bound to a host protein in SV40-transformed cells. *Nature* **278**:261-263.
146. **Linzer, D. I. and A. J. Levine.** 1979. Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* **17**:43-52.
147. **DeLeo, A. B., G. Jay, E. Appella, G. C. Dubois, L. W. Law, and L. J. Old.** 1979. Detection of a transformation-related antigen in chemically induced

sarcomas and other transformed cells of the mouse. *Proc.Natl.Acad.Sci.U.S.A.* **76**:2420-2424.

148. **Finlay, C. A., P. W. Hinds, and A. J. Levine.** 1989. The p53 proto-oncogene can act as a suppressor of transformation. *Cell* **57**:1083-1093.
149. **Malkin, D., F. P. Li, L. C. Strong, J. F. Fraumeni, Jr., C. E. Nelson, D. H. Kim, J. Kassel, M. A. Gryka, F. Z. Bischoff, M. A. Tainsky, and .** 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**:1233-1238.
150. **Hesketh, R.** 1997. *The Oncogene and Tumour Suppressor Gene Facts Book.* Academic press.
151. **Prives, C.** 1998. Signaling to p53: breaking the MDM2-p53 circuit. *Cell* **95**:5-8.
152. **Jin, S. and A. J. Levine.** 2001. The p53 functional circuit. *J.Cell Sci.* **114**:4139-4140.
153. **Lane, D. P.** 1992. Cancer. p53, guardian of the genome. *Nature* **358**:15-16.
154. **Xiong, Y., G. J. Hannon, H. Zhang, D. Casso, R. Kobayashi, and D. Beach.** 1993. p21 is a universal inhibitor of cyclin kinases. *Nature* **366**:701-704.
155. **Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi, and S. J. Elledge.** 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**:805-816.
156. **Santarosa, M. and A. Ashworth.** 2004. Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. *Biochim.Biophys.Acta* **1654**:105-122.
157. **Haupt, Y., R. Maya, A. Kazaz, and M. Oren.** 1997. Mdm2 promotes the rapid degradation of p53. *Nature* **387**:296-299.

158. **Kubbutat, M. H., S. N. Jones, and K. H. Vousden.** 1997. Regulation of p53 stability by Mdm2. *Nature* **387**:299-303. "
159. **Michael, D. and M. Oren.** 2003. The p53-Mdm2 module and the ubiquitin system. *Semin.Cancer Biol.* **13**:49-58.
160. **Perry, M. E., J. Piette, J. A. Zawadzki, D. Harvey, and A. J. Levine.** 1993. The mdm-2 gene is induced in response to UV light in a p53-dependent manner. *Proc.Natl.Acad.Sci.U.S.A.* **90**:11623-11627.
161. **Juven, T., Y. Barak, A. Zauberman, D. L. George, and M. Oren.** 1993. Wild type p53 can mediate sequence-specific transactivation of an internal promoter within the mdm2 gene. *Oncogene* **8**:3411-3416.
162. **Sherr, C. J.** 1998. Tumor surveillance via the ARF-p53 pathway. *Genes Dev.* **12**:2984-2991.
163. **Zhang, Y. and Y. Xiong.** 2001. Control of p53 ubiquitination and nuclear export by MDM2 and ARF. *Cell Growth Differ.* **12**:175-186.
164. **Hakem, R. and T. W. Mak.** 2001. Animal models of tumor-suppressor genes. *Annu.Rev.Genet.* **35**:209-241.
165. **Leslie, N. R. and C. P. Downes.** 2002. PTEN: The down side of PI 3-kinase signalling. *Cell Signal.* **14**:285-295.
166. **Trotman, L. C. and P. P. Pandolfi.** 2003. PTEN and p53: who will get the upper hand? *Cancer Cell* **3**:97-99.
167. **Liu, M. C. and E. P. Gelmann.** 2002. P53 gene mutations: case study of a clinical marker for solid tumors. *Semin.Oncol.* **29**:246-257.
168. **Hirohashi, S. and Y. Kanai.** 2003. Cell adhesion system and human cancer morphogenesis. *Cancer Sci.* **94**:575-581.
169. **Esteller, M.** 2003. Relevance of DNA methylation in the management of cancer. *Lancet Oncol.* **4**:351-358.

170. **Herman, J. G. and S. B. Baylin.** 2003. Gene silencing in cancer in association with promoter hypermethylation. *N.Engl.J.Med.* **349**:2042-2054.
171. **Feinberg, A. P.** 2002. Genomic imprinting and cancer, p. 43-55. *In* B. Vogelstein and K. W. Kinzler (eds.), *The Genetic Basis of Human Cancer*. McGraw-Hill, New York.
172. **Beatty, J. A., J. J. Callanan, A. Terry, O. Jarrett, and J. C. Neil.** 1998. Molecular and immunophenotypical characterization of a feline immunodeficiency virus (FIV)-associated lymphoma: a direct role for FIV in B-lymphocyte transformation? *J.Virol.* **72**:767-771.
173. **Hardy, W. D., Jr., P. W. Hess, E. G. MacEwen, A. J. McClelland, E. E. Zuckerman, M. Essex, S. M. Cotter, and O. Jarrett.** 1976. Biology of feline leukemia virus in the natural environment. *Cancer Res.* **36**:582-588.
174. **Jarret, O.** 1994. Feline leukaemia virus, p. 473-487. *In* E. A. Chandler, C. J. Gaskell, and R. M. Gaskell (eds.), *Feline medicine and therapeutics*. Blackwell Scientific Publications Ltd., Oxford, United Kingdom.
175. **Shelton, G. H., K. D. McKim, P. L. Cooley, P. F. Dice, R. G. Russell, and C. K. Grant.** 1989. Feline leukemia virus and feline immunodeficiency virus infections in a cat with lymphoma. *J.Am.Vet.Med.Assoc.* **194**:249-252.
176. **Sabine, M., J. Michelsen, F. Thomas, and m. Zheng.** 1988. Feline AIDS. *Aust.Vet.Practit.* **18**:105-107.
177. **Rosenberg, M. P., A. E. Hohenhaus, and R. E. Matus.** 1991. Monoclonal gammopathy and lymphoma in a cat infected with feline immunodeficiency virus. *J.Am.Anim.Hosp.Assoc.* **27**:335-337.
178. **Poli, A., F. Abramo, F. Baldinotti, M. Pistello, L. Da Prato, and M. Bendinelli.** 1994. Malignant lymphoma associated with experimentally induced feline immunodeficiency virus infection. *J.Comp Pathol.* **110**:319-328.

179. **Ishida, T., T. Washizu, K. Toriyabe, S. Motoyoshi, I. Tomoda, and N. C. Pedersen.** 1989. Feline immunodeficiency virus infection in cats of Japan. *J.Am.Vet.Med.Assoc.* **194**:221-225.
180. **Callanan, J. J., I. A. McCandlish, B. O'Neil, C. E. Lawrence, M. Rigby, A. M. Pacitti, and O. Jarrett.** 1992. Lymphosarcoma in experimentally induced feline immunodeficiency virus infection [corrected]. *Vet.Rec.* **130**:293-295.
181. **Buracco, P., R. Guglielmino, O. Abate, V. Bocchini, E. Cornaglia, D. B. DeNicola, M. Cilli, and P. Ponzio.** 1992. Large granular lymphoma in a FIV-positive and FeLV-negative cat. *J.Small Anim.Pract.* **33**:279-284.
182. **Alexander, R., W. F. Robinson, J. N. Mills, C. R. Sherry, E. Sherard, A. J. Paterson, S. E. Shaw, W. T. Clark, and T. Hillingsworth.** 1989. Isolation of feline immunodeficiency from three cats with lymphoma. *Aust.Vet.Pract.* **19**:93-97.
183. **Shelton, G. H., C. K. Grant, S. M. Cotter, M. B. Gardner, W. D. Hardy, Jr., and R. F. DiGiacomo.** 1990. Feline immunodeficiency virus and feline leukemia virus infections and their relationships to lymphoid malignancies in cats: a retrospective study (1968-1988). *J.Acquir.Immune.Defic.Sydr.* **3**:623-630.
184. **Hutson, C. A., B. A. Rideout, and N. C. Pedersen.** 1991. Neoplasia associated with feline immunodeficiency virus infection in cats of southern California. *J.Am.Vet.Med.Assoc.* **199**:1357-1362.
185. **Feder, B. M. and A. I. Hurvitz.** 1990. Feline immunodeficiency virus infection in 100 cats and association with lymphoma. *Proc.ACVIM Forum* **8**:1112.
186. **Bennett, M. and N. R. Smyth.** 1992. Feline immunodeficiency virus: a brief review. *Br.Vet.J.* **148**:399-412.
187. **Richards, J. R.** 2005. Feline immunodeficiency virus vaccine: implications for diagnostic testing and disease management. *Biologicals* **33**:215-217.

188. **Ackley, C. D., J. K. Yamamoto, N. Levy, N. C. Pedersen, and M. D. Cooper.** 1990. Immunologic abnormalities in pathogen-free cats experimentally infected with feline immunodeficiency virus. *J.Virol.* **64**:5652-5655.
189. **English, R. V., C. M. Johnson, D. H. Gebhard, and M. B. Tompkins.** 1993. In vivo lymphocyte tropism of feline immunodeficiency virus. *J.Virol.* **67**:5175-5186.
190. **Dean, G. A., G. H. Reubel, P. F. Moore, and N. C. Pedersen.** 1996. Proviral burden and infection kinetics of feline immunodeficiency virus in lymphocyte subsets of blood and lymph node. *J.Virol.* **70**:5165-5169.
191. **Callanan, J. J., H. Thompson, S. R. Toth, B. O'Neil, C. E. Lawrence, B. Willett, and O. Jarrett.** 1992. Clinical and pathological findings in feline immunodeficiency virus experimental infection. *Vet.Immunol.Immunopathol.* **35**:3-13.
192. **Flynn, J. N., C. A. Cannon, C. E. Lawrence, and O. Jarrett.** 1994. Polyclonal B-cell activation in cats infected with feline immunodeficiency virus. *Immunology* **81**:626-630.
193. **Lane, H. C., H. Masur, L. C. Edgar, G. Whalen, A. H. Rook, and A. S. Fauci.** 1983. Abnormalities of B-cell activation and immunoregulation in patients with the acquired immunodeficiency syndrome. *N.Engl.J.Med.* **309**:453-458.
194. **Ammann, A. J., D. Abrams, M. Conant, D. Chudwin, M. Cowan, P. Volberding, B. Lewis, and C. Casavant.** 1983. Acquired immune dysfunction in homosexual men: immunologic profiles. *Clin.Immunol.Immunopathol.* **27**:315-325.
195. **McChesney, M. B. and M. B. Oldstone.** 1989. Virus-induced immunosuppression: infections with measles virus and human immunodeficiency virus. *Adv.Immunol.* **45**:335-380.

196. **Bennett, M., C. McCracken, H. Lutz, C. J. Gaskell, R. M. Gaskell, A. Brown, and J. O. Knowles.** 1989. "Prevalence of antibody to feline immunodeficiency virus in some cat populations. *Vet.Rec.* **124**:397-398.
197. **Pedersen, N. C., J. K. Yamamoto, T. Ishida, and H. Hansen.** 1989. Feline immunodeficiency virus infection. *Vet.Immunol.Immunopathol.* **21**:111-129.
198. **Yamamoto, J. K., H. Hansen, E. W. Ho, T. Y. Morishita, T. Okuda, T. R. Sawa, R. M. Nakamura, and N. C. Pedersen.** 1989. Epidemiologic and clinical aspects of feline immunodeficiency virus infection in cats from the continental United States and Canada and possible mode of transmission. *J.Am.Vet.Med.Assoc.* **194**:213-220.
199. **Tompkins, M. B., P. D. Nelson, R. V. English, and C. Novotney.** 1991. Early events in the immunopathogenesis of feline retrovirus infections. *J.Am.Vet.Med.Assoc.* **199**:1311-1315.
200. **Torten, M., M. Franchini, J. E. Barlough, J. W. George, E. Mozes, H. Lutz, and N. C. Pedersen.** 1991. Progressive immune dysfunction in cats experimentally infected with feline immunodeficiency virus. *J.Virol.* **65**:2225-2230.
201. **Davidson, M. G., J. B. Rottman, R. V. English, M. R. Lappin, and M. B. Tompkins.** 1993. Feline immunodeficiency virus predisposes cats to acute generalized toxoplasmosis. *Am.J.Pathol.* **143**:1486-1497.
202. **Barlough, J. E., C. D. Ackley, J. W. George, N. Levy, R. Acevedo, P. F. Moore, B. A. Rideout, M. D. Cooper, and N. C. Pedersen.** 1991. Acquired immune dysfunction in cats with experimentally induced feline immunodeficiency virus infection: comparison of short-term and long-term infections. *J.Acquir.Immune.Defic.Sydr.* **4**:219-227.
203. **Siebelink, K. H., I. H. Chu, G. F. Rimmelzwaan, K. Weijer, R. van Herwijnen, P. Knell, H. F. Egberink, M. L. Bosch, and A. D. Osterhaus.** 1990. Feline immunodeficiency virus (FIV) infection in the cat as a model for

HIV infection in man: FIV-induced impairment of immune function. *AIDS Res.Hum.Retroviruses* **6**:1373-1378. "

204. **Sparger, E. E., B. L. Shacklett, L. Renshaw-Gegg, P. A. Barry, N. C. Pedersen, J. H. Elder, and P. A. Luciw.** 1992. Regulation of gene expression directed by the long terminal repeat of the feline immunodeficiency virus. *Virology* **187**:165-177.
205. **English, R. V., P. Nelson, C. M. Johnson, M. Nasisse, W. A. Tompkins, and M. B. Tompkins.** 1994. Development of clinical disease in cats experimentally infected with feline immunodeficiency virus. *J.Infect.Dis.* **170**:543-552.
206. **Callanan, J. J., B. A. Jones, J. Irvine, B. J. Willett, I. A. McCandlish, and O. Jarrett.** 1996. Histologic classification and immunophenotype of lymphosarcomas in cats with naturally and experimentally acquired feline immunodeficiency virus infections. *Vet.Pathol.* **33**:264-272.
207. **Levine, A. M.** 1992. Acquired immunodeficiency syndrome-related lymphoma. *Blood* **80**:8-20.
208. **Herndier, B. G., L. D. Kaplan, and M. S. McGrath.** 1994. Pathogenesis of AIDS lymphomas. *AIDS* **8**:1025-1049.
209. **Ramsay, A. D., J. Giddings, A. Baskerville, and M. P. Cranage.** 1991. Phenotypic analysis of malignant lymphoma in simian immunodeficiency virus infection using anti-human antibodies. *J.Pathol.* **164**:321-328.
210. **Feichtinger, H., P. Putkonen, C. Parravicini, S. L. Li, E. E. Kaaya, D. Bottiger, G. Biberfeld, and P. Biberfeld.** 1990. Malignant lymphomas in cynomolgus monkeys infected with simian immunodeficiency virus. *Am.J.Pathol.* **137**:1311-1315.
211. **Hernandez, A. M. and D. Shibata.** 1995. Epstein-Barr virus-associated non-Hodgkin's lymphoma in HIV-infected patients. *Leuk.Lymphoma* **16**:217-221.

212. **Ballerini, P., G. Gaidano, J. Z. Gong, V. Tassi, G. Saglio, D. M. Knowles, and R. Dalla-Favera.** 1993. Multiple genetic lesions in acquired immunodeficiency syndrome-related non-Hodgkin's lymphoma. *Blood* **81**:166-176.
213. **Beatty, J. A., C. E. Lawrence, J. J. Callanan, C. K. Grant, E. A. Gault, J. C. Neil, and O. Jarrett.** 1998. Feline immunodeficiency virus (FIV)-associated lymphoma: a potential role for immune dysfunction in tumourigenesis. *Vet.Immunol.Immunopathol.* **65**:309-322.
214. **Terry, A., J. J. Callanan, R. Fulton, O. Jarrett, and J. C. Neil.** 1995. Molecular analysis of tumours from feline immunodeficiency virus (FIV)-infected cats: an indirect role for FIV? *Int.J.Cancer* **61**:227-232.
215. **Laurence, J. and S. M. Astrin.** 1991. Human immunodeficiency virus induction of malignant transformation in human B lymphocytes. *Proc.Natl.Acad.Sci.U.S.A.* **88**:7635-7639.
216. **Astrin, S. M., E. Schattner, J. Laurence, R. I. Lebman, and C. Rodriguez-Alfageme.** 1992. Does HIV infection of B lymphocytes initiate AIDS lymphoma? Detection by PCR of viral sequences in lymphoma tissue. *Curr.Top.Microbiol.Immunol.* **182**:399-407.
217. **Herndier, B. G., B. T. Shiramizu, N. E. Jewett, K. D. Aldape, G. R. Reyes, and M. S. McGrath.** 1992. Acquired immunodeficiency syndrome-associated T-cell lymphoma: evidence for human immunodeficiency virus type 1-associated T-cell transformation. *Blood* **79**:1768-1774.
218. **Neer, E. J., C. J. Schmidt, R. Nambudripad, and T. F. Smith.** 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371**:297-300.
219. **Li, D. and R. Roberts.** 2001. WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell Mol.Life Sci.* **58**:2085-2097.
220. **Dualan, R., T. Brody, S. Keeney, A. F. Nichols, A. Admon, and S. Linn.** 1995. Chromosomal localization and cDNA cloning of the genes (DDB1 and

- DDB2) for the p127 and p48 subunits of a human damage-specific DNA binding protein. *Genomics* **29**:62-69. "
221. **Esteller, M., G. Gaidano, S. N. Goodman, V. Zagonel, D. Capello, B. Botto, D. Rossi, A. Glohini, U. Vitolo, A. Carbone, S. B. Baylin, and J. G. Herman.** 2002. Hypermethylation of the DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients with diffuse large B-cell lymphoma. *J.Natl.Cancer Inst.* **94**:26-32.
222. **Tang, J. and G. Chu.** 2002. Xeroderma pigmentosum complementation group E and UV-damaged DNA-binding protein. *DNA Repair (Amst)* **1**:601-616.
223. **Takao, M., M. Abramic, M. Moos, Jr., V. R. Otrin, J. C. Wootton, M. McLenigan, A. S. Levine, and M. Protic.** 1993. A 127 kDa component of a UV-damaged DNA-binding complex, which is defective in some xeroderma pigmentosum group E patients, is homologous to a slime mold protein. *Nucleic Acids Res.* **21**:4111-4118.
224. **Lee, T. H., S. J. Elledge, and J. S. Butel.** 1995. Hepatitis B virus X protein interacts with a probable cellular DNA repair protein. *J.Virol.* **69**:1107-1114.
225. **Groisman, R., J. Polanowska, I. Kuraoka, J. Sawada, M. Saijo, R. Drapkin, A. F. Kisselev, K. Tanaka, and Y. Nakatani.** 2003. The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**:357-367.
226. **Tan, T. and G. Chu.** 2002. p53 Binds and activates the xeroderma pigmentosum DDB2 gene in humans but not mice. *Mol.Cell Biol.* **22**:3247-3254.
227. **Fujiwara, Y., C. Masutani, T. Mizukoshi, J. Kondo, F. Hanaoka, and S. Iwai.** 1999. Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *J.Biol.Chem.* **274**:20027-20033.
228. **Aboussekhra, A., M. Biggerstaff, M. K. Shivji, J. A. Vilpo, V. Moncollin, V. N. Podust, M. Protic, U. Hubscher, J. M. Egly, and R. D. Wood.** 1995.

Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80**:859-868. "

229. **Mu, D., C. H. Park, T. Matsunaga, D. S. Hsu, J. T. Reardon, and A. Sancar.** 1995. Reconstitution of human DNA repair excision nuclease in a highly defined system. *J.Biol.Chem.* **270**:2415-2418.
230. **Bessho, T., A. Sancar, L. H. Thompson, and M. P. Thelen.** 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *J.Biol.Chem.* **272**:3833-3837.
231. **Wakasugi, M., M. Shimizu, H. Morioka, S. Linn, O. Nikaido, and T. Matsunaga.** 2001. Damaged DNA-binding protein DDB stimulates the excision of cyclobutane pyrimidine dimers in vitro in concert with XPA and replication protein A. *J.Biol.Chem.* **276**:15434-15440.
232. **de Laat, W. L., N. G. Jaspers, and J. H. Hoeijmakers.** 1999. Molecular mechanism of nucleotide excision repair. *Genes Dev.* **13**:768-785.
233. **Fitch, M. E., I. V. Cross, and J. M. Ford.** 2003. p53 responsive nucleotide excision repair gene products p48 and XPC, but not p53, localize to sites of UV-irradiation-induced DNA damage, in vivo. *Carcinogenesis* **24**:843-850.
234. **Mellon, I., G. Spivak, and P. C. Hanawalt.** 1987. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**:241-249.
235. **Nichols, A. F., P. Ong, and S. Linn.** 1996. Mutations specific to the xeroderma pigmentosum group E Ddb- phenotype. *J.Biol.Chem.* **271**:24317-24320.
236. **Hwang, B. J., J. M. Ford, P. C. Hanawalt, and G. Chu.** 1999. Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. *Proc.Natl.Acad.Sci.U.S.A.* **96**:424-428.

237. **Tang, J. Y., B. J. Hwang, J. M. Ford, P. C. Hanawalt, and G. Chu.** 2000. Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis. *Mol.Cell* **5**:737-744.
238. **Hayflick L. and Moorhead PS.** 1961. The serial cultivation of human diploid cell strains. *Exp.Cell Res.* **25**:585-621.
239. **Foley G.E., Lazarus H., Farber S., Uzman B.G., Boone B.A., and McCarthy R.E.** 1965. Continuous culture of human lymphoblasts from peripheral blood of a child with acute leukemia. *Cancer* **18**:522-529.
240. **Sundstrom, C. and K. Nilsson.** 1976. Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int.J.Cancer* **17**:565-577.
241. **Koeffler, H. P., R. Billing, A. J. Lysis, R. Sparkes, and D. W. Golde.** 1980. An undifferentiated variant derived from the human acute myelogenous leukemia cell line (KG-1). *Blood* **56**:265-273.
242. **Lozzio, C. B. and B. B. Lozzio.** 1975. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood* **45**:321-334.
243. **Weiss, A., R. L. Wiskocil, and J. D. Stobo.** 1984. The role of T3 surface molecules in the activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level. *J.Immunol.* **133**:123-128.
244. **Asou, H., S. Tashiro, K. Hamamoto, A. Otsuji, K. Kita, and N. Kamada.** 1991. Establishment of a human acute myeloid leukemia cell line (Kasumi-1) with 8;21 chromosome translocation. *Blood* **77**:2031-2036.
245. **Leibovitz, A., J. C. Stinson, W. B. McCombs, III, C. E. McCoy, K. C. Mazur, and N. D. Mabry.** 1976. Classification of human colorectal adenocarcinoma cell lines. *Cancer Res.* **36**:4562-4569.
246. **Baxter, E. W., K. Blyth, E. R. Cameron, and J. C. Neil.** 2001. Selection for loss of p53 function in T-cell lymphomagenesis is alleviated by Moloney

- murine leukemia virus infection in myc transgenic mice. *J.Virol.* **75**:9790-9798.
247. **Ralph, P.** 1973. Retention of lymphocyte characteristics by myelomas and theta + -lymphomas: sensitivity to cortisol and phytohemagglutinin. *J.Immunol.* **110**:1470-1475.
248. **Todaro G.J. and Green H.** 1963. *Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines.* *J.Cell Biol.* **17**:299-313.
249. **Sambrook, J., E. Fritsch, and T. Maniatis.** 1989. *Molecular Cloning : A Laboratory Manual.* Cold Spring Harbor, New York.
250. **Ausubel, F., R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, and K. Struhl.** 1994. *Current protocols in molecular biology,* New York.
251. **Birnboim, H. C. and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.
252. **Chung, C. T., S. L. Niemela, and R. H. Miller.** 1989. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc.Natl.Acad.Sci.U.S.A.* **86**:2172-2175.
253. **Grunstein, M. and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc.Natl.Acad.Sci.U.S.A.* **72**:3961-3965.
254. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**:5294-5299.
255. **Chomczynski, P. and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal.Biochem.* **162**:156-159.
256. **Löffert, D., S. Karger, M. Berkenkopf, N. Seip, and Kang J.** 1997. PCR optimization: Primer design. *QIAGEN News* 3-6.

257. **Kwok, S., D. E. Kellogg, N. McKinney, D. Spasic, L. Goda, C. Levenson, and J. J. Sninsky.** 1990. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. *Nucleic Acids Res.* **18**:999-1005.
258. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J.Mol.Biol.* **98**:503-517.
259. **Marais, R., J. Wynne, and R. Treisman.** 1993. The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell* **73**:381-393.
260. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal.Biochem.* **72**:248-254.
261. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-685.
262. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc.Natl.Acad.Sci.U.S.A.* **76**:4350-4354.
263. **Durrant, I., L. C. Benge, C. Sturrock, A. T. Devenish, R. Howe, S. Roe, M. Moore, G. Scozzafava, L. M. Proudfoot, T. C. Richardson, and .** 1990. The application of enhanced chemiluminescence to membrane-based nucleic acid detection. *Biotechniques* **8**:564-570.
264. **Ioshikhes, I. P. and M. Q. Zhang.** 2000. Large-scale human promoter mapping using CpG islands. *Nat.Genet.* **26**:61-63.
265. **Hannenhalli, S. and S. Levy.** 2001. Promoter prediction in the human genome. *Bioinformatics.* **17 Suppl 1**:S90-S96.
266. **Prestridge, D. S.** 1995. Predicting Pol II promoter sequences using transcription factor binding sites. *J.Mol.Biol.* **249**:923-932.

267. **Heinemeyer, T., E. Wingender, I. Reuter, H. Hermjakob, A. E. Kel, O. V. Kel, E. V. Ignatieva, E. A. Ananko, O. A. Podkolodnaya, F. A. Kolpakov, N. L. Podkolodny, and N. A. Kolchanov.** 1998. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Res.* **26**:362-367.
268. **Liu, H. and L. Wong.** 2003. Data mining tools for biological sequences. *J.Bioinform.Comput.Biol.* **1**:139-167.
269. **Guda, C. and S. Subramaniam.** 2005. pTARGET [corrected] a new method for predicting protein subcellular localization in eukaryotes. *Bioinformatics.* **21**:3963-3969.
270. **Cullen, B. R., P. T. Lomedico, and G. Ju.** 1984. Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis. *Nature* **307**:241-245.
271. **McDougall, A. S., A. Terry, T. Tzavaras, C. Cheney, J. Rojko, and J. C. Neil.** 1994. Defective endogenous proviruses are expressed in feline lymphoid cells: evidence for a role in natural resistance to subgroup B feline leukemia viruses. *J.Virol.* **68**:2151-2160.
272. **Stewart, M., A. Terry, M. Hu, M. O'Hara, K. Blyth, E. Baxter, E. Cameron, D. E. Onions, and J. C. Neil.** 1997. Proviral insertions induce the expression of bone-specific isoforms of PEBP2alphaA (CBFA1): evidence for a new myc collaborating oncogene. *Proc.Natl.Acad.Sci.U.S.A.* **94**:8646-8651.
273. **Hopp, T. P. and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc.Natl.Acad.Sci.U.S.A.* **78**:3824-3828.
274. **Kyte, J. and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. *J.Mol.Biol.* **157**:105-132.
275. **Hopp, T. P. and K. R. Woods.** 1983. A computer program for predicting protein antigenic determinants. *Mol.Immunol.* **20**:483-489.

276. **Harlow, E. and D. Lane.** 1988. Immunizations, p. 53-137. *In* Antibodies. Cold Spring Harbor Laboratory.
277. **Smith, D. B. and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**:31-40.
278. **Marston, F. A.** 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem.J.* **240**:1-12.
279. **Frankel, S., R. Sohn, and L. Leinwand.** 1991. The use of sarkosyl in generating soluble protein after bacterial expression. *Proc.Natl.Acad.Sci.U.S.A.* **88**:1192-1196.
280. **Remm, M., C. E. Storm, and E. L. Sonnhammer.** 2001. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J.Mol.Biol.* **314**:1041-1052.
281. **Gentile, M., L. Latonen, and M. Laiho.** 2003. Cell cycle arrest and apoptosis provoked by UV radiation-induced DNA damage are transcriptionally highly divergent responses. *Nucleic Acids Res.* **31**:4779-4790.
282. **Itoh, T., A. Nichols, and S. Linn.** 2001. Abnormal regulation of DDB2 gene expression in xeroderma pigmentosum group E strains. *Oncogene* **20**:7041-7050.
283. **Nichols, A. F., T. Itoh, J. A. Graham, W. Liu, M. Yamaizumi, and S. Linn.** 2000. Human damage-specific DNA-binding protein p48. Characterization of XPE mutations and regulation following UV irradiation. *J.Biol.Chem.* **275**:21422-21428.
284. **Gavin, A. C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edlmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B.**

- Kuster, G. Neubauer, and G. Superti-Furga.** 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**:141-147.
285. **Edelmann, W., K. Yang, A. Umar, J. Heyer, K. Lau, K. Fan, W. Liedtke, P. E. Cohen, M. F. Kane, J. R. Lipford, N. Yu, G. F. Crouse, J. W. Pollard, T. Kunkel, M. Lipkin, R. Kolodner, and R. Kucherlapati.** 1997. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell* **91**:467-477.
286. **Wiesendanger, M., B. Kneitz, W. Edelmann, and M. D. Scharff.** 2000. Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. *J.Exp.Med.* **191**:579-584.
287. **Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman, and E. K. O'Shea.** 2003. Global analysis of protein localization in budding yeast. *Nature* **425**:686-691.
288. **Hesketh, J.** 2004. 3'-Untranslated regions are important in mRNA localization and translation: lessons from selenium and metallothionein. *Biochem.Soc.Trans.* **32**:990-993.
289. **Wolff, B., J. J. Sanglier, and Y. Wang.** 1997. Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. *Chem.Biol.* **4**:139-147.
290. **Campbell, M. R., P. N. Nation, and S. E. Andrew.** 2005. A lack of DNA mismatch repair on an athymic murine background predisposes to hematologic malignancy. *Cancer Res.* **65**:2626-2635.
291. **Yang, M., Z. Wu, and S. Fields.** 1995. Protein-peptide interactions analyzed with the yeast two-hybrid system. *Nucleic Acids Res.* **23**:1152-1156.
292. **Amundson, S. A., K. T. Do, L. Vinikoor, C. A. Koch-Paiz, M. L. Bittner, J. M. Trent, P. Meltzer, and A. J. Fornace, Jr.** 2005. Stress-specific

signatures: expression profiling of p53 wild-type and -null human cells. *Oncogene* **24**:4572-4579.

293. **Elbashir, S. M., J. Harborth, K. Weber, and T. Tuschl.** 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **26**:199-213.
294. **Natrajan, R., J. Louhelainen, S. Williams, J. Laye, and M. A. Knowles.** 2003. High-resolution deletion mapping of 15q13.2-q21.1 in transitional cell carcinoma of the bladder. *Cancer Res.* **63**:7657-7662.
295. **Fleischmann, C., S. Bevan, J. C. Neil, A. Terry, and R. S. Houlston.** 2003. Mutations in the candidate tumour suppressor gene FLJ12973 on chromosome 15q15 are rare in colorectal cancer. *Cancer Lett.* **196**:65-67.
296. **Keays, D. A., T. G. Clark, and J. Flint.** 2006. Estimating the number of coding mutations in genotypic- and phenotypic-driven N-ethyl-N-nitrosourea (ENU) screens. *Mamm.Genome* **17**:230-238.

