CHARACTERISATION OF THE *FIT*-1 COMMON INTEGRATION LOCUS

Nighean I Barr

A thesis submitted for the degree of Doctor of Philosophy

Department of Veterinary Pathology
Faculty of Veterinary Medicine
University of Glasgow
April, 1999

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

List of Contents

List of figures and tables

Chapter 1

Chapter 3

Chapter 4

Chapter 7

Chapter 8

Abbreviations

ALV	avian leukosis virus
A-MuLV	Abelson murine leukaemia virus
bp	base pairs
cDNA	complementary DNA
cM	centi Morgan
C-terminus	carboxy terminus
dCTP	deoxycytosine triphosphate
EDTA	ethylene-diamine-tetra-acetic acid
EUCIB	European interspecific backcross
FCS	foetal calf serum
FeLV	feline leukaemia virus
F-MuLV	Friend murine leukaemia virus
gapdh	glyceraldehyde 3-phosphate dehydrogenase
HPRT	hypoxanthine phosphoribosyl-transferase
IPTG	isopropylthio-β-galactosidase
kb	kilobase pairs
LTR	long terminal repeat
LZ	leucine zipper
Mo-MuLV	Moloney murine leukaemia virus
MOPS	sodim morpholinopropane sulphonic acid
NRD	negative regulatory domain
N-terminus	amino terminus
OD	optical density
ORF	open reading frame
PAC	P1 derived artificial chromosome
PCR	polymerase chain reaction
RFLP	restriction fragment length polymorphism
rpm	revolutions per minute
RPMI	Rosewall Park Memorial Institute medium
RT-PCR	reverse transcribed PCR

SDS	sodium dodecyl sulphate
STS	sequence tagged site
TEMED	N,N,N,N,-tetramethylethylenediamine
Tris	2-amino-2(hydroxymethyl)-1,3-propandiol
UTR	untranslated region
X-gal	5-bromo 4-chloro 3-indolyl β-galactosidase
YAC	yeast artificial chromosome

Acknowledgements

I would firstly like to thank my supervisors Professor Jim Neil and Dr Monica Stewart for their help and encouragement for which I am very grateful. Thanks also go to all the members of the Molecular Oncology Laboratory especially Dr Euan Baxter, Dr Francoise Vaillant, Dr Ming Hu, Dr Moyra Campbell and of course Tom McPherson. Special thanks go to Anne Terry for her endless help and suggestions.

I also acknowlede the assistance of Dr Angus Lauder and Dr Kathy Weston (ICR, London) for their initial suggestions, advice and probes for c-*myb* analysis; Dr Linda Wolff (NIH, Bethesda, Maryland, USA) for the *Mml*1 probe; Dr Naoya Yuhki (NCI Frederick, Maryland, USA) for screening his feline PAC library with the *Fit*-1 probe. I specifically acknowledge Professor Hajime Tsujimoto (University of Tokyo, Japan) for sending the FT1 cell line. The work was supported by the Leukaemia Research Fund.

Finally, special thanks go to Andy Stevenson for his seemingly eternal patience, encouragement, support and sanity, all of which is greatly appreciated.

Declaration

The studies described in this thesis were performed in the Department of Veterinary Pathology at the University of Glasgow Veterinary School. The author is responsible for all the results contained in this thesis except where it is stated otherwise. In particular, the initial analysis of the *Fit*-1 locus was carried out by Dr Ruth Fulton, Professor Hajime Tsujimoto and Dr Christos Tsatsanis and is described in section 3.1.2.

Sections of this thesis have been reproduced in the following publication and abstracts:

Barr NI, Stewart M, Tsatsanis C, Fulton R, Hu M, Tsujimoto H, and Neil JC. 1999. The *Fit*-1 common integration locus in human and mouse is closely linked to *MYB*. Mammalian Genome (In press). (Reference 157 and Appendix 2).

Barr NI, Stewart M, Neil JC. *Fit*-1, an integration locus defining a putative *myc*-collaborating gene, is closely linked to c-*myb*. Abstract: Cold Spring Harbor Meeting on Cancer Genetics and Tumor Suppressor Genes, August 1998 and the Fourth International Feline Retrovirus Research Symposium, May 1998.

Nighean I Barr, April 1999.

Summary

The aim of this study was to identify the gene which was affected by proviral integration within the *Fit*-1 common integration locus. The *Fit*-1 locus was identified in the cat as a common site of insertion for feline leukaemia virus (FeLV) in thymic lymphomas induced by FeLV-*myc* recombinant viruses. Thus precedent suggested that it harbours a gene which cooperates with c-*myc* in T-cell leukaemogenesis. Using a zooblot positive probe from the major insertion cluster, the mouse homologue was mapped to chromosome 10 and the human homologue to chromosome 6 (N. Barr *et al*, in press and Appendix 2). A cosmid clone containing homologous human sequences was isolated and used to generate a PAC clone contig around the human *Fit*-1 locus, revealing that it is only 100kb upstream of the c-*myb* gene.

A number of retroviral insertion sites have been mapped close to c-*myb* on mouse chromosome 10, *Mml*1, *Ahi*-1 and *Mis*-2. Thus, the close proximity of c-*myb* and *Fit*-1 in the human genome and the lack of obvious transcribed sequences at these loci might suggest that retroviral insertions within this cluster of loci activates the c-*myb* gene by long-range activation. However, analysis of c-*myb* expression in tumours rearranged at *Ahi*-1, *Mis*-2 and *Mml*1 failed to show evidence of long-range activation. In contrast, careful analysis of the FT1 cell line which carries an insertion at *Fit*-1 and c-*myc* revealed evidence of approximately 2-fold over-expression of c-*myb* which may, at least in part, be due to a decreased turnover of c-*myb* RNA. This subtle change in expression may have been missed in previous analyses of primary tumour RNA from tumours rearranged at *Fit*-1, *Ahi*-1 or *Mis*-2. These results raised the unexpected prospect that Myc and Myb can collaborate in T-lymphoma development.

In addition to over-expression of the c-*myb* gene in the FT1 cell line, there was also evidence of an altered c-*myb* transcript structure in this cell line. Further analysis revealed that this was due to alternative splicing of the c-*myb* gene involving the alternative exon 9A, a form which was approximately 5-fold over-

expressed in FT1 cells. Although this form has been associated with both normal and activated forms of c-*myb*, levels seen in the FT1 cell line were higher than in previous studies.

Thus, it appears that c-*myb* may be the true target of insertions at the *Fit*-1/*Mml*1/*Ahi*-1/*Mis*-2 cluster and that, at least for *Fit*-1, retroviral integration may promote the alternative splicing of c-*myb*, possibly favouring a more oncogenic isoform.

CHAPTER 1

Chapter 1

1. Introduction

Cancer is essentially a genetic disease of which around 200 different forms have been recognised in man. Despite the vastly different pathologies of cancers, they all share the same key features at the molecular level, namely uncontrolled cell division and inappropriate survival. Tumour growth is typically a slow, multi-step process in which 5 or 6 separate genetic, or epigenetic, events must occur to produce the fully malignant phenotype (reviewed in (1)). The multi-step nature of this disease can be illustrated by the correlation of an increased risk of cancer with age. The incidence of cancer increases exponentially with time as individuals are exposed to mutagens thus increasing the risk of cancer as we get older.

There are basically two categories of genes which are the frequent targets of somatic mutations during the development of cancer. Firstly, proto-oncogenes, which generally exhibit dominant behaviour when mutated (becoming an oncogene). Proto-oncogenes are often involved in proliferation and cell signaling pathways and it is usually their deregulation which is responsible for their tumorigenic potential. Secondly, tumour suppressor genes are frequently implicated in cancer where their function is lost by the acquisition of mutation(s). Tumour suppressor genes are involved in, for example, the repair and maintenance of DNA and regulation of the cell cycle. They are generally recessive where loss of function of both copies of the gene is required for their tumorigenic potential.

Most cancers will have lost suppressor gene and gained oncogene functions. This could occur by, for example, exposure to mutagens such as chemicals, ionising radiation or tumour viruses. Such mutations often lead to deregulated proliferation. However, equally important is the deregulation of the normal processes leading to programmed cell death (apoptosis) or terminal differentiation.

A major goal of cancer research has been to identify the genes which contribute to tumour development and which genes can collaborate in this multi-step process. A major contribution to the identification of cellular oncogenes have come from retroviruses. Retroviral oncogenesis is associated with either transduction or insertional mutagenesis of a host cell proto-oncogene. The isolation of the sequence which has been transduced by the virus or the sequence flanking the integrated provirus has provided a vast amount of knowledge of genes and pathways involved in tumorigenesis.

1.1 Retroviral oncogenesis

Retroviral infection is causally associated with a wide range of lymphoid malignancies, of which lymphomas of T cell origin are the most common and best characterised in the domestic cat (2). The malignant potential of such lymphomas can be directly attributed to a transduced oncogene (acute transforming retroviruses) or to proviral integration in the host genome which activates a host cell proto-oncogene (chronic transforming retroviruses). This forms the basis of two loosely categorised divisions of oncogenic retroviruses which have examples in cats, birds and rodents. However, it must be emphasised that retroviruses which are classified as chronic may also fall into the acute transforming category, indeed acute transforming retroviruses are usually formed *de novo* from the chronic type (2). Nevertheless, it is a useful way to separate the two major ways in which retroviruses promote cancer in animal models.

a) Acute transforming retroviruses

Some oncogenic retroviruses can induce tumours in animal models with a relatively short latency period and can transform cells *in vitro* (3-5). Further analysis reveals that these retroviruses have replaced some of their replicative genes with sequences (v-*onc*) which are related to proto-oncogene sequences from the host cell genome (c-*onc*). Such transduction of proto-oncogenes occurs very rarely in nature however, it is relatively prevalent in sporadic lymphomas in cats associated with feline leukaemia virus (FeLV) infection (6). It is this, often

mutated or de-regulated transduced cellular gene, which is responsible for transformation with such short latency. For example, the proto-oncogene c-*myc* is a frequent target for transduction by FeLV in naturally occuring feline T cell leukaemias (7-9) resulting in a replication deficient, recombinant FeLV carrying the *Myc* oncogene (v-*myc*). The acquisition of *Myc* by FeLV induces T cell lymphomas with relatively short latency (8-12 weeks after infection), however it does not expand the spectrum of tumours induced by FeLV alone, possibly due to tissue tropism encoded by FeLV (3).

b) Chronic/slow transforming retroviruses

Retroviruses which do not carry cellular genes induce tumours with a much longer latency, often months or years after the viral challenge. Furthermore, the tumours are mono- or oligoclonal and the viruses cannot transform cells in tissue culture (reviewed in (10)). The tumorigenic potential of this type of retrovirus can be directly attributed to integration of the provirus in the cellular genome. Most studies suggest that retroviral integration within the host genome is essentially random. If the retrovirus integrates within a proto-oncogene in such a way as to activate (or inactivate) its expression, then that cell will have a selective growth advantage and will grow and divide in a clonal fashion. An oncogene or suppressor gene mutated in this way is effectively tagged by the presence of a virus nearby or within the gene. Therefore identifying regions of common proviral integration in independently derived tumours represents a method for cloning novel oncogenes or suppressor genes in the immediate vicinity of the integrated provirus which are involved in tumour development.

Retroviruses have been used for a number of years to identify genes involved in tumorigenesis. The first example was in 1981 when Hayward *et al* found that cellular sequences with homology to v-*myc* were activated in chicken bursal lymphomas by insertion of avian leukosis virus (ALV) (11). Since then greater than 60 genes have been identified as targets for insertional mutagenesis in retrovirus-induced tumors of the mouse, chicken and cat (Table 1.1 and reviewed in (12)). Most of these mutations result in the activation of dominant

Table 1.1 Common insertion sites in retroviral leukaemogenesis.

Gene/locus	*Retrovirus* [a]	*Tumour*	*Mode of activation*[b]	*Function / other information*	*References*
Ahi-1	A-MuLV	pre-B lymphoma	?	Function unknown; linkage to *Myb*[c] (35kb 3` of c-*myb*)	(17;18)
Bmi-1/*Flvi*-2	M-MuLV, FeLV, FeLV-*myc*	pre-B lymphoma T lymphoma	5` and 3` E; 3` P	Transcription factor of *Pc*-G family.	(19-25)
Cb-1/*Fim*3	F-MuLV; M-MuLV	Myeloid leukaemia	LRA of *Evi*1.	Hypermethylation and enhanced expression of *Evi*1.	(26-28)
Dsi-1	M-MuLV	T lymphoma (rat)	?	?	(29)
Evi-1	AKXD-MuLV	Myeloid leukaemia	5` and 3` P; 5` E	Zinc finger transcription factor.	(30;31)
Evi-2 / *Nf*1	BXH2-MuLV	Myeloid leukaemia	Truncation of NF1 tumour suppressor.	GTPase activating protein.	(32;33)
Evi-3	AKXD-MuLV	B lymphoma	?P	Function unknown.	(34)
Evi-5/*Tmi*-1	AKXD-MuLV	T lymphoma	Truncation of 3` coding region; LRA of *Gfi*1.	Homology to cell cycle regulators.	(35-39)
Fis-1/*Cyl*-1	F-MuLV	Myeloid leukaemia T lymphoma	LRA of Cyclin D1.	G_1 phase Cyclin D1.	(40;41)
Fit-1	FeLV, FeLV-*myc*	T lymphoma	LRA of c-*myb*?	Linkage to c-*myb*.	(25;42)
Fli-1 / *Sic*-1	F-MuLV; Cas-Br-E	Erythroleukaemia; Non-T/B lymphoma	5`E; 3`P	*ets* related transcription factor.	(43-46)
Fli-2 / *Nfe*2	F-MuLV	Erythroleukaemia	Inactivation and loss of wt allele.	Transcription factor (NF-E2).	(47;48)
Flvi-1	FeLV	Splenic lymphomas	?	?	(49;50)
Frat-1	M-MuLV	Progressed T lymphoma	Truncation at 3` UTR; E.	Unknown function in tumour progression.	(51)
*Fre*2	F-MuLV	Erythroleukaemia	Chromosomal translocation.	Unknown function.	(52)
Gfi-1	M-MuLV	T lymphoma (rat)	P	Zinc finger repressor.	(38;53;54)
Gfi-2 / IL-9R	M-MuLV	T lymphoma (rat)	Truncation at 3` UTR	Growth factor receptor.	(55)
Gin-1	G-MuLV	T lymphoma	?	?	(56)
Lck	M-MuLV	T lymphoma	5` and 3` P	*src*-related tyrosine kinase.	(57;58)
Mis-2	M-MuLV	T lymphoma (rat)	?	Linkage to c-*myb* (160kb 3` of c-*myb*)[c]	(59)
Mlvi-1 / *Pvt*-1 / *Mis*-1	M-MuLV	T lymphoma (rat)	E	*cis* activation of c-*myc* and *Mlvi*-1 (*Mlvi*-1 270kb 3` of c-*myc*).	(60-64)
Mlvi-4	M-MuLV	T lymphoma (rat)	5` P (*Mlvi*-4); E (c-*myc* and *Mlvi*-1).	*cis* activation of c-*myc*, *Mlvi*-1 and *Mlvi*-4 (*Mlvi*-4 30kb 3` of c-*myc*).	(63-65)

Gene/locus	*Retrovirus* [a]	*Tumour*	*Mode of activation*[b]	*Function / other information*	*References*
*Mml*1	Ampho 4070A	myeloid leukaemia	?	Linkage to c-*myb* (25kb 5` of c-*myb*) [c].	(66;67)
c-*myb*	M-MuLV, F-MuLV, ALV	Myeloid leukaemia B lymphoma	Truncation of 5` and 3` coding sequence.	Transcription factor.	(68-72)
c-*myc*	M-MuLV, F-MuLV, FeLV, ALV	T & B lymphoma erythroleukaemia	3` P; 5` and 3` E.	Transcription factor.	(7;11;73-75)
N-*myc*	M-MuLV	T lymphoma	Truncation of 3`UTR; 3`E.	Transcription factor.	(76;77)
p53	F-MuLV	Erythroleukaemia	Inactivation of one or both alleles.	Transcription factor; tumour suppressor.	(75;78)
Pal-1	Eμ-*myc* + M-MuLV	pre-B lymphoma	LRA of *Gfi*-1.	Unknown function; Linked to *Gfi*-1 / *Evi*-5.	(20;38;39)
Pim-1	M-MuLV, F-MuLV; Eμ-*myc* +M-MuLV; FeLV;FeLV-*myc*.	T and B lymphoma; Erythroleukaemia	3` E; truncation of 3` UTR.	Serine/threonine protein kinase.	(20;25;75; 79; 80)
Pim-2	Eμ-*myc*/Pim-1$^{-/-}$ +MMuLV	T and B lymphoma	Mainly 5` E	Serine/threonine protein kinase with homology to *Pim*-1.	(14;81;82)
Pim-3	Eμ-*myc*/Pim-1/2$^{-/-}$ +MMuLV	T and B lymphoma	?	Serine/threonine protein kinase with homology to *Pim*-1.	(14)
Prlr	M-MuLV	T lymphoma (rat)	5` P	Prolactin receptor	(83)
Tblvi-1	TBLV	T lymphoma	?E	Unknown function.	(84)
Tiam-1	M-MuLV	T lymphoma line (invasive)	Mainly truncation of 5` and 3` coding sequence.	Homology to GDP-GTP exchanger.	(85)
Tic-1 (formerly *Pim*-2)	M-MuLV	Progressed T lymphoma	?	?	(86;87)
Til-1 / CBFA1	M-MuLV & CD2-*myc*	T lymphoma	5` E of a variant promoter	Transcription factor.	(88;89)
Tpl-1 / *Ets*-1	M-MuLV	Progressed T lymphoma (rat)	?E and P of *Ets*-1	Genetically linked and homologous to the *Ets*-1 transcription factor.	(90)
Tpl-2 / *Cot*	M-MuLV; Eμ-*myc*/*Pim*-1/2$^{-/-}$ + M-MuLV	Progressed T lymphoma (rat); B and T lymphoma	Truncation of 3` coding sequence; 3`E and mRNA stabilisation.	MAP kinase kinase kinase.	(14;91-93)
Vin-1/CyclinD2	M-MuLV; RadLV	T lymphoma	5` E	G_1 to S-phase cell-cycle protein	(94;95)

Table 1.1: key to table

[a] Retroviral mutagens: M-MuLV, murine leukaemia virus;, F-MuLV, Friend murine leukaemia virus; A-MuLV, Abelson murine leukaemia virus; G-MuLV, gross passage A murine leukaemia virus; FeLV, feline leukaemia virus; FeLV-*myc*, feline leukaemia virus which carries a transduce *myc* gene; Ampho 4070A, amphotrophic retrovirus 4070A; TBLV, type B leukaemogenic retrovirus; ALV, avian leukosis virus; RadLV, radiation leukaemia virus.

[b] Predominant activating mechanism: 5`P, promoter insertion where the 5` LTR is utilised; 3`P, promoter insertion where the 3` LTR is utilised; E, enhancer insertion; 5`E, enhancer insertion where the provirus is situated upstream and in the opposite transcriptional orientation to the gene; 3`E, enhancer insertion where the provirus is situated downstream and in the same transcriptional orientation to the gene; LRA, long range *cis* activation; UTR, untranslated region; wt, wild type.

[c] No evidence for long range activation of c-*myb*.

oncogene(s) which act in growth signalling pathways, thus leading to inappropriate survival of these cells. The set of genes identified in this way overlaps significantly with those subject to retroviral transduction and non-viral mutational processes such as chromosomal translocations and gene amplification, illustrating the value of retroviruses as tags for the genes which are central to neoplastic cell growth. The utility of this approach has been demonstrated further by the identification of collaborating genes in oncogene transgenic mice and in elucidating functionally related genes in knockout mice (13;14).

1.1.1 The structure and life-cycle of FeLV

Feline leukaemia virus (FeLV) is a member of the oncovirinae subfamily of exogenous retroviruses which are horizontally transmitted among the domestic cat population (15). FeLV infection has been causally associated with a number of lymphoid malignancies as well as degenerative diseases of the domestic cat (16). There are three FeLV subgroups: A, B and C. FeLVs of subgroup A are ecotropic and found in all infected cats, while FeLV B and C are amphotropic and are found in 50% and 1% of infected cats respectively. Most oncoviruses are type C and include avian leukosis virus (ALV), murine leukaemia virus (MuLV) as well as FeLV, around which this study is centred. However, an exception to this is the mouse mammary tumour virus (MMTV) which is a type B retrovirus that causes carcinomas of mammary epithelial tissue. In MMTV induced carcinomas, the *int* genes or oncogenes that belong to the *Wnt* family or the *Fgf* family of growth factors are frequent targets of proviral insertions. However, MMTV induced mammary carcinomas will not be discussed further.

Type C retroviruses share similar structures and mode of replication (15). The structural gene order of FeLV is *gag-pol-env* (Figure 1.1A). *gag* and *env* encode the major structural proteins which form the internal non-glycosylated components and the glycosylated virion surface glycoproteins respectively (96). The *pol* region is responsible for the protease, integrase and reverse transcriptase products. The FeLV genome, like other retroviruses, also consists of 5` and 3` long terminal repeat (LTR) sequences containing U3, U5 and R regions which are

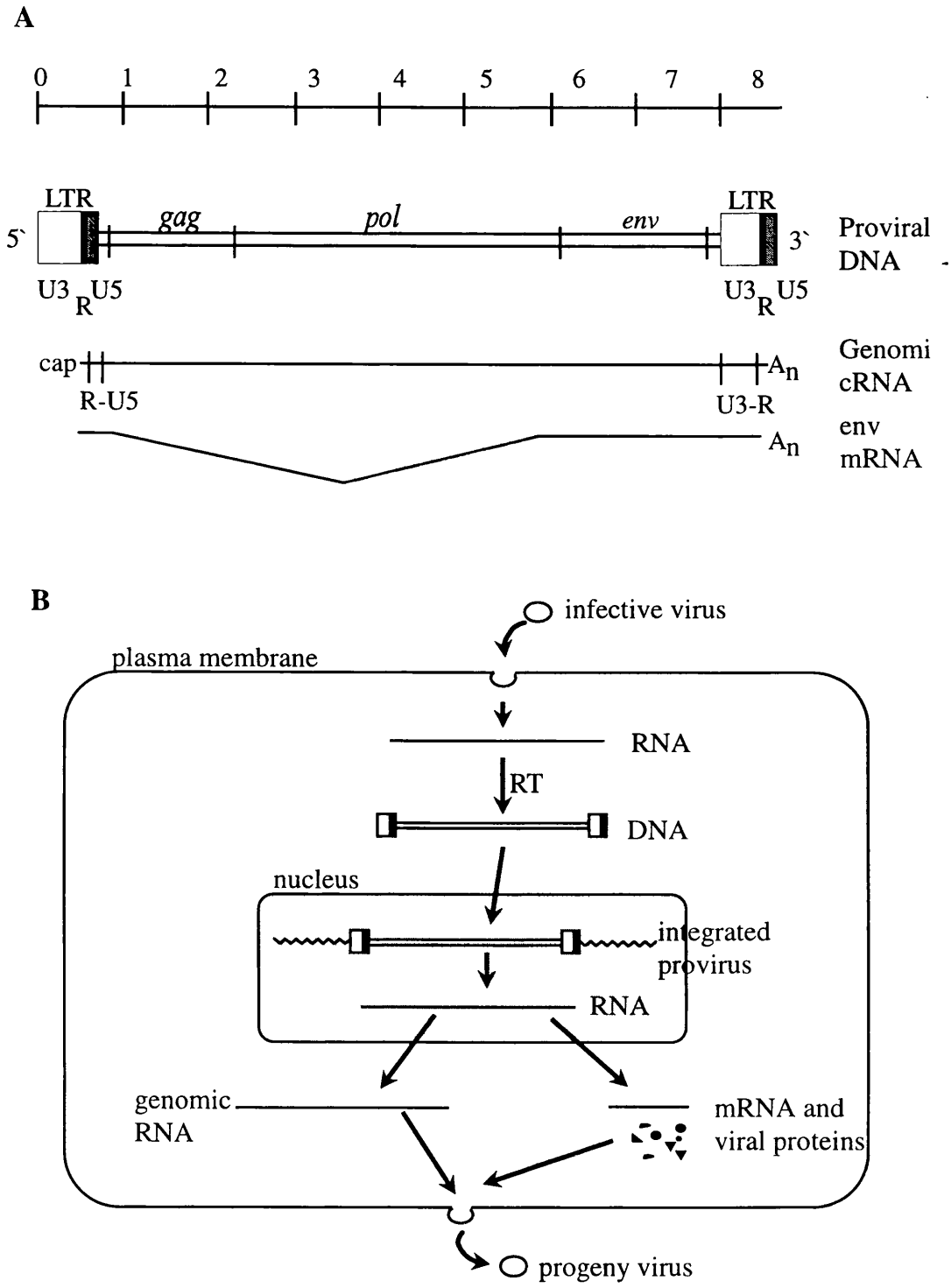


Figure 1.1 Genetic structure and life-cycle of FeLV. (A) The FeLV genome structure and expression. (B) Schematic overview of the FeLV life-cycle. RT, reverse transcriptase; LTR, long terminal repeat; An, polyadenylation; cap, the capped nucleotide at the 5` end of the viral RNA.

copied to form the proviral LTR. The LTRs contain sequence responsible for the initiation of transcription as well as for polyadenylation of viral RNA. The U3 region of the LTR contains enhancer elements which promote viral (and cellular) transcription.

Retroviruses can exert their mutagenic effect during a natural step in their life-cycle as the provirus integrates into the host genome where it may disrupt or deregulate a cellular oncogene involved in growth or development. The first stage of the life-cycle (Figure 1.1B) involves the synthesis of double stranded DNA and the integration of the provirus in the host cell chromosome. This is accomplished without any help from the host cell and involves the virally encoded enzymes, reverse transcriptase and integrase. However, once integrated the provirus relies on the host cell machinery for transcription and processing of RNA. Finally, new virus particles are assembled and released by budding through the host cell membrane resulting in an envelope surrounding the viral core which is composed of host cell membrane enriched with viral glycoproteins.

1.1.2 Mechanisms of retroviral insertional mutagenesis

Viral integration within or adjacent to cellular proto-oncogenes can alter the structure or level of expression of such genes, thereby conferring a selective growth advantage to the target cell and its progeny. The ability of the integrated provirus to activate the transforming capacity of the adjacent gene is an inherent property of the viral LTR which contains the powerful viral transcriptional control sequences. As retroviral integration shows little sequence specificity, the identification of preferred integration sites in more than one independently derived virus-induced cancer has generally been interpreted as evidence of a selective advantage arising from these insertions and the presence of a nearby target gene. That cell will then grow and divide in a clonal fashion contributing to the outgrowth of a tumour. Thus, an oncogene activated by the virus is effectively tagged by the presence of the virus nearby or within the gene.

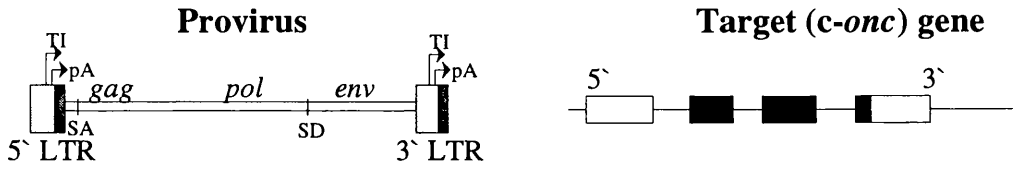
This property has been used to detect both novel and previously identified oncogenes such as v-*myc*, v-*myb*, v-*mos* and v-*ras* which were identified as components of acutely transforming viruses (10;12). Screening both naturally occuring and experimentally derived tumours for the presence of common sites of retroviral integrations is the first step in the identification of novel common insertion loci.

The major mechanisms by which a proto-oncogene is activated by the insertion of a retrovirus within or nearby the coding sequence are described below (and reviewed in (97)). It is important to point out, however, that each mechanism is not mutually exclusive and more than one mechanism of activation may occur at any one time.

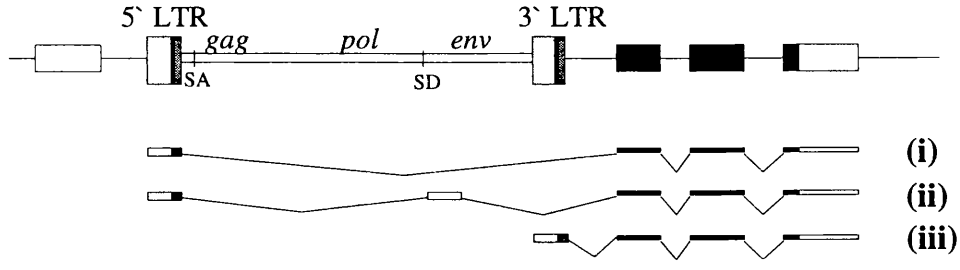
1.1.2.1 Promoter insertion

Tumours which have proviral insertions at the 5` end of a gene in the same transcriptional orientation are predominantly activated by LTR promotion. This generally takes two forms: 5` and 3` promotion. In 5` promotion, transcription initiates at the relatively strong 5` promoter of the proviral LTR rather than the normal cellular promoter (Figure 1.2A i and ii). This results in abnormally high levels of the hybrid transcript and often involves cryptic splicing of the virus to a splice acceptor site in the adjacent proto-oncogene. 5` promoter insertion can be exemplified by the classic case of avian B cell lymphoma where c-*myc* transcription is driven by the viral promoter situated 5` to the gene (11). 3` LTR promoter insertion is similar to 5` promotion except that transcription is initiated at the 3` promoter of the LTR (Figure 1.2A iii). The proviruses in tumours activated in this way often suffer large deletions and aberrant splicing patterns. For example, the *Lck* gene activated by 5` and 3` promoter insertion in Moloney murine leukaemia virus (M-MuLV) induced T cell lymphomas in rats, involved aberrant splicing which removed 5` regulatory sequences (58).

Figure 1.2 Mechanisms of retroviral insertional mutagenesis

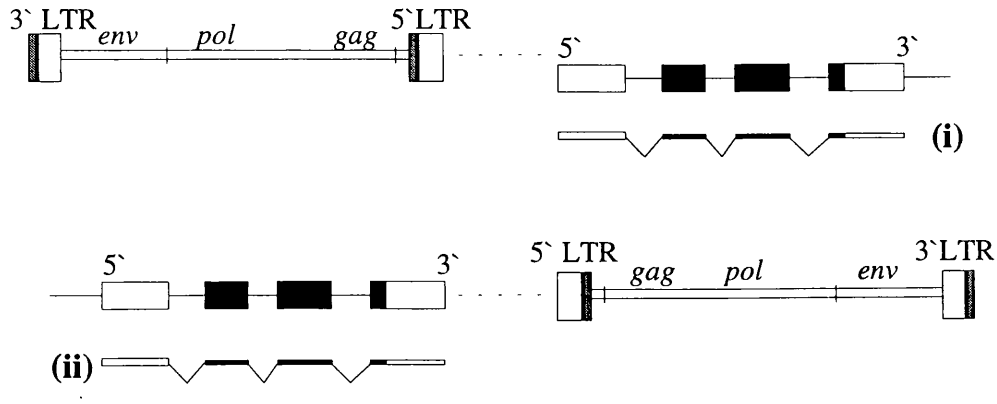


(A) Promoter insertion



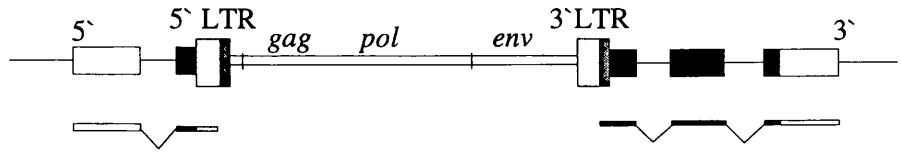
High levels of altered c-*onc* mRNA

(B) Enhancer insertion



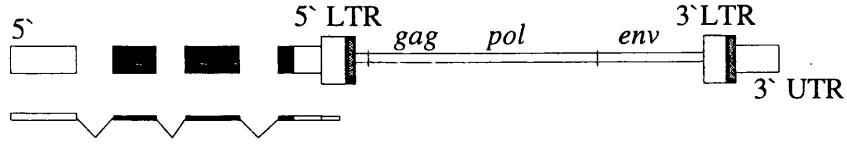
High levels of unaltered c-*onc* mRNA

(C) Truncation of the gene product



High levels of altered c-*onc* mRNA

(D) Effects on RNA stability



High levels of unaltered c-*onc* mRNA

Figure 1.2 Mechanisms of retroviral insertional mutagenesis. The examples shown illustrate the mechanisms involved in the insertional activation of proto-oncogenes by retroviral insertion, however more complex retroviral deletions and splicing events are not shown. Shown at the top are diagrams of the structural features of a provirus, and a proto-oncogene (c-*onc*) which consists of 3 coding exons as well as 5` and 3` untranslated, regulatory regions. See text for details of activation for each of the mechanisms illustrated.
(A) Activation by promoter insertion. (i) transcription initiated from the 5` LTR promoter; (ii) transcription initiated from the 5` LTR promoter and splicing involving a cryptic splice donor in the proviral genome. (B) Activation by enhancer insertion. The broken line in this figure indicates that the provirus may be situated close to the proto-oncogene, or far away to exert enhancment of the gene. (i) provirus situated upstream and in the opposite transcriptional orientation to c-*onc*; (ii) provirus situated downstream and in the same transcriptional orientation to c-*onc*. (C) Truncation of the gene product by proviral insertion within coding sequence, also often involves enhancement of the truncated product, or results in the inactivation of the gene. (D) Stabilisation of an unaltered c-*onc* by removal of RNA stabilising motifs in the 3` UTR.
Abbreviations and key: LTR, long terminal repeat; TI, transcription initiation; SD, splice donor; SA, splice acceptor; pA, polyadenylation signal; UTR, untranslated region. ■ protein coding DNA; □ untranslated DNA; ═ mRNA; ⋁ splicing.

1.1.2.2 Enhancer insertion

Eukaryotic enhancer elements can activate transcription of a linked gene by a *cis* acting mechanism. Their effect on transcription is independent of orientation and they preferentially stimulate transcription from the most proximal promoter but have the ability to function over long distances (reviewed in (98)). The U3 domains of viral LTRs contain powerful enhancer elements which act as targets for cellular transcription factors. The enhancer elements are responsible for the tissue specificity seen with some retrovirally induced cancers, presumably due to their interaction with tissue specific factors. FeLV proviruses isolated from naturally occuring feline thymic lymphomas revealed a high degree of heterogeneity in their LTR enhancer sequences, with up to three repeats of the LTR enhancer in one tumour (99;100). In situations where enhancer duplications were observed, there was a concomitant increase in functional binding sites for transcription factors.

Thus, viral enhancers play an important role in FeLV induced leukaemogenesis (2) where they activate transcription by *cis* acting mechanisms from the promoter of a nearby cellular proto-oncogene, often over-riding normal cellular transcriptional control. Enhancer insertional activation is characterised by the integration of a provirus either at the 5` end of a proto-oncogene in the opposite transcriptional orientation, or at the 3` end in the same transcriptional orientation (Figure 1.2B i and ii). The net result of such insertions in the tumour is the upregulation of an otherwise normal oncogene.

Enhancer insertion is by far the most common mode of retroviral activation, possibly because of the flexibility of position and orientation of provirus to produce an oncogenic effect. Some examples of genes activated by enhancer insertion are listed in Table 1.1.

1.1.2.3 Long distance effects

Some proviruses appear to activate genes at distances of up to 300kb away. This again is due to the proviral LTR enhancer which, as eluded to above, can exert

their effect on promoters located long distances from the provirus (Figure 1.2B). Evidence in favour of this mechanism comes from the activation of the c-*myc* and *Evi*-1 genes by proviruses situated at great distances from these genes. Insertions at the *Mlvi*-1 and *Cb*-1/*Fim*-3 loci activate transcription from the normal promoters of c-*myc* and *Evi*-1 respectively, even though the proviruses are located up to 300kb away (28;64). This demonstrates why it can, in some cases, be very difficult to locate a transcription unit at a common proviral insertion site.

1.1.2.4 Truncation of the gene product

If a provirus integrates within the coding sequence of a gene, this can have profound effects on the resulting protein, inactivating or mutating it in such a way as to produce an oncogenic protein (Figure 1.2C). For example, c-*myb* has frequently been found as the target of proviral integration in myeloid leukaemias of pristane primed BALB/c mice infected with M-MuLV (69;101;102). Proviral insertions were found at both the 5` and 3` ends of the *Myb* gene, resulting in amino- or carboxy-terminal truncations respectively. The resultant truncated proteins are oncogenic most likely because both termini of the gene negatively regulate Myb activity. Therefore the removal of these regulatory elements will result in the inappropriate expression of an oncogenic Myb protein. Furthermore, enhancer or promoter insertion may also play a role to increase the levels of the truncated, oncogenic form of the protein.

1.1.2.5 Effects on RNA stability

When proviral integration occurs within the 3` untranslated region (UTR) of a proto-oncogene in the same transcriptional orientation, transcription is prematurely terminated at the polyadenylation signal in the 5` LTR (Figure 1.2D). The net result of integrations of this type 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. This frequently occurs in concert with enhancer insertion to up-regulate transcription of the oncogene. *Frat*-1 is activated by stabilisation of mRNA together with enhancer insertion in M-MuLV induced transplanted lymphomas of

Pim-1 or *Myc* transgenic mice (51). Similarly, N-*myc* has been shown to be activated by proviral insertions in the 3` untranslated region of the gene in MuLV induced T cell lymphomas (77).

1.1.2.6 Inactivation of a gene

Inactivation of a gene by proviral integration is a relatively rare event because it requires insertion at both alleles of a gene. However, it has occurred on at least three occasions, resulting in the inactivation of p53, *Evi*-2/*Nf*1 or *Fli*-2. Friend murine leukaemia virus (F-MuLV) induced erythroleukaemia was shown to have integrated proviruses in both alleles of the tumour suppressor gene p53 resulting in no detectable expression of p53 protein (78). An alternative mode for the inactivation of a gene is by proviral insertion in one allele and loss of the other allele by, for example deletion or point mutation. A common insertion locus in myeloid leukaemia of BXH-2 mice, *Evi*-2, has been shown to have provirus inserted in both allelles (33). This results in the truncation and inactivation of the NF1 tumour suppressor protein, a GTPase activating protein. *Fli*-2, a common integration site in F-MuLV erythroleukaemias, also has a provirus integrated in one allele of the NF-E2 transcription factor and this is accompanied by deletion of the other allele resulting in the inactivation of NF-E2 in this erythroleukaemia cell line (48).

1.1.3 Transduction

The transduction of cellular genes by retroviruses is a rare event in nature, possibly due to the number of complicated recombination events which must take place in order to generate such a hybrid. However, transduction of the c-*myc* proto-oncogene by FeLV (FeLV-*myc*) has been found in a number of naturally occurring thymic lymphomas (7-9). In all *myc*-transducing FeLV isolates the entire coding sequence of c-*myc* from exons 2 and 3 have been acquired by the hybrid virus (2;6) suggesting that the N- and C-termini of v-*myc* are crucial for oncogenesis by FeLV-*myc*. The potency of the truncated *Myc* allele may also be enhanced as the v-*myc* gene is placed outwith normal, cell specific controls and is expressed under the control of viral signals.

Figure 1.3 shows a possible mechanism for the production of an FeLV-*myc* recombinant virus based on the model proposed by Swanstrom *et al* (103) for the genesis of Rous sarcoma virus which has transduced a cellular gene (c-*src*). The first step involves the integration of a provirus upstream and in the same transcriptional orientation as c-*myc*. A recombination event between the retrovirus and c-*myc* followed by a deletion fuses the viral and c-*myc* coding sequences. Following transcription and splicing, the hybrid transcript is co-packaged with helper virus involving the 5` LTR. On subsequent infection a second recombination event generates a retrovirus with a 3` viral LTR resulting in a novel hybrid FeLV genome which includes the transduced portion of c-*myc*. This model is supported by the fact that the v-*onc* gene has been processed to remove intron sequences.

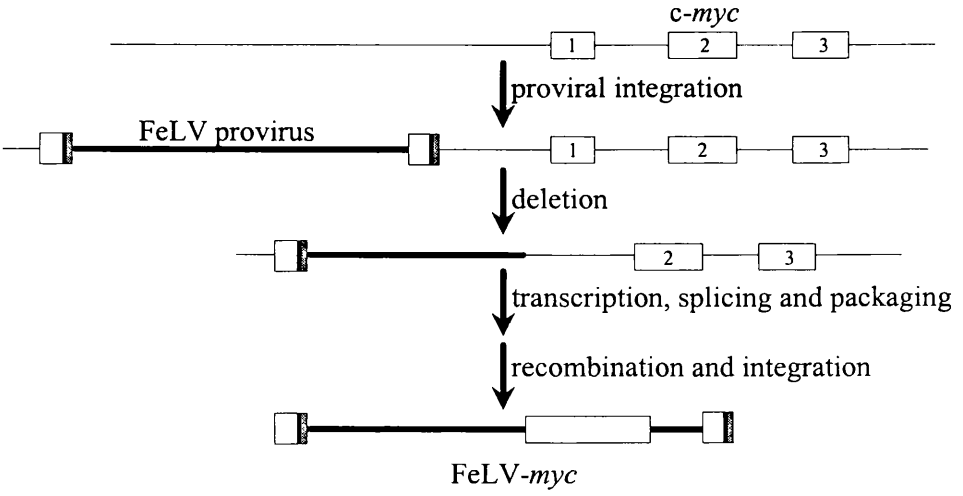


Figure 1.3 Model for the generation of FeLV-*myc* recombinant virus. The proposed mechanism is based on the model of Swanstrom *et al* (103).

1.1.4 Concerted activation in retroviral leukaemogenesis

Cancer is a multi-step process of genetic and epigenetic events where tumours require the accumulation of multiple mutations before the fully malignant

phenotype can develop. Although the activation of a cellular oncogene by, for example proviral integration, is a crucial step in this process, secondary changes are required for progression through multi-step carcinogenesis. A second proviral insertion could be one such event which offers a further selective advantage to the initiating tumour cell.

Thus, there are several lines of direct evidence that proviral activation of two oncogenes can occur in the same tumour cell with synergistic effects on tumour development. For example, both c-*myc* and *Pim*-1 have been shown to be activated by proviral insertion in the same tumour cell clone (74;104). In these studies MuLV induced T cell lymphomas in mice resulted in the activation of one or both of these genes in tumours which developed after a short latency. Although some were oligoclonal, at least one of the tumours was found to be monoclonal on transplantation into a syngenic mouse (74). Together this suggests the presence of a modified c-*myc* and *Pim*-1 allele by proviral insertion within the same cell population and that this is associated with tumour onset with relatively short latency. This was one of the first illustrations of synergy between c-*myc* and *Pim*-1 in T cell lymphomagenesis.

1.2 Identification of genes at retroviral insertion sites

The multistep evolution of tumour development can explain why most tumours are monoclonal in origin: only an individual cell which has accumulated the necessary mutations will form the fully malignant phenotype. Those which have not will die or become overwhelmed by the rapid proliferation of the tumour clone. Our understanding of the events involved in this process is constantly growing as an increasing number of genes are identified which are mutated or deregulated in cancer (see (105) for a comprehensive listing of oncogenes and tumour suppressor genes).

The molecular cloning of genes which have been activated by proviral insertion as well as the identification of genes which have been transduced by a retrovirus,

have together been responsible for the isolation of the majority of proto-oncogenes to date (see Table 1.1, and reviewed in (12)). Examples can be drawn from various growth signalling pathways where they generally have regulatory roles in cellular growth and development. Examples include growth factors (*sis* (PDGF)), growth factor receptors (c-*erb*B/(EGF-R), *Gfi*-2/IL-9R), cytoplasmic protein kinases (*Pim*-1, *Pim*-2, *Tpl*-2), GTPases (*ras*), GDP exchangers (*Tiam*-1), nuclear transcription factors (c-*myc*, c-*myb*, *Bmi*-1, *Gfi*-1, *Fli*-2/*Nfe*2) and cell cycle regulators (*Fis*-1/Cyclin D1, *Vin*-1/Cyclin D2). Mutations which alter the expression of one or more of these genes may lead to deregulated growth and development. The elucidation of the pathways involved in this have been central to cancer research and the more that is known about the various genes involved in this process the greater our understanding will be.

Following the discovery of a common insertion site, the arduous task of identifying the gene which is affected by proviral integration begins. Table 1.2 provides an overview of a selection of common leukaemia virus insertion loci at which the relevant genes have been identified, a selection of which are discussed in more detail below. Following the identification of a transcript (most commonly by northern blot analysis) conventional cDNA cloning is invariably used to characterise the gene. However, a number of common insertion loci remain uncharacterised with respect to the genes involved (see Table 1.1), including *Bla*-1, *Dsi*-1, *Flvi*-1, *Gin*-1 and *Tic*-1.

Table 1.2 Identification of genes at type C retroviral insertion sites

Locus	*Method*	*Other clues*	*References*
Ahi-1	Exon trapping	CpG island	(17;18)
Bmi-1/*Flvi*-2	Northern	CpG island	(19;20;23)
Cb-1/ *Fim*-3	Linkage to *Evi*-1[a]		(26-28)
Evi-1	Northern		(30;31)
Evi-2/*Nf*-1[b]	Northern	Zooblot	(32;33)
Evi-3	Northern	CpG; Zooblot	(34)
Evi-5	Linkage to *Gfi*-1[a]	Northern	(35-39)
Fis-1/*Cyl*-1	Linkage to CyclinD1[a]	CpG island	(40;41)
Fit-1	Linkage to c-*myb* [a, b]		(25;42)
Fli-1/*Sic*-1	Northern		(43-45)
Fli-2/*Nfe*-2 [c]	Northern		(47;48)
Frat-1	Northern	Zooblot	(51)
Gfi-1	Northern		(53)
Gfi-2/IL-9R	Northern		(55)
Mis-2	Linkage to c-*myb*		(18)
Mlvi-1	Linkage to c-*myc* [a]	Northern; Zooblot	(60;63;64)
Mlvi-4	Linkage to c-*myc* and *Mlvi*-1[a]	Northern	(63-65)
Pal-1/*Evi*-5 / *Eis*-1	Linkage to *Gfi*-1[a]		(20;37-39)
Pim-1	Northern		(25;79)
Tblvi-1	Northern	Zooblot	(84)
Tiam-1	Northern	Zooblot	(85)
Til-1/CBFA1	Northern	Zooblot	(88;89)
Tpl-1	Northern		(90)
Tpl-2/*Cot*	Northern		(91;92)
Vin-1/*CCND2*	Northern	CpG island	(94;95)

[a] Evidence for long range activation.
[b] This thesis
[c] Insertional inactivation.

1.2.1 Methods for the identification of genes at retroviral insertion sites

Retrovirally induced leukaemogenesis is characterised by the outgrowth of a clonal cell population carrying retroviral insertions at distinct regions of the chromosome. The term 'common insertion site' refers to a region of the chromosome which carries a proviral insertion in more than one independently derived tumour. Generally, a common insertion site would be identified by isolating a single copy DNA probe from the flanking sequence which would then be used on Southern blot analysis of DNAs from various tumour and control samples. A region of chromosomal DNA which is frequently targeted by proviral

integration would be identified by the presence of a rearranged band in addition to the genomic band in more than one independently derived tumour sample.

The majority of common insertion sites are then characterised by northern blot analysis using unique sequence probes from the DNA flanking the provirus cluster. The techniques involved in this are discussed in more detail in section 3.1.3 and Figure 3.3. Usually a tumour specific transcript pattern or over-expression of the normal transcript is detected in tumours with proviral insertion.

Other methods which have been used in the identification of genes affected by proviral integration include 1) Chromosomal localisation where the common insertion locus is linked to known genes, some of which show evidence of long range activation (Table 1.2 and section 1.2.6). 2) Exon trapping (106) was developed as a strategy whereby fragments of cloned genomic DNA which have been spliced during cycling of the vector in the host cell may be recovered as cDNA. This technique has been successfully used to identify a gene at the *Ahi*-1 common site of Abelson-MuLV integration in pre-B lymphoma (P. Jolicoeur, Clinical Research Institute of Montreal, Canada, personal communication). 3) Zooblot analysis to identify DNA which is evolutionarily conserved using unique sequence probes from the insertion cluster which hybridise to DNA from various species immobilised on a Southern blot (a so-called zooblot). A unique sequence probe which is positive on zooblot analysis suggests conservation of this sequence during evolution which may be due to the presence of coding sequence.

4) CpG islands often serve as landmarks in the search for transcription units (107). CpG islands are clusters of non-methylated G-C dinucleotides which are relatively rare in eukaryotic DNA, with only a fifth of the number expected. This composition can be accounted for by the tendancy of methylated cytosines to decay to TpG and CpA while non-methylated cytosines do not (107;108). CpG islands span an average of 1-2kb and are often associated with the 5` ends of genes, particularly house-keeping genes (108;109). They often extend beyond the promoter region and into the first or even second exon. Therefore, they

represent ideal tags to identify transcripts. Restriction enzymes have been used to detect potential gene sequences by detecting clusters of sites for certain rare-cutting restriction enzymes which specifically recognise non-methylated, C-G containing sequences (110). Thus, by generating a long-range restriction map with such restriction enzymes using pulsed-field gel electrophoresis (PFGE), a number of proto-oncogenes have been detected by association with CpG islands (see Table 1.2).

Discussed below is an overview of a selection of genes identified at common insertion sites, principally by northern blot analysis with DNA probes from the locus. Other evidence for the presence of a gene at these loci will also be discussed.

1.2.2 The *Pim* family

The *Pim*-1 gene, which was first identified as a common proviral integration site in M-MuLV induced murine T cell lymphomas (79), encodes a serine/threonine protein kinase (80). The gene affected by M-MuLV insertion at *Pim*-1 was identified by northern blot analysis with probes derived from DNA flanking the insertion cluster. Proviral insertions were 3` and in the same transcriptional orientation as the gene, therefore enhancer insertion was proposed as the mechanism of activation of *Pim*-1. *Pim*-1 was also identified as the target of FeLV integration in feline T cell lymphomas where rearrangements were observed in both spontaneously and experimentally induced lymphomas with FeLV or FeLV-*myc* isolates (25).

Pim-1 knockout mice exhibit very subtle phenotypes (14;81) implying functional redundancy between *Pim*-1 and other genes (discussed further in section 1.3.3). Indeed two additional *Pim* family members were identified in this system which may be able to compensate for the loss of *Pim*-1. *Pim*-2, not to be confused with *Tic*-1 which was formerly known as *Pim*-2 (86), was identified in *Pim*-1 null mutant mice using degenerate PCR primers (81). Furthermore, *Pim*-2 was a frequent site of proviral integration in tumours induced by M-MuLV infection in

these mice. cDNA cloning revealed that *Pim*-2 is 53% identical to *Pim*-1 at the amino acid level and northern blot analysis demonstrated that the normal gene is over-expressed by enhancer insertion in lymphomas with proviral insertions at *Pim*-2 (81). The oncogenic potential of *Pim*-2 was demonstrated when an activated copy of this gene was introduced into the germline of mice (82). In this study, *Pim*-2 transgenic mice were predisposed to T-cell lymphomas similar to those of *Pim*-1 transgenic mice. Furthermore, these mice exhibited a strong synergy with an Eμ-*myc* transgene in pre-B cell leukaemia with a latency of 3 to 4 weeks.

In an analogous system, another *Pim* family member, *Pim*-3, was detected as a common M-MuLV insertion site in tumours from *Pim*-1/2 knockout mice again in conjunction with a degenerate PCR approach (14).

1.2.3 *Bmi*-1/*Flvi*-2

Bmi-1 was identified as a target for retroviral insertion in Eμ-*myc* transgenic mice infected with M-MuLV (19;20) and in FeLV-*myc* induced feline lymphomas (23) resulting in B and T cell lymphomas respectively. The coding potential of this common insertion locus was identified primarily by screening for CG-rich sequences using rare cutting restriction endonucleases (20). Probes flanking the CpG island were then used on northern blot analysis to identify a transcript which was over-expressed in tumours carrying insertions near *Bmi*-1. The enhanced expression was due to either enhancer insertional activation or promoter insertion where the relatively strong 3` LTR drove transcription leaving the coding domain intact (19;20;24). The oncogenic potential of *Bmi*-1 was formally established in bitransgenic Eμ-*Bmi*-1/*myc* mice (111), substantiating the synergy between *Bmi*-1 and *Myc*. However, this relationship may not be exclusive (1;25) implicating activation of *Bmi*-1 as an early event in the development of lymphoma.

Bmi-1 is highly conserved in evolution and encodes a protein with homology to a family of nuclear transcription factors which contain a cysteine-rich zinc finger binding domain, termed RING finger (19;20). The *Drosophila* homologue of

Bmi-1 was subsequently found to share homology with the *polycomb* group (*Pc*-G) of transcription factors which are responsible for maintaining the segment specific repression of homeotic genes, such as *Posterior Sex Combs*, during *Drosophila* development (21;22).

1.2.4 The *Gfi*-1 locus

The *Gfi*-1common insertion locus was identified in a rat T lymphoma interleukin 2 (IL-2) dependent cell line where it is involved in the progression to IL-2 independence (53). The gene affected by proviral integration was identified by northern blot analysis and was found to be highly expressed in cell lines carrying insertions at *Gfi*-1, primarily owing to transcription initiating from the proviral LTR. However, additional tumour cell lines with no *Gfi*-1 rearrangement also exhibited high levels of *Gfi*-1 transcript. Subsequent studies have shown that proviral integrations at *Gfi*-1 and its surrounding loci, *Pal*-1/*Evi*-5/*Eis*-1, all appear to upregulate *Gfi*-1 transcription suggesting that *Gfi*-1 is the target for insertions at these loci (37-39). Insertions at these loci which are located within a 50kb domain, are thought to activate *Gfi*-1 by a long range enhancer insertion mechanism operating over distances of up to 25kb.

1.2.5 The *Evi* genes

A number of common ecotropic MuLV insertion sites have been identified in tumours in AKXD or BXH-2 recombinant inbred mice: *Evi*-1, 2, 3 and 5 (30;32;34;35).

Evi-1 was identified as a common site of integration in AKXD-23 mice which developed myeloid tumours at a high frequency (30). Subsequently, integration at *Evi*-1 was shown to activate a novel gene with homology to the zinc finger family of transcription factors in IL-3 dependent myeloid leukaemia cell lines which have failed to differentiate (31).

Upregulation of *Evi*-1 transcription was shown to be due to promoter insertional activation where transcription initiates in the 5` or 3` LTR. However,

integrations at the *Cb*-1/*Fim*-3 common F-MuLV insertion locus 90kb from *Evi*-1, were also shown to activate transcription of *Evi*-1 (26-28). Evidence that integration at *Cb*-1/*Fim*-3 activates *Evi*-1 transcription in *cis* is several-fold. Firstly, no transcript sequences have been identified at the *Cb*-1/*Fim*-3 locus. Secondly, the proviral insertions at *Fim*-3 are all orientated 5` and in the opposite transcriptional orientation to *Evi*-1 suggesting an enhancer insertional activation mechanism. Indeed, enhancers have been shown to be active from distances of up to 300kb (41;64). Furthermore, in cells which are rearranged at *Evi*-1, one allele of the *Evi*-1 gene was shown to be hypermethylated while none of the myeloid cell lines lacking *Evi*-1 rearrangements showed hypermethylation. However, *hyper*methylation is normally associated with gene *inactivity* (112). Bartholemew *et al* (28) proposed that hypermethylation of *Evi*-1 might allow the interaction with proteins which specifically bind to methylated DNA to realise their tumorigenic potential.

Transcripts at the *Evi*-2 locus were identified by northern blot analysis using a unique sequence probe which was found to be evolutionarily conserved on zooblot analysis (32). cDNA cloning identified the *Evi*-2 gene as a leucine zipper transmembrane protein which is the murine homologue of the human neurofibromatosis type 1 (*NF*-1) tumour suppressor gene. NF-1 encodes a GTPase activating protein, neurofibromin, which is involved in the development of neurofibrosarcomas and myeloid leukaemia (32). Integration at *Evi*-2 resulted in the production of truncated *Nf*-1 transcripts and no stable, full length neurofibromin (33). Furthermore, the other *Nf*-1 allele was shown to be inactivated in one case by a second proviral insertion at *Evi*-2.

A cluster of proviral insertions at *Evi*-3 were shown to occur within a CpG island suggesting that they may lie within the 5` region of a gene. Further evidence that a gene is located in this region came from zooblot analysis with unique sequence probes which were found to be evolutionarily conserved. Promoter insertion is the most likely mode of activation since proviruses were inserted in the same

transcriptional orientation within a CpG island resulting in the inappropriate expression of *Evi*-3 transcripts during B cell lymphomagenesis (34).

As discussed above, proviral insertion at the *Evi*-5 common insertion site has been shown to activate the *Gfi*-1 oncogene which is located 18kb downstream of *Evi*-5 (35;37-39). However, a recent study has identified a novel transcript by northern blot analysis which is activated by proviral integration within the 3` UTR of the *Evi*-5 gene (36). Zooblot analysis showed that, not only is *Evi*-5 well conserved, but that it is a member of a multigene family consisting of three homologous genes. Although the coding potential of *Evi*-5 has not been directly demonstrated, the *Evi*-5 transcript does show homology to cell cycle regulators thus suggesting a role for *Evi*-5 in cell cycle control, the over-expression of which may be involved in the development of T cell lymphoma.

1.2.6 *Mlvi*-1 and *Mlvi*-4

M-MuLV induced rat T cell lymphomas has led to the identification of two common insertion loci which map close to the c-*myc* proto-oncogene on rat chromosome 7: *Mlvi*-1 which mapped 270kb 3` of c-*myc* and was found to be identical to the *pvt*-1/*Mis*-2 locus (60;61); and *Mlvi*-4 which mapped 30kb 3` of c-*myc* (65). Genes at both loci have been identified which are activated by proviral insertion. At the *Mlvi*-1 locus a tumour specific RNA transcript was identified by northern blot analysis using an evolutionarily conserved probe from immediatedly 5` of the insertion cluster (63). At the *Mlvi*-4 locus a tumour specific transcript was also identified by northern analysis, which was activated by promoter insertion and is transcribed in the same transcriptional orientation as c-*myc* (65).

In addition to activation of *Mlvi*-1 transcripts, provirus insertion at *Mlvi*-1 has also been shown to activate c-*myc* expression 270kb away (63). Similarly, insertion at *Mlvi*-4 activates c-*myc*, *Mlvi*-1 and *Mlvi*-4 expression (63;65). Thus, provirus insertion in the *Mlvi*-4 locus affects the expression of at least three genes within 300kb.

The ability of proviruses at either locus to exert a long range *cis* effect on *Myc* expression was demonstrated by cell hybrid analysis (64). Cell hybrids between two rat T lymphoma cell lines carrying a provirus in either *Mlvi*-1 or *Mlvi*-4 and a murine T lymphoma line demonstrate that enhanced expression of rat c-*myc* co-segregated with the rearranged *Mlvi*-1 or 4 loci whereas expression of murine c-*myc* was unaffected. This provides direct evidence of the ability of viral enhancer elements to exert their effects in *cis* over large distances of up to 300kb and is substantiated by other reports of regulatory elements which can act over long distances in *cis*. For example, a translocation (6;15) in murine plasmacytomas, termed *pvt*-1, which joins c-*myc* to the immunoglobulin κ locus results in the over-expression of c-*myc* over a distance of at least 72kb (113). This translocation is functionally equivalent to M-MuLV integration at *Mlvi*-1, illustrating the value of this approach in the identification of genes involved in the development of human neoplasia.

1.2.7 Progression loci: *Tiam*-1, *Tpl*-1 and *Tpl*-2

Retroviral tagging is not only useful in the identification of dominant oncogenes, but can also provide insight into subtle functions involved in tumour progression. By a method of proviral tagging in combination with *in vitro* selection for invasiveness, *Tiam*-1 was identified in T cell lymphoma (85). Invasive variants of a T cell lymphoma line were generated by MuLV infection and selection for their ability to infiltrate monolayers of fibroblasts. An evolutionarily conserved probe identified the *Tiam*-1 transcript which was activated by proviral insertions resulting in both N- and C-terminal truncated proteins. The predicted *Tiam*-1 protein shows homology with GDP-GTP exchangers. Transfection studies of non-invasive cells showed that the invasive phenotype was induced with both N- and C-terminally truncated *Tiam*-1 fragments suggesting a direct role for *Tiam*-1 in invasion during T cell lymphomagenesis.

Two additional progression loci have been identified by prolonged culture of cells derived from M-MuLV induced rat T cell lymphomas: *Tpl*-1 (90) and *Tpl*-2

(92). Genes have been identified at both these common insertion loci by northern blot analysis where they were found to encode respectively an *ets* transcription factor and a *cot* serine/threonine protein kinase involved in the MAP kinase pathway (90-93).

1.2.8 *Myc*

Discussed above is an overview of a selection of oncogenes which have been identified by retroviral insertional mutagenesis (Table 1.2). However, many common insertion loci have been found to correspond to proto-oncogenes which have already been identified by other means. For example, *Myc* has been identified by chromosomal translocation in Burkitt's lymphoma (114), gene amplification (115), as an oncogenic sequence transduced by acute-transforming retroviruses (7-9;116) as well as by retroviral insertional activation by ALV (11), MuLV (73;74) and FeLV (7). Both promoter and enhancer insertional activation have been observed in retrovirally induced T and B cell lymphomas. In general, the enhanced expression of an otherwise normal Myc protein is significant for tumorigenesis, rather than a structural change.

The expression of c-*myc* is a key feature of proliferating cells (117) and the over-expression of *Myc* due to the loss of the normal control predisposes to tumours (118;119). A significant role for *Myc* in feline thymic lymphomagenesis was demonstrated when FeLV carrying a transduced *myc* gene (FeLV-*myc*) was inoculated into newborn kittens. The induction of T cell lymphoma occurred with a relatively short latency, usually within 12-14 weeks, compared to approximately 2 years with standard FeLV strains (3;4). However, these tumours were invariably clonal or oligoclonal in origin (4) suggesting that additional events are required to induce complete transformation. This observation is consistent with both the hypothesis of multistep tumorigenesis (120) and previous studies suggesting that *Myc* alone is insufficient to cause disease (121). A significant emphasis has, therefore, been placed on the identification of genes which could collaborate with *Myc* in tumours induced with FeLV-*myc* (23;25;42;49). This system is analogous to proviral tagging of Eμ-*myc* transgenic

mice to identify collaborating oncogenes discussed in section 1.3 (19;20). It is worth noting, however, that the outbred domestic cat has one major advantage over highly inbred mouse models in that it gives greater insight into natural disease processes.

The v-*myc* gene was first detected in avian **my**elo**c**ytomatosis virus MC29 as a transduced gene which was responsible for the transforming potential of this retrovirus (116). The c-*myc* gene is now known to be a member of a family of proto-oncogenes which also includes N-*myc* (122) and L-*myc* (123). The proteins encoded by the various *Myc* genes are nuclear phosphoproteins with very short half-lives which can bind DNA (reviewed in (124)). Myc contains, in addition to the N-terminal transactivation domain, a number of C terminal motifs reminiscent of a nuclear transcription factor, namely a basic region, a helix-loop-helix motif and leucine zipper domain (Figure 1.4). These regions confer the ability of Myc to bind DNA in both a sequence specific and non-specific manner (125).

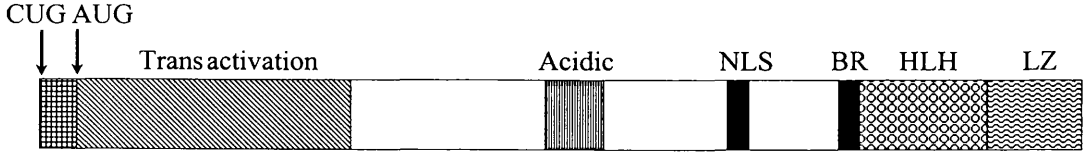


Figure 1.4 Structure of the Myc protein. The major Myc domains are illustrated: transactivation domain; acidic domain, NLS, nuclear localisation domain; BR, basic region; HLH, helix-loop-helix domain; LZ, leucine zipper domain. Also shown are the alternative N-terminal regions which initiate translation from a CUG codon in exon 1 (p67 Myc product) or from a AUG codon in exon 2 (p64 Myc product).

Furthermore, the C-terminal domain is responsible for heterodimerisation with the Myc binding partner, Max (126). Max is a highly stable protein which also

contains a helix-loop-helix motif and leucine zipper domain but which lacks the transactivation domain. Max is capable of forming homodimers but preferentially forms heterodimers with Myc. Moreover, Max binding has been shown to be required for efficient DNA binding and transformation by Myc (127;128). Therefore, the transcriptionally inactive, stable Max homodimer which can bind DNA acts as a repressor of Myc in quiescent cells. Max is relieved of its repression by the over-expression of c-*myc* during proliferation by shifting the equilibrium towards the active Myc-Max heterodimers (129). Max can also form heterodimers with its alternative partners, Mad and Mxi-1, and these complexes also serve as antagonists of the active Myc/Max heterodimer by sequestering Max and through competition for DNA binding targets (130).

Myc was shown to be directly involved in the control of growth factor dependent proliferation when microinjection of c-Myc protein into the nuclei of quiescent cells resulted in the stimulation of DNA synthesis (117). Furthermore, the expression of c-*myc* increases as quiescent cells enter the G_1/S phase of the cell cycle (Figure 1.5). Although the levels of c-*myc* increase dramatically when resting cells enter the cell cycle, c-*myc* expression is invariant throughout the cell cycle (131). Downregulation of c-*myc* correlates with the onset of differentiation and enforced expression has been shown to inhibit differentiation. Therefore, it was proposed that Myc acts as a molecular switch directing cells either to terminal differentiation or continued proliferation (132).

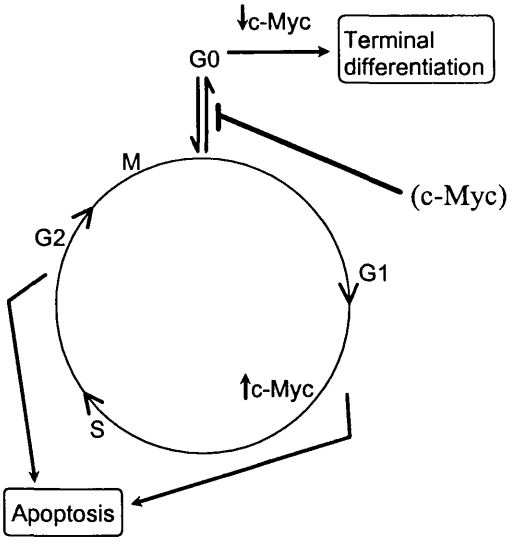


Figure 1.5. c-Myc and the cell cycle. The levels of cellular c-Myc increase during the G_1/S phase of the cell cycle and decrease as cells enter G_0. The effects of over-expression of c-Myc (in parenthesis) on terminal differentiation is also indicated.

Myc is tightly controlled by both growth factors and the repressors Max and Mad (130). Tight control of Myc is paramount because of its involvement in the control of proliferation where the inappropriate expression of c-*myc* could lead to uncontrolled growth and cancer. The risk of *Myc* over-expression may be lessened by its tendency to promote apoptosis instead of proliferation in the absence of suitable survival factors (Figure 1.6). High levels of Myc allow the cell to enter the cell cycle and divide when stimulated by mitogens, but in the absence of survival factors, the cell will die by apoptosis (133). However, in cancer cells the over-expression of *Myc* often leads proliferation even when no exogenous mitogenic signal is present. The apoptotic cell death which would normally occur in response to Myc can be prevented by the co-expression of the anti-apoptotic gene, *Bcl*-2 (134;135). Other genes which cooperate with *Myc* in tumorigenesis could also influence this decision of life or death by rescuing cells from Myc-induced apoptosis, thus maintaining the continued survival of the cell. Therefore, the identification of genes which synergise with *Myc* in this process represents a major goal in cancer research and is discussed further in section 1.3.

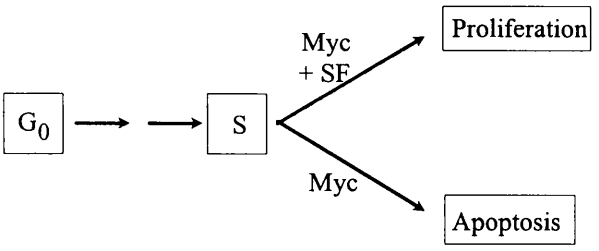


Figure 1.6 Myc links the decision between proliferation and apoptosis by the presence or absence of survival factors (SF).

1.3 Retroviruses as tools to study multistage tumorigenesis

1.3.1 Proviral tagging in transgenic mice

Proviral tagging in mouse model systems has made an important contribution to our understanding of the sequential steps involved in the accumulation of mutations resulting in cancer (for reviews see (12;136)). When an oncogene such as c-*myc* is transferred to the germline of the mouse genome under the control of the immunoglobulin heavy chain enhancer (Eμ-*myc*), the transgenic animals invariably develop B cell lymphomas (19;20). However, the tumours occur with variable latency and are generally monoclonal, suggesting that the abnormal expression of the transgene in itself is not sufficient for the development of malignancy but that other additional events are also required for full transformation.

Thus, the expression of an oncogene in transgenic mice can be viewed as the initiating event in the multistage process. In an effort to identify the events which cooperate with the transgene in tumorigenesis, proviral tagging has successfully been employed (19;20;81;87;88;137). M-MuLV infection of newborn mice which carry the transgene in their genome characteristically show a dramatic increase in the incidence and acceleration of the development of tumorigenesis compared to their non-transgenic littermates. Thus, identifying genes at common insertion sites in these tumours represents an efficient way of identifying gene(s) which synergise with the transgene. However, even the tumours induced in this

way are generally mono- or oligo-clonal suggesting that further events are involved in the development of the fully malignant phenotype. This is in agreement with the conjecture that 3-6 separate events are required for tumorigenesis (reviewed in (1)).

The first reported use of proviral tagging in transgenic mice was in 1989 by van Lohuizen *et al.* Eμ-*Pim*-1 transgenic mice neonatally infected with M-MuLV showed a rapid acceleration of T lymphomagenesis compared to their non-transgenic littermates (137). The oncogene *Pim*-1 encodes a cytoplasmic serine/threonine protein kinase (80) and was originally discovered as a common insertion site in M-MuLV induced thymic lymphomas (79). Eμ-*Pim*-1 transgenic mice showed a low incidence of clonal tumours after a long latency of around 7 months. However, on infection with M-MuLV, clonal T lymphomas developed after 7-8 weeks in transgenic animals compared to 22 weeks in non-transgenic controls. In all these lymphomas either c-*myc* (80%) or N-*myc* (20%) was activated by proviral insertion resulting in high levels of expression of the respective *Myc* genes. This strongly suggests that *Pim*-1 can cooperate with both c- and N-*myc* in T cell lymphomagenesis. Formal confirmation that *Pim*-1 and c-*myc* directly synergise in lymphomagenesis came when Eμ-*myc* mice were crossed with Eμ-*Pim*-1 mice to yield double transgenic mice over-expressing both oncogenes (138). These mice developed pre-B cell lymphomas *in utero*, and represent the strongest collaboration observed between c-*myc* and another oncogene to date. However, despite their severe phenotype and the speed of onset, the lymphomas of these double transgenic mice were frequently clonal indicating that further events had occured during the development of the cancer.

The system described above led to the identification of genes which collaborate with *Pim*-1 in lymphomagenesis. To determine whether genes which collaborate with c-*myc* in B cell lymphomagenesis could be identified in the same way, Eμ-*myc* transgenic mice were infected neonatally with M-MuLV (20). In addition to the anticipated activation of *Pim*-1, three novel common insertion sites were identified: *Bmi*-1, *Pal*-1 and *Bla*-1. Together they encompassed 75% of these

tumours suggesting there are still a number of genes activated by proviral insertion in these lymphomas which can also synergise with c-*myc*. A similar study led to the identification of not only *Pim*-1 and *Bmi*-1 as potential collaborators with c-*myc*, but also the novel loci *Tic*-1 (previously known as *Pim*-2) and *Emi*-1 (19).

In another *Myc/Pim*-1 bitransgenic system, *Gfi*-1 and its neighbouring loci, *Pal*-1/*Evi*-5/*Eis*-1, were identified as targets of proviral integration in a high percentage of the tumours which arose following neonatal retroviral infection of these mice (37-39). All of these integrations resulted in the over-expression of *Gfi*-1 suggesting that the zinc finger protein efficiently cooperates with both *Myc* and *Pim*-1 in the development of T cell lymphomagenesis. This system also identified *Tiam*-1 as a target for M-MuLV integration in 12% of the tumours, suggesting that this gene can also act in concert with *Myc* and *Pim*-1 in tumorigenesis.

Other transgenic systems have been generated to identify novel genes involved in tumorigenesis. For example, c-*myc* has been linked to a T cell specific regulatory element, the locus control region from the human CD2 gene. This results in the over-expression of *Myc* specifically in the T cell compartment (139). In M-MuLV accelerated lymphomas of CD2-*myc* transgenic mice, the *Til*-1 locus was identified as a common site of retroviral integration in 33% of the T cell tumours (88). Viral insertions at *Til*-1 have subsequently been found to transcriptionally activate the CBFA1 (PEBP2αA) transcription factor, suggesting that this gene also cooperates in the development of T cell lymphoma (89).

Proviral tagging of transgenic mice and subsequent transplantation of primary tumours to syngeneic hosts represents another variation on the theme. This provides an opportunity to identify novel genes contributing to tumour progression. The novel proto-oncogene *Frat*-1 was identified in this way (51). Primary tumours of *Pim*-1 or *Myc* transgenic mice induced by M-MuLV were transplanted and the resultant tumours analysed to identify common insertion

loci. *Frat*-1 was found to be activated in a minor fraction of primary tumours, while the transplanted tumours were rearranged in 17% and 30% of tumours in the *Pim*-1 and *Myc* transgenic mice respectively. Thus a novel oncogene has been identified which collaborates with both *Pim*-1 and *Myc* in T cell lymphomagenesis which is specifically involved in tumour progression.

By analogy to studies of proviral tagging of transgenic mice to identify genes which collaborate with the transgene, the initiating event may be a transduced oncogene introduced with the viral inoculum. For example, Tsatsanis *et al* showed that FeLV carrying a transduced *Myc* allele led to the identification of genes which act in concert with *Myc* in the accelerated T cell lymphomas (25). Thus, *Flvi*-2 (the feline homologue of *Bmi*-1), *Pim*-1 and *Fit*-1 can collaborate with c-*myc* in T cell lymphomagenesis in the cat.

1.3.2 Complementation groups

Certain combinations of activated oncogenes confer a selective growth advantage on a tumour cell, while others do not. As discussed above, proviral tagging has been a very useful strategy for the identification of genes which can synergise during tumorigenesis in transgenic mice. However, some combinations of activated oncogenes are never observed together in these systems. For example, in Eμ-*Pim*-1 transgenic mice infected with MuLV, the tumours frequently had activated c- and N-*myc* genes (137). However, the activation of both genes was never observed in a single cell. This suggests that not only is *Myc* an efficient collaborator of *Pim*-1, but that the activation of c-*myc* does not offer any further selective growth advantage to a cell which is already carrying an activated N-*myc* gene, and vice versa.

From this, and other experiments, the concept of complementation groups evolved where c-, N- and L-*myc* belong to the same complementation group (Figure 1.7). Genes which have the ability to complement or substitute for another gene are often part of the same gene family which have similar functions or act in the same growth control pathway, for example the *Myc* gene family.

However, as discussed below, some examples exist where genes appear to fall into the same complemention group but which are structurally unrelated, for example *Bmi*-1 and *Gfi*-1. The observation that many oncogenes show functional overlap suggest that a limited number of growth control pathways are involved. The elucidation of the functions of redundant genes in growth and development will lead to an increased understanding of these processes and how they are deregulated during tumorigenesis.

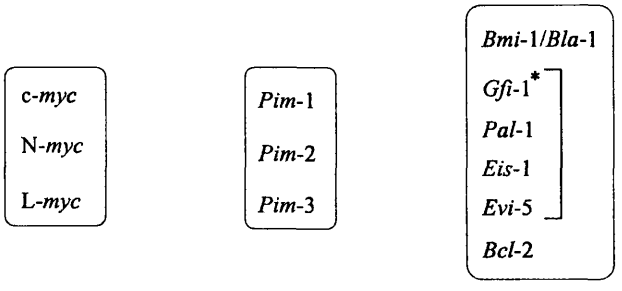


Figure 1.7 Oncogene complementation groups. There are three main complementation groups of structuarlly or functionally related oncogenes which were originally described by A. Berns *et al* (39) by establishing synergy in double transgenic mice. The asterisk next to *Gfi*-1 signifies that this gene is the target of insertions at *Pal*-1/*Evi*-5/*Eis*-1.

Integrations within the *Gfi*-1/*Pal*-1/*Evi*-5/*Eis*-1 locus are mutually exclusive with activation of *Bmi*-1 indicating that they fall within the same complementation group (19;20;38;39). Furthermore, activation of *Bmi*-1 in lymphomas of M-MuLV infected Eμ-*bcl*-2 transgenic mice is never seen, although c-*myc* is a frequent target of proviral insertion in this system. Thus, *Bmi*-1 and *Bcl*-2 may also fall within the same complementation group along with *Gfi*-1, all of which can synergise with *Myc* in lymphoma development.

This strong synergy with *Myc* can be rationalised by looking at the functions of *Bmi*-1, *Bcl*-2 and *Gfi*-1. The *Gfi*-1 proto-oncogene encodes a zinc finger

repressor protein which collaborates with *Myc* in lymphomagenesis (111). A recent report from Grimes *et al* demonstrates that over-expression of the *Gfi*-1 proto-oncogene represses *Bax* expression and inhibits T cell death (54). They show that the promotion of interleukin 2 (IL-2) independence by *Gfi*-1 in an IL-2 dependent T cell line is not only due to stimulation of cell cycle progression, but also to inhibition of apoptosis. This anti-apoptotic property can be directly attributed to the repression of *Bax*, a member of the *Bcl*-2 gene family. The over-expression of *Bcl*-2 inhibits apoptosis (140), whereas *Bax* can induce apoptosis (141). Therefore the repression of *Bax* by *Gfi*-1 is thought to inhibit apoptosis, possibly by modulating the balance between *Bax* and *Bcl*-2. Thus the over-expression of *Gfi*-1 is equivalent to the over-expression of *Bcl*-2 during oncogenesis. This could explain the strong collaboration of *Gfi*-1 and *Myc* during tumorigenesis since c-*myc* induced apoptosis is inhibited by *Bcl*-2 (134;135), for which *Gfi*-1 may be able to substitute. Furthermore, this might explain the functional redundancy between activation of *Gfi*-1 and *Bcl*-2 because both have an inhibitory effect on apoptosis when over-expressed.

Bmi-1 is another oncogene which has been shown to collaborate with *Myc* in tumorigenesis (111) and has been placed within the same complementation group as *Gfi*-1 and *Bcl*-2 (39;87). As for *Gfi*-1 and *Bcl*-2, it has been suggested that activation of *Bmi*-1 may act to inhibit apoptosis (87), although this has not been formally demonstrated. Again this might explain the strong collaboration between *Myc* and *Bmi*-1. Thus the activation of *Bmi*-1, *Bcl*-2 and *Gfi*-1 in any combination is rarely seen during tumorigenesis since this would not offer that cell any further growth advantage.

1.3.3 Complementation tagging

Complementation tagging is a further approach which has been developed to investigate the pathways in which oncogenes act in multi-step tumorigenesis (14;81). This technique involves proviral tagging in mice where the gene of interest has been deleted. Complementation tagging has been successfully used to identify genes which can substitute for the *Pim*-1 gene (81). As discussed

before *Pim*-1 and c-*myc* are strong collaborators in lymphomagenesis. Therefore, proviral tagging in Eμ-*myc* mice which lack a functional *Pim*-1 gene (Eμ-*myc*/*Pim*-$1^{-/-}$) should identify gene(s) which act downstream or in parallel to *Pim*-1 in lymphomagenesis and which can collaborate with c-*myc*. Together with a PCR-based screen, *Pim*-2 was identified as a common insertion site in 80% of these tumours (81). *Pim*-2 was found to be 53% identical to *Pim*-1 at the amino acid level and Eμ-*Pim*-2 mice were very similar to Eμ-*Pim*-1 mice in oncogenic behaviour. In addition, the activation of *Pim*-2 in mice lacking functional *Pim*-1 could rescue the *Pim*-1 null phenotype (82). Furthermore, the functional redundancy between *Pim*-1 and *Pim*-2 was demonstrated by the lack of activation of *Pim*-2 in M-MuLV induced lymphomas of Eμ-*Pim*-1 mice (81). Both genes can collaborate with c-*myc* and *Gfi*-1 or *Bmi*-1 (39;82).

In a parallel study, transgenic Eμ-*myc* mice which were homozygous null for both *Pim*-1 and *Pim*-2 were infected with M-MuLV (14). Again *Bmi*-1 and the *Gfi*-1/*Pal*-1/*Evi*-5/*Eis*-1 locus were frequent targets of proviral insertional activation. However, two novel common insertion sites were also identified: 1) *Pim*-3, which is closely related in sequence to the other two *Pim* family members; and, 2) *cot,* a serine/threonine kinase which was identified as the mouse homologue of the *Tpl*-2 common insertion site in rats (92). Thus, complementation tagging of *Myc* collaborating oncogenes which are functionally equivalent to *Pim*-1 has led to at least three oncogenes which may act downstream or in parallel to *Pim*-1 in tumorigenesis.

In conclusion, proviral tagging in transgenic mice has led to the assignment of at least three major complementation groups which can all collaborate with each other but not with members of the same group (Figure 1.7 and (39)). 1) The *Myc* family of structurally and functionally related oncogenes (c-*myc*, N-*myc* and L-*myc*). 2) The *Pim* family of structurally and functionally related of oncogenes (*Pim*-1, *Pim*-2 and the recently identified *Pim*-3). 3) The group of structurally unrelated, putative anti-apoptotic oncogenes (*Bcl*-2, *Gfi*-1 and *Bmi*-1). Although

Bmi-1 has been assigned to the latter group, its function as an inhibitor of cell death is circumstantial (87).

1.4 Aims

The aim of this study was to identify the gene which is targeted by retroviral integration at the *Fit*-1 locus in FeLV-*myc* induced T cell lymphomas of the domestic cat. The *Fit*-1 common insertion locus was previously found to be occupied by an FeLV provirus in 16% of tumours screened using a single copy probe flanking the insertion site from one of these tumours (25;42). Despite the isolation of a conserved open reading frame from within the major insertion cluster and 30kb of cloned genomic DNA, no transcription units affected by proviral insertion in these tumours was identified. However, the recently generated FT1 cell line which is rearranged at both *Fit*-1 and c-*myc* allowed a more thorough analysis of the *Fit*-1 locus to search for a gene affected by insertion by screening for sequences that are single copy, conserved in evolution and expressed as RNA.

CHAPTER 2

Chapter 2

2. Materials and Methods

2.1 Materials

Described below are a number of the most commonly used reagents.

2.1.1 Eukaryotic cell lines

FT1: A feline T cell line which was established from a spontaneous thymic lymphosarcoma with a FeLV integration at c-*myc* (142) and at *Fit*-1.

3201: FeLV negative feline T cell line with a germ line rearrangement of c-*myc* derived from a naturally occurring thymic lymphosarcoma (143).

F422: Feline T cell line, containing an FeLV transduced with v-*myc*, established from a thymic lymphosarcoma of a kitten inoculated with the second passage of the Rickard strain of FeLV (144).

T3: FeLV positive lymphoid tumour cell line established from a naturally occurring thymic lymphosarcoma, contains a replication-competent FeLV helper and recombinant FeLV provirus with a v-*myc* oncogene (7).

T17: FeLV positive cell line established from a naturally occurring thymic lymphosarcoma, contains transduced c-*myc* and T cell receptor β chain, as well as helper virus (145).

FL74: FeLV positive cell line established from a naturally occurring thymic lymphosarcoma (146).

MyaI: Interleukin 2-dependent feline T-lymphoblastoid cell line (147).

CEM: Human T lymphoblastoid cell line derived from peripheral buffy coat of a 4 year old Caucasian female with acute lymphoblastic leukaemia (148)

Jurkat: The Jurkat E61 clone is a human acute T cell leukaemia cell line (149).

SupT1: Human T cell lymphoblastic lymphoma line (150).

2.1.2 Bacterial strains

E. coli DH5α: MAX Efficiency DH5α Competent Cells, Life Technologies, GibcoBRL.

E. coli INVαF`: One Shot Competent cells, Invitrogen. F` *end*A1 *rec*A1 *hsd*R17 (r_k^-, m_k^+) *sup*E44 *thi*-1 *gyr*A96 *rel*A1 ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 λ^-.

E. coli TOP10: One Shot Competent cells, Invitrogen. F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *deo*R *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G.

2.1.3 Cloning vectors

pCR2.1: *lacZ*$^+$, *amp*^{r}, *kan*^{r}, T7 promoter and priming site, M13 forward and reverse primer annealing sites. Plasmid vector for the direct insertion of a PCR product that has been generated by *Taq* polymerase which has a non-template dependent activity that adds a single deoxyadenosine to the 3` ends of PCR products (Invitrogen).

pCRII: As above but also has SP6 promoter and priming sites (Invitrogen).

pUC18/19: pBR322 based cloning vector. *lacZ*$^+$, *amp*^{r}, multiple cloning site and M13 forward and reverse primer annealing sites (GibcoBRL).

pBluescriptII KS/SK: A pUC19 derived phagemid. *lacZ*$^+$, *amp*^{r}, multiple cloning site and M13 forward and reverse primer annealing sites (Stratagene).

2.1.4 Chemicals and enzymes

Most of the chemicals used were purchased from Sigma, BDH, Boehringer Mannheim, Pharmacia or GibcoBRL unless otherwise stated. All chemicals were of Analar quality. Enzymes and the appropriate buffers were purchased from GibcoBRL Life Technologies unless otherwise stated.

2.1.5 Radiochemicals

$\alpha(^{32}P)$ dCTP, specific activity 3000Ci/mmol was supplied by Amersham Life Science.

2.1.6 Media and antibiotics

L. broth: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride in ddH_2O autoclaved and stored at room temperature.

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

2 X YT: Bacto-tryptone, bacto-yeast extract, NaCl (GibcoBRL). 31g made up in 1litre of ddH_2O, autoclaved and stored at 4 °C.

Ampicillin: 50mg/ml in ddH_2O . Filtered through a 0.22μm filter, aliquotted and stored at -20°C. Used at a final concentration of 100μg/ml.

Kanamycin: As above but used at a final concentration of 25μg/ml.

RPMI-1640: Tissued culture medium supplemented with 10-20% foetal calf serum (FCS), 100units/ml penicillin, 10μg/ml streptomycin and 2mM glutamine (all GibcoBRL).

Ciprofloxacin: Anti-mycoplasma antibiotic added to tissue culture medium as necessary to a final concentration of 10μg/ml (Miles, Bayer Diagnostics).

2.1.7 Stock solutions and general reagents

Ammonium persulphate: 10% (w/v) stock solution in ddH_2O, freshly made.

Denaturation buffer: 1.5M NaCl, 0.5M NaOH in dH_2O. Stored at room temperature.

Denhardt's solution (50X): 1% bovine Serum Albumin (BSA), 1% Ficol, 1% polyvinyl pyrollidone in ddH_2O. Aliquoted and stored at -20°C.

DNA size markers: ϕX174 RF DNA/*Hae*III fragments and λDNA/*Hin*dIII fragments (GibcoBRL). 5kb ladder and λ ladder (Bio-Rad).

DEPC-treated ddH_2O: For RNase free H_2O, ddH_2O was treated with 0.05% diethyl pyrocarbonate (DEPC) overnight at room temperature. To remove any trace DEPC, the DEPC-H_2O was autoclaved for 30 minutes at 15lb.

Ethidium bromide: 10mg/ml stock in ddH_2O. Working concentration 0.5μg/ml. Stored at room temperature protected from light.

Gel loading buffers:

1. DNA- 50% glycerol, 0.5% bromophenol blue, 0.5% xylene cyanol, 100mM EDTA in ddH_2O. Stored at room temperature and used at a 1:10 dilution.
2. RNA - 50% formamide, 2.2M formaldehyde, 1X MOPS, in ddH_2O. Made fresh.

Gene screen (20X): 0.5M NaH_2PO_4, 0.5M Na_2HPO_4, pH6.5. Stored at room temperature.

High Prime: Premixed solution for random primed DNA labelling using ($\alpha^{32}P$) dCTP and random oligonucleotides as primers (Boehringer Mannheim).

Hybond N/N^+: Nylon membrane for the transfer of nucleic acids (Amersham Life Science).

MOPS buffer (10X): 200mM MOPS pH7.0, 50mM potassium acetate, 10mM EDTA. Stored at 4°C in the dark.

Neutralisation buffer: 1.5M NaCl, 0.5M Tris-HCl pH8.0 in dH_2O. Stored at room temperature.

Nick column: Sephadex G-50, Pharmacia Biotech, stored at room temperature.

Oligonucleotides: Unlabeled and IRD-41 fluorescent labeled oligonucleotides were synthesised by MWG Biotech. Unlabeled oligonucleotides resuspended to 1mg/ml in ddH_2O, IRD-41 labeled oligonucleotides to 1pmol/µl. Stored at -20°C (IRD-41 labeled oligos were also protected from light).

Polyacrylamide solution: Acrylamide/Bis acrylamide stock solution. 30% (w/v) acrylamide, 1.579% (w/v) bis acrylamide, ratio 19:1. Severn Biotech Ltd., stored in the dark at 4 °C.

Polyacrylamide (4%) solution with urea: 4.8ml Sequagel XR ultra pure concentrate (National Diagnostics), 25.2g urea, 7.2ml 10X TBE in a total volume of 60ml with ddH_2O. 400µl of 10% ammonium persulphate and 40µl of TEMED was added for each sequencing gel using the 66cm Li-Cor gel apparatus.

Phosphate buffered saline (PBS): 100mM NaCl, 80mM di-sodium hydrogen orthophosphate, 20mM sodium di-hydrogen orthophosphate, adjusted to pH7.5. Autoclaved and stored at 4°C.

Pre-hybridisation buffer:

1) Southern blots: 6 X SSC, 0.5% SDS, 5X Denhardt's solution and 30µg/ml heat denatured, sheared salmon sperm DNA.

2) Northern blots: 50% deionised formamide, 4 X Denhardt's solution, 4 X SSC, 1.6 X Gene Screen, 0.1% SDS, 8% Dextran Sulphate and 30μg/ml heat denatured, sheared salmon sperm DNA.

RNA ladders: 0.24-9.5kb RNA ladder, stored at -70°C. 3-5μg/lane (GibcoBRL).

RNase A: Prepared as a 10mg/ml stock in 10mM Tris-HCl pH7.5 and 15mM NaCl. Boiled for 15 minutes and cooled slowly. Stored at 20°C.

RNAzol B: Total RNA was extracted from cell lines or tissue using RNAzol B (Biogenesis). Stored at 4°C in the dark.

SSC (20X): 3M NaCl, 0.3M Sodium Citrate in dH_2O and adjusted to pH7.0. Stored at room temperature.

STET: 8% sucrose, 50mM Tris-HCl pH8, 50mM EDTA pH8, 5% triton X100 made up in ddH_2O, filtered (0.22 μm). Aliquots stored at 4°C or -20°C.

TAE (50X): 2M Tris-HCl pH8.15, 1.5M NaOAc, 1M NaCl, 0.1M EDTA. Stored at room temperature.

TBE (10X): 0.9M Tris-HCl, 0.9M Boric acid, 25mM EDTA pH8.3. Stored at room temperature.

TE: 10mM Tris-HCl, 1mM EDTA adjusted to pH8.0. Autoclaved and stored at room temperature.

X-Gal (5-bromo-4-chloro-3-indolyl-β-galactoside): 160mg/ml in dimethyl-formamide (DMF) added to L-agar to a final concentration of 80μg/ml. Stored in aliquots at -20°C in the dark.

2.2 Methods

2.2.1 Eukaryotic cell culture

2.2.1.1 Maintenance of eukaryotic cells

Eukaryotic feline and human T cell suspension cells were grown in plastic flasks (Costar Corporation, USA) at 37^0C in an atmosphere of 5% CO_2 in air. Growth medium is as outlined in 2.1.6 with the exception of the Mya 1 and T17 cell lines which were supplemented with 100units/ml of recombinant human IL-2 and 2 x 10^{-5} moles/ml of 2-mercaptoethanol. Cells were subcultured every 2-3 days and maintained at densities of between 5 x 10^5 and 1 x 10^6 cells/ml. Cells numbers were determined using a Coulter Counter ZM, Coulter Electronics Ltd.

2.2.1.2 Separation of healthy from non-viable cells

The separation of live cells from dead or non-viable cells was performed by phase separation in Ficoll-paque (Pharmacia). Briefly, cells were resuspended in 5ml of medium in a 15ml Falcon tube and underlayed with 5ml Ficoll followed by centrifugation at 2000rpm for 10 min. Live cells form an interphase between the medium and the Ficoll while dead cells form a pellet. The interphase was then removed and the live cells diluted to 50% with medium to remove any remaining Ficoll. Live cells were pelleted at 1500rpm for 5 min and cultured as above.

2.2.1.3 Long term storage of eukaryotic cells

Cells were resuspended at approximately 3 x 10^6 cells/ml in growth medium supplemented with 10% dimethyl sulphoxide (DMSO, Koch-Light Laboratories Ltd.), frozen at -1^0C/min to -70^0C and stored in liquid nitrogen.

2.2.2 Preparation of DNA and RNA

2.2.2.1 Isolation of plasmid, cosmid and PAC DNA

Recombinant plasmid DNA was isolated using a number of different procedures depending on the intended use as outlined below. In all cases, a single colony was inoculated into L-broth containing the appropriate antibiotic unless otherwise stated.

2.2.2.2 Small scale purification of plasmid DNA

'Miniprep' DNA was isolated by the STET method as outlined in Maniatis *et al* (151). Briefly, 1.5ml of an overnight culture was pelleted and resuspended in 100µl STET containing 3mg/ml lysozyme and 100µg/ml RNaseA. The samples were boiled for 45 sec and centrifuged at 14000rpm for 15 min. The DNA was precipitated with 100µl propan-2-ol, pelleted by centrifugation, washed in 70% ethanol, dried and resuspended in the desired volume of TE pH8.0. DNA was stored at -20^{0}C.

2.2.2.3 Purification of sequencing grade plasmid and cosmid DNA

Large quantities of high quality plasmid or cosmid DNA was purified from 100ml or 500ml respectively of an exponentially growing overnight culture of recombinant bacteria using the QIAGEN Plasmid Purification Maxi kit according to the manufacturer's instructions. Small scale purification of sequencing grade plasmid DNA was purified from 1.5 - 3ml of an overnight culture using the QIAwell 8 Plasmid Purification kit according to the manufacturer's instructions. 500ng of DNA was used per cycle sequencing reaction. DNA was stored at -20^{0}C.

2.2.2.4 Purification of bacteriophage P1-derived (PAC) constructs

PAC DNA was isolated using a variation of the QIAGEN Plasmid Maxi Kit protocol. Briefly, 0.5ml of an overnight culture of the recombinant PAC clone was inoculated into 500ml of 2 X YT containing 25µg/ml kanamycin. The culture was grown with vigorous shaking at 37°C until an A_{550} of 0.15 was reached. 5ml of freshly made, filter-sterilised 0.1M IPTG was then added to the 500ml culture and grown at 37°C until an A_{550} of 1.3-1.5 was reached. The remainder of the protocol is identical to the QIAGEN Plasmid Maxi Kit protocol.

2.2.2.5 Isolation of genomic DNA

High molecular weight genomic DNA was prepared from 1 x 10^7 - 10^8 cells using the Nucleon BACC2 kit (Scotlab) according to the manufacturer's instructions, except for the DNA precipitation stage: the precipitated DNA was spooled, washed in 70% ethanol, briefly air dried and resuspended in an appropriate volume of TE pH8.0. The concentration of DNA was calculated on the basis that a 50μg/ml solution of double stranded DNA has an OD_{260} measurement of 1. DNA was stored at 4^0C and remains stable for a number of years.

2.2.2.6 Isolation of total cellular RNA

Total cellular RNA was prepared using RNAzol B (Biogenesis) which is based on the guanidinium thyocyanate-phenol chloroform extraction method of Chomczynski and Sacchi (152). The manufacturer's instructions were followed using 2ml of RNAzol B per 10^7 cells. All procedures were performed at 4^0C using DEPC treated ddH_2O and RNase free tubes and pipette tips. The concentration of RNA was calculated on the basis that a 40μg/ml solution of RNA has an OD_{260} measurement of 1. RNA was stored at -70^0C in DEPC treated ddH_2O.

2.2.2.7 Isolation of poly A^+ mRNA

Using total RNA as the starting material (above), the poly(A) mRNA fraction was isolated using PolyATtract mRNA Isolation Systems (Promega) according to the manufacturers instructions. All procedures were performed at 4^0C using DEPC treated ddH_2O and RNase free tubes. The concentration of RNA was calculated on the basis that a 40μg/ml solution of RNA has an OD_{260} measurement of 1. RNA was stored at -70^0C in DEPC treated ddH_2O.

2.2.3 Gel electrophoresis

Gels containing 0.6 - 2% agarose (w/v) in 1 X TBE or 1 X TAE were used to separate DNA (151). 50, 100, 200 or 400ml gels were poured in perspex tanks

and wells were cast using appropriate combs. Gel loading buffer was added to samples at a final concentration of 1 X, samples were then electrophoresed in 1 X TBE or 1 X TAE. When TAE gels were used, the running buffer was changed frequently to prevent buffer exhaustion. Known concentrations of DNA size markers were run alongside the samples in order to gauge both product size and yield. Gels were stained either by adding 0.25μg/ml ethidium bromide to the molten agarose prior to casting, or by staining the gel after electrophoresis for 20 min in running buffer or dH_2O containing 0.5μg/ml ethidium bromide. Gels were viewed on a short wave UV transilluminator (UVP Inc.) and photographed using the MWG 2000i gel documentation system (MWG-Biotech). Further analysis of band intensities and molecular weights of product was carried out as necessrary using Phoretix 1D Plus (Phoretix International, NonLinear Dynamics Ltd.).

2.2.4 Purfication of DNA from agarose gels

Following electrophoresis, a DNA fragment of interest was excised from an agarose gel using a clean scalpel. Highly purified DNA was obtained with the QIAquick Gel Extraction kit (Qiagen) according to the manufacturers instructions.

2.2.5 Restriction endonuclease digestion

Restriction endonucleases (GibcoBRL) were used with the buffers supplied in accordance with the manufacturer's recommendations. 10 units of enzyme was generally used to digest 1μg of miniprep DNA at 37^0C for 1 hour, and 100 units per 20μg of genomic DNA for 12-24 hours.

2.2.6 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) allows the rapid, specific amplification of DNA sequences between two specific oligonucleotides using very little starting material. The Perkin Elmer GeneAmp PCR Core Reagents kit was used according to the manufacturers instructions. The kit utilises a recombinant, thermostable *Taq* DNA polymerase encoded by a modified form of the *Thermus aquaticus* (*Taq*) DNA polymerase gene. The thermostable property of *Taq*

allows the cyclical heat denaturation of double stranded DNA at 95^{0}C. This step is followed by the annealing of the two complementary primers to the 3` boundaries of the target sequence, typically carried out at 55^{0}C unless otherwise stated. Finally the primers are extended at 72^{0}C. The target temperature at each stage is maintained for 1 minute and carried out for 20 - 35 cycles. Standard reaction conditions were, depending on the template, 10ng - 1µg of template DNA in a 50µl reaction mixture containing 10mM Tris-HCl pH8.3, 50mM KCl, 1.5 - 3 mM $MgCl_2$, 100µM of each nucleotide, 1µg of each primer and overlaid with 25µl of mineral oil. Thermal cycling was carried out in 0.5ml reaction tubes in a programmable thermal cycler (Perkin Elmer DNA Thermal Cycler). 5µl of the product was analysed by acrylamide or agarose gel electrophoresis.

2.2.7 Cloning of hybrid DNA molecules

2.2.7.1 Ligation of DNA fragments

Fragments of DNA generated by restriction digestion were ligated into approximately 50ng of vector using T4 DNA ligase (GibcoBRL) according to the manufacturers instructions. Vector DNA was linearised using an appropriate restriction enzyme. To prevent re-circularisation of the vector DNA, both 5`-phosphate groups were hydrolysed with 0.5 units of calf intestinal alkaline phosphatase (CIAP, Promega) at 37$^0$C for 30 min. If one or both ends generated by a restriction enzyme digest must be converted into blunt ends for cloning, the 5` or 3` protruding ends can be 'filled-in' or removed using T4 DNA polymerase (GibcoBRL) which has a 5`-3` polymerase activity as well as a 5`-3` exonuclease activity. Ligation of DNA fragments generated by PCR were carried out according to the manufacturer's instructions with the Original TA Cloning Kit or the TA-TOPO Cloning Kit (both from Invitrogen). Generally, 50ng of vector DNA and a three fold molar excess of insert DNA were ligated. Ligations were carried out overnight at 14^{0}C except when using the TA-TOPO kit where ligations were carried out at room temperature for 5 min.

2.2.7.2 Bacterial transformation

1-2µl of a 1:5 dilution of ligation mix was added to 20µl of competent bacterial cells (DH5α, GibcoBRL), or 1µl of TA cloned DNA was added to 50µl of INVαF` or TOP10 One Shot Competent cells (Invitrogen) along with 2µl of 2-mercaptoethanol. In all cases the mixtures were then incubated on ice for 30 min. Cells were subsequently heat shocked at 42^{0}C for exactly 45 seconds and returned to ice for 2 min. 250µl of SOC medium (Invitrogen) was then added and incubated with shaking for 30 - 60 min. Cells were then plated onto L-agar plates containing 100µg/ml ampicillin and 80µg/ml X-gal. These plates were incubated overnight at 37^{0}C.

2.2.7.3 Identification of recombinants

For transformation using *lacZ* complementation, white colonies were picked from plates containing X-gal and used to inoculate 5ml of L-broth containing ampicillin (100µg/ml) and grown overnight at 37^{0}C with shaking. Colonies containing the appropriate inserts were identified by small-scale plasmid purification and restriction digestion, or by performing PCR on a colony which has been added to 50µl of ddH_2O and boiled for 10 min then centrifuged for 10 min. Alternatively, when a 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 in 1975 (153). Briefly, white or pale blue colonies were picked and spotted on nitrocellulose filters over L broth agar plates containing 100µg/ml ampicillin and on a reference plate. Colonies were grown overnight at 37^{0}C then the DNA was denatured, neutralised and crosslinked to the filter as for Southern blots (see below). Colonies which contain the sequence of interest were identified by hybridisation with an appropriate probe (see below) and the relevant colony picked from the reference plate for further analysis.

2.2.8 DNA and RNA hybridisation analysis

2.2.8.1 Southern blot transfer of DNA

This was carried out essentially as described by Southern (154). Generally, 20µg of high molecular weight genomic DNA was digested with 100 units of the

appropriate restriction endonuclease overnight, precipitated with half volume of 7.5M ammonium acetate and two volumes of 100% ethanol and the DNA pellet dried and resuspended for several hours in TE pH8.0. DNA samples were electrophoresed as described above on a 0.6 - 0.8% agarose gel in 1 X TAE buffer overnight at 25V. $5x10^3$ counts per minute of end labeled marker (5kb ladder and/or λ HindIII DNA) together with 500ng of unlabeled marker were also loaded on the gel. Following electrophoresis, the gels were stained and photographed as described above, submerged in denaturation buffer for 30 min, followed by submersion in neutralisation buffer for 30 min. The gel was equilibrated in 20 X SSC and the DNA was transferred onto Hybond-N (Amersham) nylon membrane by capillary blotting overnight in 20 X SSC. The membrane was removed and rinsed briefly in 2 X SSC and the DNA crosslinked to the membrane (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation).

2.2.8.2 Northern blot transfer of RNA

10-40μg of total cellular RNA or 1-5μg of $polyA^+$ mRNA was lyophylised (VR-1 Hetovac, Heto) then resuspended in 20μl of RNA loading buffer and denatured for 15 min at 65^0C. Subsequently, 3-5μl of RNA running dye was added and the samples electrophoresed for 3 hours at 100V in 200ml of a 1% agarose gel containing 1 X MOPS and 2.2M formaldehyde in 1X MOPS buffer that was continually recirculated. 5μg of RNA ladders (GibcoBRL), which were treated in the same way as the RNA samples, were used as molecular weight markers. Following electrophoresis, the marker lanes were removed using a clean scalpel and stained in 3μg/ml ethidium bromide for 30 min, then destained overnight and photographed. The rest of the gel was washed twice in ddH_2O for 20 min to remove formaldehyde, then equilibrated in 10 X SSC and transferred overnight onto Hybond-N in 10 X SSC. The membrane was removed and rinsed briefly in 2 X SSC and the RNA crosslinked to the membrane (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation). All northern blots were probed with a rat gapdh probe to control for RNA loading and integrity. The gapdh probe is a 750bp *Eco*RI fragment purified from the plasmid pGapdh.

2.2.8.3 Preparation of radiolabeled DNA probes

PCR or restriction fragment DNA probes were gel-purified from non-specific or vector sequence and radioactively labeled using a 'random prime' DNA labeling kit (155) (High Prime, Boehringer Mannheim), and $\alpha(^{32}P)$ dCTP, specific activity 3000Ci/mmol (Amersham). Generally, 20-50ng of heat denatured DNA was radiolabeled using 50μCi (1.85Mbq) of $\alpha(^{32}P)$ dCTP and 4μl of High Prime in a final volume of 20μl, following the manufacturer's instructions. Unincorporated nucleotides were removed by gel-filtration through Sephadex-G50 beads (Nick Columns, Pharmacia) and labeled fragments were eluted in 400μl of TE buffer and the activity of 2μl was determined. Typically, incorporations of 10^8 - 10^9 cpm/μg were achieved.

2.2.8.4 End labeling of DNA size markers

1μg of DNA size marker was dephosphorylated using 0.5 units of calf intestinal alkaline phosphatase (CIAP, GibcoBRL), 1 X CIAP buffer and 10mM Tris-HCl pH8.0 for 15 min at 37^0C followed by 15 min at 56^0C. The DNA was extracted with phenol/chloroform and precipitated using 0.3M NaOAc and 100% ethanol. The DNA was resuspended in ddH_2O and 6μl of 5 X forward reaction buffer (GibcoBRL) and denatured at 90^0C for 3 min. 170μCi of $\gamma(^{32}P)$ dCTP (Amersham) and 10 units of T4 polynucleotide kinase (GibcoBRL) was added to a total volume of 30μl and incubated at 37^0C for 1 hour. The reaction was stopped with 1μl of 0.5M EDTA. Unincorporated nucleotides were removed as above. The activity of 2μl was determined and the sample diluted appropriately. Typically, 10^3 cpm per lane were loaded alongside unlabeled markers

2.2.8.5 Hybridisation of labeled probes to membrane bound nucleic acids

Standard high stringency conditions for the hybridisation of specific radiolabeled probes on nucleic acids immobilised on nylon membranes were as follows: membranes which had been pre-wetted in 2 X SSC and rolled into Hybaid hybridisation bottles were pre-hybridised in 10-20ml of the appropriate pre-

hybridisation buffer (see section 2.1.7) at 65^{0}C (Southern blots) or 42^{0}C (northern blots) for at least 2 hours in a Hybaid oven with continual rotation. Freshly boiled DNA probe was added to the pre-hybridisation solution: 10^6 cpm/ml was used for genomic Southerns and northerns and 10^5 cpm/ml was used for plasmid DNA Southern blots. Filters were hybridised overnight as above. After rinsing briefly with 2 X SSC, the membrane was washed for 20 min with three changes of 0.1X SSC, 0.5% SDS at 60^{0}C, unless otherwise stated. Membranes were then sealed in polythene and exposed to X-ray film (Kodak). Following hybridisation, filters were stripped of probe by continuous shaking in boiling dH_2O containing 0.1%SDS until the solution has cooled to room temperature.

2.2.9 DNA sequence analysis

Cycle sequencing reactions were carried out using IRD41-labeled primers (MWG-Biotech) and the ThermoSequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Life Science). Cycle sequencing is based on the chain termination method of Sanger (156), but uses a thermostable DNA polymerase to give multiple rounds of high temperature DNA synthesis. Briefly, 500ng of plasmid DNA, 1pmol of IRD-41 labeled primer (MWG-Biotech) and 2µl of reaction mix which contains 45µM each of dGTP, dATP, dTTP and dCTP, reaction buffer and thermostable DNA polymerase, were mixed in a 0.5ml reaction tube. The reaction is made up to 8µl with ddH_2O and overlaid with 20µl of mineral oil and denatured for 5 min at 95^{0}C then passed through 25 cycles of 95^{0}C for 30 seconds, 48-60^{0}C for 30 seconds and 72^{0}C for 30 seconds. The mineral oil was removed by running each sample down a parafilm gradient and collecting the aqueous phase at the bottom of the slope. 4µl of formamide loading buffer supplied with the kit was added and the reactions separated on denaturing gels (Sequagel XR ultra pure concentrate (National Diagnostics). Data recording was performed on a Li-Cor model 4000 DNA sequencer (MWG-Biotech). Typically 800-1000bp of sequence was generated from one run.

DNA sequence was analysed using the GCG package (Wisconsin Package Version 9.1, Genetics Computer Group (GCG), Madison, Wisconsin) and database searches were carried out by FastA searches in GCG or BLAST searches (Basic Local Alignment Search Tool, National Center for Biotechnology Information (NCBI): http://www.ncbi.nlm.nih.gov/BLAST/).

CHAPTER 3

Chapter 3

3. Feline *Fit*-1 and the search for transcribed sequences in the FT1 cell line

3.1 Introduction

Cancer is a multi-step process in which a normal cell develops into a fully malignant tumour via a mechanism of clonal expansion triggered by genetic events leading to deregulated proliferation. Retroviruses have been used for a number of years to identify genes involved in this process (12;97). As discussed in chapter 1, the basis of retrovirus-induced tumours appears to be an accumulation of somatic mutations, at least a number of which can be directly attributed to proviral integration within the host DNA. Although most integrations will have no significant effect on the host cell, some will inevitably affect critical cellular gene(s) involved in the regulation of cellular proliferation and differentiation. Therefore, a cell which has undergone an insertional mutation event affecting such an (onco)gene will have an increased survival potential allowing that cell to grow and divide in a clonal fashion, often leading to the outgrowth of a tumour.

Thus, retroviruses serve as useful tools to identify novel oncogenes in the vicinity of an integrated provirus. Indeed, retroviruses have been used to identify greater than 60 cellular genes as targets for insertion in tumours of the mouse, rat, chicken and cat (Table 1.1 in chapter1 and reviewed in (12)). This study will focus on retrovirus-induced lymphomas of the domestic cat where a number of parallels can be drawn from the better characterised murine system.

Cats inoculated with feline leukaemia virus (FeLV) isolates which carry a transduced *Myc* gene (FeLV-*myc*) were shown to develop thymic lymphomas with a relatively short latency period, 12-14 weeks post infection (3;4). Despite this short latency, the tumours were clonal, suggesting that *Myc* alone is not sufficient to induce full malignant transformation and that other genetic events may be required. Therefore, the possibility that helper proviral insertional mutagenesis played a role by activating a gene which collaborates with *Myc* was

investigated. This method of 'proviral tagging' has been frequently used to identify novel genes involved in tumorigenesis, whereby host DNA sequence flanking the provirus in one tumour is used to identify regions of chromosomal DNA interrupted commonly in independent tumours (7;11;12). This technique led to the identification of the *Fit*-1 common insertion locus (FeLV integration site in T cell lymphoma) (25;42). Identification of a gene at *Fit*-1 was expected to lead us to a gene which collaborates with *Myc* in the development of T cell neoplasia.

3.1.2 Background on the *Fit*-1 common integration locus

Fit-1 was cloned by inverse PCR from tumour F422-1T (Figure 3.1) and used to screen a panel of 63 thymic lymphomas. 16% of these tumours were found to be rearranged at *Fit*-1 (25;42). This is highly significant considering the low probability of 10 independent proviral integration events occuring at random in the same 10kb genomic fragment (10). Eight of these lymphomas were induced by naturally-occurring FeLV strains carrying a v-*myc* oncogene (25), and two were field case thymic lymphomas: T21-T and FT1, from which a cell line of the same name was established (142). Interestingly, the FT1 tumour and cell line also has a proviral insertion at c-*myc* ((100) and discussed in section 3.1.3). Furthermore, all the other tumours which were rearranged at *Fit*-1 contained an activated *Myc* gene except for tumour T21-T. Together, this suggests that the *Fit*-1 common integration locus harbours a gene which is activated by FeLV insertion and collaborates with c-*myc* in T cell leukaemogenesis, although this relationship may not be exclusive. Moreover, occupation of *Fit*-1 by FeLV was implicated as a late progression step in tumour development (25).

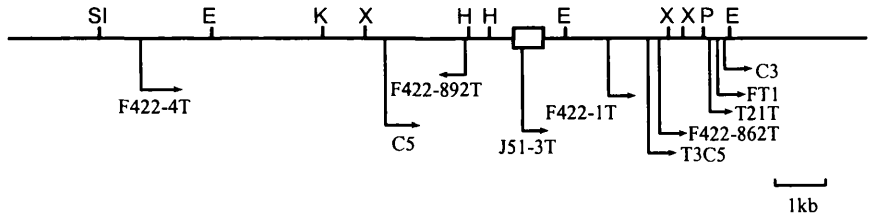


Figure 3.1 Restriction map and sites of proviral integration in the *Fit*-1 locus. This figure is based on information described previously (25;42). The positions and orientations of proviral integrations for each tumour and cell line (FT1) are shown. The position of the conserved *Fit*-1 probe is shown (open box). Restriction enzmye abbreviations: E, *Eco*RI; SI, *Sst*I; K, *Kpn*I; X, *Xba*I; H, *Hin*dIII; P, *Pst*I.

The *Fit*-1 locus has been mapped to feline chromosome B2 which carries several known oncogenes including *Pim*-1 and c-*myb*, both of which have been implicated in T cell leukaemogenesis (42). However, restriction mapping indicated no close linkage to these genes and northern blot analysis of RNA from primary tumours rearranged at *Fit*-1 revealed no evidence of long-range activation. Moreover, initial analysis of two other loci identified as common FeLV integration sites in T cell lymphomas, *Flvi*-1 (49) and *Flvi*-2 (23) (subsequently found to encode *Bmi*-1, also a c-*myc* collaborating gene (19;24)), excluded the possibility that these genes were linked to *Fit*-1.

A stretch of conserved sequence from the *Fit*-1 locus allowed the mapping of the murine homologue using the European Interspecific Backcross between *Mus musculus* and *Mus spretus* (EUCIB) (157). Three recombinants were detected in 119 backcross progeny screened for both the *Fit*-1 and *Ahi*-1 loci, a murine retroviral integration locus which is closely linked to c-*myb* on chromosome 10 (17;18). This total of 119 mice represented an initial random screening panel and a second panel of known chromosome 10 recombinants. Based on EUCIB algorithms these data placed *Fit*-1 approximately 1cM proximal to *Ahi*-1 on

mouse chromosome 10 (Figure 3.2). However, preliminary Southern analysis with a c-*myb* probe in the feline tumours with rearrangements at *Fit*-1 excluded the possibility that insertions at this common insertion locus actually represented insertions within the c-*myb* oncogene (42). Thus, *Fit*-1 was assumed to be a novel locus.

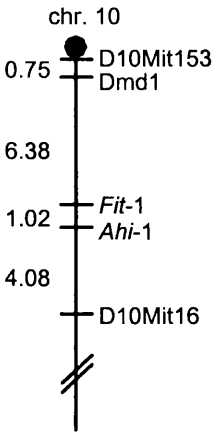


Figure 3.2 Genetic linkage map of the *Fit*-1 and *Ahi*-1 common insertion loci on mouse chromosome 10. The positions of these loci were determined relative to each other in 119 interspecific backcross progeny (157). Also shown are the positions of the anchor markers D10Mit153, Dmd1 and D10Mit16. Recombination distances between loci are given in centiMorgans and are listed on the left.

To search for a gene affected by insertion at *Fit*-1, DNA flanking the cluster of proviral insertions was screened for sequences that are single copy, conserved in evolution and expressed as RNA in a tumour specific fashion. To this end, three *Fit*-1 positive λ phage clones were isolated spanning approximately 30kb of genomic sequence (λRF1, λHT301B2 and λHT58A3). Southern blot hybridisation with a probe derived from the major insertion cluster revealed a region of unique sequence which showed evidence of cross-species hybridisation, the *Fit*-1 probe (Figure 3.1). This region also contained a long open reading

frame with some features of an internal coding exon. However, the *Fit*-1 probe failed to show any evidence of a transcription unit which was altered as a result of viral insertion at *Fit*-1. In addition, the application of the exon trapping technique revealed several candidate exons none of which showed evidence of homology to published sequences or of expression in either normal or tumour material. These preliminary results, together with an illustration of the *Fit*-1 λ phage clones and subclones, are summarised in Figure 3.3.

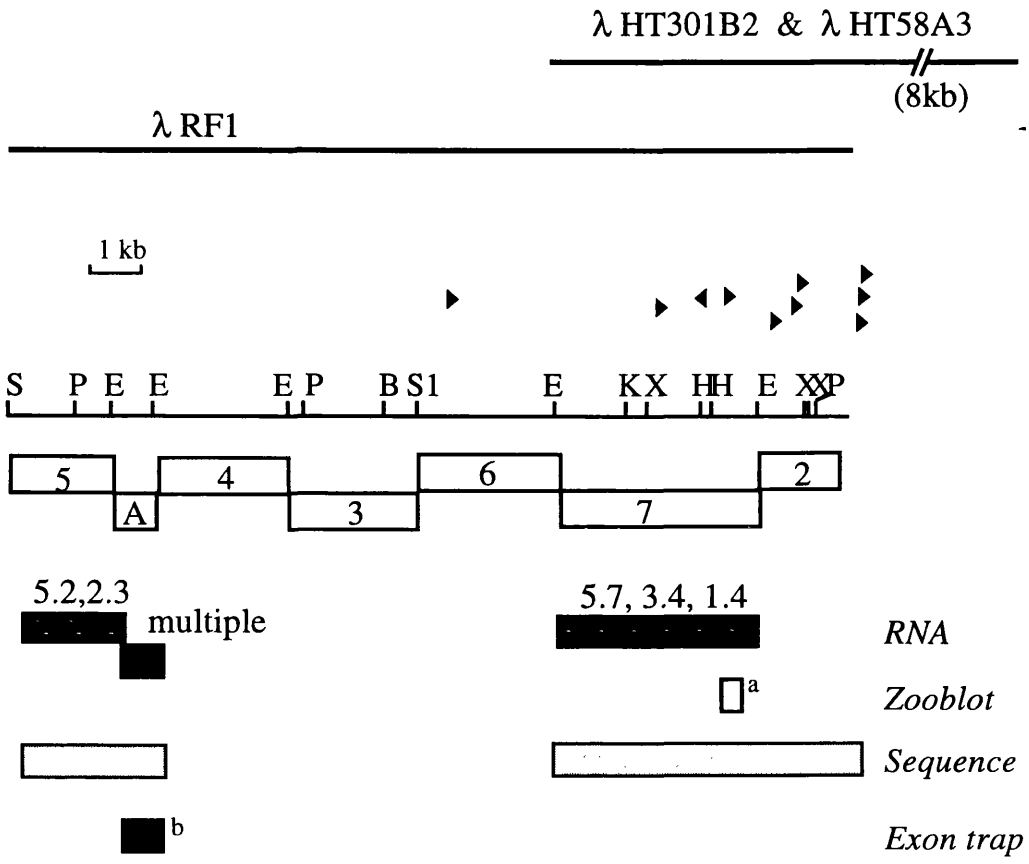


Figure 3.3 Summary of the *Fit*-1 locus illustrating the status of the *Fit*-1 locus prior to the start of this study. The lines above the restriction map show the extent of the phage clones obtained from the locus, only the λRF1 clone is shown in detail. The arrowheads indicate the positions and orientations of the integrated FeLV proviruses in individual tumours. The open boxes represent the λRF1 subclones while the shaded boxes below show domains for which DNA *sequence* is available, which are conserved across species (*zooblot*), or which hybridise to cellular *RNA*. (a) the conserved *fit*-1 probe, (b) putative exons identified by *exon trapping*. Restriction enzyme abbreviations: E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; P, *Pst*I; S, *Sal*I; SI, *Sst*I; X, *Xba*I.

A 1.4 Kb tumour specific RNA was identified in the FT1 cell line using a probe derived from the 3` end of the insertion cluster (H. Tsujimoto, Tokyo, Japan, unpublished results). However, reproducible signal could not be obtained on northern analysis. Furthermore, screening of a cDNA library did not identify any cDNA clones which hybridise to the probe. In addition, several weakly hybridising transcripts were identified by northern blot analysis using probes from around the insertion cluster, however none were expressed in a tumour specific manner. The search for transcripts up until this point has been severely hampered by the limited availability and relatively poor quality of primary tumour RNA.

3.1.3 The FT1 cell line

The FT1 T cell line was established from a cat with a spontaneous thymic lymphosarcoma (142). This cell line was shown to be rearranged at c-*myc* by Southern blot analysis. The c-*myc* rearrangement is due to the direct integration of an FeLV provirus immediately upstream and in the opposite transcriptional orientation to c-*myc* (100), illustrated in Figure 3.4. The LTR of this integrated provirus has three copies of an enhancer-like sequence (100). Therefore, the high level of c-*myc* transcript seen in FT1 RNA by northern blot analysis (142) is most likely due to enhancer insertional activation of c-*myc*.

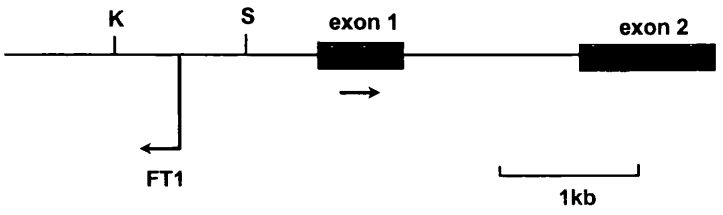


Figure 3.4 Partial restriction map illustrating the site of proviral insertion at c-*myc* in the FT1 cell line (100). The orientation of the provirus is upstream and in the opposite orientation to the direction of transcription of the c-*myc* gene (arrow). Only exons 1 and 2 of the c-*myc* gene are shown (closed boxes). Restriction enzyme abbreviations: K, *Kpn*I; S, *Sal*I.

In addition to a proviral integration at c-*myc*, the FT1 cell line also has an insertion at *Fit*-1 (H. Tsujimoto, unpublished) and it is this property which has made this cell line invaluable in the search for transcribed sequence at the *Fit*-1 locus. Previous studies were limited by the availablility and the viability of primary tumour RNA. However, FT1 cells represent a renewable source of both RNA and DNA with a rearrangement at *Fit*-1 for further analysis of the *Fit*-1 locus.

3.2 Results

3.2.1 Strategy for identifying transcription units at the *Fit*-1 locus

The evidence strongly suggested the presence of a gene at *Fit*-1 because 16% of tumours analysed for rearrangements at *Fit*-1 had an FeLV provirus inserted within this region. However, northern blot analysis with unique sequence probes from the insertion cluster failed to show any evidence of a transcription unit which was activated by viral insertion in primary tumour RNAs. Using the FT1 cell line, an unlimited supply of rearranged DNA and RNA, I set about to systematically search for transcript(s) which are altered as a result of viral insertion at *Fit*-1. For this analysis, probes derived from the λRF1 genomic

phage clone which spans the region surrounding the insertion cluster were isolated. Figure 3.5 outlines the strategy used to search for transcribed sequence in the FT1 cell line which were altered compared to a control cell line (3201, a feline T cell line with no rearrangement at *Fit*-1). The λRF1 subclone 2 is used as an example in this figure (see also section 3.2.1). However, this analysis was performed accross the entire λRF1 clone.

1. Identification of single copy sequence

It is likely that DNA which encodes a protein product will be single copy, unless it is part of a multi-gene family. Therefore, the DNA surrounding the proviral insertion cluster was screened to identify such sequences which could represent a gene targeted by these proviruses. To this end, λRF1 subclones (see Figure 3.3) were digested with various restriction enzymes (Figure 3.5A) and blotted on nylon membrane as outlined in section 2.2.8. The Southern blot was then hybridised with [α^{32}P] dCTP labelled total genomic DNA from the 3201 cell line.

Any DNA fragments which strongly hybridised to 3201 DNA after autoradiography were discarded since they represent repetitive sequences (Figure 3.5A and B, asterisk). Fragments which did not appear to hybridise on a short exposure were identified as single or low copy number sequences and were used for further analysis. For example the 1kb *Eco*RI/*Pst*I fragment indicated by the arrow in Figure 3.5A, did not hybridise strongly to total DNA on the Southern blot of this gel (Figure 3.5B, arrow) suggesting that this fragment did not contain repetitive sequences. Other DNA fragments from subclone 2 hybridised strongly to total DNA suggesting that these sequences occur frequently in the genome, for example the 400bp *Eco*RI/*Pst*I fragment (Figure 3.5A and B, asterisk). A putative single-copy fragment was then gel purified (section 2.2.4) and used to probe a Southern blot of 3201 DNA digested with various restriction enzymes (Figure 3.5C). In this example, the 1kb *Eco*RI/*Pst*I fragment from subclone 2 is confirmed as single copy as determined by the single band hybridising to this probe on Southern analysis (Figure 3.5C).

Figure 3.5 Summary of the procedure used to identify transcription units at the *Fit*-1 locus

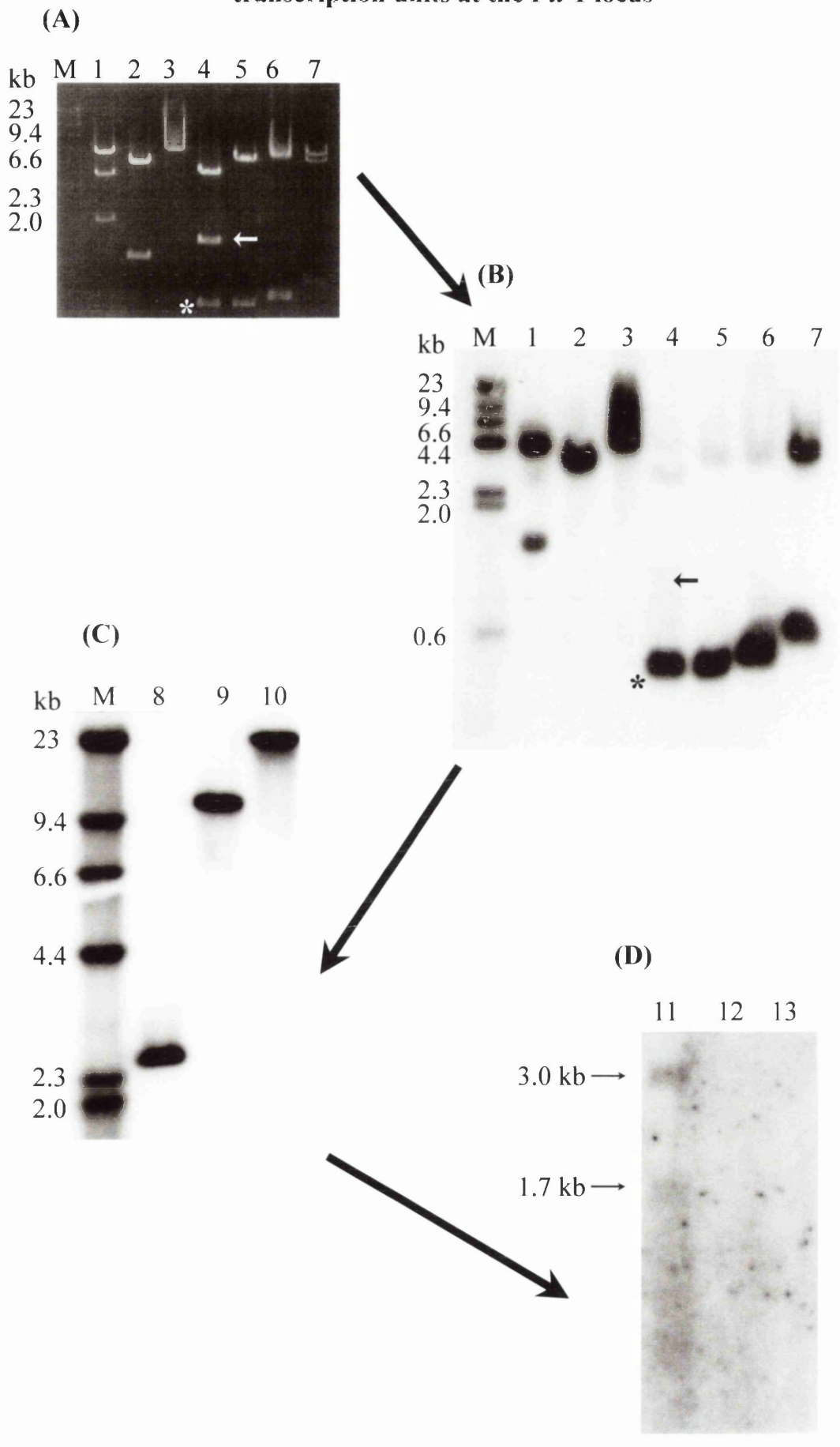


Figure 3.5 Strategy to identify transcription units at the *Fit*-1 locus. (A) Ethidium bromide stained agarose gel of subclone 2 DNA digested with various combinations of restriction endonucleases described below. (B) The Southern blot hybridised with genomic DNA (3201). The asterisks in A and B represent the 400bp *Eco*RI/*Pst*I repetitive sequence and the arrows in A and B highlight the single copy 1kb *Eco*RI/*Pst*I fragment from subclone 2. (C) 3201 cell DNA digested with restriction enzymes (below) and hybridised with the single copy probe identified above. (D) Northern blot of FT1 (11), 3201 (12) and T3 (13) total cellular RNA probed with the single copy probe identified above.
M, λHindIII markers; 1, *Eco*RI/*Sal*I; 2, *Eco*RI/*Xba*I; 3, *Xba*I; 4, *Eco*RI/*Pst*I; 5, *Xba*I/*Pst*I; 6, *Pst*I/*Sal*I; 7, *Xba*I/*Sal*I; 8, *Eco*RI; 9, *Sst*I; 10, *Bam*HI.

2. Detection of transcripts

A DNA probe which hybridised to a single band (or very few bands) on Southern blot analysis of total cat DNA (Figure 3.5C) was then used as a probe on northern blot analysis of FT1 RNA (Figure 3.5D). A probe which hybridises to FT1 RNA on northern would then be used to screen a cDNA library to isolate and characterise the putative gene.

3.2.2 Further analysis of the λRF1 clone failed to identify a transcription unit which was altered as a result of viral insertion at *Fit*-1

Initial mapping of repetitive sequences within the λRF1 clone revealed a number of single and low copy number sequences which were then used as probes on FT1 northern blots to search for any sequences which were transcribed. The results are summarised in Figure 3.6.

The entire subclone 7 which contains the conserved open reading frame has been extensively analysed on northern blot previously, identifying a number of RNAs which hybridise to probes from this region (Figure 3.6, pale blue boxes). However, none of these RNA species were expressed in a tumour specific fashion and therefore were not pursued (R. Fulton, unpublished). Further analysis of FT1 RNA with the conserved *Fit*-1 probe (derived by PCR using the orl f and orl r primers detailed in Appendix 1) and a series of other unique sequence probes from this subclone also failed to show evidence of transcription units which were activated by viral insertion (not shown). Similarly, probes from subclones 3 and 6 did not hybridise to any transcripts on FT1 or control RNAs (not shown).

Probes from subclones 5 (*Pst*I/*Xho*I fragment); A (*Xho*I/*Eco*RI fragment); and 2 (*Eco*RI/*Pst*I fragment, discussed above) hybridised to very faint bands at 3.0kb and 1.7kb on FT1 total RNA (Figure 3.6, dark blue boxes and Figure 3.5D). However, the signal washed off at high stringency, was not reproducible, nor was there any signal seen on Poly $(A)^+$ selected mRNA. If this represented a true gene then one would expect the signal to be enhanced on mRNA.

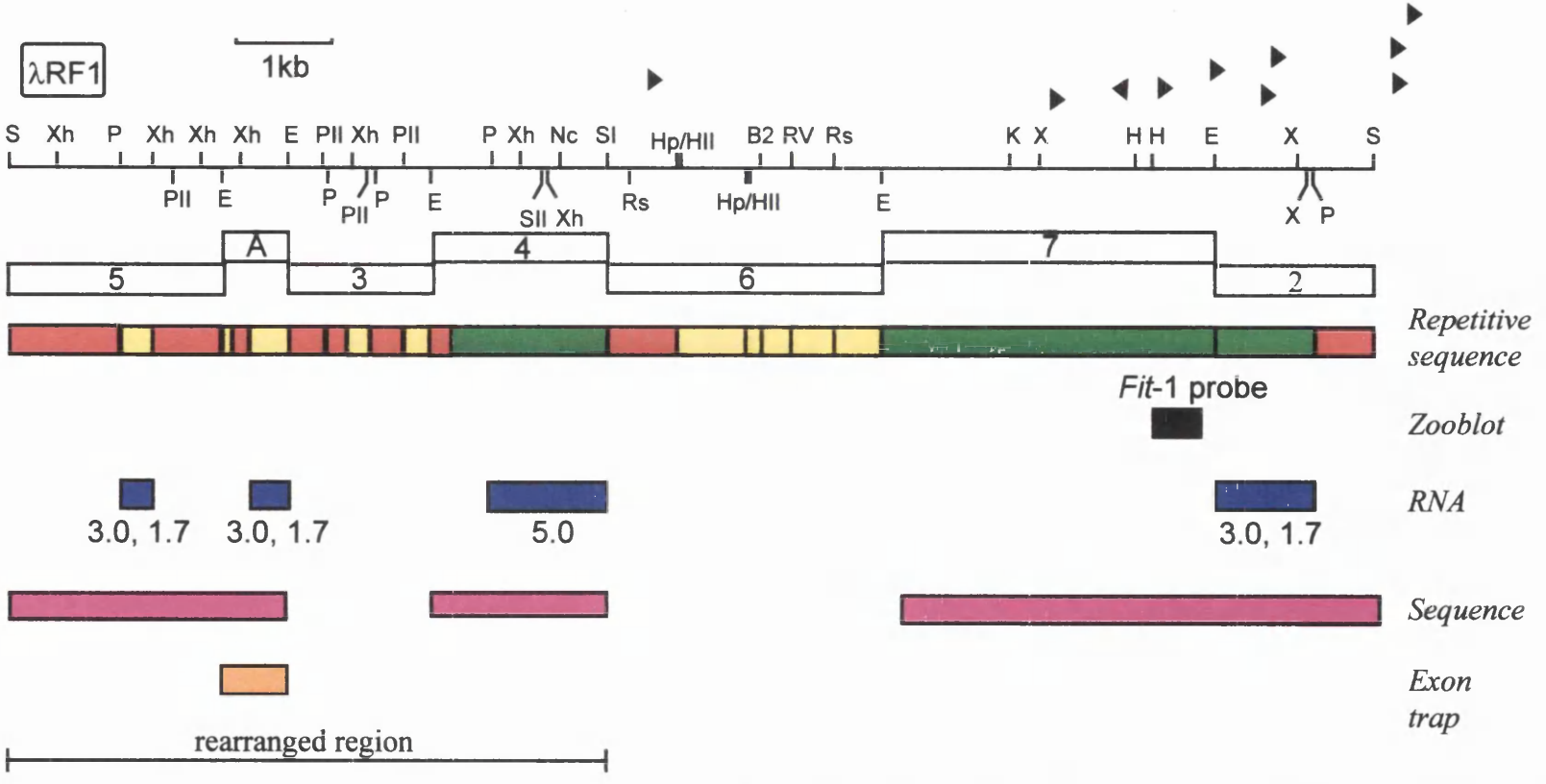


Figure 3.6 Summary of results illustrating the current data from the search for transcribed sequence at the *Fit*-1 locus. The arrowheads indicate the positions and orientations of the integrated FeLV proviruses in individual tumours. The open boxes represent the λRF1 subclones. The red boxes represent highly repetitive DNA sequences, while the yellow and green boxes indicate middle repetitive and single copy sequence respectively. The locations of conserved and transcribed sequences are shown below. Also shown are domains for which DNA sequence is available or which have been identified as putative exons by exon trapping and the 5` rearranged region of λRF1. Restriction enzyme abbreviations: B2, *Bgl*II; E, *Eco*RI; RV, *Eco*RV; HII, *Hin*cII; H, *Hin*dIII; Hp, *Hpa*I; K, *Kpn*I; Nc, *Nco*I; P, *Pst*I; PII, *Pvu*II; Rs, *Rsa*I; S, *Sal*I; SI, *Sst*I; SII, *Sst*II; X, *Xba*I; Xh, *Xho*I.

Subclone A contains two putative exons identified by exon trapping (R. Fulton, unpublished). Therefore, to further characterise sublone A for the presence of functional exons, DNA sequencing of the subclone and northern blot analysis with the putative exons was carried out. The 648bp subclone A was fully sequenced confirming the presence of putative splice acceptor and donor sites defining the supposed exon 1 and exon 2 regions (Figure 3.7, underlined sequences). These putative exons were also found to flank a region with homology to the feline repetitive DNA satellite monomer, FA-SAT. The 483bp FA-SAT monomer is a family of relatively G/C rich sequence which comprise 1-2% of the feline genome and have been found to be localised at the telomeres of chromosomes (158).

In addition, a possible open reading frame was predicted in exon 2, but not from exon 1 through to exon 2 (Figure 3.7). Sequence analysis using the computer algorithm, Gene ID (BioMolecular Engineering Research Center, Boston University, geneid@darwin.bu.edu), predicted with medium probability the presence of a first exon within the putative exon 2 sequence. Furthermore, low copy number probes corresponding to exon 1 (a 69bp oligonucleotide probe, fitA3, Appendix 1) and exon 2 (generated by PCR with the primers fitA1 and fitA2, Appendix 1) hybridised to the faint bands again at 3 and 1.7kb in FT1 total RNA (not shown). However, as before the signal disappeared with a high stringency wash, it was not reproducible, nor was it enhanced on Poly $(A)^+$ selection of mRNA. Moreover, zooblot analyses with both exon 1 and 2 probes were negative (not shown). Together these results suggest that the putative exons 1 and 2 do not represent true exons.

3.2.3 Analysis of subclone 4 reveals that the λRF1 clone is rearranged

Several probes from subclone 4 were found to be single copy by Southern analysis including the *Sst*II/*Nco*I fragment (Figure 3.6). However, this probe gave unexpected results on Southern blot analysis. Therefore, to clarify the restriction map around the λRF1 subclone 4, the restriction maps of FT1 and λRF1 were compared. These results are summarised in Table 3.1.

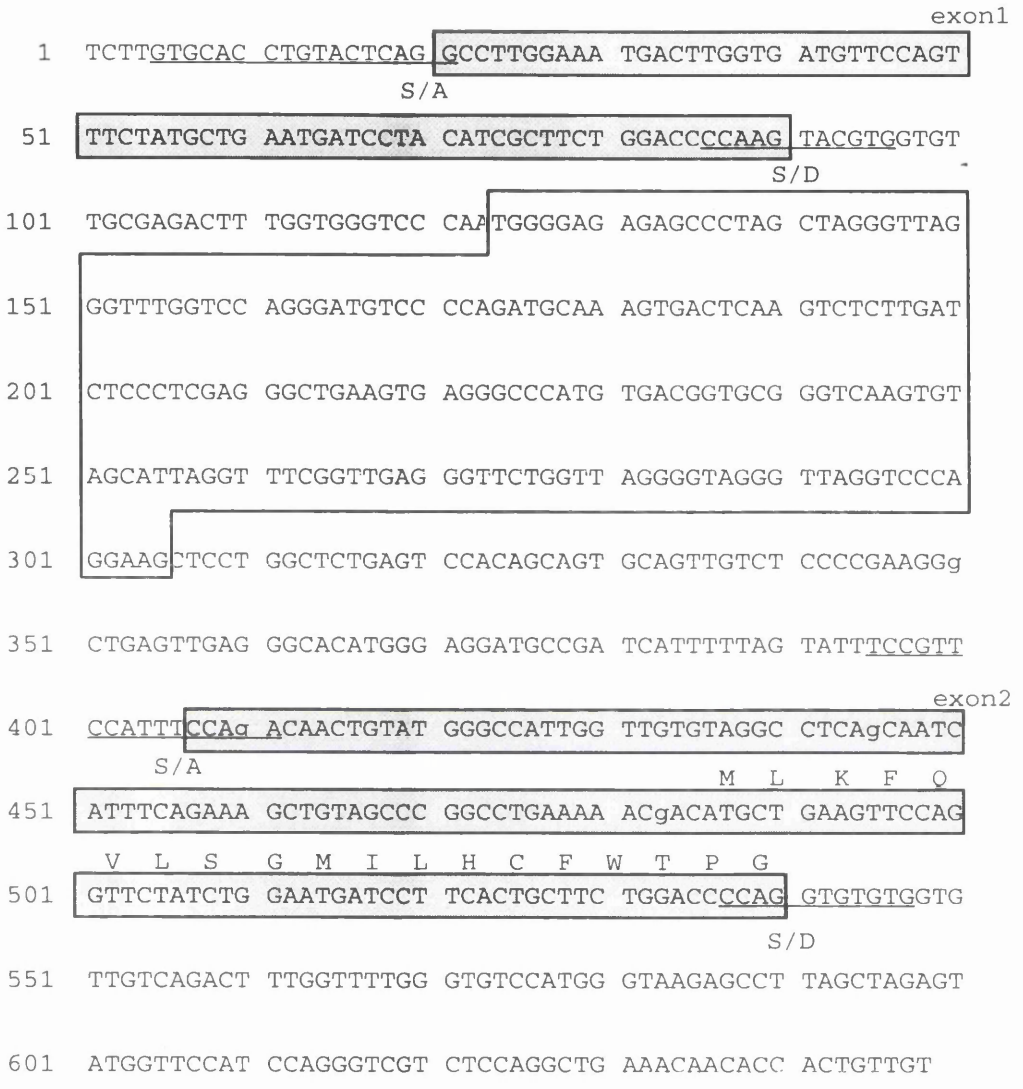


Figure 3.7 Nucleotide sequence and predicted amino acid sequence of the *Fit*-1 subclone A. The shaded areas indicate the 2 potential exons identified by exon trapping, exon1 and exon2. The potential splice acceptor (S/A) sites and splice donor (S/D) sites are underlined. The boxed area indicates the region with homology to the FA-SAT repetitive satellite monomer. The predicted amino acid shown above the nucleotide sequence indicates the potential first exon predicted by the computer algorithm, Gene-ID.

For example, the *Sst*II/*Nco*I probe from the λRF1 subclone 4 hybridises to a 10.5kb *Eco*RI fragment in FT1 genomic DNA. However, the same probe hybridised to a 5.5kb band in λRF1 as predicted from the restriction map of λRF1 shown in Figure 3.6. The same Southern blot, when probed with the conserved *Fit*-1 probe, hybridised to the expected sized fragments in both the λRF1 clone and genomic DNA (not shown). This suggests that the λRF1 clone is chimaeric, possibly due to a recombination event during the original cloning manipulations so that the 5` end of the clone is rearranged (Figure 3.6). Therefore all results 5` of the *Sst*I site at the junction between subclones 4 and 6 may be invalid since this region could potentially be derived from anywhere in the genome.

Table 3.1. Comparison of restriction fragment sizes with the subclone 4 *Sst*II/*Nco*I probe in FT1 and λRF1 DNA.

	*Eco*RI	*Sst*I	*Kpn*I	*Eco*RI/*Sst*I
FT1 DNA	10.5kb	20kb	23kb	7kb
λRF1 DNA	5.5kb	6.5kb	9kb	1.8kb

3.3 Discussion

Ten out of 63 thymic lymphomas screened were rearranged at *Fit*-1, strongly suggesting that a gene involved in tumorigenesis is affected by viral insertion at *Fit*-1. However, despite systematic searching across the entire λRF1 clone for a transcription unit which was altered as a result of viral insertion in the FT1 cell line, no definitive evidence of the presence of a gene at the *Fit*-1 locus was identified (summarised in Figure 3.6). There are a number of possible explanations why none have been detected and these are discussed below.

1. The cluster of proviral insertions at the *Fit*-1 locus may influence a gene which is outwith the region covered by the phage clones. The ability of retroviral insertions to affect gene expression over long distances has been demonstrated (28;41;64;65;159). For example, provirus integration at the *Mlvi*-4 and *Mlvi*-1 loci which map 30kb and 270kb downstream of *Myc* respectively (64;65) have

been shown to enhance *Myc* expression by *cis*-acting mechanisms operating over large distances (64). Clearly, this would make the search for a gene affected by proviral integration at *Fit*-1 a very difficult task.

2. The *Fit*-1 gene may consist of a number of very short exons which could easily be missed by the method utilised to search for transcription units at *Fit*-1. For example, the *Ahi*-1 gene was originally identified as a common integration site in 16% of Abelson pre-B cell lymphomas (17), and has subsequently been found to encode a gene which consists of a number of short exons (P. Jolicoeur, unpublished). Exon trapping was used to identify these short exons, a technique which identified two candidate *Fit*-1 exons (Figure 3.7). DNA sequence from the putative exons 1 and 2 showed no evidence of matches to published sequences, evolutionary conservation or of expression in either normal or tumour RNA, despite the presence of putative splice acceptor and donor sites. Therefore, I concluded that exons 1 and 2 were not true exons. This illustrates the caution which must be used when interpreting results from this system since spurious splicing events may occur at cryptic splice acceptor and donor sites present in the sequence which is being analysed.

3. Proviral integrations at *Fit*-1 may result in inactivation of gene expression which may have been missed by northern blot analysis. Such inactivation by proviral insertion has been observed in myeloid leukaemias of BXH-2 mice where integration in both alleles of *Evi*-2 results in the inactivation of the NF1 tumour suppressor protein (33). It is unlikely that integration at *Fit*-1 inactivates a tumour suppressor because rearrangements have not been observed in both alleles of *Fit*-1 in any of the tumours or cell lines (42). However, it is possible that one allele has been inactivated by proviral integration while the other allele has undergone additional (epi)genetic mutations resulting in its inactivation. This could occur in a number of ways including point mutation or hypermethylation of the DNA which has been associated with gene inactivity (160).

4. Another possible explanation for the lack of an obvious transcription unit at the *Fit*-1 locus is that it may be present only at very low levels which are undetectable by northern blot analysis. An alternative approach to examine this would be to perform PCR on reverse transcribed cDNA (RT-PCR).

CHAPTER 4

4. Cloning and characterisation of the human homologue of *Fit*-1

4.1 Introduction

16% of feline thymic lymphomas screened were found to be rearranged at *Fit*-1 (25;42) suggesting that a gene involved in tumorigenesis is affected by viral insertion at *Fit*-1. However, no evidence of a transcription unit activated by proviral insertion has been found in the 30kb of DNA flanking the insertion cluster (chapter 3 and H. Tsujimoto, unpublished).

As discussed, a gene which is distant from the insertion cluster may be the target of viral integration at *Fit*-1. However, because of the constraints of working with a species where relatively little is known about its genome, and in view of the conservation of the *Fit*-1 domain across species, it was decided to isolate the human homologue of *Fit*-1. There is a far greater wealth of resources available for large-scale genomic analysis in humans compared to cats, primarily owing to the Human Genome Project. Therefore, by isolating large human genomic clones, the search for adjacent genes in human rather than feline DNA could be extended.

4.2 Results

4.2.1 Isolation of a human cosmid clone containing the *Fit*-1 region

Fit-1 has been mapped to feline chromosome B2 (42), much of which has been shown to be homologous to human chromosome 6 (161). Therefore, to isolate the human homologue of *Fit*-1, a human chromosome 6 specific cosmid library (RessourcenZentrum/PrimarDatenbank (RZPD), Berlin, Germany (162)) was screened using the conserved feline *Fit*-1 probe.

The library consists of a single gridded high density hybridisation filter, representing approximately 36,844 clones with an average insert size of 40-50kb. The filter was hybridised essentially as described in section 2.2.8 using the

conserved *Fit*-1 probe (not shown). One confirmed positive clone was identified; clone ICRFc1090188Q5 (cosmid 1).

4.2.2 Restriction map of the *Fit*-1 positive human chromosome 6 cosmid clone

A crude restriction map of the human *Fit*-1 region was generated by digesting DNA from the 45kb cosmid 1 clone with various restriction enzymes. Restriction analysis was performed with a series of single and double digests using the restriction endonucleases *Sal*I, *Xho*I, *Cla*I, *Xba*I, *Sst*I, *Pst*I, *Hin*cII, *Hin*dIII and *Eco*RI. A Southern blot of the size separated DNA fragments was hybridised with the conserved *Fit*-1 probe (Figure 4.1). These restriction mapping data are summarised in Figure 4.2A.

To generate a detailed partial restriction map flanking the region hybridising to the *Fit*-1 probe in the human genome, an 11.5kb *Eco*RI fragment was subcloned in pUC18 and characterised by restriction mapping with the following restriction enzymes: *Sst*I, *Hin*dIII, *Pst*I, *Xba*I, *Pvu*II, *Hin*cII and *Eco*RV and hybridised with the conserved *Fit*-1 probe (not shown). The flanking sequences of the *Fit*-1 probe were mapped to a 1.2kb *Pvu*II fragment (Figure 4.2B). In this physical map only the closest restriction sites to the *Fit*-1 probe could be identified with a number of enzymes (indicated in parenthesis).

4.2.3 The feline *Fit*-1 open reading frame is conserved in humans

In the cat, a putative long ORF was identified within λRF1 subclone 7; the *Fit*-1 probe (R. Fulton, unpublished and Figure 3.6). To isolate the region of the human genome which has homology to the feline *Fit*-1 probe and to establish if this putative ORF was maintained, the 1.2kb *Pvu*II fragment which hybridised to the *Fit*-1 probe was subcloned and sequenced. This fragment was blunt end cloned in pBluescript SK^+ as described in section 2.2.7. Positive colonies containing the *Fit*-1 homologous *Pvu*II fragment were identified by Grunstein and Hogness hybridisation analysis (153). A *Fit*-1 positive clone was sequenced with the M13 universal forward and reverse primers (Appendix 1) as described in section 2.2.9.

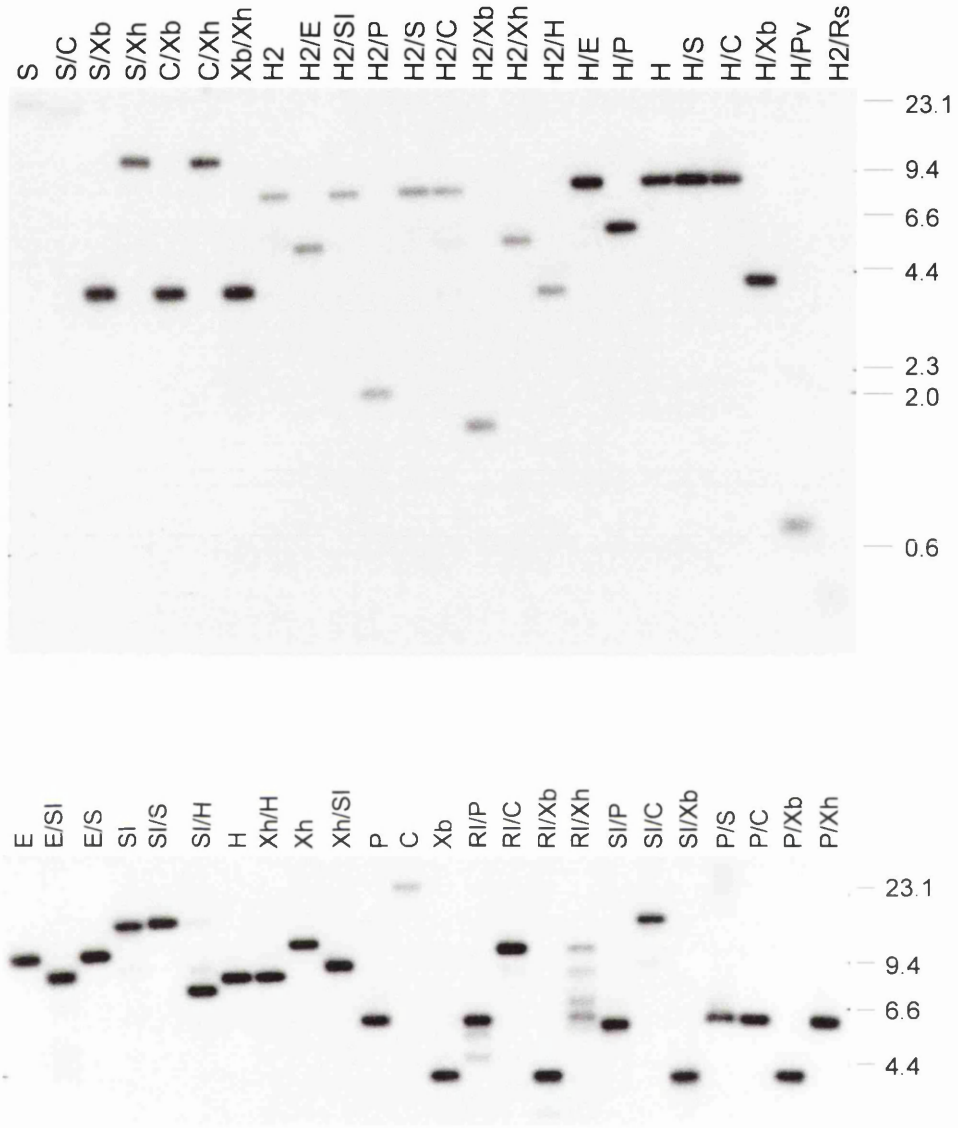


Figure 4.1 Southern blot analysis and restriction mapping of cosmid 1. 1μg of cosmid 1 DNA was digested with the restriction enzymes shown, separated on 0.6% agarose gels and hyridised with the human *Fit*-1 probe. Southern blots were washed at high stringency (0.1 X SSC, 0.5% SDS, 60oC). Molecular size markers were λ *Hin*dIII restriction fragments and are indicated in kilobases. Restriction enzyme abbreviations: C, *Cla*I; E, *Eco*RI; H2, *Hin*cII; H, *Hin*dIII; P, *Pst*I; Pv, *Pvu*II; Rs, *Rsa*I; S, *Sal*I; SI, *Sst*I; Xb, *Xba*I; Xh, *Xho*I

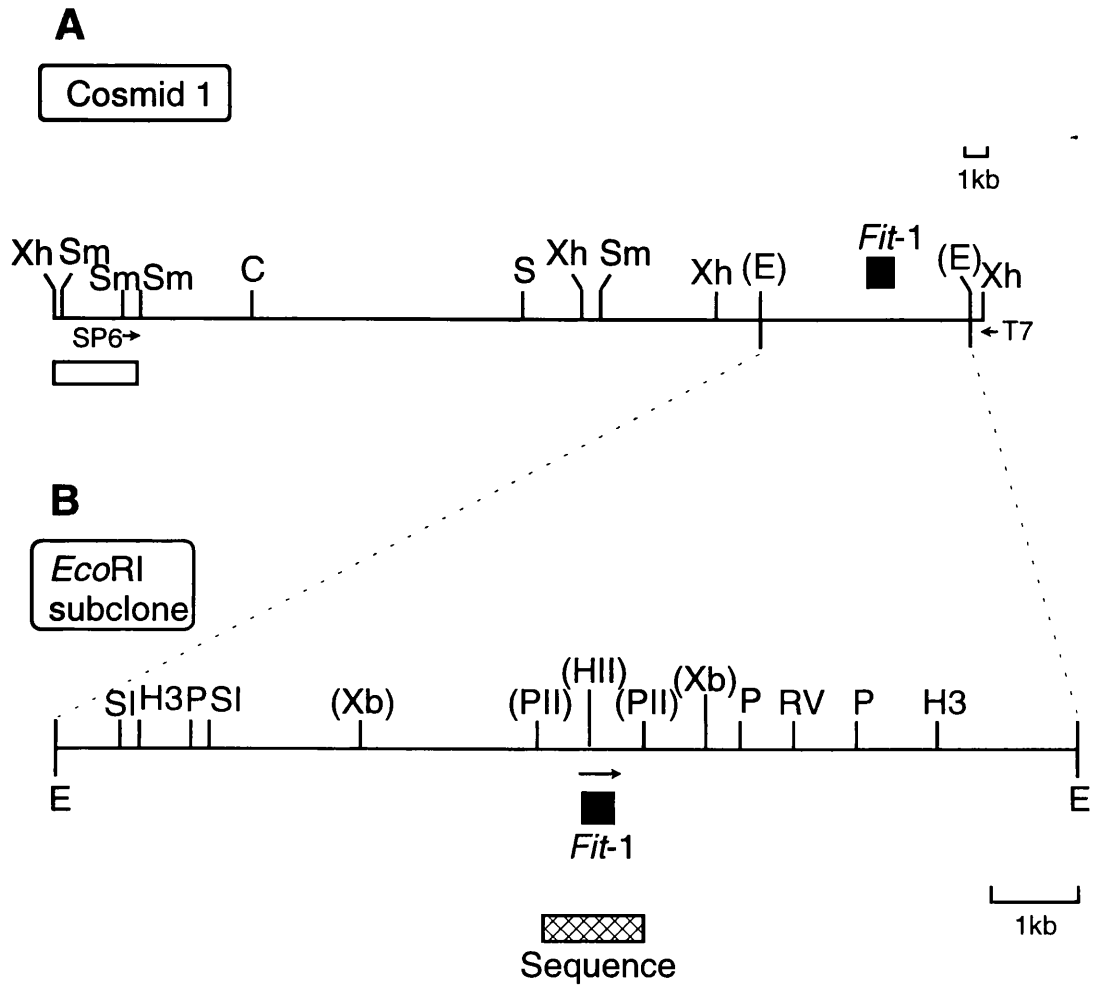


Figure 4.2 Restriction map of the human *Fit*-1 region from the human chromosome 6 cosmid clone, cosmid 1 (A); and the *Eco*RI subclone (B). The position of the *Fit*-1 probe is indicated (closed box), the arrow indicates the direction of the open reading frame. The location of the vector (Lawrist 4) is also indicated (open box) with the positions and orientations of the SP6 and T7 primer sequences. The hatched box represents the *Fit*-1 positive *Pvu*II subclone which has been fully sequenced. Restriction enzyme abbreviations: C, *Cla*I; E, *Eco*RI; HII, *Hin*cII; H3, *Hin*dIII; P, *Pst*I; PII, *Pvu*II; RV, EcoRV; S, *Sal*I; SI; *Sst*I; Xb, *Xba*I; Xh, *Xho*I. Partial restriction sites are indicated in parenthesis.

A comparison of the DNA sequence from the feline and human *Fit*-1 domains was carried out using the GCG (Genetics Computer Group, Wisconsin) computer algorithm. This alignment reveals that the human *Fit*-1 sequence is approximately 80% identical to the feline DNA sequence over 517bp (Figure 4.3A and (157)). Moreover, the potential splice acceptor and donor signals are highly conserved between the cat and human DNA sequences. At the amino acid level, the open reading frame present in the feline peptide sequence is maintained in the human genome (Figure 4.3B). The feline and human peptide sequences are 76% identical over 41 amino acids at the start of the sequence falling to 16% identity over 25 amino acids in the middle of the sequence. The homology rises again to 77% over 22 amino acids at the C terminus. Furthermore, a number of the changes are conservative amino acid substitutions.

To generate a human *Fit*-1 probe, PCR was performed using the *Pvu*II clone as template and primers designed from this sequence: primers huorf-f and huorf-r (Appendix 1 and Figure 4.3A). Both the human and the feline *Fit*-1 probes hybridise to the same fragment on a feline genomic DNA Southern blot (not shown) confirming that this probe represents the true human equivalent of the cat locus.

4.2.4 The human *Fit*-1 probe is conserved but is not transcribed

The feline *Fit*-1 probe was previously shown to be conserved in mammalian evolution (157). Therefore, owing to the homology of the human and feline *Fit*-1 DNA sequences, one would expect the human *Fit*-1 probe to show evidence of cross species conservation. Consequently, zooblot analysis was performed on DNA from various species with with the human *Fit*-1 probe (Figure 4.4). This identified discrete bands in human, monkey, mouse, dog, cow and rabbit DNA which hybridised to the probe, thus confirming the conservation of the human *Fit*-1 locus. This analysis suggests that it might play an important role in the biological function of the cell. However, despite this intriguing conserved open reading frame at *Fit*-1, no transcripts were identified on northern blot analysis

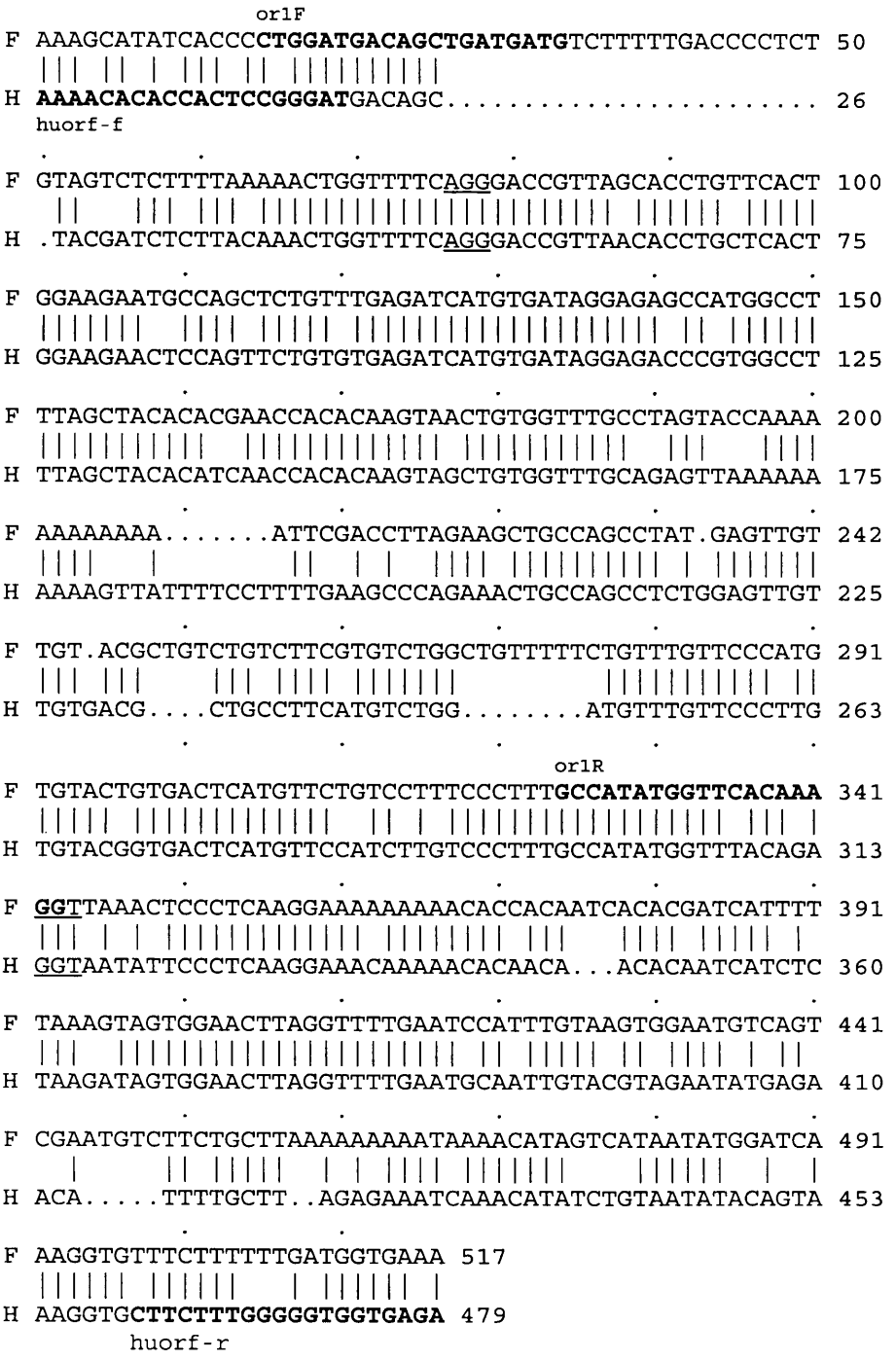


Figure 4.3A. DNA sequence comparison of the human and feline *Fit*-1 regions. A comparison of the DNA sequences between the feline (F) and human (H) *Fit*-1 DNA sequence derived from the *Pvu*II *Fit*-1 positive subclone is shown. Accession numbers: feline *Fit*-1, AF113885; human *Fit*-1, AF113886. Gaps have been introduced for maximum alignment. The putative splice donor and splice acceptor sites which are conserved between feline and human are underlined; the huorf-f and huorf-r oligonucleotide primers which amplify the human *Fit*-1 probe and the or1F and or1R primers which amplify the feline *Fit*-1 probe are indicated in bold.

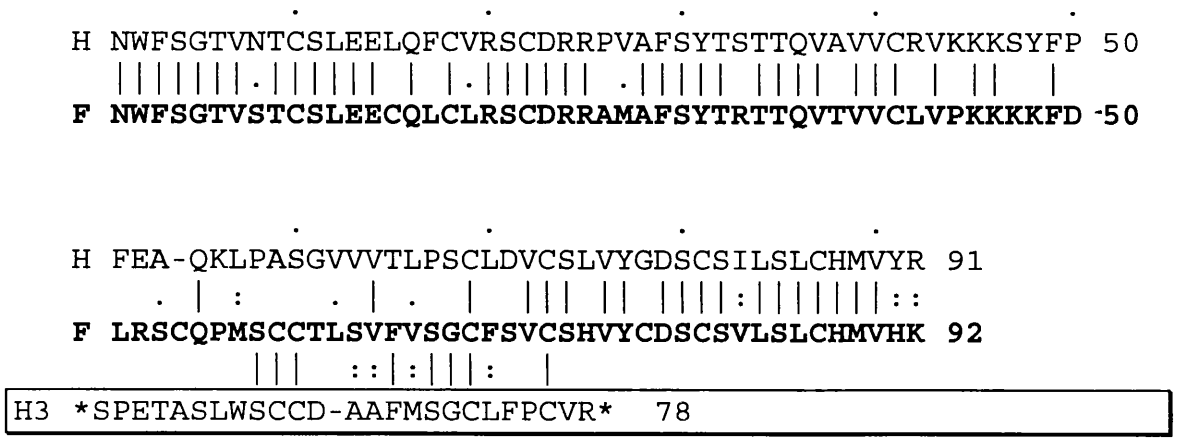


Figure 4.3B. Predicted peptide sequence comparison of the human and feline *Fit*-1 regions. A comparison of the predicted peptide sequences from the feline (F) *Fit*-1 open reading frame (bold type) and the human (H) *Pvu*II clone. The upper sequence is the predicted human peptide sequence in reading frame 1. The boxed sequence is the predicted human peptide sequence in reading frame 3 (H3). Gaps have been introduced for maximum alignment. Key: (|), identity; (:), conservative amino acid change; (.), neutral change. This figure shows that the open reading frame in the feline *Fit*-1 sequence is conserved in the human, however maximum alignment requires a short frame shift from reading frame 1 to 3.

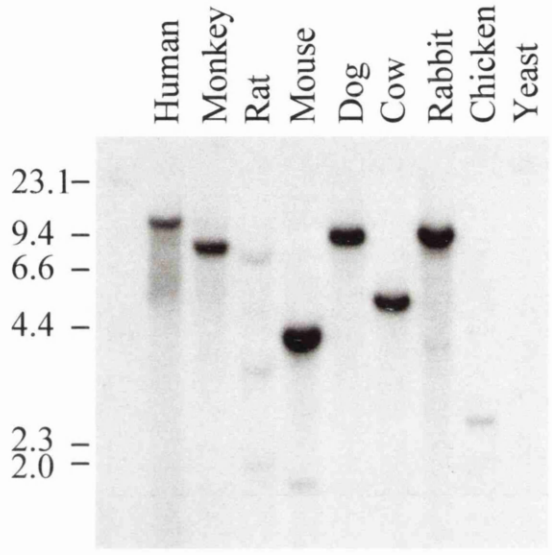


Figure 4.4 Zooblot analysis with the human *Fit*-1 probe. Interspecies zooblot (Clontech Laboratories, Inc.) containing 4μg of genomic DNA per lane digested with *Eco*RI was probed with the human *Fit*-1 probe. The blot was washed at low stringency (three times in 2 X SSC, 0.5% SDS, 60oC). The sizes of the λHindIII DNA size markers are indicated (kb).

using poly A^+ selected mRNA extracted from various human T cell leukemia lines (not shown).

4.2.5 Relevance of *Fit*-1 to human tumours

As discussed in section 3.1, *Fit*-1 has been mapped to feline chromosome B2 (42) and mouse chromosome 10, 1cM from *Ahi*-1, a common insertion locus which has been mapped 35kb downstream of c-*myb* (18;157). The localisation of *Fit*-1 to human chromosome 6 has now been confirmed by the isolation of the homologous region from a human chromosome 6 cosmid library. Thus, we can infer that the likely position of *Fit*-1 on chromosome 6 is somewhere in the vicinity of c-*myb* at 6q23.3-q24 (163). The long arm of chromosome 6 (6q) is often affected by genetic changes where deletions of 6q (6q-) are among the most frequent chromosomal aberrations observed in malignant lymphomas and lymphoblastic leukaemias (164;165). This suggests the existence of sequences involved in the control of cellular growth within the long arm of chromosome 6.

A human T cell leukaemia line which is not 6q-, the CEM cell line, has been reported to have chromosomal changes at 6q and rearrangements at the c-*myb* promoter region (166). Therefore, to establish whether the CEM cell line is also rearranged at *Fit*-1, a Southern blot of CEM and control (Jurkat) DNA digested with *Pst*I, *Eco*RI and *Sst*I was hybridised with the human *Fit*-1 probe (Figure 4.4A). Interestingly, this identified two closely linked restriction fragment length polymorphisms (RFLP) in CEM DNA at the *Fit*-1 locus. The *Pst*I genomic fragment which hybridises to *Fit*-1 in Jurkat DNA is 6.5kb, whereas in CEM DNA there is an additional *Fit*-1 hybridising fragment of approximately 9kb (Figure 4.5A, arrowhead). Furthermore, the *Eco*RI/*Sst*I double digest has an additional *Fit*-1 hybridising band at 10kb which is not present in the control (Figure 4.5A, arrow). This suggests that both the *Pst*I and the *Sst*I restriction enzyme sites are lost in a subset of CEM cells. These data are summarised in Figure 4.5B.

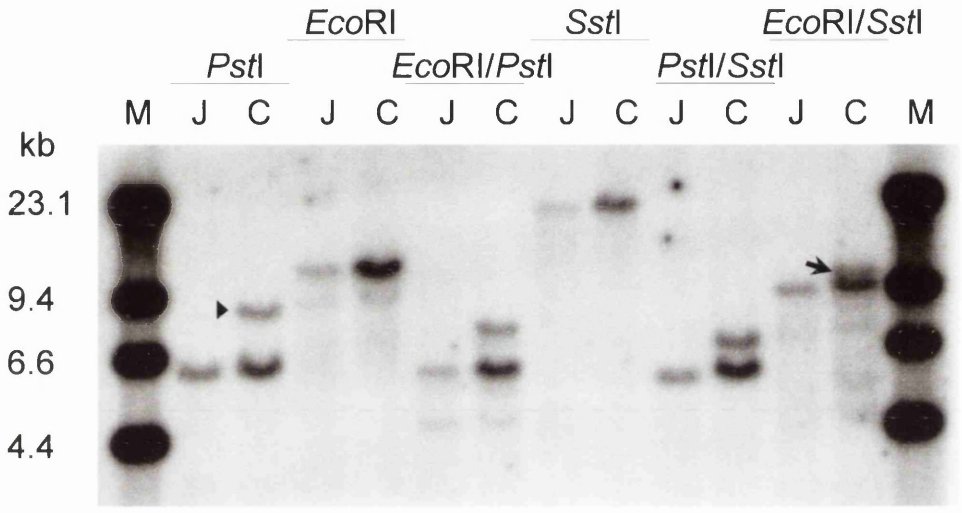


Figure 4.5A. The CEM cell line has two RFLPs at *Fit*-1. 20μg of CEM (C) and control (Jurkat, J) DNA was digested with the restriction enzymes as shown and hybridised with the human *Fit*-1 probe. Southern blots were washed at high stringency (0.1 X SSC, 0.5% SDS, 60oC). Molecular size markers were λ HindIII restriction fragments and are indicated in kilobases. This figure demonstrates that CEM DNA has an RFLP at *Fit*-1 with a *Pst*I digest (arrowhead), and with *Sst*I (arrow).

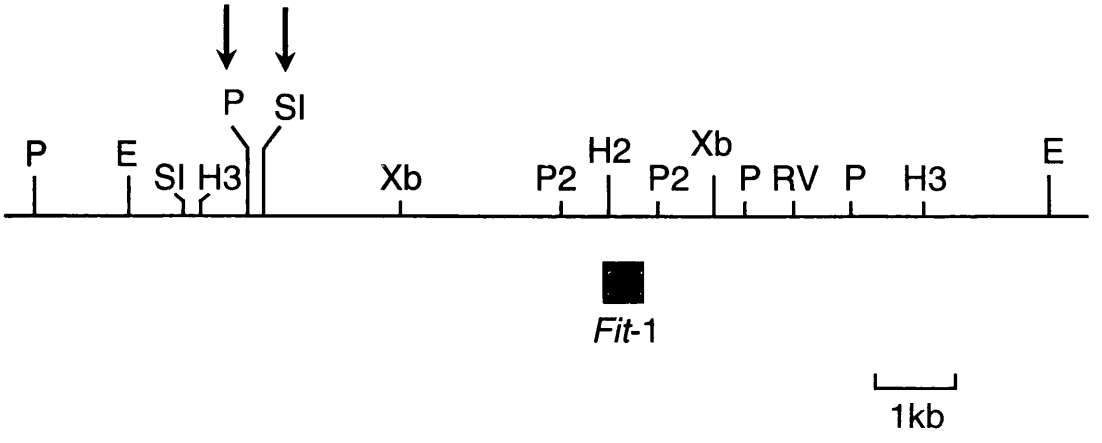


Figure 4.5B. Partial restriction map of the human *Fit*-1 region highlighting the positions of the *Pst*I and *Sst*I sites which are lost in a subset of CEM cells (arrows). The position of the *Fit*-1 probe is also illustrated (closed box). Restriction enzyme abbreviations: E, *Eco*RI; H2, *Hin*cII; H3, *Hin*dIII; P, *Pst*I; P2, *Pvu*II; RV, *Eco*RV; SI; *Sst*I; Xb, *Xba*I.

To further characterise the apparent rearrangement seen in CEM DNA with the *Fit*-1 probe, CEM DNA was analysed by Southern blot with a number of restriction enzymes and hybridised with the human *Fit*-1 probe. This analysis failed to reveal any further RFLPs at *Fit*-1 in CEM DNA (not shown). Furthermore, cloning of the region surrounding the *Pst*I and the *Sst*I restriction sites in CEM DNA, and subsequent sequence analysis confirmed that these RFLPs were simple base change polymorphisms, only differing from Jurkat DNA at the *Pst*I and/or *Sst*I restriction enzyme sites (Table 4.1).

Table 4.1 Summary of the CEM clones which have a *Pst*I and *Sst*I restriction enzyme site.

clone	*Pst*I site	*Sst*I site
5.1	-	-
5.3	+	?[a]
5.4	-	-
30.1	-	+
30.3	+	+
8f	-	-

[a] clone 5.3 did not extend as far as the *Sst*I genomic site.

4.3 Discussion

In order to extend the search for adjacent genes in a species where there are greater resources available for genomic analysis, the human *Fit*-1 homologous region was cloned from a chromosome 6 cosmid library. From analysis of DNA sequence from a *Pvu*II subclone homologous to *Fit*-1, the putative long open reading frame identified in the cat within the λRF1 subclone 7 (Figure 3.6) was found to be conserved in humans (Figure 4.3B). Apart from a short frame shift, the differences were mainly conservative amino acid changes, further emphasising the importance of the *Fit*-1 region. Despite this high degree of conservation, the human *Fit*-1 probe did not identify any transcripts on northern blot analysis. However, the human *Fit*-1 probe will serve as a useful tool to identify novel proto-oncogene(s) which are closely linked on chromosome 6.

The long arm of chromosome 6 is the site of frequent genetic alterations, most of which are concentrated around 6q23 (164;165;167-169), the likely position of human *Fit*-1 based on linkage analysis in the mouse (157). Together, these examples lend support to the presence of a tumour suppressor gene or a gene involved in the control of cellular growth at this region of chromosome 6.

Although it has been reported that c-*myb* gene activation is not involved in 6q- abnormalities of lymphoid tumours (170), an acute lymphoblastic leukaemia cell line (CEM) has been identified with a submicroscopic deletion within chromosome 6q involving c-*myb* (166). This deletion results in the loss of a highly conserved region of the c-*myb* promoter in this cell line. However, no effect has been noted on c-*myb* promoter activity. Therefore, it is unclear what the effect of this deletion is. It is possible that it may induce subtle changes in the expression of c-*myb* or that it may have an effect on another loci in this region of chromosome 6.

Therefore, the involvement of *Fit*-1 in the CEM cell line was investigated revealing an apparent rearrangement at *Fit*-1 with *Pst*I and *Sst*I digests (Figure 4.5A). This was of great interest because of the relatively close proximity of c-*myb* to *Fit*-1 implied by linkage analysis in the mouse. However, a series of additional digests failed to reveal any further rearranging bands with the *Fit*-1 probe in CEM DNA. Furthermore, sequence analysis showed that CEM sequence only diverged from control (Jurkat) sequence by a single base substitution at either or both of the closely linked *Pst*I or *Sst*I restriction sites. Although this RFLP was found to be due to a simple base change polymorphism in the restriction sites, a possible role for *Fit*-1 in the CEM cell line is still worth bearing in mind.

CHAPTER 5

Chapter 5

5. Generation of a cosmid and PAC contig of the human *Fit*-1 region

5.1 Introduction

The conserved feline *Fit*-1 probe and a series of other unique sequence probes from 30kb of DNA flanking the insertion cluster failed to show any evidence of a transcription unit which was activated as a result of viral insertion. Furthermore, the human *Fit*-1 probe also failed to identify transcripts in either normal or tumour RNA. This was surprising considering that 16% of the feline lymphomas screened were rearranged at *Fit*-1, suggesting that this region of the chromosome harbours a gene which is involved in lymphomagenesis.

However, as alluded to in chapter 3, proviral integration at *Fit*-1 may influence a gene which is outwith the region covered by the phage or cosmid clones. Therefore, I entered upon a wider gene hunting investigation in the human genome by generating a contiguous physical map around the *Fit*-1 locus. Initially this was achieved by screening a chromosome 6 specific cosmid library for cosmid clones which overlap with cosmid 1, and subsequently by isolating overlapping PAC clones which can accommodate much larger inserts.

5.1.1 Contig assembly and genomic mapping studies

A number of cloning vectors have been designed for large scale genomic mapping studies of complex genomes. The relative merits of each of these are summarised in Table 5.1. A major consideration for the choice of cloning vector is the size of exogenous insert DNA that is maintained. YACs (yeast artificial chromosomes) were developed as high-capacity cloning vectors that can be transformed into yeast where they are maintained as artificial chromosomes (171). YACs can propagate exogenous DNA of up to 1,000kb in length. Whereas, cosmids can only accommodate an average of 40kb of insert DNA. Of intermediate size are BAC (bacterial artificial chromosmes) and P1 derived bacteriophage based cloning systems. BAC vectors are based on the *E.coli* fertility plasmid (F factor) and can propagate exogenous DNA inserts of up to

300kb (172). The traditional bacteriophage P1 packaging system can accept inserts of 70-100kb in length (173). More recently, a modification of the bacteriophage P1 cloning vector was developed, the PAC (P1-derived artificial chromosome), which can carry inserts of up to 300kb, but more typically around 130-150kb (174). Each of these cloning vectors have their advantages and disadvantages which are summarised in Table 5.1.

Table 5.1 The relative merits of various vehicles of large scale genomic mapping studies

Cloning vector	*Advantages*	*Disadvantages*	*Reference*
YAC	Very large inserts (up to 1000kb).	Low transformation efficiency. Chimaerism [a]. Insert instability. Difficulties in DNA manipulation.	(171;175)
BAC	Large inserts (> 300kb). No chimaerism or instability. Electroporation [b]. Easy DNA manipulation [c].	DNA recovery low. No positive selection of recombinants.	(172)
PAC	Large inserts (100-300kb). No chimaerism or instability. Electroporation [b]. Easy DNA manipulation [c]. Can select for recombinants.	Small insert size Relative to YACs.	(174)
P1 derived	Reasonably large inserts (70-100kb). Can select for recombinants. No chimaerism or instability.	Smaller insert size than PACs. Elaborate *in vitro* packaging system required.	(173)
Cosmid	Relatively small inserts (35-50kb). Easy DNA manipulation. Chromosme specific libraries available.	Relatively small inserts (35-50kb). Frequent insert instability.	(176)

[a] Chimaeric clones are generated by either co-ligation of unlinked sequences, or by *in vivo* recombination following the transformation of multiple YACs into the same yeast cell (175).
[b] The advantage of electroporation based transfer of DNA is that it avoids the use of elaborate packaging systems and reduces site-specific recombination events.
[c] Transformation in *E.coli* allows for simple alkaline lysis based DNA recovery (89).

5.1.2 Sequence-tagged sites (STS) and genomic mapping

Following the identification of a clone from a genomic library which contains the sequence of interest, a contiguous physical map surrounding this region can be generated. A number of fingerprinting procedures have been utilised to determine overlap and to generate an ordered representation of clones (a so-called contig).

The mapping of complex genomes has been revolutionised by the advent of sequence-tagged site (STS) content mapping as a strategy for identifying clones which overlap (177). STSs are single copy sequences of typically 200-500bp in length which are used to construct genetic and physical maps of the genome. Each site is defined by a unique set of PCR primers which amplify a specific single copy product under specified PCR conditions (reviewed in (178)). STS content mapping is based on the assumption that any two clones which share the same STS must overlap. Thus, any unique sequence in the genome can be used as an STS to identify overlapping clones which represent a specific region of the genome. Furthermore, because a STS is simply described by a set of specific PCR primers, STS based mapping allows for community access through a database (177).

A parallel to STS content mapping involves the use of primers for *Alu* sequences (*Alu*-PCR) which was first described by Nelson (179). Inter-*Alu* PCR is a method whereby DNA sequences which fall between *Alu* repeats in the human genome are amplified to generate random probes or STSs which can be used to identify clones that overlap. In this way, two clones which share a similar pattern of *Alu*-PCR fragments after gel electrophoresis must overlap. Similarly, in chromosomal walking, *Alu* specific primers and primers derived from the vector can be used to amplify sequence from the end of a particular clone, again to identify overlapping clones (178;179).

The ultimate goal of the Human Genome Project is to sequence the entire human genome. Toward this end, genetic and physical maps of the chromosomes are being constructed using techniques such as those discussed above. A vast and ever increasing number of genomic regions have been mapped, perhaps most notably a YAC contig map which covers 75% of the human genome (180). Emphasis in the earlier stages has been to characterise regions containing genes. There are various ways in which regions coding for a gene may be identified such as those discussed in section 1.2.1. In addition, ESTs (expressed sequence tags), which are STSs derived from a cDNA clone, have also aided in the mapping of the human genome and have the added advantage of directly pointing to an expressed gene (181).

5.2 Methods

5.2.1 Screening of a PAC library by PCR

The human PAC library (number 704) from RZPD, consists of 39 primary pools of clones which have been pooled in a three dimensional system from a series of 96 well microtitre plates. The library is screened by performing PCR on 1μl of primary pool using standard cycling conditions (section 2.2.6) and either the huorf-f / huorf-r primers (to amplify the human *Fit*-1 probe sequence) or the cos1SP66 / cos1SP67 primers (to amplify the cos7 probe sequence). Positive pools were identified by visualisation of the PCR product following gel electrophoresis and ethidium bromide staining. A positive pool was confirmed by hybridisation of a Southern blot of the PCR products with a probe derived from the PCR primers used (either the human *Fit*-1 probe or the cos7 probe). Based on the positves identified from the primary pools, secondary pools are then obtained from RZPD and a second round of PCR screening performed. In this way the specific X, Y coordinates of the corresponding clone(s) were identified and requested from RZPD. Positive clones were confirmed by Grunstein and Hogness colony hybridisation analysis (153).

5.2.2 Pulsed Field Gel Electrophoresis (PFGE)

1μg of PAC DNA was digested with various combinations of rare-cutting restriction enzymes according to the manufacturer's instructions and loaded into each well of a 200ml 1% agarose gel in 0.5 X TBE. Pulsed-field gel electrophoresis (CHEF-DR II Pulsed Field Electrophoresis Systems, Bio-Rad) was run at 6V/cm for 16 h with a pulse time of 1-6 s in 0.5 X TBE kept at a constant temperature of 14^0C. As molecular weight markers, 5Kb Ladder, λ 50Kb Ladder (Bio-Rad) and λ *Hind*III fragments (GibcoBRL) were used. Following electrophoresis, the gels were stained with ethidium bromide.

5.2.3 Southern blot transfer of large DNA fragments

Since large DNA fragments cannot be transferred efficiently, DNA fragments separated by PFGE were cleaved before transfer onto membranes by UV irradiation at 60mJoules (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation). Following cleavage, the size separated fragments were equilibrated in transfer buffer (0.4N NaOH, 1.5M NaCl) for 15 min and the DNA transferred onto Hybond-N^+ (Amersham) by capillary blotting in transfer buffer for at least 48 hours. The membrane was neutralised in 0.5M Tris-HCl pH7.0 for 5 min, followed by equilibration in 2 X SSC. The complete transfer of DNA from the gel to the membrane was confirmed by staining the gel post transfer with ethidium bromide and viewed on a short wave UV transilluminator (UVP Inc.). Specific DNA fragments were detected by hybridisation as described previously (section 2.2.8.5) using the human *Fit*-1 or c-*myb* exon 2 probes.

5.3 Results

5.3.1 Isolation of an overlapping cosmid clone

In order to extend the search for genes flanking the *Fit*-1 locus in the human genome, the RZPD chromosome 6 cosmid library was screened for clones which overlap with the *Fit*-1 positive cosmid clone (cosmid 1). To this end, probes from the termini of cosmid 1 were isolated and used to re-probe the cosmid library filter. The cosmid vector (Lawrist 4) contains the SP6 and T7 promoter primer sites which flank the cloning site. Therefore, by performing cycle sequencing reactions with IRD41 fluorescent labelled SP6 and T7 primers (see

section 2.9) and using cosmid 1 DNA as template, DNA sequence from the termini of the cosmid 1 DNA insert was obtained (not shown). By comparing the DNA sequence with published sequences (NCBI Blast search), repetitive elements, such as *Alu* repeats, were identified and excluded. Any sequence which did not reveal any homology with known repetitive elements was used to design STS PCR primers. If the sequence obtained with the SP6 or T7 primers was repetitive, another round of cycle sequencing was performed by designing primers based on the available sequence. Figure 5.1 shows two sets of STS PCR primers from cosmid 1 termini: cos1SP66 and cos1SP67 from the SP6 end of the clone; and cos1T75 and cos1T76 from the T7 end of the clone (Appendix 1).

Following the identification of the STS primers described above, the DNA between these primers was amplified by PCR (Figure 5.1) using standard conditions (section 2.2.6). To confirm that these sequences were single copy, the PCR products were gel purified, labelled with [α^{32}P] dCTP and used to probe a Southern blot of total human DNA (Jurkat) digested with various restriction enzymes. Two probes, cos1SP667 and cos1T756 from the SP6 and T7 termini of cosmid 1 respectively (Figure 5.1) were identified and confirmed as single copy by Southern blot analysis. However, when used to probe the RZPD chromosome 6 cosmid library filter, only the cos1SP667 probe from the SP6 terminus of cosmid 1 identified an overlapping clone, ICRFc109B0915Q5 (cosmid 7), and from then on this probe was renamed the cos7 probe. Thus, the generation of an overlapping contig extending in both directions from the *Fit*-1 region was not possible.

5.3.2 Isolation of overlapping PAC clones

To extend genomic mapping in both directions around the *Fit*-1 region, it was decided to isolate PAC clones positive for *Fit*-1 and/or cos7. PACs offer several advantages over cosmids, most notably the significantly larger insert size which can be accommodated compared to cosmids (Table 5.1). The human PAC library (number 704) from RZPD was screened by two rounds of PCR as outlined in section 5.2.1. The PCR primers used were huorf-f and huorf-r (Appendix 1) to

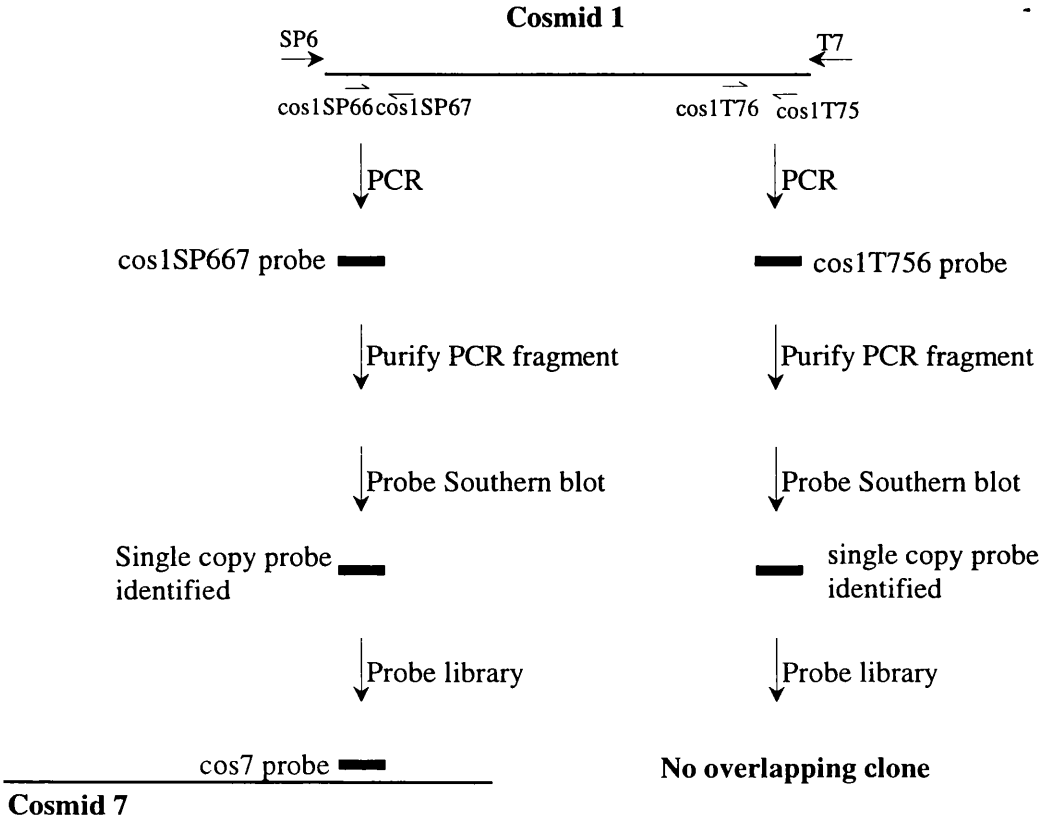


Figure 5.1 Strategy to identify clones which overlap with cosmid 1. Not to scale. The arrows indicate the positions of the SP6 or T7 cycle sequencing primers from the vector sequence. The half arrows indicate STS primers (see Appendix 1 for sequence and Table 5.3 for nomenclature). The solid boxes represent probe fragments derived from STS PCR.

amplify the conserved human *Fit*-1 open reading frame; and cos1SP66 and cos1SP67 (Appendix 1) to amplify the cos7 probe sequence. These primer sets amplify single copy sequences of 478bp and 498bp respectively and represent a pair of STSs from the *Fit*-1 region. Following primary and secondary PCRs, four PAC clones were confirmed as positive for one or both of the above STSs (Table 5.2).

Table 5.2 The *Fit*-1 and cos7 positive PAC clones

RZPD clone name	*Clone*	*Fit-1*	*cos7*	*Size (kb)*
LLNLP704L0612	PAC b	+	+	140
LLNLP704P0794	PAC e	+	-	130
LLNLP704L01315	PAC n	+	+	135
LLNLP704M01112	PAC p	-	+	125

5.3.3 STS content mapping of PAC and cosmid clones

To generate a contiguous map of the region encompassing the *Fit*-1 and cos7 positive PAC and cosmid clones, end-specific STSs were designed to demonstrate overlap within the contig. DNA sequence information derived from both insert ends of each PAC and cosmid clone was generated by sequencing with the SP6 and T7 IRD-41 labelled cycle sequencing primers from the cloning sites of the cosmid (Lawrist 4) or PAC (pCYPAC-2N) vectors.

This allowed the development of PCR assays specific for STSs from both ends of each clone, except for PAC b where only one end-specific STS could be generated due to incomplete sequence data. STS primer pairs from the SP6 and T7 insert ends of each clone are shown on Table 5.3. Using standard PCR cycling conditions (2.2.6), each terminal STS primer pair was used to analyse DNA from the four PAC and two cosmid clones to determine which contained these STSs and hence establish overlap. These data are summarised in Table 5.3 and Figure 5.2. For example, the PAC b end-specific STS from the T7 terminus of this clone (bT72 and bT73) is present in PAC clones b, e and n (Figure 5.2, red). Therefore, the T7 insert end of PAC b overlaps with PACs e and n, but not with PAC p or cosmids 1 and 7.

Table 5.3 Summary of STS content mapping

Clone	*Fit-1*	*cos7*	*Terminal STS primer pairs* [a]	*Positive PAC / cosmid clones*
cosmid 1	+	+	cos1T75 & cos1T76	b, e, n, 1
			cos1SP66 & cos1SP67	b, n, p, 1, 7
cosmid 7	-	+	cos7T71 & cos7T72	b, p, 7
			cos7SP61 & cos7SP62	b, e, n, p, 1, 7
PAC b	+	+	bT72 & bT73	b, e, n
PAC e	+	-	eT71 & eT72	b, e, n, p, 1, 7
			eSP61 & eSP62	e, n
PAC n	+	+	nT71 & nT72	b, n, p, 7
			nSP61 & nSP62	e, n
PAC p	-	+	pT71 & pT72	b, p
			pSP62 & pSP63	b, e, n, p, 1

[a] Nomenclature: each primer is named after the clone from which it is derived, from either the T7 or SP6 terminus and a specific number depending on how many rounds of sequencing/PCR were required to identify single copy sequence. For all primer sequences, see Appendix 1.

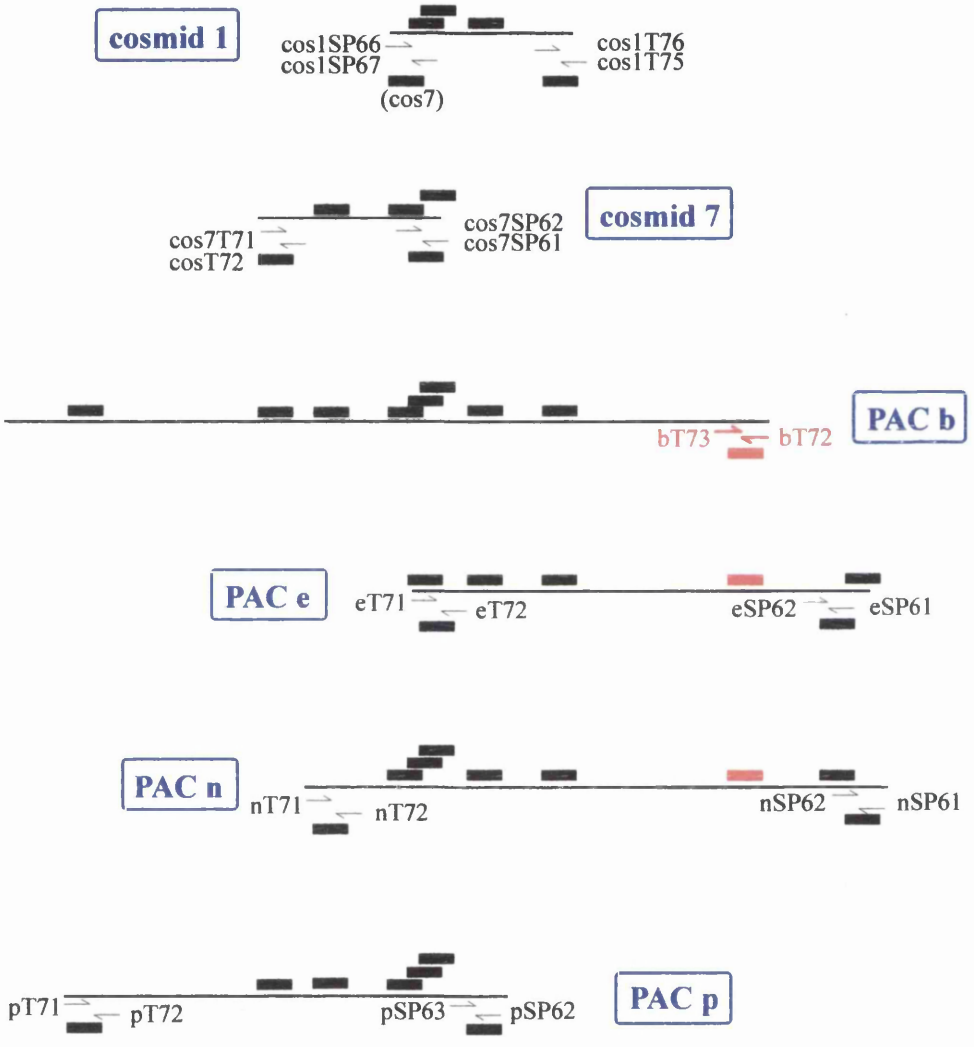


Figure 5.2 STS content mapping of the cosmid and PAC clones around *Fit*-1. This figure shows the positions of the end-specific STS primer pairs from each clone (half arrows). The solid line below each STS primer pair represents the PCR product. The other clones which also contain this STS (as determined by PCR) are also shown directly above or below this. In this way clones which overlap were determined. As an example, the PAC b STS primer pair (bT72 and bT73) and PCR product are highlighted in red This figure is not drawn to scale. For nomenclature, see Table 5.3 and for primer sequences, see Appendix 1.

5.3.4 Sequence analysis of PAC clones reveals homology to c-*myb*

Whilst sequencing from the T7 end of PAC p, a comparison of this sequence with the database of published sequences (NCBI Blast search) revealed 99% homology to the c-*myb* oncogene over 727bp (Figure 5.3). This places c-*myb* and *Fit*-1 within the same contiguous region of the human genome. The region of homology falls between the nucleotides 19081 to 19807 of the c-*myb* gene. This is synonymous with the region around exon 10 of this gene (Figure 5.3).

5.3.5 Long-range restriction mapping of overlapping PAC clones by PFGE places *Fit*-1 100kb upstream of c-*myb*

Following the identification of *Fit*-1 and c-*myb* on the same PAC clone, three overlapping PAC clones spanning the entire cloned region (PACs b, n and p) were characterised by pulsed-field gel electrophoresis (PFGE) to establish the proximity of *Fit*-1 and c-*myb* on human chromosome 6. DNA was digested with the rare-cutting restriction enzymes *Not*I (which excises the insert from the vector), *Bss*HII, *Cla*I, *Ecl*XI, *Sal*I and *Sst*II in various combinations and separated by PFGE (section 5.2.2 and 5.2.3). Southern blots of the size separated DNA fragments were hybridised with the human *Fit*-1 probe (generated by PCR with the huorf-f and huorf-r primers, Appendix 1). Following stripping of the probe, the filter was hybridised with a human c-*myb* exon 2 specific probe (generated by PCR with the exon2f and exon2r primers, Appendix 1). One representative Southern blot of the PAC b clone hybridised with *Fit*-1 (Figure 5.4A) or c-*myb* (Figure 5.4B) is shown. This result confirms the sequencing data which places c-*myb* within the same contiguous region. Southern blot analysis of PACs n and p are not shown. A long-range restriction map of approximately 200kb of cloned DNA surrounding the *Fit*-1 region is presented in Figure 5.5. These data placed *Fit*-1 and c-*myb* on a single PAC clone, approximately 100kb apart.

5.4 Discussion

To extend the search for a closely linked proto-oncogene which may be the target of insertion at *Fit*-1, overlapping cosmid and PAC clones extending in both directions were isolated and characterised. An overlapping contig map spanning

```
  727 TTTACCTTGGAATTGTTTAAGGAACACCTAGTTTTTTCCCAGAACTTTAGTGTTCTGAGA P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19081 TTTACCTTGGAATTGTTTAAGGAACACCTAGTTTTTTCCCAGAACTTTAGTGTTCTGAGA M

  667 GACACAGTTTATTATTCAAGAATTTTCCATATAATGTAAACCTTCTGTAATACTAATTTT P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19141 GACACAGTTTATTATTCAAGAATTTTCCATATAATGTAAACCTTCTGTAATACTAATTTT M

  607 GGTTCATAGATTATTAAAGAGAAAAAACAAAGTATGAAAGGAAACTATTAATTATTAACT P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19201 GGTTCATAGATTATTAAAGAGAAAAAACAAAGTATGAAAGGAAACTATTAATTATTAACT M

  547 ATTATTTTCCCCAAAATCACTAAAGCCTTTTAAATGTTAATAAAGAACACATCTCATTCT P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19261 ATTATTTTCCCCAAAATCACTAAAGCCTTTTAAATGTTAATAAAGAACACATCTCATTCT M

  487 CTTTAATGACAAGTGTCATTTATTGAGAGTATTTTCTCACAGTTTTGCAAGTTTTTCAGC P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19321 CTTTAATGACAAGTGTCATTTATTGAGAGTATTTTCTCACAGTTTTGCAAGTTTTTCAGC M

  427 AATGACCCTAAATTTTATTTCTTGGAAACCACACCCAAATCCTGGGTGCTCAATTGAAAT P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19381 AATGACCCTAAATTTTATTTCTTGGAAACCACACCCAAATCCTGGGTGCTCAATTGAAAT M

  367 AAAACAATACCCTAGTCAATATAACATGCTTGTTAGCATCTTTATTATTTAATATAGTGG P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19441 AAAACAATACCCTAGTCAATATAACATGCTTGTTAGCATCTTTATTATTTAATATAGTGG M

  307 GTCAGGAAAACATTCTTGTTATCTAGGTGTGATTTTTAAATTGGGGAGAATAAAGAATTA P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19501 GTCAGGAAAACATTCTTGTTATCTAGGTGTGATTTTTAAATTGGGGAGAATAAAGAATTA M

  247 CCATCTAATCTAAGTATTTTTTCTTTCTCTCCATATTTAGTTCTTAAACACTTCCAGTAA P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19561 CCATCTAATCTAAGTATTTTTTCTTTCTCTCCATATTTAGTTCTTAAACACTTCCAGTAA M

  187 CCATGAAAACTCAGACTTGGAAATNCCTTCTTTAACTTCCACCCCCCTCATTGGTCACAA P
      |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
19621 CCATGAAAACTCAGACTTGGAAATGCCTTCTTTAACTTCCACCCCCCTCATTGGTCACAA M

  127 ATTGACTGTTACAACACCATTTCATAGAGACCAGACTGTGAAAACTCAAAAGGAAAATAC P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19681 ATTGACTGTTACAACACCATTTCATAGAGACCAGACTGTGAAAACTCAAAAGGAAAATAC M

   67 TGTGTAAGTCTTTGGTTCAGGAATAAAATGATTTATCTTATTTACTTATATTTAATTAAT P
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19741 TGTGTAAGTCTTTGGTTCAGGAATAAAATGATTTATCTTATTTACTTATATTTAATTAAT M

    7 GGAATAT P
      |||||||
19801 GGAATAT M
```

Figure 5.3 Blast search results with PAC p terminal sequence. This figure shows the homology between sequence derived from PAC p (P) with human c-*myb* (M) sequence between 19081 and 19807 bp. Exon 10 lies between 19601 bp and 19743 bp of the human c-*myb* gene (accession number U22376, Westin EH and Gorse KM, unpublished).

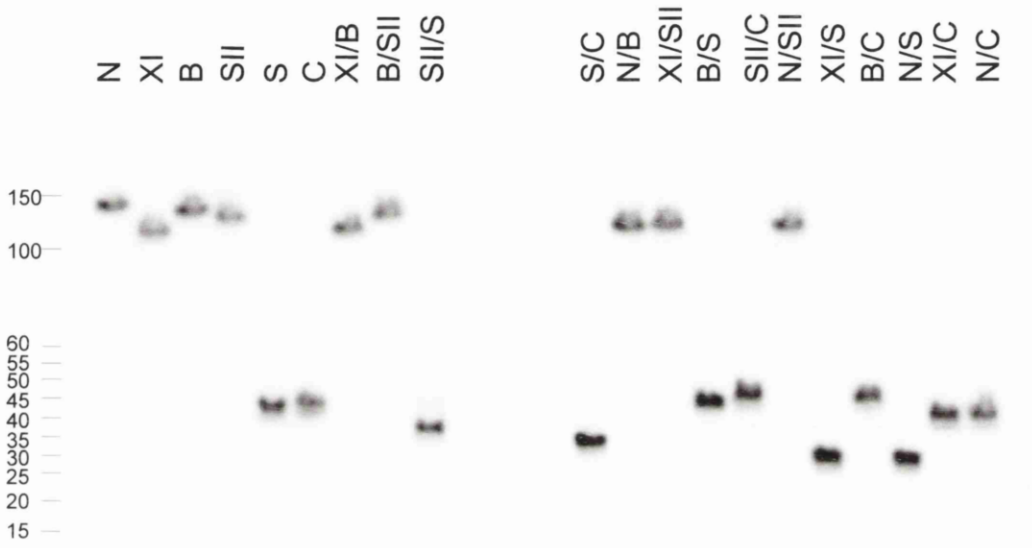


Figure 5.4A Southern blot analysis of PAC b DNA hybridised with the human *Fit*-1 probe. The approximate positions of the DNA size markers (5kb ladder and λ ladder) are shown on the left in kb. Restriction enzyme abbreviations: N, *Not*I; XI, *Ecl*XI; B, *Bss*HII; SII, *Sst*II; S, *Sal*I; C, *Cla*I. Washing conditions: three times at 0.1 X SSC, 0.5% SDS, 60oC, 20 min.

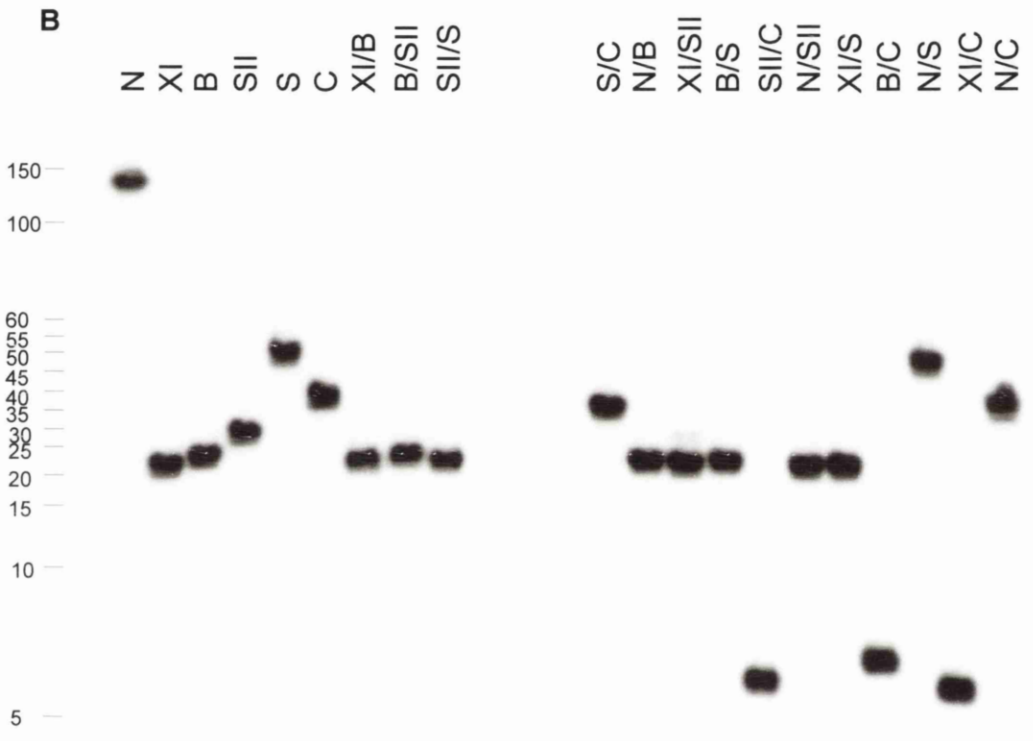


Figure 5.4B Southern blot analysis of PAC b DNA hybridised with the c-*myb* exon2 probe. The approximate positions of the DNA size markers (5kb ladder and λ ladder) are shown on the left in kb. Restriction enzyme abbreviations: N, *Not*I; XI, *Ecl*XI; B, *Bss*HII; SII, *Sst*II; S, *Sal*I; C, *Cla*I. Washing conditions: three times at 0.1 X SSC, 0.5% SDS, 60oC, 20 min.

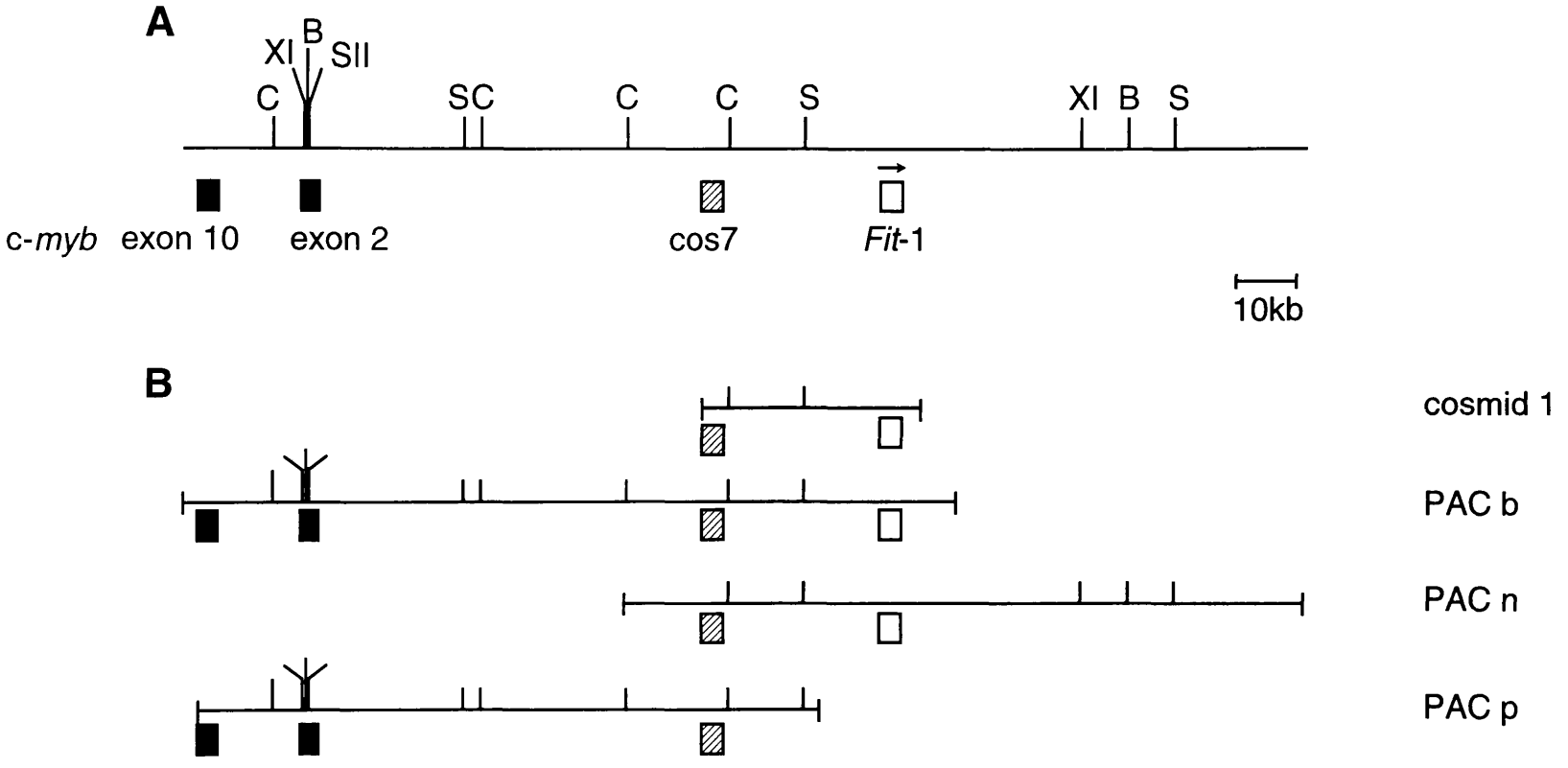


Figure 5.5 Human genomic map around *Fit*-1 and c-*myb*. A restriction map of the genomic sequence around the human *Fit*-1 probe is shown in (A). This map is derived from restriction analysis carried out on the individual clones shown in (B). The regions homologous to c-*myb* are also indicated (closed boxes). The position of the *Fit*-1 probe (open box) and the direction of the putative ORF is illustrated (the direction of the homologous feline ORF is the same as the majority of FeLV integrations at *fit*-1); the position of the cos7 probe is indicated (hatched box). Restriction enzyme abbreviations: B, BssHII; C, ClaI; XI, EclXI; N, *Not*I; S, SalI; SII, *Sst*II.

approximately 200kb of genomic sequence was generated using a combination of STS content mapping and long-range restriction mapping by PFGE (Figure 5.5).

A high degree of homology with the c-*myb* proto-oncogene was identified in PAC p (Figure 5.3) suggesting that *Fit*-1 and c-*myb* are closely linked on chromosome 6. Southern blot analysis confirmed this linkage and placed *Fit*-1 approximately 100kb upstream of c-*myb* (Figure 5.4B). Recent genetic linkage data in the mouse placed *Fit*-1 1cM proximal to the Abelson MuLV common insertion site, *Ahi*-1, which has been previously mapped to mouse chromosome 10, 35kb downstream of c-*myb* (Figure 3.2 and (18;157)). 1cM is approximately equivalent to 2Mb in physical distance. Clearly, this figure is inconsistent with the human physical mapping data and this disparity is discussed further in chapter 7.

Three lines of evidence suggest that *Fit*-1 maps *upstream* of c-*myb*. Firstly, sequence analysis of the T7 terminus of PAC p reveals that it carries sequence homologous to exon 10 of c-*myb* (Figure 5.3). Secondly, PAC b DNA hybridised specifically with an exon 2 c-*myb* probe (Figure 5.4B). Thirdly, a cluster of rare cutting restriction enzymes which cleave DNA at CpG containing sites (*Sst*II, *Bss*HII and *Ecl*XI, Figure 5.5A), is consistent with previous results which suggest the presence of a CpG island at the 5` end of the c-*myb* gene (170).

A possible explanation for the close linkage of *Fit*-1 to c-*myb* is that proviral integration within this locus activates c-*myb* expression by a long-range enhancer insertion mechanism. Such long-range *cis* activation of gene expression has been demonstrated previously for c-*myc*, *Evi*-1 and Cyclin D1 (28;41;64). Although initial analysis did not detect any abnormalities in c-*myb* expression in primary tumours rearranged at *Fit*-1 (42), the FT1 cell line with a proviral insertion at *Fit*-1 made it possible for subtle changes in c-*myb* expression to be investigated.

CHAPTER 6

Chapter 6

6. Analysis of c-*myb* expression in the FT1 cell line

6.1 Introduction

The *Fit*-1 locus has now been mapped close to c-*myb* in the human and mouse genomes (Figures 5.5 and 3.1 respectively and (157)). If the physical map of *Fit*-1 and c-*myb* is conserved in the feline genome, it is conceivable that proviral insertion within the *Fit*-1 region affects expression of the c-*myb* gene. This possibly deregulated expression may be responsible for the genesis of the observed T cell lymphomas, a hypothesis which is further supported by the observation of activated c-*myb* in other lymphomas of T cell origin (182;183).

Long-range *cis*-activation by proviral insertion has been observed previously for the c-*myc* oncogene where proviral insertions at the 3` end of the gene up to 300kb away have been shown to activate its expression (63;64). Therefore, due to the close proximity of *Fit*-1 and c-*myb* and the lack of transcription units activated by proviral insertion within 30kb of cloned feline genomic DNA, the expression of this oncogene was analysed in the FT1 cell line which is rearranged at *Fit*-1. The expression of c-*myb* in primary tumours carring proviral insertions at *Fit*-1 has been analysed previously (42). However, its expression has never been determined in a cell line where more subtle alterations in expression can be explored.

6.1.1 The c-*myb* oncogene

c-*myb* is a member of a family of nuclear transcription factors which also contains A- and B-*myb*. c-Myb plays a central role in proliferation, differentiation and apoptosis, primarily in haematopoietic cells (184), but also in non-haematopoietic cells such as bovine smooth muscle cells (185). The loss of N- or C-terminal sequences or its deregulated expression can convert c-Myb into a potent transforming protein in haematopoietic cells (186).

6.1.2 c-Myb in differentiation, proliferation and apoptosis

The expression of c-*myb* is predominantly confined to undifferentiated haematopoietic cells and as the cells mature expression disappears. The role of c-Myb in differentiation can be demonstrated in homozygous null mutant mice which do not survive past day 15 of foetal development and where haematopoiesis is severely impaired (187). Furthermore, immature myeloid or erythroid cells constitutively expressing wild type or truncated c-*myb* continue to proliferate in the presence of differentiation inducing agent, implying a crûcial role for c-*myb* in haematopoietic cell development (188-190). c-Myb has also been shown to regulate a wide range of genes which are important for differentiation (reviewed in (184)).

A large body of evidence exists suggesting a significant role for c-*myb* in proliferation as illustrated by its frequent activation in neoplasia (186). c-Myb activity is directed to the G_1/S phase of the cell cycle (Figure 6.1) (191;192) and it has been demonstrated that c-*myb*, when co-microinjected with c-*myc*, can drive quiescent bovine smooth muscle cells to progress into the S phase of the cell cycle (193). As cells enter terminal differentiation, c-*myb* expression is suppressed (Figure 6.1). This is supported by the observation that myeloblastic leukaemia M1 cells, which can be induced to terminally differentiate with interleukin 6 (IL6), continue to proliferate in response to IL6 when they constitutively express c-*myb* (190). Initial studies with anitisense c-*myb* oligonucleotides which inhibit c-Myb function resulted in the inhibition of growth of haematopoietic cell lines and blocked cells in the G_1/S phase of the cell cycle, providing direct evidence of a role for c-Myb in proliferation (194;195). These results have been consolidated with observations confirming that c-Myb is required for cell cycle progression: firstly, an inducible, dominant negative c-Myb construct resulted in the inhibition of proliferation in T cells and arrest in G_1 (196); and secondly, ribozymes specifically directed against c-*myb* result in arrest of cell growth in smooth muscle cells (197).

That c-Myb plays an important role in regulating proliferation can be further demonstrated by the genes which it has been reproted to transactivate. These

include c-*myc* which also has a central role in proliferation (see section 1.2.8). A number of studies support a role for c-*myb* in the transactivation of the c-*myc* promoter *in vitro* by direct binding to the c-*myc* promoter region (198;199), and that the constitutive expression of c-*myb* abolished the downregulation of c-*myc* during differentiation (190).

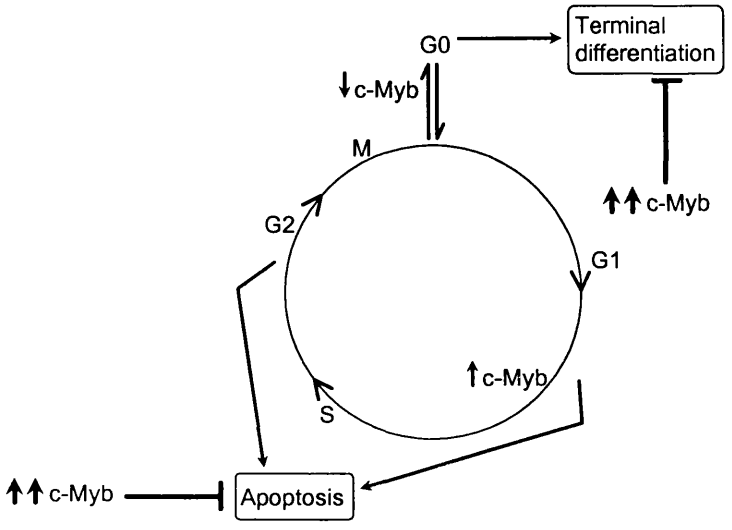


Figure 6.1. c-Myb and the cell cycle. The levels of cellular c-Myb increase during the G_1/S phase of the cell cycle and decrease as cells enter G_0 (indicated by single arrows). The effects of over-expression of c-Myb (as indicated by the double arrows) on apoptosis and terminal differentiation are also shown.

In addition to proliferation and differentiation, a role for c-Myb in apoptosis has recently been described in T cells and myeloid cells (183;200). The first study utilised an inducible, dominant negative c-Myb protein in the EL4 T cell lymphoma line and in transgenic mice. Activation of this mutant resulted in an increase in apoptosis which directly correlated with the downregulation of the anti-apoptotic protein, Bcl-2 (183). The second study used a temperature-sensitive mutant of the E26 avian leukosis virus encoding *myb-ets*. In this system culture of myeloid cells at the non-permissive temperature inactivates v-*myb* and the cells die by apoptosis and again this correlated with downregulation of Bcl-2

(200). In both systems the forced expression of *Bcl*-2 rescued cells from apoptosis. A further report confirmed that the constitutive expression of c-*myb* in the CTLL-2 T cell line results in resistance to apoptosis and again this was associated with the direct regulation of the *Bcl*-2 promoter (201). However, a study in primary haematopoietic cells transformed with an inducible C-terminal truncated c-Myb protein which also resulted in the downregulation of *Bcl*-2 expression, was not associated with an increase in apoptosis (202). This disparity could be due to the proposed cell-type specific anti-apoptotic effect of c-Myb (200;202) and/or that other additional mechanisms are involved in c-Myb induced apoptosis, such as different apoptotic pathways.

6.1.3 Structure and regulation of c-Myb

The c-Myb transcriptional transactivator protein has three major domains (Figure 6.2) encoded by 15 functional exons: an amino-terminal functional DNA binding domain; a central transactivation domain; and a carboxy-terminal negative regulatory domain (NRD) involved in the inhibition of transactivation by c-Myb (203).

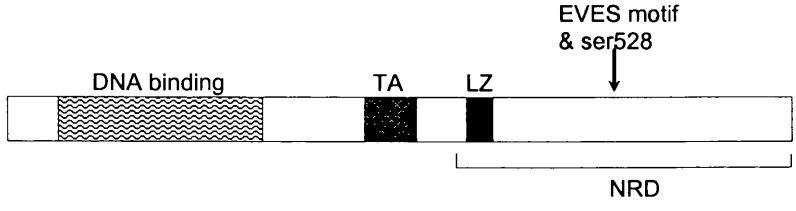


Figure 6.2. c-Myb protein structure and functional domains. The locations of the DNA binding, transactivation (TA), leucine zipper (LZ) and negative regulatory (NRD) domains are shown for the normal (p75) c-Myb protein. Also shown are the approximate positions of the EVES motif and serine 528 (ser528).

The regulation of c-Myb transcriptional transactivation activity is complex with transcriptional and post-translational regulation occurring. The NRD plays an important role in the regulation of c-Myb activity and transformation as

demonstrated by deletions of this region which result in an increase in transcriptional transactivation (203-207).

The NRD contains a putative leucine zipper (LZ) motif which may be responsible, at least in part, for the negative regulation of c-Myb (204;207). This property may be through protein-protein interactions mediated by the LZ motif (207;208). It has also been suggested that this region is involved in the autoregulation of c-Myb activity by homodimerisation resulting in a protein unable to bind DNA (209). These interactions are impaired in c-Myb mutants carrying a defective LZ motif. The negative regulatory role of the LZ domain can be further demonstrated by the frequent proviral insertion in exon 9 which lies within this domain (68;102), and by deletion of exon 9 which is associated with enhanced transcriptional activation (210). Furthermore, an alternatively spliced form of c-Myb (p85) which has an additional 121 amino acids in humans encoded by the alternative exon 9A, also disrupts the LZ motif (211). Again this has been shown to result in increased transactivation compared to wild-type p75 c-Myb (210;212).

The FAETL motif lies within the LZ domain and by deletion analysis has been shown to be required for transactivation and transformation by c-Myb. The disruption of this motif results in the increase of both the transactivating and transforming capacities of c-Myb, possibly through loss of interaction with a negative regulating factor mediated by the LZ domain (213).

Another C-terminal NRD motif, the EVES motif, is involved in the regulation of c-Myb by mediating intramolecular interactions with the N-terminal DNA binding domain (214). Furthermore, this motif was shown to mediate *inter*molecular interactions with the transcriptional coactivator p100 which also contains a highly related EVES domain. Interestingly, the EVES motif also contains a highly conserved serine 528 which, when phosphorylated has been shown to negatively regulate the activity of the c-Myb protein (215). Dash *et al* suggest a model in which c-Myb interacts with p100 via its EVES motif resulting

in an activated form of c-Myb. c-Myb could then be inactivated by an interaction with the C-terminal EVES motif of c-Myb with the N-terminus of c-Myb, a reaction which is likely to be regulated by phosphorylation of serine 528 (214). A recent study suggests that the coactivator p100 was phosphorylated and activated by the serine/threonine protein kinase Pim-1 which in turn stimulated c-Myb transcriptional activity (216).

An important means of regulation of c-*myb* expression is by attenuation, thus providing a sensitive mechanism of regulation by blocking transcriptional elongation at the onset of terminal differentiation (217;218). The attenuation site has been mapped to a position within intron 1 of the gene (219). A further level of control is by increasing the stability of c-Myb either at the RNA level (220) or the protein level (221).

6.1.4 c-Myb and cancer

Deregulated expression as well as structural alterations of c-*myb* are associated with a range of haematopoietic neoplasms (188-190). The oncogenic activation of c-*myb* by insertional mutagenesis has been implicated in myeloid leukaemias (68;69), B cell lymphomas (71;72) as well as promonocytic leukaemias (102) and truncation of either end of c-Myb has been shown to be sufficient for transformation (222).

Proviral integration often occurs within the first three exons of c-*myb* resulting in the activation of a potent transforming protein (69;72;102;102;223). Thus transcription from the relatively strong LTR promoter together with the loss of the attenuation site within intron 1 is thought to be responsible for the deregulation of transcriptional control. The minimal region of c-Myb truncation in B cell lymphomagenesis has been identified and results in the deletion of only 20 amino acids by ALV insertion producing a very potent transforming protein (72). C-terminal truncation by retroviral insertion has been shown to occur within exon 9 of c-*myb*, within or near the leucine zipper motif (68;102) possibly reflecting the putative negative regulatory role of this motif discussed above.

Retroviral insertions have also been observed which result in the truncation of c-Myb within exon 14 (101).

Aberrant c-*myb* expression has been observed in human neoplasia of haematopoietic and non-haematopoietic origin (224). Amplification of c-*myb* has been detected in a number of patients with acute myeloid leukaemia (AML) (225;226) while in another study defective c-*myb* and c-*myc* mRNA turnover was observed in cells from patients with AML (227). A reduction in the number of malignant haematopoietic colony-forming units in patients with chronic myelogenous leukaemia (CML) was demonstrated following exposure to c-*myb* antisense oligodeoxynucleotides (194;228) and more recently by transfection with an intracellular anti-c-*myb* single-chain antibody, an approach which is thought to be more specific for the functional knockout of the c-*myb* gene (229). In addition, the immature T cell acute lymphoblastic leukaemia (ALL) line, CEM, has lost a large portion of the c-*myb* promoter (166). Although amplification of *Myb* appears to be the most common abnormality of this gene in human leukaemias, truncated expression has been observed in the human CML cell line TK-6 (230). This form showed evidence of increased transactivation relative to wild-type c-*myb* probably due to the lack of the negative regulatory C terminal domain.

Despite the relatively common occurrence of c-*myb* abnormalities in human neoplasia, the large-scale molecular mapping of the c-*myb* locus at chromosome 6q23.3-6q24 did not reveal any rearrangements or deletions within 1Mbp of the gene (170). However, a study of haematopoietic cells cultured from a patient with AML showed that the amplification of normal sized c-*myb* mRNA correlated with chromosomal abnormalities in the region 6q22-24 (226).

6.2 Methods

6.2.1 c-*myb* RNA stability studies

To determine the stability of c-*myb* RNA, FT1 and 3201 cells ($5x10^5$/ml) were treated with 5μg/ml of the RNA synthesis inhibitor, actinomycin D (Sigma) for 0,

0.25, 0.5, 1, 2, 3, 4 and 8 hours. Actinomycin D was dissolved in 100% ethanol as a 1mg/ml stock solution and 5μg/ml was added to the culture medium. Total RNA was extracted from treated cells and northern blot analysis was performed on 20μg as described in sections 2.2.2.6 and 2.2.8.2. The filters were sequentially hybridised with the probes described in the text. The half-life of the c-*myb* message was estimated from densitometric scans of autoradiographs using Phoretix 1D Plus (NonLinear Dynamics Ltd.) and gapdh RNA as standard.

6.3 Results

6.3.1 c-*myb* gene expression is upregulated in FT1 cells

To determine if provirus integration within the *Fit*-1 locus in the FT1 cell line enhances c-*myb* expression by long-range activation, northern blot analysis was performed on 5μg of poly $(A)^+$ enriched mRNA and 40μg of total RNA extracted from exponentially growing cells (FT1 and 3201). This analysis was performed by hybridisation with a feline c-*myb* exon 2 specific probe which was generated using the conserved human exon 2 primers (exon2f and exon2r, Appendix 1). The filter was then stripped of probe and hybridised with a gapdh probe to control for RNA loading.

Expression of c-*myb* in the FT1 cell line and control (3201) is shown in Figure 6.3. This reveals that the feline c-*myb* transcript is similar in size to the human and murine homologues, approximately 4kb. A three day exposure of the northern blot (Figure 6.3A) shows that the absolute level of c-*myb* mRNA in the FT1 cell line (lane 1) is elevated approximately two fold over the control sample (lane 2). This analysis also revealed a 7.5kb transcript which hybridises to the c-*myb* specific probe in FT1 cells (lane 1) which is not present in 3201 cells (lane 2). Furthermore, a shorter exposure time (Figure 6.3B) reveals a 4.4kb species in FT1 mRNA which is not present in 3201 mRNA. Thus, c-*myb* mRNA from the FT1 cell line with a *Fit*-1 rearrangement is over-expressed compared to the 3201 cell line and also exhibits two larger transcripts which are not present in the control sample.

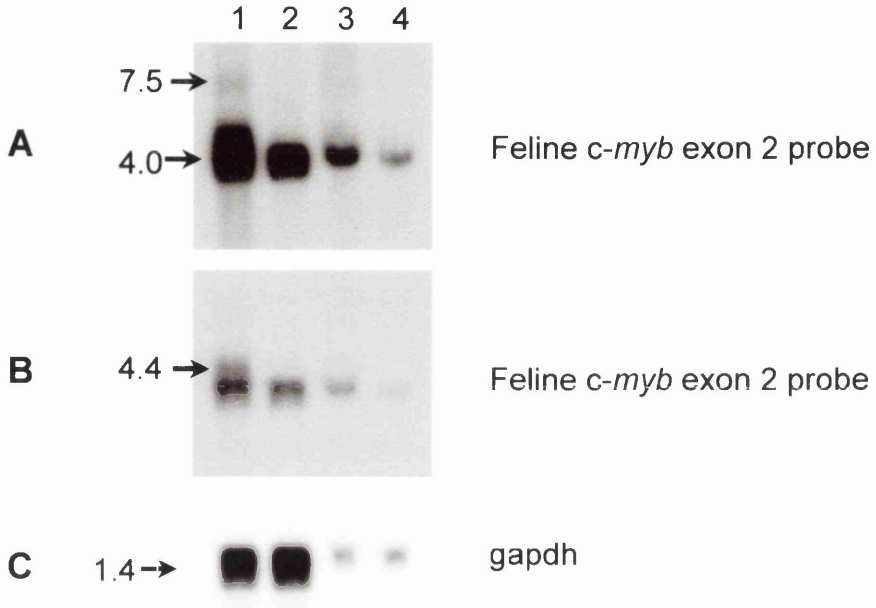


Figure 6.3 c-*myb* expression levels in FT1 cells. (A) Northern blot analysis of 5μg poly (A)+ FT1 mRNA (lane 1) or 3201 mRNA (lane 2); and 40μg of FT1 (lane 3) or 3201 (lane 4) total RNA. A feline c-*myb* exon 2 probe was used in this analysis and the filter was washed 3 times in 0.5 X SSC, 0.5% SDS at 60oC. (B) a shorter (overnight) exposure of the same blot highlighting the 4.4kb FT1 specific band. (C) The same northern blot probed with gapdh and washed 3 times in 0.1 X SSC, 0.5% SDS at 60oC.

A number of other feline T cell lines not rearranged at *Fit*-1 were analysed for expression levels of c-*myb* message to determine whether the enhancement of the c-*myb* transcript was due to an abnormality specific for the FT1 cell line. Total RNA was extracted from the following cell lines: FT1, 3201, Mya1, T3, F422, FL74, FL4 and T3C1. Northern blot analysis was performed on 40μg of each RNA using the feline c-*myb* exon 2 probe (Figure 6.4A) and a gapdh probe to control for RNA loading and integrity (Figure 6.4B). Densitometric analysis of a scanned image of the autoradiograph was performed using gapdh RNA as standard. The relative abundance of the c-*myb* message in each cell line is shown in Figure 6.4C. The results of this analysis confirmed that expression of the c-*myb* gene is upregulated in the FT1 cell line approximately two fold over 3201, T3, F422, FL74 and T3C1 cells, suggesting that proviral insertion within the *Fit*-1 locus in the FT1 cell line may indeed have a long-range effect on c-*myb* gene expression.

6.3.2 The c-*myb* gene is not rearranged in the FT1 cell line

To confirm that the over-expression of c-*myb* observed in the FT1 cell line is not due to proviral insertion within the c-*myb* gene itself, genomic DNAs from FT1 and 3201 cells were examined by Southern blot analysis for polymorphisms in restriction fragments which hybridise to c-*myb* probes. The feline c-*myb* exon 2 probe did not reveal any restriction fragment length polymorphisms (RFLP) in FT1 DNA (Figure 6.5A, lanes 1, 3 and 5), suggesting that the 5` end of c-*myb* is not rearranged. Furthermore, FT1 Southern blot analysis using mouse cDNA probes spanning the entire c-*myb* gene (the 2kb *Nco*I and the 2.5 kb *Eco*RI fragments from the plasmid pT7β/c-*myb* (231)) did not reveal any RFLPs (not shown).

In addition, fragments which hybridised to the c-*myb* exon 2 probe (Figure 6.5A) did not coincide with fragments which hybridise to the conserved feline *Fit*-1 probe on the same filter (Figure 6.5B). Although a very similar sized *Kpn*I genomic *Fit*-1 fragment also appeared to hybridise to the c-*myb* probe (Figures 6.5 A and B, lanes 5 and 6), the rearranged *Fit*-1 hybridising fragment in FT1

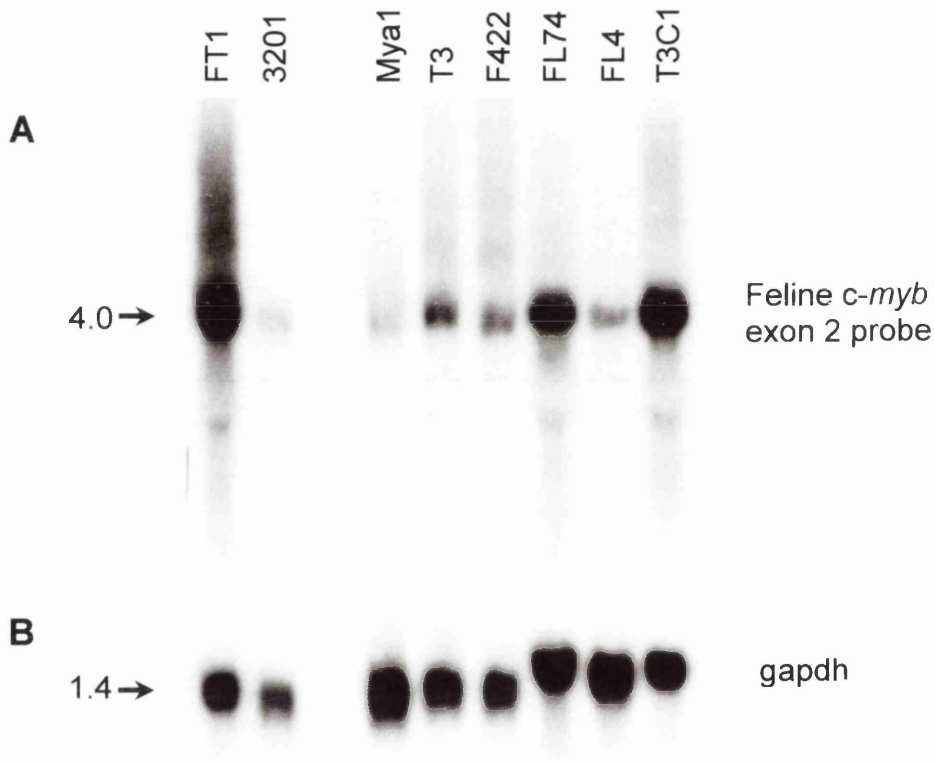


Figure 6.4 c-*myb* expression levels in feline T cell lines. Total RNA (40μg) from the various feline T cells indicated was extracted and analysed by northern blotting with a feline c-*myb* exon 2 specific probe (A). The filter was washed 3 times in 0.5 X SSC, 0.5% SDS at 60oC. After removing the probe, the same filter was hybridised to a gapdh probe to determine the approximate amount of RNA per lane (B).

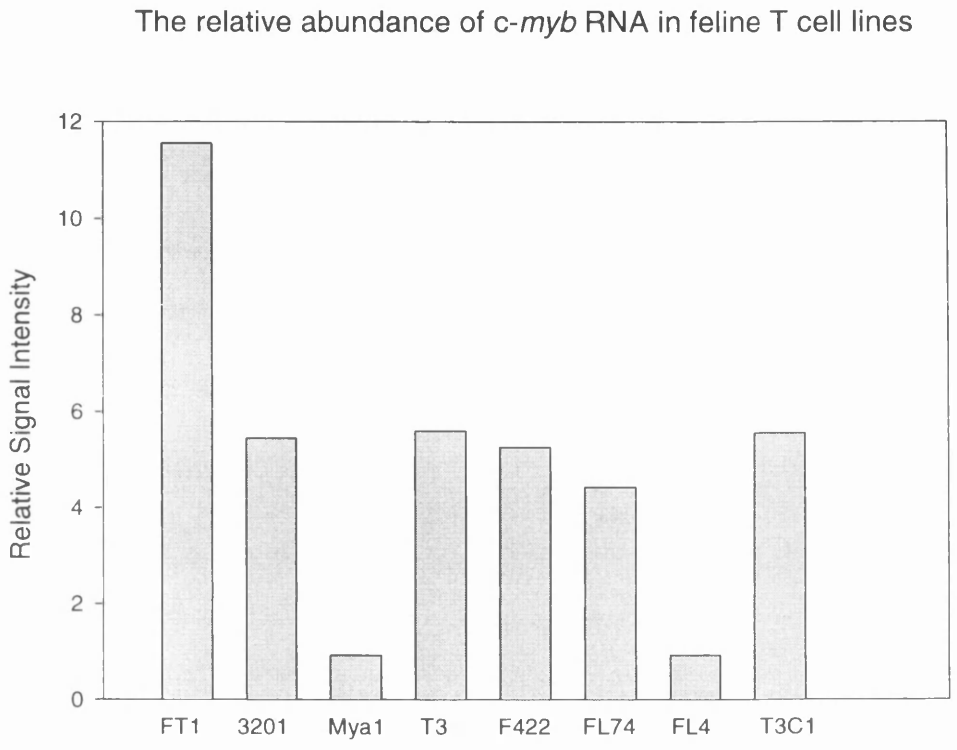


Figure 6.4C The relative abundance of c-*myb* RNA in feline T cell lines. Densitometry of a scanned image of the autoradiograph shown in Figure 6.4A was performed using Phoretix 1D Plus, NonLinear Dynamics Ltd. The relative signal intensity was calculated as a ratio of the gapdh signal shown in Figure 6.4B

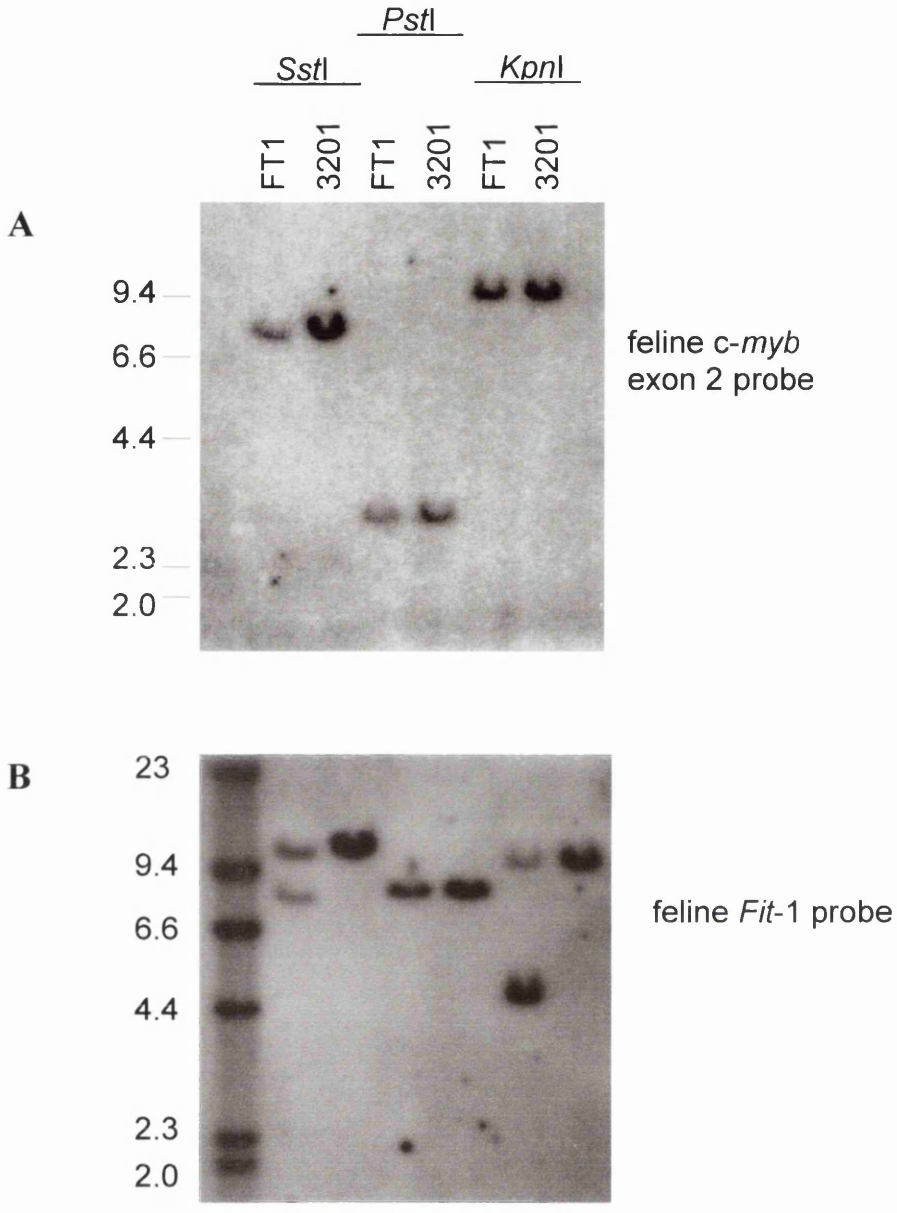


Figure 6.5 Southern blot analysis of c-*myb* and *Fit*-1 in the FT1 cell line. FT1 and 3201 DNA digested with *Sst*I, *Pst*I and *Kpn*I was immobilised on a Southern blot and probed with the feline c-*myb* exon 2 probe (A) or the feline *Fit*-1 probe (B). Both filters were washed at high stringency (0.1 X SSC, 0.5% SDS at 60oC).

DNA did not hybridise to the c-*myb* probe (compare lanes 5 in Figures 6.5A and B). This suggests that two similarly sized but distinct *Kpn*I fragments hybridise to both probes. Therefore, the FT1 cell line is not rearranged at c-*myb*, suggesting that it is proviral insertion within *Fit*-1 rather than c-*myb* which is responsible, at least in part, for the over-expression of c-*myb*.

6.3.3 c-*myb* RNA is more stable in the FT1 cell line

c-*myb* transcription is controlled at the level of transcriptional arrest and mRNA turnover (217;220;232). To determine if the increased abundance of c-*myb* mRNA observed in the FT1 cell line is due to a decrease in the turnover rate of this message, the half-life of c-*myb* RNA was determined. FT1 and 3201 cells were treated with actinomycin D to inhibit *de novo* RNA synthesis for 0, 0.25, 0.5, 1, 2, 3, 4 and 8 hours as described in section 6.2.1. Cytoplasmic RNAs were then isolated and analysed by northern blot using the feline c-*myb* exon 2 probe (Figure 6.6A). As controls, the filters were also hybridised with a gapdh probe (to control for RNA loading and integrity) and a histone H3.2 probe (to verify that RNA synthesis has been blocked). The latter probe was generated by PCR using the primers hisf and hisr described in Appendix 1 (233). Autoradiographs of these northern blots are shown in Figure 6.6B and C respectively.

Densitometric analysis of a scanned image of the autoradiographs shown in Figure 6.6A and C was performed using gapdh RNA as standard. The relative abundance of the c-*myb* and histone H3.2 RNAs at each time point are shown in Figures 6.6D and E respectively. Histone H3.2 RNA degraded with a half-life of approximately one hour (Figure 6.6C and E) confirming that RNA synthesis has been blocked in both cell lines. The results also show that c-*myb* RNA turned over with a half-life of approximately 3 hours in the 3201 cell line, consistent with the reported half-life of this transcript (193;220). However, in the FT1 cell line the half-life of c-*myb* RNA was significantly increased to approximately 8 hours (Figure 6.6D). This could account for higher steady state levels of c-*myb* transcript in FT1 cells.

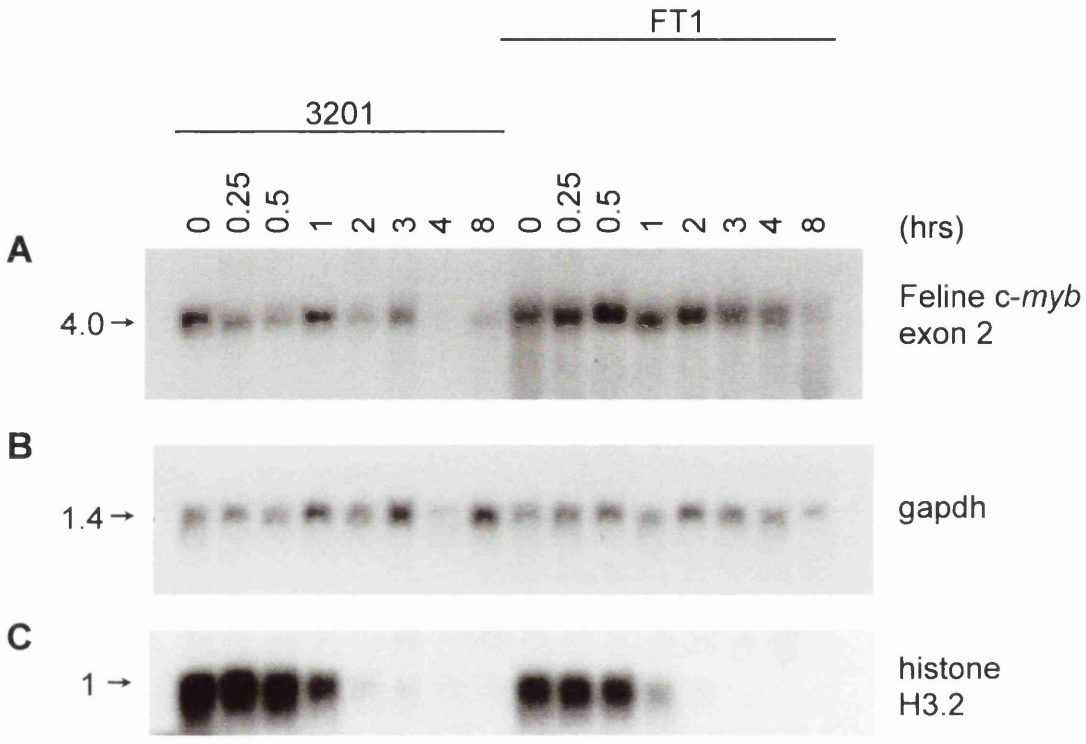
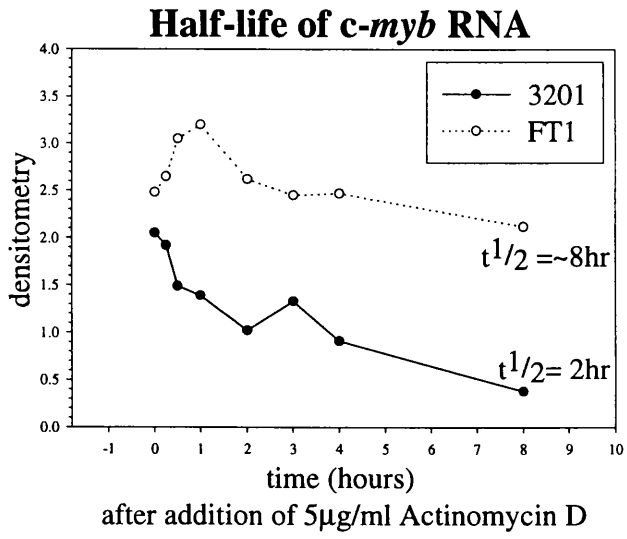


Figure 6.6 c-*myb* RNA half life. 3201 and FT1 cells were treated with actinomycin D for the indicated times (hours) and total RNA (20μg) was extracted. The northern blot was sequentially hybridised with feline c-*myb* exon 2 (A), gapdh (B) and histone H3.2 (C) This figure shows that the c-*myb* transcript in the FT1 cell line has a longer half life compared to the 3201 cell line. The half life of histone H3.2 RNA is not significantly different between the two cell lines.

D



E

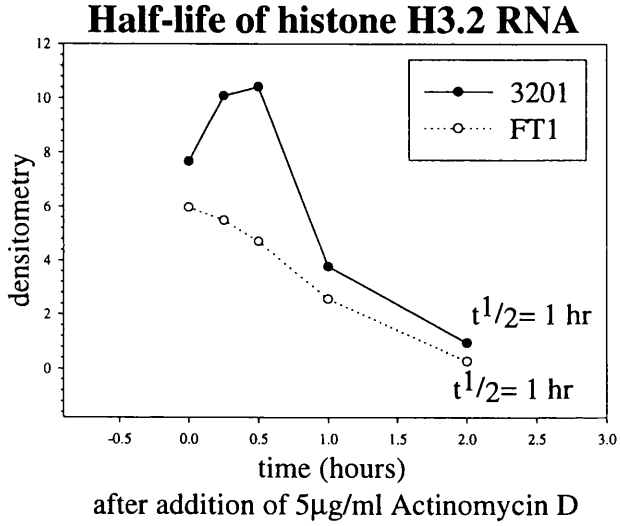


Figure 6.6 (continued) c-*myb* RNA half-life. Densitometry of a scanned image of the autoradiograph shown in Figure 6.6A (c-*myb*) was performed using Phoretix 1D Plus and is expressed as a ratio of the gapdh signal (D). Densitometry of a scanned image of the autoradiograph shown in Figure 6.6C (histone H3.2) was performed as above (E). The half-lives of the c-*myb* and histone H3.2 RNAs were estimated from the time taken (in hours) for half the message to be degraded.

6.4 Discussion

The location of *Fit*-1 approximately 100kb 5` to the human c-*myb* gene suggests that proviral integrations within this locus may activate c-*myb* gene expression in a manner analagous to the long-range *cis* activation observed with the c-*myc* gene (63;64). Northern blot analysis indicates that the relative level of c-*myb* mRNA is perturbed in the FT1 cell line with a rearrangement at *Fit*-1, suggesting that proviral integration at *Fit*-1 may activate its expression in *cis* (Figure 6.3). This analysis also revealed two larger mRNA species which are present in FT1 but not control cells, discussed further in chapter 7.

The enhanced c-*myb* expression observed in the FT1 cell line is not due to FeLV integration within this gene since Southern blot analysis revealed no rearrangements at c-*myb* in FT1 DNA using probes spanning the entire gene (Figure 6.5A and not shown). This result is in agreement with previous Southern analysis on primary tumour DNA hybridised with a c-*myb* probe which was carried out by our collaborator, H. Tsujimoto (42). In this study, tumours rearranged at *Fit*-1 revealed no RFLPs except for one, tumour T3-1T, where the rearranged fragment did not coincide with the rearranged *Fit*-1 fragment. It is possible that tumour T3-1T has proviral insertions at both the *Fit*-1 and c-*myb* loci, however lack of intact primary tumour material has not allowed the clarification of this issue.

c-*myb* activity is regulated at several levels (discussed in section 6.1.3). At the transcription level, control is primarily through attenuation of transcriptional elongation, a property which has been mapped to a region within intron 1 of the gene (217-219). c-*myb* gene expression is also regulated by a change in the stability of the message: as cells are artificially induced to enter differentiation, there is a rapid decline in mRNA levels due to a reduced half-life of this transcript as well as increased attenuation (220). In FT1 cells which are rearranged at *Fit*-1, the c-*myb* message appears to be more stable than control cells and this may contribute to the over-expression of this gene observed in the FT1 cell line (Figure 6.6). Indeed, the increase in the half-life is so significant

that it may account for the entire increased steady-state level in FT1 RNA. However, the possibility that other factors, such as reduced transcriptional pausing are also involved cannot be ruled out.

Two leukaemia specific c-Myb proteins which are truncated within the C-terminus by 248 or 96 amino acids have been shown to have significantly increased half-lives compared to full length c-*myb* (221;234). Wolff *et al* suggest that these observed increases of c-Myb protein stability due to proviral integration within exons 9 or 13 may contribute to the inappropriate expression of c-Myb in monocytic leukaemia. Furthermore, c-*myb* mRNA has been shown to be upregulated in patients with acute myeloid leukaemia and this was shown to be associated with enhanced stability of this transcript (227).

Higher steady-state levels of c-*myb* may result in an increased survival of the FT1 T cell lymphoma line. c-*myb* has been shown to regulate *Bcl*-2 expression resulting in protection from apoptosis (183;200). Therefore, it is conceivable that changes in c-*myb* RNA stability which result in the inappropriate expression of c-*myb* might confer increased survival on the FT1 cell line due to a decrease in apoptosis. Alternatively, the FT1 specific 4.4kb c-*myb* mRNA (Figure 6.3A) could represent an alternatively spliced product or one which has been initiated at a variant promoter and it may be this transcript which is more stable (discussed in greater detail in chapters 7 and 8).

CHAPTER 7

Chapter 7

7. Analysis of FT1 specific c-*myb* transcripts

7.1 Introduction

Northern blot analysis revealed that not only is the relative level of c-*myb* mRNA enhanced in the FT1 cell line, but that two larger mRNA species are present in FT1 which could not be detected in control cells (Figure 6.3). This observation raises the possibility that proviral insertion at *Fit*-1 may favour the expression of the larger transcripts. An attractive candidate for the 4.4kb FT1 specific c-*myb* message is the alternatively spliced form of c-*myb* previously identified in human, mouse and chicken which encodes a larger protein of approximately 85-89kDa (211;235;236). The 89kDa protein is generated from an alternatively spliced mRNA that contains an additional 121 amino acids in human and mouse which is derived from an alternative exon between exons 9 and 10, termed exon 9A. Therefore, it is possible that the 4.4kb RNA which is specific to FT1 cells contains these additional 363bp of sequence, a figure which is consistent with the increase in size of this product from approximately 4.0kb to 4.4kb (Figure 6.3).

The c-*myb* exon 9 sequence spans the putative leucine zipper domain, a motif which is thought to be involved in the negative regulation of c-Myb activity (204;207). The transcript encoding the larger protein which contains exon 9A in addition to exon 9 disrupts this motif, possibly resulting in a protein which is more resistant to negative regulation or one which has defective inter- or intra-molecular interactions. Although the function of this form is currently unknown, transfection studies have shown that it results in increased transcriptional activation of a luciferase reporter gene in avian myelomonoblasts (212). This study also showed that, although this form of c-Myb was able to mediate leukaemic transformation, there was no direct correlation of transcriptional activation and transformation efficiency.

7.2 Methods

7.2.1 Semi-quantitative reverse transcriptase PCR (RT-PCR) assays

Total RNA was extracted from cells as described in section 2.2.2.6. First-strand cDNA synthesis was generated using the First-strand cDNA Synthesis Kit (Amersham Pharmacia Biotech) which utilises M-MuLV reverse transcriptase. 5μg of total RNA and 5μl of bulk first-strand reaction mix was used according to the manufacturers' instructions. The first-strand reaction was primed with 0.2μg of the *Not*I-d$(T)_{18}$ bifunctional primer provided with the kit (see Appendix 1 for primer sequence). Following synthesis of the first-strand cDNA, the double-stranded RNA:cDNA heteroduplex was heat-denatured at 90^0C for 5 min and 3μl aliquots were amplified by PCR using sets of primers specific for c-*myb* (exon9f and exon11r, Appendix 1) or hypoxanthine phosphoribosyl-transferase (HPRT5` and HPRT3`, Appendix 1). PCR was performed essentially as described in section 2.2.6 except that 3mM $MgCl_2$ was used and cycling was performed for 25 cycles (c-*myb*) or 18 cycles (HPRT). The PCR products were run in a 1.5% agarose gel (section 2.2.3) and transferred to Hybond N membranes (Amersham) by standard Southern blotting procedures (section 2.2.8.1). The Southern blots were probed with c-*myb* (exon 9 or FE9A; see text for details) or HPRT (a 312bp PCR fragment amplified from residues 269 - 580 of *Mus spretus* cDNA using the HPRT5` and HPRT3` primers, see Appendix 1).

7.3 Results

7.3.1 Analysis of an alternatively spliced form of c-*myb* in the FT1 cell line

To determine if the 4.4kb FT1 specific c-*myb* message was derived from an alternative splicing event involving exon 9A, a probe specific for exon 9A was generated. A pair of PCR primers (huex9af and huex9ar, Appendix 1) were designed from two regions of exon 9A which were relatively well conserved in human, mouse and chicken. A feline exon 9A fragment of the expected size (357bp) was generated following PCR using 3201 DNA as template and standard cycling conditions (section 2.2.6, not shown). This fragment was gel purified (section 2.2.4) and used as a probe, the FE9A probe.

The FE9A probe was hybridised to the same poly $(A)^+$ mRNA northern blot which was previously probed with the c-*myb* exon 2 probe (chapter 6). The

FE9A probe hybridised to the 4.4kb mRNA which was previously seen with the exon 2 probe, but not to the 4.0kb species (Figure 7.1, lanes 1 and 2). This confirmed that the larger 4.4kb message, which was specific to FT1 RNA, was indeed derived from an alternative splicing event involving exon 9A. Moreover, the exon 9A containing c-*myb* species was more abundant in FT1 RNA (lane 1) compared to 3201 RNA (lane 2). Densitometry was performed on the northern blot shown in Figure 7.1 and this showed that exon 9A is approximately five-fold more highly expressed in FT1 than 3201 cells.

The structure of feline exon 9A was further analysed by PCR amplification of cDNAs derived from FT1 and 3201 cells (section 7.2.1). Using oligonucleotide primers corresponding to DNA sequences from the 5` end of human c-*myb* exon 9 (exon9f) and the 3` end of exon 11 (exon11r), PCR amplification yielded products of 499bp and 862bp (Figure 7.2A). The major band at 499bp is the size expected for a product which does not contain exon 9A, while the 862bp product corresponds to the size expected for a fragment containing the entire 363bp exon 9A.

Southern blot analysis of the PCR products was performed using a c-*myb* exon 9 specific probe generated by PCR using the oligonucleotide primers exon9f and exon9r (Appendix 1) and standard cycling conditions (section 2.2.6). The exon 9 probe hybridised to both the 499bp and the 862bp products confirming that they are both specific to the c-*myb* gene (Figure 7.2B). A duplicate Southern blot of the PCR products was hybridised with the FE9A probe identifying the 862bp product as containing exon 9A sequences, but not the 499bp product (Figure 7.2C). A 650bp product also hybridised to the exon 9A probe. The likely origin of this product is from another splicing event which utilises the 3` portion of exon 9A, termed exon 9A`. This species has previously been identified in chicken mRNA and is 210 bp shorter than exon 9A (237). An identical aliquot of first-strand cDNA was amplified using the HPRT primers (HPRT5` and HPRT3`) to control for RNA loading. The Southern blot of the PCR products probed with an HPRT probe is shown in Figure 7.2D.

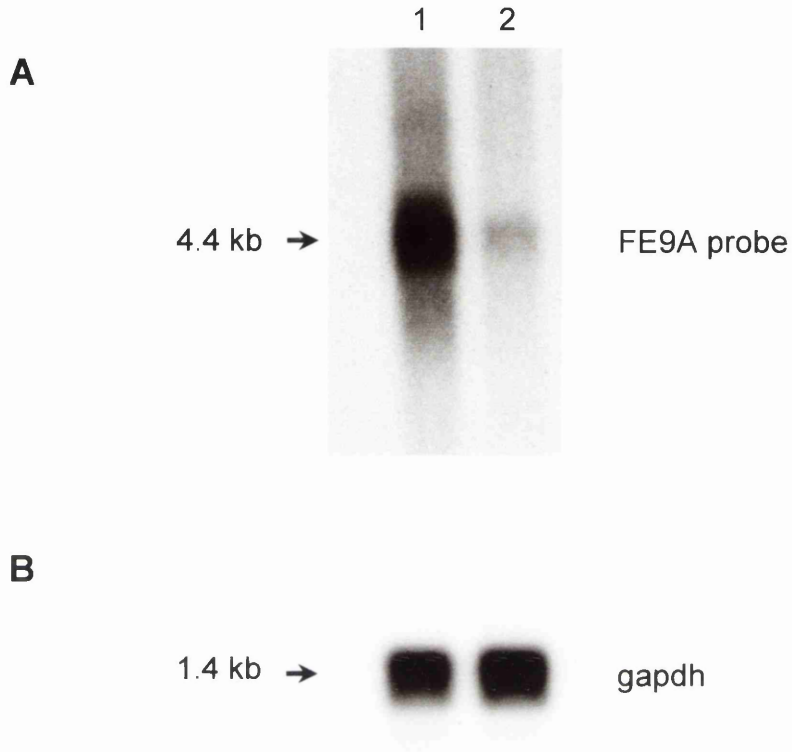


Figure 7.1 c-*myb* exon 9A expression levels in FT1 cells. (A) Northern blot analysis of 5μg poly (A)+ FT1 mRNA (lane 1) or 3201 mRNA (lane 2). A feline c-*myb* exon 9A specific probe (FE9A) was used in this analysis and the filter was washed three times in 0.1 X SSC, 0.5% SDS at 60oC. (B) The same northern blot probed with gapdh, washing conditions were as for FE9A.

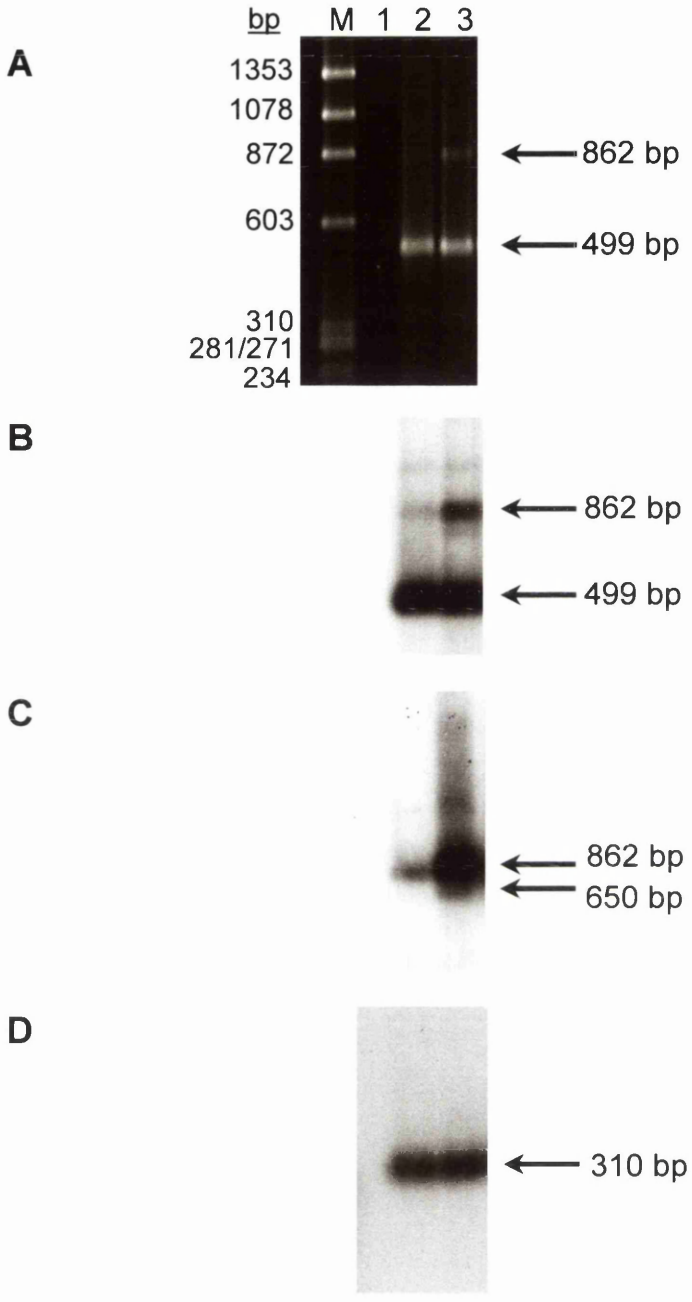


Figure 7.2 RT-PCR analysis of FT1 cells. (A) Ethidium bromide stained agarose gel of PCR products following amplification of cDNAs derived from the following RNAs (each panel is identical): (M) ϕX174/HaeIII DNA markers; (1) no RNA negative control; (2) 3201 RNA; (3) FT1 RNA. Primer sets for the PCR reactions are as follows: exon9f and exon11r (A, B and C); HPRT5` and HPRT3` (D). (B) Southern blot hybridisation of the gel shown in (A) using the human c-*myb* exon 9 probe. The filter was washed three times in 0.5 X SSC, 0.5% SDS at 60oC. (C) A duplicate Southern blot of the gel shown in (A) hybridised with the FE9A probe. The filter was washed three times in 0.1 X SSC, 0.5% SDS at 60oC. (D) Southern blot hybridisation using the HPRT probe. The filter was washed three times in 0.1 X SSC, 0.5% SDS at 60oC.

To calculate the level of exon 9A-containing c-*myb* message, densitometry was performed on the Southern blot shown in Figure 7.2B which has been probed with a c-*myb* exon 9 probe. Since this probe detects RNA species which contain as well as those which lack exon 9A sequence, the amount of probe hybridised is proportional to the amount of each isoform of c-*myb*. By calculating the area under the curve, exon 9A-containing cDNA was found to be present at approximately 9% of the total c-*myb* signal from 3201 cells. In contrast, exon 9A-containing cDNA was present at levels of approximately 25% in FT1 cells.

7.3.2 Feline c-*myb* exon 9A sequence

To determine the DNA sequence of feline c-*myb* exon 9A, the FT1 RT-PCR products shown in Figure 7.2A were cloned in pCRII TOPO (Invitrogen) as outlined in section 2.2.7. Colonies containing c-*myb* sequence with and without exon 9A were isolated by performing Grunstein and Hogness colony hybridisation (153) using the FE9A and the exon 9 probes, respectively. Several clones containing exon 9 hybridising sequences from FT1 and 3201 cDNAs were obtained and sequenced using the universal M13 forward and reverse primers (section 2.2.9, not shown). However, clones containing exon 9A sequence were only obtained from FT1 cDNA, possibly due to the under-representation of this form in 3201 cells. The sequence of feline exon 9A is presented in Figure 7.3. As for the human and mouse c-*myb* exon 9A, the feline exon was found to contain 363 bases which encode 121 amino acids (211;238).

The nucleotide sequences (Figure 7.4A) as well as the deduced amino acid sequences (Figure 7.4B) for feline, human, murine and chicken exon 9A were aligned to demonstrate homology. These comparisons show that feline and human c-*myb* exon 9A share the highest degree of homology at the nucleotide level (89% identity) and at the amino acid level (80% identity). The regions of maximal homology appear to be clustered into three boxes. Most notably, the C-terminal 12 amino acids of exon 9A are identical in all organisms (Figure 7.4B). These residues were previously shown to be highly conserved between human, mouse and *Drosophila* (211), chicken (237) and now the cat (Figure 7.4B). This

exon9f

1 ACACAGAACC ACACATGCAG CTACCCCGGG TGGCACAGCA CCACCATCGC

51 TGACCACACC AGACCTCATG GAGACAGTGC ACCTGTTTCC TGTTTGGAAG

101 AACACCACTC CACTCCCTCT CTGCCTGTGG ATCCTGGCTC CCTACCTGAA

151 GAAAGTGCCT CTCCAGCAAG GTGCATGATC GTCCACCAAG GCACCATTCT

201 GGATAATGTT AAGAACCTTT TAGAATTTGC AGAAACACTT CAATTTATAG

SA

251 ATTCT**GATTC TTCATCATGG cGTGATCTCA GCAGTTTTGA ATTCTTTGAA**

301 **GAAGCAGAGT TTTCACCTAG CCAACATCAT GCAGGCAGAG CCCTAAAGCT**

351 **TCAGCAAAGA GAGGGCAGTC TGAATAGACC TGCAGGAGAG CCTAGCACAA**

401 **GGGTGAACGC TCTCAAGTTG AGTGAGGGTT CACTTGACCC GCTCAAGCCC**

451 **TTACCTCCcT CGAGGCACGG CACAGTTCCA CTGGTCATCC TTCGAAAAAA**

501 **GCGCGGCCAG GCAAGCCCCT TAGCCACTGG AGGCTCTAGC TCCTTCCTAT**

551 **TTGCTGGCGT CAGCAGCTCA ACTCCCAAGC GTTCCCCTGT CAAAAGCCtA**

SD

601 **CCCTTCTCCC CTTCACAG**TT CTTAAACACT TCTGGTAACC ATGAAAACTT

651 AGACTTGGAA ATGCcTTCTT TAACTTCTAC CCCTCTCAAT GGTCACAAAT

701 TGACTGTTAC aACACCATTT CATAGAGACC AGACTGTGAA AACTCAAAAG

751 GAAAATACTA TTTTTAGAAC TCCAGCTATC AAAAGGTCAA TCCTAGAAAG

801 CTCTCCAAGA ACTCCTACAC CGTTCAAACA TGCACTTGCA GCTCAAGAAA

exon11r

851 TTAAATACGG TCC

Figure 7.3 Sequence of the feline exon 9A. The PCR products from FT1 cDNAs using the oligonucleotide primers exon9f and exon11r were cloned and sequenced as outlined in section 2.2.7. The positions of these primers are underlined. The boundaries of exon 9A are shown by the location of the splice acceptor (SA) and splice donor (SD) sites; the entire exon 9A sequence is shown in bold. The DNA sequence was obtained from both strands of two clones; ambiguities are shown in small letters.

```
ex9afe       1 TTTATAGATTCTGATTCTTCATCATGGCGTGATCTCAGCAGTTTTGAATTCTTTGAAGAA
ex9ahu       1 TTTATAGATTCTGATTCTTCATCATGGTGTGATCTCAGCAGTTTTGAATTCTTTGAAGAA
ex9amu       1 TTTATAGATTCTGATTCT...TCGTGGTGTGATCTCAGCAGTTTTGAATTCTCTGAAGAA
ex9ach       1 ......TTTCAGGATCCTTCATCATGGGGTGATCTCAGCAGTTTTGAATTCTTTGAAGA.
consensus    1 tttatagaTtctGATtCTtcaTCaTGG GTGATCTCAGCAGTTTTGAATTCTtTGAAGAa

ex9afe      61 ...GCAGAGTTTTCACCTAGCCAACATCATGCAGGCAGAGCC......CTAAAGCTTCAG
ex9ahu      61 ...GCAGATTTTTCACCTAGCCAACATCACACAGGCAAAGCC......CTACAGTTTCAG
ex9amu      58 GCGGCAGCTTTTTCACCTAGCCAGCAGCCCACAGGCAAAGCCTTCCAGCTTCAGCTTCAG
ex9ach      54 ..CACAGACATTCTGGCTGGCAAAGCTACCTCAGGCACAGCC......GTGCAGCTGCAG
consensus   61    gCAGa tTTtcacCTaGCcAacatc cgCAGGCAgAGCC      cTacAGcTtCAG

ex9afe     112 CAAAGAGAGGGCAGTCTGAATAGACCTGCAGGAGAGCCTAGCACAAGGGTGAACGCTCTC
ex9ahu     112 CAAAGAGAGGGCAATGGGACTAAACCTGCAGGAGAACCTAGCCCAAGGGTGAACAAACGT
ex9amu     118 CAAAGAGAGGGCCATGGGACTAGATCTGCAGGAGAGCCTAGCCTGAGGGTGACCAGGCGA
ex9ach     106 CATGGAGGGGCCAGTGCTTGTAGACCTCCAGGACTCCCCATCTCAAACCTGAGCAAAACC
consensus  121 CAaaGAGaGGgCagTg ga TAgAcCTgCAGGAgagCCtAgC caAgggTGAaCaaac

ex9afe     172 AAGTTGAGTGAGGGTTCACTTGACCCGCTCAAGCCCTTACCTCCCTCGAGGCACGGCACA
ex9ahu     172 ATGTTGAGTGAGAGTTCACTTGACCCACCCAAGGTCTTACCTCCTGCAAGGCACAGCACA
ex9amu     178 GTGCTGAGCGAGGCATCGCTCGGCCCAC.......ACTCACCCCAAGCGAGGCACAGCAAG
ex9ach     166 ATGTCAAGTCAGAGCCCTCCTGGCTCACCAAAGTCCTTGTCTGCCTCGCAGGGCAGTGTG
consensus  181 atGttgAGtgAGgg tCaCttGaCcCaC  aag  CTtacCtcc  CgagGcaCaGca a

ex9afe     232 GTTCCACTGGTCATCCTTCGAAAAAAGCGCGGCCAGGCAAGCCCCTTAGCCACTGGAGGC
ex9ahu     232 ATTCCACTGGTCATCCTTCGAAAAAAACGGGGCCAGGCCAGCCCCTTAGCCACTGGAGAC
ex9amu     232 GTTCGGCTGGTCGTCCTACGAAAAAGGCGGGGCCAGGCCAGCCCCCTAGCCGCTGGAGAG
ex9ach     226 GCTCCATGGG...TCCTTCGCAAAAGGAGAGGGCATGCCAGCCCCTTAGCCAGTGGCCCC
consensus  241 gtTCcactGGtcaTCCTtCGaAAAAagcGgGGcCAgGCcAGCCCCtTAGCCacTGGaggc

ex9afe     292 TCTAGCTCCTTCCTATTTGCTGGCGTCAGCAGCTCAACTCCCAAGCGTTCCCCTGTCAAA
ex9ahu     292 TGTAGCTCCTTCATATTTGCTGACGTCAGCAGTTCAACTCCCAAGCGTTCCCCTGTCAAA
ex9amu     292 CCTAGCCCCTCCCTCTTTGCTGACGTCATCAGCTCAACTCTCAAGCATTCCCCTGTCAAA
ex9ach     283 AGTAGCACCTTGGGATTGGCTGACGGCAGCAGCTCAACTTCTAAGCACTCCCCTGTCAAA
consensus  301   TAGC CCTtc taTTtGCTGaCGtCAgCAGcTCAACTcccAAGCgtTCCCCTGTCAAA

ex9afe     352 AGCCTACCCTTCTCCCCTTCACAGTTCTTAAACACTTCTGGTAACCATGAAAACTT
ex9ahu     352 AGCCTACCCTTCTCTCCCTCGCAGTTCTTAAACACTTCCAGTAACCATGAAAACTC
ex9amu     352 AGCCTACCCTTCTCTCCCTCGCAGTTCTTGAACACTTCCAGCAACCATGAAAACTC
ex9ach     343 AGCCTGCCCTTCTCTCCCTCGCAGGTAGATCAAGACTTCTGCACAGATGTTACTAA
consensus  361 AGCCTaCCCTTCTCTcCCcTCgCAGtTcttaaAcactTccgG AaccATGaaAact
```

Figure 7.4A Comparison of feline exon 9A DNA sequence with homologous sequences from various organisms. The alignments were performed with the computer programme GCG and the most highly conserved residues were highlighted using Boxshade (http://www.ch.embnet.org). Red indicates identity; periods indicate gaps in the alignment. ex9a, exon 9A; fe, feline; hu, human; mu, murine; ch, chicken.

```
ex9afe       1 DSSSWRDLSSFEFFEE.AEFSPSQHHAGRA..LKLQQREGSLNRPAGEPSTRVNALKLSE
ex9ahu       1 DSSSWCDLSSFEFFEE.ADFSPSQHHTGKA..LQFQQREGNGTKPAGEPSPRVNKRMLSE
ex9amu       1 D.SSWCDLSSFEFSEEAAAFSPSQQPTGKAFQLQLQQREGHGTRSAGEPSLRVTRRVLSE
ex9ach       1 DPSSWGDLSSFEFFED.TDILAGKATSGTA..VQLQHGGASACRPPGLPISNLSKTMSSQ
consensus    1   SSW DLSSFEF Ee           G A  l  Q   g   r  G P   v     S

ex9afe      58 GSLDPLKPLPPSRHGTVPLVILRKKRGQASPLATGGSSSFLFAGVSSSTPKRSPVKSLPF
ex9ahu      58 SSLDPPKVLPPARHSTIPLVILRKKRGQASPLATGDCSSFIFADVSSSTPKRSPVKSLPF
ex9amu      60 ASLGPHS..PQARHSKVRLVVLRKRRGQASPLAAGEPSPSLFADVISSTLKHSPVKSLPF
ex9ach      58 SPPGSPKSLSASQGSVAPWV.LRKRRGHASPLASGPSSTLGLADGSSSTSKHSPVKSLPF
consensus   61                    V LRKkRG ASPLA G  S    A   SST KrSPVKSLPF

ex9afe     118 SPSQ
ex9ahu     118 SPSQ
ex9amu     118 SPSQ
ex9ach     117 SPSQ
consensus  121 SPSQ
```

Figure 7.4B Comparison of feline exon 9A predicted amino acid sequence with homologous sequences from various organisms. The alignments were performed with the computer programme GCG and the most highly conserved residues were highlighted using Boxshade (http://www.ch.embnet.org). Red indicates identity; blue indicates similarity; periods indicate gaps in the alignment. ex9a, exon 9A; fe, feline; hu, human; mu, murine; ch, chicken.

high degree of conservation between such diverse species might suggest an essential biological function for an exon 9A containing c-*myb* gene.

7.4 Discussion

An alternatively spliced transcript of the c-*myb* gene involving the alternative exon 9A was first identified in a mouse myeloid tumour cell line (238). Since this initial report, c-*myb* exon 9A has also been identified in the human and chicken genomes (211;237). Sequences homologous to c-*myb* exon 9A have also been identified in the murine B-*myb* gene (239) and the *Drosophila* c-*myb* gene (211;240). The results presented here indicate that c-*myb* exon 9A sequence also exists in the cat (Figure 7.2). Comparisons of feline, human, murine and chicken exon 9A sequences show a high degree of conservation at both the nucleotide and protein levels (Figure 7.4). Previous analyses have shown that the last 12 amino acids in exon 9A are highly conserved in a very diverse group of organisms including human, mouse, chicken, sea urchin and *Drosophila* (211;212;237). The preservation of exon 9A sequences through evolution suggests an essential biological function of this form of the c-*myb* gene.

It has previously been observed that, although exon 9A was expressed in all chicken haematopoietic tissues tested, cells from yolk sac contained the highest levels of this alternatively spliced c-*myb* mRNA (236). Exon 9A-containing c-*myb* RNA in yolk sac was present at 15% of the normal 4.0kb c-*myb* message, while the other tissues contained only 1-3%. Results presented in this chapter have shown that the level of exon 9A-containing RNA in 3201 cells was estimated to be about 9% of the amount of total c-*myb* RNA (Figure 7.2B). This is in agreement with values seen previously (235;236). However, a substantially larger proportion of c-*myb* from FT1 cells contained exon 9A, approximately 25% of the total c-*myb* signal (Figure 7.2B). These data are in reasonable agreement with the levels of the 4.4kb transcript observed on the northern blot hybridised with a feline c-*myb* exon 2 probe shown in chapter 6 (Figure 6.3B). Densitometry of this blot showed that the 4.4kb c-*myb* message is present at

approximately 30% of the total c-*myb* signal in FT1 mRNA, while in 3201 mRNA this form was virtually undetectable. The slight variation seen between these figures may be due to disproportionate amplification of cDNA products during PCR or the indistinct nature of the exon 9A and normal forms of c-*myb* bands seen on the northern blot shown in Figure 6.3B on which densitometry was performed.

CHAPTER 8

Chapter 8

8. General Discussion

Proviral tagging has proved to be a powerful tool for the identification of novel genes in retrovirus-induced tumours of the mouse, chicken and cat where more than 60 genes have been found as targets for insertional mutagenesis (Table 1.1). The set of genes identified in this way overlaps significantly with those identified by non-retroviral means such as oncogenes present at chromosomal translocations and those which have been amplified in tumorigenesis. This illustrates the value of this approach in identifying genes which are involved in naturally occurring neoplasia.

The *Fit*-1 locus was initially identified by proviral tagging in cats using a naturally-occurring FeLV strain carrying a v-*myc* oncogene. This locus was found to be occupied by helper provirus in 16% of feline T cell lymphomas screened (25;42). Thus, precedent suggests that the *Fit*-1 locus harbours a gene which collaborates with Myc in T cell neoplasia. Preliminary analysis failed to identify a transcription unit which was altered as a result of proviral insertion at *Fit*-1, despite the isolation of 30kb of genomic sequence and the identification of a stretch of unique sequence which is conserved in mammalian evolution (157). Therefore, the purpose of this study was to extend the search for transcripts which were affected by proviral integration at *Fit*-1 using the FT1 T cell lymphoma line which is rearranged at both *Fit*-1 and c-*myc*, a resource which was not previously available for this analysis.

To search for a gene affected by insertions at *Fit*-1, a 15kb feline genomic λ phage clone (λRF1) which spans the major insertion cluster (Figure 3.3) was screened for sequences that are single copy, conserved in evolution and expressed as RNA specifically in the FT1 cell line. This strategy (outlined in Figure 3.5) was utilised across the entire clone. No clear evidence of a transcription unit affected by viral insertion was identified. However, this analysis did reveal a

region of the λRF1 which was not contiguous in genomic DNA, possibly due to a spurious recombination event which may have occurred during cloning.

Possible reasons for the failure to identify transcripts in the vicinity of the clustered integration sites were discussed previously (section 3.3). The most likely explanation is that the cluster of proviral insertions at *Fit*-1 may be influencing the expression of a gene which is outwith the 30kb of cloned sequence. A precedent for such an event is seen with the c-*myc* oncogene which was found to be the target of proviral insertions at the *Mlvi*-1/*pvt*-1/*Mis*-1 and *Mlvi*-4 loci which map 270kb and 30kb downstream respectively (61;63;64). Therefore, in view of the conservation of the *Fit*-1 domain, we chose to extend the search for adjacent genes in human DNA where there are greater resources available for genomic analysis.

The *Fit*-1 locus was previously mapped to feline chromosome B2 (42). Using the EUCIB panel of mice and the conserved *Fit*-1 domain, the locus was subsequently mapped to mouse chromosome 10 close to *Ahi*-1, a murine retroviral integration locus which is 35kb downstream of c-*myb* (18;157). Much of feline chromosome B2 has been shown to be homologous to human chromosome 6 (161), and the human c-*myb* gene has been mapped to human chromosome 6 at 6q23.3-24 (163). Therefore, to isolate the human homologue of *Fit*-1, a human chromosome 6 specific cosmid library was screened with the conserved *Fit*-1 probe identifying a single clone containing this sequence. Nucleotide sequence analysis of the human *Fit*-1 homologous domain revealed that the long open reading frame present in the feline sequence is 80% identical at the DNA level to the corresponding region in the human genome (Figure 4.3A). The presence of conserved sequences would suggest that this region has been maintained through evolution because it plays an important biological function. However, probes encompassing these open reading frames (the feline and human *Fit*-1 probes) failed to reveal any expressed RNAs on northern blot analysis. This was surprising considering their conservation and the presence of features of

an internal coding exon such as the putative splice acceptor and splice donor sites highlighted on Figure 4.3A.

The physical linkage of *Fit*-1 and c-*myb* was demonstrated when a *Fit*-1 positive PAC clone was found to contain sequences homologous to c-*myb* (Figure 5.4B). Long-range restriction mapping and pulsed field gel electrophoresis mapped the *Fit*-1 region approximately 100kb upstream of c-*myb* on human chromosome 6 (Figure 5.5). Murine c-*myb* has been mapped to chromosome 10 which also harbours a number of common murine leukaemia virus (MuLV) integration loci including *Ahi*-1, *Mis*-2 and more recently *Mml*1 (Figure 8.1). *Ahi*-1 was identified as a common site of proviral insertion in Abelson induced pre-B lymphomas (17) and was mapped approximately 35kb downstream of c-*myb* (18). The *Mis*-2 locus was identified as a common integration site in Moloney MuLV induced thymic lymphomas in rats and was subsequently found to map 160kb downstream of c-*myb* (59). Finally, *Mml*1 was identified as a novel integration site in MuLV induced promonocytic myeloid leukaemias. This locus was mapped 20-25kb upstream of c-*myb* (66).

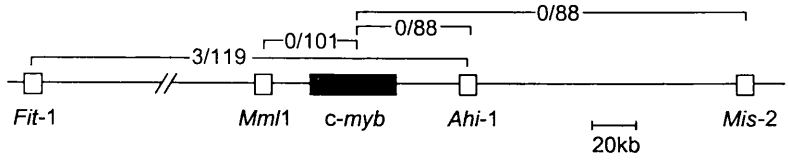


Figure 8.1. Comparative map of the mouse *Fit*-1, *Mml*1, c-*myb*, *Ahi*-1 and *Mis*-2 regions combining both the physical and recombination data between each of these loci on chromosome 10 (18;59;66). See text for more detail. Note that the physical position of *Fit*-1 on this figure is an estimate based on the human data and also that it is drawn with the c-*myb* gene in a 5`-3` orientation.

Genetic linkage analysis using the EUCIB panel of mice also placed the *Fit*-1 locus within this cluster of loci on mouse chromosome 10 (157). Of 119 backcross progeney screened for both loci only three recombinants were detected. This total of 119 mice represented an initial random screening panel and a second panel of known chromosome 10 recombinants. Based on EUCIB algorithms these data placed *Fit*-1 approximately 1cM from *Ahi*-1 (Figure 3.2).

Although I have been unable to map the physical location of *Fit*-1 in the mouse due to difficulties in cloning the mouse homologue of the *Fit*-1 probe, we can tentatively construct an integrated recombination/physical map of this cluster illustrating the putative position of *Fit*-1 based on the human data which places it 100kb upstream of c-*myb* (Figure 8.1). Clearly, the genetic linkage analysis in the mouse which suggested a linkage of 1cM, approximately equivalent to 2000kb, is inconsistent with the physical analysis in the human genome. Although a true genomic difference between the species may exist, this discrepancy may reflect the limited resolution of the EUCIB map in this domain and/or the possible existence of a meiotic recombinational hot spot between the *Fit*-1 and *Ahi*-1 regions. To clarify this apparent discrepancy, cloning of the mouse homologue of the *Fit*-1 domain using a mouse PAC library will allow the physical mapping relative to c-*myb*. The recent availability of feline PAC and BAC libraries will also serve as useful tools for the physical mapping of *Fit*-1 and c-*myb* in the cat and this work is currently underway in collaboration with N. Yuhki (NCI Frederick, Maryland, USA).

The *Fit*-1 locus was shown to be distinct from the *Mml*1 common insertion locus by the failure of the conserved *Fit*-1 probe to hybridise to the murine *Mml*1 P1 clone number 5362, Genome Systems Inc., St. Louis, MI (66), even under low stringency conditions (not shown). However, as the *Mml*1 probe did not hybridise to human DNA, we have been unable to map the physical linkage between *Fit*-1 and *Mml*1 in the human genome (L. Wolff, NIH, Bethesda, Maryland, USA, personal communication). The generation of cross-species probes from the *Mml*I locus should help resolve this issue.

The most significant explanation for the clustering of integration sites in this region is that they have a common consequence, possibly the activation of c-*myb* gene expression by a long-range enhancer insertional activation mechanism. Indeed, the results of this thesis have shown that c-*myb* gene expression is upregulated in the FT1 cell line with a proviral insertion approximately 100kb upstream of the c-*myb* gene (Figure 6.3). A precedent for such a clustering of insertions affecting a common target has been seen at the *Gfi*-1/*Pal*-1/*Evi*-5/*Eis*-1 locus. Insertions at any one of these loci which lie within a 50kb domain, appear to activate *Gfi*-1 transcription in B and T cell lymphomas (37-39). Thus, independent proviral integrations in the *Gfi*-1/*Pal*-1/*Evi*-5/*Eis*-1 locus can activate *Gfi*-1 expression over distances of up to 25kb, probably by an enhancer insertion mechanism.

Perhaps the best characterised example of long-range activation is the activation of the c-*myc* oncogene by retroviral insertion at the *Mlvi*-1 or *Mlvi*-4 loci within a 300kb domain (63;64). Proviral insertions within either loci were shown to activate, in addition to *Mlvi*-1 and *Mlvi*-4, the c-*myc* gene. Activation was shown to occur in *cis* using cell hybrids between rat thymic lymphoma cell lines rearranged at *Mlvi*-1 or *Mlvi*-4 and a murine T cell line (64). From this analysis, the over-expression of rat c-*myc* cosegregated with the rearranged loci while there was no effect on murine c-*myc*, providing strong evidence that provirus insertion exerts its long-range effect on c-*myc* in *cis*.

Other examples of long-range activation include the *Evi*-1 gene which has also been shown to be activated in *cis* by retroviral insertions 90kb 5` of the gene at the *Cb*-1/*fim*-3 locus, possibly by affecting the methylation state of the *Evi*-1 gene (28). Proviral integration at the *Fis*-1 locus is closely linked to the cyclin D1 gene (*Cyl*-1) on mouse chromosome 7 where it has been mapped within 300kb on

the 5` side of this gene. Further analysis revealed that proviral integration at *Fis*-1 can activate the expression of *Cyl*-1 in mouse lymphomas and that this may be functionally equivalent to the *BCL*1 translocation in human B cell lymphoma (41). More recently, a woodchuck hepatitis virus common integration locus (*win*) was mapped 150-180kb from N-*myc*2 on the X chromosome in woodchuck liver tumours and was shown to be associated with enhanced expression of this gene (159).

It is possible that an analogous situation is operative in the *Fit*-1/*Mml*1/c-*myb*/A*hi*-1/*Mis*-2 cluster where activation of a common target (most likely c-*myb*) may be occurring in *cis* over distances of up to 160kb. This distance is within those seen previously, where enhancer elements have been shown to activate transcription from a gene up to 300kb away (discussed above). In favour of this hypothesis are the results of this study which reveal that the FT1 cell line which carries a proviral insertion within the *Fit*-1 locus shows evidence of enhanced c-*myb* expression (Figure 6.3). As there are no proviral insertions within the c-*myb* gene itself (Figure 6.4A), this suggests that proviral insertion at *Fit*-1 is affecting the expression of c-*myb*. These results appear contrary to previous northern blot analysis of RNA from tumours carrying proviral insertions at *Fit*-1 (42). However, as these analyses were conducted on primary tumour RNA of variable quality, subtle changes in the expression of c-*myb* may have been missed.

Although proviral occupation of the *Fit*-1 locus may be affecting c-*myb* expression in *trans*, a number of lines of evidence suggest that this is not the case; most notably the failure to identify any evidence of a transcription unit which is affected by proviral integration at *Fit*-1. Furthermore, long-range restriction mapping of the human *Fit*-1/c-*myb* region did not identify a CpG island in the intervening DNA. As discussed previously, CpG islands are frequently associated with the 5` ends of genes and serve as useful indicators of the presence of actively transcribed sequence (107). However, they have only

been associated with approximately 60% of genes. Therefore, the lack of evidence of a CpG island in the sequence flanking the *Fit*-1 locus does not necessarily exclude the possibility of another nearby gene which has yet to be identified.

Of particular interest when considering possible mechanisms for the long-range activation of c-*myb* by integration at *Fit*-1, is the observation that all the proviruses except one are integrated in the same orientation (Figure 3.3). This pattern is not typical of promoter insertion, but signifies that an enhancer insertional activation mechanism may be occurring, a mechanism which is known to occur over long distances (discussed above). Moreover, the integrations mainly occur in regions of unique sequence and at least one of these regions is conserved in evolution (Figure 3.1, tumour J51-3T). This observation may reflect the presence of regulatory elements in this region that influence c-*myb* at a long distance and which may be disrupted by proviral insertion. Conversely, disruption of chromatin structure in the vicinity of the proviral insertion may have a downstream effect possibly by disturbing regulatory interactions between the *Fit*-1 region and c-*myb*. Such a postulate has been raised previously in a hypothetical model for the mode of activation of N-*myc*2 by integrations at the *win* common insertion locus (159).

Analysis of mRNA from the FT1 cell line which is rearranged at *Fit*-1 has now provided evidence of enhanced expression of the c-*myb* gene in this T cell lymphoma line compared to a number of other feline T cell lymphoma lines without rearrangements at *Fit*-1 (Figure 6.4). However, do proviral insertions at *Mml*1, *Ahi*-1 and *Mis*-2 also activate the expression of c-*myb*? This hypothesis is consistent with the lack of evidence of transcription units at *Mml*1 and *Mis*-2, although there is some evidence of a novel gene at *Ahi*-1 (18). Northern blot analyses of RNA from tumours carrying proviral insertions at *Ahi*-1 or *Mis*-2 showed no evidence of strongly enhanced c-*myb* expression (18;59). However, as in the initial *Fit*-1 studies, these analyses were conducted on primary tumour RNA and subtle changes in the expression of c-*myb* may have been missed. It is

interesting to note that two murine T cell lymphoma cell lines (p/m 16i and G1500 44i) were recently shown to carry insertions at *Ahi*-1 (M. Stewart, unpublished). Therefore, these will serve as useful tools to analyse subtle effects on c-*myb* expression as a result of proviral integration at *Ahi*-1. Analysis of c-*myb* RNA and protein from murine myeloid leukaemias and cell lines with *Mml*1 rearrangements also showed no obvious evidence of over-expression (66). To conclude, therefore, it is conceivable that these clustered insertions have a subtle effect on the expression of c-*myb*. However the possibility that these insertions affect another closely linked gene which has yet to be identified cannot be ruled out.

Northern blot analysis revealed not only that c-*myb* expression was enhanced in the FT1 cell line, but that there were also two additional mRNA species which hybridised to the c-*myb* probe with sizes of approximately 4.4kb and 7.5kb (Figure 6.3). These observations raise the possibility that proviral insertion at *Fit*-1 favours expression of these larger transcripts. Further northern blot analysis of FT1 mRNA revealed that the 4.4kb message hybridised specifically to an exon 9A (FE9A) probe (Figure 7.1). This analysis also showed that the exon 9A form of c-*myb* was approximately five fold over-expressed in FT1 compared to 3201 mRNA.

Although over-expression of c-*myb* containing exon 9A has previously been shown to increase transcriptional transactivation of a reporter gene in avian myelomonoblasts (212), the function of this isoform is currently poorly characterised. Previous analysis showed that, although exon 9A was expressed in all chicken haematopoietic tissues tested, cells from yolk sac contained the highest levels of this alternatively spliced c-*myb* mRNA (236). Exon 9A-containing c-*myb* RNA in yolk sac was present at 15% of the normal, 4.0kb c-*myb* message, while the other tissues tested contained only 1-3% of the exon 9A form. Schuur *et al*, therefore suggested that this isoform may play a role in differentiation of immature haematopoietic cells (236). However, neither the FT1 nor the 3201 cell lines can be classed as immature cell lines as their T cell

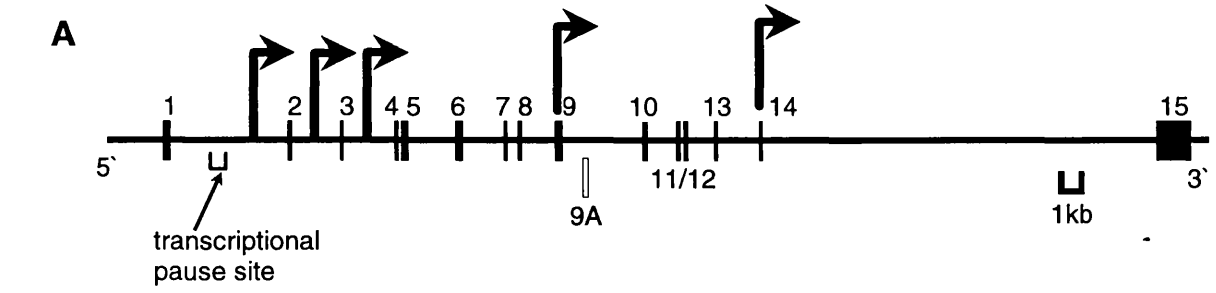
receptor Cβ chains are rearranged as assessed by Southern blot analysis (not shown). Further analysis of cell surface markers on these cell lines will be necessary to support this conclusion. Therefore, the exon 9A isoform of c-*myb* may play another role, possibly in the development of T cell lymphoma.

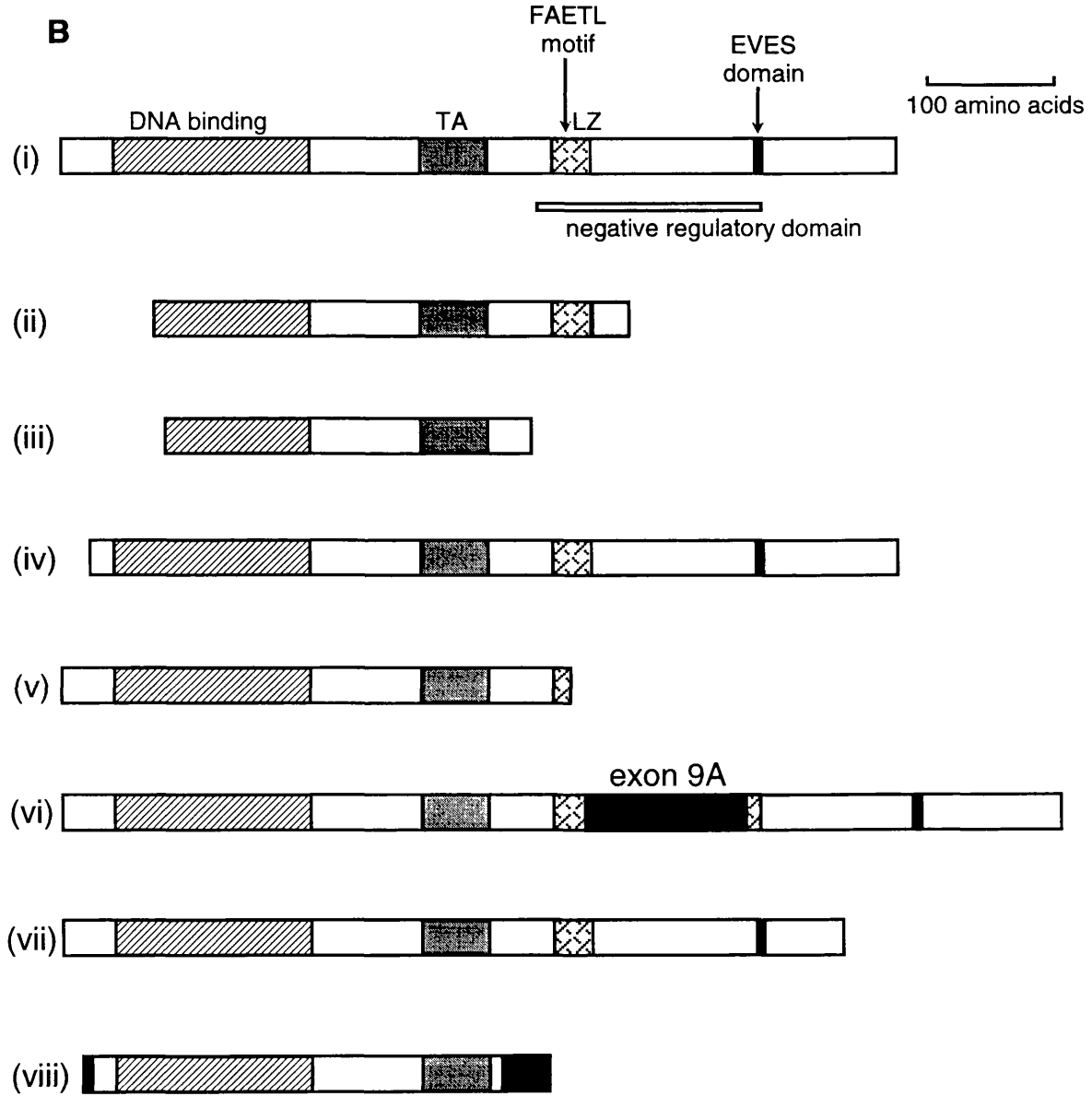
The importance of c-*myb* exon 9A is inferred from the conservation of this exon through evolution. In particular the 12 C-terminal amino acid residues are conserved in a very diverse group of organisms and genes including mouse B-*myb*; c-*myb* from human, mouse, chicken, sea urchin, *Drosophila* and now the cat (Figure 7.4B and (211;212;237;239)). Thus exon 9A has been retained through evolution, suggesting a significant biological function of this isoform of c-*myb*.

Further evidence that this region of the c-*myb* gene plays an important function comes from the demonstration that exon 9 has been the site of insertional mutagenesis in at least two cases in myelomonocytic and promonocytic leukaemias (Figure 8.2(v) and (68;102)). Furthermore, a number of lines of evidence suggest that disruption of the c-*myb* gene in the region around exon 9 which encompasses the putative leucine zipper domain results in increased transcriptional activity. Firstly, a bovine T cell line in which exon 9 is deleted was shown to have a three fold increase in transcriptional activity compared to wild-type c-*myb* when cotransfected with reporter plasmids containing *myb* binding sites (210). Secondly, a study which involved the deletion of part of the negative regulatory domain encompassing the putative leucine zipper motif resulted in a significant increase in transcriptional transactivation which was also associated with an increased transforming activity (204). More recently, a human leukaemia cell line, TK-6, was found to have a c-*myb* gene which was truncated at the 3` end of exon 9, possibly due to a chromosomal translocation (241). Again, this form was associated with an increase in transactivation activity. Therefore, it appears that the disruption of c-*myb* within or near the putative leucine zipper which spans exons 9 and 10, is improtant for transactivation by c-*myb* and therefore may be important during transformation.

Figure 8.2 Structure of the c-*myb* gene and protein. (A) Diagrammatic representation of the c-*myb* gene structure. Exons 1 to 15 are shown as solid bars; the alternative exon 9A is shown as an open bar. The arrows represent the sites and orientations of the most common retroviral insertions in murine promonocytic leukaemias and avian B cell lymphomas. The approximate position of the transcriptional pause (attenuation) site in intron 1 is indicated. The positions of the exons are based on the published human c-*myb* gene (GenBank accession number U22376). (B) A comparison of the various c-Myb protein structures is shown. The positions of the DNA binding domain, the transcriptional transactivation domain (TA), the putative leucine zipper domain (LZ) are shown as well as the approximate positions of the negative regulatory domain and the FAETL and EVES motifs. This figure is based on the human c-Myb protein.

(i) Wild-type 640 amino acid (p75) c-Myb protein structure. (ii) Avian myeloblastosis virus (AMV) v-Myb protein; (iii) the avian leukaemia virus E26 v-Myb protein which is expressed as a *gag-myb-ets* fusion protein; (iv) N-terminal truncated protein. The structure shown illustrates the minimal N-terminal truncation seen following retroviral integration within intron 1 resulting in a protein which is truncated by 20 amino acids, (72), more commonly integration occurs within introns 2 and 3 resulting in truncation of 47 or 71 amino acids respectively, promoter insertion (68;69); (v) C-terminal truncation following retroviral integration within exon 9 resulting in truncation of 240 or 248 amino acids following use of the proviral polyadenylation signal (68;102); (vi) c-Myb protein arising form the use of the alternative exon 9A (solid box) (243;237;235); (vii) proviral integration within intron 14 results in truncation of c-Myb by 38 amino acids (101); (viii) structure of c-Myb from clone pMbm-2 which contains unique 5` sequences which replaces exon 1 following use of the second c-*myb* promoter. This clone also contains an alternatively spliced exon between exons 8 and 9 (exon 8A, solid box) which results in the premature termination of these transcripts (166;243).


A
1
2
3
4 5
6
7 8
9
10
13
14
15
5`
3`
11/12
9A
1kb
transcriptional
pause site


B
FAETL
motif
EVES
domain
100 amino acids
DNA binding
TA
LZ
(i)
negative regulatory domain
(ii)
(iii)
(iv)
(v)
exon 9A
(vi)
(vii)
(viii)

As discussed above, the *Fit*-1 locus lies within a cluster of other common retroviral insertion loci on mouse chromosome 10 (Figure 8.1). Therefore, it will be interesting to see if the exon 9A isoform of c-*myb* is also upregulated in cell lines which are rearranged at *Ahi*-1 and *Mml*I. The p/m 16i and G1500 44i mouse cell lines which are rearranged at *Ahi*-1 will serve as useful tools in this analysis. Furthermore, the discovery of an additional cell line derived from a feline thymic lymphoma which is rearranged at *Fit*-1, the FT-G cell line (H. Tsujimoto, unpublished), will clearly reinforce the conclusions on the effects of proviral insertion at *Fit*-1 seen in the FT1 cell line if it proves to also upregulate exon 9A-containing c-*myb* expression.

Evidence presented in this thesis suggest that the enhanced expression observed in the FT1 cell line is, at least in part due to an increase in the half-life of the c-*myb* transcripts (Figure 6.5). As discussed in chapter 6, c-*myb* has been shown to be controlled by attenuation of transcriptional elongation as well as by changes in RNA stabilty (217;218;220). Thus by decreasing the turnover of c-*myb* mRNA in the FT1 cell line, over-expression of the transcript would be expected, an event which has been shown to be associated with tumorigenesis (189). It would be interesting to determine whether, in addition to an increased stability, c-*myb* is less prone to attenuation in the FT1 cell line, or whether an increase in the mRNA half-life is solely responsible for the observed increase in steady-state levels of c-*myb* transcript.

It is difficult to imagine how proviral insertion at *Fit*-1 could result in an increase in stability of the c-*myb* message, particularly when the majority of insertional activation of the c-*myb* gene seen previously involves truncation of the gene product (186). However, the enforced expression of full-length c-*myb* has been associated with increased transformation *in vitro* (189). A possible conjecture is that proviral insertion affects promoter usage resulting in a transcript which is more stable; or perhaps retroviral insertion promotes alternative splicing using exon 9A, again resulting in a more stable transcript. Interestingly, a c-Myb protein associated with leukaemia which is truncated within exon 9 due to

retroviral integration exhibited a four-fold increased half-life compared with wild type c-Myb protein (221). Such C-terminally truncated c-Myb proteins have been shown to produce a higher level of transformation than cells expressing wild type c-Myb (189). More recently, it has been shown that an instability determinant is located within the putative leucine zipper domain (234). Therefore, it is possible that protein stabilisation may represent a common mechanism of oncogenic activation. It will be informative to screen the RNAs from actinomycin D treated FT1 cells shown in Figure 6.6 for expression levels of exon 9A to examine whether this isoform of c-*myb* is more stable in the FT1 cell line. Furthermore, the development of a c-*myb* exon 9A specific antibody within our laboratory (M. Campbell) will also serve as a useful to study the stability of the exon 9A-containing protein in the FT1 cell line.

As alluded to above, FeLV integration at *Fit*-1 may affect promoter usage in the c-*myb* gene. Indeed, a second promoter has been identified in the human c-*myb* gene (242). This promoter was identified from the cDNA clone pMbm-2 from the T-ALL cell line CEM which was previously found to contain a unique 26bp sequence replacing exon 1 (243). The second promoter (P2) was mapped to a highly conserved region at the 3` terminus of intron 1 (242). Interestingly, this clone from CEM cells was also shown to have undergone alternative splicing between exons 8 and 9 resulting in a C-terminal truncated Myb protein (Figure 8.2(viii) and (243)). It will be interesting to examine the level of exon 9A containing c-*myb* expression in this cell line since the use of an alternative promoter may also be influencing alternative splicing using exon 9A of the c-*myb* gene. In addition, cDNA cloning of the exon 9A-containing form of c-*myb* from FT1 cells to analyse any differences in the 5` end of c-*myb* genes including exon 9A would also tell whether this is a general phenomenon associated with alternative splicing of c-*myb*. Alternatively, an analogy with the c-*myc* oncogene may be drawn where retroviral insertion appears to favour transcripts initiated from the P1 promoter in thymic lymphomas (244;245). Although the functional significance of this is unknown, it was suggested that alterations in P1/P2 ratios may serve as good indicators of c-*myc* deregulation in thymic lymphomas.

Proviral insertion at *Fit*-1 may conceivably affect the methylation state of the c-*myb* gene in the FT1 cell line. CpG dinucleotides are frequently modified at the 5` position of the cytosine ring by methylation (108), and hypomethylation has been shown to be associated with an increase in gene activity (160). Furthermore, the ability of a provirus to affect the methylation state of flanking sequence has been demonstrated previously and this was found to be associated with decreased gene activity (112). Thus, using the PAC contig, it will be informative to identify regions where analysis could be performed using methylation sensitive restriction enzymes to detect evidence of heterogeneity which corresponds with activity of c-*myb*.

Whatever the mode of activation of the c-*myb* oncogene in the FT1 cell line, its enhanced expression raises the novel prospect of cooperation between c-*myc* and c-*myb* in T cell lymphomagenesis since this cell line is also rearranged at c-*myc*. Support for this observation comes from several observations: (i) a role for c-Myb in apoptosis has recently been described in T cells and myeloid cells which directly correlates with the downregulation of the Bcl-2 anti-apoptotic protein, a known Myc collaborator (134;140;183;200); (ii) recent data show that in bovine vascular smooth muscle cells c-*myc* and c-*myb* can collaborate to promote cell cycle progression to S-phase (193); (iii) both c-*myb* and c-*myc* have been shown to be upregulated in patients with acute myeloid leukaemia and this was shown to be associated with enhanced stability of both RNA species (227); (iv) the protein product of the *Pim*-1 oncogene, a known c-*myc* collaborator in lymphoid transformation, has been shown to bind to the transcriptional co-activator, p100, stimulating c-Myb transcriptional activity (216); (v) finally, mice carrying a v-*myb* transgene directed to the T cell compartment develop T cell lymphomas, many of which show evidence of insertional activation at c-*myc* when accelerated by retroviral infection (K. Weston, ICR, London, personal communication). This last observation provides more direct evidence for synergy between c-*myb* and c-*myc* in T cell lymphoma development.

This evidence raises the important questions of the mechanisms by which Myb and Myc synergise to promote the transformation of haematopoietic cells and why the activation of both genes never been observed before in c-*myc* transgenic mice. Leverson *et al* have attempted to address the latter question by suggesting that c-*myb* activation is a complex process involving multiple activating mutations which are unlikely to be observed by retroviral insertional mutagenesis of this gene. They also suggested that *Pim*-1 may stimulate a number of transcription factors so that its activation of c-*myb* alone is not sufficient to promote *myc*-initiated transformation (216).

To examine the synergy between c-*myb* and c-*myc* further, a transgenic cross has been set up in our laboratory (M. Campbell) in collaboration with K. Weston (ICR, London) which should lead to a better understanding of the role of c-*myb* in *Myc*-induced tumorigenesis. It would also be interesting to cross *Pim*-1 and c-*myb* transgenic mice to examine whether there is an acceleration of tumour onset. If the *Pim*-1 and c-*myb* oncogenes do indeed belong to the same functional pathway then no acceleration should be expected.

The details of c-*myb* and c-*myc* cooperation in lymphomagenesis remain mysterious, particularly in view of their similar roles in the regulation of growth and differentiation and the fact that both proteins can inhibit apoptosis when over-expressed (133;201). However, the recent report which shows evidence of a collaboration between *Pim*-1 and p100 to enhance c-*myb* activity suggests that a novel signal transduction pathway linking c-*myb,* p100 and *Pim*-1 may be involved (216).

Transgenic studies have allowed the assignment of oncogenes to different complementation groups as discussed in chapter 1 (Figure 1.7 and (20)). Is it possible to assign *Fit*-1 and c-*myb* to any of these groups? Previous analyses have suggested a strong association between insertion at *Fit*-1 and c-*myc* activation (25). This analysis also identified proviral integration at both the *Fit*-1 and *Flvi*-2 loci (the feline homologue of *Bmi*-1) in one tumour induced by FeLV-

myc. Furthermore, although one tumour has previously been shown to carry an insertion at both *Fit*-1 and *Pim*-1 (25), the *Fit*-1 insertion was submolar while the *Pim*-1 insertion was more highly represented. This suggested that they probably belong to different tumour clones. Therefore, it is unlikely that *Fit*-1 (and c-*myb*) belong to any of the complementation groups outlined in Figure 1.7 suggesting that they may form a novel complemenation group which can collaborate with c-*myc* in T cell lymphomagenesis.

In conclusion, I have reported the chromosomal location of the *Fit*-1 common integration locus on human chromosome 6 in close proximity to c-*myb* and a number of other common insertion sites identified in the mouse, *Mml*1/*Ahi*-1/*Mis*-2. Furthermore, I have identified that the FT1 cell line which is rearranged at both *Fit*-1 and c-*myc* but not c-*myb,* shows evidence of over-expression of the c-*myb* oncogene. This cell line was also shown to over-express the exon 9A isoform of c-*myb*. Future studies which will be aimed at increasing our understanding of the synergistic relationship between *Fit*-1, c-*myb* and c-*myc* in T cell lymphomagenesis.

REFERENCES

References

1. **Vogelstein, B. and K.W. Kinzler**. 1993. The multistep nature of cancer. Trends in Genetics **9**:138-141.

2. **Neil JC, Fulton R, M. Rigby, and Stewart M**. 1991. Feline Leukaemia Virus: Generation of Pathogenic and Oncogenic Variants. Current Topics in Microbiology and Immunology **171**:67-93.

3. **Levy, L.S., R.E. Fish, and G.B. Baskin**. 1988. Tumorigenic potential of a *myc*-containing strain of feline leukemia virus *in vivo* in domestic cats. Journal of Virology **62**:4770-4773.

4. **Onions, D., G. Lees, D. Forrest, and J. Neil**. 1987. Recombinant feline viruses containing the *myc* gene rapidly produce clonal tumours expressing T-cell antigen receptor gene transcripts. International Journal Of Cancer **40**:40-45.

5. **Bonham, L., P.A. Lobellerich, L.A. Henderson, and L.S. Levy**. 1987. Transforming potential of a *myc*-containing variant of feline leukemia-virus *invitro* in early-passage feline cells. Journal Of Virology:3072-3081.

6. **Neil JC, Forrest D, Doggett DL, and J.I. Mullins**. 1987. The role of feline leukaemia virus in naturally occurring leukaemias. Cancer Surveys **6**:117-137.

7. **Neil, J.C., D. Hughes, R. Mcfarlane, N.M. Wilkie, D.E. Onions, G. Lees, and O. Jarrett**. 1984. Transduction and rearrangement of the *myc* gene by feline leukemia-virus in naturally occurring T-cell leukemias. Nature **308**:814-820.

8. **Levy, L.S., M.B. Gardner, and J.W. Casey**. 1984. Isolation of a feline leukemia provirus containing the oncogene *myc* from a feline lymphosarcoma. Nature **308**:853-856.

9. **Mullins, J.I., D.S. Brody, R.C. Binari, and S.M. Cotter**. 1984. Viral transduction of c-*myc* gene in naturally occurring feline leukemias. Nature **308**:856-858.

10. **Peters, G.** 1990. Oncogenes at viral integration sites. Cell Growth & Differentiation **1**:503-510.

11. **Hayward, W.S., B.G. Neel, and S.M. Astrin**. 1981. Activation of a cellular *onc* gene by promoter insertion in ALV- induced lymphoid leukosis. Nature **290**:475-480.

12. **Jonkers, J. and A. Berns**. 1996. Retroviral insertional mutagenesis as a strategy to identify cancer genes. Biochimica et Biophysica Acta - Reviews on Cancer **1287**:29-57.

13. **Berns, A.** 1991. Tumorigenesis in transgenic mice - identification and characterization of synergizing oncogenes. Journal Of Cellular Biochemistry **47**:130-135.

14. **Allen, J.D. and A. Berns**. 1996. Complementation tagging of cooperating oncogenes in knockout mice. Seminars In Cancer Biology **7**:299-306.

15. **Vogt VM**. 1997. Retroviral Virions and Genomes, p. 27-70. In J.M. Coffin, Hughes SH, and H. Varmus (eds.), Retroviruses. Cold Spring Harbor Laboratory Press.

16. **Hardy WD Jr**. 1993. Feline Oncoretroviruses, p. 109-180. In Levy JA (ed.), The Retroviridae. Plenum Press.

17. **Poirier, Y., C. Kozak, and P. Jolicoeur**. 1988. Identification of a common helper provirus integration site in abelson murine leukemia virus-induced lymphoma DNA. Journal Of Virology **62**:3985-3992.

18. **Jiang, X., L. Villeneuve, C. Turmel, C.C. Kozak, and P. Jolicoeur**. 1994. The *myb* and *ahi*-1 genes are physically very closely linked on mouse chromosome-10. Mammalian Genome **5**:142-148.

19. **Haupt, Y., W.S. Alexander, G. Barri, S.P. Klinken, and J.M. Adams**. 1991. Novel zinc finger gene implicated as *myc* collaborator by retrovirally accelerated lymphomagenesis in e-mu-*myc* transgenic mice. Cell **65**:753-763.

20. **Vanlohuizen, M., S. Verbeek, B. Scheijen, E. Wientjens, H. Vandergulden, and A. Berns**. 1991. Identification of cooperating oncogenes in e-mu-*myc* transgenic mice by provirus tagging. Cell **65**:737-752.

21. **Vanlohuizen, M., M. Frasch, E. Wientjens, and A. Berns**. 1991. Sequence similarity between the mammalian *bmi*-1 protooncogene and the *Drosophila* regulatory genes *psc* and su(z)2. Nature **353**:353-355.

22. **Brunk, B.P., E.C. Martin, and P.N. Adler**. 1991. Drosophila genes posterior sex combs and suppressor 2 of zeste encode proteins with homology to the murine *bmi*-1 oncogene. Nature **353**:351-353.

23. **Levy, L.S. and P.A. Lobellerich**. 1992. Insertional mutagenesis of *flvi*-2 in tumors induced by infection with LC-FeLV, a *myc*-containing strain of feline leukemia-virus. Journal Of Virology **66**:2885-2892.

24. **Levy, L.S., P.A. Lobellerich, and J. Overbaugh**. 1993. *Flvi*-2, a target of retroviral insertional mutagenesis in feline thymic lymphosarcomas, encodes *bmi*-1. Oncogene **8**:1833-1838.

25. **Tsatsanis, C., R. Fulton, K. Nishigaki, H. Tsujimoto, L. Levy, A. Terry, D. Spandidos, D. Onions, and J.C. Neil**. 1994. Genetic-determinants of feline leukemia virus-induced lymphoid tumors - patterns of proviral insertion and gene rearrangement. Journal of Virology **68**:8296-8303.

26. **Bordereaux, D., S. Fichelson, B. Sola, P.E. Tambourin, and S. Gisselbrecht**. 1987. Frequent involvement of the *fim*-3-region in friend murine leukemia virus-induced mouse myeloblastic leukemias. Journal Of Virology **61**:4043-4045.

27. **Sola, B., D. Simon, M.G. Mattei, S. Fichelson, D. Bordereaux, P.E. Tambourin, J.L. Guenet, and S. Gisselbrecht**. 1988. *Fim*-1, *fim*-2/c-*fms*, and *fim*-3, 3 common integration sites of friend murine leukemia-virus in myeloblastic leukemias, map to mouse chromosome-13, chromosome-18, and chromosome-3, respectively. Journal of Virology **62**:3973-3978.

28. **Bartholomew, C. and J.N. Ihle**. 1991. Retroviral insertions 90 kilobases proximal to the *evi*-1 myeloid transforming gene activate transcription from the normal promoter. Molecular and Cellular Biology **11**:1820-1828.

29. **Vijaya, S., D.L. Steffen, C. Kozak, and H.L. Robinson**. 1987. *Dsi*-1, a region with frequent proviral insertions in moloney murine leukemia virus-induced rat thymomas. Journal of Virology **61**:1164-1170.

30. **Mucenski, M.L., B.A. Taylor, J.N. Ihle, J.W. Hartley, H.C. Morse, N.A. Jenkins, and N.G. Copeland**. 1988. Identification of a common ecotropic viral integration

site, *evi-* 1, in the DNA of AKXD murine myeloid tumors. Molecular and Cellular Biology **8**:301-308.

31. **Morishita, K., D.S. Parker, M.L. Mucenski, N.A. Jenkins, N.G. Copeland, and J.N. Ihle**. 1988. Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid-leukemia cell-lines. Cell **54**:831-840.

32. **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. Molecular And Cellular Biology **10**:4658-4666.

33. **Largaespada, D.A., J.D. Shaughnessy, 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. Journal of Virology **69**:5095-5102.

34. **Justice, M.J., H.C. Morse, N.A. Jenkins, and N.G. Copeland**. 1994. Identification of *evi*-3, a novel common site of retroviral integration in mouse AKXD B-cell lymphomas. Journal of Virology **68**:1293-1300.

35. **Liao, X.B., A.M. Buchberg, N.A. Jenkins, and N.G. Copeland**. 1995. *Evi*-5, a common site of retroviral integration in AKXD T-cell lymphomas, maps near *gfi*-1 on mouse chromosome-5. Journal Of Virology **69**:7132-7137.

36. **Liao, X., Y. Du, H.C. Morse III, N.A. Jenkins, and N.G. Copeland**. 1997. Proviral integrations at the *Evi5* locus disrupt a novel 90 kDa protein with homology to the *Tre2* oncogene and cell-cycle regulatory proteins. Oncogene **14**:1023-1029.

37. **Schmidt, T., M. Zornig, R. Beneke, and T. Moroy**. 1996. MoMuLV proviral integrations identified by sup-f selection in tumors from infected *myc/pim* bitransgenic mice correlate with activation of the *gfi*-1 gene. Nucleic Acids Research **24**:2528-2534.

38. **Zornig, M., T. Schmidt, H. Karsunky, A. Grzeschiezek, and T. Moroy**. 1996. Zinc-finger protein *gfi*-1 cooperates with *myc* and *pim*-1 in T- cell lymphomagenesis by reducing the requirements for IL-2. Oncogene **12**:1789-1801.

39. **Scheijen, B., J. Jonkers, D. Acton, and A. Berns**. 1997. Characterization of *pal*-1, a common proviral insertion site in murine leukemia virus-induced lymphomas of c-*myc* and *pim*-1 transgenic mice. Journal Of Virology **71**:9-16.

40. **Silver, J. and C. Kozak**. 1986. Common proviral integration region on mouse chromosome-7 in lymphomas and myelogenous leukemias induced by friend murine leukemia-virus. Journal of Virology **57**:526-533.

41. **Lammie, G.A., R. Smith, J. Silver, S. Brookes, C. Dickson, and G. Peters**. 1992. Proviral insertions near cyclin D1 in mouse lymphomas - a parallel for *bcl1* translocations in human B-cell neoplasms. Oncogene **7**:2381-2387.

42. **Tsujimoto, H., R. Fulton, K. Nishigaki, Y. Matsumoto, A. Hasegawa, A. Tsujimoto, S. Cevario, S.J. Obrien, A. Terry, D. Onions, and J.C. Neil**. 1993. A common proviral integration region, *fit*-1, in T-cell tumors induced by *myc*-containing feline leukemia viruses. Virology **196**:845-848.

43. **Bendavid, Y., E.B. Giddens, and A. Bernstein**. 1990. Identification and mapping of a common proviral integration site *fli*- 1 in erythroleukemia-cells induced by friend murine leukemia- virus. Proceedings of the National Academy of Sciences of the USA **87**:1332-1336.

44. **Bergeron, D., L. Poliquin, C.A. Kozak, and E. Rassart**. 1991. Identification of a common viral integration region in cas-br-e murine leukemia virus-induced non-T-cell, non-B-cell lymphomas. Journal Of Virology **65**:7-15.

45. **Ott, D.E., J. Keller, and A. Rein**. 1994. 10A1 MuLV induces a murine leukemia that expresses hematopoietic stem-cell markers by a mechanism that includes *fli*-1 integration. Virology **205**:563-568.

46. **Bendavid, Y., E.B. Giddens, K. Letwin, and A. Bernstein**. 1991. Erythroleukemia induction by friend murine leukemia-virus - insertional activation of a new member of the *ets* gene family, fli-1, closely linked to c-*ets*-1. Genes & Development **5**:908-918.

47. **Bendavid, Y., M.R. Bani, B. Chabot, A. Dekoven, 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. Molecular and Cellular Biology **12**:4449-4455.

48. **Lu, S.J., S. Rowan, M.R. Bani, and Y. Bendavid**. 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. Proceedings of the National Academy of Sciences of the USA **91**:8398-8402.

49. **Levesque, K.S., L. Bonham, and L.S. Levy**. 1990. *Flvi*-1, a common integration domain of feline leukemia-virus in naturally-occurring lymphomas of a particular type. Journal of Virology **64**:3455-3462.

50. **Levesque, K.S., M.G. Mattei, and L.S. Levy**. 1991. Evolutionary conservation and chromosomal localization of *flvi*-1. Oncogene **6**:1377-1379.

51. **Jonkers, J., H.C. Korswagen, D. Acton, M. Breuer, and A. Berns**. 1997. Activation of a novel proto-oncogene, *Fratl*, contributes to progression of mouse T-cell lymphomas. EMBO Journal **16**:441-450.

52. **Friedrich, R.W., M. Veit, D. Eisel, U. Friedich, M. Pass, and C.A. Kozak**. 1997. *Fre2*, a proviral integration site of Friend murine leukemia virus that is closely linked to Fv2. Leukemia **11**:619-623.

53. **Gilks, C.B., S.E. Bear, H.L. Grimes, and P.N. Tsichlis**. 1993. Progression of interleukin-2 (IL-2)-dependent rat T-cell lymphoma lines to IL-2-independent growth following activation of a gene (*gfi*- 1) encoding a novel zinc finger protein. Molecular And Cellular Biology **13**:1759-1768.

54. **Grimes, H.L., C.B. Gilks, T.O. Chan, S. Porter, and P.N. Tsichlis**. 1996. The *gfi*-1 protooncoprotein represses bax expression and inhibits T- cell death. Proceedings of the National Academy of sciences of the USA **93**:14569-14573.

55. **Flubacher, M.M., S.E. Bear, and P.N. Tsichlis**. 1994. Replacement of interleukin-2 (IL-2)-generated mitogenic signals by a mink cell focus-forming (MCF) or xenotropic virus-induced IL- 9- dependent autocrine loop - implications for MCF virus-induced leukemogenesis. Journal Of Virology **68**:7709-7716.

56. **Villemur, R., Y. Monczak, E. Rassart, C. Kozak, and P. Jolicoeur**. 1987. Identification of a new common provirus integration site in gross passage a murine leukemia virus-induced mouse thymoma DNA. Molecular and Cellular Biology 7:512-522.

57. **Voronova, A.F. and B.M. Sefton**. 1986. Expression of a new tyrosine protein-kinase is stimulated by retrovirus promoter insertion. Nature **319**:682-685.

58. **Shin, S. and D.L. Steffen**. 1993. Frequent activation of the *lck* gene by promoter insertion and aberrant splicing in murine leukemia virus-induced rat lymphomas. Oncogene **8**:141-149.

59. **Villeneuve, L., X. Jiang, C. Turmel, C.A. Kozak, and P. Jolicoeur**. 1993. Long-range mapping of *Mis*-2, a common provirus integration site identified in murine leukemia virus-induced thymomas and located 160 kilobase pairs downstream of *myb*. Journal Of Virology **67**:5733-5739.

60. **Tsichlis, P.N., P.G. Strauss, and L.F. Hu**. 1983. A common region for proviral DNA integration in MoMuLV-induced rat thymic lymphomas. Nature **302**:445-449.

61. **Koehne, C.F., P.A. Lazo, K. Alves, J.S. Lee, P.N. Tsichlis, and P.V. Odonnell**. 1989. The *mlvi*-1 locus involved in the induction of rat T-cell lymphomas and the *pvt-1/mis*-1 locus are identical. Journal Of Virology **63**:2366-2369.

62. **Tsichlis, P.N., M.A. Lohse, C. Szpirer, J. Szpirer, and G. Levan**. 1985. Cellular DNA regions involved in the induction of rat thymic lymphomas (*mlvi*-1, *mlvi*-2, *mlvi*-3, and c-*myc*) represent independent loci as determined by their chromosomal map location in the rat. Journal Of Virology **56**:938-942.

63. **Tsichlis, P.N., B.M. Shepherd, and S.E. Bear**. 1989. Activation of the *mlvi-1/mis-1/pvt*-1 locus in moloney murine leukemia virus-induced T-cell lymphomas. Proceedings of the National Academy of Sciences of the USA **86**:5487-5491.

64. **Lazo, P.A., J.S. Lee, and P.N. Tsichlis**. 1990. Long-distance activation of the *myc* protooncogene by provirus insertion in *mlvi*-1 or *mlvi*-4 in rat T-cell lymphomas. Proceedings of the National Academy of Sciences of the USA **87**:170-173.

65. **Tsichlis, P.N., J.S. Lee, S.E. Bear, P.A. Lazo, C. Patriotis, E. Gustafson, S. Shinton, N.A. Jenkins, N.G. Copeland, K. Huebner, C. Croce, G. Levan, and C. Hanson**. 1990. Activation of multiple genes by provirus integration in the *mlvi*- 4 locus in T-cell lymphomas induced by moloney murine leukemia- virus. Journal of Virology **64**:2236-2244.

66. **Koller, R., M. Krall, B. Mock, J. Bies, V. Nazarov, and L. Wolff**. 1996. *Mml1*, a new common integration site in murine leukemia virus- induced promonocytic leukemias maps to mouse chromosome-10. Virology **224**:224-234.

67. **Bies, J, Richard Koller, Barbara Hoffman, Arshad Amanullah, Beverly Mock, and Linda Wolff**. 1997. MuLV-insertional mutagenesis of c-*myb* and *Mml1* in a murine model for promonocytic leukemia. Leukemia (supplement) **11**:247-250.

68. **Shenong, G.L.C., H.C. morse, M. Potter, and J.F. Mushinski**. 1986. 2 modes of c-*myb* activation in virus-induced mouse myeloid tumors. Molecular and Cellular Biology **6**:380-392.

69. **Shenong, G.L.C. and L. Wolff**. 1987. Moloney murine leukemia virus-induced myeloid tumors in adult balb/c mice - requirement of c-*myb* activation but lack of v-*abl* involvement. Journal of Virology **61**:3721-3725.

70. **Wolff, L., R. Koller, and W. Davidson**. 1991. Acute myeloid-leukemia induction by amphotropic murine retrovirus (4070A) - clonal integrations involve c-*myb* in some but not all leukemias. Journal Of Virology **65**:3607-3616.

71. **Kanter, M.R., R.E. Smith, and W.S. Hayward**. 1988. Rapid induction of B-cell lymphomas - insertional activation of c-*myb* by avian-leukosis virus. Journal Of Virology **62**:1423-1432.

72. **Jiang, W.P., M.R. Kanter, I. Dunkel, R.G. Ramsay, K.L. Beemon, and W.S. Hayward**. 1997. Minimal truncation of the c-*myb* gene product in rapid-onset B- cell lymphoma. Journal Of Virology **71**:6526-6533.

73. **Corcoran, L.M., J.M. Adams, A.R. Dunn, And S. Cory**. 1984. Murine T-lymphomas in which the cellular *myc* oncogene has been activated by retroviral insertion. Cell **37**:113-122.

74. **Selten, G., H.T. Cuypers, M. Zijlstra, C. Melief, and A. Berns**. 1984. Involvement of c-*myc* in MuLV-induced T-cell lymphomas in mice - frequency and mechanisms of activation. EMBO Journal **3**:3215-3222.

75. **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.

76. **Dolcetti, R., S. Rizzo, A. Viel, R. Maestro, V. Dere, G. Feriotto, and M. Boiocchi**. 1989. N-*myc* activation by proviral insertion in MCF 247-induced murineT- cell lymphomas . Oncogene **4**:1009-1014.

77. **Van Lohuizen, M., M. Breuer, and A. Berns**. 1989. N-*myc* is frequently activated by proviral insertion in MuLV- induced T cell lymphomas. EMBO Journal **8**:133-136.

78. **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 . Journal Of Virology **62**:4752-4755.

79. **Cuypers, H.T., G. Selten, W. Quint, M. Zijlstra, E.R. Maandag, W. Boelens, P. Vanwezenbeek, 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.

80. **Saris, C.J.M., J. Domen, and A. Berns**. 1991. The *pim*-1 oncogene encodes two related protein-serine/threonine s by alternative initiation at AUG and CUG. EMBO Journal **10**:655-664.

81. **Vanderlugt, N.M.T., J. Domen, E. Verhoeven, K. Linders, H. Vandergulden, 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 Journal **14**:2536-2544.

82. **Allen, J.D., E. Verhoeven, J. Domen, M. Vandervalk, and A. Berns**. 1997. *Pim*-2 transgene induces lymphoid tumors, exhibiting potent synergy with c-*myc*. Oncogene **15**:1133-1141.

83. **Barker, C.S., S.E. Bear, T. Keler, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, R.S. Yeung, and P.N. Tsichlis**. 1992. Activation of the prolactin receptor gene by promoter insertion in a moloney murine leukemia virus-induced rat thymoma. Journal of Virology **66**:6763-6768.

84. **Mueller, R.E., L. Baggio, C.A. Kozak, and J.K. Ball**. 1992. A common integration locus in type-B retrovirus-induced thymic lymphomas. Virology **191**:628-637.

85. **Habets, G.G.M., E.H.M. Scholtes, D. Zuydgeest, R.A. Vanderkammen, J.C. Stam, A. Berns, and J.G. Collard**. 1994. Identification of an invasion-inducing gene, *tiam*-1, that encodes a protein with homology to gdp-gtp exchangers for rho- like proteins. Cell **77**:537-549.

86. **Breuer, M.L., H.T. Cuypers, and A. Berns**. 1989. Evidence for the involvement of *pim*-2, a new common proviral insertion site, in progression of lymphomas. EMBO Journal **8**:743-747.

87. **Shinto, Y., M. Morimoto, M. Katsumata, A. Uchida, K. Aozasa, M. Okamoto, T. Kurosawa, T. Ochi, M.I. Greene, and Y. Tsujimoto**. 1995. Moloney murine leukemia-virus infection accelerates lymphomagenesis in e-mu-*bcl*-2 transgenic mice. Oncogene **11**:1729-1736.

88. **Stewart, M., A. Terry, M. O'Hara, E. Cameron, D. Onions, and J.C. Neil**. 1996. *Til*-1 - a novel proviral insertion locus for moloney murine leukemia- virus in lymphomas of CD2-*myc* transgenic mice. Journal of General Virology **77**:443-446.

89. **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 pebp2 alpha a (cbfa1): evidence for a new myc collaborating oncogene. Proceedings of The National Academy of Sciences of the USA **94**:8646-8651.

90. **Bear, S.E., A. Bellacosa, P.A. Lazo, N.A. Jenkins, N.G. Copeland, C. Hanson, G. Levan, and P.N. Tsichlis**. 1989. Provirus insertion in *tpl*-1, an *ets*-1-related oncogene, is associated with tumor progression in moloney murine leukemia virus-induced rat thymic lymphomas. Proceedings of The National Academy of Sciences of the USA **86**:7495-7499.

91. **Makris, A., C. Patriotis, S.E. Bear, and P.N. Tsichlis**. 1993. Genomic organization and expression of *tpl*-2 in normal-cells and moloney murine leukemia virus-induced rat T-cell lymphomas - activation by provirus insertion. Journal of Virology **67**:4283-4289.

92. **Patriotis, C., A. Makris, S.E. Bear, and P.N. Tsichlis**. 1993. Tumor progression locus-2 (*tpl*-2) encodes a protein-kinase involved in the progression of rodent T-cell lymphomas and in T- cell activation. Proceedings of The National Academy of Sciences of the USA **90**:2251-2255.

93. **Salmeron, A., T.B. Ahmad, G.W. Carlile, D. Pappin, R.P. Narsimhan, and S.C. Ley**. 1996. Activation of MEK-1 and SEK-1 by *Tpl*-2 proto-oncoprotein, a novel MAP kinase kinase kinase. EMBO Journal **15**:817-826.

94. **Tremblay, P.J., C.A. Kozak, and P. Jolicoeur**. 1992. Identification of a novel gene, *vin*-1, in murine leukemia virus- induced T-cell leukemias by provirus insertional mutagenesis. Journal of Virology **66**:1344-1353.

95. **Hanna, Z., M. Jankowski, P. Tremblay, X.Y. Jiang, A. Milatovich, U. Francke, and P. Jolicoeur**. 1993. The *vin*-1 gene, identified by provirus insertional mutagenesis, is the cyclin D2. Oncogene **8**:1661-1666.

96. **Neil, J.C. and D.E. Onions**. 1985. Feline leukemia viruses - molecular-biology and pathogenesis. Anticancer Research **5**:49-63.

97. **Kung, H.J., C. Boerkoel, and T.H. Carter**. 1991. Retroviral mutagenesis of cellular oncogenes: A review with insights into the mechanisms of insertional activation. Current Topics in Microbiology and Immunology **171**:1-25.

98. **Serfling, E., M. Jasin, and W. Schaffner**. 1985. Enhancers and eukaryotic gene-transcription. Trends in Genetics **1**:224-230.

99. **Fulton, R., M. Plumb, L. Shield, and J.C. NEIL**. 1990. Structural diversity and nuclear-protein binding-sites in the long terminal repeats of feline leukemia-virus. Journal of Virology **64**:1675-1682.

100. **Miura, T., M. Shibuya, H. Tsujimoto, M. Fukasawa, and M. Hayami**. 1989. Molecular-cloning of a feline leukemia provirus integrated adjacent to the c-*myc* gene in a feline T-cell leukemia-cell line and the unique structure of its long terminal repeat. Virology **169**:458-461.

101. **Nazarov, V. and L. Wolff**. 1995. Novel integration sites at the distal 3' end of the c-*myb* locus in retrovirus-induced promonocytic leukemias. Journal of Virology **69**:3885-3888.

102. **Mukhopadhyaya, R. and L. Wolff**. 1992. New sites of proviral integration associated with murine promonocytic leukemias and evidence for alternate modes of c-*myb* activation. Journal Of Virology **66**:6035-6044.

103. **Swanstrom, R., R.C. Parker, H.E. Varmus, and J.M. Bishop**. 1983. Transduction of a cellular oncogene - the genesis of rous- sarcoma virus. Proceedings of the National Academy of Sciences of the USA **80**:2519-2523.

104. **Cuypers, H.T.M., G.C. Selten, M. Zijlstra, R.E. Degoede, C.J. Melief, and A.J. Berns**. 1986. Tumor progression in murine leukemia virus-induced T-cell lymphomas - monitoring clonal selections with viral and cellular probes. Journal of Virology **60**:230-241.

105. **Hesketh R**. 1997. The Oncogene and Tumour Suppressor Gene Facts Book. Academic Press,Harcourt Brace & Company,

106. **Duyk, G.M., S.W. Kim, R.M. Myers, and D.R. Cox**. 1990. Exon trapping - a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. . Proceedings of the National Academy of Sciences of the USA **87**:8995-8999.

107. **Bird, A.P.** 1987. CpG islands as gene markers in the vertebrate nucleus. Trends In Genetics **3**:342-347.

108. **Bird, A.P.** 1986. CpG-rich islands and the function of DNA methylation. Nature **321**:209-213.

109. **Cross, S.H. and A.P. Bird**. 1995. CpG islands and genes. Current Opinion in Genetics and Development **5**:309-314.

110. **Lindsay, S. and A.P. Bird**. 1987. Use of restriction enzymes to detect potential gene sequences in mammalian DNA. Nature **327**:336-338.

111. **Haupt, Y., M.L. Bath, A.W. Harris, and J.M. Adams**. 1993. *Bmi*-1 transgene induces lymphomas and collaborates with *myc* in tumorigenesis. Oncogene **8**:3161-3164.

112. **Jahner, D. and R. Jaenisch**. 1985. Retrovirus-induced denovo methylation of flanking host sequences correlates with gene inactivity. Nature **315**:594-597.

113. **Cory, S., M. Graham, E. Webb, L. Corcoran, and J.M. Adams**. 1985. Variant (6-15) translocations in murine plasmacytomas involve a chromosome-15 locus at least 72 kb from the c-*myc* oncogene. EMBO Journal **4**:675-681.

114. **Dallafavera, R., M. Bregni, J. Erikson, D. Patterson, R.C. Gallo, and C.M. Croce.** 1982. Human c-*myc* onc gene is located on the region of chromosome-8 that is translocated in burkitt-lymphoma cells. Proceedings of the National Academy of Sciences of the USA **79**:7824-7827.

115. **Collins, S. and M. Groudine.** 1982. Amplification of endogenous *myc*-related DNA-sequences in a human myeloid-leukemia cell-line. Nature **298**:679-681.

116. **Sheiness D and J.M. Bishop.** 1979. DNA and RNA from uninfected vertebrate cells contain nucleotide sequences related to the putative transforming gene of avian myelocytomatosis virus. Journal of Virology **31**:514-521.

117. **Kaczmarek, L., J.K. Hyland, R. Watt, M. Rosenberg, and R. Baserga.** 1985. Microinjected c-*myc* as a competence factor. Science **228**:1313-1315.

118. **Hann, S.R. and R.N. Eisenman.** 1984. Proteins encoded by the human c-*myc* oncogene - differential expression in neoplastic-cells. Molecular And Cellular Biology **4**:2486-2497.

119. **Marcu, K.B., S.A. Bossone, and A.J. Patel.** 1992. Myc function and regulation. Annual Review Of Biochemistry **61**:809-860.

120. **Vogelstein, B. and K.W. Kinzler.** 1993. The multistep nature of cancer. Trends in Genetics **9**:138-141.

121. **Adams, J.M., A.W. Harris, C.A. Pinkert, L.M. Corcoran, W.S. Alexander, S. Cory, R.D. Palmiter, and R.L. Brinster.** 1985. The c-*myc* oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature **318**:533-538.

122. **Schwab, M., K. Alitalo, K.H. Klempnauer, H.E. Varmus, J.M. Bishop, F. Gilbert, G. Brodeur, M. Goldstein, and J. Trent.** 1983. Amplified DNA with limited homology to *myc* cellular oncogene is shared by human neuro-blastoma cell-lines and a neuro-blastoma tumor. Nature **305**:245-248.

123. **Nau, M.M., B.J. Brooks, J. Battey, E. Sausville, A.F. Gazdar, I.R. Kirsch, O.W. McBride, V. Bertness, G.F. Hollis, and J.D. Minna.** 1985. L-*myc*, a new *myc*-related gene amplified and expressed in human small cell lung cancer. Nature **318**:69-73.

124. **Meichle, A., A. Philipp, and M. Eilers.** 1992. The functions of Myc proteins. Biochimica Et Biophysica Acta **1114**:129-146.

125. **Blackwell, T.K., L. Kretzner, E.M. Blackwood, R.N. Eisenman, and H. Weintraub.** 1990. Sequence-specific DNA-binding by the c-Myc protein. Science **250**:1149-1151.

126. **Blackwood, E.M. and R.N. Eisenman.** 1991. Max - a helix-loop-helix zipper protein that forms a sequence- specific DNA-binding complex with Myc. Science **251**:1211-1217.

127. **Littlewood, T.D., B. Amati, H. Land, and G.I. Evan.** 1992. Max and c-Myc Max DNA-binding activities in cell-extracts. Oncogene **7**:1783-1792.

128. **Amati, B., M.W. Brooks, N. Levy, T.D. Littlewood, G.I. Evan, and H. Land.** 1993. Oncogenic activity of the c-Myc protein requires dimerization with Max. Cell **72**:233-245.

129. **Kretzner, L., E.M. Blackwood, and R.N. Eisenman**. 1992. Myc and Max proteins possess distinct transcriptional activities. Nature **359**:426-429.

130. **Amati B and Hartmut L.** 1994. Myc - Max - Mad: a transcription factor network controlling cell cycle progression, differentiation and death. Current Opinion In Genetics & Development **4**, 102-108.

131. **Thompson, C.B., P.B. Challoner, P.E. Neiman, and M. Groudine**. 1985. Levels of c-*myc* oncogene messenger-RNA are invariant throughout the cell-cycle. Nature **314**:363-366.

132. **Freytag, S.O.** 1988. Enforced expression of the c-*myc* oncogene inhibits cell-differentiation by precluding entry into a distinct predifferentiation state in G_0/G_1. Molecular And Cellular Biology **8**:1614-1624.

133. **Evan, G.I., A.H. Wyllie, C.S. Gilbert, T.D. Littlewood, H. Land, M. Brooks, C.M. Waters, L.Z. Penn, and D.C. Hancock**. 1992. Induction of apoptosis in fibroblasts by c-Myc protein. Cell **69**:119-128.

134. **Fanidi, A., E.A. Harrington, and G.I. Evan**. 1992. Cooperative interaction between c-*myc* and *bcl*-2 proto-oncogenes. Nature **359**:554-556.

135. **Bissonnette, R.P., F. Echeverri, A. Mahboubi, and D.R. Green**. 1992. Apoptotic cell-death induced by c-Myc is inhibited by bcl-2. Nature **359**:552-554.

136. **Berns, A., N. Vanderlugt, M. Alkema, M. Vanlohuizen, J. Domen, D. Acton, J. Allen, P.W. Laird, and J. Jonkers**. 1994. Mouse model systems to study multistep tumorigenesis. Cold Spring Harbor Symposia On Quantitative Biology **59**:435-447.

137. **Vanlohuizen, M., S. Verbeek, P. Krimpenfort, J. Domen, C. Saris, T. Radaszkiewicz, and A. Berns**. 1989. Predisposition to lymphomagenesis in *pim*-1 transgenic mice - cooperation with c-*myc* and N-*myc* in murine leukemia-virus induced- tumors. Cell **56**:673-682.

138. **Verbeek, S., M. Vanlohuizen, M. Vandervalk, J. Domen, G. Kraal, and A. Berns**. 1991. Mice bearing the e-mu-*myc* and e-mu-*pim*-1 transgenes develop pre- B- cell leukemia prenatally. Molecular and Cellular Biology **11**:1176-1179.

139. **Stewart, M., E. Cameron, M. Campbell, R. McFarlane, S. Toth, K. Lang, D. Onions, and J.C. Neil**. 1993. Conditional expression and oncogenicity of c-*myc* linked to a CD2 gene dominant control region. International Journal Of Cancer **53**:1023-1030.

140. **Vaux, D.L., S. Cory, and J.M. Adams**. 1988. *Bcl*-2 gene promotes hematopoietic-cell survival and cooperates with c-*myc* to immortalize pre-B-cells. Nature **335**:440-442.

141. **Korsmeyer, S.J.** 1995. Regulators of cell death. Trends.Genet. **11**:101-105.

142. **Miura, T., H. Tsujimoto, M. Fukasawa, Kodama,T. Shibuya, M. Hasegawa, A. and Hayami, M**. 1987. Structural abnormality and over-expression of the *myc* gene in feline leukemias. International Journal Of Cancer **40**:564-569.

143. **Snyder HW Jr, Hardy WD Jr, Zuckerman EE, and Fleissner E**. 1978. Characterisation of a tumour-specific antigen on the surface of feline lymphosarcoma cells. Nature **275 (5681)** : 656-658.

144. **Rickard CG, Post JE, Noronha F, and Barr LM**. 1969. A transmissible virus-induced lymphocytic leukemia of the cat. Journal of National Cancer Institute **42 (6)**:987-1014.

145. **Fulton R, Forrest D, McFarlane R, Onions D, and Neil JC**. 1987. Retroviral transduction of T-cell antigen receptor beta-chain and *myc* genes. Nature **326 (6109)**:190-194.

146. **Theilen GH, Dungworth DL, Kawakami TG, Munn RJ, Ward JM, and Harrold JB**. 1970. Experimental induction of lymphosarcoma in the cat with "C"-type virus. Cancer Research **30 (2)**:401-408.

147. **Miyazawa T, Furuya T, Itagaki S, Tohya Y, Takahashi E, and Mikami T**. 1989. Establishment of a feline T-lymphoblastoid cell line highly sensitive for replication of feline immunodeficiency virus. Arch Virol **108 (1-2)**:131-135.

148. **Foley GE**. 1965. Continuous culture of human lymphoblasts from peripheral blood of a child with acute leukemia. Cancer **18**:522-529.

149. **Weiss A, Wiskocil RL, and Stobo JD**. 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. Journal Of Immunology **133 (1)**:123-128.

150. **Smith SD, Shatsky M, Cohen PS, Warnke R, Link MP, and Glader BE**. 1984. Monoclonal antibody and enzymatic profiles of human malignant T-lymphoid cells and derived cell lines. Cancer Research **44 (12 Pt 1)**:5657-5660.

151. **Maniatis T, Fritsch EF, and Sambrook J**. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

152. **Chomczynski, P. and N. Sacchi**. 1987. Single-step method of rna isolation by acid guanidinium thiocyanate phenol chloroform extraction. Analytical Biochemistry **162**:156-159.

153. **Grunstein M and Hogness DS**. 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.

154. **Southern EM**. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal Of Molecular Biology **98 (3)**:503-517.

155. **Feinberg, A.P. and B. Vogelstein**. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Analytical Biochemistry **132**:6-13.

156. **Sanger F, Nicklen S, and Coulson AR**. 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A **74 (12)**:5463-5467.

157. **Barr NI, Stewart M, C. Tsatsanis, Fulton R, Hu M, Tsujimoto H, and Neil JC**. 1999. The *Fit*-1 common integration locus in human and mouse is closely linked to *MYB*. Mammalian Genome **In press**. (See Appendix 2).

158. **Fanning, T.G.** 1987. Origin and evolution of a major feline satellite DNA. Journal Of Molecular Biology **197**:627-634.

159. **Fourel, G., J. Couturier, Y. Wei, F. Apiou, P. Tiollais, and M.A. Buendia**. 1994. Evidence for long-range oncogene activation by hepadnavirus insertion. EMBO Journal **13**:2526-2534.

160. **Doerfler, W.** 1983. DNA methylation and gene activity. Annual Review Of Biochemistry **52**:93-124.

161. **O'Brien, S.J., H.N. Seuanez, and J.E. Womack**. 1988. Mammalian genome organization - an evolutionary view. Annual Review Of Genetics **22**:323-351.

162. **Lehrach, H., R. Drmanac, J. Hoheisel, Z. Larin, and et al**. 1990. Hybridization Fingerprinting in Genome Mapping and Sequencing. Genome Analusis Volume 1: Genetic and Physical Mapping, Cold Spring Harbor Laboratory Press **1**:39-81.

163. **Rao, P.H., V.V.V.S. Murty, G. Gaidano, R. Hauptschein, R. Dallafavera, and R.S.K. Chaganti**. 1993. Subregional localization of 20 single-copy loci to chromosome-6 by fluorescence insitu hybridization. Genomics **16**:426-430.

164. **Menasce, L.P., V. Orphanos, M. SantibanezKoref, J.M. Boyle, and C.J. Harrison**. 1994. Common region of deletion on the long arm of chromosome-6 in non-hodgkins-lymphoma and acute lymphoblastic-leukemia. Genes Chromosomes & Cancer **10**:286-288.

165. **Zhang, Y., P. Matthiesen, K. WeberMatthiesen, and B. Schlegelberger**. 1996. Frequent deletions of 6q23-24 in B-cell non-hodgkins-lymphomas detected by fluorescence in-situ hybridization. Blood **88**:1503-1503.

166. **Jacobs, S.M., K.M. Gorse, S.J. Kennedy, and E.H. Westin**. 1994. Characterization of a rearrangement in the c-*myb* promoter in the acute lymphoblastic-leukemia cell-line CCRF-CEM. Cancer Genetics And Cytogenetics **75**:31-39.

167. **Devilee, P., M. van Vliet, P. van Sloun, D.N. Kuipers, J. Hermans, P.L. Pearson, and C.J. Cornelisse**. 1991. Allelotype of human breast carcinoma: a second major site for loss of heterozygosity is on chromosome 6q. Oncogene **6**:1705-1711.

168. **Millikin, D., E. Meese, B. Vogelstein, C. Witkowski, and J. Trent**. 1991. Loss of heterozygosity for loci on the long arm of chromosome-6 in human-malignant melanoma. Cancer Research **51**:5449-5453.

169. **Suto, Y., Y. Sato, S.D. Smith, J.D. Rowley, and S.K. Bohlander**. 1997. A t(6;12)(q23;p13) results in the fusion of *Etv6* to a novel gene, *stl*, in a B-cell ALL cell line. Genes Chromosomes & Cancer **18**:254-268.

170. **Park, J.G. and E.P. Reddy**. 1992. Large-scale molecular mapping of human c-*myb* locus - c-*myb* protooncogene is not involved in 6q- abnormalities of lymphoid tumors. Oncogene **7**:1603-1609.

171. **Burke, D.T., G.F. Carle, and M.V. Olson**. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science **236**:806-812.

172. **Shizuya, H., B. Birren, U.J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, and M. Simon**. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *escherichia-coli* using an F-factor-based vector. Proceedings of the National Academy of Sciences of the USA **89**:8794-8797.

173. **Pierce, J.C., B. Sauer, and N. Sternberg**. 1992. A positive selection vector for cloning high-molecular-weight DNA by the bacteriophage-P1 system - improved cloning efficacy. Proceedings of the National Academy of Sciences of the USA **89**:2056-2060.

174. **Ioannou, P.A., C.T. Amemiya, J. Garnes, P.M. Kroisel, H. Shizuya, C. Chen, M.A. Batzer, and P.J. De Jong**. 1994. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nature Genetics **6**:84-89.

175. **Green, E.D., H.C. Riethman, J.E. Dutchik, and M.V. Olson**. 1991. Detection and characterization of chimeric yeast artificial- chromosome clones. Genomics **11**:658-669.

176. **Yokobata, K., B. Trenchak, and P.J. Dejong**. 1991. Rescue of unstable cosmids by *invitro* packaging. Nucleic Acids Research **19**:403-404.

177. **Olson, M., L. Hood, C. Cantor, and D. Botstein**. 1989. A common language for physical mapping of the human genome. Science **245** :1434-1435.

178. **Rose, E.A.** 1991. Applications of the polymerase chain-reaction to genome analysis. FASEB Journal **5**:46-54.

179. **Nelson, D.L., S.A. Ledbetter, L. Corbo, M.F. Victoria, R. Ramirezsolis, T.D. Webster, D.H. Ledbetter, and C.T. Caskey**. 1989. *Alu* polymerase chain-reaction - a method for rapid isolation of human-specific sequences from complex DNA sources. Proceedings of the National Academy of Sciences of the USA **86**:6686-6690.

180. **Chumakov, I.M., P. Rigault, I. Legall, C. Bellannechantelot, A. Billault, S. Guillou, P. *et al.*** 1995. A YAC contig map of the human genome. Nature **377**:175-175.

181. **Adams, M.D., J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olde, R.F. Moreno, A.R. Kerlavage, W.R. Mccombie, and J.C. Venter**. 1991. Complementary-DNA sequencing - expressed sequence tags and Human Genome Project . Science **252**:1651-1656.

182. **Badiani, P.A., D. Kioussis, D.M. Swirsky, I.A. Lampert, and K. Weston**. 1996. T-cell lymphomas in v-*myb* transgenic mice. Oncogene **13**:2205-2212.

183. **Taylor, D., P. Badiani, and K. Weston**. 1996. A dominant interfering *myb* mutant causes apoptosis in T-cells. Genes & Development **10**:2732-2744.

184. **Weston, K.** 1998. Myb proteins in life, death and differentiation. Current Opinion In Genetics & Development **8**:76-81.

185. **Brown, K.E., M.S. Kindy, and G.E. Sonenshein**. 1992. Expression of the c-*myb* proto-oncogene in bovine vascular smooth- muscle cells. Journal Of Biological Chemistry **267**:4625-4630.

186. **Wolff, L.** 1996. Myb-induced transformation. Critical Reviews In Oncogenesis **7**:245-260.

187. **Mucenski, M.L., K. McLain, A.B. Kier, S.H. Swerdlow, C.M. Schreiner, T.A. Miller, D.W. Pietryga, W.J.J. Scott, and S.S. Potter**. 1991. A functional c-*myb* gene is required for normal murine fetal hepatic hematopoiesis. Cell **65**:677-689.

188. **Clarke, M.F., J.F. Kukowskalatallo, E. Westin, M. Smith, and E.V. Prochownik**. 1988. Constitutive expression of a c-*myb* cDNA blocks friend murine erythroleukemia cell-differentiation. Molecular And Cellular Biology **8**:884-892.

189. **Ferrao, P., E.M. Macmillan, L.K. Ashman, and T.J. Gonda**. 1995. Enforced expression of full-length c-*myb* leads to density- dependent transformation of murine hematopoietic-cells. Oncogene **11**:1631-1638.

190. **Selvakumaran, M., D.A. Liebermann, and B. Hoffmanliebermann** . 1992. Deregulated c-*myb* disrupts interleukin-6-induced or leukemia inhibitory factor-induced myeloid differentiation prior to c-*myc* - role in leukemogenesis. Molecular And Cellular Biology **12**:2493-2500.

191. **Thompson, C.B., P.B. Challoner, P.E. Neiman, and M. Groudine**. 1986. Expression of the c-*myb* proto-oncogene during cellular proliferation. Nature **319**:374-380.

192. **Lyon, J., C. Robinson, and R. Watson**. 1994. The role of Myb proteins in normal and neoplastic cell- proliferation. Critical Reviews In Oncogenesis **5**:373-388.

193. **Marhamati, D.J., R.E. Bellas, M. Arsura, K.E. Kypreos, and G.E. Sonenshein**. 1997. A-*myb* is expressed in bovine vascular smooth muscle cells during the late G_1 to-S phase transition and cooperates with c-*myc* to mediate progression to S phase. Molecular And Cellular Biology **17**:2448-2457.

194. **Anfossi, G., A.M. Gewirtz, and B. Calabretta**. 1989. An oligomer complementary to c-*myb*-encoded messenger-RNA inhibits proliferation of human myeloid-leukemia cell-lines. Proceedings of the National Academy of Sciences of the USA **86**:3379-3383.

195. **Gewirtz, A.M., G. Anfossi, D. Venturelli, valpreda S, Sims R, and B. Calabretta**. 1989. G_1/S transition in normal human T-lymphocytes requires the nuclear protein encoded by c-*myb*. Science **245**:180-186.

196. **Lyon, J.J. and R.J. Watson**. 1996. Interference of *myb* transactivation activity by a conditional dominant negative protein: functional interference in a cytotoxic T- cell line results in G_1 arrest. Gene **182**:123-128.

197. **Jarvis, T.C., F.E. Wincott, L.J. Alby, J.A. McSwiggen, L. Beigelman, J. Gustofson, A. DiRenzo, K. Levy, M. Arthur, J. Matulic-Adamic, A. Karpeisky, C. Gonzalez, T.M. Woolf, N. Usman, and D.T. Stinchcomb**. 1996. Optimizing the cell efficacy of synthetic ribozymes. Site selection and chemical modifications of ribozymes targeting the proto-oncogene c-*myb*. J.Biol.Chem. **271**:29107-29112.

198. **Nakagoshi, H., C. KaneiIshii, T. Sawazaki, G. Mizuguchi, and S. Ishii**. 1992. Transcriptional activation of the c-*myc* gene by the c-*myb* and B-*myb* gene-products. Oncogene **7**:1233-1240.

199. **Evans, J.L., T.L. Moore, W.M. Kuehl, T. Bender, and J.P. Ting**. 1990. Functional-analysis of c-Myb protein in T-lymphocytic cell-lines shows that it trans-activates the c-*myc* promoter. Molecular And Cellular Biology **10**:5747-5752.

200. **Frampton, J., T. Ramqvist, and T. Graf**. 1996. v-*myb* of E26 leukemia-virus up-regulates *bcl*-2 and suppresses apoptosis in myeloid cells. Genes & Development **10**:2720-2731.

201. **Salomoni, P., D. Perrotti, R. Martinez, C. Franceschi, and B. Calabretta**. 1997. Resistance to apoptosis in CTLL-2 cells constitutively expressing c- *myb* is associated with induction of *bcl*-2 expression and *myb*-dependent regulation of *bcl*-2 promoter activity. Proceedings of the National Academy of Sciences of the USA. **94**:3296-3301.

202. **Hogg, A., S. Schirm, H. Nakagoshi, P. Bartley, S. Ishii, J.M. Bishop, and T.J. Gonda**. 1997. Inactivation of a c-*myb*/estrogen receptor fusion protein in transformed primary cells leads to granulocyte/macrophage differentiation and down regulation of c-*kit* but not c-*myc* or CDC2. Oncogene **15**:2885-2898.

203. **Sakura, H., K.I. Chie, T. Nagase, H. Nakagoshi, T.J. Gonda, and S. Ishii**. 1989. Delineation of 3 functional domains of the transcriptional activator encoded by the c-*myb* protooncogene. Proceedings of the National Academy of Sciences of the USA. **86**:5758-5762.

204. **Hu, Y.L., R.G. Ramsay, C. KaneiIshii, S. Ishii, and T.J. Gonda**. 1991. Transformation by carboxyl-deleted *myb* reflects increased transactivating capacity and disruption of a negative regulatory domain. Oncogene **6**:1549-1553.

205. **Kalkbrenner, F., S. Guehmann, and K. Moelling**. 1990. Transcriptional activation by human c-*myb* and v-*myb* genes. Oncogene **5**:657-661.

206. **Dubendorff, J.W., L.J. Whittaker, J.T. Eltman, and J.S. Lipsick**. 1992. Carboxy-terminal elements of c-*myb* negatively regulate transcriptional activation in *cis* and in *trans*. Genes & Development **6**:2524-2535.

207. **KaneiIshii, C., E.M. Macmillan, T. Nomura, A. Sarai, R.G. Ramsay, S. Aimoto, S. Ishii, and T.J. Gonda**. 1992. Transactivation and transformation by *myb* are negatively regulated by a leucine-zipper structure. Proceedings of the National Academy of Sciences of the USA. **89**:3088-3092.

208. **Favier, D. and T.J. Gonda**. 1994. Detection of proteins that bind to the leucine-zipper motif of c-Myb. Oncogene **9**:305-311.

209. **Nomura, T., N. Sakai, A. Sarai, T. Sudo, C. KaneiIshii, R.G. Ramsay, D. Favier, T.J. Gonda, and S. Ishii**. 1993. Negative autoregulation of c-*myb* activity by homodimer formation through the leucine-zipper. Journal Of Biological Chemistry **268**:21914-21923.

210. **Ishiguro, N., T. Ohzono, T. Shinagawa, M. Horiuchi, and M. Shinagawa**. 1994. A spontaneous internal deletion of the c-*myb* proto-oncogene enhances transcriptional activation in bovine T-lymphoma-cells. Journal Of Biological Chemistry **269**:26822-26829.

211. **Dasgupta, P. and E.P. Reddy**. 1989. Identification of alternatively spliced transcripts for human c-*myb* - molecular-cloning and sequence-analysis of human c-*myb* exon-9A sequences. Oncogene **4**:1419-1423.

212. **Woo, C.H., L. Sopchak, and J.S. Lipsick**. 1998. Overexpression of an alternatively spliced form of c-Myb results in increases in transactivation and transforms avian myelomonoblasts. Journal Of Virology: 6813-6821.

213. **Fu, S.L. and J.S. Lipsick**. 1996. FAETL motif required for leukemic transformation by v-Myb. Journal Of Virology:5600-5610.

214. **Dash, A.B., F.C. Orrico, and S.A. Ness**. 1996. The EVES motif mediates both intermolecular and intramolecular regulation of c-Myb. Genes Dev. **10**:1858-1869.

215. **Aziz, N., M.R. Miglarese, R.C. Hendrickson, J. Shabanowitz, T.W. Sturgill, D.F. Hunt, and T.P. Bender**. 1995. Modulation of c-Myb-induced transcription activation by a phosphorylation site near the negative regulatory domain. Proceedings of the National Academy of Sciences of the USA. **92**:6429-6433.

216. **Leverson, J.D., P.J. Koskinen, F.C. Orrico, E.M. Rainio, K.J. Jalkanen, A.B. Dash, R.N. Eisenman, and S.A. Ness**. 1998. Pim-1 kinase and p100 cooperate to enhance c-Myb activity. Mol.Cell **2**:417-425.

217. **Bender, T.P., C.B. Thompson, and W.M. Kuehl**. 1987. Differential expression of c-*myb* messenger-RNA in murine-B lymphomas by a block to transcription elongation. Science **237**:1473-1476.

218. **Watson, R.J.** 1988. A transcriptional arrest mechanism involved in controlling constitutive levels of mouse c-*myb* messenger-RNA. Oncogene **2**:267-272.

219. **Reddy CD and E.P. Reddy**. 1989. Differential binding of nuclear factors to the intron 1 sequences containing the transcriptional pause site correlates with c-*myb* expression. Proceedings of the National Academy of Sciences of the USA. **86**:7326-7330.

220. **Watson, R.J.** 1988. Expression of the c-*myb* and c-*myc* genes is regulated independently in differentiating mouse erythroleukemia-cells by common processes of premature transcription arrest and increased messenger-RNA turnover. Molecular And Cellular Biology **8**:3938-3942.

221. **Bies, J. and L. Wolff**. 1997. Oncogenic activation of c-*myb* by carboxyl-terminal truncation leads to decreased proteolysis by the ubiquitin-26S proteasome pathway. Oncogene **14**:203-212.

222. **Grasser, F.A., T. Graf, and J.S. Lipsick**. 1991. Protein truncation is required for the activation of the c-*myb* proto-oncogene. Molecular And Cellular Biology **11**:3987-3996.

223. **Nasonburchenal, K. and L. Wolff**. 1993. Activation of c-*myb* is an early bone-marrow event in a murine model for acute promonocytic leukemia. Proceedings of the National Academy of Sciences of the USA. **90**:1619-1623.

224. **Slamon, D.J., J.B. deKernion, I.M. Verma, and M.J. Cline**. 1984. Expression of cellular oncogenes in human malignancies. Science **224**:256-262.

225. **Slamon, D.J., T.C. Boone, D.C. Murdock, D.E. Keith, M.F. Press, R.A. Larson, and L.M. Souza**. 1986. Studies of the human c-*myb* gene and its product in human acute leukemias. Science **233**:347-351.

226. **Clarke, B.J., S.K. Liao, C. Leeds, P. Soamboonsrup, and P.B. Neame**. 1987. Distribution of a hematopoietic-specific differentiation antigen of K562 cells in the human myeloid and lymphoid cell lineages. Cancer Res. **47**:4254-4259.

227. **Baer, M.R., P. Augustinos, and A.J. Kinniburgh**. 1992. Defective c-*myc* and c-*myb* RNA turnover in acute myeloid-leukemia cells. Blood **79**:1319-1326.

228. **Ratajczak, M.Z., N. Hijiya, L. Catani, K. DeRiel, S.M. Luger, P. McGlave, and A.M. Gewirtz**. 1992. Acute- and chronic-phase chronic myelogenous leukemia colony-forming units are highly sensitive to the growth inhibitory effects of c-*myb* antisense oligodeoxynucleotides. Blood **79**:1956-1961.

229. **Kasono, K., Pich, J. Xiang, H.G. Kim, G. Bilbao, F. Johanning, M. Nawrath, K. Moelling, and D.T. Curiel**. 1998. Functional knock-out of c-*myb* by an intracellular anti-c-*Myb* single- chain antibody. Biochem.Biophys.Res.Commun. **251**:124-130.

230. **Tomita, A., T. Watanabe, H. Kosugi, H. Ohashi, T. Uchida, T. Kinoshita, S. Mizutani, T. Hotta, T. Murate, M. Seto, and H. Saito**. 1998. Truncated c-*Myb* expression in the human leukemia cell line TK-6. Leukemia **12**:1422-1429.

231. **Howe KM, Reakes CF, and R. Watson**. 1990. Characterization of the sequence-specific interaction of mouse c-Myb protein with DNA. EMBO Journal **9**:161-169.

232. **Boise, L.H., K.M. Gorse, and E.H. Westin**. 1992. Multiple mechanisms of regulation of the human c-myb gene during myelomonocytic differentiation. Oncogene **7**:1817-1825.

233. **Alterman, R.B.M., S. Ganguly, D.H. Schulze, W.F. Marzluff, C.L. Schildkraut, and A.I. Skoultchi**. 1984. Cell-cycle regulation of mouse H3 histone messenger-RNA metabolism. Molecular And Cellular Biology **4**:123-132.

234. **Bies, J., V. Nazarov, and L. Wolff**. 1999. Identification of protein instability determinants in the carboxy- terminal region of c-Myb removed as a result of retroviral integration in murine monocytic leukemias. Journal of Virology **73**:2038-2044.

235. **Shenong, G.L.C., B. Luscher, and R.N. Eisenman**. 1989. A 2nd c-Myb protein is translated from an alternatively spliced messenger-RNA expressed from normal and 5'-disrupted *myb* loci. Molecular And Cellular Biology **9**:5456-5463.

236. **Schuur, E.R., J.M. Rabinovich, and M.A. Baluda**. 1994. Distribution of alternatively spliced chicken c-*myb* exon 9A among hematopoietic tissues. Oncogene **9**:3363-3365.

237. **Schuur, E.R., P. Dasgupta, E.P. Reddy, J.M. Rabinovich, and M.A. Baluda**. 1993. Alternative splicing of the chicken c-*myb* exon 9A. Oncogene **8**:1839-1847.

238. **Rosson, D., D. Dugan, and E.P. Reddy**. 1987. Aberrant splicing events that are induced by proviral integration: implications for *myb* oncogene activation. Proceedings of the National Academy of Sciences of the USA. **84**:3171-3175.

239. **Kamano, H., B. Burk, K. NobenTrauth, and K.H. Klempnauer**. 1995. Differential splicing of the mouse B-myb gene. Oncogene **11**:2575-2582.

240. **Peters, C.W., A.E. Sippel, M. Vingron, and K.H. Klempnauer**. 1987. Drosophila and vertebrate myb proteins share two conserved regions, one of which functions as a DNA-binding domain. EMBO Journal. **6**:3085-3090.

241. **Tomita, A., T. Watanabe, H. Kosugi, H. Ohashi, T. Uchida, T. Kinoshita, S. Mizutani, T. Hotta, T. Murate, M. Seto, and H. Saito**. 1998. Truncated c-*Myb* expression in the human leukemia cell line TK-6. Leukemia **12**:1422-1429.

242. **Jacobs, S.M., K.M. Gorse, and E.H. Westin**. 1994. Identification of a 2nd promoter in the human c-*myb* proto-oncogene. Oncogene **9**:227-235.

243. **Westin, E.H., K.M. Gorse, and M.F. Clarke**. 1990. Alternative splicing of the human c-*myb* gene. Oncogene **5**:1117-1124.

244. **Reicin, A., J.Q. Yang, K.B. Marcu, E. Fleissner, C.F. Koehne, and P.V. O'Donnell**. 1986. Deregulation of the c-*myc* oncogene in virus-induced thymic lymphomas of AKR/J mice. Molecular and Cellular Biology **6**:4088-4092.

245. **Forrest, D., D. Onions, G. Lees, and J.C. Neil**. 1987. Altered structure and expression of c-*myc* in feline T-cell tumors. Virology **158**:194-205.