Identification of Genes Involved in Leukaemia and Differentiation Induced by Activated Mutants of the GM-CSF Receptor β subunit

Brenton James Reynolds

A thesis submitted for the degree of Doctor of Philosophy
to the University of Adelaide, School of Medicine
(Discipline of Medicine)
November, 2005

Table of Contents

Abstract

Interleukin (IL)-3, IL-5 and granulocyte-macrophage-colony-stimulating factor (GM-CSF) are cytokines that affect the growth, survival and differentiation of many cells within the haematopoietic system. The functions of these factors are mediated by membrane bound receptor complexes that are composed of specific ligand binding subunits (α) and a common signal transducing subunit (hβc). Constitutively activated mutants of hβc have been previously identified that are able to confer factor-independent signalling in a number of haematopoietic cell lines (including FDC-P1 and FDB-1). These activated mutants fall into two classes defined by the location of the mutation and their biochemical and leukaemogenic properties. In particular, the transmembrane mutant, V449E, causes an acute myeloid leukaemia *in vivo*, whereas the extracellular mutants (FIΔ or I374N) cause chronic myeloproliferative disorders.

The work described in this thesis used the activated hβc mutants to uncover novel transcriptional events induced by the GM-CSF/IL-3/IL-5 receptor complex and to define pathways associated with proliferation and differentiation. Large-scale gene expression profiling techniques were used to investigate the genes involved in these biological processes in the murine myelomonocytic cell line FDC-P1, and the bi-potent FDB-1 myeloid cell line, which are responsive to IL-3 and GM-CSF.

Membrane arrays were used to identify differences in gene expression between I374N and V449E expressing FDC-P1 cells. This technique revealed that the gene *Ptpmt1* was differentially expressed between V449E and I374N, which was subsequently confirmed by Northern blotting. This finding suggested that the phosphatase encoded by *Ptpmt1* may be involved in the different outcomes induced by these two hβc mutants. Northern analysis also revealed *Ptpmt1*, *Nab1* and *Ddx26b* to be regulated in response to human GM-CSF in FDC-P1 cells expressing human GM-CSFα and hβc.

A large-scale cDNA microarray experiment was also performed to identify genes that are selectively expressed during differentiation of FIΔ expressing FDB-1 cells, compared to proliferating V449E expressing FDB-1 cells over 24 hours. A comprehensive analysis approach was adopted to examine the microarray data and identify differentially expressed genes. Among the genes displaying differential

expression were *Btg1, S100a9, Cd24,* and *Ltf* found to be differentiation-associated and *Bnip3, Cd34, Myc, Nucleophosmin,* and *Nucleostemin* found to be proliferation-associated. *Hipk1, Klf6, Sp100,* and *Sfrs3* were also identified as potential transcriptional regulators during growth and differentiation. Northern analysis was used to confirm differences in expression for these 13 genes between FIΔ and V449E expressing FDB-1 cells. Eleven of the 13 genes examined were confirmed to be differentially expressed between FIΔ and V449E expressing FDB-1 cells over 24 hours. Furthermore, six genes (*Btg1, Hipk1, Cd24, Cd34, Klf6* and *Nucleostemin*) examined over 72 hours revealed differences in gene expression at early (6-12 hours) and late (48-72 hours) time points. Cell surface expression of CD24 protein was also shown to be induced upon FIΔ expression or GM-CSF induced differentiation of FDB-1 cells, consistent with elevated levels of *Cd24* mRNA in FIΔ cells over time. Based on their confirmed gene expression differences seen on the microarrays and Northern analysis, four genes (*Btg1, Cd24, Klf6* and *Nucleostemin*) were selected for over-expression analysis in FDC-P1 or FDB-1 cells, in order to gain insights into the function of these genes. Optimisation of the retroviral infection process was performed so that the role of these genes in proliferation and differentiation could be investigated in the FDB-1 model. Such preliminary functional experiments in FDB-1 cells will enable prioritisation of the genes for further analysis of their function in primary cells. Thus, the work in this thesis describes the first use of microarrays to identify gene expression differences between hβc mutants with differential activities affecting myeloid growth and differentiation.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed,

Brenton Reynolds

20 April 2007.

Acknowledgements

First of all I would like to thank my supervisors, Richard D'Andrea and Tom Gonda for providing me the opportunity to conduct research in their laboratories and for their advice and support throughout the entire course of my PhD. I thank the University of Adelaide for awarding me a three and half-year scholarship to further my undergraduate studies into a PhD, and for Richard for an extra 6 months of funding that allowed me time to finish my experiments and begin the writing of my thesis.

I would also like to thank my collaborators in this project including Associate Professor Patty Solomon and the many statisticians (including Gary Glonek, Patty's Honours student Julian Whiting and her PhD student Jonas Lloyd) who provided their statistical experience in the statistical analysis of membrane array and microarray experiments. Big thanks to Andrew Holloway for allowing me to perform the microarray experiments in his laboratory. Similarly, I must thank Anna Tyskin and Chris Wilkinson for their assistance with the data analysis of the microarray experiment, and Anna Brown for providing me the FDB-1 cell lines used in this study.

Many thanks to all members (past and present) of both laboratories for their support and technical assistance and I thank them for their patience in answering my seemingly never ending barrage of questions.

To my fellow PhD students at the Hanson Centre and CHRI I thank you for sharing the PhD experience with me and for everything we shared and learned. I wish you guys all the best with your studies and future careers.

Last but not least, a big thanks to my family, particularly my parents, for their love and support throughout my PhD and also to my friends for their friendship and support. I would particularly like to thank Camilla for her patience with me during the final writing stages of this thesis. My friends and family helped me remain focused and motivated at times when things were tough – for this and everything else, thank you!

Publications

Brown, A. L., Wilkinson, C. R., Waterman, S. R., Kok, C. H., Salerno, D. G., Diakiw, S. M., **Reynolds, B.**, Scott, H. S., Tsykin, A., Glonek, G. F., Goodall, G. J., Solomon, P. J., Gonda, T. J., and D'Andrea, R. J. (2006). Genetic regulators of myelopoiesis and leukemic signaling identified by gene profiling and linear modeling. J. Leukoc. Biol. 80(2): 433-447.

Conference Presentations

Oral Presentations

- Child Health Research Institute, Annual General Meeting – Young Investigator for Immunology.
  North Adelaide, November, 2002.

- 22nd Annual Lorne Genome Conference. “Identification of leukaemogenic genes involved in activated mutants of the GM-CSF receptor beta subunit”.
  Lorne, Victoria, February, 2001.

Poster Presentations

- The American Society of Hematology, 45th Annual Meeting.
  San Diego, California, USA, December, 2003.

- 2nd Australian Microarray Meeting.
  South Stradbroke Island, Queensland, July, 2002.

Abbreviations

The following abbreviations appear throughout the text of this thesis.

Ac	acetate
ALL	acute lymphoblastic leukaemia
ALV	avian leukosis virus
AMegL	Acute megakaryocytic leukaemia
AML	acute myeloid leukaemia
AMV	avian myeloblastosis virus
ANOVA	analysis of variation
(α-^{32}P)dATP	alpha-labelled 2'-deoxyadenosine-5'-triphophate
(α-^{33}P)dCTP	alpha-labelled 2'-deoxycytosine-5'-triphophate
APL	acute promyelocytic leukaemia
ATRA	all-trans-retinoic acid
B-CLL	B-cell chronic lymphocytic leukaemia
B-stat	B-statistic
βc	common β subunit
bp	base pair(s)
BSA	bovine serum albumin
cAMP	adenosine 3', 5'-cyclic monophosphate
Cdk	Cyclin-dependent kinases
cDMEM	complete Dulbecco's modified Eagle's medium
cDNA	deoxyribonucleic acid complementary to ribonucleic acid
CFU-GM	colony-forming unit granulocyte-macrophage
cIMDM	complete Iscove's modified Dulbecco's medium
CIP	calf intestinal phosphatase
CIS	common insertion site
CLL	chronic lymphoblastic leukaemia
CML	chronic myeloid leukaemia
CMML	chronic myelomonocytic leukaemia
CMP	common myeloid progenitor
CNTF	ciliary neurotrophic factor
cpm	counts per minute

cps	counts per second
CRD	cytokine receptor domain
CRM	cytokine receptor module
Cyc A	Cyclophilin A
DEPC	diethyl pyrocarbonate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide-triphosphate
DTT	dithiothreitol
EC	extracellular
E.coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
Epo	erythropoietin
EST	expressed sequence tag
F-stat	F-statistic
FACS	fluorescence-activated cell sorting
FCS	bovine foetal calf serum
FDR	false discovery rate
FITC	fluoroscein isothiocyanate
FLT3	Fms-like kinase 3
FPR	formyl peptide receptor
G-CSF	granulocyte colony-stimulating factor
G-CSFR	G-CSF receptor
GAP	GTPase-activating protein
GH	growth hormone
GM	GM-CSF e.g. hGMRα
GM-CSF	granulocyte-macrophage colony-stimulating factor
GMR	GM-CSF receptor
GMRα	GM-CSF receptor alpha subunit
GO	gene ontology
h	human
H_0	null hypothesis

Ha	alternate hypothesis
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
HMG	high mobility group
IAP	intracisternal-A particles
IL-	interleukin
IL-3Rα	IL-3 receptor alpha subunit
IMDM	Iscove's modified Dulbecco's medium
ITD	internal tandem duplications
JAK	Janus kinase
JCML	juvenile chronic myelogenous leukaemia
JMML	juvenile myelomonocytic leukemia
kb	kilobase(s)
kDa	kilo Dalton
LB	Luria-Bertani
LIF	leukaemia inhibitor factor
Lowess	locally weighted linear regression
m	murine
M	mole(s) per litre
M-CSF	macrophage colony-stimulating factor
MAP kinase	mitogen-activated protein kinase
MAPKK	MAP kinase kinase
mL	millilitre(s)
MDS	myelodysplastic syndrome
MPD	myeloproliferative disorder
MPLV	myeloproliferative leukaemia retrovirus
MOPS	3-(N-morpholino) propanesulphonic acid
MQ	Milli-Q
MuLV	murine leukaemia virus
MW	molecular weight
NF1	neurofibromatosis type 1
NIA	National Institute of Aging
NPM	Nucleophosmin
NS	Nucleostemin

NSCLC	non-small cell lung cancers
NUMA1	nuclear mitotic apparatus protein 1
OSM	oncostatin
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PE	phycoerythrin
Pen/Strep	Penicillin G (5000 U/mL) and Streptomycin Sulphate (5000 mcg/mL)
PI	phosphatidylinositol
PI3K	phosphatidylinositol 3-kinase
PKA	protein kinase A
PLZF	promyelocytic leukaemia zinc finger
PMCI	Peter MacCallum Cancer Institute
PML	promyelocytic leukemia
PMT	photo-multiplier tube
POD	PML oncogenic domains
PV	polycythemia vera
R	receptor (e.g. IL-3R)
RA	retinoic acid
RAR	retinoic acid receptor
RARE	retinoic acid response elements
Rb	retinoblastoma
rpm	revolutions per minute
RSMC	rat aortic smooth-muscle cell
RT	reverse transcription
RTK	receptor tyrosine kinase
RT-PCR	reverse transcription-polymerase chain reaction
s	standard deviation
s^2	variance
SAGE	serial analysis of gene expresion
SCF	stem cell factor
SDS	sodium dodecyl sulphate
SE	standard error
SFFV	Friend spleen focus-forming virus

SFK	Src family kinase
siRNA	small interfering RNA
SMA	Statistics for Microarray Analysis
SSC	saline sodium citrate
STAT	signal transducer and activator of transcription
t-stat	t test statistic
TAE	Tris-Acetic acid-EDTA
Taq	*Thermus aquaticus*
TE	Tris-EDTA
TEMED	N, N, N', N'-tetramethylethane-1,2-diamine
TGF	transforming growth factor
TM	transmembrane
Tpo	thrombopoietin
Tris	tris[hydroxymethyl]aminomethane (2-amino-2-(hydroxmethyl)propane-1,3-diol)
Tris-HCL	tris(hydroxymethyl)aminomethane hydrochloride
TGFβ	transforming growth factor-β
U	units
µL	microlitre(s)
uPA	urokinase type plasminogen activator
UV	ultraviolet
V	volts
% (v/v)	percentage volume per volume
% (w/v)	percentage weight per volume

Chapter 1

Introduction

1.1 Haematopoiesis

Haematopoiesis is the process by which blood cells are made. Blood cells are derived from multipotent stem cells that are found in the bone marrow in the adult. During foetal development, haematopoiesis begins in the blood islands of the yolk sac and then develops in the liver, spleen and lymph nodes before the bone marrow becomes the primary site for haematopoiesis throughout life (Gallicchio, 1998). Multipotent stem cells can proliferate in response to appropriate stimuli to produce large numbers of progenitor cells that become committed to one or another of the major cell lineages of blood. In the steady state the numbers of progenitor cells is small, but this compartment can be rapidly expanded under conditions of increased demand, such as infection or haemorrhage, or during growth of the individual. Under these conditions, alternative sites of haematopoiesis may be utilised, including the liver and spleen. This process is referred to as extramedullary haematopoiesis (Gallicchio, 1998).

Haematopoietic stem cells are referred to as multipotent because they have the ability to differentiate into mature cells of both the myeloid and lymphoid lineages, and are also capable of self-renewal (Spangrude *et al.*, 1988; Spangrude *et al.*, 1991). Importantly, haematopoietic stem cells are capable of engrafting into a myeloablated adult and reconstituting all lineages of the haematopoietic system (Spangrude *et al.*, 1988; Moore, 2005). While stem cells are multipotent, committed progenitor cells differentiate into a limited number of different cell lineages. As progenitor cells undergo further commitment, their proliferative and lineage potential is reduced until they finally form one of the mature blood cell types of the body (Figure 1).

The entire process of haematopoiesis is regulated in part by soluble mediators called cytokines, or growth factors (Watowich *et al.*, 1996; Zhu and Emerson, 2002). Both normal and leukaemic progenitor cells require growth factors for their survival, proliferation, maturation, differentiation and function (Minden, 1995; Watowich *et al.*, 1996). Many growth factors act on multiple cell lineages and frequently act in synergy with other growth factors (Watowich *et al.*, 1996). Growth factors mediate their effects by interacting with specific cell surface receptors. This interaction leads to the

activation of intracellular pathways resulting in changes in cell proliferation, survival and differentiation (Nicola, 1994; Gallicchio, 1998).

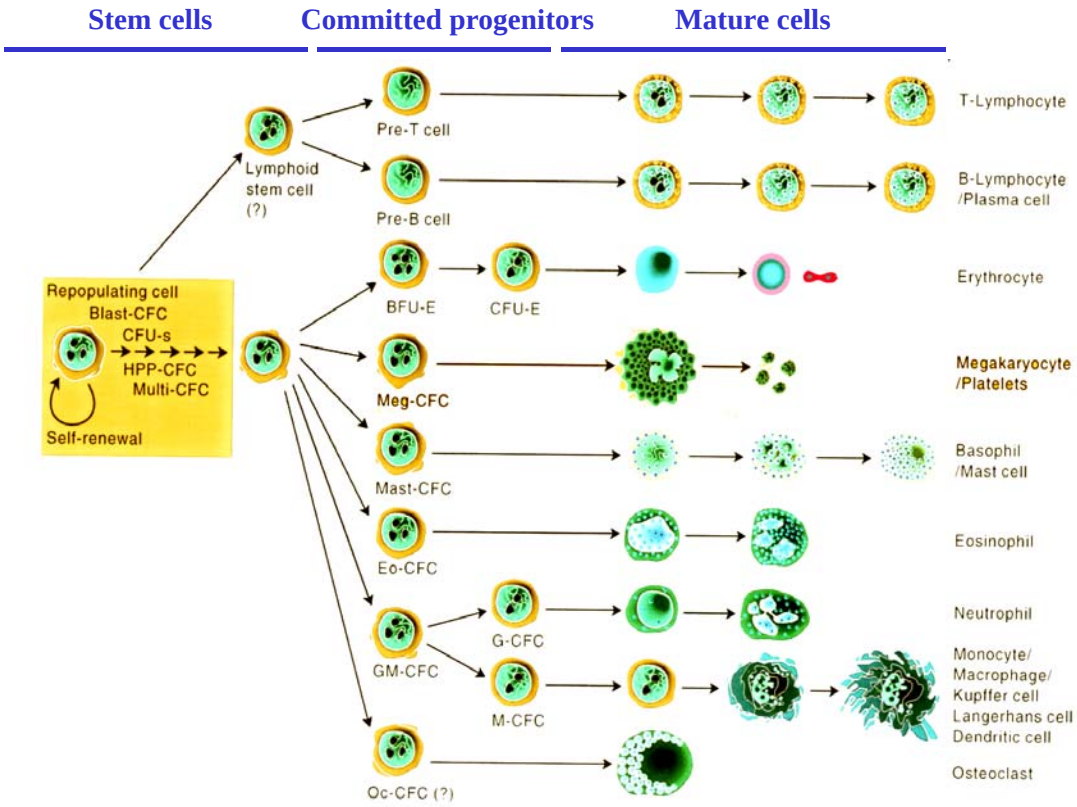


Figure 1: Outline of haematopoiesis. The haematopoietic system involves stem cells, progenitor cells and mature cells. Taken from T. Gainsford's PhD thesis, (1999) with permission.

Steady state haematopoiesis is a balance between proliferation and terminal differentiation/death of committed progenitor cells. There are numerous disorders that are the result of a breakdown of the delicate balance that exists between proliferation and differentiation of haematopoietic progenitor cells, including leukaemia and other myeloproliferative disorders (MPD). Leukaemia is the uncontrolled proliferation of white blood cells. It is usually due to genetic abnormalities that have occurred in the biochemical pathways that regulate cell proliferation and differentiation. The genetic abnormalities that can cause leukaemia involve targets including growth factor receptors, transcription factors and other regulatory molecules. Activation of oncogenes or loss of function of tumour suppressor genes, disturbance of DNA repair genes, and alteration of cell cycle genes can result in uncontrolled proliferation. (Kuby, 1997a; Brendel and Neubauer, 2000).

The work in this thesis was undertaken to increase the knowledge of the cellular pathways that control myeloid cell proliferation, survival and differentiation. It therefore concentrates on those receptors that are most closely associated with this arm of haematopoiesis. These are the receptors of the receptor tyrosine kinase and cytokine receptor families. The cytokine receptor family, which is responsible for mediating intracellular signals generated by the extracellular binding of cytokines or growth factors, will be discussed further in Section 1.3. This thesis focuses specifically on the shared subunit (hβc) of the receptors for human interleukin (IL)-3, IL-5 and granulocyte-macrophage colony-stimulating factor (GM-CSF), a member of the cytokine receptor family.

1.2 Leukaemias

Leukaemia involves an uncontrolled proliferation of progenitor cells in the absence of differentiation. Uncontrolled proliferation can arise through changes that lead to an increased proliferative capacity of the progenitor cell and/or a block in differentiation. A block in differentiation results in an accumulation of immature cells that fail to fully mature and subsequently die. Proliferating immature cells that are blocked in their differentiation and escape immune surveillance, eventually dominate the bone marrow and invade other tissues and organs leading to death (Kuby, 1997a; Evans, 1998).

While the processes of proliferation and differentiation are central to the development of leukaemia, apoptosis is also involved in the control of leukaemia. Homeostasis is maintained by balancing cell proliferation with cell death and an imbalance in either may result in cancer (Kuby, 1997a). The regulation of proliferation occurs via proteins such as growth factors acting through growth factor receptors, which induce proliferation, and tumour suppressor genes, which inhibit proliferation. The regulation of apoptosis can occur by mutation of genes that either block or induce apoptosis. For example, in B-cell follicular lymphoma a chromosomal translocation of the anti-apoptotic gene *bcl-2* occurs leading to its over-expression and inhibition of apoptosis (Golub, 1999). Inhibition of apoptosis enables a leukaemic cell to escape its usual programmed cell death and can enable the cell to live almost indefinitely. However, other events that enable the cell to proliferate and maintain its self renewal capacity must also occur for immortalisation of a cell. As stated above, a block in differentiation can also lead to leukaemia. In myeloid leukaemias, chromosomal translocations occurring in genes involved in the regulation of transcription are a common mechanism leading to arrested differentiation (Warner *et al.*, 2004). The role of transcription factors in leukaemia will be discussed in more detail below (see Section 1.2.2). Warner *et al*., (2004) suggests that "alterations in transcriptional regulation leading to aberrant myeloid differentiation may be a necessary step in leukaemogenesis," (p.7169) but are not sufficient for progression to acute myeloid leukaemia (AML) without additional events providing proliferative and survival advantages.

The concept that multiple steps are involved in the process of leukaemogenesis is supported by several lines of evidence (Hanahan and Weinberg, 2000; Gilliland *et al.*, 2004; Warner *et al.*, 2004). This concept involves a clonal evolution process, such that no single event fully transforms a cell but that additional genetic alterations accumulate and synergize in promoting proliferation or inhibiting apoptosis (Loeb, 1998; Warner *et al.*, 2004). It has been demonstrated in experimental systems that neither suppression of differentiation, nor increased proliferation on their own can lead directly to leukaemia (Testa *et al.*, 2004; Moore, 2005). However, leukaemic murine cell lines can be generated by the combined actions of an oncogene that suppresses differentiation (usually transcription factors such as *Myb*, *Myc*, *Ski*, *Rel* and *Ets*) together with a growth factor or activated tyrosine kinase (Gonda *et al.*, 1989). These findings demonstrate that the development of leukaemia, as with other cancers, is a multi-step process.

In human leukaemia, gross chromosomal changes occur quite frequently resulting in the fusion of unrelated genes creating chimeric proteins with altered function(s) (Brendel and Neubauer, 2000) (see Table 1). In many cases, the disruption of each gene's normal function by the formation of the novel chimeric protein has been demonstrated to be responsible for the disease (Pandolfi, 2001). The mutations that cause leukaemia impact on genes that fall into two broad groups. The first group of genes are those that provide a proliferative or survival advantage on the haematopoietic cells by the aberrant activation of signal transduction pathways (Gilliland *et al.*, 2004). The second group are genes that inhibit differentiation or provide self-renewal properties to the haematopoietic cells during development through the mutation of transcription factors or transcription co-activators that mediate their effects on cell differentiation or self-renewal (Gilliland *et al.*, 2004). Examples of genes from both of these groups will be discussed below.

NOTE: This table is included on page 5 of the print copy of the thesis held in the University of Adelaide Library.

Table 1: Human leukaemias showing a selection of chromosomal translocations involved in this family of diseases. Examples of genes involved in acute lymphoblastic leukaemia (ALL), acute myeloid leukaemia (AML) and chronic myelogeneous leukaemia (CML) are given along with the chromosomal regions and translocations involved in these diseases.

Adapted from Golub, (1999).

1.2.1 Animal models of leukaemia – retrovirus mediated leukaemia

Animal models provide a means to examine multiple features of human disease including; the development and testing of new approaches to disease prevention and treatment, identification of early diagnostic markers and novel therapeutic agents, and an understanding of the *in vivo* biology and genetics of tumour initiation, promotion, progression and metastasis (Wolff, 1997; Hann and Balmain, 2001).

Mouse models of cancer are useful because mice are a genetically tractable organism that shares an immense degree of genetic similarity to humans (Hann and Balmain, 2001). While rarely occurring naturally in humans, a number of models of retrovirally induced myeloid leukaemia exist in mice and chickens, which are described in the following paragraph. The discovery of retroviral systems that are capable of causing leukaemia provide models that can be used to identify and then investigate the role of specific genes and their oncogenic fusion proteins in disease.

The identification of oncogenes and tumour suppressor genes involved in leukaemia has come from animal studies of retrovirally induced leukaemia. Acutely transforming retroviruses such as avian myeloblastosis virus (AMV) are generally defective due to the incorporation of transforming sequences from the host at the expense of viral sequences, leading to a loss of replication competency (Wolff, 1997). The identification of the sequences responsible for the transforming potential of the acutely transforming retroviruses identifies novel oncogenes. AMV causes acute myeloblastic leukaemia in chickens and contains the oncogene *v-myb* derived from the proto-oncogene *c-myb* via recombination with the avian leukosis virus (ALV) (Klempnauer *et al.*, 1982; Wolff, 1997; Kanei-Ishii *et al.*, 2004). The *v-myb* oncogene is capable of transforming haematopoietic cells, thereby revealing *c-myb* as a proto-oncogene (Kanei-Ishii *et al.*, 2004). Replication competent viruses, such as murine leukaemia viruses (MuLVs), induce disease via insertional mutagenesis without needing to carry cellular proto-oncogenes (Wolff, 1997). In monocytic leukaemia, (induced in mice by Moloney-MuLV, Amphotropic virus 4070A or Friend-MuLV) the proto-oncogene most frequently targeted was *c-myb* (Wolff, 1997). The random insertion of the virus into host DNA allows the investigation of genes important to the transformation process and has been responsible for the identification of several cellular proto-oncogenes that are summarised in Table 2.

Gene product involved	Locus	Virus
Transcription factor		
c-Myb	*c-Myb*	Moloney-MuLV Amphotropic 4070A Friend-MuLV
Evi1	*Evi-1*	Endogenous ecotropic virus Moloney-MuLV Cas-Br-E Friend-MuLV
Evi1	*Cb-1/fim-3*	MuLV
Meis1	*Meis1*	Endogenous ecotropic virus
Hoxa7	*Evi-6*	Endogenous ecotropic virus
Hoxa9	*Evi-7*	Endogenous ecotropic virus
Hoxb8	*Hoxb8(Hox2.4)*	IAP
Fli1	*Fli1*	Cas-Br-E
Scl (Tal1)	*Scl/Tal1*	IAP
Signal transduction		
Nf1	*Evi-2*	Endogenous ecotropic virus
Cyclin		
Cyl-1 (*Ccnd1*)	*Fis-1*	Friend-MuLV
Growth factor		
M-CSF	M-CSF	Endogenous ecotropic virus
GM-CSF	GM-CSF	IAP
IL-3	IL-3	IAP
Growth factor receptor		
c-fms/Csf1r	*Fim-1*	Friend-MuLV

Table 2: Oncogenes found to be targets of retroviral insertion in murine myeloid leukaemias. A list of transcription factors, signal transduction molecules, growth factors and growth factor receptors identified via insertional mutagenesis in murine myeloid leukaemias. Adapted from Wolff (1997). MuLV = murine leukaemia virus, IAP = intracisternal-A particles, and Cas-Br-E is a wild mouse ecotropic virus.

Retroviral gene tagging uses viral specific sequences to determine the integration site of a retrovirus within the host's genomic DNA (Neil and Cameron, 2002). When multiple insertion events have occurred in the same genomic region in independently derived tumours – it is called a common integration site (CIS). These CISs suggest the involvement of that region and nearby genes in the oncogenic process

(Neil and Cameron, 2002; Akagi *et al.*, 2004) The identification of novel disease genes by retroviral tagging has been advanced through the development of high-throughput polymerase chain reaction (PCR)-based methods for cloning retroviral integration sites from tumours, and the publication of sequence of the mouse genome (Mikkers and Berns, 2003; Akagi *et al.*, 2004). The use of retroviral tagging has enabled Suzuki *et al*., (2002) to identify genes involved in myeloid leukaemia, and B- and T-cell lymphoma. Suzuki *et al*., (2002) identified candidate genes for 149 CISs and these genes encoded proteins that matched classes of proteins previously associated with cancer. These classes of proteins included; transcription factors or proteins that regulate transcription, proteins that remodel chromatin, kinases and phosphatases, cytokines and their receptors, Ras proteins and their regulators, signalling molecules, proteins involved in apoptosis, actin binding proteins, splicing factors and proteins that control the cell cycle. Several high-throughput insertional mutagenesis screens have recently been published, leading to the identification of 236 CISs (Hwang *et al.*, 2002; Suzuki *et al.*, 2002; Mikkers and Berns, 2003). This, along with the fact that many more candidate genes are likely to be identified in the future, has promoted Akagi *et al*., (2004) to develop the Retroviral Tagged Cancer Gene database (RTCGD) to accumulate insertional mutagenesis data and allow researchers to identify integrations of interest and compare results across multiple tumour models.

The retroviral systems described above led to the identification of various classes of genes found to be involved in leukaemogenesis. These retroviral systems provide models for the diverse types of leukaemia that exist in humans and can be used to investigate the role of specific genes and oncogenic fusion proteins in disease initiation and progression.

1.2.2 Human myeloid leukaemias

There are a number of genes that are involved in both animal models of leukaemia (described above and as shown in Table 2) and human leukaemias (see Table 3). Identification of these genes reveals conserved pathways and mechanisms that lead to leukaemia in humans. Human myeloid leukaemias often involve chromosomal translocations that result in gene activation through the formation of chimeric transcription factors with novel properties (Look, 1997; Wolff, 1997; Rowley, 1998; Golub, 1999) (see examples in Table 1 and Table 3). The involvement of these chimeric

transcription factors supports the notion that AML occurs through transcription dysregulation (Golub, 1999). Over-expression of transcription factors due to retroviral insertion is commonly responsible for causing disease in animal models (as discussed above), but this mechanism plays a small role in human disease (Wolff, 1997). The genes encoding the transcription factors EVI1, HOXA9 and a number of ETS-related genes have been found to be involved in both human and animal diseases, which is consistent with transcription factors being an important class of proteins capable of altering cellular properties (Wolff, 1997; Moore, 2005) (see examples in Table 2 and Table 3). Studies in animals and humans have demonstrated that disruption of transcription factor function can alter differentiation and lead to leukaemia (Tenen, 2003; Rosenbauer *et al.*, 2005). Key transcription factors implicated in the disruption of differentiation include GATA binding protein 1 (GATA1), the *ets* family member PU.1, and CCAAT/enhancer binding protein α (CEBPα), all of which have critical roles in controlling haematopoiesis (Orkin, 2000). Other abnormal expression or function of growth factors, their receptors or downstream signalling molecules can also lead to haematopoietic disorders (examples are discussed in Sections 1.4 and 1.5) (Lang *et al.*, 1987; Gonda and D'Andrea, 1997; Goyal and Longmore, 1999; Ward *et al.*, 2000). This section will focus on examples of human myeloid diseases including chronic myelogeneous leukaemia (CML) and acute myeloid (or myelogenous) leukaemia (AML) and the genes involved. Other leukaemia-associated genes such as *Scl* (also known as. *Tal1*) will not be discussed as they are most relevant to leukaemias of the lymphoid lineage e.g. *Scl* in acute T-cell lymphoblastic leukaemia (ALL).

Gene function	Affected genes	Chromosomal region / abnormality	Disease
Signal transducer	BCR / ABL	t(9;22)(q34;q11)	CML
	TEL / PDGFRβ	t(5;12)(q33;p13)	MDS (JMML)
	TEL / ABL	t(9;12;14)	undifferentiated leukaemia
Transcription factor	PML / RARα	t(15;17)	AML M3 (APL)
	PLZF /RARα	t(11;17)	AML M3 (APL)
	NPM / RARα	t(5;17)	AML M3 (APL)
	AML1(CBFα) / ETO	t(8;21)(q22;q22)	AML M2
	AML1 / EAP / MDS1 / EVI1	t(3;21)(q26.2;q22)	CML, MDS
	CBFβ-MYH11	inv(16)(p13;q22)	AML M4Eo
		t(16;16)(p13;q22)	AML M4Eo
	EVI1	t(3;3) and inv(3q26)	AML M0, M4
		inv(3)(q21;q26)	AML, CML, MDS
		ins(3)(q21,25;27)	AML
		t(2;3)(q21;q27)	AML
	MLL	11q23	AML M4, M5
	MLL / AF9(MLLT3)	t(9;11)(p21-22;q23)	AML M1, M5
	MLL / ENL(LTG19)	t(11;19)(q23;p13.3)	AML, mixed acute leukaemia
	MLL / ENL(MEN)	t(11;19)(q23;p13.1)	AML M4, M5
	MLL / AF4(FEL)	t(4;11)(q21;q24)	Mixed acute leukaemia
	MLL / AF6(MLLT4)	t(6;11)(q27;q23)	AML M1, M4, M5
	DEK / NUP214	t(6;9)(p23;q34)	AML M1, M2, M4
	NUP98 / HOXA9	t(7;11)(p15;p15)	AML M1, M2
	TLS(FUS) / ERG	t(16;21)(p11;q22)	AML, CML (blast crisis), MDS
	MN1 / TEL	t(12;22)(p13;q11)	AML, MDS
Transcriptional co-activator	MOZ / CBP	t(8;16)(p23;p13)	AML M5
	MOZ / TIF2	inv(8)(p11;q13)	AML
Unknown		t(1;22)	AML M7

Table 3: Chromosomal translocations occurring in human myeloid leukaemias. This table states the function of genes involved in a number of myeloid diseases along with their recurrent chromosomal abnormalities (translocations and rearrangements). Where known, the particular AML FAB subtypes (M0-M7) is stated. Genes designated in brackets "( )" are more recent names, with the exception that AML1 (known as CBFα) is also referred to as RUNX1, which is not stated in the above table. Abbreviations - acute myeloid leukaemia (AML), acute promyelocytic leukaemia (APL), chronic myelogeneous leukaemia (CML), juvenile myelomonocytic leukemia (JMML) and myelodysplastic syndrome (MDS). Adapted from Wolff (1997), Lowenberg *et al*., (1999) and Golub (1999).

1.2.2.1 Chronic myelogeneous leukaemia (CML)

Disease progression in CML can be divided into three phases; a chronic phase involving a myeloproliferative disease, an accelerated phase, and a terminal blast crisis phase where a block in differentiation causes an acute leukaemia (Chu *et al.*, 2005; Randolph, 2005a; Randolph, 2005b). The clinical course of CML involves "the chronic phase lasting approximately four years, the accelerated phase between six to eighteen months and the blast crisis phase terminating in death in less than eight months" (Randolph, 2005a, p. 42). Around 65-75% of patients that enter blast crisis develop AML while 25%-35% develop ALL (Randolph, 2005a). In the initial stage, there is an expansion of multipotent stem cells leading to hypercellular bone marrow (elevated myeloid/erythroid ratio and a shift in the myeloid lineage to immature forms), splenomegally, and a high white blood cell count with increases in the number of granulocytes, eosinophils, basophils and lymphocytes (Wolff, 1997). 90 to 95% of the cases of CML are due to the translocation known as the Philadelphia chromosome [t(9;22)(q34;q11)], which generates a fusion protein involving the genes *c-ABL(ABL1)* and *BCR* (Heisterkamp *et al.*, 1983; Wolff, 1997; Randolph, 2005a). This fusion protein called BCR-ABL has aberrant tyrosine kinase activity and is thought to activate many of the same pathways as haematopoietic growth factors (Sawyers, 1997). Recently it was shown that a specific chemical inhibitor (STI-571 / imatinib mesylate / Gleevec^{TM}) could successfully halt the course of this disease if given in chronic phase (Druker *et al.*, 2001) and now is successfully used as a therapy (Saglio *et al.*, 2004; Ren, 2005; Usui, 2005). Imatinib mesylate acts by binding to the constitutively active BCR-ABL kinase and inhibiting its kinase activity (Randolph, 2005b).

Constitutive activation of tyrosine kinases (such as c-ABL in the fusion protein BCR-ABL) is a common mechanism of transformation in myeloproliferative disorders, including CML (Golub, 1999). The tyrosine kinase activity of the BCR-ABL fusion protein leads to increased phosphorylation of adapter proteins, which stimulate several signalling pathways including Ras and JAK/STAT (Janus kinase/signal transducers and activators of transcription) pathways (Randolph, 2005a). Ras and JAK/STAT pathways are discussed in Section 1.3.2. The effect of the aberrant activation of these signalling pathways is to transform myeloid cells via reduction in apoptosis, increased proliferation, and incomplete differentiation (Randolph, 2005a).

1.2.2.2 Acute myeloid leukaemia (AML)

AML is a clonal disorder of poorly differentiated haematopoietic progenitor cells (Evans, 1998; Brendel and Neubauer, 2000). It is a heterogeneous disease caused by a variety of mechanisms, and is characterised by an increase in the number of myeloid cells in the bone marrow that are arrested in their maturation (Lowenberg *et al.*, 1999). AML can be divided into 9 distinct subgroups (M0-M7) based on a classification method developed by the French-American-British (FAB) Cooperative Group, primarily using morphologic appearance of the blasts (Bennett *et al.*, 1985; Cheson *et al.*, 1990) (see Table 4).

NOTE: This table is included on page 12 of the print copy of the thesis held in the University of Adelaide Library.

Table 4: The French-American-British (FAB) classification system for dividing AML into 9 distinct subgroups. The FAB classification, which is based on the morphologic appearance of the blasts, is stated in the first column along with its common name and the percentage of AML cases for each FAB subtype.

Adapted from Lowenberg *et al.*, (1999).

As stated above, a common mechanism of transformation in CML is the constitutive activation of tyrosine kinases (for example, c-ABL) (Golub, 1999). This is in contrast to AML where the transformation event is generally a balanced chromosomal translocation that results in the formation of chimeric proteins commonly involving transcription factors (Golub, 1999). This is highlighted by the number of transcription factors listed in Table 3 that have been found to be involved in AML.

A particularly well characterised subtype of AML called acute promyelocytic leukaemia (APL) (AML M3) will be discussed as it highlights a number of key points in the pathogenesis and treatment of leukaemia, namely the occurrence of chromosomal translocation and the use of differentiation inducing agents. It is the only FAB subtype shown to be molecularly distinct from the others (Golub, 1999). APL is characterised by a clonal expansion of tumour cells that have a promyelocytic morphological characteristic (Pandolfi, 2001). The majority of cases of APL involve a chromosomal translocation of chromosome 17 and in almost 99% of cases the patients have a reciprocal and balanced translocation between chromosomes 15 and 17 [t(15;17)(22;q11.2-12)](Pandolfi, 2001). This translocation results in the retinoic acid receptor α (*RARα*) gene on chromosome 17 being fused to the promyelocytic leukemia (*PML*) gene from chromosome 15. Other less common fusion partners, include the promyelocytic leukaemia zinc finger (*PLZF*) gene on chromosome 11, and other even rarer events include the genes nucleophosmin (*NPM*) (chromosome 5), nuclear mitotic apparatus protein 1 (*NUMA1)* (chromosome 11) and signal transducer and activator of transcription 5B *(STAT5B)* (chromosome 17) (Lin *et al.*, 1999; Pandolfi, 2001).

The PML-RARα fusion protein has multiple oncogenic functions through its ability to block both the PML and RARα pathways (Pandolfi, 2001). Due to the reciprocal nature of the translocation, the result is two distinct fusion proteins of the form PML-RARα and RARα-PML. These two fusion proteins are capable of interfering with both the PML and RARα pathways, and have a broad range of biological effects. These include effects on cellular differentiation, proliferation and survival (Pandolfi, 2001). RARα is a transcription factor and like other retinoic acid receptors (RARs) is a member of the nuclear hormone receptor superfamily, incorporating the receptors for steroid hormone, thyroid hormone, vitamin D_3 and the orphan receptors (Chambon, 1996; Dilworth *et al.*, 1999). PML has been shown to act *in vivo* as a tumour suppressor through control of apoptosis and cellular differentiation (Rego *et al.*, 2001).

Treatment of APL with retinoic acid (RA) induces disease remission by *in vivo* differentiation of APL blast cells, however, the effects are transient and patients frequently relapse if treatment is not supported with chemotherapy (He *et al.*, 1999; Melnick and Licht, 1999; Pandolfi, 2001). It is somewhat surprising that RA has this effect given that its receptor (RARα) is involved in the APL translocation. “Both PML-

RARα and PLZF-RARα interact with the nuclear receptor co-repressors SMRT or N-CoR and recruit the histone deacetylase complex" (Lin *et al.*, 1999, p.180). Inhibitors of histone deacetylases have been demonstrated to act as potent anti-tumour agents that can induce apoptosis of leukaemic blasts (Insinga *et al.*, 2005). Histone deacetylase inhibitors in combination with RA have the ability to overcome the repressor activity of these two fusion proteins and enhance the terminal differentiation of APL leukaemic cells (Grignani *et al.*, 1998; Guidez *et al.*, 1998; He *et al.*, 1998; Lin *et al.*, 1998). This finding demonstrates the importance of histone deacetylase recruitment in APL pathogenesis and suggests that the transforming potential of these fusion proteins lies not with their activation function but with their repressor functions (Lin *et al.*, 1999). Thus the treatment of APL with RA has become a paradigm for cancer treatment utilizing differentiation agents.

1.3 Cytokines and their receptors

Cytokines are regulatory peptides, which are produced by various cells in the body. They have effects on haematopoietic cells and many other cell types that regulate the immune system and promote repair processes (Thomson, 1998). Cytokines are distinct from hormones in that they usually act locally in an autocrine or paracrine manner rather than systemically via the circulation (Kuby, 1997b). The exceptions to this include macrophage colony stimulating factor (M-CSF), stem cell factor, erythropoietin (Epo) and transforming growth factor-β (TGF-β), which are all found in the circulatory system (Thomson, 1998). Cytokines mediate their effects by interacting with specific cell surface receptors – the cytokine receptor superfamily (Watowich *et al.*, 1996; Kuby, 1997b). Binding of cytokines to their receptors activates downstream intracellular signalling pathways that regulate important cellular responses such as survival, proliferation and differentiation (Watowich *et al.*, 1996; Kuby, 1997b).

The process of haematopoiesis relies on the interaction of haematopoietic cytokines, their receptors and downstream pathways (de Groot *et al.*, 1998; Zhu and Emerson, 2002; Martinez-Moczygemba and Huston, 2003; Tenen, 2003). While the precise role of haematopoietic cytokines in the development of leukaemia is not fully understood, experimental evidence suggests that growth factors and their receptors may play a critical role in the development of leukaemia (Baird *et al.*, 1995; Tomasson *et al.*, 2001). There are two classes of receptors for the haematopoietic cytokines and growth

factors, the receptor tyrosine kinases and the cytokine receptors. As the work described in this thesis concerns a member of the cytokine receptor family, this family of receptors will be discussed in detail below. The role of growth factors and their receptors in the development of leukaemia will be discussed in Sections 1.4 and 1.5. What follows is a description of the structure of the cytokine receptor superfamily and the signalling pathways they utilize.

1.3.1 Structure

Members of the cytokine receptor superfamily share a number of common structural features, including; the cytokine receptor module consisting of two fibronectin type III domains (containing 7 β strands), 4 spatially conserved cysteine residues and a membrane proximal WSXWS motif in the extracellular domain (Bagley *et al.*, 1997) (see Figure 2). They lack intrinsic tyrosine kinase activity, which distinguishes them from the receptor tyrosine kinase family (Watowich *et al.*, 1996). Some of the cytokine receptors have additional intracellular conserved motifs called Box 1 and Box 2, and these regions of the receptor are necessary to generate a proliferative signal (Murakami *et al.*, 1991; Doyle and Gasson, 1998) (see Figure 2). The Box 1 motif has specifically been shown as the site of interaction for the primary signal transducing kinase JAK2 (Quelle *et al.*, 1994; Zhao *et al.*, 1995).

The cytokine receptor superfamily includes the receptors for growth hormone (GH), Epo, thrombopoietin (Tpo), interleukins 2-7, 9, 11 and 15, as well as the colony stimulating factors (GM-CSF, G-CSF), leukaemia inhibitor factor (LIF), oncostatin (OSM), prolactin, ciliary neurotrophic factor (CNTF), Cardiotropin, the P40 subunit of IL-12 and the leptin receptor (Tartaglia *et al.*, 1995; Watowich *et al.*, 1996). Cytokine receptors function as complexes consisting of between 2-4 homologous or heterologous receptor chains. In the case of the homodimeric receptors (e.g. GH, Epo, Tpo and G-CSF), the subunits fulfil the dual roles of binding and signalling. In multi-subunit receptors, the different subunits can perform either of these roles. The receptors for GM-CSF, IL-3 and IL-5 share a common β subunit (βc) that associates with specific α subunits resulting in high affinity ligand binding. In this receptor system, both chains are involved in signalling (Bagley *et al.*, 1997). Heterodimerization of the α and β subunits activates a signalling cascade involving the JAK/STAT pathway, Ras/Raf/mitogen-activated protein (MAP) kinases, (Birnbaum *et al.*, 2000)

phosphatidylinositol 3-kinase (PI3K)/Akt pathway (Guthridge *et al.*, 1998) and the Src family kinases (SFKs) (Dahl *et al.*, 2000).

NOTE: This figure is included on page 16 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2: The cytokine receptor family. Schematic of the cytokine receptor superfamily divided into subfamilies with a representative member shown with other members stated below. The extracellular cytokine receptor module (CRM) is indicated and comprises two fibronectin-like β barrel structures called cytokine receptor domains (CRD). The WSXWS region is indicated by a dark blue ellipse, red and yellow bars in the cytoplasmic tail of some receptor components represent Box 1/Box 2 motifs respectively. The IL-2Rα is marked in grey as it is not a member of the cytokine receptor superfamily.

Adapted from Watowich *et al*., (1998)

1.3.2 Generation of intracellular signals by cytokine receptors

Signalling from the cytokine receptor families involves a number of key pathways that will be discussed in the following sections and are illustrated for the GM-CSF receptor in Figure 3.

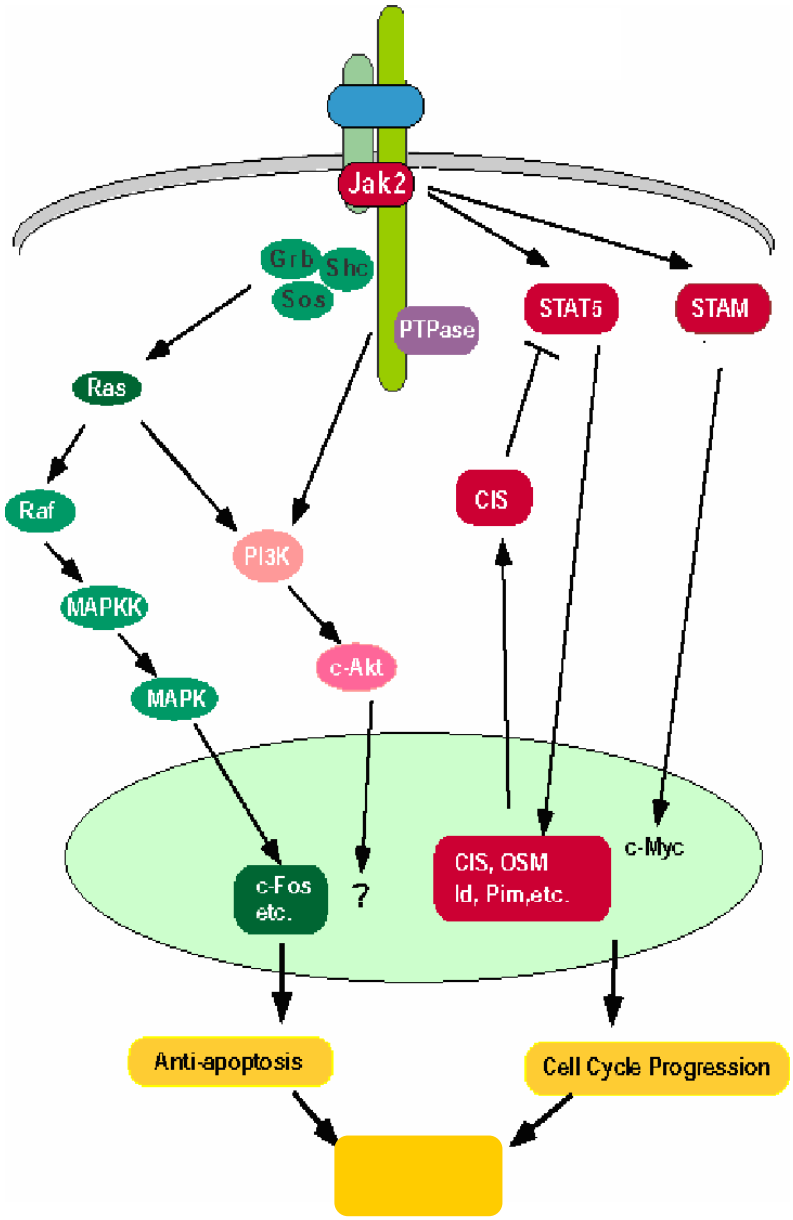


Figure 3: Cytokine receptor signalling pathways emanating for the GM-CSF receptor (GM-CSFR). GM-CSF binding to the GM-CSFR causes activation of the receptor and results in phosphorylation of receptor associated Jak2 and signalling via the JAK/STAT pathway. Other signalling pathways activated upon receptor activation include the Ras/Raf/MAP kinase pathway and the PI3K/Akt pathway (for review see de Groot *et al*., (1998)). The activation of these pathways leads to a variety of physiological events including cell cycle progression and the inhibition of apoptosis which results in cell survival and/or proliferation.

1.3.2.1 The JAK/STAT pathway

The cytokine receptor superfamily interacts with non-receptor tyrosine kinases, the Janus kinases (JAKs), which are critical for signalling (reviewed in O'Shea, 1997). Ligand binding to the receptor causes activation of the receptor and results in phosphorylation of receptor associated JAKs (Miyajima *et al.*, 1993). JAK activation (thought to occur by transphosphorylation) results in tyrosine phosphorylation of intracellular signalling molecules, and phosphorylation of tyrosine residues on the receptor (O'Shea, 1997). In the case of the GM-CSF/IL-3/IL-5 receptors, this phosphorylation is mediated by JAK2 which is essential for signalling for all three receptors (Quelle *et al.*, 1994; Watanabe *et al.*, 1997; Parganas *et al.*, 1998). Recently, an activating JAK2 mutation has been found in patients suffering from myeloproliferative disorders (MPD) (James *et al.*, 2005; Kralovics *et al.*, 2005a; Kralovics *et al.*, 2005b; Levine *et al.*, 2005). Mutation of JAK2 in MPD is consistent with observations implicating the JAK/STAT signal transduction pathway in pathogenesis of MPD (Mahon, 2005; Tefferi and Gilliland, 2005).

Receptor tyrosine phosphorylation provides binding sites for proteins with SH2 or PTB domains (Pawson and Scott, 1997; Hunter, 2000). One such group of SH2 containing proteins includes the signal transducers and activators of transcription (STATs) (O'Shea, 1997). Many of the recruited proteins become phosphorylated by JAKs or other activated kinases (e.g. SFKs) (Ward *et al.*, 2000). Once phosphoryled, STATs dissociate from the receptor, enabling the formation of dimers that are able to translocate to the nucleus and activate transcription (Ihle, 1996). STATs contribute to the specificity of cytokine signalling by selectively binding to tyrosine phosphorylated residues on specific cytokine receptors and binding to specific DNA elements to mediate gene regulation (O'Shea, 1997). STAT5 activation plays a key role in proliferative responses to IL-3, IL-5 and GM-CSF *in vitro* and *in vivo* (Mui *et al.*, 1995; Mui *et al.*, 1996; Teglund *et al.*, 1998; Ilaria *et al.*, 1999). Evidence suggests that inappropriate activation of STAT5 may contribute to the development of leukaemia (Gouilleux-Gruart *et al.*, 1997; Kelly *et al.*, 2003). Persistent signalling of STATs (e.g. constitutively activated STAT5) has been demonstrated to directly contribute to oncogenesis by stimulating cell proliferation and preventing apoptosis (Buettner *et al.*, 2002; Moore, 2005). Nyga *et al.*, (2005) showed that the association of STAT5 with the scaffolding protein Gab2, plays an essential role in the constitutively activated STAT5-

induced cell proliferation and survival by activating the Ras/Raf/MAP kinase and PI3K/Akt pathways.

1.3.2.2 The Ras/Raf/MAP kinase pathway

Ras proteins are plasma membrane proteins that bind guanine nucleotides and are highly conserved throughout evolution. They regulate differentiation, progression into S phase and immunological responses (Moodie and Wolfman, 1994). The localization of Ras to the plasma membrane is critical for its function, since it is here that it interacts with upstream activators. One such activator is the Grb2-Sos nucleotide exchange factor complex, which converts Ras to its active GTP bound state following transmembrane receptor activation (Magee and Marshall, 1999). Raf proteins, which are activated by Ras, are a family of serine/threonine kinases that act downstream of Ras to deliver a proliferative signal to the cell (Moodie and Wolfman, 1994). Activation of the MAP kinase family can occur by stimulation of the Ras-dependent signalling system or via expression of an activated Ras protein (Moodie and Wolfman, 1994). Using βc as an example, SHC and Grb2 adapter proteins bind to phosphorylated tyrosine residues of the βc forming a complex that recruits the guanine nucleotide exchange factor, Sos, enabling activation of Ras and downstream members of the MAP kinase pathway (Guthridge *et al.,* 1998). Active Raf protein activates MEK (a MAP kinase kinase or MAPKK) which in turn can activate MAP kinases (Moodie and Wolfman, 1994). The MAP kinase pathway can also be activated by a Ras-dependent but Raf-independent method (Moodie and Wolfman, 1994).

Ras mutations are common in cancer, occurring in approximately 30% of human tumours, as well as many other forms of human cancer, including myeloid leukaemias (Wittinghofer, 1998; Birnbaum *et al.*, 2000). Oncogenic point mutations in *Ras* genes cause constitutive activation of the Ras pathway that leads to elevated levels of the active guanosine triphosphate bound form of Ras (Ras-GTP) (Aricò *et al.*, 1997; Birnbaum *et al.*, 2000).

1.3.2.3 The PI3K/Akt pathway

Phosphoinositide 3-kinase (PI3K) consists of a regulatory subunit (p85) containing both SH2 and SH3 domains and a catalytic subunit (p110) (Guthridge *et al.*, 1998; Scheid *et al.*, 2001). There are multiple forms of PI3K that are able to phosphorylate phosphatidylinositols and generate secondary messenger molecules (Franke *et al.*, 1997). Phosphatidylinositol (PI) is a membrane phospholipid that can be phosphorylated

on the 3, 4, and 5 positions of its inositol ring to form seven unique lipid signalling molecules called phosphoinositides (Vanhaesebroeck *et al.*, 2001). One class of secondary messenger molecules is generated by the phosphorylation of phosphatidylinositides on the 3 position by PI3K. These second messengers generated by PI3K act on numerous downstream effector molecules including Src homology-2 (SH2) and Pleckstrin homology (PH) domains of serine/threonine kinase and tyrosine kinases as well as cytoskeletal proteins (Franke *et al.*, 1997). For example, PI3K is involved in the conversion of phosphatidylinositol 4,5-bisphophosphate ($PI(4,5)P_2$) to phosphatidylinositol 3,4,5-trisphophosphate ($PI(3,4,5)P_3$) and increases in $PI(3,4,5)P_3$ leads to increased activation of the downstream kinase Akt/PKB (Carricaburu et al 2003). Many of the diverse signalling responses of the PI3K pathway are mediated by the serine/threonine kinase Akt/PKB (protein kinase B) (Franke *et al.*, 1997; Scheid *et al.*, 2001). PI3K is necessary and sufficient for growth factor-dependent Akt activation (Franke *et al.*, 1997; Scheid *et al.*, 2001). Ras binds and activates PI3K causing the activation of Akt/PKB, but Akt/PKB does not appear to be involved in the pathway leading to MAP kinase activation (Franke *et al.*, 1997). This finding is consistent with the existence of separate signalling pathways for Ras that are distinct from the Raf/MEK/MAP kinase pathway.

Serine and threonine phosphorylation offers a mode of coupling signal transduction pathways that are distinct from tyrosine phosphorylation. In particular, this has been demonstrated in the βc where Ser-585, when phosphorylated, mediates the association of βc with 14-3-3 and activation of the PI3K/Akt pathway (Guthridge *et al.*, 2000). The interaction of 14-3-3 with Ser-585 of βc can initiate cell survival signals. 14-3-3 binds to the pro-apoptotic protein Bad and thus plays a role in mediating cell survival (Pawson and Scott, 1997; Guthridge *et al.*, 1998). 14-3-3 also binds to Cdc25C, a phosphatase, which controls the cell cycle progression by regulating the activity of Cdc2, a protein kinase that controls cell entry into mitosis (Pawson and Scott, 1997). The phosphorylation of Ser-585 was shown to be mediated by protein kinase A (PKA), implicating cAMP in IL-3 mediated cell survival, and suggests a role for serine phosphorylation and phosphoserine binding proteins in mediating some of the biological responses of receptors (Guthridge *et al.*, 2000). This would be consistent with observations that cytokine receptors with mutated tyrosine residues can still mediate cellular responses albeit at reduced levels (Okuda *et al.*, 1997; Longmore *et al.*, 1998; Okuda *et al.*, 1999).

1.4 Growth factor receptors in haematopoiesis and leukaemia

Mutations may occur in growth factor receptors and some of these may result in chronic activation of the receptor, which can in turn overcome the requirement for growth factors in maintaining cell viability and in driving proliferation. Mutation of both tyrosine kinase (e.g. FLT3) and cytokine receptors (EpoR, c-Mpl and G-CSFR) have been shown to occur in leukaemia and these receptors are discussed below. The involvement of the GM-CSFR in leukaemia will be discussed in Section 1.5.

1.4.1 Tyrosine kinase receptors – FLT3

Fms-like kinase 3 (FLT3) is a member of the receptor tyrosine kinase family which has been demonstrated to play an important role in the development of multipotent stem cells and B cells. This role has been demonstrated through studies of mice deficient in Flt3 (Mackarehtschian *et al.*, 1995) and by its expression in normal bone marrow being restricted to early progenitors, including $CD34^+$ cells and those expressing high levels of c-KIT (Rasko *et al.*, 1995; Rosnet *et al.*, 1996). FLT3 is expressed at high levels in AML, ALL, some T-cell ALL and CML in lymphoid blast crisis (Drexler, 1996; Rosnet *et al.*, 1996). Nakao *et al.*, (1996) discovered the presence of internal tandem duplications (ITD) in the juxtamembrane domain of FLT3 in AML. FLT3-ITD were reported by Gilliland and Griffin (2002) to occur in 24% of AML patients and a lower frequency in myelodysplastic syndrome (MDS) but rarely in ALL. FLT3-ITD have been detected in all FAB subtypes of AML, with a higher occurrence in M3 and a lower occurrence in M2 (Gilliland and Griffin, 2002). It has also been discovered that mutations in the activation loop of FLT3 have been detected in 7% of AML, 3% of MDS and 3% of ALL patients (Gilliland and Griffin, 2002). Activation loop mutations have also been detected in leukaemias involving other receptor tyrosine kinases including c-Kit, MET and RET (Gilliland and Griffin, 2002; Reilly, 2003; Moore, 2005). As such, approximately 30% of AML patients have acquired mutations in FLT3 leading to constitutive activation and making it the single most commonly mutated gene in AML. However, additional mutations are likely to be necessary for development of AML. This is due to the fact that neither FLT3 expression or FLT3 mutations are sufficient to cause AML, but rather act as secondary events (not necessary for initiation but) required for progression to overt AML (Fenski *et al.*, 2000; Yamamoto *et al.*, 2001; Kelly *et al.*, 2002; Moore, 2005).

It has been suggested that cooperation between chromosomal translocations involving transcription factors (e.g. those found in mixed lineage leukaemia; AML/ETO, CBFB/SMMHC, PML/RARA) and FLT3 mutations may be required for AML (Gilliland and Griffin, 2002; Moore, 2005). This suggestion, along with other evidence, has led to the proposal of a "two hit" model of leukaemogenesis (Gilliland and Griffin, 2002). In the two hit model, one class of mutations (e.g. FLT3 mutation or N-Ras mutations) provides a proliferative and survival advantage to the cells, and the second class of mutations interfere with haematopoietic differentiation and the eventual apoptosis of the differentiated cells. As mutations in Ras or FLT3 account for 50% of AML cases, it suggests that there may be other haematopoietic tyrosine kinases, other growth factor receptors or their downstream effectors that contribute to the pathogenesis of the remaining 50% of AML cases (Gilliland and Griffin, 2002).

1.4.2 Cytokine receptor mutations in leukaemia

1.4.2.1 Erythropoietin receptor

Erythropoietin (Epo) binds to the erythropoietin receptor (EpoR) to stimulate proliferation and differentiation of erythroid progenitor cells (Chretien *et al.*, 1994). The Friend spleen focus forming virus (SFFV) induces erythroleukaemia in mice and erythroblastosis in bone marrow cultures (Yoshimura *et al.*, 1990; Constantinescu *et al.*, 1999) and references therein). A key event in this transformation is the activation of the EpoR by gp55, the glycoprotein of the *env* gene of SFFV leading to the proliferation of erythroblasts characteristic of early Friend virus induced murine erythroleukaemia (Yoshimura *et al.*, 1990; Winkelmann *et al.*, 1995). There are two strains of SFFV, the anemic (A) and polycythemic (P) strains, so called because the latter but not the former can also induce polycythemia (increased numbers of red blood cells) in mice (Tambourin *et al.*, 1979). In addition the gp55-A variant is much less effective in activating the murine EpoR (Constantinescu *et al.*, 1999). The transmembrane region of EpoR was shown by Zon *et al.*, (1992) to be critical for activation by gp55 and mutation of Leu238 in human EpoR to its murine equivalent (Ser) results in its activation by gp55-P (Constantinescu *et al.*, 1999). The initial stage of Friend disease shows Epo-independent polyclonal erythroblastosis and is followed by malignant transformation of erythroid progenitors that are inhibited in their erythroid differentiation and promote transformation to erythroleukaemia (Constantinescu *et al.*, 1999 and references therein).

These studies implicate the EpoR in the erythroleukaemia associated with Friend disease.

In transgenic mice, activation of the EpoR by high levels of Epo causes erythrocytosis but does not induce leukaemic cells (Semenza *et al.*, 1990). The EpoR can be constitutively activated by a point mutation, R129C, that causes an intermolecular disulphide bond between the extracellular membrane proximal segments (Watowich *et al.*, 1994). Furthermore, expression of this mutant receptor in progenitors results in Epo-independent red-cell formation and leukaemias of several lineages (Longmore and Lodish, 1991; Longmore *et al.*, 1992; Longmore *et al.*, 1993). These results indicate that constitutive activation of the EpoR can lead to cell transformation in some contexts and implicate the EpoR in leukaemia (Constantinescu *et al.*, 1999).

1.4.2.2 c-Mpl

Thrombopoietin (Tpo) is the primary regulator of the growth and differentiation of megakaryocytic progenitor cells and the major modulator of platelet production (Ooi *et al.*, 1998; Gaur *et al.*, 2001 and references therein). However, it also expands erythroid and granulocyte progenitors in mice and stimulates a significant proportion of the primary AML cells to grow (Matsumura *et al.*, 1995; Kaushansky *et al.*, 1996). The Tpo receptor is encoded by the cellular homologue of the *v-Mpl* oncogene (*c-Mpl*) and loss of either the receptor or ligand leads to at least an 85% decrease in circulating platelets (Gurney *et al.*, 1994; Alexander *et al.*, 1996; de Sauvage *et al.*, 1996). v-Mpl consists of the cytoplasmic and transmembrane sequences of c-Mpl, but lacks all but the last 43 amino acids of the extracellular domain of mouse c-Mpl.

The biological properties of the myeloproliferative leukaemia virus (MPLV) are due to v-Mpl (Souyri *et al.*, 1990). MPLV is a defective murine retrovirus that, in the presence of helper virus, induces an acute myeloproliferative disease characterized by hepatosplenomegaly and polycythemia (Wendling *et al.*, 1986; Penciolelli *et al.*, 1987). *In vitro* infection of bone marrow cells with MPLV promotes growth and terminal differentiation of haematopoietic progenitor cells in the absence of growth factors (Wendling *et al.*, 1989a). Similarly, infection of growth factor dependent cell lines with MPLV abrogates their requirement for growth factors (Wendling *et al.*, 1989b). Cysteine substitutions in the dimer interface domain of c-Mpl can also lead to constitutive activation, presumably mediated through receptor dimerization (Alexander *et al.*, 1995; D'Andrea *et al.*, 1996). Thus in some contexts, Tpo receptor mutations can affect cell growth and intracellular signalling and result in leukaemia.

1.4.2.3 Granulocyte colony-stimulating factor receptor

Granulocyte colony-stimulating factor (G-CSF) plays a central physiological role in the regulation of neutrophil production in health, and particularly in emergency responses to infections and bone marrow aplasia (Roberts, 2005). Consequently, G-CSF-deficient mice and G-CSF receptor (G-CSFR)-deficient mice are characterised by; a resting neutropenia, deficiency in granulocyte and macrophage progenitors, a markedly reduced capacity to mount a neutrophilic response to bacterial and fungal infections, an increased mortality from these infections, a susceptibility to bacterial pneumonia, and a propensity to develop reactive amyloidosis with age (Lieschke *et al.*, 1994; Liu *et al.*, 1996; Seymour *et al.*, 1997; Basu *et al.*, 2000). Autocrine production of G-CSF occurs in CML (Jiang *et al.*, 1999). Transgenic mice expressing G-CSF or transplantation of murine bone marrow cells retrovirally transduced to produce G-CSF causes myeloproliferative disease in mice (Chang *et al.*, 1989; Yamada *et al.*, 1996). G-CSFR is expressed by all neutrophils and their precursors, and expression increases with the progress of cell maturation (Nicola and Metcalf, 1985; McKinstry *et al.*, 1997; Shinjo *et al.*, 1997). G-CSFR is also expressed on myeloid leukemia cells and cell lines and many non-haematopoietic tumour cell lines (Kimura and Sultana, 2004; Roberts, 2005).

Some patients with severe congenital neutropenia (a rare disorder of granulopoiesis) with progression to myelodysplastic syndrome (MDS)/AML were reported to have mutations of G-CSFR (Dong *et al.*, 1995a; Dong *et al.*, 1997). Activating mutations of G-CSFR have been found in AML (Dong *et al.*, 1995b; Forbes *et al.*, 2002) and truncation of the receptor has been recognised as an acquired event in severe congenital neutropenia that can be associated with subsequent progression to AML (Roberts, 2005). These truncations interfere with the binding of negative (and positive) regulatory elements and allow the truncated receptor to act in a dominant negative manner (Dong *et al.*, 1997). An activating mutation in the transmembrane region of the G-CSFR was found in two cases of AML and this mutation conferred factor independent growth on Ba/F3 cells (Forbes *et al.*, 2002).

1.5 GM-CSF, IL-3 and their shared receptor (hβc) in leukaemia

1.5.1 Biology of GM-CSF, IL-3 and IL-5

The cytokines IL-3, IL-5 and GM-CSF affect the survival, proliferation, differentiation and functional activation of cells within the haematopoietic system (Metcalf and Nicola, 1995; Martinez-Moczygemba and Huston, 2003). The effects of IL-5 are restricted to the eosinophilic lineage, GM-CSF acts on neutrophil and monocyte lineages, whereas IL-3 influences all five myeloid lineages and the early maturation of erythroid cells (Metcalf and Nicola, 1995; Bagley *et al.*, 1997; de Groot *et al.*, 1998). As mentioned above, these cytokine receptors consist of a ligand specific α subunit (IL-3Rα, IL-5Rα and GMRα) which, on ligand binding, associates with the shared βc receptor subunit to form an active complex. GM-CSF binds to the GMRα with low affinity as a monomer, but does not bind βc. Formation of high affinity complex occurs through association of GM-CSF, the GMRα and βc. Both the GMRα and βc are necessary for signal transduction and activation of downstream pathways (see above) (Doyle and Gasson, 1998). In mice there is also an additional IL-3 specific β chain ($\beta_{IL\text{-}3}$) that dimerizes only with IL-3Rα, resulting in two different receptors that are capable of responding to IL-3 in the mouse (Gorman *et al.*, 1990; Hara and Miyajima, 1992; de Groot *et al.*, 1998). Relatively closely related species such as rats do not have a second β subunit, suggesting that the two subunits have arisen by chromosomal duplication late in evolution (Gorman *et al.*, 1992; Appel *et al.*, 1995). The two subunits share 91% amino acid sequence identity with most changes being homologous (Hara and Miyajima, 1992). Targeted inactivation of each subunit suggests that they deliver similar signals upon IL-3 stimulation (Nishinakamura *et al.*, 1995; Nicola *et al.*, 1996) and is discussed in more detail below.

1.5.1.1 IL-3/IL-5/GM-CSF system deficient mice

GM-CSF-deficient mice are viable and display normal haematopoiesis indicating that GM-CSF plays a non-essential role in steady state haematopoiesis (Dranoff *et al.*, 1994; Stanley *et al.*, 1994). However, these mice developed a progressive accumulation of surfactant lipids and proteins in the alveolar space and extensive lymphoid hyperplasia associated with lung airways and blood vessels, revealing a role for GM-CSF in pulmonary homeostasis (Dranoff *et al.*, 1994; Stanley *et al.*, 1994). GM-CSF null animals have reveal an important role for GM-CSF in surfactant clearance by alveolar macrophages (Huffman *et al.*, 1996; Ikegami *et al.*, 1996; Nishinakamura *et al.*, 1996b).

Null animals also display compromised antigen-specific and LPS induced T-cell responses and IFNγ production which may be dendritic cell mediated (Wada *et al.*, 1997; Noguchi *et al.*, 1998). GM-CSF null animals have defects in macrophage function (Basu *et al.*, 1997; Scott *et al.*, 1998; Zhan *et al.*, 1999; Enzler *et al.*, 2003) and are susceptible to various infections agents (Zhan *et al.*, 1998; LeVine *et al.*, 1999; Paine *et al.*, 2000). Recently, GM-CSF has been shown to have an important accessory role in radioprotection by donor haematopoietic cells (Katsumoto *et al.*, 2005).

IL-3 deficient mice have defects in mast cell and basophil production, as well as in their immune response to parasitic infection (Lantz *et al.*, 1998). This indicates that IL-3 plays a role in host defense against infection and suggests it has a function as an emergency haematopoietic regulator in states that involve T lymphocyte activation.

IL-5 deficient mice generated by Kopf *et al.*, (1996) showed reduced levels of $CD5^+$ B cells (B-1 cells) in young mice, which returned to normal in adult mice. Eosinophils with normal morphology were produced in these IL-5 deficient mice, however, basal levels were reduced 2- to 3-fold (eosinopenia) compared to control mice. Under conditions of emergency haematopoiesis, caused by parasitic infection, no eosinophilia was seen (Kopf *et al.*, 1996). The reduced numbers of B-1 cells and eosinophils observed in IL-5 deficient mice were similar to mice with a null mutation of the IL-5 receptor alpha chain (Yoshida *et al.*, 1996). Thus, IL-5 appears to have important roles in both steady state and emergency eosinophilopoiesis and in B cell development *in vivo*.

In βc null ($\beta c^{-/-}$) mutant mice, eosinophil numbers were reduced in the peripheral blood and bone marrow, and there was a lack of eosinophilic response to parasites. This indicated a role for the βc in eosinophil function, growth differentiation and survival (Nishinakamura *et al.*, 1995; Robb *et al.*, 1995). Additionally, in clonal cultures of $\beta c^{-/-}$ bone marrow cells, IL-5 and GM-CSF failed to stimulate colony formation (Robb *et al.*, 1995). This defined an essential role for βc in mediating GM-CSF and IL-5 stimulation. It was shown, however, that IL-3 function was normal in these mice, consistent with IL-3 signalling through receptors containing β_{IL3} (Nishinakamura *et al.*, 1995; Robb *et al.*, 1995). βc null mice also developed a lung pathology involving lymphocytic infiltration and characteristics resembling alveolar proteinosis, similar to the phenotype of GM-CSF-deficient mice (discussed above) (Nishinakamura *et al.*, 1995; Robb *et al.*, 1995; Scott *et al.*, 1998).

Inactivation of the β_{IL3} gene resulted in no obvious haematopoietic abnormalities and suggests that each subunit (βc and β_{IL3}) delivers similar signals upon IL-3 stimulation (Nishinakamura *et al.*, 1995; Nicola *et al.*, 1996). Therefore, to examine the role of the entire IL-3/IL-5/GM-CSF system in haematopoiesis, double knockout mice deficient for IL-3 ligand and βc were produced by Nishinakamura et al., (1996a). These mice displayed the same features as the βc mutant mice (Nishinakamura *et al.*, 1995; Robb *et al.*, 1995). To investigate the role of the IL-3/IL-5/GM-CSF system in emergency haematopoiesis, Nishinakamura *et al.*, (1996) challenged the double knockout mice with *Listeria monocytogenes* and saw no defects in innate or acquired immunity. Also, treatment with the cytotoxic drug, 5-fluorouracil, revealed no differences in haematopoietic recovery in the double knockout mice compared to βc knockout or control mice (Nishinakamura *et al.*, 1996a). These findings led Nishinakamura *et al.*, (1996a) to suggest that the GM-CSF/IL-3/IL-5 system is dispensable for both normal haematopoiesis and haematopoiesis under demand (i.e. in times of emergency and immunity). The report by Nishinakamura *et al.*, (1996a) contrasted earlier studies (Dranoff *et al.*, 1994; Stanley *et al.*, 1994; Nishinakamura *et al.*, 1995) suggesting that the GM-CSF/IL-3/IL-5 system is required for haematopoiesis under demand (i.e. in times of infection). Taken together with the more recent findings from the GM-CSF null mice discussed above (concerning defective macrophage function and susceptibility to infection), it seems apparent that the GM-CSF/IL-3/IL-5 system is required for haematopoiesis under demand.

1.5.2 Involvement of GM-CSF and IL-3 in human diseases

1.5.2.1 Autocrine production and over-expression of GM-CSF and IL-3 in leukaemia

Autocrine production and over-expression of GM-CSF and IL-3 has been linked with leukaemia in both human disease and animal models. Autocrine GM-CSF production occurs in acute lymphoblastic leukaemia (ALL), acute myeloblastic leukaemia (AML) and JMML (also known as juvenile chronic myelogenous leukaemia (JCML)) (Young and Griffin, 1986; Gualtieri *et al.*, 1989; Freedman *et al.*, 1993). Confirmation of the role autocrine GM-CSF production plays in mediating the proliferation of JMML cells was suggested by Geissler *et al.*, (1996), who showed that the growth of JMML cells could be inhibited by the use of antibodies against GM-CSF. Similar findings have

come from studies that demonstrated that the growth of JMML cells could be stimulated by GM-CSF and inhibited by the GM-CSF antagonist, E21R (Iversen *et al.*, 1997; Ramshaw *et al.*, 2002). In a mouse model, Iversen *et al*., (1997) showed that E21R could prevent engraftment and induce remission of human JMML in an *in vivo* setting. A fatal myeloid disorder with some features of JMML can be induced by transplanting bone marrow over-expressing GM-CSF into mice (Johnson *et al.,* 1989). The specific involvement of GM-CSF in JMML will be discussed in more detail below (see Section 1.5.2.2).

Autocrine production of IL-3 occurs in CML (Jiang *et al.*, 1999; Jiang *et al.*, 2000). This was noted by Jiang *et al*., (1999) to provide an explanation for the similarities between human CML and murine models of myeloproliferative disease resulting from transplantation of mouse bone-marrow retrovirally transduced to produce IL-3 (Chang *et al.*, 1989; Wong *et al.*, 1989; Just *et al.,* 1993). $CD34^+$ leukaemic cells from CML patients proliferate *in vitro* in the absence or serum or growth factors (Jiang *et al.*, 1999; Jiang *et al.*, 2000). The mechanism by which this occurs involves an autocrine production of IL-3 (and G-CSF) and the stimulation of STAT5 phosphorylation in $CD34^+$ leukaemic cells (Jiang *et al.*, 1999).

1.5.2.2 Involvement of GM-CSF in juvenile myelomonocytic leukemia (JMML)

JMML is a rare myeloproliferative disorder (MPD), which affects infants and young children, and is characterised by an overproduction of myeloid cells and tissue infiltration (reviewed in Aricò *et al.*, 1997). There is evidence that hypersensitivity to GM-CSF may contribute to the expansion of cells in JMML. JMML cells, when stimulated with low concentrations of GM-CSF, form excessive numbers of colony-forming unit granulocyte-macrophage (CFU-GM) colonies on methylcellulose compared to normal controls (Emanuel *et al.*, 1991; Emanuel *et al.*, 1996).

Loss of the neurofibromatosis type 1 (*NF1*) tumour suppressor gene encoding neurofibromin, a GTPase-activating protein (GAP) that negatively regulates Ras, may be critical in the abnormal GM-CSF response in JMML. Individuals with neurofibromatosis type 1 (NF1) are predisposed to specific cancers, including fibrosarcoma, pheochromocytoma, astrocytoma, and JMML (Birnbaum *et al.,* 2000). 7% of JMML cases are associated with NF1 (Aricò *et al.,* 1997). JMML cells from patients with NF1 frequently show deletion of the normal *NF1* allele leading to hyperactive Ras signalling (Birnbaum *et al.,* 2000). Mutated *Ras* genes have been

detected in 30% of JMML patients who do not have NF1, but are absent in patients with NF1 (Aricò *et al.*, 1997; Birnbaum *et al.*, 2000). Human and mouse data suggests that the tumour suppressor function of the *NF1* gene in myeloid cells is mediated through the negative regulatory action of neurofibromin on Ras (Birnbaum *et al.*, 2000), and that the deregulated signalling through the Ras pathway is the central event in the abnormal growth of JMML progenitor cells (Aricò *et al.*, 1997). Furthermore, Birnbaum *et al.*, (2002) show that GM-CSF plays a key role in establishing and maintaining the MPD that results from transplantation of $Nf1^{-/-}$ haematopoietic cells into irradiated mice. They show that physiological levels of GM-CSF are sufficient to drive excessive proliferation, in particular myeloid proliferation, in their mouse model of JMML. In addition, recipient mice of $Nf1^{-/-}$ $GM\text{-}CSF^{-/-}$ haematopoietic cells are hypersensitive to exogenous recombinant GM-CSF. Since the neurofibromatosis type 1 (*NF1*) tumour suppressor gene encodes a GAP that negatively regulates Ras, this suggests an inability to down-regulate Ras-GTP in response to GM-CSF in human JMML and the similar mouse MPD (Birnbaum *et al.*, 2000).

1.5.3 β_c receptor activated mutants

1.5.3.1 Overview

A series of activating mutations in the common signal-transducing β subunit of the human GM-CSF/IL-3/IL-5 receptor (hβc) have been isolated based on their ability to induce cytokine independent proliferation of FDC-P1 cells (D'Andrea *et al.*, 1994; Jenkins *et al.*, 1995; D'Andrea *et al.*, 1996; Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000). These mutants were experimentally generated and to date have not been found to occur in human leukaemias (Freeburn *et al.*, 1998). The mutants can be classified based on the transmembrane or extracellular position of the mutation (Figure 4).

The hβc equivalent truncation to v-Mpl (called ΔH) has a 45 amino acid extracellular region. The ΔH mutation gives rise to factor independent growth when expressed in FDC-P1 cells (D'Andrea *et al.*, 1996). A truncation of domains 1-3 of hβc (ΔQP) could also confer factor independent growth, but not the truncation which removed only domains 1 and 2 (D'Andrea *et al.*, 1996).

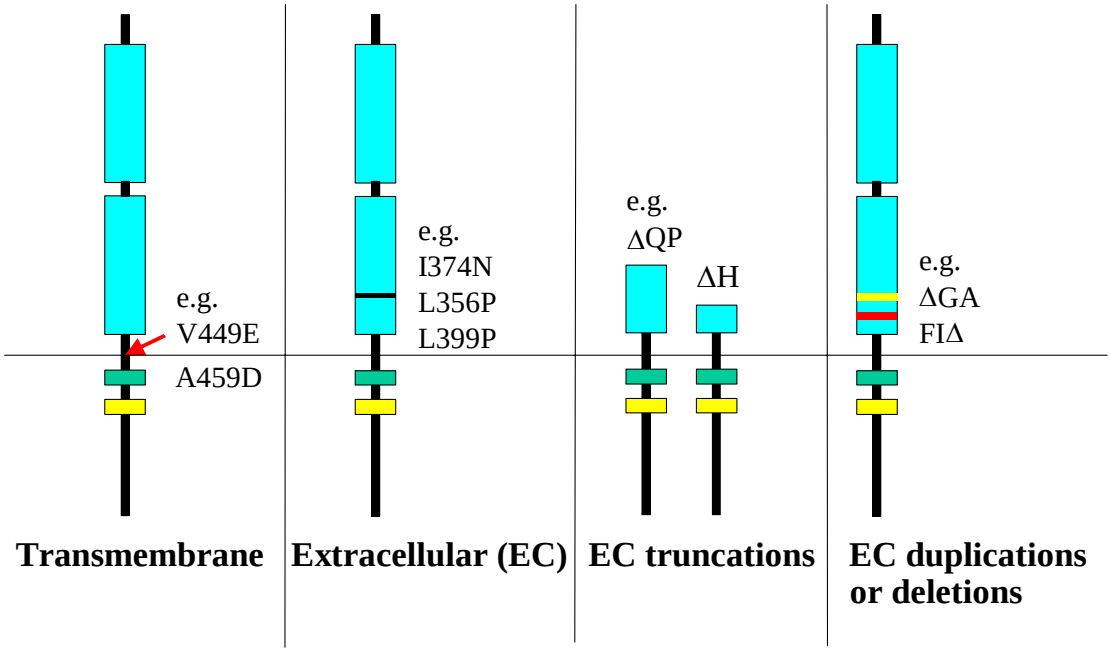


Figure 4: Activated mutants of the hβc. Activated mutants can be classed based on the position of the mutation (either a point mutation, truncation, duplication or deletion of a part of the hβc). Positions of activating mutations in hβc are shown schematically. Extracellular (EC) truncation mutants ΔQP and ΔH contain all or part of the membrane proximal cytokine receptor domain (CRD4) respectively. For review see D'Andrea *et al*., (2000).

The various mutants described here suggest a general mechanism whereby receptors can be activated by a variety of mutations (e.g. truncation, duplication or point mutation) of the receptor. The activation is believed to cause a conformational change in the receptor that mimics ligand induced receptor activation and leads to cell transformation (D'Andrea *et al.*, 1996; for review see Gonda and D'Andrea, 1997).

The hβc mutants can be further distinguished based on their requirement for the GMRα and their signalling properties (e.g. whether the tyrosine residues of the receptor are phosphorylated) (see Table 5 and D'Andrea, *et al*., 2000). As will be discussed in detail below, several of these mutants can induce haematopoietic disease when expressed in bone marrow by transgenesis or by retroviral transduction (D'Andrea *et al.*, 1998; McCormack and Gonda, 1999).

1.5.3.2 Isolation and in vitro properties of the hßc mutants in haematopoietic cells

(a) Isolation of hβc mutants in FDC-P1 cells

FDC-P1 is a myeloid progenitor cell line that was derived from an acute myeloid leukaemia and has an unknown block in differentiation (Dexter *et al.,* 1980). It is dependent on exogenous growth factors for growth and survival (eg. mIL-3 or mGM-CSF) and expresses murine GMRα, IL-3Rα, and β subunits (Lang *et al.*, 1985; Hapel *et al.*, 1986). To isolate activating mutations in hβc, random PCR mutagenesis of hβc cDNA was combined with retroviral expression cloning to screen for clones that could confer factor independent growth on FDC-P1 cells (Jenkins *et al.,* 1995). This strategy led to the isolation of a valine to glutamic acid substitution at residue 449 (V449E) within the transmembrane domain that was due to an A to T mutation at nucleotide 1374. Similarly, a T to A substitution of the nucleotide at position 1149 resulted in the change of isoleucine to asparagine at residue 374 (I374N). The I374N mutation is positioned in the extracellular region of the hβc and like the V449E mutation conferred factor independence on FDC-P1 cells (Jenkins *et al.,* 1995). Previous studies by D'Andrea *et al.*, (1994) also identified an activated hβc mutant that contained a 37 amino acid duplication in the membrane proximal region of extracellular domain called FIΔ that confers factor independent proliferation on FDC-P1 cells.

(b) Properties of the hßc mutants in BAF-B03, CTLL-2 and CTL-EN cells

The hβc mutants were introduced into another murine IL-3 dependent cell line, BAF-B03, which is a subline of the pro-B cell line Ba/F3 (Jenkins *et al.,* 1995). This tested whether the factor independence of these mutants was specific to the FDC-P1 cell line. The infection of BAF-B03 cells with constructs expressing activated hβc mutants revealed that the V449E mutant enabled factor independent growth but extracellular mutants (I374N or FIΔ) did not confer independent growth in this cell line (Jenkins *et al.,* 1995). The different abilities of the mutants to confer factor independence on the BAF-B03 cells indicate that there may be cell-type specific molecules that interact with mutant receptors and lead to their activity.

CTLL-2 is an IL-2-dependent T-cell line (Cerottini *et al.*, 1974) that does not express endogenous GM-CSF/IL-3/IL-5 receptor chains and could be used to investigate the function of the activated hβc mutants in the absence of any endogenous GM-CSF/IL-3/IL-5 receptor subunits (Jenkins *et al.*, 1995). Factor independent cells

were not produced by expression of either I374N, FIΔ or V449E activated hβc mutants in CTLL-2 cells (Jenkins *et al.*, 1995). The fact that expression of the V449E mutant in BAF-B03 cells produced factor independent cells, but did not give rise to factor independent cells in CTLL-2 cells, indicates that the expression of cell type specific molecules is also a requirement for the function of the V449E mutant in CTLL-2 cells (Jenkins *et al.*, 1995).

CTL-EN cells are derived from CTLL-2, by the introduction of murine ecotropic receptor (Baker *et al.*, 1992), which renders them more susceptible to retroviral infection. CTL-EN cells were used to investigate the requirements of hβc mutants for cell type specific molecules, specifically the requirement for the GMRα. A study by Jenkins *et al.*, (1999) identified that the mGMRα (but not the hGMRα) subunit cooperates with I374N in CTL-EN cells to enable factor independent growth. In contrast, expression of V449E and mGMRα in CTL-EN cells did not produce factor independent cells (Jenkins *et al.*, 1999).

The role of mGMRα in mediating the cell type specific factor independence of I374N in CTL-EN cells was examined in BAF-B03 cells to determine whether this effect was limited to CTL-EN cells. BAF-B03 cells (which do not express the GMRα) were shown to give rise to factor independent cells when the mGMRα subunit was co-expressed with the I374N activated mutant (Jenkins *et al.*, 1999). It was also found that both the hGMRα and mGMRα could cooperate with FIΔ in BAF-B03 cells to enable factor independent growth (Jones *et al.*, 2001). Thus, extracellular hβc mutants (I374N and FIΔ) require mGMRα for their factor independence in BAF-B03 or CTL-EN cells. In contrast, the transmembrane mutant (V449E) does not require co-expression of other receptor components to produce factor independent BAF-B03 cells, but its expression in CTL-EN cells does not produce factor independent cells (Jenkins *et al.*, 1995; Jenkins *et al.*, 1999). Thus, extracellular and transmembrane mutants are activated by fundamentally different mechanisms and it has been proposed that they represent different intermediates of the active receptor complex (Jenkins *et al.*, 1995; Jenkins *et al.*, 1999; D'Andrea and Gonda, 2000).

(c) Expression of activated hβc in primary haematopoietic cells

To investigate the effect of expression of three constitutively active mutants of hβc, McCormack and Gonda, (1997) introduced these mutants into primary haematopoietic cells (foetal liver). Expression of activated hβc mutants (I374N, FIΔ and V449E) in primary murine foetal liver cells was achieved using retroviral transduction via the RufNeo retrovirus (Rayner and Gonda, 1994). It was shown that expression of either FIΔ or I374N confers factor independent proliferation and differentiation on cells of the neutrophil and monocyte lineages. The expression of V449E conferred factor independence on the same lineages as the extracellular mutants and also conferred factor independence on cells of eosinophil, basophil, megakaryocyte and erythroid lineages (McCormack and Gonda, 1997). The lineage specificity of the extracellular mutants is consistent with the requirement of the FIΔ and I374N mutants for the GMRα subunit, which is restricted to the cells of the myeloid lineage (Jones *et al.,* 2001). Furthermore, the different factor independent cell types generated by the transmembrane and extracellular mutants suggests there are signalling differences between these two classes of hβc mutants (McCormack and Gonda, 1997).

(d) Expression of activated hβc in novel bi-potent myeloid cell lines

Novel haematopoietic cell lines were created from the retroviral infection of murine foetal liver cells with the RufNeo retroviral vector containing the wild type hβc cDNA and grown in long term culture in the presence of mIL-3, mGM-CSF and hEpo (McCormack and Gonda, 1997; McCormack and Gonda, 2000). From two cultures, two novel murine haematopoietic cell lines (FDB-1 and FDB-2) were generated that were resistant to G418, indicating the incorporation of RufNeo-hβc provirus into these cells (McCormack and Gonda, 2000). Southern blot analysis reveals that the FDB-1 and FDB-2 cell lines contained two and eight copies of the RufNeo hβc provirus respectively and in the case of FDB-1 cell line, one of the two integration sites contained a 1.2 kb deletion (McCormack and Gonda, 2000). Cells of the two cell lines died in the absence of growth factors (mIL3, mGMCSF or hEpo) but proliferated in the presence of mIL-3 with a small amount of spontaneous differentiation to neutrophils, monocytes and megakaryocytes. In the presence of mGM-CSF both cell lines differentiated along the neutrophil and monocyte lineages (McCormack and Gonda, 2000). The response to IL-3 and GM-CSF for the FDB-1 cell line is shown in Figure 5.

The FDB cell system has potential for delineating the signals involved in myeloid differentiation and proliferation mediated by IL-3 and GM-CSF receptors.

The effects of expression of constitutively active hβc mutants in FDB-1 and FDB-2 cell lines were investigated by McCormack and Gonda, (2000). They showed that when the activating mutants are expressed in these cell lines in the absence of exogenously added mIL-3 and mGM-CSF, the extracellular mutants delivered a differentiation signal (like mGM-CSF), whereas the transmembrane mutant, V449E, conferred factor-independent proliferation, maintaining an undifferentiated phenotype showing pro-myelocyte morphology (McCormack and Gonda, 2000). The response of the FDB-1 cell line grown in the absence of mIL-3 and mGM-CSF but in the presence of either FIΔ or V449E expression is shown in Figure 5. The key findings from the use of the FDB cell lines were that the extracellular mutants (FIΔ and I374N) deliver a differentiation signal, whereas the transmembrane mutant (V449E) confers factor independent proliferation on the FDB cell lines (McCormack and Gonda, 2000). As was shown in primary haematopoietic cells expressing the activated mutants (I374N, FIΔ and V449E), the results in the FDB cell lines also suggest that there are signalling differences between these two classes of hβc mutants.

Due to the extensive characterisation of the three mutants I374N, FIΔ, and V449E, and the differences between the extracellular and transmembrane mutants revealed from *in vitro* and in particular *in vivo* studies, they were chosen for further investigation in this thesis.

Figure 5: Using FDB-1 cells as a model of myeloid proliferation and differentiation. Cytospin profiles on day 5 of parental FDB-1 cells +GM-CSF or IL-3 and FDB-1 V449E and FDB-1 FIΔ cells in the absence of added growth factors. The factor dependent growth of parental FDB-1 cells is shown in the presence of IL-3 leading to continual proliferation of the cells and in the presence of GM-CSF where these parental FDB-1 cells differentiate along the neutrophil and monocyte lineages. Expression of FLAG-tagged hβc mutants leads to factor independent proliferation of the V449E expressing FDB-1 cells or factor independent differentiation of the FIΔ expressing FDB-1 cells to granulocytes and macrophages. Thus the V449E mediated proliferation is similar to IL-3 stimulated proliferation of the parental FDB-1 cells and the FIΔ mediated differentiation is comparable to the GM-CSF stimulated differentiation of the parental FDB-1 cells.

Adapted from Brown *et al*., (2004).

1.5.3.3 Signalling and biochemical properties of the hßc activated mutants

Consistent with their differing requirements for the mGMRα and their different capacity to induce FDB differentiation and proliferation described above, the mutants also show differences in their downstream signalling properties (see Table 5). A number of other activating point mutations have been isolated from a saturating screen of the hβc (Jenkins *et al.,* 1998). Comparisons between the various examples of activating mutations from the two classes of mutants (extracellular and transmembrane) enabled the generality of their signalling properties to be established. While it was found that the ERK1/2 MAP kinases JAK2 and STAT signalling molecules were constitutively activated by all mutants, only some of the mutants were constitutively tyrosine phosphorylated in the absence of growth factor (Jenkins *et al.,* 1998). Specifically, extracellular mutants (such as the I374N mutant), in the absence of hGMRα and hGM-CSF, were found to lack tyrosine phosphorylation of hβc in contrast to the transmembrane mutants (e.g. the V449E mutant) that showed phosphorylation of hβc (Jenkins *et al.,* 1998). As stated above and reported previously by Jenkins *et al*., (1995), not all mutants are capable of delivering a proliferative signal in other cell lines in the absence of ligand and other receptor components. Jenkins *et al*., (1998) demonstrated that two transmembrane domain mutants (V449E and A459D) identified in their screen where the only mutants capable of factor independent growth in BAF-B03 cells, in contrast to all the other extracellular, transmembrane and cytoplasmic domain mutants tested. The factor independent growth of both V449E and A459D is likely to be a consequence of their distinct signalling properties (constitutive hβc phosphorylation) in the BAF-B03 cell line and because they do not require GMRα to form a functional signalling complex (Jenkins *et al.,* 1998; D'Andrea and Gonda, 2000).

NOTE: This table is included on page 37 of the print copy of the thesis held in the University of Adelaide Library.

Table 5: Summary of the biological and biochemical properties conferred by constitutively active hβc mutants. Mutations of hβc found to confer constitutive activation of the receptor in FDC-P1 cells have their signalling properties defined in terms of their ability to proliferate in BAF-B03 cells, tyrosine phosphorylation status of hβc and activation of MAP kinase and STAT5. The location of the activating mutations within the CRD4 has been defined to the level of the seven β strands (designated A-G) within the CRD. TM = Transmembrane, EC = extracellular.

Adapted from D'Andrea, *et al*., (2000).

1.5.3.4 Leukaemogenic potential of activated hβc mutants– in vivo models

Given the signalling and biochemical differences between the mutants in primary and transformed murine haematopoietic cells, the next approach was to examine whether the function of the transmembrane and extracellular hβc mutants differed *in vivo*. To investigate the *in vivo* effects of the mutants two approaches have been used by others in the laboratory.

(a) Transgenic mice

Transgenic mice were generated in which expression of FIΔ complimentary DNA (cDNA) was driven by a promoter derived from the widely-expressed *phosphoglycerate kinase gene* (*PGK-1*), and expression was demonstrated in several tissues including haematopoietic cells (D'Andrea *et al.,* 1998). Transgenic mice displayed a myeloproliferative disorder characterized by splenomegaly, erythrocytosis, and granulocytic and megakaryocytic hyperplasia that resembled the human disease polycythemia vera. In addition the mice developed a sporadic, progressive neurological disease that was associated with an accumulation of macrophages in the brain stem. The results indicate that chronic low level expression of the extracellular FIΔ mutant *in vivo* causes a myeloproliferative disorder.

(b) Mouse bone marrow reconstitution

Another approach used to investigate the effects of the mutants *in vivo* involved the retroviral infection of mice donor bone marrow cells with mutant hβc cDNAs and their injection into lethally irradiated recipient mice (McCormack and Gonda, 1999). The mice carrying the extracellular FIΔ and I374N mutants suffered from a chronic multilineage myeloproliferative disorder characterized by increases in the neutrophil, monocyte, erythrocyte and platelet lineages (McCormack and Gonda, 1999). However, the V449E mutant caused an acute disease that was haematologically similar to leukaemias such as acute myeloid leukaemia, but did not exhibit factor independent erythropoiesis as seen with the extracellular mutants (McCormack and Gonda, 1999). Due to the occurrence of the disease in only some of the V449E expressing mice and the long latency of the disease, it was suggested that further oncogenic events are required for, in addition to, the V449E signal (McCormack and Gonda, 1999).

The experiments described above raise the possibility that mutations in the extracellular part of the hβc are associated with chronic myeloproliferative disorders, while mutations in the transmembrane region may result in an acute leukaemia. Therefore, the distinct signalling properties of the two classes of hβc mutants are likely to result in gene expression differences which mediate their differing biological activities.

1.6 Large scale gene expression profiling techniques

Changes in gene expression underlie most fundamental biological changes and thus differential gene expression patterns can be extremely informative for understanding a range of biological and pathological phenomena. Methods available for detecting and quantifying global gene expression levels include differential display, subtractive hybridisation, sequencing of cDNA libraries, serial analysis of gene expression (SAGE) and gene arrays (Duggan *et al.*, 1999). Gene arrays have the capability to examine the simultaneous changes in expression of thousands of genes in a given cell type or sample under different conditions.

Gene arrays have been used for various applications ranging from examining gene expression pattern in yeast under different growth/stress conditions (Spellman *et al.*, 1998; Natarajan *et al.*, 2001; Yale and Bohnert, 2001) to cancer profiling of tumours (Alizadeh *et al.*, 2000; Bittner *et al.*, 2000). They now have widespread application in basic scientific research as well as pharmaceutical and clinical research (Dudoit *et al.*, 2000; Jafari and Azuaje, 2006). The power of this technology has been increased greatly by the rapid accumulation and annotation of sequencing data from the genome and cDNA clones of humans as well as model organisms such as yeast (*Saccharomyces cerevisiae*), worms (*Caenorhabditis elegans*), flies (*Drosophila melanogaster*) and mice (*Mus musculus*).

1.6.1 Overview of membrane arrays and microarrays

Gene arrays rely on the principle of sequence complementarity exhibited by the two strands of the DNA duplex (i.e. the selective binding of nucleic acids) (Southern *et al.*, 1999; Dudoit *et al.*, 2000). Gene arrays containing hundreds to thousands of

immobilized DNA samples are hybridised with a probe[1] sample in a manner similar to blotting methods such as dot blotting and Southern blotting (Cheung *et al.*, 1999; Southern *et al.*, 1999). Gene arrays come in two basic types – membrane arrays printed onto nylon or nitrocellulose filters and microarrays using a solid glass matrix (Duggan *et al.*, 1999). These substrates can be printed with either cDNA or oligonucleotides (Duggan *et al.*, 1999; Southern *et al.*, 1999). Oligonucleotides can also be synthesized *in situ* on to the solid supports as exemplified by the GeneChip® arrays produced by Affymetrix (Lipshutz *et al.*, 1999).

Solid substrates, such as glass, have many of the same advantages as porous membranes in their ability to bind nucleic acids (Cheung *et al.*, 1999). In addition, glass has a number of specific benefits. The first is a consequence of enhanced kinetics as the probe can bind to target nucleic acids molecules without diffusing into the pores found in porous membranes. This enhances the rate of hybridisation and reduces background levels (Southern *et al.*, 1999; Holloway *et al.*, 2002). The second is that the flatness, transparency and rigidity of glass improves image acquisition and analysis (discussed below) due to better definition of the location of the immobilised targets on a solid support compared to a flexible membrane (Southern *et al.*, 1999; Holloway *et al.*, 2002). The third, and one of the most important advantages, is the ability to simultaneously analyse two differently labelled probes on the one slide. In contrast, membrane arrays are limited to serial or parallel hybridisations (Cheung *et al.*, 1999; Holloway *et al.*, 2002). Dual hybridisation (i.e. two probes hybridised to the same glass support) is achieved through the use of cDNA probes labelled with two different fluorophores, such as the cyanine dyes, Cy3 and Cy5 (Schena *et al.*, 1996). Membrane arrays typically use radioactive detection methods (e.g. ^{33}P generally used in preference to more energetic emitters) (Duggan *et al.*, 1999) and require replicate arrays to be performed in order to compare gene expression between samples (Holloway *et al.*, 2002). Using replicate arrays introduces an additional source of variability associated with the membrane array format.

The comparative nature of gene expression measurements performed with either membrane arrays or two-colour competitively hybridised microarrays generates quantitative but relative levels of gene expression. On the other hand, Affymetrix GeneChips provide absolute levels of gene expression as a consequence of the fact that

[1] Note that the terminology used in this thesis is such that the labelled sample will be termed the 'probe' while the immobilised DNA sequences located on the arrays will be termed the 'target'.

each sample is hybridised to a separate GeneChip array, and the use of mismatch and perfect match probes on the arrays (Lipshutz *et al.*, 1999; Hill *et al.*, 2001; Holloway *et al.*, 2002).

Membrane arrays and microarrays will be the focus of the following sections as they were the techniques utilised in this thesis. Affymetrix GeneChip oligonucleotide arrays were not pursued due to their higher costs compared to printed cDNA arrays. Furthermore, the decision to use microarrays was the result of this new technology undergoing development and setup at the Adelaide Microarray Facility.

1.6.2 Overview of data acquisition and image analysis

There are a number of commercially available software programs to perform the data extraction and image analysis to determine the intensity of the spots on the array (for review see Web Tables K and L from Holloway *et al.*, 2002). Some of these software programs, such as GenePix (Axon Instruments) and Quantarray (Packard Bioscience) can perform the entire image analysis – from the scanning of the images to the intensity extraction and data analysis (discussed below) and are supplied with commercial scanners (Holloway *et al.*, 2002). Other software programs are focused on just the data analysis, such as Genespring (Silicon Genetics), and require extraction of the data from the images by another application dedicated for this task e.g. Spot (CSIRO) (Holloway *et al*., 2002). The process of image analysis to extract the data from an array experiment consists of a number of steps which are discussed below.

Image analysis is a critical aspect of microarray experiments as it can have a large effect on subsequent analysis including the identification of differentially expressed genes (Yang *et al.*, 2001a; Yang *et al.*, 2001b). The first step of this process involves scanning the arrays – either directly using a confocal laser scanner (microarrays) or indirectly via a phosphorimager screen and phosphorimager (membrane arrays) – to generate an image of the array. These images are then analysed using image analysis software to measure the relative amounts of probe that have hybridised to the individual elements on the array. The relative amounts of probe are proportion to the amount of fluorescence or radioactivity detected, which is in turn related to the pixel intensity measured for each spot on the array (Cheung *et al.*, 1999).

1.6.2.1 Data acquistion

In the case of microarray slides, the data extraction process begins with the hybridised slides being imaged using a microarray scanner that produces two images representing the hybridisation of the green (G) Cy3 and red (R) Cy5 labelled probes. These two images are usually combined for visualisation purposes into a single image. At the time of scanning, the operator can make adjustments for any overall brightness differences between the red and green probes so that majority of the genes, which are not differentially expressed, appear yellow. (This relies on an assumption that the majority of the genes investigated on the array do not change i.e. they are not differentially expressed). This assumption is fundamental to the normalisation steps discussed in the proceeding section. The operator minimises these differences between the red and green channels by adjusting either the laser power or the PMT (photo-multiplier tube) voltage for each channel (Cy3 and Cy5) thereby reducing dye biases. The two channels are generally scanned at high intensity so that the fainter spots are within the usable range (to increase sensitivity) while not overexposing the brighter spots beyond the maximum detection limits of the PMT (reducing specificity) (Yang *et al.*, 2001a; Lyng *et al.*, 2004). This process achieves a normalisation step by adjusting either the laser power or photomultiplier voltage for each channel so that the measurements in each channel occur over the same dynamic range (Duggan *et al.*, 1999). Keeping the PMT gain constant and adjusting the voltage settings, or keeping the voltage constant and adjusting the PMT gain, has been reported to give similar results (Dudley *et al.*, 2002) especially when the scanners are used within their linear range (Lyng *et al.*, 2004). A report by Lyng *et al.*, (2004) recommends that two scans of each channel be performed, the first to obtain the lowest intensities in the useable range and a second scan adjusted so that the brightest spots are just below the level of saturation. This is said to increase the accuracy of the data and the number of recorded genes within each experiment (Lyng *et al.*, 2004).

1.6.2.2 Processing the scanned image

The processing stages that follow the scanning of the arrays can be divided into three stages: addressing or gridding, segmentation and intensity extraction (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Addressing is the process of defining the coordinates of each spot or target on the array. This is achieved by knowing the structure or layout of the spots on the array, but is affected by the separation between rows and columns of spots within grids and between the grids themselves, individual variation in each spot's

location and overall position of the array in the image (Yang *et al.*, 2001a; Yang *et al.*, 2001b). While complete automation of this process is preferred, user intervention is useful to enhance the reliability of the process (Yang *et al.*, 2001a; Yang *et al.*, 2001b). As mentioned above there are differences between membrane arrays and microarrays when it comes to image analysis. One of the problems with membrane arrays is that they frequently show non-linear warping of the matrix during the imaging stage after hybridisation and washing (Duggan *et al.*, 1999; Yang *et al.*, 2001b). This results in the membrane array no longer having the strict geometric regularity of a rigid glass microarray, making it difficult to specify exact target locations (Duggan *et al.*, 1999).

Segmentation is the classification of pixels as either contributing to foreground (i.e. the spot) or to background, so that intensities can be calculated (Yang *et al.*, 2001a; Yang *et al.*, 2001b). The segmentation methods used for microarray images can be summarised into four groups according to the method they use; fixed circle segmentation, adaptive circle segmentation, adaptive shape segmentation and histogram segmentation (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Fixed circle segmentation uses a circle of fixed diameter for all the spots on the array to define the foreground and is used in GenePix and Quantarray programs, whereas adaptive circle segmentation changes the diameter of the circle for each spot on the array and is utilised in GenePix software (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Adaptive shape segmentation used in Spot can generate non-circular shaped regions to define the spots on the array (Yang *et al.*, 2001a; Yang *et al.*, 2001b). This method is useful for printed arrays from non commercial sources as these frequently exhibit oval or doughnut shaped spots (Eisen and Brown, 1999). Histogram segmentation is used in the Quantarray software and uses a target mask larger than any spot on the array and determines foreground and background intensity values from the histogram of pixel values for pixels in the mask area (Yang *et al.*, 2001a; Yang *et al.*, 2001b).

Finally, the intensity extraction step calculates the intensity of all the pixels within the foreground and background of a spot in terms of red and green intensity pairs (R,G) (Yang *et al.*, 2001a; Yang *et al.*, 2001b). This involves determining the spot intensity and background intensity for each target on the array. Typically a spot's foreground intensity is calculated as the mean intensity of the pixel values measured within the segmented spot mask (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Alternatively, median values can be used, which are seen to be more robust than mean values (Yang *et al.*, 2001b). An intensity contribution attributed to the background is then calculated in

the belief that a spot's intensity is made up of contribution due to non specific binding (Yang *et al.*, 2001a; Yang *et al.*, 2001b). However, there is no consensus on the best approach to background subtraction (Holloway *et al.*, 2002). There are four background adjustment methods used; local background, constant background, morphological opening and the histogram technique describe above (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Local background adjustment methods use the region around the foreground spot mask to calculate an estimate of the background intensity – typically the median of pixel values in this region, whereas constant background adjustment subtracts a fixed background intensity from all spots on the array (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Morphological opening is applied in Spot and generates an estimate of background intensities across the entire array, which enables the estimation of background intensity to be performed within the foreground spot mask (Yang *et al.*, 2001a; Yang *et al.*, 2001b).

The goal of image analysis techniques described above is to generate reliable background adjusted foreground intensity values for both red and green channels to produce an intensity ratio (e.g. red over green) for each spot on the array. This is the starting point for subsequent statistical analysis to determine differentially expressed genes.

1.6.3 Overview of statistical methods used in the analysis of microarray experiments

There are a number of statistical methods that are involved in analysing data obtained from microarray experiments in order to identify candidate differentially expressed genes. Statistical methods correct for systematic sources of variation and determine statistically significant differentially expressed genes. Methods that correct for non biological sources of variation and that enable differentially expressed genes to be determined are discussed below.

1.6.3.1 Handling variation in a microarray experiments

In a microarray experiment, there are numerous sources of variation (Churchill, 2002; Holloway *et al.*, 2002; Yang *et al.*, 2002). The factors that cause these variations include dye biases from the efficiency of dye incorporation, experimental variability in the hybridisation and processing steps, scanner introduced biases, and differences between replicate slides that lead to print-tip or spatial effects (Holloway *et al.*, 2002; Yang *et al.*, 2002). It is essential in the first step (before statistical analysis can proceed) to

identify and remove systematic sources of variation so that any biological differences can be revealed (Yang *et al.*, 2002). This is achieved via an appropriate normalisation method which aims to remove non-biological sources of variation (see Dudoit *et al.*, 2000; Quackenbush, 2002). Such normalisation methods include global normalisation, which corrects for effects across an entire array by adjusting the red and green channels for each spot on the array by a constant factor, so that the overall mean ratio of red to green is equal to one (Quackenbush, 2002). Other normalisation methods include linear regression and rank invariant methods proposed by Tseng *et al.*, (2001). The problem with these normalisation methods is that they are based on assumptions that are too simplistic to account for the multiple sources of error typically encountered in an array experiment (Dudoit *et al.*, 2000; Yang *et al.*, 2002). Locally weighted linear regression (lowess) normalisation has been proposed as a method to remove intensity dependent effects in the red/green ratio values (Dudoit *et al.*, 2000; Quackenbush, 2002; Yang *et al.*, 2002). Without the removal of these systematic effects there is no statistical control and hence no basis for examining differential gene expression (Dudoit *et al.*, 2000).

1.6.3.2 Determination of differential gene expression

It is clear that with massive amounts of data and the multiple sources of variability involved in a microarray experiment, appropriate statistical analysis is critical to a successful experimental outcome. However, the development of statistical methods for microarray data analysis is an ongoing process in order to find more reliable answers for more difficult questions (Slonim, 2002). The analysis of the membrane and microarray experiments described in this thesis sought to answer a simple question; in the two systems under investigation, which genes were differential expressed?

The early methods for determining differentially expressed genes measured the ratio of expression levels between two samples. Genes with values above a fixed cut-off (e.g. greater than a 2 fold increase/decrease) were considered differentially expressed (Dudoit *et al.*, 2000; Slonim, 2002). Cut-offs using simple expression fold changes, despite their use in the literature, (Schena *et al.*, 1995; DeRisi *et al.*, 1996; Schena *et al.*, 1996; Wodicka *et al.*, 1997; Chu *et al.*, 1998) are not a reliable measure of significant expression changes as they rely on arbitrary cut-off values for which there is no consideration given to the variability in the data set (Quackenbush, 2001). The problem with this procedure is that it will overlook highly reproducible genes with a small fold change, below the cut-off, in favour of large fold changes, irrespective of their statistical significance obtained through replicate experiments. When using replicate experiments

with fold changes as the selection criteria, both false positives and negatives can be wrongly identified as differentially expressed (see Tanaka *et al*., 2000). Recent methods have used probability models with different distributional assumptions to the intensity pairs (e.g. R,G) to create rules for deciding differential expression (Dudoit *et al.*, 2000) (see Sections 1.6.3.3 and 1.6.3.4). Commonly the ratio of red over green (R/G) is used to estimate expression level, but other methods also include the product of red and green (RG) which reflects overall transcript abundance (Dudoit *et al.*, 2000).

There are two broad types of microarray experiments. First, single-slide cDNA microarray experiments comparing two mRNA samples on the one slide. Second, multiple-slide experiments where two or more mRNA samples are hybridised to multiple slides. For the analysis of membrane arrays, the images from two filters can be considered as similar to the red and green channels of a dual labelled microarray hybridisation. When the two filters are analysed together they can be treated as a 'single slide' that compares both samples. This makes the analysis of membrane arrays amenable to statistical methods designed for single-slide microarray analysis.

1.6.3.3 Single-slide methods of statistical analysis

Single-slide methods have been proposed by a number of authors (Chen *et al.*, 1997; Roberts *et al.*, 2000; Sapir and Churchill, 2000; Newton *et al.*, 2001). The methods proposed by Chen *et al*., (1997), Sapir and Churchill, (2000) and Newton *et al*., (2001) will be described as they were utilised in the membrane array analysis presented in Chapter 3. These three methods derive a rule for deciding whether a gene is differentially expressed by using different distributional assumptions about the intensity pairs (R,G) (Dudoit *et al.*, 2000). The method of Chen *et al*., (1997) examines the gene expression ratio (R/G) of the individual genes to determine significance based on the assumption of a normal distribution (Dudoit *et al.*, 2000). Using an iterative procedure Chen *et al*., (1997) calculate a constant representing the variation present over the entire data set (Quackenbush, 2001; Yang *et al.*, 2002). This measurement of the variation present in the data is used to produce absolute cut-offs in the gene expression ratios that assign significance to the observed gene expression values (Quackenbush, 2001; Smyth *et al.*, 2003). The method of Sapir and Churchill, (2000) uses linear regression of the intensity signals (logR versus logG) to identify residuals (Yang *et al.*, 2002). These residuals are modelled with an algorithm that applies Bayes' rule (see Moore and McCabe, 1998) to calculate the variation present in the data and to separate differentially expressed genes (Sapir and Churchill, 2000). After background correction,

the residues in this method are essentially normalized log ratios (Dudoit *et al.*, 2000; Yang *et al.*, 2002). A limitation with these two methods is that they do not consider the product RG, which reflects overall transcript abundance (Dudoit *et al.*, 2000). In an effort to address this problem, Newton *et al*., (2001) use a hierarchical model based on a gamma distribution (Gamma-Gamma-Bernouli) (Dudoit *et al.*, 2000) The method of Newton *et al*., (2001) takes into account both the R/G ratio and the product RG to identify differentially expressed genes. The key difference between the method proposed by Newton *et al*., (2001) and that of Chen *et al*., (1997) and Sapir and Churchill, (2000) is that incorporated within the measurement error model is a model explicitly reflecting differences in the true level of gene expression across the genes on the array (Dudoit *et al.*, 2000). The limitation with single slide analysis methods is that with imprecise (i.e. variable) gene expression data, the identification of differentially expressed genes without replication may not be possible via any single-slide method (Dudoit *et al.*, 2000). Ideally a good experimental design would consist of biological replicates to ascertain the underlying biological variability in the experimental system, and technical replication to obtain increased precision and confidence in the results (Churchill, 2002).

1.6.3.4 Multiple-slide methods of statistical analysis

The analysis of most microarray experiments require multiple-slide methods for the examination of differentially expressed genes due to the comparison of two or more samples on multiple slides and the desire to more accurately determine biological variability in order to determine significance. The statistical analysis of microarray experiments (including the single slide methods discussed above) is achieved by using variants of common statistical tests (Slonim, 2002). The use of these statistical tests involves two parts. The first part involves calculating a test statistic and the second part involves determining the significance of the test statistic (i.e. assigning P-values) (Slonim, 2002). The purpose of the test statistic is to rank the genes in order of their evidence for differential expression, since only a limited number of genes can be followed up in confirmatory studies (Smyth and Speed, 2003; Smyth *et al.*, 2003). Slonim (2002) state that the t-test is a standard statistical test for "detecting significant change between repeated measurements of a variable in two groups" (p.502). The t-test can be generalised via the analysis of variation (ANOVA) F-statistic and numerous other variations of the t-test statistic for microarray analysis have been defined (Golub *et al.*, 1999; Model *et al.*, 2001; Tusher *et al.*, 2001; Slonim, 2002). In analysing the

data from a microarray experiment presented in Chapter 4, three statistics were used, the t-test statistic, the B-statistic (Lonnstedt and Speed, 2002) and the F-statistic (Smyth, 2004). Note that like the t-test statistic, the F-statistic comes in many forms and as such the ANOVA F-statistic is different to the F-statistic discussed below (Moore and McCabe, 1998).

(a) T-test statistic, t-stat

To determine differentially expressed genes a t-test can be applied using the null hypothesis (H_0) and the alternate hypothesis (H_a). The null hypothesis is a statement of no effect or no change and defines that the mean (μ) of the sample expression levels is equal to zero (H_0: $\mu=0$), while the alternative hypothesis states that the mean is not equal to zero (H_a: $\mu\neq0$) (Moore and McCabe, 1998). Genes are ranked using a t-test statistic (t-stat) where the t-stat equals the difference in logarithm base 2 ($\log_2$) ratios (M) divided by the standard deviation (s)

i.e. $t = \frac{M}{s} = \frac{\log_2(\frac{R}{G})}{s}$ (Smyth *et al.*, 2003). A large t-stat suggests that a given gene has different levels of expression in the two samples under investigation (Dudoit *et al.*, 2000; Yang *et al.*, 2002). The statistical significance of the results is determined by calculating P-values, which state the probability that the null hypothesis (H_0) is true (Moore and McCabe, 1998). Smaller P-values indicate stronger evidence against the null hypothesis (Moore and McCabe, 1998) and thus stronger evidence of a gene being differentially expressed.

The problem with the standard t-stat is that it is a poor estimator of differentially expressed genes, as many non differentially expressed genes have small standard deviation and thus large t-stats, which is used as the defining criteria for differential expression (Yang *et al.*, 2002). This leads to the presence of a large number of false positive genes being identified as differentially expressed. Given this problem, attempts have been made to moderate the t-stat (Efron *et al.*, 2001; Tusher *et al.*, 2001; Lonnstedt and Speed, 2002), typically by adjusting the standard deviation (Smyth *et al.*, 2003).

Another problem when using the t-stat for microarray data is due to the large amount of multiple hypothesis testing that is performed during the statistical analysis of a microarray experiment (Caudill and Hill, 1995; Shaffer, 1995). The probability of differential expression (P-values) can be modified to adjust for the large level of multiple hypothesis testing that is performed when analysing microarray data (Yang *et*

al., 2002). However, the standard Bonferroni correction method (that multiplies the uncorrected P-value by the number of genes tested) (see Shaffer, 1995) is overly restrictive in the microarray setting as few genes are likely to be identified with adjusted P-values that are significant, following the use of this correction method (Dudoit *et al.*, 2000; Slonim, 2002). An alternative approach is to control the family-wise error rate (which is the overall likelihood of falsely selecting at least one gene incorrectly as a differentially expressed gene) (Shaffer, 1995; Slonim, 2002). Smyth *et al.*, (2003) state that this approach is also overly stringent because in a microarray analysis context falsely selecting a few genes as differentially expressed amongst a majority of correctly chosen genes will not be a serious problem. A less stringent and more powerful method is to control the false discovery rate (FDR) defined by Benjamini and Hochberg (1995) (Smyth *et al.*, 2003). It is important to correct for the FDR because when examining large microarrays, containing ten thousand genes, a 1% FDR will lead to 100 genes being classed as differentially expressed, when the observed expression differences can be explained by chance variation. Having ranked the genes by the t-stat, the adjusted P-values are then used to choose a cut-off value in the probabilities obtained for the genes listed as differentially expressed (Smyth *et al.*, 2003).

(b) Bayesian estimates, B-statistic

An alternative to the t-stat approach is the application of Bayes theorem (see Moore and McCabe, 1998) using an empirical Bayes approach (Efron *et al.*, 2001; Efron, 2003) performed by Lönnstedt and Speed (2002). The parametric empirical Bayes approach used by Lönnstedt and Speed, (2002) to determine differentially expressed genes defines a B-statistic, or B-stat, which is an estimator of the posterior log-odds of differential expression, where values greater than zero correspond to a greater than 50-50 chance that a given gene is differentially expressed (Efron, 2003; Smyth *et al.*, 2003).

$$B = \log\left(\frac{\text{Probability gene is differentially expressed}}{\text{Probability gene is not differentially expressed}}\right)$$

i.e. a 50:50 ratio is = $\log \frac{1}{1} = 0$.

For the purpose of ranking genes, the B-stat is equivalent to the penalized t-statistic

$t = \frac{\overline{M}}{\sqrt{(a + s^2)/n}}$ where $\overline{M}$ is the mean of the M-values (the log intensity ratio) for any particular gene across replicate arrays, "s" is the standard deviation of the M-values across arrays, "n" is the number of measurements or arrays and the penalty "a" is estimated from the mean and standard deviation of the sample variances "s^2" (Smyth *et al.*, 2003). The B-stat enables differentially expressed genes to be ranked according to their observed statistical significance of being differentially expressed (Smyth *et al.*, 2003).

(c) The F-statistic, F-stat

The approach of Lönnstedt and Speed (2002) to define the B-stat was based on a single sample replicated experiment. Smyth (2004) formalised the calculation of an adjusted t-stat by using an empirical Bayes approach (Efron *et al.*, 2001; Efron, 2003) so that this approach could be used more widely for the analysis of microarray experiments. The original model of Lönnstedt and Speed (2002) involved comparisons between two samples. This model was extended and established in the context of a general linear model (where several parameters are estimated) by Smyth (2004). The extended approach by Smyth (2004) allows the use of F statistics (F-stats) to simultaneously study multiple parameters of interest, in comparison to the standard t-test statistic which can only be applied to a single parameter (Moore and McCabe, 1998; Smyth, 2004). Use of the F-stat allows statistically relevant cut-offs to be assigned to the data and hence produce lists of genes of interest based on their differential expression.

(d) Additional multiple slide methods for use in large microarray studies

Once a set of differential expressed genes has been obtained, there are other statistical tools that can be used for pattern discovery and class prediction in large microarray data sets. Pattern discovery uses dimension-reduction methods including principal components analysis and multidimensional scaling to allow the visualisation of large data sets (Slonim, 2002). Other methods include clustering (or cluster analysis), are design to find groups within a set of data, which can be used to classify objects into groups for diagnostic purposes (Slonim, 2002). Clustering approaches are particularly useful for analysing the results of several microarray experiments in order to identify genes with similar patterns of expression, or to provide clues to the function of an unknown gene due to it clustering within a group of genes with a known gene function (Quackenbush, 2001; Slonim, 2002). Cluster analysis methods have been applied to the

task of cancer cell profiling and this work is typified by Golub *et al*., (1999), who were able to demonstrate the feasibility of using gene expression profiling to classify cancers between AML and ALL subtypes and predict unknown cancers into one of these two subtypes. Other multiple-slide clustering methods include hierarchical clustering and self-organizing maps to group genes with similar expression profiles such as those used by Eisen *et al*., (1998), Alizadeh *et al*., (2000), Ross *et al*., (2000) and Tamayo *et al*., (1999). In these four studies the clustering methods were used for: (1) gaining gene functions for unknown genes, (2) the molecular classification of tumours based on gene expression patterns, (3) finding features of gene expression patterns in common between tumours and cell lines, and (4) to reveal novel hypothesises about haematopoietic differentiation used in treatment of APL. These four examples indicate the diverse utility of clustering approaches in microarray analysis.

1.7 Project rationale

As described above (Section 1.2) there are many genes that when disrupted by mutation or translocation can lead to leukaemia. The underlying central theme is that such mutations lead to activation of key pathways which maintain a proliferative or survival advantage on the cells, and prevent their terminal differentiation (Gilliland *et al.*, 2004). Hence, models that allow the dissection of events that regulate proliferation and differentiation are extremely useful for identification of key genes involved in haematopoiesis and in leukaemogenesis. This in turn may reveal genes and gene products, which could serve as targets for therapy in the future.

Activating mutations of hβc differentially activate signalling in myeloid cell lines such as FDC-P1 and FDB-1. Biochemical studies have established signalling differences between the extracellular and transmembrane mutants in FDC-P1 cells (Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000) and FDB-1 cells (Brown *et al.*, 2004). The FDB-1 cell line model in particular, allows the investigation of genes involved in the processes of proliferation, survival as well as myeloid differentiation (McCormack and Gonda, 2000). Activating mutations in either the transmembrane or extracellular region of the hβc receptor, when introduced into FDB-1 cells, result in the alternative phenotypes (McCormack and Gonda, 2000). The findings from McCormack and Gonda, (2000) and Brown *et al*., (2004) have demonstrated that the extracellular and

transmembrane mutations have selective effects on signalling pathways which modulate proliferation, survival and differentiation. The differential activation of these pathways leads to changes in expression of downstream genes, which in turn regulate differentiation or proliferation. *In vivo* experiments are consistent with a set of genes induced specifically by V449E that results in uncontrolled proliferation and/or blocked differentiation that contributes to leukaemia (McCormack and Gonda, 1999). Alternatively, there may be genes repressed by V449E, or activated by I374N and FIΔ signalling, that have a tumour suppressor function and/or promote differentiation.

It is proposed that analysis of the downstream gene expression changes in response to the leukaemogenic (V449E) and non-leukaemogenic (I374N and FIΔ) hβc mutants in the myeloid cell lines, FDC-P1 and FDB-1, will reveal key genes associated with myeloid differentiation, proliferation, survival and the leukaemic induction by the V449E mutant.

1.7.1 Project aims

The overall aim of the work described in this thesis was to identify genes whose expression was changed in response to the different mutant receptors. The genes that were of particular interest, were those that may be involved in differentiation in response to I374N or FIΔ, or which may be involved in proliferation, survival or blocking differentiation in the case of V449E. Membrane arrays and glass slide microarrays were used to profile gene expression in FDC-P1 and FDB-1 cells and to monitor changes associated with expression of I374N, FIΔ and V449E in those cells.

Consequently, the specific aims of this research were;

1. To identify genes that are differentially expressed between the I374N and V449E hβc mutants in FDC-P1 cells.
2. To identify genes that are differentially expressed between the FIΔ and V449E hβc mutants in FDB-1 cells.
3. To conduct molecular and functional characterisation of differentially regulated genes identified from aims 1 and 2.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals, reagents and consumables

All chemicals and reagents used were of analytical grade and purchased from Ajax Chemicals (Auburn, NSW, Australia), BDH Laboratory Supplies (Poole, UK) and Sigma Chemical Company (St Louis, Mom, USA).

2.1.2 Solutions and Buffers – commonly used

MQ water	Milli-Q system (Millipore Australia, North Ryde, NSW, Australia) purified water.
DEPC water	Milli-Q water treated with 0.1% (v/v) of diethyl pyrocarbonate (DEPC) (Sigma) overnight before autoclaving.
80% glycerol	80% v/v of glycerol in Milli-Q water.
Phosphate buffered saline (PBS)	130 mM sodium chloride, 10 mM di-sodium hydrogen orthophosphate, 10 mM sodium di-hydrogen orthophosphate, pH 7.2.
1× SSC	150 mM sodium chloride, 15 mM tri-sodium citrate, pH 7.4.
1× TAE buffer	40 mM Tris-acetate pH 8.2, 1 mM EDTA (ethylenediaminetetraacetic acid).
1×TBE buffer	90 mM Tris-borate pH 8.2, 2mM EDTA.
FACS fix	1% (v/v) formaldehyde, 2% (w/v) glucose, and 0.02% (v/v) sodium azide in PBS.

10× Northern Buffer (MOPS buffer)

42 g MOPS (3-(N-morpholino)propanesuphonic acid), 20mL of 0.5 M EDTA, 16.6 mL of 3M sodium acetate, pH adjusted to 7 by adding approximately 6 g of sodium hydroxide and adjusting final volume up to 1 L with MQ water. 1 mL of DEPC is added to the solution and treated overnight before autoclaving.

50× Denhardt's solution

1 g Bovine Serum Albumin (BSA), 1 g Ficoll and 1 g of Polyvinylpyrrolidone in 100 mL of water.

TE pH 7.5 — 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0.

TELT buffer — 50 mM Tris-HCl pH 7.5, 62.5 mM EDTA pH 8.0, 0.4% Triton X-100, 2.5 M lithium chloride.

2.1.2.1 DNA and RNA Loading Buffers

6× Loading Buffer Type III

0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 30% (v/v) glycerol.

5× RNA loading buffer

0.5% (w/v) SDS, 25% (v/v) glycerol, 25 mM EDTA and 0.25% (w/v) bromophenol blue.

10× Ficoll loading buffer

0.25% (w/v) bromophenol blue, 25% (w/v) Ficoll.

Formamide Loading Solution

95% Formamide, 20 mM EDTA, 0.05% (w/v) bromophenol blue, and 0.05% (w/v) xylene cyanol.

RNA sample buffer (500 µL)

50 µL of 10× Northern Buffer, 250 µL of formamide, 89 µL of formaldehyde, 10 µL of 10 mg/mL stock ethidium bromide and 101 µL of MQ water.

2.1.2.2 Bacterial culture media

Luria-Bertaini (LB) broth

1% (w/v) Bacto-tryptone, 1% (w/v) Bacto-yeast extract and 1% (w/v) NaCl in Milli-Q water, pH adjusted to 7.5 with sodium hydroxide and sterilised by autoclaving.

LB-agar — LB broth containing 1.5% (w/v) Bacto-agar (Difco Laboratories, Detroit, MI, USA) (e.g. 6 g agar in 400 mL of LB broth).

2.1.3 Restriction endonucleases

Restriction endonucleases and appropriate dilution buffers used in this study were obtained from New England Biolabs (Beverley, MA, USA), and Promega (Madison, WI, USA). These included *Eco*R I (with a restriction endonuclease site of GAATTC), *Hin*d III (AAGCTT), *Kpn* I (GGTACC), *Mlu* I (ACGCGT), *Not* I (GCGGCCGC), *Sac* I (GAGCTC), *Sal* I (GTCGAC) and *Xho* I (CTCGAG).

2.1.4 Growth factors

Recombinant murine IL-3 (mIL-3) and GM-CSF (mGM-CSF) were produced in house from baculovirus vectors supplied by Dr Andrew Hapel (John Curtin School of Medical Research, Canberra, ACT, Australia).

Recombinant human GM-CSF (hGM-CSF) produced in *Eschericia coli* (*E.coli*) and purified by anion exchange and reverse phase high-performance liquid chromatography was a gift from Professor Angel Lopez (Hanson Centre for Cancer Research, Adelaide, SA, Australia).

2.1.5 Radiochemicals (Radionucleotides)

(α-^{33}P)dCTP was purchased from Geneworks (Adelaide, SA, Australia). (α-^{32}P)dATP was purchased from either Geneworks or Perkin Elmer Life Sciences (Rowville, VIC, Australia).

2.1.6 Bacteria

For the propagation of plasmids, the *E.coli* strains DH5α, DH10β (Invitrogen, Carlsbad, CA, USA) and JM109 (Promega) were used and were grown in LB broth. For the selection of bacteria harbouring plasmids containing ampicillin, chloramphenicol or

kanamycin resistance genes, the LB broth was supplemented with ampicillin (100 μg/mL), chloramphenicol (50 μg/mL) or kanamycin (50 μg/mL). Long term storage of *E.coli* was performed by the addition of 600 μL of 80% glycerol solution to 600 μL of fresh overnight culture, vortexing thoroughly, and storage at -80°C.

For the selection of individual colonies, LB-agar plates were used and were supplemented with antibiotics (100 μg/mL of ampicillin or 50 μg/mL chloramphenicol) when antibiotic selection was required.

2.1.6.1 Preparation and transformation of competent Escherichia coli (E. coli)

A single colony of JM109 strain *E.coli* was used to prepare chemically competent bacteria by inoculating 2 mL of Ψ broth (2% w/v Bacto-tryptone (Difco), 0.5% w/v Bacto-yeast extract (Difco), 20 mM $MgSO_4 \cdot 7H_2O$, 10 mM NaCl, pH 7.6). The inoculated broth was incubated at 37°C for 16 hours with shaking. 330 μL of this culture was subcultured into 10 mL of Ψ broth and incubated for approximately 90 minutes at 37°C until the OD at 600 nm reached 0.6, as determined using a DU-64 spectrophotometer (Beckman Coulter, Bucks, UK). Bacteria were chilled on ice for 5 minutes, centrifuged at 4000 rpm, at 4°C for 5 minutes and the bacterial cell pellet was resuspended in 40 mL of Tfb 1 buffer (30 mM KAc, 100 mM RbCl, 10 mM $CaCl_2 \cdot 2H_2O$, 50 mM $MnCl_2 \cdot 4H_2O$, 15% glycerol, pH 5.8). This step was repeated before the bacterial cell pellets were resuspended in 4 mL Tfb 2 buffer (10mM MOPS, 10 mM RbCl, 75 mM $CaCl_2 \cdot 2H_2O$, 15% glycerol, pH 6.5) and incubated on ice for a further 15 minutes. Chemically competent bacteria were frozen in liquid nitrogen in either 125 or 250 μL aliquots and stored at –80°C until required.

2.1.7 Tissue culture solutions and media

Tissue culture media where indicated was supplemented with antibiotics, Penicillin G (5000 U/mL) and Streptomycin Sulphate (5000 mcg/mL) (Pen/Strep), and L-glutamine (200 mM) from JRH Biosciences (Victoria, Australia).

Sterile phosphate buffered saline (PBS) used for tissue culture applications was purchased from Gibco BRL and Trypsin was supplied from JRH Biosciences.

Foetal Calf Serum (FCS) from CSL (Victoria, Australia) was heat inactivated by treatment at 56°C for 1 hour, prior to storage at -20°C.

Dulbecco's modified Eagle's medium (DMEM) supplied by Gibco BRL (Victoria, Australia) was supplemented with 2 mM L-glutamine, 1:100 dilution of

Pen/Strep and 10 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid). Complete DMEM (cDMEM) was prepared by adding 10% FCS to the above supplemented DMEM.

Iscove's modified Dulbecco's medium (IMDM) was purchased from Gibco BRL and supplemented with 2 mM L-glutamine and 1:100 dilution of Pen/Strep. Complete IMDM (cIMDM) was prepared by adding 10% FCS to the above supplemented IMDM. Puromycin selection was achieved by adding puromycin (Calbiochem. San Diego, CA, USA) to cIMDM to a final concentration of 1 µg/mL.

2.1.8 Molecular weight standards

1 kb Plus DNA ladder, purchased from Invitrogen, was supplied at 1 µg/µL in 10 mM Tris-HCl (pH 7.5), 50 mM sodium chloride and 1 mM EDTA. This was used at final concentration of 150 ng/µL in 1× Loading Buffer Type III. Fragment sizes of the ladder were 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1.65, 1, 0.85, 0.65, 0.5, 0.4, 0.3, 0.2, and 0.1 kb.

1 kb DNA ladder, purchased from Invitrogen, was supplied at 1 µg/µL in 10 mM Tris-HCl (pH 7.5), 50 mM sodium chloride and 0.1 mM EDTA. This was used at final concentration of 150 ng/µL in 1× Loading Buffer Type III. Fragment sizes of the ladder were 12.2, 11.2, 10.2, 9.2, 8.1, 7.1, 6.1, 5.1, 4.1, 3.1, 2.0, 1.64, 1.02, 0.517, 0.506, 0.396, 0.344, 0.298, 0.220, 201, 0.154, 0.134 and 0.075 kb.

EcoR1-digested bacteriophage SPP-1 DNA markers (high range), purchased from Geneworks (Adelaide, SA, Australia), were supplied at 500 ng/µL in 10 mM Tris, 1 mM EDTA, pH 7.6. These markers were used at a final concentration of 100 ng/µL in 1× Loading Buffer Type III using 2.5 µL (250 ng) per lane. It contained 16 fragments of the following sizes (kb): 8.56, 7.43, 6.11, 4.90, 3.64, 2.80, 1.95, 1.86, 1.52, 1.41, 1.16, 0.99, 0.71, 0.49 and 0.36 and 0.08.

0.25-9.5 Kb RNA ladder purchased from GibcoBRL® Life Technologies, was supplied at 1 µg/µL in 10 mM HEPES (pH 7.2), 2 mM EDTA, and contained the following sized fragments (kb): 9.49, 7.46, 4.40, 2.37, 1.36 and 0.24.

2.1.9 Vectors and cDNAs

The hβc mutants, FIΔ (D'Andrea et al., 1994), V449E and I374N (Jenkins et al., 1995) have been described previously.

Clones from the Research Genetics membrane arrays were purchased from the I.M.A.G.E. Consortium through Invitrogen, and cDNA clones from the NIA 15K microarray were obtained from the Australian Genome Research Foundation (AGRF).

The Gateway™ pENTR11 (Invitrogen) vector was kindly supplied by Dr Steven Woods (Child Health Research Institute (CHRI), North Adelaide, SA, Australia).

The pEQEco (referred to as pEQ in this thesis) ecotropic packaging vector (Persons *et al.*, 1998) and the MSCV-IRES-GFP vector (Van Parijs *et al.*, 1999) were both kind gifts from Dr Elio Vanin (St Jude Children's Research Hospital, Memphis, TN, USA).

The pVPackVSV-G and pVPackGP retroviral infection plasmids were generously donated by T. Gonda (Centre for Immunology and Cancer Research, University of Queensland, Queensland, Australia).

2.2 Methods for manipulating DNA

2.2.1 Small scale plasmid preparations

Plasmid minipreps from bacterial cultures were prepared by alkaline lysis method according to Sambrook *et al.*, (1989), using an UltraClean™ 6 minute Mini Plasmid Prep Kit (Mo Bio Laboratories, Carlsbad, CA,. USA) following the manufacturer's instructions, or via the TELT method (Medina-Acosta and Cross, 1993). For all methods, a sterile toothpick was used to inoculate 2 mL of LB-broth (containing the appropriate antibiotic for selection of the plasmid) with a single fresh colony of *E.coli*. The culture was incubated for 16 hours at 37°C with shaking. The TELT method was performed by chilling 1.5 mL of culture on ice and centrifuging at 13000 rpm for 1 minute. The supernatant was removed and the pellet resuspended in 250 µL TELT buffer. Bacteria were lysed with the addition of 20 µL lysozyme (20 mg/mL) and incubated on ice for 10 minutes. Lysates were heated at 95°C for 2 minutes, cooled on ice for 10 minutes and centrifuged at 13000 rpm for 20 minutes. The bacterial pellet was removed with a toothpick, and DNA precipitated by the addition of 500 µL ethanol. The precipitate was pelleted by centrifugation at 13000 rpm for 10 minutes at 4°C. Pellets were washed with 1 mL 70% ethanol and air dried at 55°C for 10 minutes. DNA was dissolved in 32 µL of MQ water.

2.2.2 Large scale plasmid preparations

100 mL of LB, containing the appropriate antibiotic for plasmid selection, was inoculated with a fresh culture of *E.coli* bacteria containing the plasmid of interest and grown overnight at 37°C with shaking. .Midi scale plasmid preparations were performed on the 100 mL of overnight culture using a HiSpeed™ Midi Plasmid Purification Kit (QIAGEN, Chatsworth, CA, USA) according to the manufacturer's protocol.

2.2.3 Agarose gel electrophoresis

Electrophoresis was performed using agarose gels composed of 0.8 - 2% (w/v) agarose dissolved in 1× TAE. The gels were submerged in a tank (Bio-Rad, Hercules, CA, USA) containing 1× TAE buffer and were run at up to 100 V. Nucleic acids were visualised by staining gels in a 10 µg/mL ethidium bromide solution for 5 minutes, rinsed with water and exposed to ultraviolet (UV) light. Gels were photographed using Polaroid positive land film, type 667 (Polaroid, Hertfordshire, England) or an ImagemasterTM Gel Documentation System (Amersham Pharmacia Biotech, NJ, USA).

2.2.4 Restriction endonuclease digestion

DNA was digested with the desired restriction endonuclease at a concentration of 1-5 units of enzyme per µg of DNA and incubated at 37°C for at least one hour (up to a maximum of three hours). The digestion was performed in a total volume of between 10-45 µL which contained 1× digestion buffer (specific for the enzyme used), 0.1 mg/mL of BSA (when appropriate) and sterile MQ water. Reactions were terminated by either 65°C heat inactivation for 20 minutes, or by the addition of 6× Loading Buffer Type III (to a final concentration of 1×) and analysis by gel electrophoresis.

2.2.5 Preparation of inserts and cloning vectors

DNA was digested with the appropriate restriction endonuclease and separated by agarose gel electrophoresis. Under low power UV light, to minimise UV damage to the DNA, the required DNA fragment was quickly excised from the gel with a sterile scalpel. DNA was purified using the UltraClean™ DNA extraction kit (Mo Bio Laboratories Inc.) following the manufacturer's instructions. The yield of the DNA was

calculated by comparing the density of the DNA band with the density of the bands due to the closest molecular weight standards run on the same agarose gel.

2.2.5.1 Phosphatase treatment of cloning vectors

To minimise re-ligation of vectors in the absence of an insert, vectors digested with a single restriction endonuclease (e.g. EcoR1) were treated with calf intestinal alkaline phosphatase (CIP) (Promega), by the addition of 2 µL of a 1:10 dilution of CIP (in 1× enzyme digestion buffer and MQ water) at 37°C for 1 hour. The enzyme was heat inactivated by treatment at 55°C for 5 minutes and vectors were then gel purified.

2.2.6 Ligation of DNA into plasmid vectors

Ligation reactions were performed using an insert DNA:vector DNA molecule ratio of 3:1. Reactions were carried out in a 10 µL volume containing insert DNA, vector DNA, 1× ligation buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 10 mM DTT, 50 µg/mL of BSA, 1 mM ATP) (Promega, WI, USA), 1 unit of T4 DNA ligase (New England Biolabs, MA, USA) and sterile MQ water. Reactions were incubated for 16 hours at 4°C for optimum yield, or 4 hours at room temperature for transformation the same day.

2.2.7 Bacterial transformation

2.2.7.1 Heat shock method

A 125 µL aliquot of JM109 chemically competent bacteria was thawed on ice and 50 - 100 µL of the bacterial cells were incubated for 25 minutes with 5-10 µL of DNA ligation mix. Bacteria were heat shocked for 1 minute at 42°C, cooled on ice for 2 minutes and incubated for 45 minutes at 37°C with SOC buffer (2% Bacto-tryptone (w/v), 0.5% Bacto-yeast extract (w/v), 10 mM NaCl, 2.5 mM KCl, 20 mM MgCl2, 100 mM D-glucose). Bacteria were centrifuged at 5000 rpm for 1 minute and the pellet was resuspended in 100µl SOC buffer. The resuspended bacteria were spread onto LB-agar plates containing an appropriate antibiotic for selection of the plasmid and incubated overnight at 37°C.

2.2.8 RNase A treatment and phenol chloroform extraction of plasmid mini-preparations for sequencing

Plasmid mini-preparations that failed to sequence were treated with RNase A to remove contaminating RNA and purified via a phenol chloroform extraction. 35 µL of the plasmid mini-preparations were mixed with an equal volume of 0.2 mg/mL RNase A

and incubated at 37°C for 1 hour. RNase A treated samples were mixed with 25 µL of MQ water, 5 µL of 4 M sodium chloride and 100 µL of phenol:chloroform:isoamylalcohol 25:24:1 (Life Technologies, MD, USA), and centrifuged at 13000 rpm for 2 minutes. 90 µL of the aqueous phase was pipetted into a new eppendorf and mixed with 250 µL of 100% ethanol. The DNA was precipitated by centrifugation at 14000 rpm for 30 minutes at 4°C. Pellets were washed twice with 500 µL of 75% ethanol, dried briefly at 65°C and resuspended in 22 µL of MQ water.

2.2.9 DNA sequencing

For DNA sequencing, ABI Prism® Dye BigDye™ Terminator Cycle Sequencing Ready Reaction Kit versions 2 and 3 (Applied Biosystems) and Ampli*Taq* DNA polymerase FS (Perkin Elmer) were used. A sequencing reaction contained 500 ng of plasmid DNA, 100 ng of primer and 4 µL of Terminator Ready Reaction Mix in a 20 µL volume. Reactions were performed in a thermal cycler (Perkin Elmer) using 25 cycles of 96°C for 30 seconds, 50°C for 15 seconds and 60°C for 4 minutes. This was followed by a 4°C hold step. DNA reaction products were precipitated by the addition of 80 µL of 75% isopropanol, 1 µL of 20mg/mL glycogen and a 30 minute incubation at room temperature. Samples were centrifuged at 14000 rpm for 20 minutes at 4°C, and the supernantant was discarded. The pellets were washed with 70% ethanol and air dried using a heater block at 55°C. Gel separation of sequencing reaction products were performed by the Institute for Medical and Veterinary Science Sequencing Facility (Adelaide, SA) using an ABI Prism 377 DNA Sequencer (Applied Biosystems, CA, USA) or a Perkin Elmer automated sequencer.

Alternatively, plasmids (1.25 µg/reaction) and sequencing primers (2 µL per reaction at 3.2 *p*mole/µL) were sent to the Flinders Medical Centre (FMC) Sequencing Core Facility (Bedford Park, SA) where the entire sequencing reactions, purification and separation of the products were performed.

2.2.10 Polymerase Chain Reaction (PCR)

2.2.10.1 Standard PCR conditions:

PCRs were performed using 1 µL of cDNA (or plasmid) template, 17 µL of sterile MQ water, 2.5 µL of 10× PCR buffer, 1.5 µL of magnesium chloride (25 mM), 1 µL of dNTPs (10mM), 100 ng of forward primer, 100 ng of reverse primer and 0.2 µL of

AmpliTaq® DNA polymerase or AmpliTaq® Gold™ DNA polymerase (Applied Biosystems) in a total volume of 25 µL. Reactions were denatured for 30 seconds at 94°C (or 5 minutes at 95°C with AmpliTaq Gold DNA polymerase), then subjected to 30 cycles of 95°C for 30 seconds, 50-60°C for 20 seconds, and 72°C for 1 minute. This was followed by a final 10 minute extension step at 72°C and a 4°C hold step. Reactions were run in either a MiniCycler™ from MJ Research (supplied by Geneworks), a Perkin Elmer DNA Thermal Cycler or a Px2 ThermoHybaid (Thermo Electron Corporation,Waltham, MA, USA).

2.2.10.2 cDNA synthesis

1 µg of total RNA prepared from TRIzol treated cell lines (e.g. FDB-1) was used as a template on which cDNA was synthesised (see RT-PCR section below).

2.2.10.3 Cloning

Platinum Pfx DNA polymerase (Invitrogen) was used for high fidelity cloning PCRs. DNA was amplified in a 25 µL reaction containing 1 µL of cDNA template, 1× amplification buffer, 1 mM magnesium sulphate, 100 ng of each forward and reverse primer, 0.3 mM dNTP, 1 unit Platinum Pfx DNA polymerase and MQ water. Reactions were performed with a 2 minute 94°C denaturation step, followed by 35 cycles of 15 seconds denaturation at 94°C, 30 seconds primer annealing at 50-60°C (depending on primer composition) and 1-5 minutes of primer extension (based on at least 1 minute per kb of DNA to be amplified) at 68°C. The final extension step of 10 minutes at 68°C was followed by a 4°C hold step.

2.2.10.4 Primers

Oligonucleotide primers were synthesised by Geneworks (Adelaide, South Australia) and were resuspended to a final concentration of 100 ng/µL in MQ water. What follows is a list of primers used in this study. The four genes whose entire open reading frames were amplified for cloning have their primers shown with restriction endonuclease sites (GAATTC, EcoR1 and CTCGAG, Xho1) in bold and the start (ATG) and stop codons underlined.

GAPDH 5' 5’-ACCACAGTCCATGCCATCAC-3’

GAPDH 3' 5’-TCCACCACCCTGTTGCTGTA-3’

⇨ 452 bp product from human or mouse cDNA

HprtRNAR 5’-GTCAAGGGCATATCCAACAACAAAC-3’

HprtRNAF 5’-CCTGCTGGATTACATTAAAGCACTG-3’

⇨ 352 bp product from mouse cDNA

P1 STRAIT 5’-AAGTAGTCGCCAGAGAAGG-3’

P2 STRAIT 5’-TGCCACAGGACTACTGTAC-3’

⇨ 466 bp product from mouse cDNA

5' DIPP 5’-ATGAAGTTCAAGCCCAACCA-3’

3' DIPP 5’-CTATCTTATACTGGATGGCA-3’

⇨ 540 bp product from mouse cDNA

5' MKP 5’-ATGACGCGCCGGCTGGTA-3’

3' MKP 5’-TTAATTCTTTGCTGCCCTTGC-3’

⇨ 420 bp product from mouse cDNA

5' Nedd5 5’-ATGTCTAAGCAACCAACTCA-3’

3' Nedd5 5’-TCACACATGCTGCCCGAG-3’

⇨ 1086 bp product from mouse cDNA

5' NAB1 5’-ATGGCCACAGCCTTACCT-3’

P2 NAB1 5’-GTAGGTAACTTCTCGAGAAA-3’

⇨ 927 bp product from mouse cDNA

5' orf DEAD 5’-AGGCGCTGTGGAGTTATTCTT-3’

3' orf DEAD 5’-GATCCGGTCCTGTTTTCTCA-3’

⇨ 638 bp product from mouse cDNA

P1 Rap2A 5’-CCTCTGGCTGGCTGTGTTT-3’

P2 Rap2A 5’-GGCCATGCTGACCAATAAGT-3’

⇨ 545 bp product from mouse cDNA

5' orf Hox#2 5’-CCAATCCGCCGCATCCCC-3’

3' orf Hox#2 5’-CCCCCGCCCCCAACAGC-3’

⇨ 805 bp product from mouse cDNA

5' orf RAB1 5’-GCTAAGAACGCAACGAATGT-3’

3' est RAB1 5’-TACCAGGAAGAACTAGCAAATCAG-3’

⇨ 620 bp product from mouse cDNA

P1 AI449120 5’-TCCCAAACCCACAGAGC-3’

P2 AI449120 5’-ATCCTAGGGTAGTCACAGC-3’

⇨ 384 bp product from mouse cDNA (a.k.a. Ube2i – see Chapter 3)

5'Btg1-EcoR1 5’-TA**GAATTC**GCCACCATGCATCCCTTCTACACC-3’

3'Btg1-Xho1 5’-TACCG**CTCGAG**TTAACCTGATACAGTCATCATATTG-3’

⇨ 541 bp product from mouse cDNA – BTG1

3'Btg1HA-Xho1 5’-TACCG**CTCGAG**CTAAGCGTAATCTGGAACATCGTATGGGTAACCTGATACAGTCATCATATTG-3’

⇨ 568 bp product from mouse cDNA – BTG1-HA

5'Cd24-EcoR1 5’-TA**GAATTC**GCGGACATGGGCAGAGCG-3’

3'Cd24-Xho1 5’-TACCG**CTCGAG**CTAACAGTAGAGATGTAGAAGAG-3’

⇨ 256 bp product from mouse cDNA – CD24

3'Cd24HA-Xho 5’-TACCG**CTCGAG**CTAAGCGTAATCTGGAACATCGTATGGGTAACAGTAGAGATGTAGAAGAG-3’

⇨ 283 bp product from mouse cDNA – CD24-HA

5'Copeb-EcoR 5’-TA**GAATTC**CCCGACATGGATGTGCTCC-3’

3'Copeb-EcoR 5’-TACCG**CTCGAG**TCAGAGGTGCCTCTTCATGTG-3’

Note: 3' primer was mislabelled when ordered– should have been “…–Xho1”

⇨ 880 bp product from mouse cDNA – Copeb

3CopebHA-Eco 5’-TACCG**CTCGAG**CTAAGCGTAATCTGGAACATCGTATGGGTAGAGGTGCCTCTTCATGTG-3’

Note: 3' primer was mislabelled when ordered– should have been “…–Xho”

⇨ '907 bp product from mouse cDNA – Copeb-HA

5'Ns-EcoR1 5’-TA**GAATTC**GGCCAGGATGAAGAGGCC-3’

3'Ns-Xho1 5’-TACCG**CTCGAG**TTATATAATCTGTGGTGAAGTC-3’

⇨ 1643 bp product from mouse cDNA – Ns

3'Ns-HA-Xho1 5’-TACCG**CTCGAG**CTAAGCGTAATCTGGAACATCGTAT GGGTATATATAATCTGTGGTGAAGTC-3’

⇨ 1670 bp product from mouse cDNA – Ns-HA

MSCV Forward 5’-GAACCTCCTCGTTCGAC-3’

MSCV Reverse 5’-GCCTTATTCCAAGCGGC-3’

2.3 Methods for manipulating RNA

2.3.1 Total RNA isolation

Total RNA extractions were prepared using TRIzol reagent (Invitrogen) according to the manufacturer’s recommendation and with RNase free reagents and solutions. Cells were pelleted by centrifugation at 1500 rpm for 5 minutes. The pellet was resuspended in TRIzol at a concentration of 10^7 cells/mL and surplus cells were stored at -70°C for processing at a later time. Eppendorf tubes containing 1 mL aliquots of resuspended cells were incubated at room temperature for 5 minutes and 0.2 mL of chloroform was added to each sample. Samples were mixed vigorously by shaking for 15 seconds and incubated at room temperature for at least 2 minutes before centrifugation at 13000 rpm for 15 minutes. 400 µL of the aqueous (upper) phase was transferred to a fresh eppendorf tube, combined with 500 µL of isopropanol, mixed by inversion and incubated for 10 minutes at room temperature. Total RNA was pelleted by centrifugation at 13000 rpm for 10 minutes at 4°C. Pellets were rinsed with 1 mL of 75% ethanol solution, centrifuged at 8000 rpm for 5 minutes at 4°C and dried on a 37°C dry heat block. The RNA was resuspended in 100 µL of DEPC treated MQ water. A sample of each RNA preparation was diluted 1:100 and examined via spectrometry at wavelengths of 260 nm and 280 nm to determine concentration and purity respectively. Total RNA samples were examined via agarose gel electrophoresis to ensure the RNA was not degraded. RNA samples were mixed in RNA loading buffer and run on either 1% TAE agarose gels (see Section 2.2.3) or formaldehyde-agarose gels run in 1× Northern Buffer (see Section 2.3.5.1).

For the membrane array experiment, total RNA samples obtained from FDC-P1 cells were purified by a sodium acetate and ethanol precipitation. RNA samples were precipitated by the addition of $1/10^{th}$ volume of 3 M sodium acetate and 3 volumes of 100% ethanol. Samples were centrifuged at 13000 rpm for 20 minutes at 4°C and the

supernatant discarded. Pellets were rinsed in 75% ethanol and centrifuged again at 13000 rpm for 15 minutes at 4°C. RNA pellets were dried on a 37°C dry heat block and resuspended in 200 µL of DEPC treated MQ water.

2.3.2 Poly A^+ RNA preparation from FDC-P1 cells for Northern Blotting

Frozen cell pellets were resuspended at a concentration of 10^7 cells/mL in STE buffer (0.1 M sodium chloride, 10 mM Tris pH 7.4, 1 mM EDTA). Proteinase K (Sigma) and SDS were added to the solution at a final concentration of 200 µg/mL and 0.5% respectively. Samples were then homogenized and incubated at 37°C for 90 minutes.

After washing the oligo(dT)-cellulose (Roche) in 0.5 M sodium hydroxide for 5 minutes, followed by washing several times in water, the oligo(dT)-cellulose was resuspended in Binding Buffer (0.5 M sodium chloride, 10 mM Tris pH 7.4, 1 mM EDTA, 0.1% SDS). RNA samples were mixed by inversion with the oligo(dT)-cellulose suspension for 2.5 hours at room temperature on a rotating wheel using 0.5 mL of oligo(dT)-cellulose (packed column volume) per sample (approximately 2×10^8 cells). Samples were centrifuged at 2000 rpm for 5 minutes, supernatant discarded and the oligo(dT)-cellulose resuspended in 2 mL Binding Buffer. The resuspended oligo(dT)-cellulose was added to polypropylene Econo Columns (Bio-Rad, Hercules, CA, USA), which had been pre-treated with 0.5 M sodium hydroxide and rinsed in MQ water before use.. The RNA solution was allowed to run through the columns by gravity flow, then the column was washed twice with 3 mL of Binding Buffer, followed by 3 mL of Wash Buffer (0.1 M sodium chloride, 10 mM Tris pH 7.4, 1 mM EDTA, 0.1% SDS). The column bound RNA was then eluted with 2 mL of Elution Buffer (10 mM Tris pH 7.4, 1 mM EDTA, 0.1% SDS) into a RNase free 10 mL polypropylene tube. Samples were analyzed using a GeneQuant spectrophotometer (Amersham Pharmacia Biotech, Uppsala, Sweden) to determine yield and concentration. 2 mL samples were mixed with 200 µL of 3 M sodium acetate, 4 mL of cold 100% ethanol, 40 µg glycogen and stored at -20°C. The RNA was precipitated and washed with 70% ethanol before loading RNA samples on a formaldehyde-agarose gel.

Total RNA prepared from FDC-P1 cells was also used to prepare Poly A^+ RNA for use on Northern blots. Total RNA samples (approximately 500 µg) were mixed with Binding Buffer and oligo(dT)-cellulose and processed as described above.

2.3.3 RNA preparation from FDB-1 cells for Northern Blotting and Microarrays

Poly A^+ RNA extractions from FDB-1 cells used the QuickPrep mRNA Purification Kit (Amersham Pharmacia Biotech) according to the manufacturer's instructions. Total RNA was prepared from FDB-1 cells using TRIzol and RNeasy columns as described in Section 2.6.2.

2.3.4 Reverse Transcription Polymerase Chain Reaction (RT –PCR)

Total RNA samples (1 µg) were reverse transcribed at 42°C for 1 hour using an oligo(dT) primer (Geneworks) together with the Omniscript reverse transcriptase kit (Qiagen) following the manufacturer's protocols.

Polymerase Chain Reaction (PCR) was performed to determine the presence of contaminating genomic DNA in the reverse transcribed samples and to confirm that cDNA had been synthesized. This was performed in two separate reactions. The first PCR used primers designed to amplify a 452 bp fragment of the gene *glyceraldehyde-3-phosphate dehydrogenase* (*Gapdh*). Using the reverse transcribed RNA samples and negative control samples (samples prepared without adding reverse transcriptase enzyme) as templates, contaminating genomic DNA was indicated by the amplification of a 452 bp fragment in the negative control samples. The second PCR used primers that spanned an intron of the *hypoxanthine guanine phosphoribosyl transferase 1* (*Hprt*) gene such that an appropriate sized fragment would only be generated from cDNA and not from genomic DNA. Both PCRs were performed under standard cycling conditions (see Section 2.2.10.1) with *Taq*man DNA polymerase (Roche, Mannheim, Germany) using 30 amplification cycles.

2.3.5 Northern Blotting

2.3.5.1 Formaldehyde-agarose gel electrophoresis

RNA samples were run on a denaturing agarose gel. For 100 mL gel, 1 g of agarose was dissolved in 10 mL of 10× Northern Buffer, and 72 mL of DEPC treated MQ water by heating in a microwave oven until boiling. Once cooled to below 60°C, 18 mL of formaldehyde (35-40% solution) was added and the gel poured immediately under a fume hood. Prior to pouring the gel, the gel tank, combs and gel tray were either soaked in a 0.5 M sodium hydroxide solution for at least 30 minutes and then rinsed with MQ water, or cleaned with RNaseZap (Ambion, Austin, TX, USA), washed in 0.1% SDS

and rinsed with MQ water. The gel tank was filled with 1× Northern Buffer and the solidified gel placed into the tank.

To check the integrity of the RNA, samples were mixed with an equal volume of Formamide Loading Solution, heated at 65°C for 5 minutes, cooled on ice and loaded onto the gel. The gel was run at 100 V until the bromophenol blue dye had travelled ¾ the length of the gel. The RNA was visualised as described in Section 2.2.3.

For Northern blotting, each lane was loaded with either 2 µg of Poly A^+, 20 µg of total RNA or 2 µg of MW marker (RNA ladder 0.25-9.5 kb (Invitrogen)). The RNA samples and MW marker for each lane were combined with 12 µL of RNA sample buffer and heated at 65°C for 5 minutes, placed on ice and then mixed with 3 µL of 10× Ficoll loading buffer. Samples and MW marker were loaded onto the gel. The gel was run at 100 V under a fume hood until bromophenol blue dye had travelled ¾ the length of the gel. The gel was soaked in MQ water for 30 minutes (changing water once) prior to photography under UV light or scanning on Typhoon 8600 Variable Mode Imager (Molecular Dynamics, part of Amersham Pharmacia Biotech) using the fluorescence setting.

2.3.5.2 Northern transfer

The RNA contained in the gel was transferred to Hybond-N membrane (Amersham Pharmacia Biotech) (pre-soaked in MQ) using Whatman 3MM paper and paper towelling via capillary action with 10× SSC overnight (12-24 hours). After transfer, the membrane was rinsed in 2× SSC and, while still damp, the RNA was UV crosslinked to the membrane using a GS Gene Linker (Bio-Rad) and stored dry prior to use.

2.3.5.3 Northern probing

Prehybridisation

The membrane filter was placed in a 35 × 250 mm Hybaid bottle (Thermo Hybaid Ltd., Middlesex, UK) with 15 mL of ULTRAhyb™ (Ambion) (that was prewarmed to 68°C to allow precipitated material to dissolve before use). Prehybridisation was performed at 42°C for at least 30 minutes with slow rotation of the Hybaid bottle in a Hybaid oven (Thermo Hybaid Ltd.).

Probe Synthesis

DNA probes were labelled using Strip-EZ DNA Probe Synthesis and Removal Kit (Ambion). 20-100 ng of DNA was made up to a final volume of 9 µL with MQ water and heated to 95°C for 5 minutes before being snap cooled on dry ice. Samples were

mixed with 2.5 µL 10× Decamer solution, 5 µL of 5× -dATP Reaction Buffer (-dATP and –dCTP), 2.5 µl of 10× modified dCTP, 1 µL of Exonuclease free Klenow and 5 µL of ^{32}P dATP (3000 Ci/mmol, Perkin Elmer). Reactions were allowed to proceed for 30 minutes at 37°C, and were stopped by adding 1 µL of 0.5 M EDTA. Reaction volumes were made up to a final volume of 51 µL by the addition of 25 µL of water. The reactions were loaded onto ProbeQuant G-50 Micro Columns (Amersham Pharmacia Biotech) and centrifuged at 3000 rpm for 2 minutes. 1 µL samples of each probe were taken before and after spin column purification to determine the percentage incorporation of ^{32}P into the DNA probe by measurement in a beta-counter. The purified probe was denatured (90°C for 10 minutes) and cooled on ice prior to its addition to the prehybridisation buffer (ULTRAhyb™) and the membrane filter ensuring $\geq 1\times10^6$ cpm/mL.

The first round of Northern Blots described in Chapter 4 used the Megaprime DNA labelling system (Amersham Pharmacia Biotech) following the manufacturer's protocols, with 5 µL of ^{32}P dATP (Geneworks).

Hybridisation and washing

Hybridisation of the probe to the membrane was performed overnight at 42°C in a Hybaid oven with slow rotation of the 35 × 250 mm Hybaid bottle containing 15 mL of ULTRAhyb™. Radioactive probe solutions were discarded and membranes washed three times for at least 10 minutes in 2× SSC/0.1% SDS at 42°C, then twice at higher stringency for at least 20 minutes in 2× SSC/0.1% SDS at 65°C. Note, Northern blots prepared from FDB-1 RNA were washed at 65°C using 0.2× SSC/0.1% SDS wash instead of 2× SSC/0.1% SDS for improved removal of background signal without an apparent effect on the intensity of the target bands. After the washes, membrane filters were sealed in plastic bags with 1-2 mL of 2× SSC/0.1% SDS to prevent them from drying out. Membranes were then exposed to a phosphorscreen (Molecular Dynamics) over a period of a few hours to 3 days and the screen was scanned on a phosphorimager (Typhoon 8600 Variable Mode Imager (Molecular Dynamics)). ImageQuant® version 5 (Molecular Dynamics) software was used to quantitate the hybridised probe signal on the membrane. Loading control comparisons were achieved by comparing the target gene intensity in each lane to the intensity of a house keeping gene (either *Gapdh* or Cyclophilin A (*Ppia*)) in the same lane on the same blot.

Stripping of the Strip-EZ probed membranes was performed with 20-25 mL of 1× Probe Degradation Buffer, for 2 minutes at room temperature and 10 minutes at 68°C, then 20-25 mL of 1× Reconstitution Buffer at 68°C for 10 minutes, then 0.1% SDS at 68°C for 10 minutes. Membranes were heat sealed in plastic bags with 1-2 mL of 2× SSC/0.1% SDS, and exposed to a phosphorimager screen to verify the blots had been stripped prior to their reuse.

Stripping of the membranes probed with Megaprime labelled probes was performed by the addition of boiling 0.1× SSC/0.1% SDS solution to the membranes and allowing the solution to cool to room temperature with rocking. The membrane was scanned with a Geiger-counter to determine the quantity of probe still bound. If the radioactivity was greater than 5 cps, the stripping process was repeated. Membranes were sealed in plastic bags with 1-2 mL of 2× SSC/0.1% SDS and exposed to a phosphorimager screen to verify the blots had been stripped prior to their reuse.

2.4 Mammalian cell culture

2.4.1 Cell lines and culture media

All cells used in this study were grown and maintained at 37°C in a mixture of air with 5% CO_2 in a humidified tissue culture incubator (Forma Scientific, Marietta, Ohio, USA).

The murine IL-3/GM-CSF dependent cell line, FDC-P1, was grown in complete Dulbecco's modified Eagle's medium (cDMEM) containing mGM-CSF (50 U/mL). FDC-P1 cells expressing both the human GM-CSFR alpha and beta subunits were maintained in cDMEM supplemented with hGM-CSF (2 ng/mL). FDC-P1 cells expressing hGM-CSFR alpha subunit, together with either the I374N or V449E mutant hβc, were grown in the absence of exogenously added GM-CSF (as they grow factor independently).

The murine myeloid cell line, FDB-1, was grown in complete Iscove's modified Dulbecco's medium (cIMDM) containing 500 U/mL mIL-3. FDB-1 cells containing FIΔ or V449E mutant hβc constructs were grown in cIMDM supplemented with puromycin (1 μg/mL) and mIL-3 (500 U/mL), however, these cells were assayed after washing three times in IMDM and growing them in cIMDM without mIL-3. The

presence of the mutant hβc was confirmed by antibody staining and analysis by flow cytometry (see Section 2.4.3.1).

HEK 293T cells were maintained in DMEM supplemented with 10% FCS and 1:100 dilution of Pen/Strep.

NIH 3T3 cells were maintained in DMEM supplemented with 10% FCS and 1:100 dilution of Pen/Strep.

2.4.2 Long term storage and thawing of cells

Cell lines were stored in liquid nitrogen according to standard techniques. Cells growing in logarithmic phase (determined by cell counting – see Section 2.4.5.1) were centrifuged (1500 rpm for 5 minutes at 4°C) and resuspended at ($\geq 1\times10^6$) in cold freezing medium (10% (v/v) of dimethyl sulfoxide (DMSO) in FCS). 1 mL aliquots were dispensed into sterile pre-chilled cryotubes (CELLSTAR®, Greiner Bio-One GmbH, Frickenhausen, Germany). The cryotubes were placed into a NALGENE® Cryo 1°C Freezing Container (Nalge Nunc International, Hereford, UK), an isopropanol rate cooling device, and placed at -80°C for at least 24 hours before the vials were transferred to liquid nitrogen storage tanks.

Stored cells were thawed rapidly in a 37°C waterbath and transferred to a 15 mL sterile polypropylene tube (Corning, NY, USA). The cell suspension was mixed with 9 mL of growth media added dropwise over 5 minutes whilst mixing. Cells were centrifuged (1500 rpm for 5 minutes at room temperature) and resuspended in fresh growth media.

2.4.3 Flow cytometry and antibody staining

2.4.3.1 Staining of GM-CSF receptor

Staining of GM-CSF receptor was achieved with 1C1 (anti hβc) or 4H1 (anti alpha chain) antibodies, both used at a final concentration of 20 µg/mL, and supplied by Professor Angel Lopez (Hanson Centre for Cancer Research, Adelaide, South Australia). 5×10^5 – 1×10^6 cells in 50 µL of cell culture medium were mixed with 50 µL of primary antibody diluted in PBS, or with 50 µL PBS as a negative control (i.e. no primary antibody) and incubated at room temperature for 6 minutes. Samples were washed with PBS, by the addition of 4 mL of PBS to each tube and centrifugation at 1500 rpm for 5 minutes. The supernatant was discarded and the PBS wash step was

repeated. Cell pellets were resuspended in 50 µL of secondary antibody (biotinylated anti-mouse antibody, 1:50 dilution in PBS (Vector Lab Inc., Burlingame, CA, USA)) and incubated at room temperature for 6 minutes. Tubes were washed once with PBS. Cell pellets were resuspended in 50 µL of fluorochrome conjugate (streptavidin-phycoerythrin (PE), 1:100 dilution in PBS (Caltag Laboratories, San Francisco, CA, USA)) and incubated at room temperature for 6 minutes. Samples were washed once with PBS and resuspended in cell culture media for immediate FACS analysis, or resuspended in 250-500 µL of FACS fix and stored at 4°C, protected from the light by covering the tubes in aluminium foil, for later analysis. Fluorescence was measured using an Epics-Profile II analyser (Coulter Electronics, Hialeah, FL, USA).

2.4.3.2 Cell surface antigen analysis and GFP expression analysis

5×10^5–1×10^6 cells in 50 µL of cell culture medium were mixed with 50 µL of cold primary antibody (diluted in IMDM with 10% FCS) and incubated on ice for 20 minutes. Cells were washed twice in IMDM and centrifuged at 1500 rpm for 5 minutes at 4°C. Pellets were resuspended in 100 µL of FITC (or other fluorochrome) conjugated secondary antibody (diluted in IMDM with 10% FCS) and incubated on ice for 20 minutes. Samples were washed once and then resuspended in maintenance media (e.g. cIMDM) for immediate FACS analysis/sorting or resuspended in 250-500 µL of FACS fix and stored at 4°C protected from the light by covering the tubes in aluminium foil for later analysis. Antibodies were used at previously determined optimum concentrations.

To determine mutant hβc expression (FIΔ, V449E) on the surface of the FDB-1 cells staining of the FLAG–tag epitope was performed. The primary antibody used in this case was either the anti-FLAG® M2 monoclonal antibody (Sigma, St Louis, MO) or FLYTAG ™ antibody (Amrad Biotech, Richmond, VIC, Australia). This was followed by an anti-mouse fluorescein isothiocyanate (FITC)-conjugated IgG antibody (Silenus, Hawthorn, VIC, Australia).

CD24 staining was performed using rat anti-mouse CD24 conjugated PE antibody (BD –Australia, North Ryde, NSW, Australia). For the negative control, an isotype matched control antibody rat $IgG2b_\kappa$ (Pharmingen, San Diego, CA, USA) was used together with a labelled secondary antibody.

To determine GFP expression, samples were analysed live in cell culture media (e.g. for sorting) or stored at 4°C in FACS fix protected from the light.

Analysis and/or sorting of the stained cells or GFP positive cells described above were performed by an Epics Elite ESP (Coulter Electronics, Hialeah, FL, USA) flow cytometer.

2.4.4 Retroviral infection of haematopoietic cells

FDB-1 cells infected with the mutant hβc constructs were obtained from Dr A. Brown (CHRI, North Adelaide, SA, Australia).

Transient viral producer cell lines (used to infect FDC-P1 and FDB-1 cell lines) were generated by transfecting equal amounts of pEQ ecotropic packaging vector and retroviral constructs derived from the MSCV-IRES-GFP vector. Alternatively, the pantropic vectors pVPackVSV-G and pVPackGP were transfected together with retroviral constructs derived from the MSCV-IRES-GFP vector, with all three vectors used in equal amounts. These vectors and retroviral constructs were transfected into HEK 293T cells using Lipofectamine 2000 Reagent (Invitrogen) according to the manufacturer's instructions.

2.4.4.1 Retrovirus supernatant production

Following transfection, HEK 293T cells were incubated for one day prior to the addition of fresh media for virus collection. The virus containing media was collected after 24–48 hours of retrovirus production. The supernatant was cleared of any cells or debris by centrifugation at 1500 rpm for 5 minutes and filtered through a 0.45 µm filter. Retrovirus containing supernatant was stored at 4°C for up to 1 week prior to use. The highest retroviral titres produced in the supernantant occurred at 48 hours.

2.4.4.2 Co-culture

As an alternative to producing retroviral supernatant to infect target cells, cells were incubated in direct contact with the transient retrovirus producer cell lines. Following transfection, HEK 293T cells were incubated for one day prior to the addition of FDB-1 cells (5×10^5–1×10^6 cells per T25 flask) or FDC-P1 cells (5×10^5 cells per T25 flask). Cells were grown in the specific medium of the target cells (e.g. cIMDM for FDB-1 cells) together with 4 µg/mL of polybrene. After a total of 48 hours of culture with the transient retrovirus producer cells, the target cells were harvested, centrifuged at 1500 rpm for 5 minutes and resuspended in fresh media. The target cells were collected and analysed via FACS to determine the percentage infection and sorted to obtain a purified infected population of cells based on GFP expression. Sorted GFP positive cells were

assayed for proliferation, differentiation and survival in comparison to empty vector (MSCV-IRES-GFP) infected cells.

2.4.4.3 Infection of cells with retrovirus supernatant

FDB-1 cells (5×10^4 cells in a 24 well plate) were plated in a 1:1 dilution of fresh media and media containing retrovirus, plus 4 µg/mL of polybrene, for 24 hours. Cells were then centrifuged (1500 rpm, 5 minutes) and placed in fresh growth media for 48 hours prior to analysis via FACS for GFP positive cells, as described for the co-cultured cells above.

For comparing retrovirus supernatant infection to co-culture conditions, the co-culture was performed in a 24 well plate by scaling down the usual procedure performed in a T25 flask. 5×10^4 FDB-1 cells were used in the co-culture setup and were placed in fresh growth media for 48 hours following co-culture prior to analysis via FACS, as described for the supernatant infected cells.

2.4.4.4 Titering viral supernatant

NIH 3T3 cells were used to determine the number of infectious viruses in the supernatant produced by transient retrovirus producer lines. NIH 3T3 cells were plated out 24 hours prior to infection at 1×10^4 cells per well of a 24 well plate. After removal of media, NIH 3T3 cells were grown for 24 hours in a serial dilution of media containing retroviruses in fresh media. Following this infection period, the supernatant was removed, and the NIH 3T3 cells were grown in fresh media for a further two days. Cells were removed by trypsin treatment and resuspended in FACS fix or fresh media for analysis via flow cytometry. The number of infected cells was determined by the percentage of GFP positive cells. The number of infective virus particles in the supernatant was calculated on the number of cells plated, the dilution of the retroviral supernatant, and by assuming that infection by one virus gives rise to a GFP positive cell.

An identical procedure was used when HEK 293T cells were used in parallel with NIH 3T3 cells, to compare the relative titres of the ecotropic and pantropic retrovirus containing supernatants.

2.4.5 Cell assays

2.4.5.1 Viable cell counts

Cell viability was assessed by 0.2% trypan blue exclusion and cell number determined by standard haemocytometer counting procedure.

2.4.5.2 Cell proliferation assays

Cells were washed three times in serum free media and cultured in either 500 U/mL of mIL-3 or mGM-CSF for FDB-1 cells, or 100 U/mL mIL-3 for FDC-P1 cells, or without factor. Cells were seeded at 5000 cells/well in 96 well flat-bottom plates. Growth was assessed using CellTiter 96 Aqueous One Solution Cell Proliferation Assay according to the manufacturer's instructions (Promega, Madison, WI).

2.4.5.3 Differentiation assays

FDB-1 cells were washed three times in serum free IMDM and cultured in cIMDM containing either 500 U/mL of mIL-3 or 500 U/mL of mGM-CSF, or in the absence of mIL-3 and mGM-CSF for five days. Differentiation was assessed by cytocentrifugation and Jenner-Giemsa staining. Cytospins were performed daily with 5×10^4-1×10^5 cells loaded into each chamber (minimum volume 100 µL, maximum 500 µL) and centrifuged at 5000 rpm for 5 minutes on a Cytospin 3 machine (Shardon, Cheshire, UK). Glass microscope slides (Knittel Glaser, Germany) were allowed to dry for 5 minutes before Jenner-Giemsa staining was performed using a Shandon Varistain 12 staining machine by the Diagnostic Services Laboratory of the Institute for Medical and Veterinary Science (Hanson Institute, Adelaide, SA, Australia). Slides were mounted with a drop of DePeX mounting reagent (Gurr, BDH Laboratory Supplies, Poole, UK) and a 24 mm square glass coverslip. Cells were inspected via light microscopy (Olympus model BH-2) to determine the proportion of differentiated cells.

2.5 Materials and Methods for Membrane arrays

2.5.1 Membrane arrays

The two membrane arrays used in this study were the mouse GeneFilters® (GF400a) arrays purchased from Research Genetics (Huntsville, AL, USA) that were spotted with approximately 5000 clones.

2.5.2 RNA labelling

10 µg of FDCP-1 α/I374N and α/V449E total RNA were used for each labelling reaction according to the protocols supplied by Research Genetics. Briefly, 2 µL of Oligo dT (Research Genetics) was combined with 10 µg of total RNA and heated for 10 minutes at 70°C before being chilled on ice for at least 2 minutes. This was combined with 3 µL of 10× Buffer RT (Qiagen), 1 µL of 100 mM DTT (Clontech), 1.5 µL of dNTP mixture (20mM of dATP, dGTP, dTTP), 1.5 µL of Omniscript Reverse Transcriptase (Qiagen), 3 µL of DEPC treated MQ water and 10 µL (α-^{33}P)dCTP (Geneworks) and incubated at 37°C for 90 minutes. Probes were mixed with 20 µL DEPC treated water to give a final volume of 50 µL. Unincorporated nucleotides were removed from the reactions by loading the probes onto ProbeQuant™ G-50 Micro Columns (Amersham Pharmacia Biotech) and centrifugation for 2 minutes at 3000 rpm. To confirm labelling of the probe, percentage incorporation was determined by measuring radioactivity levels before and after spin column purification using a TopCount scintillation counter (Packard Instrument Company, Meriden, CT, USA (now part of Perkin Elmer)).

2.5.3 Prehybridisation

Before first use, boiling 0.5% SDS solution was poured over the membrane arrays and left on a rocking platform for 5 minutes while the solution cooled. Microhyb solution (Research Genetics, Huntsville, AL, USA) was preheated to 65°C until contents dissolved. The two membrane arrays were added to separate 35 × 150 mm Hybaid roller bottles (Thermo Hybaid Ltd., Middlesex, UK) and then 5 mL of Microhyb solution was added to each bottle along with 5 µg of Poly dA (Research Genetics) and 5 µg of (pre-boiled for 10 minutes) mouse Cot-1 DNA (Gibco BRL Life Technologies). The prehybridisation solution was incubated with the membranes for 2.5 hours at 42°C with the bottles revolving in a Hybaid oven (Thermo Hybaid Ltd.).

2.5.4 Hybridisation

Probes were denatured for 5 minutes at 100°C and the entire probe was added to the Hybaid bottle containing the prehybridisation solution and the membrane. Membranes and radiolabelled probes were hybridised overnight at 42°C with bottles slowly rotating in a Hybaid oven.

2.5.5 Washing and scanning

Filters were washed twice in the Hybaid bottles with a solution of 2× SSC/1% SDS at 50°C (30 minutes and 25 minutes) and twice with 0.5× SSC/1% SDS at 55°C (for 20 minutes and 25 minutes). Washed filters were placed (DNA side up) on a piece of Whatman paper moistened with MQ water, wrapped in cling wrap and exposed to a phosphorimager screen (for between 3 hours to overnight). The screen was scanned using a Typhoon 8600 Variable Mode Imager (Molecular Dynamics, part of Amersham Pharmacia Biotech) at 10 micron resolution and the resulting image was visualized using ImageQuant® version 5 (Molecular Dynamics).

2.5.6 Membrane stripping

Filters were stripped after each hybridisation by washing in 0.5% SDS (heated until boiling) for 1 hour on a rocking platform at room temperature. Membranes were placed on a piece of Whatman paper moistened with MQ water, wrapped in cling wrap and exposed to a phosphorimager screen. The screen was scanned (as described above) to confirm that the intensity levels were less than that of the hybridised image.

2.6 Materials and Methods for Microarrays

2.6.1 Microarray slides

The microarray slides were arrayed with a complementary DNA (cDNA) library from the National Institute of Aging (NIA), containing over fifteen thousand clones (hereafter referred to as NIA 15K cDNA microarrays) (Tanaka *et al*., 2000). Superamine slides (Telechem, Sunnyvale, CA, USA) were contact printed in a 150 mM sodium phosphate buffer, pH 8.0, using a robotic arrayer with Telechem Stealth III pins at the Clive and Vera Ramaciotti Centre at the University of NSW.

2.6.2 RNA

Total RNA preparations, used for the microarray labelling reaction (see Section 2.6.3), were isolated using a phenol-chloroform extraction using TRIzol (Invitrogen) and column chromatography performed with an RNeasy kit (Qiagen). Cell pellets (10^7 cells) were solubilised in 1 mL TRIzol and homogenized through a 19 gauge needle before another 0.5 mL of TRIzol was added. Samples were mixed with 300 µL of chloroform and centrifuged for 15 minutes at 13000 rpm at 4 °C. 800 µL of the upper aqueous phase was transferred to a fresh eppendorf tube and mixed with 800 µL 70% ethanol

and loaded onto RNeasy midi column. A total of 4 mL of the same RNA sample was applied to each column. Columns were centrifuged at 3500 rpm for 5 minutes in a Beckman GS-6R centrifuge. The eluate was discarded and columns were reloaded with the remaining 4 mL of sample and centrifuged again. Columns were washed with 4 mL of RW1 buffer, centrifuging for 5 minutes (3500 rpm), and washed twice with 2.5 mL of RPE buffer, centrifuging for 2 minutes and 5 minutes respectively. RNA was eluted from columns using two rounds of 250 µL of RNase free water and a 3 minute 3500 rpm centrifugation, to yield approximately 500 µL of eluted RNA.

Each 500 µL RNA sample was precipitated with 50 µL of 3M sodium acetate and 1.2 mL of 100% ethanol. The precipitated RNA in 0.3 M sodium acetate and 70% ethanol solution was sent on dry ice to Melbourne, where it was stored at -80 °C prior to labelling.

2.6.3 Labelling

Labelling was performed via the indirect method of aminoallyl reverse transcription labelling at the Peter MacCallum Cancer Institute (PMCI) Microarray Core facility (Melbourne, Victoria, Australia) (http://www.ccgpm.org). For each sample, 50 µg of total RNA was used in the procedure.

The RNA samples stored at -80°C in 0.3 M sodium acetate and 70% ethanol solution were centrifuged at 14000 rpm for 30 minutes at 4°C. The RNA pellets were washed with 70% ethanol and centrifuged at 14000 rpm for 10 minutes at 4°C. After removal of the supernatant, the pellets were air dried at room temperature and resuspended in RNase free water. The RNA concentration of the samples was determined by measuring the absorption at 260 nm on a spectrophotometer by diluting 1 µL of the sample in 99 µL of 10 mM Tris pH 7.5.

The reverse transcription reaction was performed by combining 50 µg of RNA in a volume of 17.4 µL with 2 µL of Oligo dT primers (2 µg/µL) and annealed by incubation at 75°C for 5 minutes followed by cooling to 4°C. The RNA samples were mixed with 6 µL of 5× Superscript II buffer, 2 µL of 0.1 M DTT, 0.6 µL of dNTP mix (25 mM dATP, 25 mM dGTP, 25 mM dCTP, 10 mM dTTP and 15 mM aa dUTP) and 2 µL of Superscipt II (Invitrogen) reverse transcriptase enzyme and incubated at 42°C for 2.5 hours.

Hydrolysis of the samples was performed by adding 5 µL of 0.5 M EDTA and 10 µL of 0.25 M sodium hydroxide and incubation at 65°C for 15 minutes. Once cooled

to room temperature, 15 µL of 0.2 M acetic acid was added to the cDNA samples and they were stored at -20°C overnight.

The cDNA samples were thawed at 60°C for 1 minute and purified via the Qiagen QIAquick PCR purification kit (Qiagen). 300 µL of Buffer PB was added to the 60 µL cDNA samples and pipetted onto the Qiagen columns placed on a vacuum manifold. Columns were washed twice with 750 µL of Buffer PE using the vacuum manifold and then centrifuged at 13000 rpm for 1 minute. The cDNA was eluted into fresh eppendorf tubes by adding 50 µL of sterile water to the columns, incubating the samples at room temperature for 1 minute and centrifuging at 13000 rpm for 1 minute. The cDNA samples were dried at 60°C using a speedy-vac, Martin Christ RVC 2-25 (Quantum Scientific).

The coupling reaction was performed by resuspending the dried cDNA in 9 µL of water, adding 1 µL of 1 M sodium bicarbonate buffer pH 9.0, and then combining the cDNA with either the lyophilized CyScribe Cy™3 or Cy™5 reactive dyes (CyDye™ from Amersham Pharmacia Biotech). Reactions were incubated in the dark for 2 hours before a second round of Qiagen QIAquick PCR purifications. 90 µL of sterile water was added to each coupling reaction followed by 500 µL of Buffer PB and samples were added to the columns. Using a vacuum manifold, columns were washed twice with 750 µL of Buffer PE. Columns were centrifuged at 13000 rpm for 1 minute. Labelled cDNA was eluted into fresh eppendorf tubes by adding 50 µL of Buffer EB to the columns, incubating the samples at room temperature for 1 minute and centrifuging at 13000 rpm for 1 minute.

2.6.4 Hybridisation and washing

Cy3 and Cy5 labelled samples were combined in pairs and mixed with 3 µL of tRNA (4 mg/mL from Sigma), 3 µL of Cot-1 DNA (10 mg/mL from Invitrogen), 0.75 µL of Poly dA (8 mg/mL, as 50 mer oligo from Geneworks), 0.75 µL of 50× Denhart's solution containing filtered herring sperm DNA and 2.0 µL of luciferase mRNA control (Cy3 and Cy5 labelled). Samples were dried in a speedy-vac at 60°C for 55 minutes to obtain final volumes of less than 8 µL.

Probe samples were increased to a total volume of 8 µL with sterile water and combined with 8 µL of 12.5× SSC. These probes were mixed with 16 µL of 100% de-ionised formamide (filtered and stored at -20°C in single use aliquots), denatured at 100°C for 3 minutes and transferred directly to ice. Probes were combined with 0.4 µL

of 10% filtered SDS and centrifuged at 13000 rpm for 5 minutes. The entire volume (32.4 µL), avoiding any precipitant, was pipetted onto a clean 24×50 mm cover slip and applied to a NIA 15K mouse slide. Slides were placed in a humidified chamber (slide box with strips of 3MM paper moistened with 2 mL of 2× SSC) and placed in a 42°C incubator for 16 hours.

Stock preparations of the wash solutions were filtered through a 0.22 µm bottle top filter (Millipore). Washes were performed in a light proof box with shaking. Slides were immersed in wash solution 1 (0.5× SSC, 0.01% SDS) contained in a copelin jar until coverslips were dislodged, before being soaked in wash solution 1 for 2 minutes. Slides were soaked in wash solution 2 (0.5× SSC) for 3 minutes and placed in wash solution 3 (0.06× SSC) for 3 minutes. Slides were rinsed in MQ water prior to drying via centrifugation (Megafuge 1.0R - Heraeus Instruments) at 800 rpm for 5 minutes.

2.6.5 Scanning

Slides were scanned using ScanArray 5000XL (Perkin-Elmer). The initial scan was performed at 80% PMT and laser power, at a resolution of 50 µm. If white pixels were observed in this preview scan (indicating saturation) the PMT gain on Cy5 channel was lowered and the laser power on the Cy3 channel was increased until both channels had similar intensities overall. A high quality scan was performed at a resolution of 10 µm, resulting in an output file in tiff (.tif) format.

ScanAlyze was used to combine the Cy3 and Cy5 images to obtain a superimposed image (http://rana.lbl.gov/EisenSoftware.htm) in bitmap (.bmp) file format for easy viewing and printing. At PMCI, Quantarray (GSI Lumonics) and GeneSpring (Silicon Genetics/Agilent Technologies) were used to perform a preliminary analysis and quantitation of the slides.

At the Adelaide Microarray Facility (University of Adelaide, SA, Australia), slides were rescanned using a GenePix 4000B (Axon Instruments). Once again, a preview scan was performed to adjust PMT voltage (typically the 635 nm PMT voltage was 750 volts and the 532 nm PMT voltage was 550 volts) however, the laser power was not altered. A final scan was performed at 10 µm resolution. The resulting scanned slide images and data obtained from both the ScanArray machine at the PMCI and the GenePix scanner at the Adelaide Microarray Facility were used for the analysis of the experiment.

2.6.6 Gel analysis of labelled probes

Two of the 8 µL combined Cy3 and Cy5 labelled samples (prior to mixing with 12.5× SSC – see above) were analysed via agarose gel electrophoresis to confirm the quality of the labelled cDNA. The two samples were FIΔ time 0 plus V449E time 24 and V449E time 0 plus FIΔ time 24 samples. 1 µL of the combined probes from each of the two samples was mixed with 6 µL of 5× Ficoll Loading Buffer (prepared without the addition of bromophenol blue) and run on a 1% TBE gel. The gel was scanned at 532 nm (Cy3 channel) and 635 nm (Cy 5 channel) using a Molecular Imager FX Pro (Bio-Rad, Hercules, CA, USA). The gel was subsequently stained (using a 1:5000 dilution of SYBR Gold) for 20 minutes and photographed to visualise the cDNA and MW markers.

Chapter 3

Mouse GeneFilters® Analysis of Gene Expression in FDC-P1 cells Expressing Activated hβc Mutants

3.1 Introduction

The aim of the work described in this chapter was to examine the signalling differences between the extracellular and transmembrane mutants of hβc by large scale gene profiling in FDC-P1 cells. FDC-P1 cells are a murine myelomonocytic cell line that was derived from mouse bone marrow cells. The cells are dependent on exogenous growth factors (e.g. mouse IL3 or GM-CSF) for their growth and survival (Dexter *et al.*, 1980; Lang *et al.*, 1985; Hapel *et al.*, 1986). There are a number of constitutively active human βc (hβc) mutants which lead to factor independent growth when expressed in FDC-P1 cells (D'Andrea *et al.*, 1994; D'Andrea *et al.*, 1996; Jenkins *et al.*, 1998). The hβc mutants can be separated into two classes – those involving a mutation in the extracellular region of hβc, and those involving a mutation in the transmembrane region. These two classes differ in their signalling and biochemical properties. For example, β chains with an extracellular mutation require an associated alpha chain in order to provide functional signals (Jenkins *et al.*, 1999; D'Andrea and Gonda, 2000). When expressed in FDC-P1 cells, the β chain of the extracellular mutants (such as I374N) does not become tyrosine phosphorylated and neither does the adaptor protein SHC which binds to hβc (Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000, T Blake personal communication). In contrast, the transmembrane mutant V449E is constitutively tyrosine phosphorylated and does not require an alpha subunit in order to form a signalling complex (Jenkins *et al.*, 1995; Jenkins *et al.*, 1998).

In this respect V449E is similar to the transmembrane mutation (V644E) in the Neu/ErbB2 receptor tyrosine kinase (RTK) (Bargmann *et al.*, 1986) which induces constitutive receptor dimerization (Sternberg and Gullick, 1989; Weiner *et al.*, 1989; D'Andrea and Gonda, 2000). Receptor dimerization is a key step in ligand induced signalling by RTKs (Ullrich and Schlessinger, 1990) and by analogy it is suggested that the V449E mutation acts by inducing ligand-independent homodimerization of the βc subunit leading to the generation of intracellular proliferative signals (Jenkins *et al.*,

1995; Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000). That is, the signalling from the transmembrane mutant, V449E, is likely to result from homodimerization of hβc whereas the signalling from the extracelular mutant, I374N, is likely to result from α and hβc heterdimerization (for review see D'Andrea and Gonda, 2000). Given that the FDC-P1 cell line was used to initially identify the hβc mutants, and as a result of the fact that the transmembrane and extracellular mutants show signalling differences in this cell line, it follows that the FDC-P1 cell line has the potential to be used as a model for determining the results of these signalling differences between the hβc mutants. Such differences are predicted to be important in cellular contexts, such as proliferation and differentiation, and may play a role in determining the differential leukaemogenic potential of the two classes of mutants (see Section 1.5.3.4).

The primary aim of this work was to determine the consequence of the signalling differences between the extracellular mutant, I374N, and the transmembrane mutant, V449E, expressed in FDC-P1 cells, in terms of differentially expressed genes. In order to examine the gene expression differences between the FDC-P1 cells a technique previously utilised in the laboratory, called membrane arrays was used (see Bartley *et al.*, 2001). These flexible nylon membranes offer the ability to be probed several times in order to determine the relative gene expression differences between two (or more) radioactively labelled samples. Membrane arrays use small amounts of RNA for labelling and are cost effective to use, since they do not require specialised equipment to perform the experiment (Holloway *et al.*, 2002).

3.2 Results

3.2.1 Probe synthesis and hybridisation of the GeneFilters® arrays

3.2.1.1 Preparation of the FDC-P1 probes

The expression of mutant $h\beta_c$ on the cell surface of FDC-P1 cell populations (FDC-P1 V449E and FDC-P1 I374N) was confirmed by flow cytometric analysis (Figure 6), following the growth of these two cell populations in medium without added IL-3 or GM-CSF. For simplicity, FDC-P1 V449E and FDC-P1 I374N will be referred to in this chapter as V449E and I374N respectively. The factor independent growth seen by these two cell lines was not the result of autocrine growth factor production, or residual cytokine levels, as conditioned medium from the factor-independent cell pools did not

support the growth of uninfected FDC-P1 cells (Jenkins *et al.*, 1995). Cytospin preparations were prepared from both factor independent cell lines, which showed that there were no morphological differences between I374N and V449E cells expressed in FDC-P1 cells (data not shown).

Factor independent V449E and I374N cell populations were grown in log phase growth (assessed by viable cell counting) without added IL-3 and GM-CSF, prior to their collection for the preparation of total RNA (see Sections 2.4.5.1 and 2.3.1). Total RNA was prepared from both cell lines and used as a template from which cDNA was created in a reverse transcription reaction. During the reverse transcription reaction, samples were radiolabelled with ^{33}P by incorporation of alpha-labelled dCTP (see Section 2.5.2).

3.2.1.2 Hybridisation and physical properties of the GeneFilters® membrane arrays

The two radiolabelled samples representing the genes expressed in either the I374N or V449E cell populations, were hybridised to two independent mouse GeneFilters® (Research Genetics) (as described in Section 2.5) containing 5760 spots (see Figure 7A and B). These 5760 spots represented a mixture of known genes, unknown sequences called expressed sequence tags (ESTs) and control spots. Excluding the control spots on the membrane array, there was a total of 5568 ESTs and known genes on the array.

Following the parallel hybridisations described above, the two filters were then stripped and re-probed with freshly labelled RNA by using the opposite sample on each filter (Figure 7C). This generated data for each gene based on two hybridisations for each sample. These two separate sets of hybridisations were performed using the same pool of RNA and hence represent technical replicates, rather than biological replicates in this experiment.

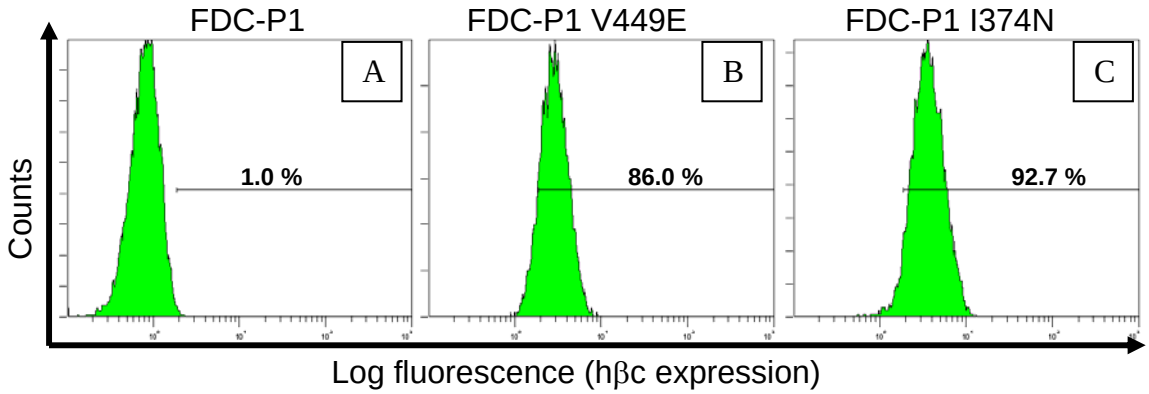
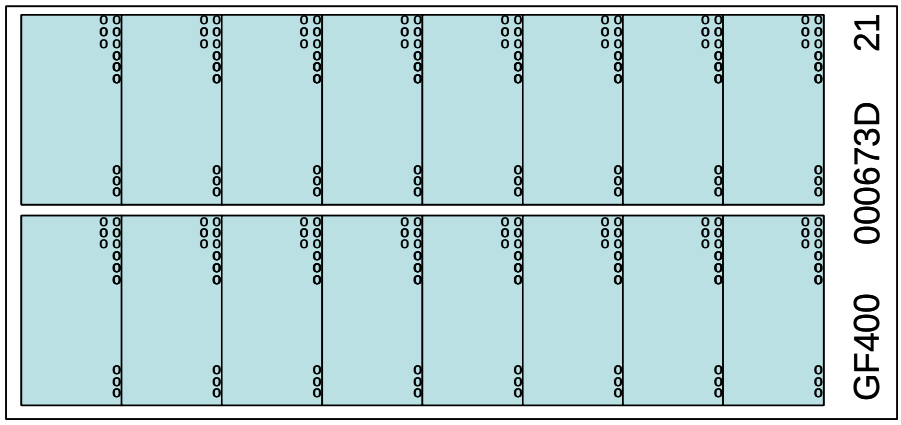


Figure 6: Cell surface expression of mutant hβc (V449E or I374N) in FDC-P1 cells. Flow cytometric analysis showing cell surface expression of the indicated hβc proteins on FDC-P1 cells expressing either the transmembrane mutant, V449E (B) or extracellular mutant, I374N (C) compared to uninfected parental cells (A). Uninfected control cells (panel A) and infected cells (panels B and C) were incubated with an anti-βc monoclonal antibody followed by an anti-mouse horse radish peroxidase secondary antibody. In each panel, the x-axis scale for the major graduations is 10^0, 10^1, 10^2 and 10^3.

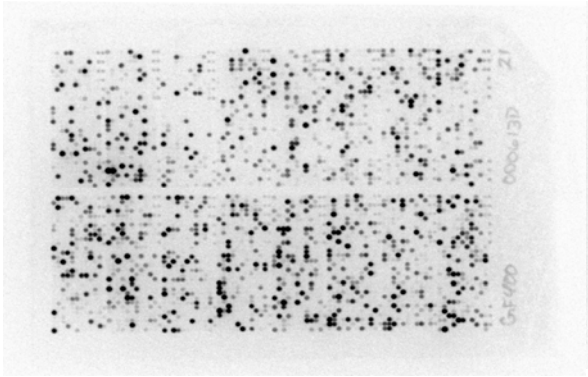
Figure 7: Layout and hybridisation of the GeneFilters® (Reseach Genetics) membrane arrays. (A) The layout of the mouse GF400A GeneFilters® membrane arrays is shown with the total genomic DNA spots indicated (circles), which act as positive controls. Each filter is numbered (e.g “F21”) and labelled with a batch number (“000673D”). (B) and (C) Hybridisation of the membrane arrays. Shown is the hybridisation of the I374N probe to the F20 membrane array and V449E probe to the F21 membrane array (B) and in the “reverse hybridisation” the F20 filter (which was originally probed with I374N cDNA) was probed with V449E cDNA and the F21 filter was probed with I374N cDNA.(C). These images were subsequently analysed to identify differentially expressed genes between these two cell populations.

A

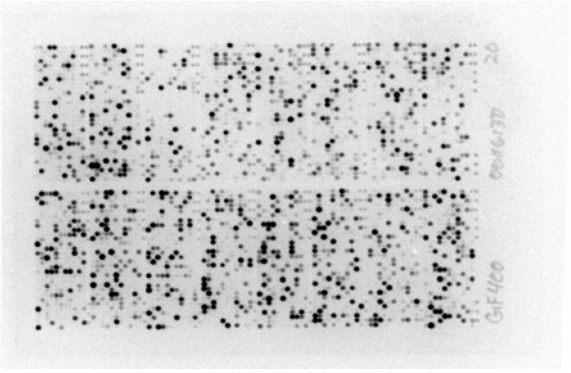


B

V449E probed filter F21

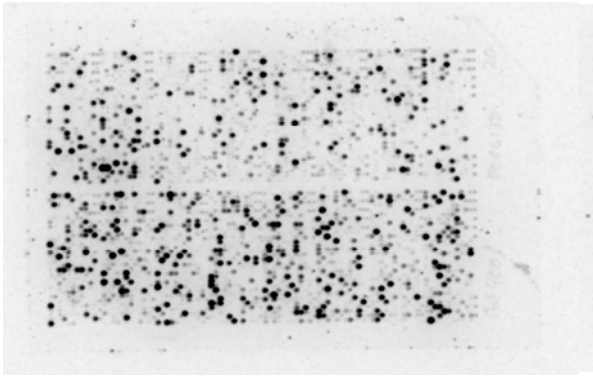


I374N probed filter F20

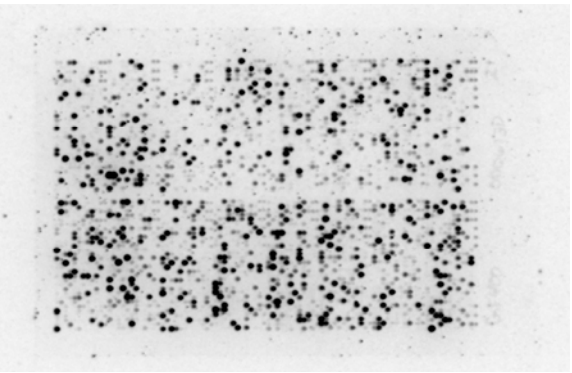


C

V449E probed filter F20



I374N probed filter F21



3.2.2 Analysis of the GeneFilters® membranes

3.2.2.1 Data acquisition

Hybridisation of the radioactively labelled probes to the arrays was visualised using a phosphorimager screen and scanner (Figure 7B and C). The relative intensity of the spots on the filter reflects the relative binding of ^{33}P-cDNA. Indirectly this gives an indication of those spots that have bound the most probe. By comparing the intensity of a spot on one filter to the equivalent spot on the other filter, the relative expression level of the gene (represented by that spot) could be determined.

To examine the intensities of the spots on each filter, image analysis software was used to extract data from each of the four hybridisations in order to identify differentially expressed genes. The initial approach for the image analysis and data extraction is represented by the shaded region of Figure 8. The first approach, represented in Figure 8A, identified differentially expressed genes based on visual inspection together with the phosphorimager software, ImageQuant, to quantify the difference in gene expression observed between the samples (Figure 8A). However, ImageQuant required manual selection of each spot for quantification and proved to be unsuitable for examining all the spots on the array at once. To examine an entire array, software designed specifically for either microarray (GenePix Pro) or large membrane array (Array Vision) analyses were tested (Figure 8B). GenePix Pro and Array Vision software were used to identify genes that were differentially expressed by examining each set of two hybridisations separately.

3.2.2.2 Background correction and normalisation of GenePix Pro and Array Vision data

Both GenePix Pro and Array Vision software programs were used to a perform background correction of the data (the ability to measure a spot's intensity and remove the effect of nearby spots or artefacts) according to the manufacturer's recommendations. It was soon apparent that even after background correction there was still an overall or 'global' difference between the two filters. For example, in the first set of hybridisations the overall average intensity of all the spots on the F20 filter was higher than the overall average for the F21 filter. Such a difference between filters is likely to be caused by differences in the level of incorporation of ^{33}P into the I374N probe compared to the V449E probe. As a consequence of this difference, and in order to compare across the two filters within each set of hybridisations, there was a need for

normalisation of the data to correct this imbalance. Normalisation of the data was achieved by dividing each individual gene's intensity by a scaling factor. For example, the median intensity of all the genes on the array was used as the scaling factor, and this number was then used to divide the background subtracted intensity value for each gene on the array to obtain a normalised value.

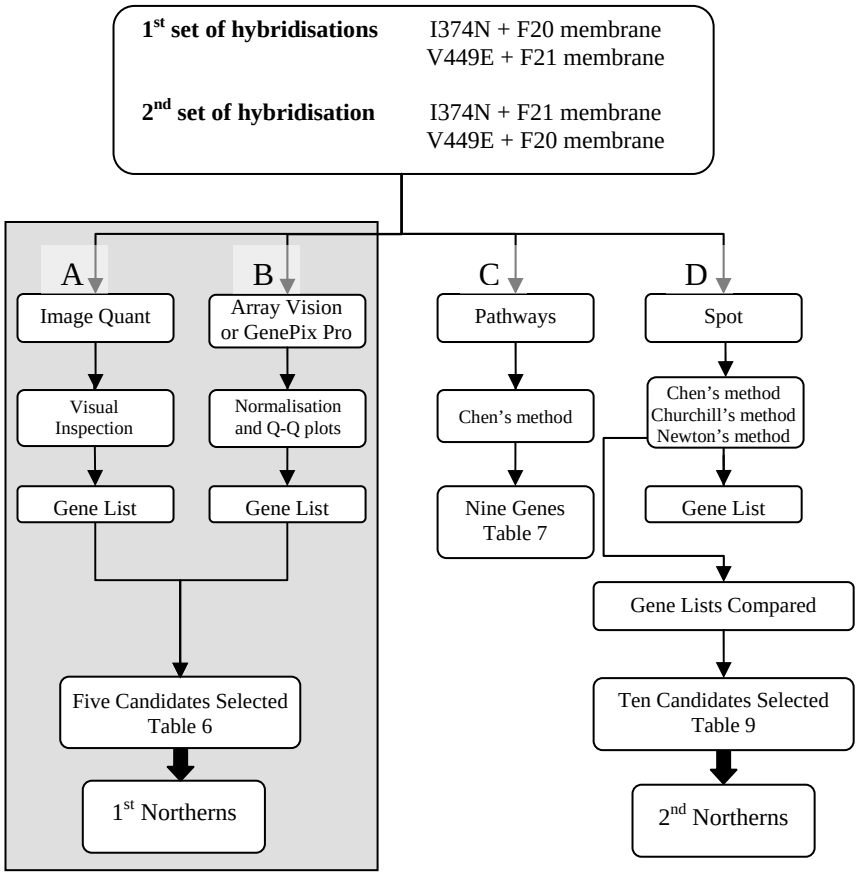


Figure 8: Flow diagram showing the four methods used in the analysis of the membrane array experiment. Two sets of two hybridisations (1st and 2nd) were performed using a pair of GeneFilters®. Each set of hybridisations were analysed individually via four methods (A-D) to quantitate the intensity of the spots on the arrays and to identify differentially expressed genes (see text for details). The initial approach is highlighted by the shaded region encompassing methods A and B. Each method generated a list of differentially expressed genes ("gene list"). The lists of differentially expressed genes identified from these two approaches were examined and a number of candidate genes were selected for investigation of their mRNA expression levels via Northern blotting ("Northerns").

3.2.2.3 Use of Quantile-Quantile plots for the initial analysis to identify differentially expressed genes

Using the normalised GenePix Pro and Array Vision data two methods were investigated to transform the data: (i) calculating the logarithm (base 2) (log_2) and (ii) calculating the square root of the data. The purpose of transforming the data is to achieve variance stabilisation – where there is equal variance across all intensity values – so that the variance seen in the data is dependent on the gene itself rather than being dependent on its intensity. The data from each set of hybridisations was then plotted using Quantile-Quantile (or Q-Q) plots. Q-Q plots show genes grouping together when their expression differences can be explained by a normal distribution and identifies differentially expressed genes as outliers (Moore and McCabe, 1998; Dudoit *et al.*, 2000). Q-Q plots were performed on both sets of hybridisations, examining data normalised by both transformation methods, to identify outlying spots which were indicative of significantly differentially expressed genes. In Figure 9 the Array Vision data for the first set of hybridisations is presented as an example of the analysis that was performed. The results from the data normalised via either log_2 or square root transformation methods are shown in Figure 9 with differentially expressed genes indicated by arrows. Similar plots were produced using the GenePix Pro data (data not shown). Plotting of the data in this way pointed to the presence of a number of outlying genes in the data set from both hybridisations, suggesting differential expression (see Figure 8B).

Using the methods described throughout Section 3.2.2, several lists of differentially genes were identified from the ImageQuant, GenePix Pro and Array Vision quantitated data. Five of the most promising genes found to be differentially expressed using these methods are listed in Table 6 and were examined via Northern blotting (see Section 3.2.2.4). The five genes were selected on the basis of their possible involvement in mediating the signalling differences between the two hβc mutants.

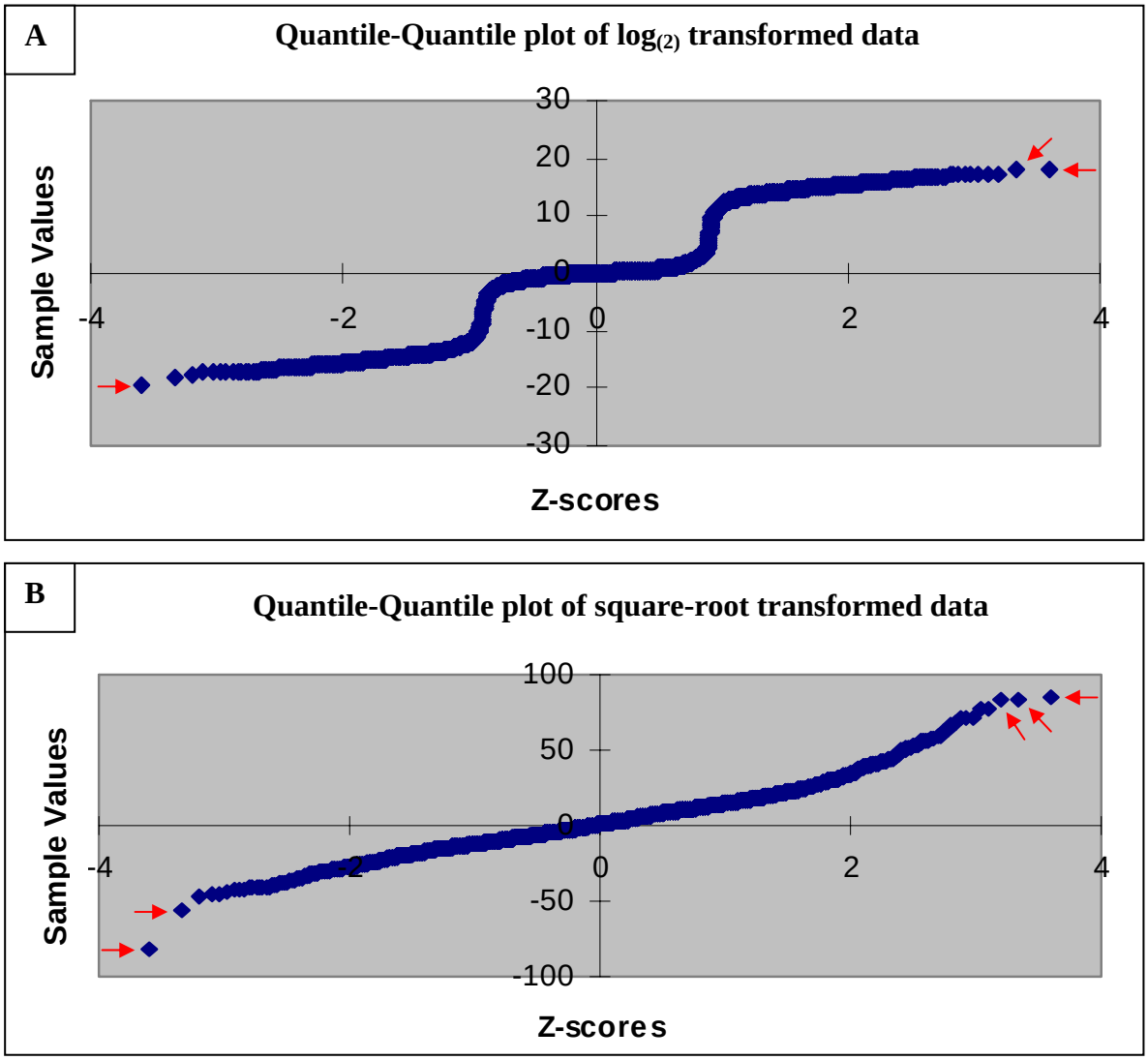


Figure 9: Log and square root transformed data plotted via Q-Q plots showing the presence of outlying genes in the data from the first set of hybridisations. The Q-Q plot shows the quantiles of the data ("Sample values") plotted against the quantiles of a standard normal distribution ("Z-scores"). The graph values on y-axis represent either the log (base 2) transformed or square root transformed ratios (V449E over I374N) obtained from the first set of hybridisations. The outlying genes (indicated by red arrows) in this plot are those genes that deviate from a normal distribution and can be considered to be statistically significant. (A) shows log (base 2) transformed data and (B) shows square root transformed data plotted via Q-Q plots.

3.2.2.4 Northern blotting of five genes identified from the initial analysis

An independent method capable of measuring relative gene transcript levels, called Northern blotting, was used to confirm the differential expression of genes identified from the membrane array experiment. The Northern blots described in this chapter used poly A^+ RNA collected from FDC-P1 cell lines expressing the mutants, V449E or I374N, or the wild type human GM-CSF receptor subunits hGMRα and hβc (αβ). RNA purification and Northern blotting were performed as described in Sections 2.3.2 and 2.3.5. FDC-P1 cells expressing either the I374N or V449E mutants were grown continuously in medium lacking IL-3 or GM-CSF. This was performed in order to maintain expression of the hβc mutants, which were required for the factor independent proliferation of the cells. The FDC-P1 αβ cells expressing the wild–type human GM-CSF receptor were washed and grown in the presence (+) or absence (-) of human GM-CSF (hGM-CSF at 2 ng/mL) for six hours. The FDC-P1 αβ cells were factor starved for six hours to decrease their level of GM-CSF receptor related signalling and were used to indicate whether the genes were regulated by the wild type receptor, in addition to any regulation seen by comparing mutant hβc signalling (i.e. I374N and V449E expressing cells).

To validate the genes identified from the initial analysis (see shaded region of Figure 8) five genes were selected for analysis of their mRNA expression levels using Northern blots. The five genes listed in Table 6 were PCR amplified to generate gene specific probes that were radiolabelled with ^{32}P and used to probe the Northern blots (see Figure 10). To normalise each gene's expression level, the blots were re-probed with a sequence specific for glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) (Figure 10). The normalisation to *Gapdh* enabled the relative gene expression difference between the FDC-P1 cell lines to be calculated and is shown in Table 6 as the normalised ratio of V449E over I374N, or αβ +hGMCSF over αβ -hGMCSF.

Protein tyrosine phosphatase, mitochondrial 1 (*Ptpmt1)* was the only gene that showed differential expression between I374N and V449E in excess of 2 fold after normalising its expression level to *Gapdh* (see Figure 10 and Table 6). Replication of the *Ptpmt1* hybridisation on another independent Northern blot revealed a similar level of differential gene expression. The combined results indicate an average increase of 1.85 fold in V449E compared to I374N (see Table 6B and Figure 11). Both *Ptpmt1* and

Ngfi-A binding protein 1 (*Nab1)* showed regulation in response to hGM-CSF on the Northern blots. Decreased mRNA levels of *Ptpmt1* and *Nab1* were seen in FDC-P1 αβ cells grown in the absence of exogenously added hGM-CSF for six hours (αβ-hGM-CSF), compared to FDC-P1 αβ cells in the presence of hGM-CSF (αβ+hGM-CSF) (Figure 10,Figure 11 and Table 6). The regulation seen in response to hGM-CSF implicates these two genes in downstream signals generated by GM-CSF signalling through hβc. However, in the case of *Nab1*, no differential expression was seen between V449E and I374N on the Northern blots, suggesting that this gene is not involved in the differential signalling properties of the V449E or I379N mutants (Figure 10 and Table 6). The *Gapdh* normalised results from *Ptpmt1* Northern blots revealed that expression of V449E results in increased expression of *Ptpmt1* compared to I374N or αβ +hGM-CSF in FDC-P1 cells. FDC-P1 αβ cells grown in the absence of hGM-CSF showed at least a 2 fold decrease in expression of *Ptpmt1* compared to I374N or αβ-hGM-CSF, and greater than a 2 fold decrease when compared to V449E (see Figure 11).

In summary, Northern blotting confirmed that *Ptpmt1* was differentially expressed between V449E and I374N as identified from the initial membrane array analysis. This suggests that expression of this gene may be a consequence of the different signalling properties of the two hβc mutants. The Northern blots also revealed that the genes, *Ptpmt1* and *Nab1*, were regulated by hGM-CSF in FDC-P1 cells expressing hGMCSFα and hβc receptor subunits. The regulation in response to hGM-CSF suggests these two genes may be involved in downstream signals generated by GM-CSF signalling through hβc.

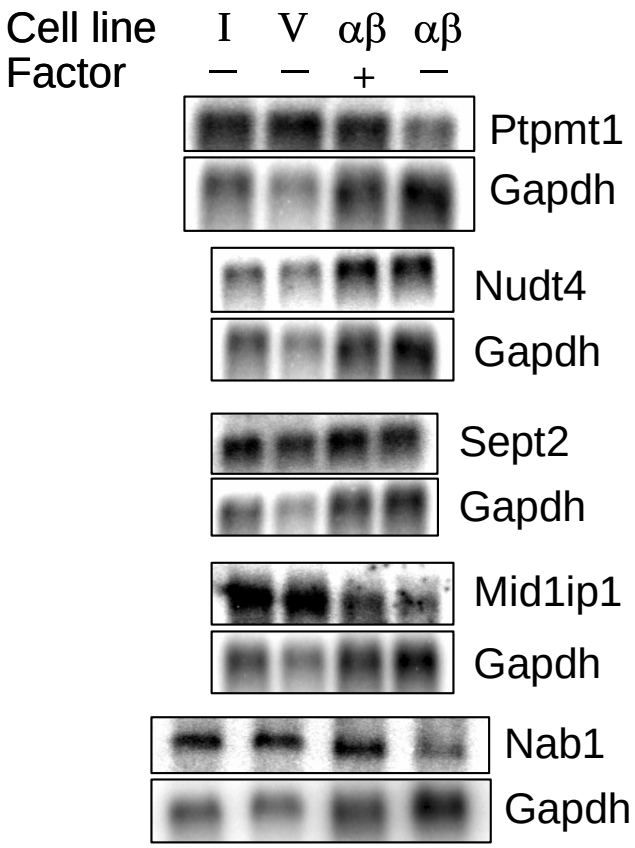


Figure 10: Northern blotting of five genes identified from the initial analysis. The five genes were PCR amplified and used as probes on Northern blots to determine their expression. The Northern blots were then probed with a housekeeping gene (*Gapdh*) to normalise for mRNA loading on the blots between each lane. V = V449E, I = I374N, αβ = FDC-P1 cells expressing α and β subunits, + = grown in the presence of hGM-CSF (2 ng/mL) and - = in the absence of hGM-CSF for 6 hours.

A

Acc	Title	Current Unigene no.	Gene	Description	Northern Quantitation Normalised ratio (V/I)	Northern Quantitation Normalised ratio (+αβ/-αβ)
AI324255	ESTs, Highly similar to GASTRULATION SPECIFIC PROTEIN G12 [Danio rerio]	Mm.29429	Mid1ip1	Mid1 interacting protein 1 (gastrulation specific G12-like (zebrafish))	1.38	1.53
AI426110	ESTs, Moderately similar to diphosphoinositol polyphosphate phosphohydrolase [H.sapiens]	Mm.24397	Nudt4	Nudix (nucleoside diphosphate linked moiety X)-type motif 4	1.40	1.25
AI430891	ESTs, Weakly similar to Similarity to Human MAP kinase phosphatase-1 [C.elegans]	Mm.23926	Ptpmt1	Protein tyrosine phosphatase, mitochondrial 1	2.21	1.93
AI451044	Ngfi-A binding protein 1	Mm.25903	Nab1	Ngfi-A binding protein 1	1.15	2.31
AI528671	Nedd 5 protein	Mm.428652	Sept2	Septin 2	1.47	1.31

B

	Northern Quantitation Normalised ratio (V/I)	Northern Quantitation Normalised ratio (+αβ/-αβ)
Ptpmt1 - First Northern blot	2.21	1.93
Ptpmt1 - Second Northern blot	1.49	2.16
Average	1.85	2.05

Table 6: Quantitation of the Northern blots for the five genes from the initial analysis. (A) The five genes that were used as probes on Northern blots had their expressions compared to a housekeeping gene (*Gapdh*) to normalise for mRNA loading on the blots between each lane. This data is then presented as the *Gapdh* normalised ratio of either V over I or +αβ over –αβ where V = V449E, I = I374N, +αβ = FDC-P1 cells expressing α and β subunits grown in the presence of hGM-CSF or -αβ in the absence of hGM-CSF for 6 hours. Acc = accession number and Title = title of the spot on the array as stated by Research Genetics. The current (January 2007) Unigene match (Current Unigene no.) for each accession number is presented along with the corresponding gene title and description that are associated with each Unigene entry. (B) The normalised results from the first and second Northern blots performed with *Ptpmt1* are presented along with the averages and standard deviations from the two Northern blots for the ratios V449E/I374N and +αβ/-αβ.

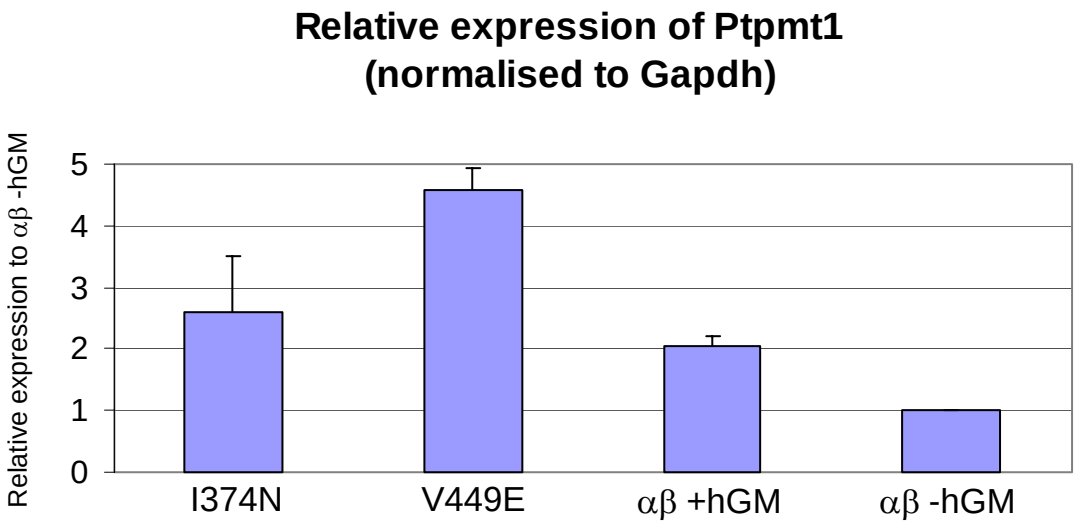


Figure 11: Relative expression of Ptpmt1 in FDC-P1 cells measured by Northern Blotting. The expression of *Ptpmt1* was assessed in FDC-P1 cells expressing I374N, V449E or wild type hβc receptor (αβ) by Northern blotting. I374N and V449E expressing FDC-P1 cells were grown continuously in medium lacking IL-3 or GM-CSF. FDC-P1 αβ cells were washed and grown in the presence of hGM-CSF (2 ng/mL) (+hGM) or in the absence of hGMCSF (-hGM) for six hours. A sequence specific for the gene *Ptpmt1* was used to probe two Northern blots. The blots were stripped and re-probed with a sequence representing *Gaphd* to normalise for mRNA loading between each sample. The data is presented as the relative expression of each sample to the *Gapdh* normalised expression of *Ptpmt1* in αβ-hGM sample. The graph shows the average of two experiments with the standard deviation for the I374N, V449E and αβ+hGM samples.

3.2.3 Statistical analysis methods used to determine a second set of differentially expressed genes

3.2.3.1 Image analysis and data acquisition using 'Pathways' software

The initial analysis methods described above (Section 3.2.2) and shown by the shaded region of Figure 8, could only examine one membrane array at a time. The software program '*Pathways*' allowed the examination of multiple filters simultaneously to obtain background subtracted intensity levels for all spots on the arrays. Pathways is a program developed and sold by Research Genetics specifically for the purpose of analysing their GeneFilters®. This software was used to combine each set of the two hybridisations and identify those genes that showed either a 2 fold up- or down-regulation between V449E and I374N (i.e. a ratio of either greater than 2 or less than 0.5) (Figure 8C). This was a generous limitation given that other researchers have applied a 3 fold cut-off on these GeneFilters® (de Haan *et al.*, 2002). The data from each set of hybridisations was normalised by multiplying by a scaling factor (see Section 3.2.2.2). From each set of hybridisations lists were obtained stating genes greater than 2 fold up- or down-regulated. Comparisons between the two lists defined a common set of nine genes which are listed in Table 7 (and shown in Figure 8C).

3.2.3.2 Using the method of Chen et al., (1997) to identify differentially expressed genes from Pathways analysed data

To define statistically significant genes in the Pathways analysed data, the first of three statistical methods described in this Chapter were used (see Figure 8C). These three statistical methods are described as single-slide methods (see Section 1.6.3.3) for their ability to identify significantly differentially expressed genes from a single microarray slide or from a pair of hybridised membrane arrays without requiring replication of the experiment. The statistical method described by Chen *et al.*, (1997) (which will be referred to as 'Chen's method') examines the gene expression ratio of each spot on the array to assesses the significance of the observed gene expression ratios (Dudoit *et al.*, 2000; Smyth *et al.*, 2003). Using Chen's method of assessing genes which are significantly differentially expressed, lists were obtained from each set of hybridisations using 90%, 95% or 99% confidence intervals. In these six lists of significantly differentially expressed genes, the presence or absence of each of the nine genes identified in Table 7 was examined. This comparison provided a measure of significance for each of the nine genes, by determining at what level a given gene in a given set of hybridisations was no longer considered significant by Chen's method

(Table 7). As can be seen in Table 7, three of the genes were only significant in one set of hybridisations at the 90% confidence level, while only one gene was significant in both sets of hybridisations at the 99% confidence level.

The nine genes identified by Pathways as differentially expressed between V449E and I374N were investigated further by visually examining the images of the four hybridised arrays. The Pathways processed images of the filters were inspected to reveal how accurately the spots had been defined (i.e. segmentation, see Section 1.7.2). Three of these nine candidate genes showed no influence by background spots and were 'visually' confirmed to be differentially expressed (see Table 7). Three other genes showed minor influences of background spots, and the last three genes were found to be false positives caused by the strong influence of a nearby spot (Table 7). The data from these nine genes presented in Table 7 suggests a lack of correlation between the visual confirmation of a differentially expressed gene and the level of significance for that gene obtained from Chen's method. For example, the only gene significant in both sets of hybridisations at a level of 99% (AI465155) was identified as a false positive, while two of the three genes which were visually confirmed (AI413927 and AI428372) were significant at a level of 95% for both sets of hybridisations (and at a level of 99% for the first set of hybridisations) (Table 7). Therefore, the presence of a gene in the 99% confidence list, assessed by Chen's method, did not infer that the gene was differentially expressed when the arrays were visually inspected. The above findings suggest that Pathways software was not adequately dealing with the influences of background spots on the arrays, which was leading to an incorrect assignment of differential expression for some genes.

Acc	Title	Best match	Unigene cluster	Normalised scaled ratio		Chen's test with a confidence interval of			Visual inspection of differential gene expression
				1st hyb (V449E/ I374N)	2nd hyb (V449E/ I374N)	90%	95%	99%	
AI413927	ESTs	Neurexophilin 3	Mm.22970	0.29	0.36	yes	yes	only 1st hyb	++
AI428372	ESTs	Pleckstrin homology domain containing, family K member 1	Mm.209844	0.30	0.33	yes	yes	only 1st hyb	++
AI450172	ESTs	IQ motif and WD repeats 1	Mm.227605	2.19	3.35	yes	no	no	+
AI465377	EST	Zinc finger protein 148	Mm.256809	0.41	0.41	only 1st hyb	only 1st hyb	no	++
AI324092	ESTs, Weakly similar to serine protease [H.sapiens]	Serine protease HTRA3	Mm.274255	0.41	0.24	only 2nd hyb	only 2nd hyb	no	-
AI324874	ESTs, Highly similar to GLANDULAR KALLIKREIN 2 PRECURSOR [Homo sapiens]	RIKEN cDNA 1110030O19 gene	Mm.20024	0.44	0.28	only 1st hyb	only 1st hyb	no	+
AI326894	DNA segment, Chr 13, Wayne State University 50, expressed	RIKEN cDNA 1700025O18 gene	Mm.52583	-0.015	0.21	yes	yes	only 1st hyb	+
	Unknown 2h1126			0.31	0.34	yes	yes	only 1st hyb	-
AI465155	ESTs	RIKEN cDNA 1110031M08 gene	Mm.347805	0.22	0.10	yes	yes	yes	-

Table 7: Summary of the nine candidates identified as statistically differentially expressed using the method of Chen *et al*., (1997) in the Pathways analysed data for both hybridisations. The normalised scaled ratios (expressed as the background subtracted V449E intensity over I374N intensity) from both sets of hybridisations (1st and 2nd) are presented along with the details of each spot's accession number, title (Research Genetics), best match title (Unigene) and the significance level at which that clone was still found to be differentially expressed using the method of Chen *et al*., (1997). Abbreviations and notes: Acc = accession number; title = title on array provided by Research Genetics; Best match = name of the Unigene cluster found to match that accession number. Visual inspection refers to the examination of the Pathways processed images of the GeneFilters® to determine if the observed differential expression was due to the influence of background spots. A +/- system has been used to indicate the confidence in the visually observed differential expression, such that ++ = confirmed differentially expression with no background influences, + = minor influence of background spots on the differential expression, - = strong influence of background spots, suggesting an absence of differential expression. Note, the negative value for one of the genes was the result of an intensity less than background on the F21 filter probed with V449E in the first set of hybridisations.

3.2.3.3 Image analysis and data acquisition using 'Spot' software

Access to recently developed CSIRO image processing software called '*Spot'*, which was designed for microarray and membrane array applications, was obtained in collaboration with Associate Professor P. Solomon and Mr. J. Whiting (Applied Mathematics Department, Adelaide University). Spot software uses sophisticated algorithms such as seeded region growing which defines non-circular shaped spots and morphological opening which determines the background at the exact location of each spot on the array (Buckley, 2000; Yang *et al.*, 2001a; Yang *et al.*, 2001b). This background calculation method is different to the methods used in the image analysis programs described earlier (Pathways, Array Vision and GenePix Pro) that used regions near to each spot to calculate background levels (Yang *et al.*, 2001a; Yang *et al.*, 2001b). Yang *et al*., (2001a, 2001b) state that the morphological opening algorithm used in Spot provides better estimates of background levels, which was expected to produce more accurate values for both the intensity and background for each gene in this study.

3.2.3.4 Significantly expressed genes identified by the statistical methods of Chen et al., (1997); Sapir and Churchill (2000); and Newton et al., (2001)

As was achieved with the Pathways image analysis software, a list of significant differentially expressed genes found in common to the first and second sets of hybridisations was identified using the data acquired from the image analysis software, Spot (see Figure 8D). This was achieved on the Spot data by using Chen's method with 99% confidence levels, to produce a list of significantly differentially expressed genes that were in common to both hybridisations (see Table 8).

The use of Spot enabled the data to be analysed by two additional statistical methods, Sapir and Churchill (2000) and Newton *et al*., (2001), which were not available for use on the Pathways acquired data. The method, proposed by Sapir and Churchill (2000) ('Churchill's method') using 99% posterior probabilities to define statistically differentially expressed genes was used to produce a list of statistically differentially expressed genes that were in common to the first and second sets of hybridisations using the Spot data (see Table 8 and Figure 8D).

A third statistical method proposed by Newton *et al*., (2001) ('Newton's method') was applied to the Spot data to identify significantly differentially expressed genes (see Table 8 and Figure 8D). The reason for using Newton's method was to introduce an intensity dependent component to the assessment of statistically significant

genes. Unlike Chen's and Churchill's methods, Newton's method adjusts the level of significance depending on the intensity of the gene, i.e. a ratio composed of low intensity measurements is regarded as more variable than the same ratio composed of higher intensity measurements. A list of significant differentially expressed genes found in common to both sets of hybridisations was identified with Newton's method. This was achieved using a significance level of log odds greater than two on the Spot data and the resulting list is stated in Table 8 (see Figure 8D).

The differences between the three methods (Chen's, Churchill's and Newton's) can best be represented on an M versus A plot (Figure 12). The M versus A plot shows the average (mean) log intensity ("A") of the two spots (one from each sample) on the x axis against the log intensity ratio ("M") of the two spots on the y axis using a $\log_2$ scale for both the x and y axis. Across each graph the three analysis methods are represented by a pair of lines showing the different cut-offs for significance for each method (Figure 12). It can be seen in Figure 12 that Chen's method and Churchill's method produce horizontal cut-offs in the M versus A plots at a fixed M (or ratio) value, whereas the cut-offs produced by Newton's method are curves, indicating decreased significance of both the low intensity and the high intensity spots.

In summary, using the Spot data, both sets of hybridisations were examined by the three statistical analysis methods, in order to identify the genes common to both sets of hybridisations that were differentially expressed between V449E and I374N (Table 8). There was some overlap in the genes identified by the three methods with both Newton's and Churchill's lists containing a subset of Chen's list (see Table 8). This overlap indicated agreement between the three methods in the genes identified as significantly differentially expressed. On this basis, the overlapping genes as well as the known genes (but not the ESTs) identified by Chen's method were chosen for analysis by Northern blotting (see shaded genes in Table 8). Examination of these genes by Northern blotting (see Section 3.2.4.3) was performed to establish if the differential gene expression identified by the membrane array analysis would be confirmed.

Chen's method -99% confidence interval

gene expression greater in I37N than V449E (I>V)

Index no.	Acc	Unigene no.	Title	Chrom
2112	AI448274	Mm.9590	Nakap95-pending Neighbor of A-kinase anchoring protein 95	17
4334	AI451623	Mm.32510	Mus musculus DC-SIGN (Dcsign) mRNA, complete cds	8

gene expression greater in V449E than I37N (V>I)

1814	AI429085	Mm.29788	ESTs	4
4358	AI451578	Mm.32498	ESTs	X
4122	U.2a613		spot on array with unknown accession no.	
4529	AI449120	Mm.32114	ESTs	17
2493	AI894208	Mm.41857	AI894218 Expressed sequence AI894218	11
4020	AI451707	Mm.17261	D19Ertd626e DNA segment, Chr 19, ERATO Doi 626 expressed	19
2018	AI426467	Mm.202311	Gna13 Guanine nucleotide binding protein, alpha 13	11
3255	AI451248	Mm.43152	2010200P20Rik RIKEN cDNA 2010200P20 gene	X
4719	AI450318	Mm.23910	Etohd2 Ethanol decreased 2	13
861	AI528700	Mm.14530	Rab1 RAB1, member RAS oncogene family	11

Churchill's method -99% posterior probability

I>V

Index no.	Acc	Unigene no.	Title	Chrom
4334	AI451623	Mm.32510	Mus musculus DC-SIGN (Dcsign) mRNA, complete cds	8

V>I

none

Newton's method -log odds for differential expression greater than 2

I>V

Index no.	Acc	Unigene no.	Title	Chrom
99	AI430764	Mm.17917	2600015J22Rik RIKEN cDNA 2600015J22 gene	4
100	AI413924	Mm.22869	ESTs	
4232	AI447409	Mm.31842	ESTs	
4322	AI451649	Mm.30642	ESTs	
2112	AI448274	Mm.9590	Nakap95-pending Neighbor of A-kinase anchoring protein 95	17
865	E.1h1110		Empty = blank spot	
4334	AI451623	Mm.32510	Mus musculus DC-SIGN (Dcsign) mRNA, complete cds	8

V>I

4358	AI451578	Mm.32498	ESTs	X

Table 8: Candidate genes in common between the two membrane array hybridisations analysed via the three statistical methods. Both hybridisations were analysed via the methods of Chen *et al*., (1997), Sapir and Churchill (2000), and Newton *et al*., (1999) to generate statistically significant gene lists. For each method the lists were separated in to those genes that had increased expression in the V449E sample (V>I) or those genes that had increased expression in the I374N sample (I>V). These lists were then compared between the two hybridisations for genes in common and are listed along with their index number (Index no.), accession number (Acc), title (Title), Unigene number (Unigene no.) and chromosome number ("Chrom") information that matched the corresponding accession number. The Unigene match stated was current as of November 2001. Shaded genes were examined using Northern blotting as described in Section 3.2.4.3.

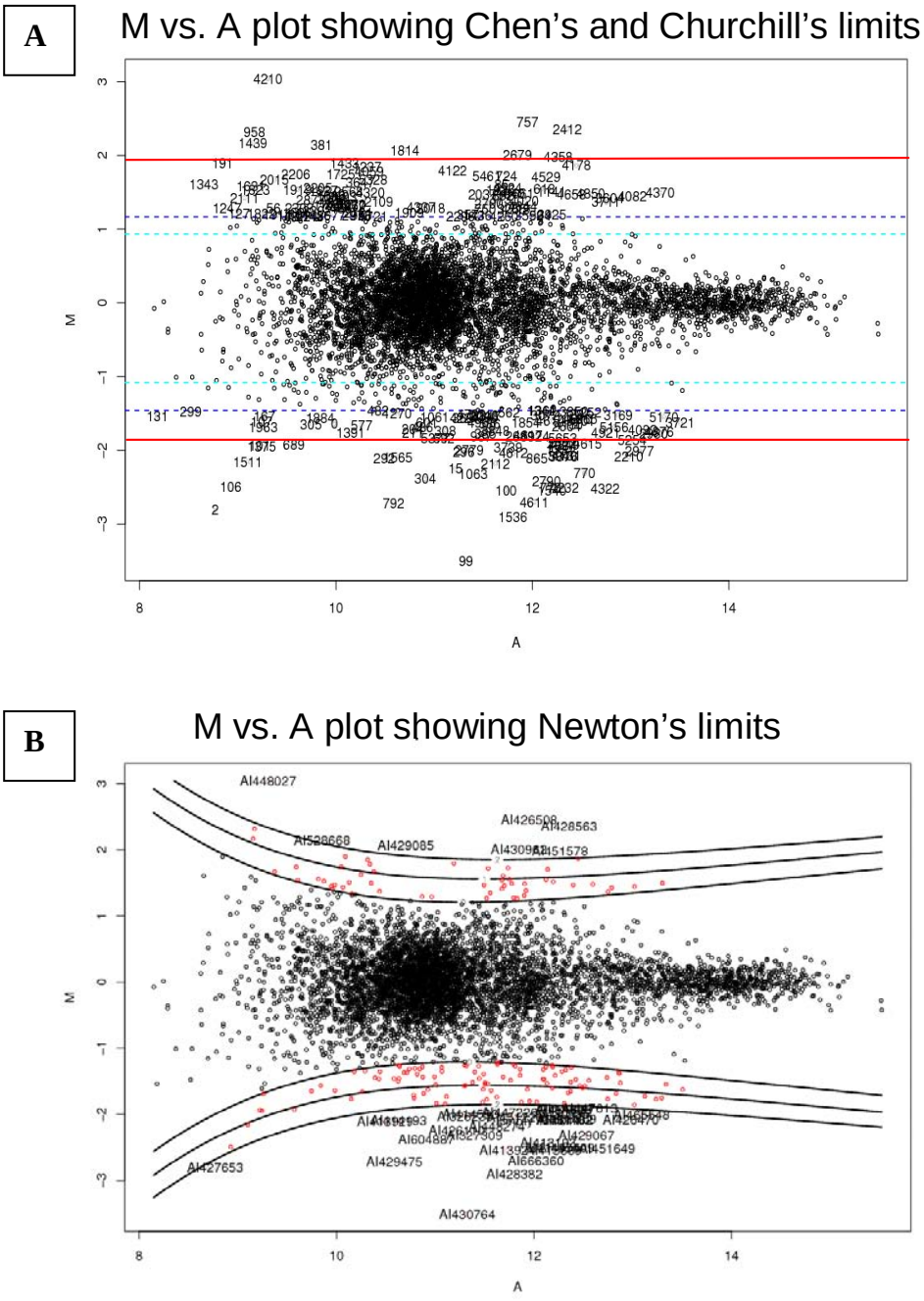


Figure 12: M versus A plots showing Chen's, Churchill's and Newton's method cut-offs on the Spot data from the first set of hybridisations. (A) M versus A plot showing 95% (dotted light blue line) and 99% (dotted blue line) confidence intervals for significance produced by Chen's method and 99% posterior probabilities (red line) for Churchill's method. **(B)** M versus A plot showing Newton's 100:1 (outer lines), 10:1 (middle lines) and 1:1 (inner lines) log odds of differential expression. M is the $\log_2$ ratio difference between the intensity of the two samples (I374N and V449E) for each gene and A is average $\log_2$ intensity for each gene. In (A), genes identified via Chen's method to be significantly differentially expressed at a confidence interval of greater than 99% are stated by their index number, whereas in (B), genes significant at a posterior probability of greater than 10:1 are indicated by red circles and those genes greater than 100:1 by their accession numbers.

3.2.4 Validation of gene expression differences using Northern blots for genes identified as differentially expressed by the three statistical analysis methods

The aim was to select a subset of genes for validation of their differences in gene expression by Northern blotting. The genes were selected from a list of 17 unique genes identified by the three statistical methods that were in common between the two sets of hybridisations (see Table 8).

3.2.4.1 Bioinformatics performed on the genes identified from the statistical analysis

Bioinformatics was used to analyse the 17 genes to rationalise the choice of genes to be examined on Northern blots. Several of the genes identified as differentially expressed in the statistical analysis were unknown genes, called expressed sequence tags (ESTs). The ESTs required the use of bioinformatics in order to identify which genes they represented, and to prioritise which ESTs to investigate further. To identify genes that were represented by ESTs on the Genefilters, searches of the Celera mouse genome database (which provided access to sequence data that was not yet available to the public) as well as public databases (e.g. the non-redundant and EST databases available from the NCBI website through BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST/)) were performed. These searches led to the association of several ESTs to known genes, which prior to their analysis, were unknown genes (see Unigene assignments and best match categories in Table 9). The sequence of the ESTs were also examined for potential open reading frames (ORFs) that were used to find putative protein matches in the databases (using blastx, tblastx, tblastn or blastp) or to find predicted conserved protein domains using SMART (ExPASy proteomics – www.expasy.org/tools/). The bioinformatics enabled prioritisation of the genes to a final set of ten genes (shaded in Table 8) that were examined using Northern blots.

3.2.4.2 Generation of probes for Northern blotting of genes identified from the statistical analysis

To confirm regulation of the genes identified from the statistical analysis (see Figure 8D), ten genes were selected for analysis of their mRNA expression levels using Northern blots. As before, poly A^+ RNA was prepared from FDC-P1 cells expressing I374N, V449E or the wild-type human GM-CSF receptor ($\alpha\beta$), except that the $\alpha\beta$ cells were washed and grown in the presence (+) or absence (-) of hGM-CSF (2 ng/mL) for eight hours.

For the analysis of the Northern blots, a gene specific probe for each of the ten genes highlighted in Table 8 was generated by either PCR amplification, or restriction enzyme digestion and purification from clones purchased from Research Genetics (see Table 9). All clones and PCR products were sequenced to confirm their identity prior to their use as Northern probes. Northern blots were re-probed with a sequence specific for *Gapdh* to enable the relative gene expression difference between the FDC-P1 cell lines to be determined (see Figure 13 and Table 9B).

3.2.4.3 Northern blotting of ten genes identified from the statistical analysis

To confirm gene expression differences between V449E and I374N on the membrane array, ten gene specific sequences were radiolabelled with ^{32}P and used as probes for Northern blots (Figure 13 and Table 9). The *RAS oncogene family member, RAB1* (*Rab1)* probe produced two bands on the Northern blot of approximately 1.7 and 2.7 kb (Figure 13). Based on the membrane array data (see Table 9) it was expected that *Rab1* would be between 2 and 4 fold over-expressed in the V449E expressing FDC-P1 cells compared to I374N expressing cells. However, it was found that after *Gapdh* normalisation, neither of the two *Rab1* bands showed significant regulation between V449e and I374N or between αβ cells +/- hGM-CSF. The 2.7 kb band that matched the predicted mRNA size for *Rab1* (as stated in public databases, NM_008996) showed less than a 1.4 fold regulation of I374N over V449E (or 0.74 fold V449E over I374N). Only after a long exposure (of 4 days) to the *Ubiquitin-conjugating enzyme E2I* (*Ube2i*) probed Northern blot were multiple faint bands visible (data not shown). These bands ranged in size from 1.3 kb up to 9 kb, including a band at 2.5 kb (the predicted size of *Ube2i* based on its mRNA sequence length in public databases, NM_011665). The membrane arrays suggested that *Ube2i* would be 3 to 4 fold over-expressed in V449E sample, however, none of the bands on the Northern blot showed differential expression in any of the four cell lines examined on the blot. Probing with *DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26B* (*Ddx26b*) gave two bright bands of approximately 4 and 7 kb and probing with *RAS oncogene family member, RAP2C* (*Rap2c*) gave three bands (the larger two were more prominent) on the Northern blots ranging in size from 3 to 4.4 kb (Figure 13). The expected sizes for *Ddx26b* and *Rap2c* were 3.8 kb (NM_173779) and 3.6 kb (NM_172413) respectively, based on their mRNA sequence length in the public databases. Neither of these genes showed any significant difference in expression levels between I374N and V449E (Figure 13 and Table 9). Examination

of the difference between FDC-P1 αβ cells grown in the presence or absence of exogenously added hGM-CSF for eight hours revealed regulation of *Ddx26b* in response to hGM-CSF (Figure 13). Normalisation of the expression levels of *Ddx26b* (using *Gapdh*) revealed a 3.2 and 1.8 fold increase in factor starved over factor stimulated αβ cells for the upper (7 kb) and lower (4 kb) bands respectively. The probes for both *Guanine nucleotide binding protein, alpha 13* (*Gna13*) and *Zinc finger, CCHC domain containing 6* (*Zcchc6*) showed a single large band around 9 kb but no regulation was seen in any of the four cell lines examined on the Northern blot (Figure 13). Probing with either *Homeo box B2* (*Hoxb2*) or *CD209a antigen* (*Cd209a*) showed no bands on the Northern blots (data not shown). The *A kinase anchor protein 8-like (Akap8l*) probe seemed to bind non-specifically to the high molecular weight RNA on the blot (data not shown). A possible gene specific band of approximately 1.2 kb (Figure 13) was seen on the *Akap8l* probed Northern blot, however, this band did not show any differential regulation between the samples. The predicted size of the *Akap8l* mRNA transcript was 2.1 kb (NM_017476), which suggests that the 1.2 kb band may be the result of non-specific binding of the *Akap8l* probe to ribosomal RNA. Probing with the *Progressive external ophthalmoplegia 1* (*Peo1*) gene sequence resulted in one bright band (approximately 4 kb) matching the predicted size (3.5 kb, NM_153796) together with a smaller less intense band (Figure 13). However, neither of these two bands were differentially expressed between the I374N and V449E samples or between FDC-P1 αβ cells +/- hGM-CSF (Table 9).

In summary, of the ten genes identified from the statistical analysis of the Spot data, seven genes showed gene specific bands on the Northern blots. However, gene expression differences between V449E and I374N expression FDC-P1 cells were detected in contrast to the findings from the membrane array analysis. The Northern blot probed with *Ddx26b* revealed that the mRNA levels of *Ddx26b* are reduced in hGM-CSF stimulated FDC-P1 αβ cells compared to V449E, I374N and αβ cells grown in the absence of hGM-CSF. The difference in *Ddx26b* mRNA levels may indicate the existence of a pathway that is activated by hGM-CSF in FDC-P1 cells expressing the wild type hGM-CSF receptor subunits (αβ), but which remains inactive in the absence of hGM-CSF in the αβ cells, or when the mutants V449E and I374N are expressed in FDC-P1 cells. The activation of this pathway in FDC-P1 αβ cells in response to hGM-CSF results in the suppression of *Ddx26b* mRNA levels.

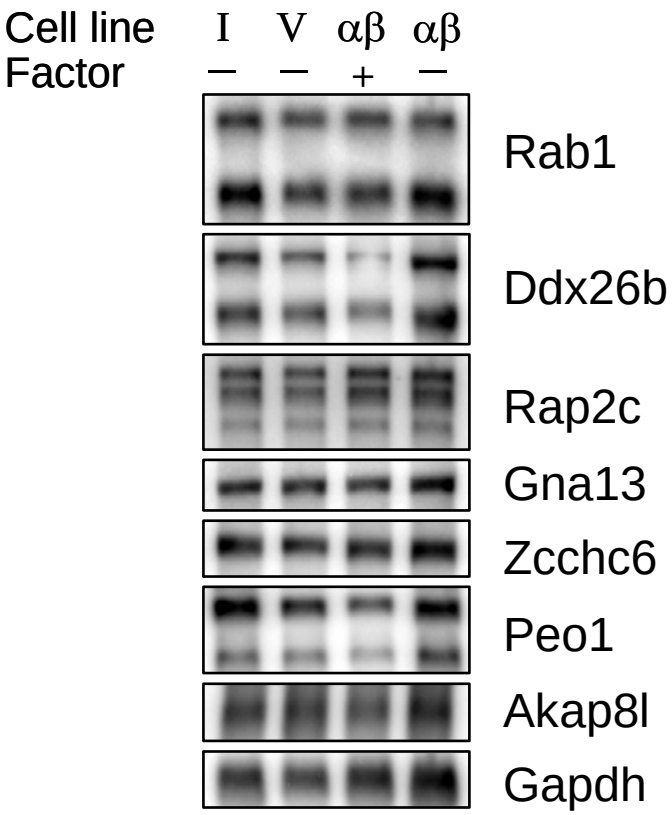


Figure 13: Northern blotting of genes identified from the statistical analysis. The ten candidates from the statistical analysis genes were either PCR amplified or restriction enzyme digested and purified from purchased clones for use as probes on Northern blots to determine their expression. The Northern blots were probed with a housekeeping gene (*Gapdh*) to normalise for mRNA loading on the blots between each lane. Probing with the *HoxB2* and *Cd209a* probes showed no specific bands, and probing with *Ube2i* revealed multiple faint bands (data not shown). Probing with *Akap8l* showed non-specific binding to larger sized MW mRNA (data not shown) as well as the band shown at approximately 1.2 kb. No differential expression between V449E and I374N expression was observed for any of the genes examined. V = V449E, I = I374N, αβ = FDC-P1 cells expressing α and β subunits, + = grown in the presence of hGM-CSF and - = in the absence of GM-CSF for 8 hours.

Table 9: Ten candidates from the statistical analysis used for Northern blots. (A) The list of ten genes that were selected for Northern analysis from the 17 genes identified from the statistical analysis in common to both sets of hybridisations. For each candidate, the current (February 2007) and previously obtained (2002) Unigene matches for each accession number are presented, along with the corresponding gene title and description that are associated with each Unigene number. Acc = accession number and Title = title of the spot on the GeneFilter® provided by Research Genetics. (B) The ten candidate genes from the statistical analysis were either PCR amplified ('PCR probe') or restriction enzyme digested and purified from purchased clones ('clone inserts') (indicated by √). The PCR probes or clone inserts were used to probe Northern blots to determine the expression of the candidate genes. The fold change from the 1st and 2nd sets of membrane array hybridisations (expressed as V449E (V) over I374N (I)) have been presented along with any differential expression ('Differential regulation') seen on the Northern blots. No differential expression between I374N and V449E expressing cells was observed for any of the genes and therefore normalised values for the Northern blots have not been presented. The *Ddx26b* probed Northern blot revealed differential expression between FDC-P1 $\alpha\beta$ cells grown with (+$\alpha\beta$) and without (-$\alpha\beta$) hGMCSF for 8 hours. The differential expression seen for *Ddx26b* between these two populations is stated for the "upper" band of approximately 7 kb and "lower" band of approximately 4 kb. The lower band matches the predicted mRNA length of Ddx26b (NM_0172779, 3767 bp).

A

Index no.	Acc.	Previous Unigene no.	Title	Current Unigene no.	Gene	Description
2493	AI894208	Mm.41857	ESTs, Highly similar to HOMEOBOX PROTEIN HOX-B2 [Homo sapiens]	Mm.281153	HoxB2	Homeo box B2
2018	AI426467	Mm.202311	GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-13 SUBUNIT	Mm.193925	Gna13	Guanine nucleotide binding protein, alpha 13
2112	AI448274	Mm.9590	ESTs, Weakly similar to zinc finger protein [M.musculus]	Mm.281005	Akap8l	A kinase (PRKA) anchor protein 8-like
4719	AI450318	Mm.23910	ESTs, Weakly similar to The KIAA0191 gene is expressed ubiquitously. [H.sapiens]	Mm.233082	Zcchc6	Zinc finger, CCHC domain containing 6
4334	AI451623	Mm.32510	ESTs	Mm.32510	Cd209a	CD209a antigen
4020	AI451707	Mm.17261	ESTs	Mm.105585	Peo1	Progressive external ophthalmoplegia 1 (human)
861	AI528700	Mm.14530	RAB1, member RAS oncogene family	Mm.271944	Rab1	RAB1, member RAS oncogene family
4529	AI449120	Mm.32114	EST	Mm.240044	Ube2i	Ubiquitin-conjugating enzyme E2I
3255	AI451248	Mm.43152	ESTs, Highly similar to unknown [H.sapiens]	Mm.43152	Rap2c	RAP2C, member of RAS oncogene family
4358	AI451578	Mm.32498	ESTs	Mm.72753	Ddx26b	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26B

B

Gene	Fold change (V/I) 1st hyb	Fold change (V/I) 2nd hyb	Northern Probe PCR product	Northern Probe Clone Insert	Northern blot Differential regulation
HoxB2	2.69	6.19		√	V/I = none +αβ/-αβ = none
Gna13	2.45	9.13		√	V/I = none +αβ/-αβ = none
Akap8l	-4.53	-6.77		√	V/I = none +αβ/-αβ = none
Zcchc6	2.41	8.82		√	V/I = none +αβ/-αβ = none
Cd209a	-3.81	-7.26		√	V/I = none +αβ/-αβ = none
Peo1	2.60	5.03		√	V/I = none +αβ/-αβ = none
Rab1	2.28	4.69	√		V/I = none +αβ/-αβ = none
Ube2i	3.25	4.08	√		V/I = none +αβ/-αβ = none
Rap2c	2.41	6.23	√		V/I = none +αβ/-αβ = none
Ddx26b	3.94	6.11	√		V/I = none +αβ/-αβ = 0.30 (upper) - 0.56 (lower) -αβ/+αβ = 3.2 (upper) - 1.8 (lower)

3.3 Discussion

3.3.1 Membrane array analysis of activated hbc mutants

The transmembrane mutant, V449E, and the extracellular mutant, I374N, are examples of the two classes of activated hβc mutants, which cause factor independent growth when expressed in FDC-P1 cells. V449E and I374N show different signalling and biological properties in FDC-P1 cells (Jenkins *et al.*, 1998; Jenkins *et al.*, 1999; D'Andrea and Gonda, 2000). The different signalling and biological properties were predicted to be the consequence of altered gene expression patterns as a result of the differing signals emanating from V449E and I374N. The aim of this chapter was to determine the consequence of these signalling differences by identifying genes that showed differential expression in V449E compared to I374N. The work described in this chapter used membrane arrays (GeneFilters®) to screen for gene expression differences between V449E and I374N expressing FDC-P1 cells. To maximise the ability to detect reproducible differences between these two populations, duplicate hybridisations of these two samples to the GeneFilters® were performed, and data was acquired and analysed using a number of approaches (Figure 8). The initial approach (shaded region of Figure 8) identified a list of five genes that showed differential expression on the membrane arrays. These five genes were examined by Northern blotting to confirm the differential expression seen on the membrane arrays. The Northern blots examined FDC-P1 cell lines expressing either V449E or I374N hβc point mutations, as well as FDC-P1 cells expressing both hGMRα and hβc subunits (αβ) grown in the presence or absence of hGM-CSF for six hours (Figure 10). In the subsequent analysis (non-shaded region of Figure 8), two different image analysis programs (Pathways and Spot) were used. These two programs were used together with three statistical analysis methods to determine the statistical significance of the observed differentially expressed genes identified from the membrane arrays. Pathways software identified nine genes in common to both sets of hybridisations, however, visual examination of these nine genes revealed the identification of three false positives by this software, despite two of these genes having significance levels of greater than 95% (see Table 7). Spot software was used in combination with three single slide statistical analysis methods to identify a list of differentially expressed genes in common to both

sets of hybridisations (Table 8). Bioinformatics was used to prioritise this list of genes stated in Table 8 to ten genes, which were examined by Northern blotting (Figure 13).

3.3.2 Findings from the Northern blots

The Northern blots examined a total of 15 genes to confirm the differential expression between V449E and I374N in FDC-P1 cells. The Northern blots revealed an increased expression of *Ptpmt1* mRNA in V449E cells compared to I374N cells (1.85 fold) (Figure 11 and Table 6B), indicating that the differential expression of this gene may be the consequence of the different signalling properties of the two hβc mutants. The other 14 genes analysed on Northern blots did not exhibit differential gene expression between these two cell lines (Figure 10, Figure 13, Table 6 and Table 9).

In addition, the Northern blots revealed differential expression of *Ptpmt1, Nab1* and *Ddx26b* in FDC-P1 cells expressing αβ. In FDC-P1 αβ cells grown in the presence of hGM-CSF, *Ptpmt1 and Nab1* showed increased expression, while *Ddx26b* showed decreased expression when compared to the cells deprived of hGM-CSF. It is interesting to observe that although the hβc mutants are able to proliferation in the absence of IL-3 or GM-CSF, they do not suppress mRNA levels of *Ddx26b*. This finding may indicate that it is the stimulation of FDC-P1 cells with hGM-CSF that causes the suppression of *Ddx26b* mRNA. The regulation of the gene expression of *Ptpmt1, Nab1* and *Ddx26b* in response to hGM-CSF suggests these genes may be involved in downstream signals generated by GM-CSF signalling through hβc. The possible roles of these three genes in hβc signalling in FDC-P1 cells are discussed below.

The *Ptpmt1* gene was selected for Northern blot analysis because of its protein similarity to the phosphatase domain of the tumour suppressor phosphatase PTEN (Pagliarini *et al.*, 2004). Ptpmt1 is a dual specificity phosphatase, which dephosphorylates both serine and threonine residues, and has a unique substrate preference for phosphatidylinositol 5-phosphate (PI(5)P) *in vitro* (Pagliarini *et al.*, 2004). PI(5)P has been linked to tumour suppression via its regulation of Akt (Pagliarini *et al.*, 2004). Increases in PI(5)P leads to increased Akt activation via the negative regulation of a phosphatidylinositol 3,4,5-trisphosphate ($PI(3,4,5)P_3$) phosphatase by PI(5)P (Carricaburu *et al.*, 2003). Recently Ptpmt1 has been identified as a mitochondrial phosphatase (Pagliarini *et al.*, 2005). In a pancreatic β cell line, Pagliarini

et al., (2005) found that disruption of *Ptpmt1* gene expression caused an increase in ATP production and enhanced insulin secretion (Pagliarini et al., 2005). The findings from this study indicate that Ptpmt1 is a potential drug target in type II diabetes and pancreatic cancer (Boisclair and Tremblay, 2005; Pagliarini *et al.*, 2005). Although the *in vivo* substrate of Ptpmt1 protein has yet to be identified (Pagliarini *et al.*, 2004; Pagliarini *et al.*, 2005), it is interesting to speculate how this phosphatase may differentially regulate the signals generated by V449E compared to I374N and wild type hβc. One possibility is that Ptpmt1 may interact and dephosphorylates a substrate that interacts with, or acts downstream of, V449E induced signalling but is not involved with I374N or hβc signalling. The consequence of Ptpmt1 interacting with V449E may be involved in the leukaemogenic properties of this mutant seen *in vivo* (McCormack and Gonda, 1999). Further studies to investigate whether the Ptpmt1 protein is over-expressed in V449E cells compared to I374N, and whether a specific interaction with V449E could be demonstrated.

The differential regulation of *Nab1* in FDC-P1 αβ cells is of interest as the Nab1 protein is a repressor of the zinc finger transcription factor, Egr1 (Russo *et al.*, 1995). Egr1 activity is involved in monocyte differentiation, where inhibition of *Egr1* leads to granulocytic development and over-expression causes macrophage differentiation (Nguyen *et al.*, 1993). The findings from the Northern blot reveal that the hβc mutants (V449E and I374N) and FDC-P1 αβ cells stimulated with hGM-CSF show increased *Nab1* levels (compared to FDC-P1 αβ cells deprived of hGM-CSF for six hours) (Figure 13 and Table 9). This increase in *Nab1* mRNA levels may lead to an increase in Nab1 protein, which would result in a decrease in Egr1 mediated transcription in FDC-P1 cells.

Ddx26b is a DEAD box protein characterised by the conserved motif Asp-Glu-Ala-Asp/His and like other DEAD box proteins are ATP-dependent RNA helicases (de la Cruz *et al.*, 1999). To date no findings have been reported for Ddx26b (human or mouse) in the literature, but at the amino acid level it shares 63% identity and 74% similarity to the human and mouse DEAD box protein, Integrator complex subunit 6 (INTS6) (also known as DDX26 and DICE1) as determined by a Blast2 alignment (www.ncbi.nih.gov/blast/bl2seq/wblast2.cgi). *DICE1* (*INTS6*) is a tumour suppressor gene reported in a critical region of loss of heterozygosity in non-small cell lung carcinomas (Wieland *et al.*, 2001; Wieland *et al.*, 2004). Ectopic expression of *DICE1*

was shown to inhibit colony formation in non-small cell lung carcinoma cell lines and a prostate carcinoma cell line and had growth suppressing activity in soft agar for an insulin-like growth factor 1 (IGF1) receptor (IGF1R) transformed tumour cell line dependent on IGF1 (Wieland *et al.*, 2004). It remains to be established whether expression of the *Ddx26b* gene has similar effects on GM-CSF signaling for GMR dependent cells such as the FDC-P1 line. If Ddx26b protein has growth suppressing effects on the GM-CSF receptor, then its observed reduced mRNA expression on the Northern blot in the FDC-P1 αβ cell line grown in the presence of hGM-CSF fits with the above data for DICE1 and the IGF1R.

3.3.3 Comparison between image analysis programs

The two software programs for image analysis and data acquisition (Pathways and Spot) used in the subsequent analysis were directly compared by examination of the lists of differentially regulated genes produced by each method. When the list obtained from the first set of hybridisations using Churchill's method on the Spot data was compared to the list obtained from the first set of hybridisations using Chen's method on the Pathways data it revealed eight genes in common (data not shown). It was expected that a greater number of genes would have been found to be differentially expressed by these two methods given that the same data (i.e. the first set of hybridisations) was used for the comparison. While this could be due to the two different statistical methods used, this seems unlikely given that Chen's method (at a 99% confidence interval) produces a similar but less stringent threshold for significance than Churchill's method (using 99% posterior probabilities) (see Figure 12). It was expected that the two lists would have considerable overlap in the genes they each identified as differentially expressed, since all the genes identified by Churchill's method should have been present in the list produced by Chen's method. From the lack of agreement it seems more likely that this discrepancy between Chen's method and Churchill's method is a consequence of the different image analysis programs used (Pathways and Spot). This suggests that using different software programmes for the image analysis of GeneFilters® can lead to differences in the data generated and consequently affect the differentially expressed genes identified by statistical methods. While the results presented in this chapter could not definitively establish which of these programs was superior, they indicate the importance of good background subtraction and spot identification/segmentation

methods in an image analysis package, in order to identify differentially expressed genes.

In the analysis of the GeneFilters® presented in this chapter, it was observed that commercial programs, such as Pathways, had difficulties in adequately removing the effects of nearby spots from a target spot. This frequently led to the identification of false positives (a gene wrongly identified as differentially expressed) (see examples in Table 7) due to the contribution of nearby spots influencing a target spot on one array more than the other. By calculating and correcting for the background contribution to each spot, it has been suggested that a more accurate determination of each spot's intensity could be performed through the use of Spot (Yang *et al.*, 2001a; Yang *et al.*, 2001b; Smyth *et al.*, 2003). Since the results obtained by any statistical method to assess statistically differentially expressed genes will rely on (1) the quality of the data and (2) the number of replicate experiments, it follows that the more reliable the data is to begin with, the better the output or result is likely to be. By comparison to Pathways, the use of Spot software identified fewer false positives, and suggests it would be the most appropriate program for analysing the GeneFilters® and performing appropriate background correction of the data.

3.3.4 Overall success of the approach and future directions

The aim of this work was to identify differentially expressed genes between V499E and I374N. The differential expression of *Ptpmt1* in V449E and I374N expressing FDC-P1 cells seen on Northern blots shows that the analysis of these cell lines using GeneFilters® was able to identify a differentially expressed gene. It was expected that more of the genes identified from the membrane array analysis would be confirmed to be differentially expressed by Northern blots, given that these two cell lines show biochemical and signalling differences (Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000). However, the results from the Northern blots revealed that of 15 genes examined, only one was found to be differentially expressed between V449E and I374N.

There may be several reasons why more genes were not confirmed to be differentially expressed between V449E and I3749 cell lines. Primarily, it may be that only a small number of genes differ in their expression levels between these two cell populations and these genes may not have been represented on the GeneFilters®. Use of larger arrays containing sequences representing the entire mouse genome would confirm

whether there are other genes differentially expressed in these two cell populations that were not present on the GeneFilters® used in this study.

Secondly, the gene expression differences that result in the biochemical and signalling differences between these cell lines may be small in magnitude and therefore difficult to detect using membrane arrays. This was suggested by the Northern blot result with *Ptpmt1* (Figure 10 and Table 6), which was the one sequence confirmed to be regulated between V449E and I374N, and its expression level differed between these cells by 1.85 fold (Table 6B). By performing more replicate experiments, this would provide additional measurements of each gene's expression levels and in turn would provide greater statistical power to identify statistically significant genes with small gene expression differences between the two samples. Replicate experiments require either several GeneFilters® to be used or the reuse of the same GeneFilters® several times. GeneFilters® have been used by others to detect differences in gene expression (de Haan *et al.*, 2002; Saito *et al.*, 2002; Schoondermark-Stolk *et al.*, 2002; Zheng *et al.*, 2002; Brejning *et al.*, 2003; Pang *et al.*, 2003; Wride *et al.*, 2003). However, while Brejning *et al*., (2003) used their arrays up to five times, Schoondermark-Stolk *et al*., (2002) used them only once as they noticed a significant decline in the quality of the filters after stripping. Furthermore, it has been reported that the largest contribution to variation in these membrane array experiments is due to variation arsing from reusing the membranes (Coombes *et al.*, 2002). Therefore, replication of these experiments would perhaps be best performed using several blots from the same batch and hybridising them only once.

Thirdly, while it was predicted that there would be gene expression differences between I374N and V449E mutants, it is possible that there are no other gene expression differences, apart from *Ptpmt1* identified above, that were detectable in FDC-P1 cells. In this scenario, the signalling and biochemical differences that exist between V449E and I374N could be explained by changes in protein levels. These changes in protein levels could occur as a result of post translation modification (such as phosphorylation) that controls the activity of key proteins, or by small gene expression changes in proteins that regulate protein degradation. For example, it has been shown that these mutants differ in the requirement for the alpha subunit and in the level of phosphorylation of the hβc (Jenkins *et al.*, 1995; Jenkins *et al.*, 1999). Therefore, it is probable that in the absence of gene expression changes, one or more proteins that interact with either the alpha subunit (required by I374N) or the phosphorylated tyrosine

residues of the hβc (seen with the V449E mutant), could mediate the differences in signalling of these two mutants.

Finally, the lack of detectable gene expression differences may be due to the choice of cell line used. It is clear that these mutants signal differently, as indicated by studies in other cell lines and *in vivo* studies (Jenkins *et al.*, 1995; McCormack and Gonda, 1997; McCormack and Gonda, 1999). However, in the FDC-P1 cells there are no morphological differences or alterations in their rate of proliferation induced by these two types of hβc mutants. Perhaps, despite upstream differences between V449E and I374N, it may be that these two hβc mutations do not induce differential changes in gene expression within this cell line. It is possible that signalling differences will be observable only in cell lines that give distinct cellular responses to these two mutants. FDC-P1 cells are blocked in their differentiation, so any genes which may alter their differentiation that are differentially expressed between V449E and I374N may not be able to be detected in FDC-P1 cells. Hence, analysis of these mutants in a cell line that is capable of different biological outcomes (e.g. proliferation and differentiation) would be of interest to investigate the gene expression differences between the mutants in such a setting. This idea is addressed in the next chapter where results from a microarray experiment were performed using a cell line that is capable of proliferation and differentiation in response to mutant hβc receptor signalling.

3.3.5 Summary

A membrane array analysis of genes differentially expressed between V449E and I374N expressing FDC-P1 cells was performed. Using a novel approach consisting of multiple image analysis programs and statistical methods, several gene lists of differentially expressed genes were identified. In total, 15 genes were analysed by Northern blots to confirm differential gene expression seen on the membrane arrays. *Ptpmt1* was found to be differentially expressed between V449E and I374N, suggesting this phosphatase may be involved in the different signalling properties of these two hβc mutants. The Northern blots also revealed *Ptpmt1*, *Nab1* and *Ddx26b* to be regulated in response to hGM-CSF in FDC-P1 cells expressing hGM-CSFα and hβc. The differential expression of genes in response to hGM-CSF indicates these genes may be involved in downstream signals generated by GM-CSF signalling through hβc. At the present time it is unclear why so many false positives were identified from the membrane array analysis, even

when sophisticated statistical analyses were used. However, it may indicate that the membrane array method was not sensitive or robust enough to detect the differentially expressed genes, or that there are no gene expression differences between the two classes of hβc mutants when expressed in FDC-P1 cells.

Chapter 4

Microarray Analysis of Gene Expression in FDB-1 Cells Expressing Activated $h\beta_c$ Mutants

4.1 Introduction

4.1.1 Using FDB-1 cells as a model of myeloid proliferation and differentiation

The aim of the work described in this chapter was to examine the signalling differences between the extracellular and transmembrane mutants of hβc by large scale gene profiling. This chapter describes the use of a dynamic cell line model that enables the examination of genes involved in hβc signalling that are associated with the biological processes of proliferation and differentiation. The cell line chosen for the study was the myeloid cell line FDB-1, a factor dependent cell line, which proliferates in response to mouse IL-3 (mIL-3), differentiates in mouse GM-CSF (mGM-CSF) and dies in the absence of either mIL-3 or mGM-CSF (McCormack and Gonda, 2000). The FDB-1 cell line arose from the retroviral infection of murine foetal liver cells with the RufNeo retroviral vector containing the wild type hβc cDNA and grown in long term culture in the presence of mIL-3, mGM-CSF and human Epo (hEpo) (McCormack and Gonda, 1997; McCormack and Gonda, 2000). The cells that arose from this culture were called FDB-1 cells (for factor dependent cell line expressing human beta common)(McCormack, 1998). FDB-1 cells are resistant to the antibiotic Neomycin (G418), which indicates the incorporation of RufNeo-hβc provirus into these cells, and express hβc protein on their cell surface, (McCormack and Gonda, 2000). Southern blot analysis revealed that the FDB-1 cell line contained two copies of the RufNeo hβc provirus (McCormack and Gonda, 2000). Analysis of another cell line (FDB-2) that arose from a similar, but independent, foetal liver culture revealed eight proviral integration sites. Hence, it seems that the RufNeo provirus may be linked to the transforming event in FDB-1 cells. Furthermore, infection with RufNeo provirus lacking hβc cDNA did not generate similar cell lines, suggesting that the transforming event may require hβc cooperation with an activated oncogene (T. Gonda, personal communication). In the presence of mGM-CSF, FDB-1 cells differentiate along the

neutrophil and monocyte lineages with complete differentiation after 5-7 days (see Introduction, Figure 5) while in mIL-3 they continue to proliferate and retain their promyelocyte morphology. In the presence of both mIL-3 and mGM-CSF, the IL-3 response is dominant leading to proliferation of the FDB-1 cells (McCormack and Gonda, 2000). Thus, comparison of the response to IL-3 and GM-CSF in the FDB-1 cell line enables examination of the switch between myeloid proliferation and differentiation in a manipulable cell model.

In FDB-1 cells, expression of the I374N, V449E and FIΔ mutant hβc receptors (Section 1.5.3) leads to factor independence (McCormack and Gonda, 2000). The effects of the hβc mutants are seen upon mIL-3 withdrawal which reveals a differential signalling capacity of the two classes of activated mutants (extracellular and transmembrane). In the absence of exogenously added growth factors (such as mIL3 or mGM-CSF), V449E expression results in continual proliferation of the FDB-1 cells, while expression of either FIΔ or I374N mutant leads to differentiation (similar to mGM-CSF stimulation of the parental FDB-1 cells) to granulocytes and macrophages (see Figure 5) (McCormack and Gonda, 2000). Note that the growth signal (either mIL-3 or V449E) is dominant over the differentiation signal (mGM-CSF or FIΔ) (McCormack and Gonda, 2000; Brown *et al.*, 2004). The differential response to V449E and I374N/FIΔ in FDB-1 cells is consistent with the findings from the biochemical analysis of FDC-P1 cells (Jenkins *et al.*, 1998) (see Section 1.5.3.3 and Table 5). It has been concluded that the extracellular and transmembrane hβc mutants deliver distinct intracellular signals leading to the different biological outcomes (reviewed in D'Andrea and Gonda, 2000).

To confirm the expression of the activated forms of hβc introduced into the FDB-1 cell line, each of the mutants (V449E, I374N and FIΔ) was engineered to express the FLAG epitope at their N-terminus (McCormack and Gonda, 2000; Brown *et al.*, 2004). This enabled expression of the hβc mutants to be distinguished from RufNeo mediated expression of the wild type hβc in FDB-1 cells and showed that FLAG-tagged versions of V449E and FIΔ were expressed at equivalent levels (McCormack and Gonda, 2000). However, the addition of the FLAG-epitope to the I374N mutant disrupted its ability to cause factor independent growth in FDB-1 cells, despite activity of the FLAG-tagged I374N mutant seen in FDC-P1 cells (McCormack, 1998). Thus, the expression level of the I374N mutant could not be compared with the other mutants

(FIΔ or V449E) in FDB-1 cells by flow cytometry and hence, this mutant was not examined further in the FDB-1 cell lines. As a consequence of an inability of the FLAG-tagged I374N mutant to function in FDB-1 cells, the FIΔ mutant was utilised as an example of an extracellular hβc mutant which was compared to the transmembrane mutant V449E in the microarray experiments reported in this chapter (McCormack and Gonda, 2000; Brown *et al.*, 2004).

The hypothesis is that comparisons between FDB-1 cells expressing the two hβc mutants will provide insights into the differentiation-associated events induced by the FIΔ mutant and the proliferative signalling and block in differentiation supported by the V449E hβc mutant. Thus, the aim of the study described in this chapter was to examine the gene expression profiles associated with the two classes of activating hβc mutations (extracellular and transmembrane) using the V449E and FIΔ hβc mutants using the NIA 15K cDNA microarrays (see Section 4.1.2). Genes that are differentially expressed are predicted to have biological activity influencing differentiation, survival and proliferation. Furthermore, changes associated with the V449E mutant may contribute to its potent leukaemia inducing activity (McCormack and Gonda, 1997; McCormack and Gonda, 1999) and therefore, this may be an approach to identifying myeloid oncogenes.

4.1.2 Microarray analysis

In order to examine the gene expression differences between the FDB-1 cells expressing the two mutant hβcs a microarray gene profiling approach was utilised. In this study, glass microarrays were chosen over the previously used membrane arrays (described in Chapter 3) as they offered a number of benefits (see Section 1.6.1) including their ability to simultaneously analyse two difference samples at once, improved image acquisition and analysis, and increased number of genes per array (Cheung *et al.*, 1999; Southern *et al.*, 1999; Holloway *et al.*, 2002). The microarrays were spotted with a complementary DNA (cDNA) clone set of fifteen thousand clones from the National Institute of Aging (NIA) (National Institutes of Health, USA) which from hereon will be referred to as the NIA 15K cDNA microarrays (Tanaka *et al.*, 2000). A cDNA clone set was chosen because at the onset of these experiments cDNA clone set were more available, more extensive in their gene coverage and were also more cost effective based on the number of genes per library than the emerging oligonucleotide clone sets. The NIA 15K clone set also contained a large number of novel cDNAs (78% novel and

22% known of the 15264 clones) and therefore, offered the potential for the discovery of previously uncharacterised genes (Tanaka *et al.*, 2000).

4.1.3 Experimental design

The design of a microarray experiment can have significant impact on the robustness of the data produced. Thus, advice was sought from statisticians for the design and analysis of this experiment and this section summarises the design issues considered. The experimental design was developed together with Associate Professor P. Solomon and Dr. G. Glonek (Applied Mathematics Department, Adelaide University) with data analysis performed by Dr. A. Tyskin, Mr. J. Lloyd and Dr. C. Wilkinson.

It was decided to concentrate on changes in gene expression in the 24 hour time period following mIL-3 withdrawal. It was predicted that this would reveal differential changes after mIL-3 withdrawal between the two forms of mutant hβc. The 24 hour time point was selected as the primary goal was to identify the transcriptional targets of receptor-associated signalling pathways. For example, it was predicted that in FIΔ cells the use of later time points may have biased the selection of differentially expressed genes to those events downstream of the early changes initiating differentiation. Earlier time points were not chosen as it was reasoned that there may have not been sufficient time (after e.g. 6 to 12 hours) for the differentiation programme to be initiated after switching over from the IL-3 proliferative signal. While a time-course experiment would have overcome some of these issues, other time points, in addition to time 0 and 24 hours, could not be investigated due to constraints on the number of microarray slides that were available for use in this experiment.

It was of interest to identify the genes differentially expressed between the two samples (i.e. the sample effect α) and those genes differentially expressed over time in the differentiating FIΔ cells (i.e. the time effect β) (see Figure 15). However, the primary aim was to identify genes displaying a significant interaction effect (γ) which represents the interaction of sample and time. Genes with a significant γ are those that changed significantly over time and between the two cell lines. Examination of γ was expected to identify those genes associated with the proliferation of V449E and the differentiation of FIΔ.

A common approach for microarray experiments is to use a reference design. In this design (shown in Figure 14) all samples are directly compared to a reference population, represented by μ, which in this experiment would be FIΔ at time zero. In the

reference design, six slides are used to perform three “dye swap” comparisons represented by the double arrows between each sample in Figure 14. A “dye swap” involves two samples being labelled with Cy3 or Cy5 and used to probe a microarray and then repeating the process with the samples labelled with the reciprocal dye on another microarray. The reference design examines the effect due to two factors (α and β) by comparing the samples back to a common sample called the reference sample. A limitiation of this design is that it does not easily permit the measurement of γ (Glonek and Solomon, 2004) (see Discussion, Section 4.3.1.1). Therefore, to measure γ, a similar design was adopted to the one proposed by Speed, (2001) where all six possible pairwise comparisons between the four conditions (FIΔ at times 0 and 24 hours and V449E at times 0 and 24 hours) are performed (see Figure 15A). For logistical reasons the microarray experiment was limited to eight slides, so the remaining two slides were used to perform dye swaps on the most critical comparisons – the cell line comparisons between FIΔ and V449E at either time zero or after 24 hours (Figure 15B). The resulting design that was used is represented in Figure 15B, with the eight arrows representing the different microarray hybridisations that were performed with the two samples (V449E and FIΔ) over the two time points (0 and 24 hours). The presence of double arrows at the 0 and 24 hour time points represents dye swap hybridisations.

In statistical terms, the design shown in Figure 15 corresponds to a 2 x 2 factorial analysis, consisting of two factors – sample and time, each with two levels. This design maximises the statistical power of the analysis using a total of eight slides and gives direct and indirect ways of estimating effects due to cell type (α), time (β) and the interaction of cell type and time (γ). The design achieves balance, replication and an estimate of dye-biases present in the microarray experiment.

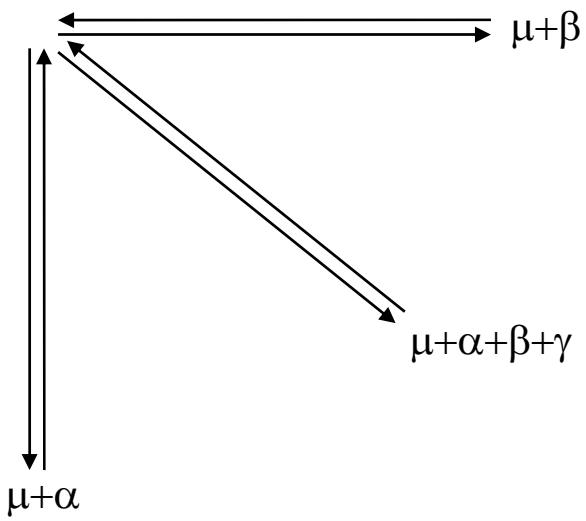


Figure 14: The common reference design showing six comparisons between two factors where three dye swap comparisons are performed. μ represents the ground state or base level of the sample that is used as a common reference for comparisons to the other samples and as we are only interested in changes, no attempt is made to measure μ. α is the effect due to one factor (e.g. sample), β is the effect due to a second factor (e.g. time). γ is the interaction parameter, which is the effect due to both factors α and β (also written as "αβ").

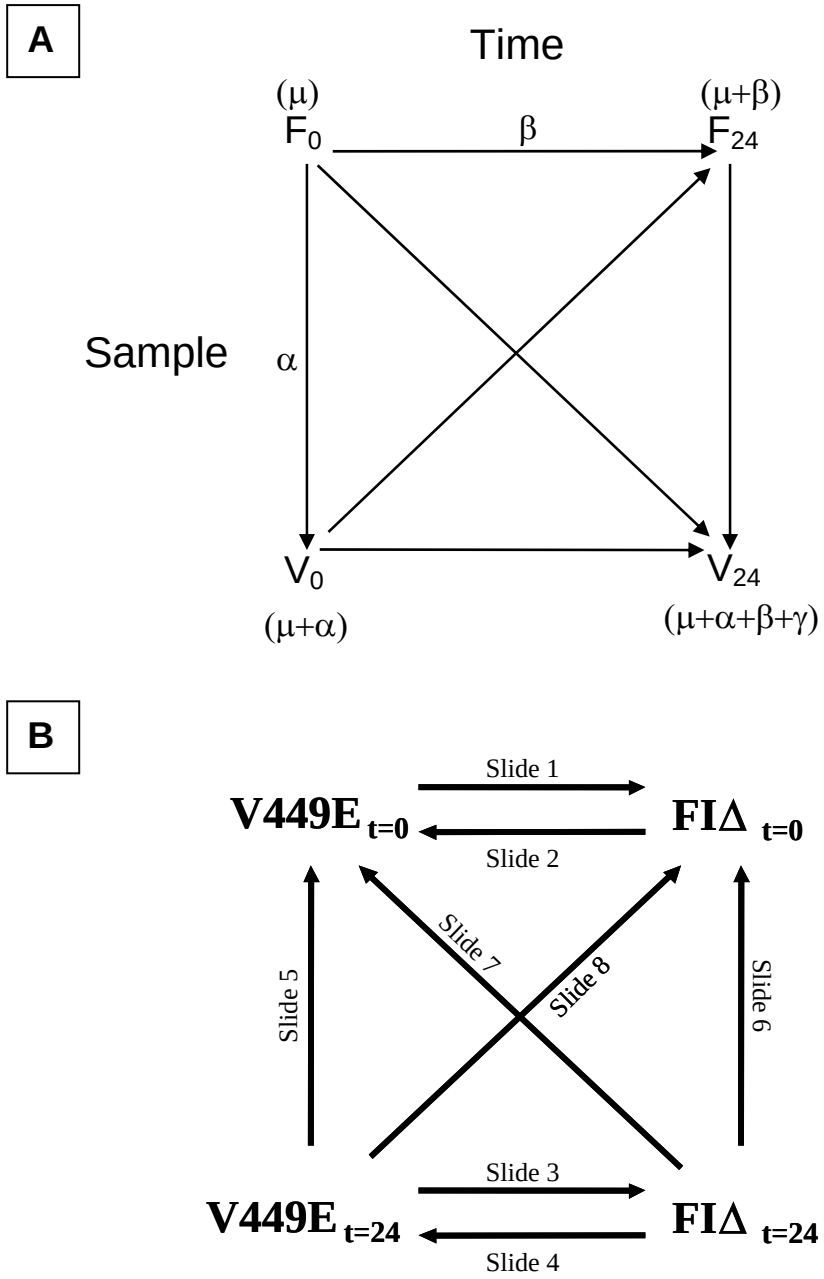


Figure 15: Diagrammatic representation of the microarray experimental design. (A) The two samples FIΔ ("F") and V449E ("V") and two time points t=0 and t=24 hours (subscript 0 and 24 respectively) combine to give a 2 x 2 factorial design. The four sample and time configurations and the six possible pairwise hybridisations are shown where α is the sample effect, β is the time effect and γ represents the interaction effect. **(B)** Shows the eight slides and hybridisations that were performed for the microarray experiment. The arrow-heads indicate the sample that was labelled with the red Cy5 dye, and the arrow-tails indicate the green Cy3 labelled samples. The double arrows at the 0 and 24 hour time points represent dye swap hybridisations.

4.1.4 Applying a linear model to analyse the microarray data

A linear modelling approach that combined the normalised log ratios of the eight arrays into information for a single clone was implemented. The terms defined in Figure 14 were applied to Figure 15A such that F_0 represents μ, V_0 represents $\mu+\alpha$, F_{24} represents $\mu+\beta$, and V_{24} represents $\mu+\alpha+\beta+\gamma$. Table 10 describes the expected log ratios for each hybridisation, including the two crossover hybridisations.

A design matrix was created such that the contribution of each pairwise hybridisation to estimating the parameters α, β and γ was explicitly stated. The three parameters can be linked to the microarray analysis as follows, $\alpha = V_0-F_0$, $\beta= F_{24}-F_0$ and $\gamma = (F_{24}-F_0)-(V_{24}-V_0)$. The linear model was fitted to each gene to estimate α, β and γ, which was tested against a null hypothesis (H_0) of $\alpha=0$ or $\beta=0$ or $\gamma=0$. Each gene was then measured for the significance of its expression in terms of α, β and γ in order to determine whether it was differentially expressed.

	Configuration		
	Green	Red	Log Ratio
1	F_0	V_0	α
2	F_0	F_{24}	β
3	F_0	V_{24}	$\alpha+\beta+\gamma$
4	F_{24}	V_{24}	$\alpha+\gamma$
5	V_0	V_{24}	$\beta+\gamma$
6	V_0	F_{24}	$\beta-\alpha$
7	V_0	F_0	$-\alpha$
8	V_{24}	F_{24}	$-\alpha-\gamma$

Table 10: Log ratio M = log(R/G) for each configuration from the 2 x 2 factorial design using eight slides. Alpha (α) is the cell line effect, beta (β) is the time effect and gamma (γ) is the interaction parameter, which represents the effect due to both sample and time. Slides 7 and 8 represent the dye swap hybridisations of slides 1 and 4 respectively.

4.2 Results

4.2.1 Preparation of the FDB-1 samples for microarray analysis

4.2.1.1 Treatment of cells

FDB-1 cells expressing either the FLAG-V449E hβc mutant (FDB-1 V449E) or FLAG-FIΔ hβc mutant (FDB-1 FIΔ) (Brown *et al.*, 2004) were grown in the presence of mouse IL-3 (mIL-3) (and absence of mGM-CSF) to induce proliferation and suppress the differentiation of FDB-1 FIΔ cells. Cells were washed three times and grown in medium without mIL3 or mGM-CSF to allow the transition to growth independent of exogenously added growth factors. Control cultures of parental FDB-1 cells were similarly washed and deprived of mIL3, to ensure insufficient levels of mIL3 remained to support growth of the factor dependent parental FDB-1 cells. Some of the washed parental FDB-1 cells were stimulated with mIL3 or mGM-CSF to monitor their proliferation/differentiation. Cytospins were performed on all cell populations, daily for five days (data not shown), so that the morphology of the cells could be examined microscopically as a way of confirming their response in the presence or absence of growth factors (mIL3 or mGM-CSF). The expression of each hβc mutant in the FDB-1 cells was confirmed by flow cytometry using the anti-FLAG antibody (Figure 16). For RNA preparation, samples were collected at time zero (in the presence of mIL-3) and after twenty-four hours in the absence of mIL-3 i.e. of mutant hβc signalling.

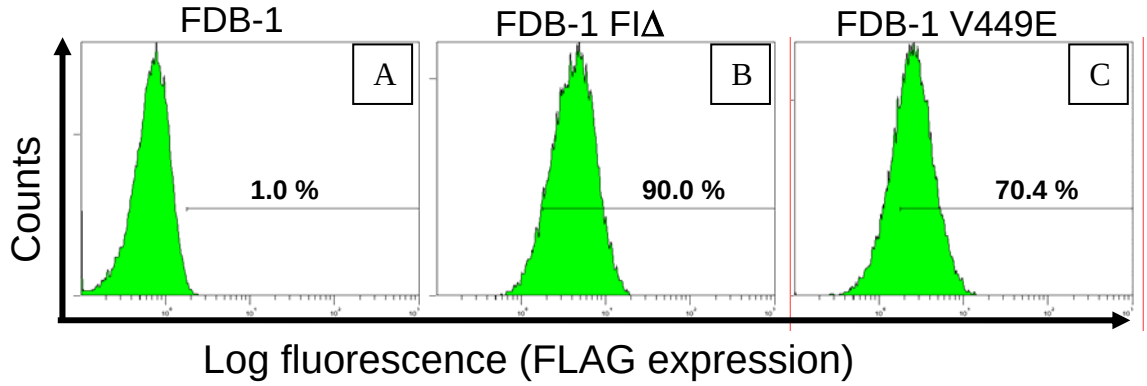


Figure 16: Cell surface expression of mutant hβc (FIΔ or V449E) in FDB-1 cells. Flow cytometric analysis showing cell surface expression of the indicated FLAG-tagged hβc proteins on FDB-1 cells expressing either the extracellular mutant, FIΔ (B) or transmembrane mutant, V449E (C) compared to uninfected parental cells (A). Infected cells (panels B and C) were stained with an anti-FLAG monoclonal antibody and compared with similarly stained uninfected cells (panel A). In each panel, the x-axis scale for the major graduations is 10^0, 10^1, 10^2 and 10^3.

4.2.1.2 RNA preparation, probe labelling and microarray hybridisation

Total RNA was isolated from the cells via a phenol-chloroform extraction method which used TRIzol and RNeasy midi columns (see Section 2.6.2) to produce RNA of sufficient quality to perform microarray experiments. RNA size and quality was assessed via UV illuminated agarose gel electrophoresis and spectrophotometery (with all samples having an OD 260/280 ratio exceeding 2.0, indicative of low protein contamination in the samples). 50μg of total RNA was reverse transcribed into cDNA using SuperScript II while incorporating either the Cy3 or Cy5 fluorophores by the indirect method of aminoallyl reverse transcription labelling (http://www.microarrays.org/pdfs/amino-allyl-protocol.pdf). The 16 labelled samples were then paired (Cy3 and Cy5), mixed and used to probe eight NIA 15K cDNA glass microarrays in the configuration shown in Figure 15. Prior to the hybridisations, two of the combined Cy3 and Cy5 labelled samples were analysed via agarose gel and imaged using a Molecular Imager FX Pro (Bio-Rad) to confirm the quality of the labelled cDNA (Figure 17). This revealed that both cDNA samples had been successfully labelled and that the labelled cDNA ranged in size from 500 bp up to 12 kb (Figure 17). Hybridisations were performed overnight (14-16 hours) in a humidified chamber at 42°C. Slides were washed and scanned to visualise and quantify the binding of the probes to the arrays. Three slides had large background effects (where hybridisation of the probe occurred to non DNA spotted regions of the chip) covering an area greater than 4 supergrids and were repeated with the same RNA samples.

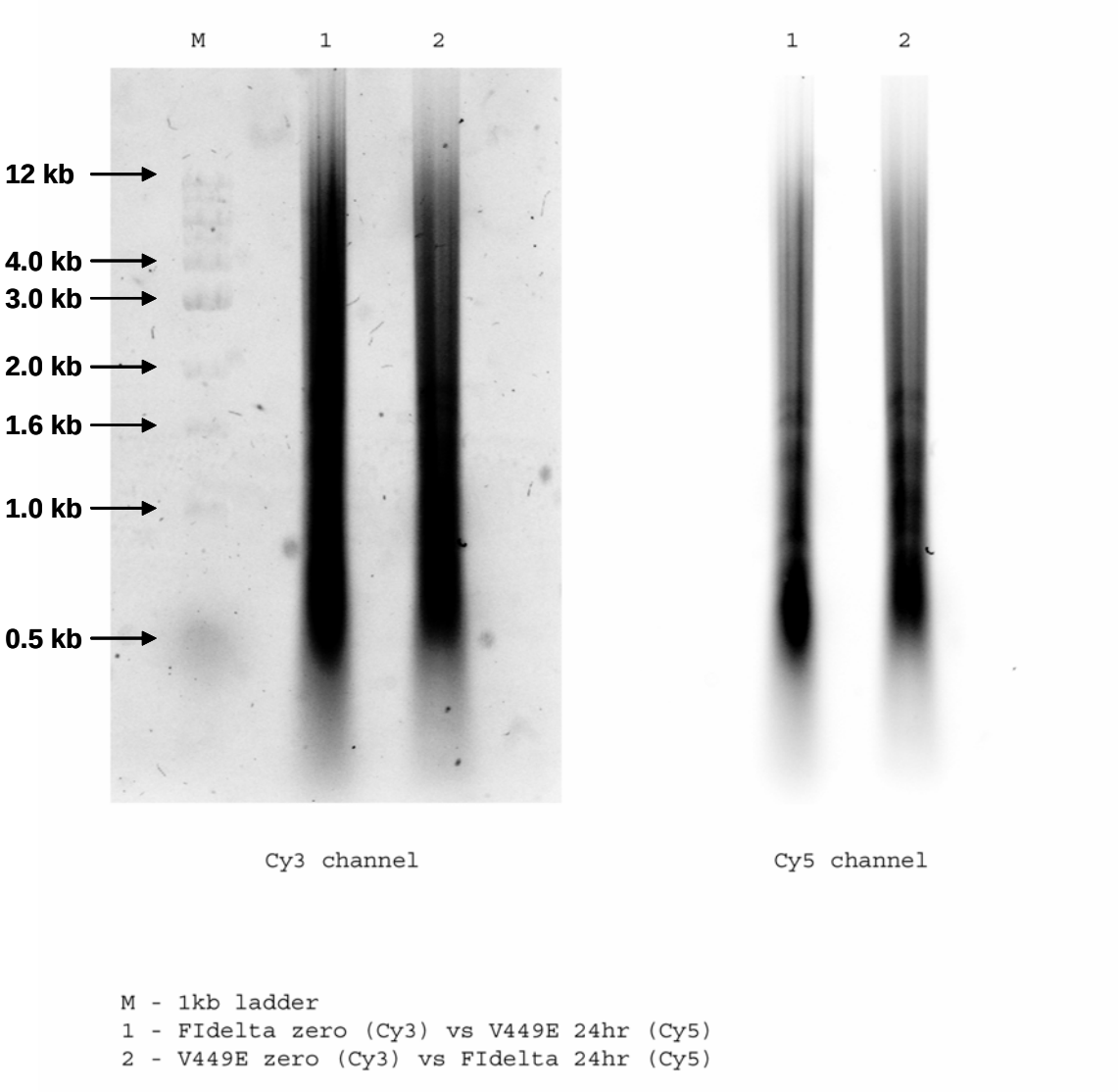


Figure 17: Indirect labelling of cDNA via the aminoallyl reverse transcription labelling process. SuperScript II reverse transcribed RNA samples were labelled with Cy5 or Cy3 and the labelled cDNA was separated via agarose gel electrophoresis and scanned at 532 nm (Cy3 channel) and 635nm (Cy5 channel) using a Molecular Imager FX Pro (Bio-Rad).

4.2.1.3 Scanning of the microarray slides

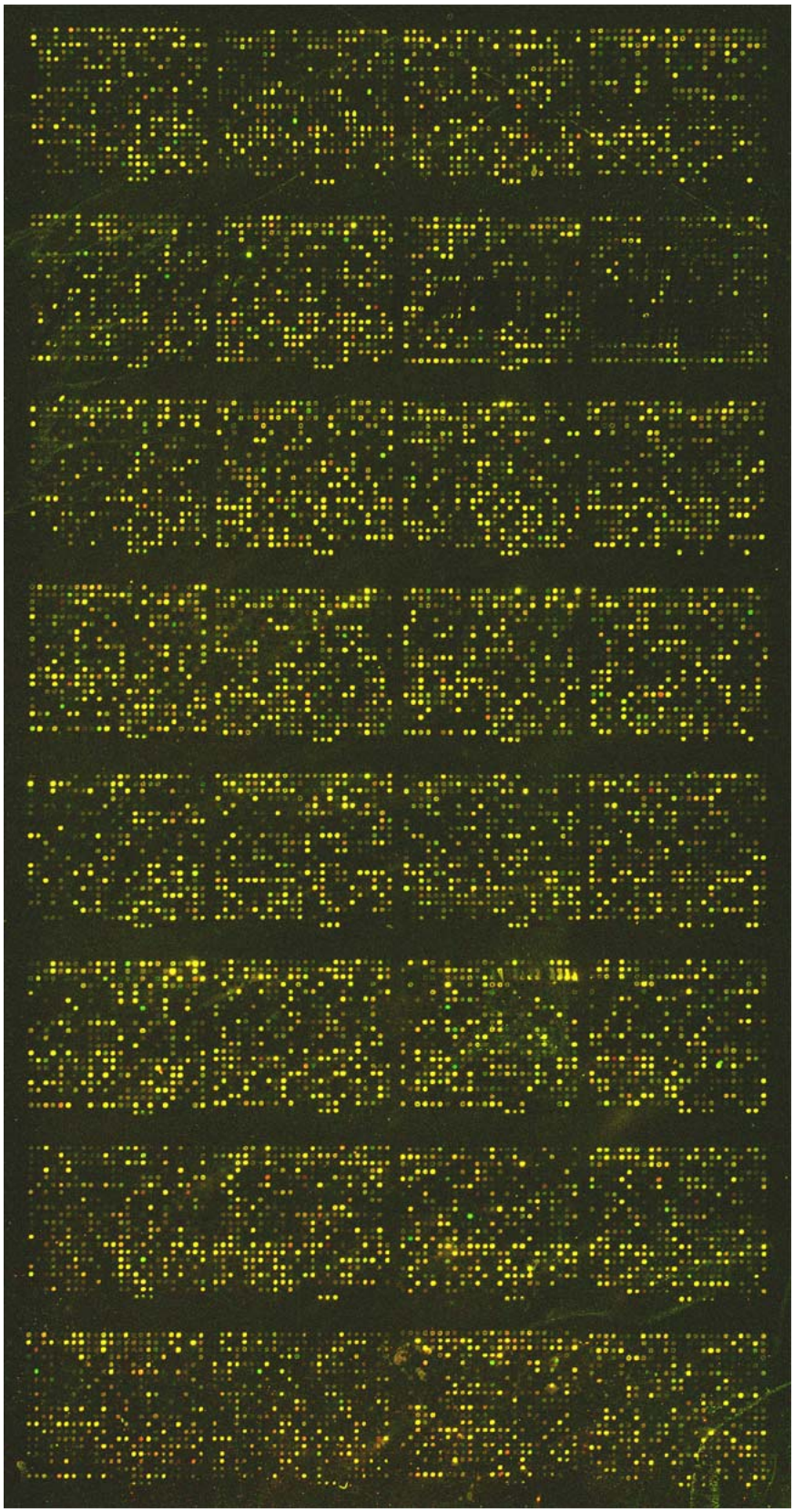
(a) Initial scanning performed at the Peter MacCallum Cancer Institute

The hybridisation of the probes to the NIA 15K cDNA microarrays (Section 4.2.1.2) and the scanning of the slides were performed at the Peter MacCallum Cancer Institute (PMCI), Melbourne. The slides were scanned using a ScanArray 5000 XL (Perkin-Elmer) microarray scanner from Packard Biochip Technologies that generated Cy3 and Cy5 images that were combined into a single image for each slide. The result for a representative slide is shown in Figure 18, revealing the two channels, Cy3 and Cy5, combined into a single image. This image shows variations of colour from the red of the Cy3, to the green of the Cy5, with the appearance of yellow regions in locations where both dyes have bound to the same spot on the array in equal proportions. All slides were visually assessed for quality (uniform hybridisation and the absence of large artefacts) and any slides with gross defects were subsequently repeated. Based on these criteria, the slides numbered 6, 7 and 8 (see Figure 15B for slide numbering) were repeated.

(b) Subsequent scanning performed at the Adelaide Microarray Facility

The microarrays were also scanned using the Adelaide Microarray Facility GenePix 4000B scanner from Axon Instruments Inc. (CA). As variations in the scanners and the software can impart subtle differences on the resulting images, even when scanning the same slide (Yang *et al.*, 2002), it was decided to use both scans for subsequent data analysis. Each set of scanned slides (PMCI and Adelaide) were treated as a separate experiment and analysed separately in the subsequent sections.

Figure 18: A representative image of the combined Cy3 and Cy5 channels from one of the eight hybridised slides from the microarray experiment. Shown is the hybridisation of the Cy3 labelled FIΔ time 24h sample with the Cy5 labelled V449E time 24h sample (slide 3 or "24hour1"). The presence of red and green spots on the microarray is indicative of differentially expressed genes, while yellow spots suggest similar expression levels in both samples.



4.2.1.4 Intensity extraction using Spot software

Information extraction from the microarray hybridisations was performed using *Spot* software (Buckley, 2000). Using this software, the foreground and background intensity (in terms of red and green (R,G) intensity pairs) were obtained for all spots on the arrays. As mentioned in Chapter 3, Spot has a number of advantages over many other commercial image analysis packages available. The key advantages of Spot are better identification of spot edges and the use of a background subtraction method that is less variable than other methods (Yang *et al.*, 2001a; Yang *et al.*, 2001b). For segmentation, Spot uses adaptive shape segmentation via the seeded region growing (SRG) algorithm, which enables irregular and non circular shaped spots to be accurately defined. Background values are estimated using a non-linear filtering technique known as morphological opening (Soille, 1999). The background estimates are lower, less variable and typically do not produce background values greater than foreground values, unlike many other programs (Yang *et al.*, 2001a; Yang *et al.*, 2001b). The values obtained from Spot were saved and then imported into the software package R (Ihaka and Gentleman, 1996) for normalisation.

4.2.2 Normalisation of the microarray data

The purpose of normalisation is to remove non-biological sources of variation. In a microarray experiment, there are many sources of variation. These include dye biases from the efficiency of dye incorporation, experimental variability in the hybridisation and processing steps and differences between experimental conditions for replicate slides. Within- and multiple-slide normalisation as described in the proceeding sections were undertaken using the statistical programming environment R (Ihaka and Gentleman, 1996) and SMA (Statistics for Microarray Analysis) package, available from http://www.R-project.org and from Professor Terry Speed's group at http://www.stat.berkley.edu/users/terry/zarray/Html/index.html. The data was then plotted using M versus A plots (described in Section 3.2.3.4) to show the log-intensity ratio ($M=\log_2 R/G$) versus the mean log intensity ($A=\log_2\sqrt{(R \times G)}$) for all the genes present on the array (see Figure 19). Both sets of scans (PMCI and Adelaide) were normalised using identical procedures (see proceeding sections and Figure 21 and Figure 24).

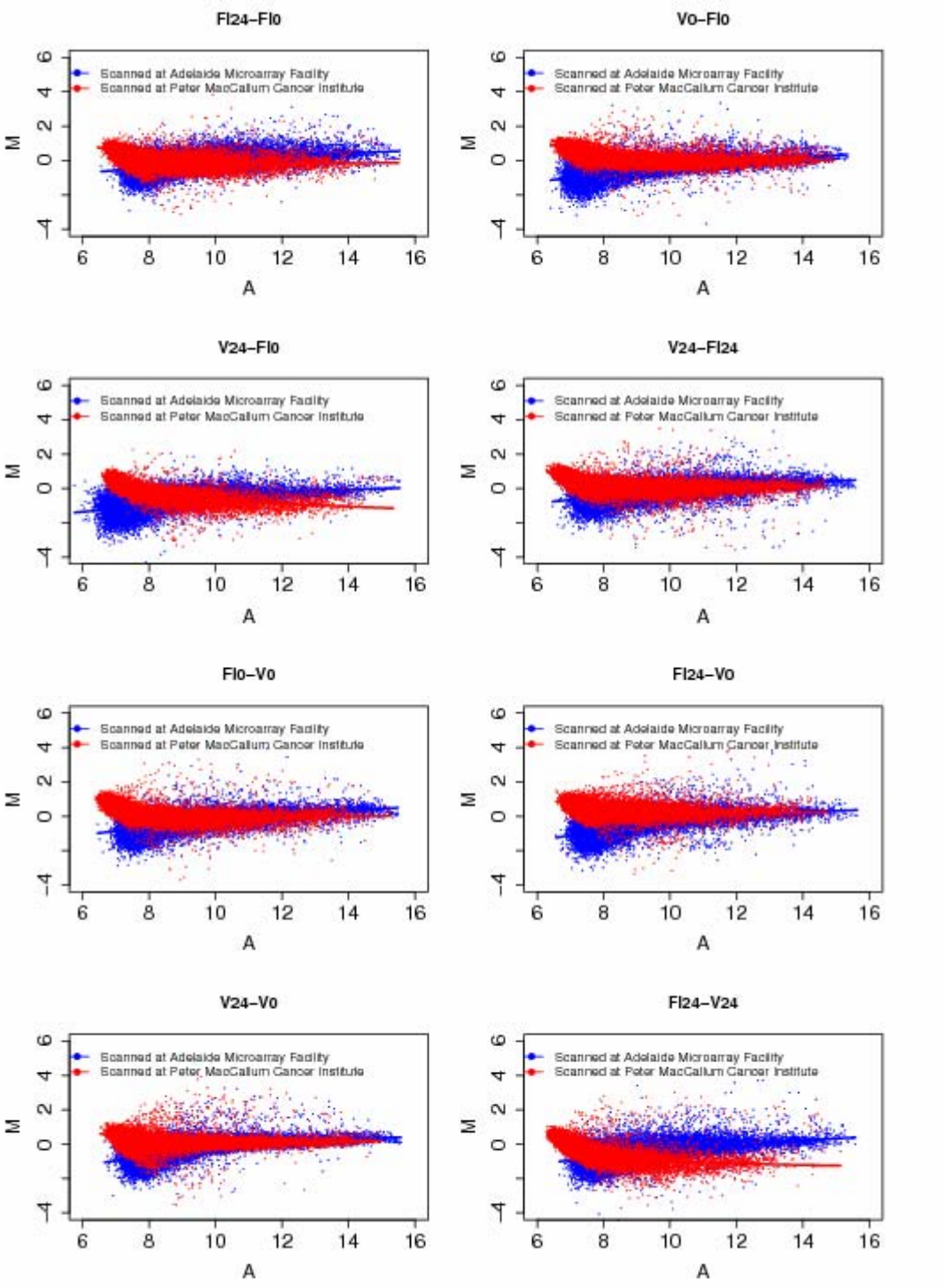


Figure 19: M versus A plots showing all eight slides prior to Lowess normalisation. For each M versus A plot, blue dots indicate Adelaide (ADL) scanned data and the red dots indicate PMCI (PM) scanned data all from the same slide. The blue and red lines in the middle of the data points are the global lowess curves for the Adelaide and PMCI data respectively for each slide. The lowess curve defines the centre of the data without being greatly affected by outliers. Abbreviations used: FI0 = FIΔ at time 0, FI24 = FIΔ at time 24h, V0 = V449E at time 0 and V24 = V449E at time 24h.

4.2.3 Lowess normalisation

The data from each slide was normalised using "lowess", which is a robust scatter plot smoother (shown in Figure 19) that removes any intensity dependent dye biases present in each of the hybridisations (Yang *et al.*, 2001c; Yang *et al.*, 2002). Compared to a global median normalisation approach, which assumes a uniform grading of systematic error across all variables in an experiment, lowess normalisation removes global dye biases in an intensity dependent way (Yang *et al.*, 2001c; Yang *et al.*, 2002). Lowess can therefore correct for intensity dependent effects such as, a strong Cy3 bias at low intensities but not at higher intensities, without overcorrecting the other data. However, lowess normalisation relies on the assumption that only a small percentage of the genes are differentially expressed and that there is not a global change across all of the genes (Yang *et al.*, 2001c; Yang *et al.*, 2002). An example of global lowess correction is shown in Figure 20 where prior to normalisation (A) an intensity dependent dye bias is evident that is removed following lowess normalisation (B).

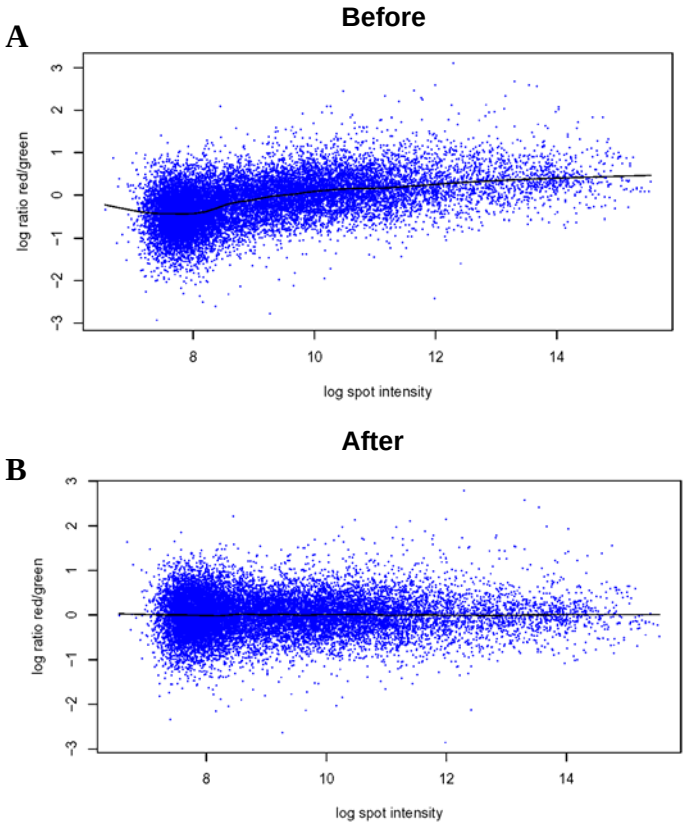


Figure 20: M versus A plot showing (A) before and (B) after global lowess normalisation of one slide from the microarray experiment. The black line in the middle is the lowess curve, which defines the centre of the data without being greatly affected by outliers. Normalisation makes the lowess curve the new zero line.

4.2.3.1 Within-slide normalisation - print-tip normalisation

Since dye biases are frequently not constant across an array and can vary with the spot intensity and location on the array (Yang *et al.*, 2001c; Yang *et al.*, 2002), a lowess normalisation for each print-tip group was performed to normalise the microarray data presented in this chapter. An example of this is shown in Figure 21 where each of the coloured lines represent the individual print-tip group lowess curve before and after normalisation. Print-tip group lowess normalisation is recommended for cDNA microarrays especially when expression changes are expected for only a small percentage of the genes (Yang *et al.*, 2001c; Yang *et al.*, 2002; Smyth and Speed, 2003). The resulting normalisation effectively becomes an intensity and spatial normalisation process, removing any spatial effect due to print-tip variations at the time of spotting the array. In addition, it reduces local effects or artefacts on the slides, usually caused by uneven hybridisation or drying out of the probe. Print-tip group lowess involves separately normalising all the genes contained in each print tip group (504 genes in 32 supergrids) to remove the spatial effect from that grid on the array. The effect of this normalisation process is shown in Figure 22 where the effect of an obvious spatial effect is removed from a particularly affected slide. This normalisation approach was performed on the data from each hybridisation.

The M versus A plot for each slide and each scan prior to normalisation is shown in Figure 19 with their global lowess curves. In Figure 23 the microarray data is represented as box plots displaying the intensity log ratio distributions for all slides and for both scans prior to normalisation. Together these two figures allow direct comparisons between the Adelaide scans and PMCI scans. The box plots illustrate the differences between the two data sets in terms of their median values and the range that the M values span for each slide. The M versus A plots show the differences in the intensities (A) and ratio (M) obtained from the two scans of each slide. In Figure 19 it can be seen that the PMCI scanned slides have higher M values at the lower intensity range (approximately 6-8). In Figure 23 the box plots show that the mean intensity of the PMCI scanned data is higher than the Adelaide scanned data for each matched slide. Figure 19 and Figure 23 reveal that the two different scanners (and their settings) had differing effects on the data obtained from identical microarray slides. Thus, it was decided to analyse each set of scans separately (Adelaide and PMCI) rather than pooling the data from both scans together. In order to combine the data from each scan, it was

necessary to apply further normalisation so that the eight slides could be compared directly (see below).

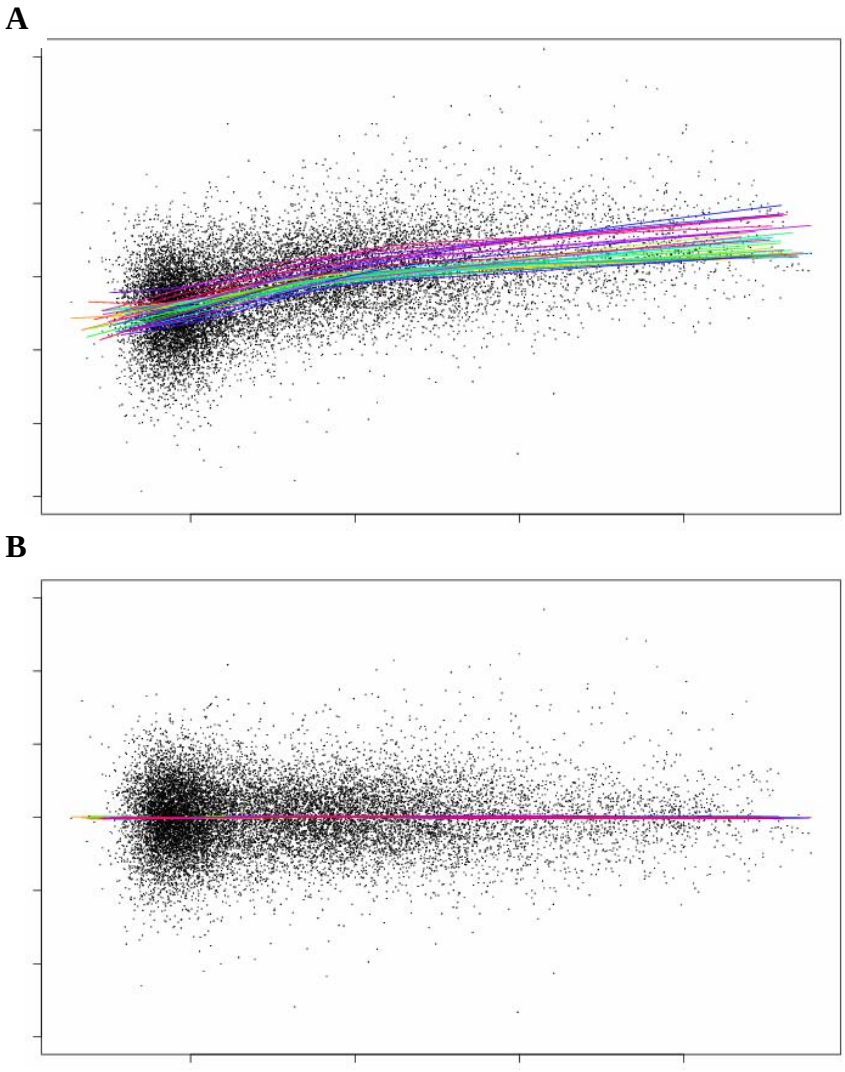


Figure 21: M versus A plot showing (A) before and (B) after print-tip lowess normalisation of one slide from the microarray experiment. The coloured lines in the middle represent each of the print-tip group lowess curves. Normalisation makes the lowess curve the new zero line and after normalisation all the individual print-tip lowess curves now line up on top of one another.

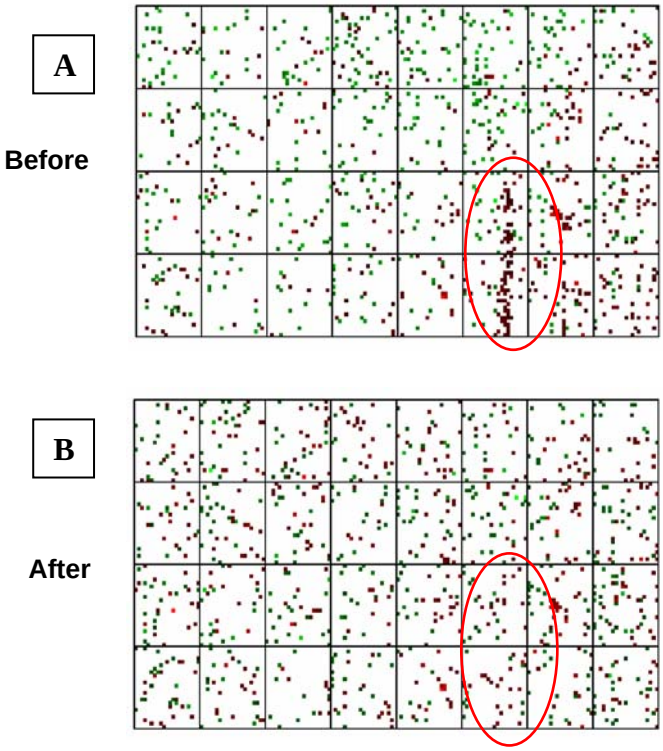


Figure 22: Spatial normalisation via within–slide print-tip normalisation. Hybridisations are seldom perfectly even. To overcome this, a lowess normalisation was performed for each print tip group (supergrid) separately and if necessary, adjusted each group to bring them to the same scale. The coloured dots represent genes that are either over- or under-expressed. The removal of a spatial artefact is shown in the region circled in red, where the cluster of red spots in the bottom two grids three from the right hand bottom corner in (A) are removed in (B).

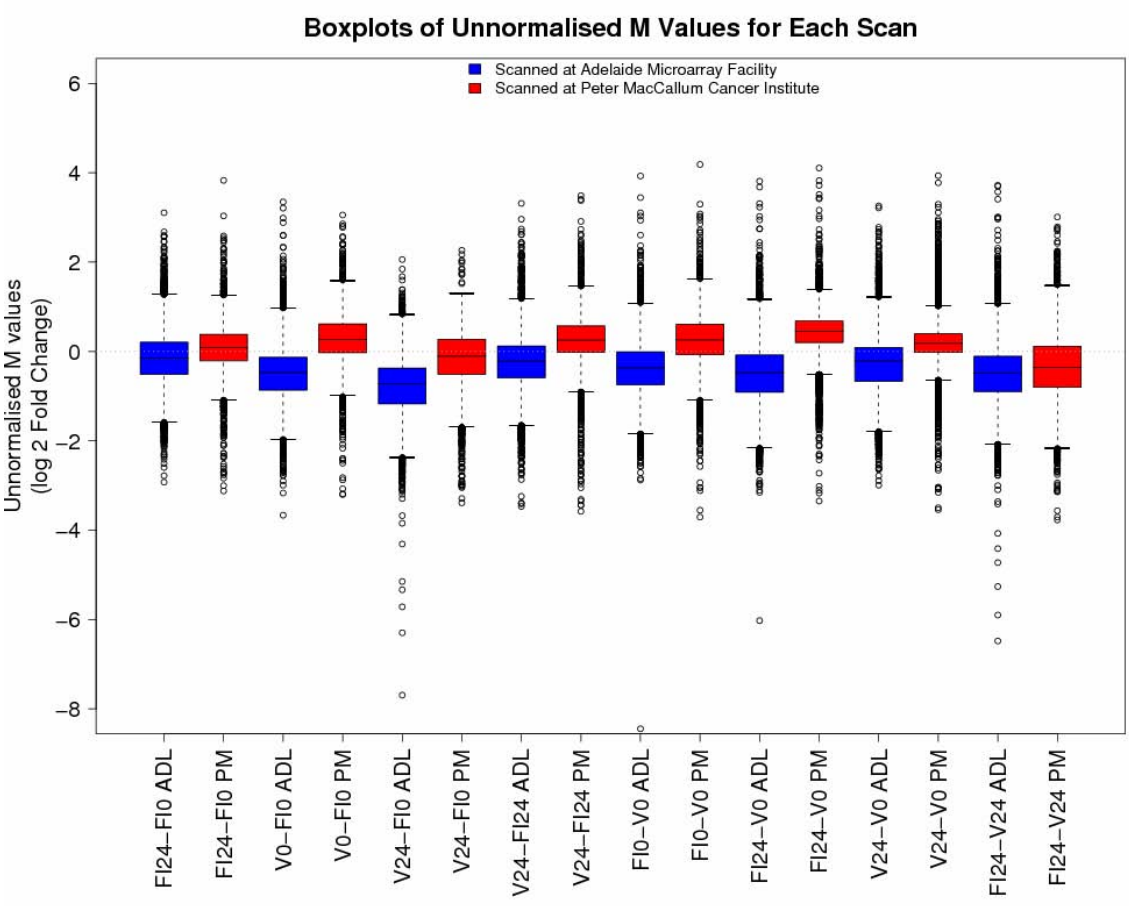


Figure 23: Box plots displaying the intensity log ratio distributions for the eight different slides prior to normalisation comparing the Adelaide scans (blue) and PMCI scans (red) for each slide. A box plot displays graphically the five-number summary of a set of numbers involving the three quartiles and the maximum and minimum values. The central box of the box plot extends from the first to the third quartile and thus encompasses the middle 50% of the data. For each of the eight slides, the data obtained from the Adelaide scan (ADL) is presented in blue alongside the PMCI scan (PM) shown in red, to give a total of 16 sets of data. Abbreviations used: FI0 = FIΔ at time 0, FI24 = FIΔ at time 24h, V0 = V449E at time 0 and V24 = V449E at time 24h.

4.2.3.2 Between-array scale normalisation

The final stage in the normalisation process, involves scaling the data so that for each slide the majority of the data covers the same range of values thereby allowing direct comparisons between slides. As stated for print-tip lowess normalisation, between-array scale normalisation also relies on the assumption that most genes are not differentially expressed in any of the slides. Failure to correct for between-array scale normalisation may result in a given slide (or slides) having an increased weight when averaging the log ratios across all slides, which would skew the results. Between-array scale normalisation is shown in Figure 24(iii) using box plots that show the intensity log ratio distribution for each of the eight slides comparing the scale normalised Adelaide and the PMCI data side by side. Figure 24 summarises the entire normalisation process by showing the data from each slide prior to normalisation (i), after print-tip lowess normalisation (ii) (Section 4.2.3.1), and after between-array scale normalisation (iii), revealing the before and after effects of normalisation at each step. After scale normalisation, all the slides now have their log ratios centred on zero and the whiskers marking the first and third quartiles for each slide in the box plots now span the same range of M values. This indicates that for each slide the majority of the data now falls within the same range of values across all slides. Between-array scale normalisation was important to enable the analysis of all eight slides at once using the linear modelling approach described in Section 4.1.4.

In summary, two normalisation methods were used on the data obtained from the microarray experiments. First, any spatial effects or dye biases present on the slides were removed via print-tip group lowess normalisation (Section 4.2.3.1). Second, between-array scale normalisation (Section 4.2.3.2) was used to enable comparisons betwen the slides. This two step normalisation process enabled the data (for a given scanner e.g. Adelaide) obtained from the eight slides to be combined and analysed using the linear model (described in Section 4.1.4).

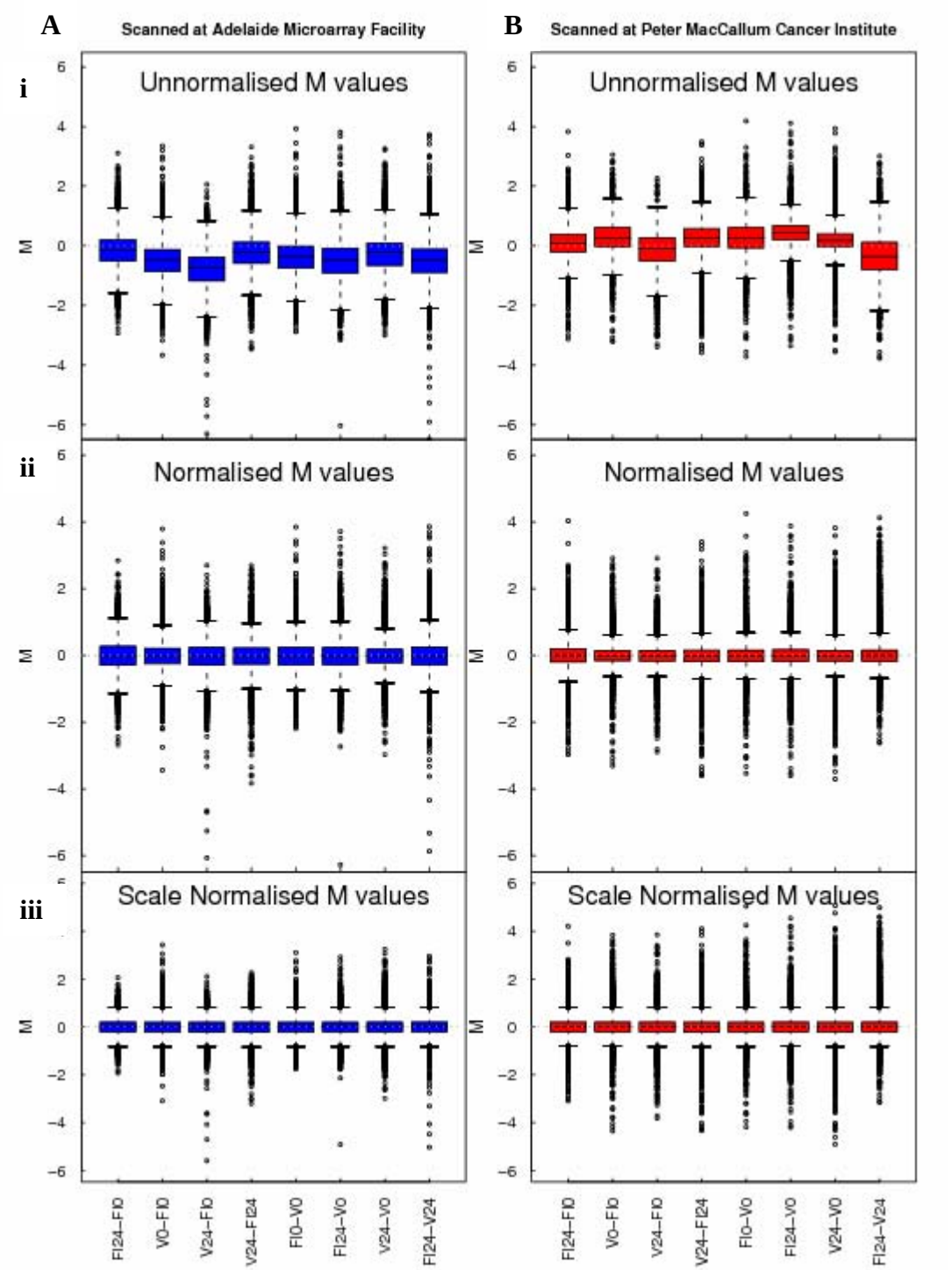


Figure 24: Box plots displaying the intensity log ratio distributions for the 8 different slides of the microarray experiment. (A) Shows the Adelaide scanned gene ratios in three separate panels for (i) prior to normalisation, (ii) after within-print tip group lowess normalisation and (iii) after between-array scale normalisation for each microarray slide. (B) Shows the same as (A) for the PMCI scanned data. Note that after lowess normalisation the eight slides have their log ratios centred on zero and after scale normalisation the whiskers marking the first and third quartiles for each slide now span the same range for every slide. Abbreviations used: FI0 = FIΔ at time 0, FI24 = FIΔ at time 24h, V0 = V449E at time 0 and V24 = V449E at time 24h.

4.2.3.3 Statistical analysis to determine differentially expressed genes

The statistical analysis presented in this chapter was a collaboration with Associate Professor P. Solomon and her group, the Microarray Analysis Group (MAG), in the Applied Mathematics Department, Adelaide University. The F-stat data analysis (see below) was performed with Dr. C. Wilkinson, School of Mathematical Sciences, Adelaide University and Child Health Research Institute. Early analyses of microarray data (Schena *et al.*, 1995; DeRisi *et al.*, 1996; Schena *et al.*, 1996; Wodicka *et al.*, 1997; Chu *et al.*, 1998) used simple expression fold change cut-offs to identify differentially expressed genes. However, these cut-offs are not a reliable measure of significant expression changes, as the selection of the cut-off point is arbitrary and no consideration is given to the variability in the data (Dudoit *et al.*, 2000; Quackenbush, 2001). Therefore, simple fold change cut-offs were not used as a basis for defining differentially expressed genes in this microarray experiment. To determine genes with a high probability of differential expression, a test statistic was used to rank the genes in order of their probability of differential expression (Slonim, 2002; Smyth and Speed, 2003; Smyth *et al.*, 2003). For the analysis of the microarray data, three statistics were used, the t-test statistic (t-stat), the empirical Bayes posterior log odds of differential expression or B-statistic (B-stat) (Lonnstedt and Speed, 2002) and the F-statistic (F-stat) (Smyth, 2004) (see Introduction, Sections 1.6.3.4(a), (b) and (c)). The B-stat was computed using the "stat.bay.est" function from the SMA package for R available from http://www.stat.berkely.edu.users/terry/zarray/Software/smacode.html and was used to provide an alternative estimator of a gene's significance of being differentially expressed rather than relying on t-values alone. The F-stat classifications were performed with the software package Limma of the Bioconductor project http://www.bioconductor.org (Gentleman *et al.*, 2004). The use of the F-stat is described in Section 4.2.6.1, while the t-stat and B-stat were applied separately to both the PMCI and Adelaide scans and used to rank the differentially expressed genes (see Table 11).

4.2.4 Ranking the genes with significant γ parameters

In an effort to generate the most robust data set a strategy was devised to combine the data from the two scans and the derived mean and median intensity values from the two analysis methods. This was performed to maximise the incorporation of information from both sets of scanned data as neither set could be excluded from the analysis. Similarly, while mean values are frequently used in microarray analysis, median values offer a robust alternative and thus were also included in this analysis. For each different analysis method, the probability of being differentially expressed was determined for each gene. Differential expression was defined based on the interaction parameter, γ. The most significant genes were ranked number one (1) with the rest ranked in ascending order. This ranking was then repeated for all the analysis combinations (see Figure 25) and generated eight ranked gene lists (see Table 11). Finally, in an attempt to integrate all the data the genes were ranked using an average score. This was performed by totalling each gene's ranked score over the eight lists to give it a 'sum' score. All the genes were then sorted based on this score, such that those genes with the lowest score were at the top, indicative of the fact that they had, on average, performed consistently well across both scans and analysis methods. As seen in Table 11, genes could also be ranked by looking at their minimum value (i.e. the best rank score a gene received in any of the rankings) or by their maximum value (the worst ranking a gene received). However, for the purpose of further analysis the sum ranking system was used because it reflected the overall performance of each gene across all ranking methods.

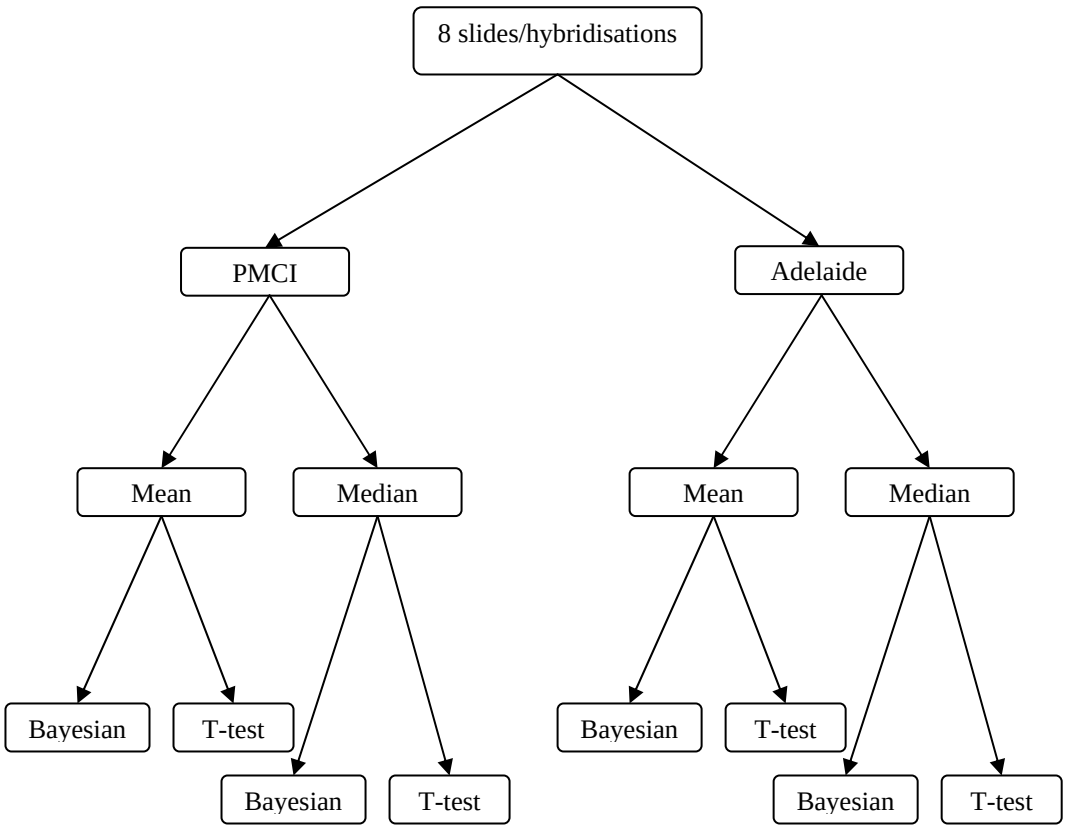


Figure 25: Flow diagram of the eight different approaches used to assign significance to the differential expression of each gene. The eight approaches (ranked gene lists) were derived from the eight slides, scanned at either Peter MacCallum Cancer Institute (PMCI) or Adelaide Microarray Facility (Adelaide), using either mean or median intensity values, which were then tested via a t-test or Bayesian test statistic, to assign a significance to each gene's observed expression level. These eight gene lists from the entire microarray data set were ranked in ascending order from the most significant (ranked one) to least significant. The eight ranked lists were subsequently used to define the Top 200 gene list (see Table 11).

Accession Number	Mean1 t rank	Mean1 bay rank	Med1 t rank	Med1 bay rank	Mean2 t rank	Mean2 bay rank	Med2 t rank	Med2 bay rank	Min	Max	Sum	Gene symbol	Gene name
BI076450	5	2	3	2	1	1	7	3	1	7	24	Ltf	Lactotransferrin
BG085146	19	3	46	3	5	2	3	2	2	46	83	Ltf	Lactotransferrin
AW546964	1	11	6	18	11	5	28	7	1	28	87	S100a9	S100 calcium binding protein A9 (calgranulin B)
BG071147	17	12	37	21	10	10	14	9	9	37	130	Gpt2	Glutamic pyruvate transaminase (alanine aminotransferase) 2
AU015908	67	4	54	4	7	3	2	1	1	67	142	Ltf	Lactotransferrin
C85808	4	1	2	1	39	4	92	6	1	92	149	Spi2-1	Serine protease inhibitor 2-1
BG084015	10	10	57	28	28	11	37	10	10	57	191	Psat1	Phosphoserine aminotransferase 1
AW537134	61	5	27	5	60	9	30	4	4	61	201	Podxl	Podocalyxin-like
AU017245	33	14	43	19	19	8	91	22	8	91	249	Eif1a	Eukaryotic translation initiation factor 1A
AU021661	25	16	148	46	32	15	4	8	4	148	294	1500005K14Rik	RIKEN cDNA 1500005K14 gene
AU014847	58	13	157	37	6	6	15	5	5	157	297	Lcn2	Lipocalin 2
AU016534	6	8	178	32	51	20	29	13	6	178	337	2610009E16Rik	RIKEN cDNA 2610009E16 gene
AW558048	28	21	67	50	86	16	69	11	11	86	348	Cd24a	CD24a antigen
C77000	38	63	23	123	21	32	16	39	16	123	355	C77032	EST C77032
AW549756	39	48	52	261	2	12	17	17	2	261	448	Dstn	Destrin
AU019271	132	6	91	7	82	7	146	14	6	146	485	1190003K14Rik	RIKEN cDNA 1190003K14 gene
AA407316	120	36	133	113	43	17	31	21	17	133	514	Neu1	neuraminidase 1
C76712	231	28	49	43	95	19	75	16	16	231	556	1500005K14Rik	RIKEN cDNA 1500005K14 gene
AW556983	2	39	1	166	122	69	104	71	1	166	574	AU020206	Expressed sequence AU020206
AU042170	113	19	66	14	199	46	98	33	14	199	588	Cd24a	CD24a antigen
AW552367	72	9	143	20	157	14	185	19	9	185	619	Stfa1	stefin A1
AW555779	35	83	26	209	137	33	85	44	26	209	652	Mapk3	Mitogen activated protein kinase 3
C77773	100	15	110	30	177	23	179	26	15	179	660	Fip1l1	FIP1 like 1 (S. cerevisiae)
AW557082	160	35	127	98	168	36	35	23	23	168	682	Psap	Prosaposin
AW558851	37	29	24	62	204	37	251	54	24	251	698	Mthfd2	Methylenetetrahydrofolate dehydrogenase (NAD+ dependent), methenyltetrahydrofolate cyclohydrolase
AW552307	157	18	398	40	45	13	18	12	12	398	701	Slc39a1	Solute carrier family 39 (iron-regulated transporter), member 1
AW558327	27	54	42	482	33	48	45	46	27	482	777	2310076K21Rik	RIKEN cDNA 2310076K21 gene

AU016154	59	96	19	386	48	31	147	52	19	386	838	S100a6	S100 calcium binding protein A6 (calcyclin)
AW536183	104	85	141	385	25	28	59	29	25	385	856	Cct3	Chaperonin subunit 3 (gamma)
AW554082	295	84	151	190	53	22	42	25	22	295	862		
C87776	54	52	74	254	135	108	172	125	52	254	974	Prkcd	Protein kinase C, delta
AU015495	136	140	71	599	9	18	5	15	5	599	993	Cyba	Cytochrome b-245, alpha polypeptide
C88302	172	53	192	263	147	35	116	32	32	263	1010	Cybb	Cytochrome b-245, beta polypeptide
AW545639	175	168	61	273	101	96	77	79	61	273	1030	Phgdh	3-phosphoglycerate dehydrogenase
C81388	177	100	124	267	160	62	129	67	62	267	1086	Slc16a1	Solute carrier family 16 (monocarboxylic acid transporters), member 1
AU021370	99	160	213	640	4	21	1	20	1	640	1158	Akr1c13	Aldo-keto reductase family 1, member C13
AW546851	141	133	86	594	37	44	78	74	37	594	1187	Cd34	CD34 antigen
AU023528	16	31	526	242	246	64	71	31	16	526	1227	Tesc	Tescalcin
AW549367	11	224	22	820	47	40	27	37	11	820	1228	Dhrs3	Dehydrogenase/reductase (SDR family) member 3
AW553460	143	131	90	458	93	39	210	73	39	458	1237	Fcgr3	Fc receptor, IgG, low affinity III
C76390	300	81	87	119	92	194	52	330	52	330	1255	Eif1a	Eukaryotic translation initiation factor 1A
AW558650	316	79	296	185	272	47	90	24	24	316	1309	0610043B10Rik	RIKEN cDNA 0610043B10 gene
AW550564	15	114	12	342	305	117	283	130	12	342	1318	2410005K20Rik	RIKEN cDNA 2410005K20 gene
AU020752	74	128	40	985	22	27	23	28	22	985	1327	Adam8	A disintegrin and metalloprotease domain 8
AW552021	178	156	65	181	267	78	272	133	65	272	1330	D11Bwg0434e	DNA segment, Chr 11, Brigham & Women's Genetics 0434 expressed
AU043052	320	25	594	73	150	26	165	27	25	594	1380	Dgat2	Diacylglycerol O-acyltransferase 2
AU020132	223	240	164	443	85	122	46	77	46	443	1400	Odc	Ornithine decarboxylase, structural
AW553604	75	201	118	646	115	43	191	70	43	646	1459	Ucp2	Uncoupling protein 2, mitochondrial
AW536664	409	74	215	392	141	59	107	69	59	409	1466	Tmem2	Transmembrane protein 2
AW537798	87	26	45	80	589	140	404	104	26	589	1475	G6pdx	Glucose-6-phosphate dehydrogenase X-linked

Table 11: Table showing the first 50 genes and their ranked lists for the eight different analysis combinations with the genes then sorted by the overall sum ranking. Shown is the ranking score obtained via the t-test (t-rank) and Bayesian estimate (bay rank) combined with the use of either the mean intensity values or median (Med) intensity values for both the Adelaide (#1) or PMCI (#2) scanned data. Also shown is the minimum and maximum score for the selected genes and the sum score generated by totalling the rank scores across all eight ranking methods.

As a first approach to further define the sequences represented, the Top 200 genes ranked by sum were examined by searching the databases using manual bioinformatic approaches to obtain information regarding each gene. Manual bioinformatics approaches included the use of BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) searches of public databases looking for similarity to clones on the microarray with published sequences, as well as updates released from the NIA website concerning the NIA 15K cloneset (http://lgsun.grc.nia.gov/cDNA/15k.html). Unigene (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene) was also extensively used to identify the clones on the array and establish whether the sequences represent known or novel genes. This limit of 200 was chosen as it represented a manageable list of genes with a high probability of differential expression. From this list of 200, genes were selected that were known (i.e. not an EST) and were deemed to be an "interesting gene" from the point of view of the original hypothesis outlined in Section 4.1. With a view to validating the differentially expressed genes identified by the statistical analysis, a small subset of genes from the Top 200 were subjected to Northern blot analysis (see Chapter 5).

4.2.5 Examining gene expression ratios from microarray Top 200 list

4.2.5.1 Selection of genes associated with either differentiation or proliferation

Using the data from the FIΔ time=24 hours / FIΔ time=0 hour (F24/F0) slide gives an indication of the expression changes (increase/decrease) associated with FIΔ signalling. Genes that are associated with an increase in expression upon differentiation in FDB-1 cells could be potential regulators of differentiation or products of differentiation i.e. genes characteristic of differentiated cells (e.g. granule enzymes). Conversely, a decrease in expression would be indicative of a potential role in proliferation or self-renewal. The Top 200 genes were divided into those that had either an increased or decreased expression associated with FIΔ signalling. These are summarized in Table 12 along with their observed fold change on the F24/F0 microarray slide. A negative fold change indicates decreased expression in the F24 sample and a positive fold change indicates increased expression in F24 sample, compared to the F0 sample (Table 12).

Table 12: The Top 200 genes and their increased or decreased expression upon FIΔ signalling. (A) The differentiation-associated list – those genes that increase expression over time in FIΔ. **(B)** The proliferation-associated list – those genes that decreased in expression over time in FIΔ. Note the 'fold change' gives the expression fold change of F24 over F0 whereby positive values indicate increased expression in the F24 population compared to F0 and negative values indicate increase expression in the F0 population compared to F24. The "Sum" score is as described in Table 2 while the "Rank" is the Global rank obtained from the B-stat analysis of the data from the Adelaide and PMCI scans using mean and median values.

(A) Differentiation-Associated list

Index	ID	Gene symbol	Gene name	Decrease/Increase in FIΔ signalling	Fold change	Sum	Rank
5834	H3109G06	Ltf	Lactotransferrin	Increase	5.41	24	1
803	H3109B05	Ltf	Lactotransferrin	Increase	5.10	83	2.5
8399	H3115B07	S100a9	S100 calcium binding protein A9 (calgranulin B)	Increase	2.64	87	10
13330	H3087A12	Ltf	Lactotransferrin	Increase	4.90	142	2.5
14850	H3082F04	1500005K14Rik	RIKEN cDNA 1500005K14 gene	Increase	3.46	294	5
5271	H3083G02	Lcn2	Lipocalin 2	Increase	3.33	297	4
9837	H3089C03	2610009E16Rik	RIKEN cDNA 2610009E16 gene	Increase	2.15	337	90
5492	H3155C06	Cd24a	CD24a antigen	Increase	2.57	348	20
11946	H3122E03	Dstn	Destrin	Increase	2.37	448	17.5
2227	H3076D09	1190003K14Rik	RIKEN cDNA 1190003K14 gene	Increase	3.17	485	45
473	H3003F05	Neu1	Neuraminidase 1	Increase	2.29	514	41
2610	H3034F01	1500005K14Rik	RIKEN cDNA 1500005K14 gene	Increase	3.28	556	6
1811	H3109A05	Cd24a	CD24a antigen	Increase	2.50	588	48
12472	H3131F12	Stfa1	stefin A1	Increase	2.77	619	37
8991	H3145H03	Mapk3	Mitogen activated protein kinase 3	Increase	1.99	652	62
14700	H3038B04	Fip1l1	FIP1 like 1 (S. cerevisiae)	Increase	1.86	660	159
8479	H3151D11	Psap	Prosaposin	Increase	2.57	682	22
380	H3131D05	Slc39a1	Solute carrier family 39 (iron-regulated transporter), member 1	Increase	3.07	701	28
7506	H3156E02	Ypel5	Yippee-like 5 (Drosophila)	Increase	1.61	777	210
217	H3087H09	S100a6	S100 calcium binding protein A6 (calcyclin)	Increase	1.72	838	245.5
2925	H3138D01	Cnn2	Calponin 2	Increase	2.47	862	19
8232	H3059B03	Prkcd	Protein kinase C, delta	Increase	1.93	974	66
9795	H3085G03	Cyba	Cytochrome b-245, alpha polypeptide	Increase	2.12	993	31
10240	H3060F11	Cybb	Cytochrome b-245, beta polypeptide	Increase	2.12	1010	36
14304	H3068B04	Tesc	Tescalcin	Increase	1.57	1227	329
13962	H3121E04	Dhrs3	Dehydrogenase/reductase (SDR family) member 3	Increase	2.68	1228	7.5
9433	H3135G11	Fcgr3	Fc receptor, IgG, low affinity III	Increase	2.07	1237	39
14042	H3157G08	Ypel3	Yippee-like 3 (Drosophila)	Increase	3.18	1309	12.5

3224	H3080C05	Adam8	A disintegrin and metalloprotease domain 8	Increase	2.09	1327	34
12981	H3129D04	D11Bwg0434e	DNA segment, Chr 11, Brigham & Women's Genetics 0434 expressed	Increase	1.91	1330	98
13394	H3111G08	Dgat2	Diacylglycerol O-acyltransferase 2	Increase	2.35	1380	99
15484	H3136E12	Ucp2	Uncoupling protein 2, mitochondrial	Increase	1.41	1459	719
2584	H3026F09	Tmem2	Transmembrane protein 2	Increase	2.04	1466	77
15178	H3032A12	G6pdx	Glucose-6-phosphate dehydrogenase X-linked	Increase	1.71	1475	429
9411	H3127G03	Samhd1	SAM domain and HD domain, 1	Increase	2.08	1528	65
11548	H3008E11	Copeb	Core promoter element binding protein	Increase	1.70	1543	156
7895	H3114A06	Sat	Spermidine/spermine N1-acetyl transferase	Increase	2.24	1545	43
10915	H3114D11	Btg1	B-cell translocation gene 1, anti-proliferative	Increase	2.00	1558	55.5
637	H3049H09	Slc6a6	Solute carrier family 6 (neurotransmitter transporter, taurine), member 6	Increase	1.68	1733	137
14586	H3004F04	Klhdc4	Kelch domain containing 4	Increase	1.57	1818	165
6999	H3154H02	AI325941	Expressed sequence AI325941	Increase	2.15	1829	92.5
5445	H3139C02	Add1	Adducin 1 (alpha)	Increase	1.58	1833	343.5
6301	H3092H10	Gyg1	Glycogenin 1	Increase	1.98	1882	44
7898	H3126G06	Dstn	Destrin	Increase	2.03	1977	103.5
11764	H3070E11	Samhd1	SAM domain and HD domain, 1	Increase	2.13	2019	72
14874	H3090F04	Chst12	Carbohydrate sulfotransferase 12	Increase	2.00	2058	52
8101	H3019H11	Atp6v0c	ATPase, H+ transporting, lysosomal 16kD, V0 subunit C	Increase	1.71	2067	175
1838	H3113G05	Ypel5	Yippee-like 5 (Drosophila)	Increase	1.41	2074	622
2158	H3052B09	Sp100	Nuclear antigen Sp100	Increase	2.00	2084	59
7188	H3056A02	Copeb	Core promoter element binding protein	Increase	1.60	2098	196
14466	H3124F04	Mgst1	Microsomal glutathione S-transferase 1	Increase	1.60	2169	333
3419	H3144A05	Slc31a2	Solute carrier family 31, member 2	Increase	2.33	2192	97
7654	H3034A10	Plaur	Urokinase plasminogen activator receptor	Increase	2.22	2215	70.5
6416	H3136D06	Scotin	Scotin gene	Increase	1.39	2273	1040
2435	H3152B05	Hipk1	Homeodomain interacting protein kinase 1	Increase	2.29	2279	47
8986	H3149B11	Mpp1	Membrane protein, palmitoylated (55 kDa)	Increase	2.30	2346	17.5
260	H3091D05	2310043N10Rik	RIKEN cDNA 2310043N10 gene	Increase	1.72	2377	121
2857	H3126H09	Ckap4	Cytoskeleton-associated protein 4	Increase	1.66	2393	367
9590	H3009G07	4632428N05Rik	RIKEN cDNA 4632428N05 gene	Increase	2.59	2400	15
7698	H3050E02	Pigb	Phosphatidylinositol glycan, class B	Increase	2.03	2493	79

209	H3075F05	6230421P05Rik	RIKEN cDNA 6230421P05 gene	Increase	1.85	2496	147
1129	H3055G09	C130099A20Rik	RIKEN cDNA C130099A20 gene	Increase	1.88	2503	82
2636	H3042D05	Scap2	Src family associated phosphoprotein 2	Increase	1.74	2541	128
5963	H3149A06	Aldh3b1	aldehyde dehydrogenase 3 family, member B1	Increase	2.40	2570	84
356	H3123D05	Pygl	Liver glycogen phosphorylase	Increase	2.02	2591	29
11453	H3136E07	Plp2	Proteolipid protein 2	Increase	1.45	2741	669
4568	H3021D06	Dhrs1	Dehydrogenase/reductase (SDR family) member 1	Increase	1.53	2762	565
2614	H3034B09	Dgat1	Diacylglycerol O-acyltransferase 1	Increase	1.94	2764	69
6333	H3100D02	Tram1	Translocating chain-associating membrane protein 1	Increase	1.67	2769	206
7374	H3120E02	Itm2b	Integral membrane protein 2B	Increase	2.20	2781	16
2475	H3008H01	Anxa1	Annexin A1	Increase	2.37	2798	53
4833	H3109D02	Anxa9	Annexin A9	Increase	1.51	2818	483
6415	H3136D10	H3136D10	Transcribed locus, weakly similar to NP_058618.2 galactose-3-O-sulfotransferase 1 [Mus musculus]	Increase	1.30	2830	1103
7400	H3128C06	Adipor1	Adiponectin receptor 1	Increase	1.64	2833	280.5
5340	H3111A02	Prdx5	Peroxiredoxin 5	Increase	2.32	2851	46
9363	H3111G03	Plekhm1	Pleckstrin homology domain containing, family M (with RUN domain) member 1	Increase	1.53	2867	309
6283	H3084D10	Iqgap1	IQ motif containing GTPase activating protein 1	Increase	1.56	2890	183
11985	H3142C03	Rgl2	Ral guanine nucleotide dissociation stimulator,-like 2	Increase	1.83	2904	157
7281	H3088C02	Daf1	Decay accelerating factor 1	Increase	1.42	3100	1366
11478	H3144E03	Zyx	Zyxin	Increase	2.31	3118	76
15241	H3056G12	1110001A05Rik	RIKEN cDNA 1110001A05 gene	Increase	1.59	3126	976
8866	H3109B11	Slc27a4	Solute carrier family 27 (fatty acid transporter), member 4	Increase	2.16	3132	101
5581	H3017G10	1500003D12Rik	RIKEN cDNA 1500003D12 gene	Increase	1.46	3153	348
2722	H3078B09	Klf5	Kruppel-like factor 5	Increase	1.78	3162	411
5330	H3111G06	Arg2	Arginase type II	Increase	1.72	3174	443
9606	H3021E03	Tegt	Testis enhanced gene transcript	Increase	1.74	3176	105
13925	H3117E08	Raf1	v-raf-1 leukemia viral oncogene 1	Increase	2.06	3183	68
10403	H3120B07	Gsn	Gelsolin	Increase	1.61	3217	232
9522	H3155E03	BC027342	cDNA sequence BC027342	Increase	1.79	3265	130
583	H3037D09	2610017G09Rik	RIKEN cDNA 2610017G09 gene	Increase	3.12	3316	38
9660	H3037A03	Actn1	Actinin, alpha 1	Increase	1.67	3334	209
1525	H3009G09	1110008P14Rik	RIKEN cDNA 1110008P14 gene	Increase	1.74	3335	291

5159	H3043A06	Gsr	Glutathione reductase 1	Increase	1.88	3344	253
3275	H3096A05	Dap	death-associated protein	Increase	1.28	3358	1850.5
5844	H3109A02	Fcer1g	Fc receptor, IgE, high affinity I, gamma polypeptide	Increase	1.70	3387	143
13729	H3053G12	Pnkp	Polynucleotide kinase 3'- phosphatase	Increase	2.23	3392	50
11758	H3058A11	H3058A11	UNKNOWN	Increase	1.76	3402	417
5491	H3155C10	Ugt1a6	UDP glycosyltransferase 1 family, polypeptide A6	Increase	1.85	3419	125
11405	H3120E07	Serpinb1a	Serine (or cysteine) proteinase inhibitor, clade B, member 1a	Increase	1.37	3488	705
75	H3039H01	Cnn2	Calponin 2	Increase	2.04	3611	60
5246	H3075G06	Mocs1	Molybdenum cofactor synthesis 1	Increase	1.54	3709	341
8105	H3019F07	Stk10	Serine/threonine kinase 10	Increase	1.50	3730	670.5
3036	H3016A01	Hif1a	Hypoxia inducible factor 1, alpha subunit	Increase	1.27	3781	1480
11252	H3060C07	A630082K20Rik	RIKEN cDNA A630082K20 gene	Increase	1.83	3809	248
15780	H3062A04	Rab32	RAB32, member RAS oncogene family	Increase	1.85	3890	108
13460	H3123C08	Sirt7	Sirtuin 7 (silent mating type information regulation 2, homolog) 7 (S. cerevisiae)	Increase	1.72	3932	204
521	H3009F05	H3009F05	Similar to phosphatidylserine decarboxylase (Mm.283387)	Increase	2.09	4056	63.5
15634	H3014A12	Capg	Capping protein (actin filament), gelsolin-like	Increase	1.49	4074	1657
8646	H3037F03	Cab39	Calcium binding protein 39	Increase	1.42	4161	334
7206	H3064E02	Atp6v1a1	ATPase, H+ transporting, V1 subunit A, isoform 1	Increase	1.08	4214	6568
11989	H3138G11	Niban	Niban protein	Increase	1.56	4223	384
4502	H3003H06	2310014H01Rik	RIKEN cDNA 2310014H01 gene	Increase	1.94	4245	249

(B) Proliferation-Associated list

Index	ID	Gene symbol	Gene name	Decrease/Increase in FIΔ signalling	Fold change	Sum	Rank
5289	H3095C02	Gpt2	Glutamic pyruvate transaminase (alanine aminotransferase) 2	Decrease	-2.01	130	106
14248	H3052F12	Spi2-1	Serine protease inhibitor 2-1	Decrease	-4.13	149	22
1251	H3095G01	Psat1	Phosphoserine aminotransferase 1	Decrease	-2.38	191	24
9633	H3029C03	Podxl	Podocalyxin-like	Decrease	-3.06	201	25
1264	H3091E09	Eif1a	Eukaryotic translation initiation factor 1A	Decrease	-3.34	249	7.5
1096	H3035E09	C77032	EST C77032 (Nucleostemin)	Decrease	-1.61	355	167
10993	H3150H11	AU020206	Expressed sequence AU020206	Decrease	-1.48	574	315
12030	H3158E03	Mthfd2	Methylenetetrahydrofolate dehydrogenase (NAD+ dependent), methenyltetrahydrofolate cyclohydrolase	Decrease	-1.67	698	225
14148	H3024B04	Cct3	Chaperonin subunit 3 (gamma)	Decrease	-1.53	856	298.5
2549	H3022F05	Phgdh	3-phosphoglycerate dehydrogenase	Decrease	-1.69	1030	115
5679	H3049G02	Slc16a1	Solute carrier family 16 (monocarboxylic acid transporters), member 1	Decrease	-1.52	1086	473
4767	H3081H02	Akr1c13	Aldo-keto reductase family 1, member C13	Decrease	-2.11	1158	26
6879	H3114H02	Cd34	CD34 antigen	Decrease	-1.82	1187	88
9662	H3033G07	Eif1a	eukaryotic translation initiation factor 1A	Decrease	-1.51	1255	337
11437	H3124G11	2410005K20Rik	RIKEN cDNA 2410005K20 gene	Decrease	-1.62	1318	155
6750	H3078F02	Odc	Ornithine decarboxylase, structural	Decrease	-1.28	1400	866
7476	H3152A02	St6gal1	Beta galactoside alpha 2,6 sialyltransferase 1	Decrease	-2.09	1662	136
11456	H3136C07	Bnip3	BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3	Decrease	-1.58	1698	317
15854	H3082G08	Myo10	Myosin X	Decrease	-1.53	1882	599
13564	H3007E12	9930111I18Rik	RIKEN cDNA 9930111I18 gene	Decrease	-2.36	1915	89
13462	H3123A12	H3123A12	UNKNOWN	Decrease	-1.18	1941	3168
5967	H3145G02	Phb	Prohibitin	Decrease	-1.47	1977	372.5
13211	H3047A08	Phgdh	3-phosphoglycerate dehydrogenase	Decrease	-1.69	2098	173
13963	H3121C12	Ankrd10	Ankyrin repeat domain 10	Decrease	-1.30	2100	1053
5360	H3119C06	Nol5	Nucleolar protein 5	Decrease	-1.62	2149	302
10698	H3042F03	Smyd5	SET and MYND domain containing 5	Decrease	-1.25	2171	2256.5
11908	H3118E11	Hbb-b1	Hemoglobin, beta adult major chain	Decrease	-3.47	2177	133

15938	H3118G08	D17H6S56E-5	DNA segment, Chr 17, human D6S56E 5	Decrease	-1.99	2186	40
12256	H3059F12	Tgtp	T-cell specific GTPase	Decrease	-1.80	2282	150
5959	H3149C10	Tpi	Triosephosphate isomerase	Decrease	-1.03	2313	15036.5
8821	H3089H11	Myc	Myelocytomatosis oncogene	Decrease	-1.05	2323	12844.5
6449	H3140F06	Zfhx1b	Zinc finger homeobox 1b	Decrease	-1.42	2333	389
9266	H3079G07	Fkbp4	FK506 binding protein 4	Decrease	-1.16	2381	2701
3053	H3024E05	Emb	Embigin	Decrease	-1.30	2417	1488
2051	H3024B05	Cct3	Chaperonin subunit 3 (gamma)	Decrease	-1.50	2440	380
9368	H3111C07	Adprt1	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase) 1	Decrease	-1.60	2563	239.5
2268	H3096B01	AU040377	Expressed sequence AU040377	Decrease	-1.72	2568	626.5
5552	H3013C06	Mat2a	Methionine adenosyltransferase II, alpha	Decrease	-1.01	2583	15855.5
5003	H3005B06	Ppat	Phosphoribosyl pyrophosphate amidotransferase	Decrease	-1.48	2634	773
10617	H3022D03	Galk1	Galactokinase 1	Decrease	-1.52	2636	272
10116	H3024B03	Cct3	Chaperonin subunit 3 (gamma)	Decrease	-1.55	2645	511
7166	H3044G06	Phgdh	3-phosphoglycerate dehydrogenase	Decrease	-1.56	2723	325
6873	H3118D02	Paics	Phosphoribosylaminoimidazole carboxylase, phosphoribosylaminoribosylaminoimidazole, succinocarboxamide synthetase	Decrease	-1.54	2783	172
14532	H3152B04	2410118I19Rik	RIKEN cDNA 2410118I19 gene	Decrease	-1.71	2876	280.5
3484	H3008E09	Ddx18	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	Decrease	-1.20	2912	2023
12043	H3154C11	Eprs	Glutamyl-prolyl-tRNA synthetase	Decrease	-1.55	2939	269
6101	H3032F06	Hbb-bh1	Hemoglobin Z, beta-like embryonic chain	Decrease	-2.54	2949	223.5
14810	H3078H08	AU020206	Expressed sequence AU020206	Decrease	-1.36	3023	806
8228	H3059D07	Aars	Alanyl-tRNA synthetase	Decrease	-1.36	3110	575
2972	H3154D05	Exosc8	Exosome component 8	Decrease	-1.40	3118	1197
4922	H3141H06	Phgdhl1	Phosphoglycerate dehydrogenase like 1	Decrease	-1.30	3121	878
1711	H3077C09	Nap1l1	Nucleosome assembly protein 1-like 1	Decrease	-1.55	3127	222
4857	H3117D02	Hbb-bh1	Hemoglobin Z, beta-like embryonic chain	Decrease	-2.53	3177	239.5
2081	H3028F05	Odc	Ornithine decarboxylase, structural	Decrease	-1.29	3187	1324
11041	H3006H11	Nsun2	NOL1/NOP2/Sun domain family 2	Decrease	-1.52	3225	871.5
2466	H3156F01	2310067B10Rik	RIKEN cDNA 2310067B10 gene	Decrease	-2.18	3243	55.5
8569	H3013H11	Ranbp1	RAN binding protein 1	Decrease	-1.48	3252	712
4596	H3029B02	Cct3	Chaperonin subunit 3 (gamma)	Decrease	-1.37	3268	789

14483	H3136B08	Npm1	Nucleophosmin 1	Decrease	-1.61	3288	298.5
13350	H3095E04	H3095E04	UNKNOWN	Decrease	-1.76	3301	85
10541	H3008F07	H2-Eb1	histocompatibility 2, class II antigen E beta	Decrease	-2.03	3314	123
13659	H3029G04	Gemin6	Gem (nuclear organelle) associated protein 6	Decrease	-1.44	3382	811
8711	H3049B07	Hsp105	Heat shock protein 105	Decrease	-1.08	3390	11617
10851	H3102H03	Iars	Isoleucine-tRNA synthetase	Decrease	-1.81	3392	212
1575	H3025G01	Sfrs3	Splicing factor, arginine/serine-rich 3 (SRp20)	Decrease	-1.30	3512	796
7650	H3034E02	Prps1	phosphoribosyl pyrophosphate synthetase 1	Decrease	-1.31	3538	823
7409	H3124E06	Atf4	activating transcription factor 4	Decrease	-1.46	3641	651
4571	H3021B06	Flnb	Filamin, beta	Decrease	-1.70	3649	241
7522	H3008A10	Mybbp1a	MYB binding protein (P160) 1a	Decrease	-1.52	3694	618
11977	H3142G11	Tars	threonyl-tRNA synthetase	Decrease	-1.41	3698	856
5063	H3011A06	Pyp	Pyrophosphatase	Decrease	-1.75	3715	213.5
1571	H3029A05	Cse1l	Chromosome segregation 1-like (S. cerevisiae)	Decrease	-1.59	3776	177
13067	H3005B08	2700029C06Rik	RIKEN cDNA 2700029C06 gene	Decrease	-1.77	3786	199
6346	H3112B10	AW545966	expressed sequence AW545966	Decrease	-1.34	3816	1060
14342	H3084H08	D330037H05Rik	RIKEN cDNA D330037H05 gene	Decrease	-1.26	3824	1979
15514	H3144A12	Ahcy	S-adenosylhomocysteine hydrolase	Decrease	-2.00	3847	61
766	H3089B09	Ddx18	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	Decrease	-1.20	3854	1860
3437	H3152E05	H3152E05	UNKNOWN	Decrease	-1.42	3903	974
10451	H3136B07	Npm1	Nucleophosmin 1	Decrease	-1.52	3904	362
7947	H3142G02	Mthfd1	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent), methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthase	Decrease	-1.34	3913	1603.5
9498	H3147E03	Statip1	Signal transducer and activator of transcription interacting protein 1	Decrease	-1.36	3930	491
5981	H3157E06	Ssfa1	Sperm specific antigen 1	Decrease	-1.19	3969	2177
8544	H3003B03	Ppid	Peptidylprolyl isomerase D (cyclophilin D)	Decrease	1.15	4038	4342
10449	H3136D03	Txndc7	Thioredoxin domain containing 7	Decrease	-1.48	4049	454
4853	H3117F06	Rbbp7	Retinoblastoma binding protein 7	Decrease	-1.46	4116	434
1871	H3121A05	Ppan	Peter pan homolog (Drosophila)	Decrease	-1.09	4131	11090
713	H3073F05	1200009K13Rik	RIKEN cDNA 1200009K13 gene	Decrease	-1.52	4171	255
16024	H3138E12	8430438D04Rik	RIKEN cDNA 8430438D04 gene	Decrease	-2.88	4190	310

4.2.5.2 Multiple clones in the Top 200 list that represent the same gene

Within the NIA clone-set there were a number of genes that were represented more than once on the array. This is apparent in Table 12 where several of the genes in the Top 200 list are represented by several different clone identifiers (Clone ID). This can arise when an EST and a known gene are later found to represent the same coding sequence. Alternatively, some genes were placed on the array more than once due to sequence redundancy, (for example two ESTs of different sequence that both represent the same gene) and others for control purposes (e.g. GAPDH). Some of the regulated genes on the array were represented by more than one clone or spot on the array. This provided increased confidence with regard to the observed expression changes when multiple clones representing the same gene give similar gene expression ratios.

Using the four dye swap slides, which compared V449E (V) and FIΔ (F) at both time zero (0) and time 24 hours, expression ratios were obtained for each of the Top 200 ranked genes. Plotting the F/V 24 hours ratio and the F/V 0 hour ratio gave an indication of the fold increase/decrease in expression over time. Genes that increased in expression over 24 hours in FIΔ were classed as differentiation-associated, whereas genes that decreased in expression over 24 hours in FIΔ were classed as proliferation-associated. This classification was based on the expectation that genes involved in differentiation would increase in expression in FIΔ cells over time, whereas genes involved in proliferation would increase in expression in V449E cells over time. For example, a gene with a F/V 0 hour ratio of one that increases to four for the F/V 24 hours ratio is indicative of a differentiation-associated gene. This increase in the F/V 24 hours ratio compared to the F/V 0 hour ratio can occur by either an increase in expression in the FIΔ sample over time, or by a decrease in the V449E sample over time. Thus, these ratios can be used to examine differentiation- and proliferation-associated genes that go beyond just looking at the change in expression of genes in the FIΔ cell line (used in Table 12) by also including the change in the V449E cell line over time. The F/V 24 hours and F/V 0 hour ratio data is presented for a selection of the genes from the Top 200 in Table 13. In Table 13 the average ratio was shown for the genes represented by more than one clone. For comparing the change in expression that occurs in each cell line over time, the F24/F0 and V24/V0 ratios are also presented (see Table 13).

Table 13: Gene expression changes expressed as ratios using the Adelaide (mean intensity) data. Using the dye swapped slides at time 0 and 24 hours, the time course slides (F24/F0 and V24/V0) and by averaging the data for multiple clones for the same gene across the dye swapped slides, gene expression ratios for the selected genes are presented. These ratios give an indication of the relative fold increase (greater than one) or decrease (less than one) in gene expression that occurs between the samples (FIΔ and V449E) over time. The shaded genes with their ratios presented in italics were the average of multiple clones.

Gene name	Symbol	F/V time 0	F/V time 24	F24/F0	V24/V0
DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	Ddx18	*1.23*	*0.82*	*0.89*	*1.23*
Chaperonin subunit 3 (gamma)	Cct3	*1.24*	*0.73*	*0.64*	*1.07*
CD24a antigen	Cd24a	*0.91*	*2.05*	*2.88*	*0.92*
Core promoter element binding protein	Copeb	*0.65*	*1.24*	*1.53*	*1.06*
Destrin	Dstn	*1.95*	*3.61*	*2.26*	*1.05*
Hemoglobin Y, beta-like embryonic chain	Hbb-y	*0.92*	*0.26*	*0.41*	*1.16*
Lactotransferrin	Ltf	*2.46*	*12.46*	*5.15*	*0.92*
Nucleophosmin 1	Npm1	*1.26*	*0.80*	*0.58*	*1.10*
Ornithine decarboxylase, structural	Odc	*1.22*	*0.73*	*0.84*	*1.12*
3-phosphoglycerate dehydrogenase	Phgdh	*1.36*	*0.77*	*0.65*	*1.03*
SAM domain and HD domain, 1	Samhd1	*0.84*	*1.39*	*2.19*	*1.17*
Yippee-like 5 (Drosophila)	Ypel5	*0.90*	*1.51*	*1.59*	*0.92*
Annexin A1	Anxa1	1.27	2.51	2.29	1.17
Annexin A9	Anxa9	0.72	1.25	1.50	0.93
Activating transcription factor 4	Atf4	0.93	0.60	0.69	1.05
BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3	Bnip3	1.07	0.54	0.70	1.12
B-cell translocation gene 1, anti-proliferative	Btg1	0.60	1.06	1.91	1.11
Chromobox homolog 5 (Drosophila HP1a)	Cbx5	0.90	0.29	0.43	1.21
CD34 antigen	Cd34	0.72	0.37	0.61	0.89
Calponin 2	Cnn2	1.16	2.02	2.22	1.34
Eukaryotic translation initiation factor 1A	Eif1a	1.17	0.70	0.62	1.44
Similar to eukaryotic translation initiation factor 1A		0.44	0.17	0.29	0.79
Fc receptor, IgE, high affinity I, gamma polypeptide	Fcer1g	0.88	1.33	1.59	1.10
Fc receptor, IgG, low affinity III	Fcgr3	1.11	1.89	2.09	1.12
Hypoxia inducible factor 1, alpha subunit	Hif1a	0.76	1.21	1.07	0.93
Homeodomain interacting protein kinase 1	Hipk1	1.20	1.97	2.35	1.13
Kruppel-like factor 5	Klf5	0.83	2.05	1.43	1.01
EST C77032 (Nucleostemin)	C77032	1.35	0.73	0.61	1.10
Mitogen activated protein kinase 3	Mapk3	1.12	2.03	2.13	1.15
MYB binding protein (P160) 1a	Mybbp1a	1.15	0.76	0.59	1.14
Myelocytomatosis oncogene	Myc	1.28	0.86	1.00	1.39
Urokinase plasminogen activator receptor	Plaur	1.70	3.15	2.54	1.10
Protein kinase C, delta	Prkcd	0.80	1.57	1.80	1.03
RAB32, member RAS oncogene family	Rab32	1.04	1.72	1.68	1.22
v-raf-1 leukemia viral oncogene 1	Raf1	1.36	2.21	2.04	1.17
Retinoblastoma binding protein 7	Rbbp7	1.08	0.72	0.67	0.93
Ral guanine nucleotide dissociation stimulator,-like	Rgl2	0.79	1.49	1.81	1.11
S100 calcium binding protein A6 (calcyclin)	S100a6	0.78	1.25	1.86	0.92
S100 calcium binding protein A9 (calgranulin B)	S100a9	2.81	7.49	2.66	0.98
Src family associated phosphoprotein 2	Scap2	1.05	1.83	1.67	1.08
Serine (or cysteine) proteinase inhibitor, clade B, member 1a	Serpinb1a	0.88	1.48	1.25	1.02
Splicing factor, arginine/serine-rich 3 (SRp20)	Sfrs3	1.14	0.78	0.79	1.05
Solute carrier family 16 (monocarboxylic acid transporters), member 1	Slc16a1	1.41	0.87	0.59	1.20
Solute carrier family 27 (fatty acid transporter), member 4	Slc27a4	1.19	2.61	2.16	1.24
Solute carrier family 31, member 2	Slc31a2	1.18	2.84	2.40	1.29
Solute carrier family 39 (iron-regulated transporter), member 1	Slc39a1	1.17	3.78	2.38	1.09
Solute carrier family 6 (neurotransmitter transporter, taurine), member 6	Slc6a6	0.75	1.23	1.65	0.95
Nuclear antigen Sp100	Sp100	0.93	1.76	1.71	1.10
Serine protease inhibitor 2-1	Spi2-1	1.04	0.16	0.23	1.44
Signal transducer and activator of transcription interacting protein 1	Statip1	0.98	0.68	0.72	1.00
Serine/threonine kinase 10	Stk10	0.78	1.46	1.38	1.04
T-cell specific GTPase	Tgtp	0.86	0.44	0.62	0.94
Vesicle transport through interaction with t-SNAREs 1B homolog	Vti1b	1.27	3.12	1.20	0.85
Yippee-like 3 (Drosophila)	Ypel3	0.73	1.37	3.27	1.47
Zinc finger homeobox 1b	Zfhx1b	1.01	0.67	0.73	1.03

In Figure 26 individual ratios of F/V at time 0 and 24 hours are shown separately for 12 genes (represented by a total of 27 clones) from the microarray. Figure 26 shows the average expression ratio hence, only the genes represented by multiple clones within the Top 200 list are shown. This figure reveals that in the Top 200 list there were 12 genes that had two or more clones represented on the array and that for each gene there was agreement in the expression levels found for the multiple clones, indicated by the small standard deviation between replicates.

Figure 26: Twelve genes identified with multiple clones from the Top 200 list. Microarray data from the FIΔ time 24 versus 0 hours slide and V449E time 24 versus 0 hours slide. Data from these slides was used to calculate the fold increase over time in the FIΔ and V449E samples. Some of the genes were represented on the microarray by more than one clone (or spot) on the array. When this occurred, the average fold increase was presented in the figure along with error bars showing the standard deviation between the two or three clones used to calculate the average fold increase.

4.2.6 Further significance testing

4.2.6.1 Applying the F-statistic to the microarray data

Of primary interest was γ (the interaction parameter) which identifies genes that are regulated over time and between samples. However, also of interest were the parameters α and β which identify those genes that were different between FIΔ and V449E (sample differences, α) and those genes that were significantly different over time (time differences, β). Gene lists derived from the analysis of α and β parameters may reveal specific genes involved in FIΔ differentiation, or those genes that are critically different between FIΔ and V449E that contribute to their different biological effects. The F-stat (see Section 1.6.3.4(c)) was used to look at genes with significant γ and either significant α or β. This was achieved by using the F-stat to test the null hypothesis (H_0) that α=β=γ=0. After identifying genes significant for γ, the F-stat examines the possibility that α and/or β are also significant. The limitation with the F-stat is that there is no formal method to perform a multiple testing correction, so conservative P-values must be used to avoid increased numbers of false positives.

4.2.6.2 Significant α and β gene lists

The ranked gene list as described in Section 4.2.4 was restricted to the genes identified on the basis of the interaction parameter (γ). To identify those genes that were significantly different between the samples (α) and those genes significantly differentially expressed over time in FIΔ (β), the data from the experiment was reanalysed with the Limma package of Bioconductor, (available from the Bioconductor project site http://www.bioconductor.org). For this analysis, the genes were ranked by calculating separately the critical values for each parameter (α, β, and γ) for all genes using the B-stat and F-stat. Using the F-stat ranked data gene lists were produced where the relative increase or decrease in expression in FIΔ compared to V449E could be examined. The significant genes defined by the F-stat (using genes with significant γ and β parameters) were split based on whether they had a significant α parameter. The gene lists were further limited by selecting a 1.5 fold cut-off for up- or down-regulation to create manageable sized lists, yielding a total of 177 genes that fitted these criteria. Therefore as long as either FIΔ or V449E was increased greater than or equal to 1.5 fold (or decreased by less than or equal to –1.5 fold) over time, a given statistically

significant gene was included in the proceeding lists. Using the process outlined above, a total of 177 genes were identified that fitted this criteria. These 177 genes as mentioned above were split into (1) a group with significant α and (2) a group with non-significant α. These two groups were further subdivided into three subgroups where each subgroup represented a relevant subclass (see Table 14). The three subgroups identified differentiation-associated genes, proliferation-associated genes, and IL-3 withdrawal-associated genes. The six lists are summarised below in Table 14 and the genes they contain are listed in Table 15.

No. of genes	List	Expression change over time
Genes that are not differentially regulated between FIΔ and V449E at time = 0 (no significantly sample effect = **α**)		
68	Differentiation-associated genes	↑FIΔ V449E unchanged
22	Proliferation-associated genes	↓FIΔ ↑V449E or V449E unchanged
3	IL-3 withdrawal-associated genes	↑FIΔ ↑V449E or ↓FIΔ ↓V449E
Genes that are significantly different between FIΔ and V449E at time = 0 (significantly **α** = sample effect)		
59	Differentiation-associated genes	↑FIΔ V449E unchanged
21	Proliferation-associated genes	↓FIΔ ↑V449E or V449E unchanged
4	IL-3 withdrawal-associated genes	↑FIΔ ↑V449E or ↓FIΔ ↓V449E

Table 14: Biologically relevant subclasses defined by the F-statistic. Using F-stat ranked data significant gene lists were produced based on significant γ and β parameters and then split by whether the α parameter was also significant. In addition each gene's fold expression over time had to be either up or down-regulated by at least 1.5 fold to be included. These lists were subdivided into biologically meaningful groups based on the relative expression of FIΔ and V449E over time, where increased expression is denoted by ↑, and decreased expression by ↓.

Table 15A-E: Biologically relevant gene lists identified from F-stat analysis. (A) Differentiation-associated genes. (B) Differentiation-associated genes that have a sample difference. (C) Proliferation-associated genes. (D) Proliferation-associated genes that have a sample effect. (E) IL-3 withdrawal-associated genes, with and without sample effects. Clones representing the same gene within a list are highlighted. Abbreviations used; ID = NIA CloneID, LLID = locus link ID, Avg F fc = average FIΔ fold change, rank = ranking in the initial analysis via γ and GO-abv-source = Gene Ontology abbreviations obtained from the Source website (http://source.standford.edu/cgi-bin/source/sourceSearch) for each gene based on its LLID. Some genes with LLID did not yield GO abbreviations from Source.

Table 15A: Differentiation-associated genes

ID	LLID	Name	Gene Name	Avg F fc	Rank	GO abv - source
H3082F04	76566	1500005K14Rik	RIKEN cDNA 1500005K14 gene	3.20	9	
H3155C06	12484	Cd24a	CD24a antigen	2.62	15	external side of plasma membrane|membrane
H3131D05	53945	Slc40a1	Solute carrier family 40 (iron-regulated transporter), member 1	2.89	20	binding|integral to membrane|iron ion transport|iron ion transporter activity|mitochondrial inner membrane|synaptic vesicle|transport
H3003F05	18010	Neu1	Neuraminidase 1	2.20	22	exo-alpha-sialidase activity|lysosome|protein binding
H3138D01	12798	Cnn2	Calponin 2	2.43	23	actin binding|calmodulin binding|smooth muscle contraction
H3145H03	26417	Mapk3	Mitogen activated protein kinase 3	2.00	29	ATP binding|MAP kinase activity|MAP kinase activity|cell cycle|cytoplasm|nucleus|nucleus|organogenesis|phosphorylation|phosphotyrosine binding|protein amino acid phosphorylation|protein binding|protein kinase activity|protein serine/threonine kinase activity|protein-tyrosine kinase activity|signal transduction|transferase activity
H3135G11	14131	Fcgr3	Fc receptor, IgG, low affinity III	2.11	32	IgG binding|IgG receptor activity|antibody-dependent cellular cytotoxicity|antigen presentation, exogenous antigen via MHC class I|cell surface receptor linked signal transduction|extracellular space|integral to membrane|mast cell activation|mast cell activation|neutrophil chemotaxis|phagocytosis, engulfment|phagocytosis, recognition|positive regulation of phagocytosis|positive regulation of tumor necrosis factor-alpha biosynthesis|positive regulation of type I hypersensitivity|positive regulation of type I hypersensitivity|positive regulation of type III hypersensitivity|positive regulation of type IIa hypersensitivity|receptor activity|regulation of immune response|serotonin secretion
H3114A06	20229	Sat1	Spermidine/spermine N1-acetyl transferase 1	2.10	34.5	N-acetyltransferase activity|acyltransferase activity|diamine N-acetyltransferase activity|spermine catabolism|transferase activity
H3085G03	13057	Cyba	Cytochrome b-245, alpha polypeptide	2.15	36	electron transport|heme binding|integral to membrane|membrane|mitochondrion|oxidoreductase activity|superoxide metabolism
H3059B03	18753	Prkcd	Protein kinase C, delta	1.84	40	ATP binding|B-cell proliferation|diacylglycerol binding|humoral defense mechanism (sensu Vertebrata)|intracellular signaling cascade|membrane|protein amino acid phosphorylation|protein kinase activity|protein serine/threonine kinase activity|transferase a
H3144A05	20530	Slc31a2	Solute carrier family 31, member 2	2.20	46	carrier activity|copper ion transport|copper ion transporter activity|integral to membrane|integral to plasma membrane|transport
H3026F09	83921	Tmem2	Transmembrane protein 2	2.00	49	biological_process unknown|integral to membrane|molecular_function unknown
H3052B09	20684	Sp100	Nuclear antigen Sp100	2.08	53	DNA binding|chromosome|nucleus
H3009G07	74048	4632428N05Rik	RIKEN cDNA 4632428N05 gene	2.54	54	extracellular space|integral to membrane
H3008H01	16952	Anxa1	RIKEN cDNA C430014K04 gene	2.63	61	arachidonic acid secretion|calcium ion binding|calcium-dependent phospholipid binding|cell cycle|phospholipase A2 inhibitor activity|phospholipase inhibitor activity|protein binding|regulation of cell proliferation|sarcolemma|signal transduction
H3149A06	67689	Aldh3b1	Aldehyde dehydrogenase 3 family, member B1	2.30	65	oxidoreductase activity
H3004F04	234825	Klhdc4	Kelch domain containing 4	1.79	73	
H3092H10	27357	Gyg1	Glycogenin 1	2.02	74	carbohydrate biosynthesis|glycogen biosynthesis|transferase activity|transferase activity, transferring hexosyl groups
H3021E03	110213	Tegt	Testis enhanced gene transcript	1.89	78	apoptosis|endoplasmic reticulum|insoluble fraction|integral to membrane|integral to plasma membrane|nucleus
H3055G09	235534	Acpl2	Acid phosphatase-like 2	2.02	79	biological_process unknown|cellular_component unknown|molecular_function unknown
H3109B11	26569	Slc27a4	Solute carrier family 27 (fatty acid transporter), member 4	2.10	84	catalytic activity|lipid transport|metabolism|transport|transporter activity
H3053G12	59047	Pnkp	Polynucleotide kinase 3'- phosphatase	2.16	86	ATP binding|DNA repair|catalytic activity|hydrolase activity|kinase activity|nucleus|transferase activity

H3090F04	59031	Chst12	Carbohydrate sulfotransferase 12	1.97	88	integral to membrane|transferase activity
H3039H01	12798	Cnn2	Calponin 2	2.10	89	actin binding|calmodulin binding|smooth muscle contraction
H3075F05	109033	6230421P05Rik	RIKEN cDNA 6230421P05 gene	1.81	94	
H3144E03	22793	Zyx	Zyxin	2.31	111	cell adhesion|cell-cell adherens junction|protein binding|zinc ion binding
H3053A12	19735	Rgs2	Regulator of G-protein signaling 2	2.41	132	G-protein coupled receptor protein signaling pathway|GTPase activator activity|cell cycle|nucleus|signal transducer activity|signal transduction
H3062A04	67844	Rab32	RAB32, member RAS oncogene family	1.82	133	ATP binding|GTP binding|mitochondrion|protein amino acid phosphorylation|protein kinase activity|protein transport|small GTPase mediated signal transduction
H3119G12	108155	Ogt	O-linked N-acetylglucosamine (GlcNAc) transferase (UDP-N-acetylglucosamine:polypeptide-N-acetylglucosaminyl transferase)	2.19	203	N-acetyltransferase activity|O-linked glycosylation|intracellular|transferase activity|transferase activity, transferring glycosyl groups
H3082B08	15257	Hipk1	Homeodomain interacting protein kinase 1	1.95	217	ATP binding|PML body|cytoplasm|kinase activity|nuclear speck|nucleus|protein amino acid phosphorylation|protein binding|protein kinase activity|protein serine/threonine kinase activity|transferase activity
H3075G06	56738	Mocs1	Molybdenum cofactor synthesis 1	1.57	219	Mo-molybdopterin cofactor biosynthesis
H3138H04	14081	Acsl1	Acyl-CoA synthetase long-chain family member 1	2.05	255	acetate-CoA ligase activity|catalytic activity|fatty acid metabolism|ligase activity|long-chain-fatty-acid-CoA ligase activity|magnesium ion binding|metabolism|mitochondrion
H3122H06	59010	Sqrdl	Sulfide quinone reductase-like (yeast)	2.13	289	disulfide oxidoreductase activity|electron transport|mitochondrion|oxidoreductase activity
H3047C12	70617	5730508B09Rik	RIKEN cDNA 5730508B09 gene	1.74	302	
H3153G05	74302	Mtmr3	Myotubularin related protein 3	1.74	334	hydrolase activity|phosphoprotein phosphatase activity|protein amino acid dephosphorylation|zinc ion binding
H3003C07	320790	Chd7	Chromodomain helicase DNA binding protein 7	1.69	365	biological_process unknown|cellular_component unknown|molecular_function unknown
H3122H08		A630082K20Rik	RIKEN cDNA A630082K20 gene	2.12	382	
H3138D03	54610	Tbc1d8	TBC1 domain family, member 8	2.65	429	GTPase activator activity|biological_process unknown|cellular_component unknown
H3108A12	223601	0910001A06Rik	RIKEN cDNA 0910001A06 gene	1.51	464	
H3048B06	227929	Pscdbp	Pleckstrin homology, Sec7 and coiled-coil domains, binding protein	2.10	469	cell cortex|cytoplasm|protein binding|regulation of cell adhesion
H3115A11	107607	Card4	Caspase recruitment domain 4	2.05	475	ATP binding|I-kappaB kinase/NF-kappaB cascade|apoptosis|caspase activation|cytoplasm|defense response|induction of apoptosis|intracellular|protein binding|regulation of apoptosis
H3077A06	245860	AU019532	Expressed sequence AU019532	1.56	479	biological_process unknown|cellular_component unknown|molecular_function unknown
H3142A02	19053	Ppp2cb	Protein phosphatase 2a, catalytic subunit, beta isoform	1.54	489	CTD phosphatase activity|cytoplasm|hydrolase activity|manganese ion binding|myosin phosphatase activity|nucleus|phosphoprotein phosphatase activity|protein amino acid dephosphorylation|protein binding|protein phosphatase type 1 activity|protein phosphatase type 2A activity|protein phosphatase type 2A activity|protein phosphatase type 2C activity|protein serine/threonine phosphatase activity
H3108A03	11810	Apobec1	Apolipoprotein B editing complex 1	1.98	643	cytidine to uridine editing|hydrolase activity|hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in cyclic amidines|lipoprotein metabolism|mRNA editing|mRNA editing|mRNA processing|zinc ion binding
H3114A07	11977	Atp7a	ATPase, Cu++ transporting, alpha polypeptide	1.64	669	ATP binding|ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism|Golgi apparatus|catalytic activity|cation transport|copper ion binding|copper ion transport|copper ion transporter activity|copper-exporting ATPase activity|hydrogen:potassium-exchanging ATPase activity|hydrogen:potassium-exchanging ATPase complex|hydrolase activity|hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances|integral to membrane|magnesium ion binding|membrane|metabolism|metal ion binding|metal ion transport|metal ion transporter activity
H3150C10		0610010D24Rik	RIKEN cDNA 0610010D24 gene	1.73	671	

H3019F12	70568	Cpne3	Copine III	1.67	693	calcium-dependent phospholipid binding|lipid metabolism|protein serine/threonine kinase activity|transferase activity|transporter activity|vesicle-mediated transport
H3144F06	11520	Adfp	Adipose differentiation related protein	2.03	710	lipid particle|lipid storage|long-chain fatty acid transport|membrane|nucleus|plasma membrane|protein binding
H3094C02	54427	Dnmt3l	DNA (cytosine-5-)-methyltransferase 3-like	1.85	758	cytoplasm|embryonic development (sensu Mammalia)|imprinting|nuclear heterochromatin|nucleus|protein binding|spermatogenesis
H3002E01	27361	Sepx1	Selenoprotein X 1	2.01	809	cytoplasm|nucleus|oxidoreductase activity|protein repair|protein-methionine-R-oxide reductase activity
H3104A04	12986	Csf3r	Colony stimulating factor 3 receptor (granulocyte)	2.03	856	cell adhesion|extracellular space|hematopoietin/interferon-class (D200-domain) cytokine receptor activity|integral to membrane|membrane|neutrophil chemotaxis|protein binding|receptor activity
H3061E12	93691	Klf7	Kruppel-like factor 7 (ubiquitous)	1.51	1045	DNA binding|nucleic acid binding|nucleus|regulation of transcription, DNA-dependent|zinc ion binding
H3085H09	52024	Ankrd22	Ankyrin repeat domain 22	1.81	1256	DNA binding|integral to membrane|regulation of transcription, DNA-dependent|transcription factor activity
H3085D09	170834	Oosp1	Oocyte secreted protein 1	1.52	1620	extracellular region|extracellular space
H3021G01	83554	Fstl3	Follistatin-like 3	1.52	1679	activin binding|activin inhibitor activity|extracellular space|extracellular space|negative regulation of BMP signaling pathway
H3008E11	23849	Copeb	Core promoter element binding protein	1.71	52	DNA binding|cytokine and chemokine mediated signaling pathway|nucleic acid binding|nucleus|regulation of transcription, DNA-dependent|zinc ion binding
H3038B04	66899	Fip1l1	FIP1 like 1 (S. cerevisiae)	1.80	16.5	
H3105D09	66824	Pycard	PYD and CARD domain containing	1.62	179.5	apoptosis|apoptotic protease activator activity|cytosol|extracellular region|induction of apoptosis|inflammatory response|intracellular|negative regulation of cell cycle|protein binding|regulation of apoptosis|regulation of caspase activation
H3068B04	57816	Tesc	Tescalcin	1.61	37	calcium ion binding|cytoplasm|lamellipodium|protein binding|sodium ion homeostasis
H3124F04	56615	Mgst1	Microsomal glutathione S-transferase 1	1.77	64	glutathione metabolism|glutathione transferase activity|glutathione transferase activity|membrane|microsome|microsome|mitochondrion|mitochondrion|transferase activity
H3126H09	216197	Ckap4	Cytoskeleton-associated protein 4	1.59	69	biological_process unknown|cell surface|endoplasmic reticulum|protein binding
H3078B09	12224	Klf5	Kruppel-like factor 5	1.80	151	DNA binding|nucleic acid binding|nucleus|regulation of transcription, DNA-dependent|zinc ion binding
H3058A11		Mm.220992	Similar to: Mus musculus similar to endogenous retroviral family W, env(C7), member 1 (syncytin); envelope protein; syncytin [Homo sapiens] (LOC234826), mRNA	1.68	99	
H3032A12	14381	G6pdx	Glucose-6-phosphate dehydrogenase X-linked	1.56	48	carbohydrate metabolism|glucose metabolism|glucose-6-phosphate 1-dehydrogenase activity|glucose-6-phosphate 1-dehydrogenase activity|glutathione metabolism|oxidoreductase activity
H3024F02	21452	Tcn2	Transcobalamin 2	1.64	278	binding|cobalt ion transport|cobalt ion transporter activity|extracellular space|transport
H3111G06	53612	Arg2	Arginase type II	1.93	87	arginase activity|arginine catabolism|arginine metabolism|catalytic activity|hydrolase activity|manganese ion binding|mitochondrion|urea cycle
H3095G06	11891	Rab27a	RAB27A, member RAS oncogene family	1.54	167.5	GTP binding|GTPase activity|blood coagulation|intracellular|protein binding|protein transport|small GTPase mediated signal transduction|vesicle-mediated transport
H3152B12	13134	Dach1	Dachshund 1 (Drosophila)	1.75	2539	DNA binding|cellular physiological process|development|nucleus|protein binding|regulation of transcription|regulation of transcription|regulation of transcription, DNA-dependent|transcription factor activity|transcription factor complex

Note: Average V fc (not shown) was neither up nor down regulated and had values between 1.49 and -1.35

Table 15B: Differentiation-associated genes with sample effect, alpha

ID	LLID	Name	Gene Name	Avg F fc	Avg sample diff	Rank	GO abv - source
H3109G06	17002	Ltf	Lactotransferrin	5.62	-2.37	1.5	extracellular region|extracellular space|ferric iron binding|iron ion homeostasis|iron ion transport|transport
H3083G02	16819	Lcn2	Lipocalin 2	3.29	-2.42	5	binding|extracellular space|transport|transporter activity
H3115B07	20202	S100a9	S100 calcium binding protein A9 (calgranulin B)	2.84	-2.95	6	calcium ion binding
H3122E03	56431	Dstn	Destrin	2.28	-1.70	16.5	actin binding|intracellular
H3149B11	17524	Mpp1	Membrane protein, palmitoylated	2.37	-1.48	77	membrane|protein binding
H3151D11	19156	Psap	Prosaposin	2.47	-1.74	19	extracellular space|lipid metabolism|lysosome|mitochondrion|sphingolipid metabolism
H3123D05	110095	Pygl	Liver glycogen phosphorylase	2.14	-1.43	67	carbohydrate metabolism|catalytic activity|glycogen metabolism|glycogen phosphorylase activity|phosphorylase activity|transferase activity|transferase activity, transferring glycosyl groups
H3080C05	11501	Adam8	A disintegrin and metalloprotease domain 8	2.13	1.49	56	cell-cell adhesion|hydrolase activity|integral to membrane|integrin-mediated signaling pathway|metalloendopeptidase activity|metallopeptidase activity|peptidase activity|plasma membrane|proteolysis and peptidolysis|zinc ion binding
H3060F11	13058	Cybb	Cytochrome b-245, beta polypeptide	2.25	-1.85	24	electron transport|integral to membrane|ion channel activity|membrane|oxidoreductase activity|voltage-gated ion channel activity
H3037D09	72141	Adpgk	ADP-dependent glucokinase	2.73	-1.31	57	extracellular space
H3111A02	54683	Prdx5	Peroxiredoxin 5	2.28	-1.59	104	antioxidant activity|mitochondrion|peroxisome|peroxisome organization and biogenesis
H3114D11	12226	Btg1	B-cell translocation gene 1, anti-proliferative	2.03	1.66	50	cell proliferation|cellular physiological process|cytoplasm|negative regulation of cell proliferation|negative regulation of cell proliferation|nucleus|positive regulation of myoblast differentiation|protein amino acid methylation|protein binding
H3048E05	16993	Lta4h	Leukotriene A4 hydrolase	2.01	-1.38	235	catalytic activity|hydrolase activity|leukotriene biosynthesis|leukotriene-A4 hydrolase activity|membrane alanyl aminopeptidase activity|metallopeptidase activity|proteolysis and peptidolysis|zinc ion binding
H3117E08	110157	Raf1	V-raf-1 leukemia viral oncogene 1	2.01	-1.42	122	ATP binding|cellular physiological process|cytosol|diacylglycerol binding|intracellular|intracellular signaling cascade|kinase activity|plasma membrane|protein amino acid phosphorylation|protein binding|protein kinase activity|protein kinase cascade|protein serine/threonine kinase activity|receptor signaling protein activity|regulation of cell cycle|signal transduction|transferase activity
H3034B09	13350	Dgat1	Diacylglycerol O-acyltransferase 1	2.03	-1.27	75	acyltransferase activity|diacylglycerol O-acyltransferase activity|endoplasmic reticulum|integral to membrane|transferase activity
H3034A10	18793	Plaur	Urokinase plasminogen activator receptor	2.11	-1.63	96.5	U-plasminogen activator receptor activity|cell surface receptor linked signal transduction|extracellular space|integral to membrane|receptor activity
H3155B04	11990	Atrn	Attractin	1.89	-1.48	403	development|inflammatory response|integral to membrane|membrane|receptor activity|regulation of body size|structural molecule activity|sugar binding
H3008A09	11852	Rhob	Ras homolog gene family, member B	2.00	-1.51	446	GTP binding|GTPase activity|Rho protein signal transduction|angiogenesis|apoptosis|cell adhesion|cell cycle|cellular physiological process|membrane|negative regulation of cell cycle|nucleus|protein binding|protein transport|small GTPase mediated signal tr
H3030E10	13030	Ctsb	Cathepsin B	1.71	-1.17	192	cathepsin B activity|cysteine-type endopeptidase activity|cysteine-type peptidase activity|extracellular space|hydrolase activity|lysosome|mitochondrion|peroxidase activity|protein targeting|proteolysis and peptidolysis
H3150G10	110829	Lims1	LIM and senescent cell antigen-like domains 1	1.68	-1.19	273	cell-matrix adhesion|focal adhesion|protein binding
H3155C10	22236	Ugt1a6	UDP glycosyltransferase 1 family, polypeptide A6	1.81	-2.00	109	glucuronate metabolism|glucuronate metabolism|glucuronosyltransferase activity|glucuronosyltransferase activity|transferase activity|transferase activity, transferring glycosyl groups

H3042D05	54353	Scap2	Src family associated phosphoprotein 2	1.76	-1.14	121	cytoplasm|negative regulation of cell proliferation|plasma membrane
H3029G05	13139	Dgka	Diacylglycerol kinase, alpha	1.80	-1.17	329	calcium ion binding|cytosol|diacylglycerol binding|diacylglycerol kinase activity|intracellular signaling cascade|kinase activity|phospholipid binding|plasma membrane|protein kinase C activation|transferase activity
H3092G05	11745	Anxa3	Annexin A3	1.80	-1.55	1047	calcium ion binding|calcium-dependent phospholipid binding|phospholipase A2 inhibitor activity|phospholipase inhibitor activity
H3146F06	20660	Sorl1	Sortilin-related receptor, LDLR class A repeats-containing	1.76	-1.27	538	cholesterol metabolism|copper ion binding|development|electron transport|electron transporter activity|endocytosis|extracellular space|integral to membrane|lipid transport|lipid transporter activity|membrane|receptor activity
H3085D10	14544	Gda	Guanine deaminase	1.94	-1.32	307.5	extracellular space|hydrolase activity
H3128C07	12192	Zfp36l1	Zinc finger protein 36, C3H type-like 1	2.27	-2.36	321	DNA binding|nucleic acid binding|nucleus
H3134E06		H3134E06		1.94	-2.06	119	
H3012H07	12505	Cd44	CD44 antigen	1.99	-2.25	424	cell adhesion|external side of plasma membrane|hyaluronic acid binding|integral to membrane|membrane|protein binding|receptor activity|ureteric bud branching
H3137F12	338467	Zcwcc3	Zinc finger, CW-type with coiled-coil domain 3	1.85	-1.70	774	
H3138G08	67712	Mscp	Mitochondrial solute carrier protein	1.71	-1.27	784	binding|integral to membrane|mitochondrial inner membrane|transport
H3036D03	84505	Setdb1	SET domain, bifurcated 1	1.56	-1.33	613.5	DNA binding|chromatin modification|histone-lysine N-methyltransferase activity|inner cell mass cell proliferation|methyltransferase activity|nucleic acid binding|nucleus|transferase activity|zinc ion binding
H3060C07	338523	A630082K20Rik	RIKEN cDNA A630082K20 gene	1.81	-0.62	141	
H3054D12		H3054D12		1.74	-2.49	271	
H3148G09	259302	Srgap3	SLIT-ROBO Rho GTPase activating protein 3	2.09	-1.27	1188.5	GTPase activator activity|biological_process unknown|cellular_component unknown|protein binding
H3049G05		Ggta1	glycoprotein galactosyltransferase alpha 1, 3 (Ggta1)	1.54	-1.94	970.5	
H3008H10	70568	Cpne3	Copine III	1.80	-1.12	676	calcium-dependent phospholipid binding|lipid metabolism|protein serine/threonine kinase activity|transferase activity|transporter activity|vesicle-mediated transport
H3146H02	218452	1110007C02Rik	RIKEN cDNA 1110007C02 gene	1.57	-1.52	662	
H3052D06		4833415F11Rik	Adult male testis cDNA, RIKEN full-length enriched library, clone:4933428M03 product:hypothetical BTB/POZ domain/Methylated-DNA--protein-cysteine methyltransferase containing protein, full insert sequence	1.71	-2.92	894	
H3136A03	67266	2900024C23Rik	RIKEN cDNA 2900024C23 gene	1.58	-1.43	3313	
H3122H07	233549	Mogat2	Monoacylglycerol O-acyltransferase 2	1.67	-1.78	841	2-acylglycerol O-acyltransferase activity|acetyltransferase activity|acyltransferase activity|diacylglycerol biosynthesis|endoplasmic reticulum|lipid metabolism|monoacylglycerol metabolism|transferase activity|triacylglycerol biosynthesis
H3028F03	13039	Ctsl	Cathepsin L	1.56	-1.63	3122	actinidain activity|cathepsin L activity|cysteine-type endopeptidase activity|cysteine-type peptidase activity|extracellular space|hydrolase activity|lysosome|peptidase activity|protein binding|proteolysis and peptidolysis
H3054F06	271221	5031414D18Rik	RIKEN cDNA 5031414D18 gene	1.53	-1.45	1066	

H3136C04		A330103N21Rik	RIKEN cDNA A330103N21 gene	1.78	-1.41	2101.5	
H3109B05	17002	Ltf	Lactotransferrin	5.26	-2.23	1.5	extracellular region|extracellular space|ferric iron binding|iron ion homeostasis|iron ion transport|transport
H3087A12	17002	Ltf	Lactotransferrin	4.80	-2.17	3	extracellular region|extracellular space|ferric iron binding|iron ion homeostasis|iron ion transport|transport
H3131F12	20861	Stfa1	Stefin A1	2.73	-2.81	8	cysteine protease inhibitor activity|endopeptidase inhibitor activity|intracellular
H3076D09		1190003K14Rik	RIKEN cDNA 1190003K14 gene	2.84	-1.40	12.5	
H3089C03	69202	2610009E16Rik	RIKEN cDNA 2610009E16 gene	1.97	2.22	18	
H3111G08	67800	Dgat2	Diacylglycerol O-acyltransferase 2	2.41	-2.20	27	acyltransferase activity|diacylglycerol O-acyltransferase activity|integral to membrane|integral to membrane|transferase activity|triacylglycerol biosynthesis|triacylglycerol biosynthesis
H3120B07	227753	Gsn	Gelsolin	1.59	1.48	169	actin binding|actin cytoskeleton|actin filament severing|barbed-end actin filament capping|calcium ion binding|cytoplasm|cytoskeleton|extracellular space|lamellipodium|structural molecule activity|vesicle-mediated transport
H3043A06	14782	Gsr	glutathione reductase 1	1.76	-1.38	100	
H3154H07		H3154H07		1.88	-1.86	127	
H3088C07	381411	Mm.9889	similar to 1-aminocyclopropane-1-carboxylate synthase	1.69	-1.48	207	
H3139F01	20652	Soat1	Sterol O-acyltransferase 1	2.07	-1.83	241	acyltransferase activity|cholesterol metabolism|endoplasmic reticulum|integral to membrane|sterol O-acyltransferase activity|transferase activity
H3069G05	114713	Rasa2	RAS p21 protein activator 2	1.50	-1.32	173	GTPase activator activity|intracellular signaling cascade
H3067H02		1200015M12Rik		1.79	-2.18	233	
H3053C11		1200015M12Rik		1.66	-2.66	804	
H3064C11		3930401B19Rik		1.51	-2.97	824	

Note: Average V fc (not shown) was neither up nor down regulated and had values between 1.43 and -1.29

Table 15C: Proliferation-associated genes

ID	LLID	Name	Gene Name	Avg F fc	Rank	GO abv - source
H3052F12		Serpina3g	Serine (or cysteine) proteinase inhibitor, clade A, member 3G	-3.88	4	
H3095G01	107272	Psat1	Phosphoserine aminotransferase 1	-2.38	10	L-serine biosynthesis|metabolism|transaminase activity|transferase activity
H3029C03	27205	Podxl	Podocalyxin-like	-2.95	7	cellular physiological process|extracellular space|integral to membrane|plasma membrane
H3095C02	108682	Gpt2	Glutamic pyruvate transaminase (alanine aminotransferase) 2	-1.81	12.5	transaminase activity|transferase activity
H3118E11		H3118E11		-3.04	31	
H3035E09	30877	C77032	EST C77032	-1.59	28	biological_process unknown|molecular_function unknown|nucleolus
H3032F06	15135	Hbb-bh1	Hemoglobin Z, beta-like embryonic chain	-2.46	68	hemoglobin complex|oxygen transport|oxygen transporter activity|transport
H3158E03	17768	Mthfd2	Methylenetetrahydrofolate dehydrogenase (NAD+ dependent), methenyltetrahydrofolate cyclohydrolase	-1.64	26	catalytic activity|extracellular space|folic acid and derivative biosynthesis|hydrolase activity|magnesium ion binding|methenyltetrahydrofolate cyclohydrolase activity|methylenetetrahydrofolate dehydrogenase (NADP+) activity|mitochondrion|one-carbon compo
H3136C07	12176	Bnip3	BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3	-1.66	42	apoptosis|integral to membrane|mitochondrion|mitochondrion|regulation of apoptosis
H3096B01		5930412K08	hypothetical Laminin-G domain/Fibronectin type III domain/EGF-like domain/Calcium-binding EGF-like domain containing protein	-1.72	110	
H3027E04	12035	Bcat1	Branched chain aminotransferase 1, cytosolic	-1.51	492	branched chain family amino acid biosynthesis|branched chain family amino acid catabolism|branched chain family amino acid metabolism|branched-chain-amino-acid transaminase activity|branched-chain-amino-acid transaminase activity|catalytic activity|cytoso
H3123E01	71846	1700013H19Rik	RIKEN cDNA 1700013H19 gene	-1.62	670	
H3096D08		H3096D08		-2.03	134	
H3007E12		H3007E12		-2.45	33	
H3137F01		H3137F01		-2.09	150	
H3012D06	26557	Homer2	Homer homolog 2 (Drosophila)	-2.50	128	GKAP/Homer scaffold activity|actin binding|cytoplasm|metabotropic glutamate receptor signaling pathway|protein binding
H3115H09	13797	Emx2	Empty spiracles homolog 2 (Drosophila)	-2.14	299	DNA binding|anterior/posterior pattern formation|brain development|development|neuron differentiation|nucleus|regulation of transcription, DNA-dependent|transcription factor activity
H3059F12	21822	Tgtp	T-cell specific GTPase	-1.83	70	GTP binding|cellular_component unknown
H3121H09	110460	Acat2	Acetyl-Coenzyme A acetyltransferase 2	-1.69	635	acetyl-CoA C-acetyltransferase activity|acyltransferase activity|mitochondrion|transferase activity
H3138E12	278279	8430438D04Rik	RIKEN cDNA 8430438D04 gene	-2.97	126	integral to membrane
H3143G03	66902	Mtap	Methylthioadenosine phosphorylase	-1.60	396	S-methyl-5-thioadenosine phosphorylase activity|nucleobase, nucleoside, nucleotide and nucleic acid metabolism|transferase activity|transferase activity, transferring glycosyl groups|transferase activity, transferring pentosyl groups
H3154F05	68539	1110006I15Rik	RIKEN cDNA 1110006I15 gene	-2.05	297.5	extracellular space|integral to membrane

Note: Average V fc (not shown) was neither up nor down regulated and had values between 1.33 and -1.28 except for the first gene where V fc = 1.64

Table 15D: Proliferation-associated genes with sample effect, alpha

ID	LLID	Name	Gene Name	Avg F fc	Avg sample diff	Rank	GO abv - source
H3026C11	237353	4831416G18Rik	RIKEN cDNA 4831416G18 gene	-1.64	-1.84	748	
H3008F07	14969	H2-Eb1	Histocompatibility 2, class II antigen E beta	-2.08	2.27	51	MHC class II receptor activity|antigen presentation, exogenous antigen|antigen presentation, exogenous antigen via MHC class II|antigen presentation, exogenous peptide antigen|antigen processing, exogenous antigen via MHC class II|extracellular space|immu
H3152A02	20440	St6gal1	Beta galactoside alpha 2,6 sialyltransferase 1	-1.90	1.69	43	Golgi apparatus|beta-galactoside alpha-2,6-sialyltransferase activity|integral to membrane|membrane|protein amino acid glycosylation|sialyltransferase activity|transferase activity|transferase activity, transferring glycosyl groups
H3138D09	140557	Smc1l2	SMC (structural maintenace of chromosomes 1)-like 2 (S. cerevisiae)	-1.57	2.34	157.5	ATP binding|ATPase activity|ATPase activity, coupled to transmembrane movement of substances|DNA binding|cell cycle|chromosome organization and biogenesis (sensu Eukaryota)|chromosome segregation|chromosome, pericentric region|meiosis|membrane|nucleus|pro
H3154E06	15903	Idb3	Inhibitor of DNA binding 3	-1.68	1.96	1580	negative regulation of transcription from Pol II promoter|nucleus|protein binding|protein domain specific binding
H3081H02	27384	Akr1c13	Aldo-keto reductase family 1, member C13	-2.14	2.35	25	aldo-keto reductase activity|oxidoreductase activity|xenobiotic metabolism
H3118G08	110956	D17H6S56E-5	DNA segment, Chr 17, human D6S56E 5	-2.07	1.52	45	viral envelope
H3156F01		H3156F01		-2.34	1.51	91.5	
H3095E04		Fv4	Friend virus susceptibility 4	-1.81	1.21	125	
H3114H02	12490	Cd34	CD34 antigen	-1.77	1.49	38	cell adhesion|external side of plasma membrane|integral to membrane|membrane|protein binding
H3041H12	16878	Lif	Leukemia inhibitory factor	-2.59	1.58	244	cytokine activity|extracellular region|extracellular space|growth factor activity|immune response|leukemia inhibitory factor receptor binding
H3017B11	53328	Pgrmc1	Progesterone receptor membrane component 1	-1.75	1.30	425.5	integral to membrane|microsome|receptor activity|steroid binding
H3097D08		H3097D08		-1.71	1.26	283	
H3150B05	14461	Gata2	GATA binding protein 2	-1.83	1.81	251	DNA binding|neuron differentiation|nucleus|regulation of transcription, DNA-dependent|transcription factor activity
H3008G04	14461	Gata2	GATA binding protein 2	-1.65	1.95	268	DNA binding|neuron differentiation|nucleus|regulation of transcription, DNA-dependent|transcription factor activity
H3021B06	286940	Flnb	Filamin, beta	-1.64	1.23	176	actin binding|cytoplasm|cytoskeleton|focal adhesion|myogenesis|protein binding|stress fiber
H3015F01	66889	Rnf128	Ring finger protein 128	-1.82	2.63	242.5	Golgi apparatus|endoplasmic reticulum|extracellular space|integral to membrane|late endosome|ligase activity|negative regulation of cytokine biosynthesis|peptidase activity|proteolysis and peptidolysis|ubiquitin-protein ligase activity|ubiquitin-protein l
H3024D03	13664	Eif1a	Eukaryotic translation initiation factor 1A	-1.77	0.54	498.5	eukaryotic translation initiation factor 4F complex|translation factor activity, nucleic acid binding
H3121B01	15122	Hba-a1	Hemoglobin alpha, adult chain 1	-1.68	1.94	1050	hemoglobin complex|mitochondrion|oxygen transport|oxygen transporter activity|transport
H3051E10	30838	Fbxw4	F-box and WD-40 domain protein 4	-1.60	1.42	535	Wnt receptor signaling pathway|development|ubiquitin cycle
H3029F02	18195	Nsf	N-ethylmaleimide sensitive fusion protein	-1.51	1.36	3761	ATP binding|ATPase activity, coupled|Golgi apparatus|endoplasmic reticulum|hydrolase activity|magnesium ion binding|nucleoside-triphosphatase activity|nucleotide binding|protein transport|syntaxin binding|transport

Note: Average V fc (not shown) was neither up nor down regulated and had values between 1.07 and -1.3 except for the first gene where V fc = 1.51

Table 15E: IL-3 associated genes with and without sample effect

ID	LLID	Name	Gene Name	Avg F fc	Avg V fc	Avg sample diff	Rank	GO abv - source
H3157G08	66090	Ypel3	Yippee-like 3 (Drosophila)	3.24	1.54		30	
H3151H12	66168	Grina	Glutamate receptor, ionotropic, N-methyl D-asparate-associated protein 1 (glutamate binding)	3.38	1.56		120	integral to membrane|receptor activity
H3154B07	16432	Itm2b	Integral membrane protein 2B	2.44	1.50		236	ATP binding|biological_process unknown|integral to membrane|membrane fraction
H3121E04	20148	Dhrs3	Dehydrogenase/reductase (SDR family) member 3	2.69	1.57	1.72	44	extracellular space|integral to membrane|metabolism|oxidoreductase activity
H3054F05	17105	Lyzs	Lysozyme	3.86	1.88	-3.23	124	carbohydrate metabolism|catalytic activity|cell wall catabolism|cytolysis|defense response to bacteria|extracellular region|extracellular space|hydrolase activity|hydrolase activity, acting on glycosyl bonds|lysozyme activity|mitochondrion
H3139E04	12266	C3	Complement component 3	2.31	1.79	-2.79	3657	complement activation|complement activation, alternative pathway|complement activation, classical pathway|endopeptidase inhibitor activity|extracellular region|extracellular space|inflammatory response|positive regulation of phagocytosis|positive regulati
H3058A06	67041	Oxct1	3-oxoacid CoA transferase 1	-2.26	-1.56	1.26	392	3-oxoacid CoA-transferase activity|CoA-transferase activity|extracellular space|metabolism|mitochondrion|mitochondrion|succinyl-CoA metabolism|transferase activity

4.2.6.3 Using Gene Ontologies (GO) to gain further biological insight into differentially expressed genes

For each of the above lists (Table 15), gene ontology (GO) was determined for each gene with the use of DAVID (Dennis *et al.*, 2003). GO annotation provides a structural language that can be applied to the differentially expressed genes on the array and used to examine functional categories in common between them (Ashburner *et al.,* 2000). Using the software application EASE (Hosack *et al.,* 2003) significance could be assigned to a given gene's ontology based on the occurrence of that ontology in the entire microarray data set. In this way, an over-representation analysis could be performed on the GO terms in common from the genes stated in Table 15.

(a) Over-representation analysis

In the entire list of differentially expressed genes from the F-stat data (i.e. all genes mentioned in Table 15), the significant biological processes identified (using a Bonferroni score of less than 0.05) included defence response, immune response and response to biotic stimulus. Using a significant EASE score of less than 0.01 the list was expanded to include inflammatory response, innate immune response, response to pest/pathogen/parasite, response to wounding, response to external stimulus and transition metal ion transport (see Table 16). These additional categories identified using the EASE cut-off score of 0.01 are consistent with the fact that the Bonferroni score is considered more stringent than the EASE score (Hosack *et al.*, 2003). Similarly, the significant molecular functions identified (using an EASE cut-off score of 0.01) involved transferase activity (in particular involving acyl groups) and cation and metal ion transporter activity (Table 16). The GO cellular component revealed extracellular and extracellular space as significant categories (using an EASE cut-off of 0.01) (Table 16). However, none of the molecular functions or cellular components were significant based on a Bonferroni score of 0.05.

(b) Reclassification of lists

To elucidate the biological processes and molecular functions within the groups specified in Table 14, three redefined lists were created. The first list contained the genes associated with differentiation that did not have a significant sample effect, the "differentiation-associated" list (68 genes). The second list contained the genes associated with proliferation that did not have a significant sample effect, the

"proliferation-associated" list (22 genes). The third list contains all the genes that had a sample effect (i.e. genes that differed between V449E and FIΔ), which included the differentiation, proliferation and IL-3 withdrawal genes from the bottom half of Table 14 (84 genes in total). The gene ontologies specific to these three redefined lists are described in Figure 27. Based on these groupings it was predicted that inferences can be made on the type of broad biological processes and molecular functions that are associated with FIΔ differentiation versus V449E proliferation.

For the proliferation-associated list examining the GO classification "biological process" revealed biosynthesis and organic acid metabolism as the only categories with three genes. Using the GO classification "molecular function" the only category with greater than three genes was transferase activity (data not shown in Figure 27). GO analyses for the other lists are as shown in Figure 27.

System	Gene Category	List Hits	EASE score	Bonferroni score
GO Biological Process	defense response	14	0.0000058	0.004
GO Biological Process	immune response	12	0.0000198	0.014
GO Biological Process	response to biotic stimulus	14	0.0000438	0.031
GO Biological Process	inflammatory response	6	0.0002133	0.149
GO Biological Process	innate immune response	6	0.0002641	0.185
GO Biological Process	response to pest/pathogen/parasite	8	0.0003884	0.272
GO Biological Process	response to wounding	6	0.0018441	1.0
GO Biological Process	response to external stimulus	14	0.0032260	1.0
GO Biological Process	transition metal ion transport	4	0.0081093	1.0
GO Cellular Component	extracellular space	28	0.0009587	0.671
GO Cellular Component	extracellular	28	0.0073935	1.000
GO Molecular Function	acyltransferase activity	7	0.0012561	0.879
GO Molecular Function	transferase activity\, transferring groups other than amino-acyl groups	7	0.0013987	0.979
GO Molecular Function	transferase activity\, transferring acyl groups	7	0.0019015	1.0
GO Molecular Function	transferase activity	25	0.0028252	1.0
GO Molecular Function	transition metal ion transporter activity	4	0.0028583	1.0
GO Molecular Function	di-\, tri-valent inorganic cation transporter activity	4	0.0036484	1.0
GO Molecular Function	O-acyltransferase activity	4	0.0067635	1.0
GO Molecular Function	acylglycerol O-acyltransferase activity	3	0.0087453	1.0

Table 16: Over-representation analysis. Gene ontology (GO) terms for all the differentially expressed genes from the F-stat analysis were determined using EASE (Hosack *et al*., 2003). "List Hits" counts the number of differentially expressed genes identified for a given GO category. A significant GO term was defined by having an EASE score of less than 0.01.

Figure 27A-D: Gene Ontology (GO) annotation to the three lists (differentiation-associated, proliferation-associated and sample effect) using EASE. Categories containing less than three genes and the unclassified category (containing all genes that could not be assigned a GO term) were not included. (A) Differentiation-associated list showing the GO classification by biological process, (B) differentiation-associated list showing the GO classification by molecular function, (C) sample effect list showing the GO classification by biological process and (D) sample effect list showing the GO classification by molecular function. For the proliferation-associated list examining the GO classification "biological process", biosynthesis and organic acid metabolism were the only categories with greater than three genes, whereas for the GO classification "molecular function" the only category was transferase activity (pie diagrams not shown).

Figure 27A

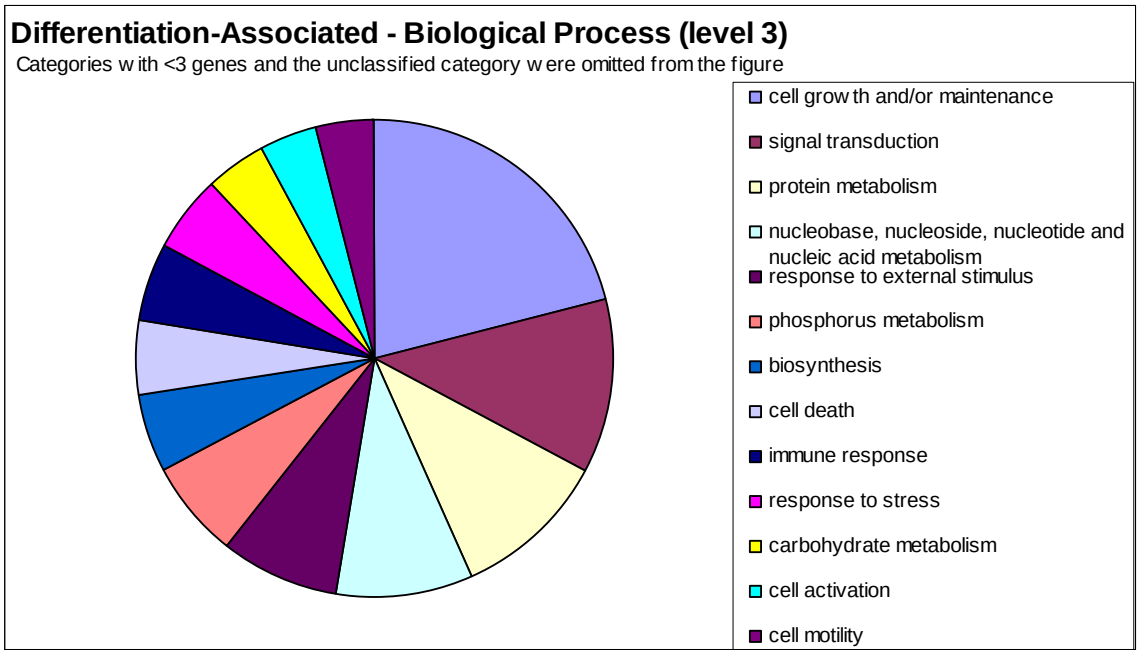


Figure 27B

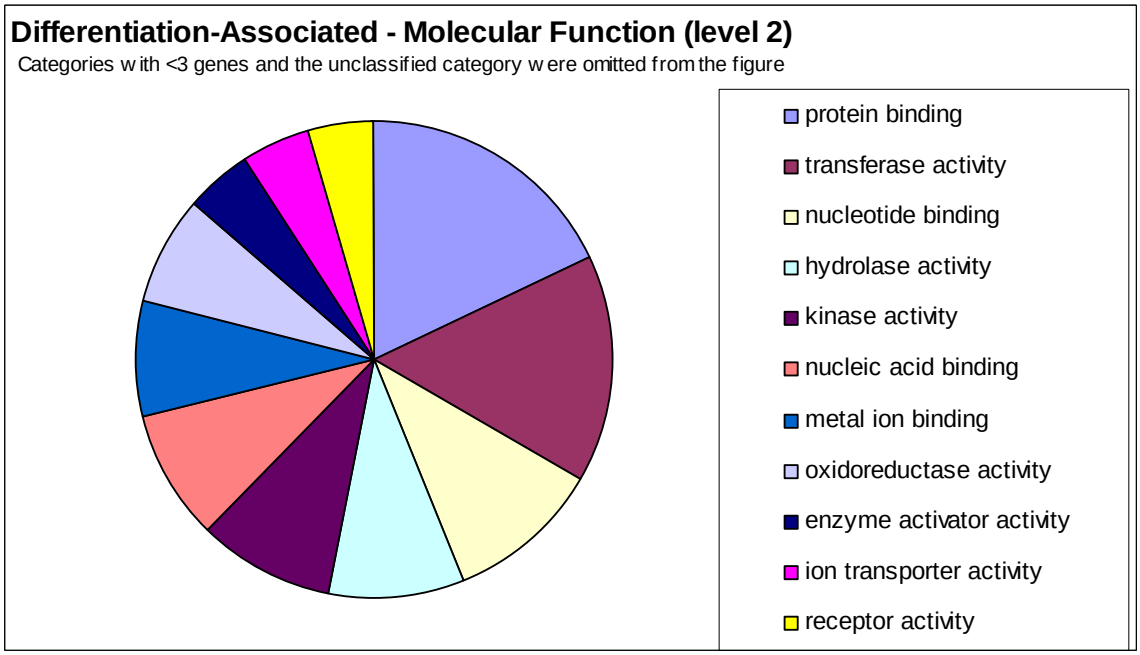


Figure 27C

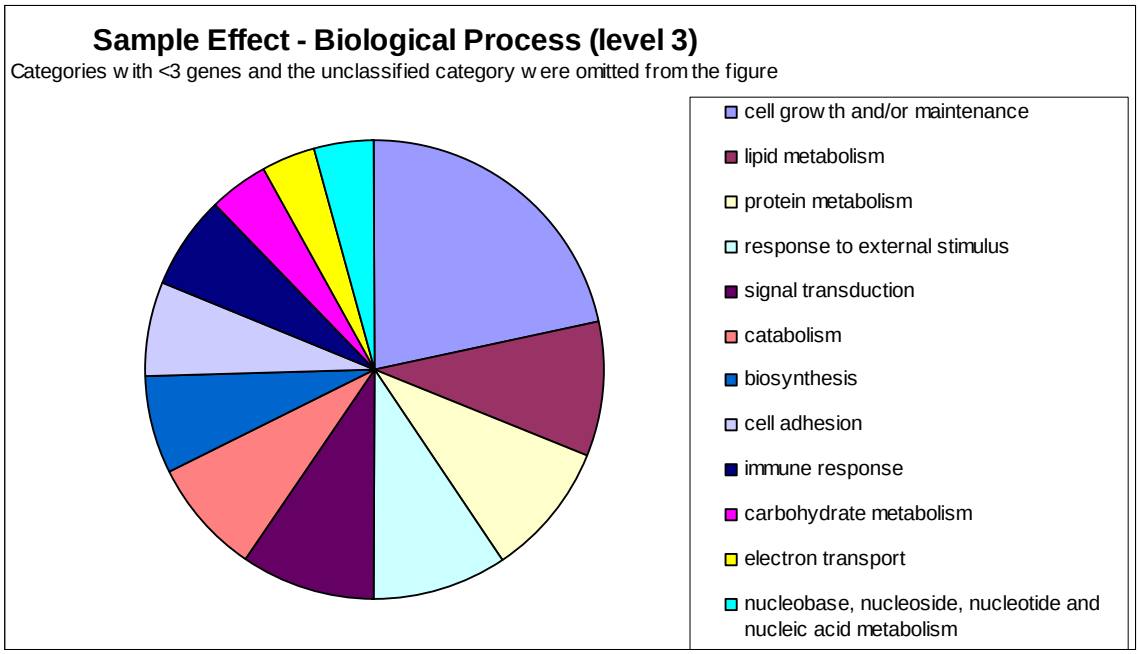
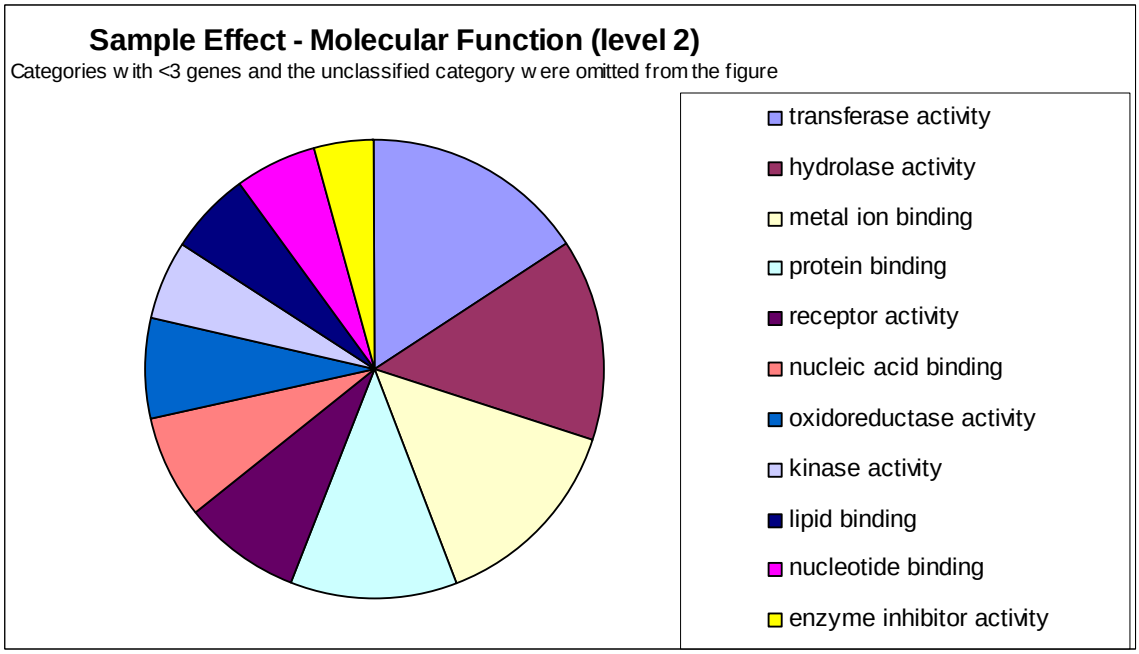


Figure 27D



4.3 Discussion

4.3.1 Experimental design and analysis of the microarray data

4.3.1.1 Experimental design

The design utilized in this experiment (Figure 15) was superior to the commonly used reference design (Figure 14) in that it achieves lower estimates of variance for all parameters measured (Glonek and Solomon, 2004). Subsequent work by Glonek and Solomon (2004) has determined that the design used in this chapter could be improved if the cross over hybridisations (slides 7 & 8) were replaced with dye swaps of slides 5 and 6 (see Glonek G.F. and Solomon P.J., (2004) for further details).

It is desirable to balance the number of times each cDNA probe is labelled with respect to the use of either the Cy3 or Cy5 probe (Glonek and Solomon, 2004). However, in this design the hybridisations were performed such that a confounding effect was introduced in the labelling reactions by (for example) labelling the V449E time 0 sample more times with the red Cy5 dye than the green Cy3 (see Figure 15). This can introduce additional variations in the data if either dye causes a dye specific effect (known as a dye bias) into the experiment. Normalisation should remove any dye specific effects (see Section 4.2.2). Furthermore, only minimal effects due to dye biases were seen in the dye swap slides of this microarray experiment, and therefore this confounding effect was unlikely to have had a significant effect on the analysis presented in this chapter.

4.3.1.2 Analysis

The development of the most appropriate methods for the analysis of microarray data is an area of rapid and continuing development (Dudoit *et al.*, 2000; Tseng *et al.*, 2001; Lonnstedt and Speed, 2002; Smyth *et al.*, 2003; Smyth, 2004). At the time the microarray data were analysed, methods were less advanced and a level of uncertainty existed over which of the available methods would be most suitable for the analysis of these experiments. Therefore, a broad approach was taken to acquire as much information as possible, accepting a potential cost of more false positives. The data was organised as a ranked list, to include the two separate scans of the arrays, the two measures of gene intensity (mean and median) and the two approaches to assess significance (t-stat and B-stat). The comprehensive approach enabled the search for genes that, on average, performed consistently well across all the scans and analysis

methods. It was a robust method of examining the gene lists for significantly differentially expressed genes, which still permitted the examination of the high ranked genes from any of the individual lists.

This microarray experiment was unique in that it combined data from two scanners in order to find differentially expressed genes. Since both scans produced high quality data (as assessed via M versus A plots) and neither could be assessed to be superior to the other, they were both used. However, as shown in Figure 19, Figure 23 and Figure 24 there were clear differences between the two scans prior to normalisation. The scans completed at PMCI were performed immediately after hybridisation, whereas the Adelaide scans were performed several days later using a different microarray scanner. Thus, differences between the two sets of scans could be due to the two scanners and settings used or due to a possible loss of signal on the slides over time. After normalisation, combining the two scans was adopted as a method of increasing the robustness of this study. A possible flaw in this approach is that if one scanner is more sensitive at detecting differential expression in the lowly expressed genes, then by combining the scans and looking for genes that performed consistently well across both scans, these lowly expressed genes could be over-looked. The advantage of this strategy is that any spurious or non-reproducible result from one scan that did not occur in the other scan meant that that gene was not found to be statistically significant.

4.3.2 Ranked gene list and F-stat analysis lists from the microarray analysis

4.3.2.1 Summary of the gene lists produced

The microarray analysis yielded (1) a ranked list of all genes based on statistical significance (assessed via the γ parameter) from which a Top 200 gene list was generated and (2) gene lists obtained from F-stat analysis that represented genes with significant β and/or α parameters (see Figure 28). With regard to the ranked list, it is difficult to establish where the cut-off is between the real and false positives. The Top 200 list was presented as a list of differentially expressed genes identified from the microarray analysis; however, it is likely that there are other significantly expressed genes that are found outside this list. To identify where the cut-off exists in the ranked data requires another method of gene expression analysis (e.g. real-time PCR or Northern Blots) to confirm the differential expression of each gene. The information obtained from these other methods could then be collated to determine at what point the

microarray analysis begins to disagree with an alternative quantitative method of measuring gene expression. Another approach would be to replicate the microarray experiment in order to gather more information on each of the genes examined on the microarrays. Replication could involve both biological replicates (additional independently prepared samples) and technical replicates, which would prove especially useful in determining the variability of each gene's expression. A powerful way to gather more information on gene expression changes within a microarray experiment is to examine expression over a time-course. Examination of other time points would provide further insights into the expression profile of the genes involved and may help validate some of the genes by revealing co-ordinated and/or sustained gene expression profiles. Such a study has now been performed using the FDB-1 system (Brown *et al.*, 2006) and there is extensive overlap between the gene lists (discussed further in Chapter 6).

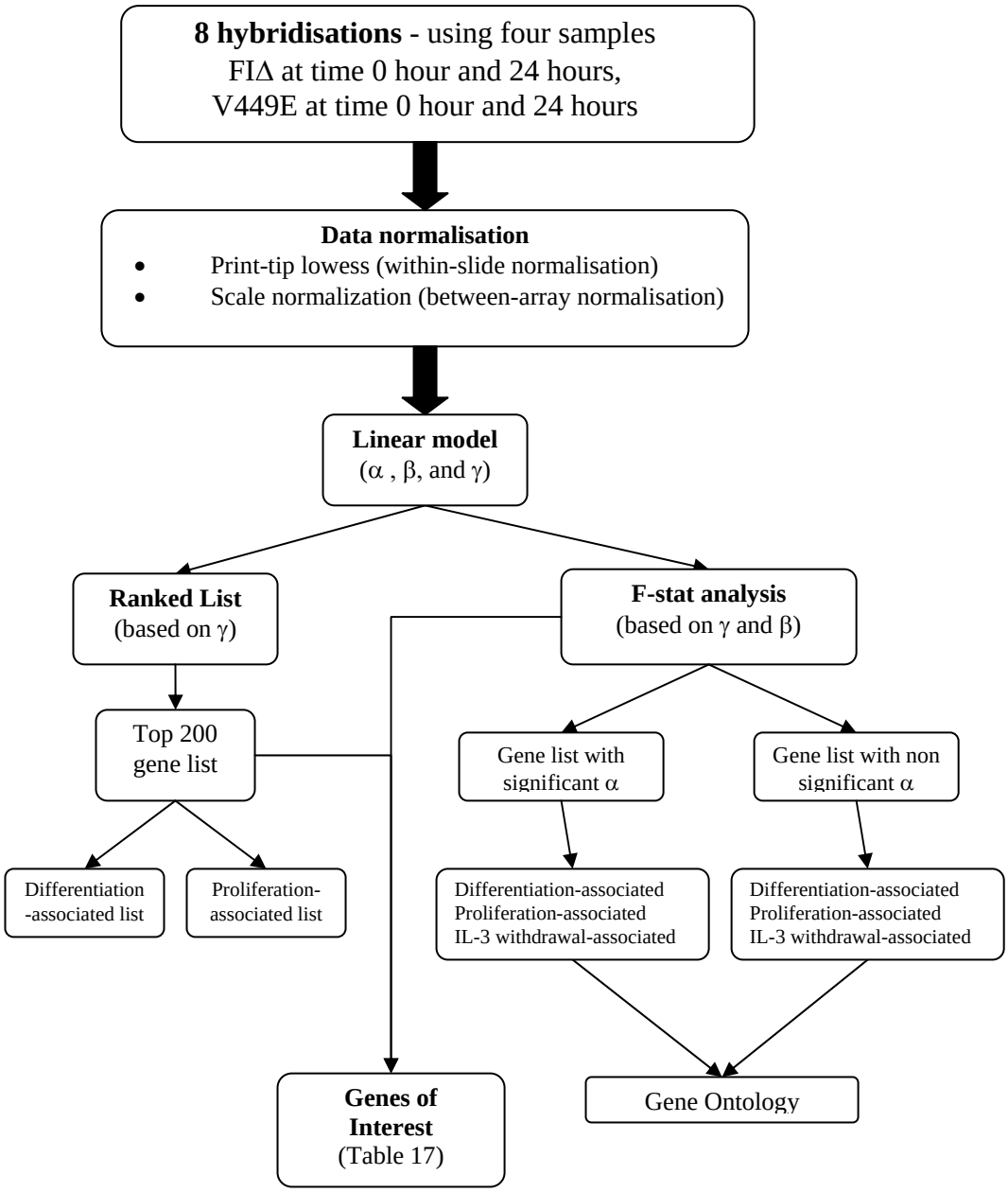


Figure 28: Flow diagram of the microarray experiment. The four samples were hybridised on eight NIA 15K microarrays and then scanned at either PMCI or Adelaide. The data was then normalised in a two-step process. The normalised data was then assessed for significantly differentially expressed genes by the use of a linear model examining the parameters α, β and γ. A ranked list of genes was obtained by examining genes with significant y parameter and from this list the first 200 genes were defined as the "Top 200" list. Alternatively the F-statistic (F-stat) was used to examine genes with significant γ and β parameters and used to define three subgroups of genes. Gene ontology was determined for all the genes identified from the F-stat analysis. Finally a summary list of interesting genes was defined from both the Top 200 gene list and F-stat analysis.

4.3.2.2 GO categorization of F-stat gene lists

The over-representation analysis using the software application EASE identified a number of common GO terms in the entire F-stat gene list (see Section 4.2.6.3(a) and Table 16). However, when these lists are redefined into three broader gene groups (i.e. differentiation-associated list, proliferation-associated list and sample effect list) these common groups are no longer apparent within these sub-groups, with the exception of the molecular function 'transferase activity' (see Section 4.2.6.3(b) and Figure 27). The lack of common groups may be a consequence of the reduced number of genes within each of these sub-groups, which on their own do not reach significance. Therefore, no key functional groups were identified that were unique to one sub-group and not another. One possible implication from this finding is that these sub-group lists (relating to proliferation and differentiation-associated genes) share many common functional pathways in order to achieve their different biological outcomes.

It was interesting that there were a number of similarities between the 'sample effect' and 'differentiation' lists. For example they both had 'cell growth and/or maintenance' as the number one group, but 'lipid metabolism' was only in the sample effect list. The molecular process 'transferase activity' was also in common to all the groups (even the proliferation group). Transferase activity as defined by GO (GO:0016740) is the "catalysis of the transfer of a group, e.g. methyl group, glycosyl group, acyl group, phosphorus-containing, or other groups, from one compound to another compound" (http://www.ebi.ac.uk/ego/). A subgroup of this term is transferase activity, transferring phosphorus-containing groups (GO:0016772). This could indicate that kinase and phosphatase activities are some of the key functions of the proteins contained in these groups, but the involvement of these types of proteins in cell signalling is well established. At this point in time it is unclear what these differences and similarities in the GO categories specifically reveal about the individual groups or genes that are contained within each group. However, taken as a whole, they indicate that both similar functioning genes and genes with different biological processes and molecular functions are responsible for the observed effects on the FDB-1 cells proliferation, survival and differentiation in the context of GM-CSF/IL-3/IL-5 receptor signalling.

The 'proliferation-associated' list represents those genes involved with V449E signalling (and perhaps leukaemic cell proliferation). The 'differentiation-associated'

gene list represents FIΔ signalling. However, the IL-3 withdrawal group is less easily defined. Both V449E and FIΔ hβc mutants would contribute to the IL-3 withdrawal group where the expression of a gene has either increased or decreased in both cell lines, but by differing amounts. This could be the result of the underlying mutant hβc signalling even in the presence of mIL-3, which is dominant over the differentiating effects of FIΔ signalling, but it is also possible that this group does not involve the mutant hβc at all. Instead, the IL-3 withdrawal group could represent the change in gene expression as a consequence of a loss of normal signalling from the mouse IL-3 receptor after these cells are grown without mIL-3. In mice there are two β subunits, the mouse βc which can be activated by GM-CSF/IL-3/IL-5 and the IL-3 specific β subunit (β_{IL-3}) that binds IL-3 with low affinity and only forms a high affinity receptor with IL-3 receptor alpha (IL-3Rα) (Gorman *et al.*, 1990; Hara and Miyajima, 1992; de Groot *et al.*, 1998; Guthridge *et al.*, 1998). In the presence of mIL-3, FDB-1 cell growth is likely to be due to signalling through both of these receptors. However, it is clear that the activated forms of hβc do not recapitulate the entire mIL-3 receptor signal in FDB-1 cells (e.g. FIΔ causes differentiation of FDB-1 cells). Thus, in the absence of mIL-3 these signalling differences will become apparent and the gene expression differences seen in the IL-3 withdrawal group may be the result. In addition, within the IL-3 withdrawal list may be genes involved in the survival of the FIΔ and V449E expressing cells, thus providing an insight into the signalling from the GM-CSF/IL-3/IL-5 receptor that contributes to cell survival.

Further experiments using microarrays examining the normal GM-CSF/IL-3/IL-5 receptor signalling in the parental FDB-1 cells stimulated with either mIL-3 or mGM-CSF over time would confirm the involvement of these genes in the survival and differentiation process of the FDB-1 cells. Recently, such a study has been reported by Brown *et al*., (2006) , which defined a set of novel genes identified in FDB-1 cells with the potential to regulate myeloid growth and differentiation (discussed in Section 6.3.1). The examination of the V449E and FIΔ mutants together with the wild type receptor in other haematopoietic cell lines would reveal the true or natural signals emanating from the mutant and wild type hβc receptor and distinguish those signals that are cell type specific. Such studies would provide a more complete understanding of the genetic events that occur downstream of the hβc receptor.

4.3.2.3 Selection of specific genes of interest

Using both the Top 200 ranked gene list and the gene lists from F-stat analysis, a list of genes of interest was compiled based on known function, literature searches, protein and DNA motifs, gene ontology (see Section 4.2.6.3), known association with other key proteins, possible role in cancer based on information obtained from the mouse Retroviral Tagged Cancer Gene Database (RTCGD) (http://rtcgd.ncifcrf.gov) or presence in other cancer databases (e.g. Cancer Gene Expression Database (CGED)). This group of genes was selected because they covered a variety of biological processes that could be postulated to be involved in aspects of myeloid differentiation (FIΔ expressing cells) or involved in the proliferation and leukaemic potential of the V449E expressing cells. Groups of genes were selected based on whether published data had shown an associated with proliferation, differentiation, or transcription (see Table 17A, B and C). The genes within those sections are discussed in the following paragraphs and summarised in Table 17.

Table 17: FDB-1 genes of interest. Genes of interest identified from the FDB-1 microarray study and subdivided into three classification groups – differentiation-associated (A), transcription-associated (B) and proliferation-associated (C). "Increase/decrease" refers to the relative increase/decrease in expression of the genes in the FIΔ cells from time 0 to time 24. Genes with proviral integration sites as identified in the Retroviral Tagged Cancer Gene Database (RTCGD) (Akagi *et al*., 2004) have been stated along with the number of integration sites found.

A

Class of gene identified	Gene name	Tumours (RTCGD)	increase/decrease associated with FlΔ induced signalling	Comment
Differentiation associated	Anxa1	3	increase	Also known as Lipocortin 1. Phosphorylated by EGFR and involved in its degradation (Futter et al., 1993; Radke et al., 2004).
	Anxa3	0	increase	Highly expressed in neutrohils, abundant in granules (Perron et al., 1997).
	Anxa9	0	increase	Does not bind acidic phospholipids in the presence of submillimolar Ca^{2+} concentrations, unlike the other family members (Goebeler et al., 2003).
	Btg1	1	increase	Anti-proliferative (Sasajima et al., 2002). Dysregulated in AML (Virtaneva et al., 2001).
	Calgranulin-b (S100a9)	0	increase	Over-expressed in AML M2 and M4 (Scoch et al., 2002)
	Cd24	1	increase	Marker of granulocyte differentiation (Elhetany and Patel, 2002)
	Csf3r	1	increase	Granulocyte colony stimulating factor receptor.
	Fcer1g	0	increase	Fc receptor, IgE, high affinity I, gamma polypeptide.
	Fcgr3	0	increase	Fc receptor, IgG, low affinity III.
	Hba-a1	0	decrease	Hemoglobin alpha, adult chain 1.
	Hbb-b1	0	decrease	Hemoglobin, beta adult major chain.
	Hbb-bh1	0	decrease	Hemoglobin Z, beta-like embryonic chain.
	Hbb-y	0	decrease	Hemoglobin Y, beta-like embryonic chain.
	Lactoferrin (lactotransferrin)	0	increase	Secondary granule protein, regulated by C/EBP (Khanna-Gupta et al., 2003) and KLF5 (Shi et al., 1999).
	Lcn2	0	increase	Component of neutrophil granules, ability to bind and transport a wide variety of small lipophilic substances (Kjeldsen et al., 2000).
	Mapk3	0	increase	Mitogen activated protein kinase 3. Role in signal transduction.
	Plaur	0	increase	Receptor for uPA. Expression up regulated by uPA, TGF-β, and KLF4 (Montuori et al., 2001; Lund et al., 1992; Wang et al., 2001). Marker of granulocytic differentiation (Elghetany et al., 2003). Role in tumour growth, cell invasion and metastasis (Wang et al., 2004; Montuori et al., 2001; Miyake et al., 1999; Mustjoki et al., 1999; Stephens et al., 1999; Pappot, 1999; Grondahl-Hansen et al., 1995; Memarzadeh et al., 2002; Dano et al., 2005) and is expressed on AML blast cells (Plesner et al., 1994; Lanza et al., 1998; Mustjoki et al., 1999).
	Prkcd	0	increase	Protein kinase C delta. Intracellular signalling cascade.
	Ypel3	0	increase	Ypel highly conserved gene family in eukaryotes. Subcellular localisation of all Ypel proteins was to centrosomes and mitotic appraratus (Hosono et al., 2004).
	Ypel5	0	increase	See description for Ypel3, but the subcellular localisation of Ypel5 was to mitotic spindle (slightly different to Ypel proteins 1-4) suggesting a distinct function (Hosono et al., 2004).

B

Class of gene identified	Gene name	Tumours (RTCGD)	increase/decrease associated with FlΔ induced signalling	Comment
Transcription associated	Cbx5	1	decrease	Chromobox homolog 5 (Drosophila HP1a).
	Dach1	0	increase	Related to Ski proto-oncogene (Hammon et al., 1998). Inhibits TGF-β signalling via repression of Smad signalling (Wu et al., 2003).
	Emx2	0	decrease	Maintenance of differentiation in endometrium (Noonan et al., 2001). Regulates neural stem cell proliferation (Cecchi, 2002).
	Gata2	0	decrease	Expressed in haematopietic stem and progenitor cells and controls early stages of haematopoiesis (Grass et al., 2003). Over-expression suppresses haematopoiesis (Persons et al., 1999; Heyworth et al., 1999; Briegel et al., 1993).
	Helios	0	decrease	Ikaros related (Rebollo and Schmitt 2003). Dysregulated in ALL (Sun et al., 2002; Tenkanashi et al., 2002)
	Hipk1	0	increase	Interacts with homeodomain proteins (Kim et al., 1998). Involved in relocalizing the transcriptional repressor Daxx from PODs (Ecsedy et al., 2003)
	Klf5	1	increase	Positive regulator of cellular proliferation (Sun et al., 2001). Interacts with retinoic acid receptor (Shindo et al., 2002). Known tumour suppressor activity (Chen et al., 2002 and 2003). Regulates the lactoferrin promoter (Shi et al., 1999).
	Klf6	0	increase	Also known as Copeb. Candidate tumour suppressor in prostate cancer (Narla et al., 2001).
	Klf7	1		Role in fate decision between proliferation and cell cycle arrest/differentiation (Bieker, 2001).
	Sfrs3	2	decrease	Regulation of cell-specific splicing (Jumaa et al., 1997). enriched in stem cell populations (Ramalho-Santos et al., 2002).
	Sp100	0	increase	Localised to PODs, interacts with PML (Seeler et al., 1998).

C

Class of gene identified	Gene name	Tumours (RTCGD)	increase/decrease associated with FlΔ induced signalling	Comment
Proliferation associated	Bnip3	0	decrease	Member of the Bcl-2 protein family, pro-apoptotic and induced by hypoxia and HIF1α (Regula et al., 2002; Crow, 2002).
	Cd34	0	decrease	Marker for immature haemopoietic progenitors.
	Lif	0	decrease	IL-6 related cytokine family member (Metcalf, 2003).
	c-Myc	65	decrease	Associated with proliferation, known oncogene.
	Nucleostemin	1	decrease	Nucleolar protein in stem cells (Tsai and McKay, 2002).
	Nucleophosmin	0	decrease	Nucleolar phosphoprotein more abundant in cancer cells (Chan et al., 1989) positive regulator of cell proliferation, and over-expression leads to malgnant transformation (We et al., 2002).
	Serpina3g (Spi2-1)	0	decrease	Serine protease inhibitor. Over expression delays myeloid differentiation in FDCP-Mix cells (Hampson et al., 1997).

4.3.2.4 Genes with a known association with myeloid differentiation

The 'associated with differentiation' list includes a number of genes that have been previously shown to increase in expression during myeloid differentiation. For example, *Lactoferrin* (*Lactotransferrin, Ltf*) encodes a secondary granule protein (Khanna-Gupta *et al.,* 2003), CD24 is a marker of granulocytic differentiation (Elghetany and Patel, 2002; Elghetany *et al.,* 2004) and Calgranulin-b (S100a9) is a neutrophil regulated protein involved in granulocytic differentiation (Cammenga *et al.,* 2003). Lipocalin 2 (Lcn2) has been demonstrated to be a component of specific granules in neutrophils via its co-localisation with lactoferrin (Kjeldsen *et al.,* 2000). Lipocalins are a diverse family of proteins that have the ability to bind and transport a wide variety of small lipophilic substances (Kjeldsen *et al.,* 2000). *Lipocalin 2* expression is regulated by glucocorticords, a fact explained by the presence of two glucocorticoid responsive core elements in the gene (Kjeldsen *et al.,* 2000). The 5' flanking region of *Lipocalin 2* gene also contains binding sites for C/EBP and PU.1 which are required for development of the myeloid lineages (Kjeldsen *et al.,* 2000).

Other differentiation-associated genes include *B-cell translocation gene 1* (*Btg1*), which is a member of the BTG anti-proliferative proteins (Rouault *et al.,* 1992; Rouault *et al.,* 1993; Sasajima *et al.,* 2002). Over-expression of *BTG1* has been shown to inhibit NIH3T3 (Rouault *et al.,* 1992), rat aortic smooth-muscle cell (RSMC) (Wilcox *et al.,* 1995) and myoblast proliferation and induce myoblast differentiation (Rodier *et al.,* 2001). BTG1 has been found to bind to Hoxb9 and enhance its transcriptional activity (Prevot *et al.,* 2000). Hoxb9 is a homeobox containing transcription factor which activates target genes that are critical for metazoan development (Prevot *et al.,* 2000). BTG1 has been shown to have an important role in erythroid differentiation, due to its regulation of protein arginine methyl transferase 1 (Kolbus *et al.,* 2003; Bakker *et al.,* 2004). Cho *et al.,* (2004) suggested that BTG1 may play a role in the differentiation of myeloid cells, which agrees with the findings for this gene in the microarray analysis presented in this chapter.

Three members of the Annexin family, *Anxa1*, *Anxa3* and *Anxa9* were also found to be differentiation-associated. The Ca^{2+} and phospholipid-binding Annexin protein family consists of cytosolic proteins, which are reversibly and Ca^{2+} dependently associated with cellular membranes. Annexins have been implicated in vesicular trafficking and membrane fusion events (Gerke and Moss, 2002). The N-terminal

domain of Annexin A1 contains phosphorylation sites for protein kinase C and several tyrosine kinases including the epidermal growth factor receptor (EGF) receptor (EGFR) (Futter *et al.*, 1993; Radke *et al.*, 2004). The phosphorylation of Annexin A1 is induced by growth factors and cytokines and may be responsible for mediating or regulating the mitogenic response (Radke *et al.*, 2004). Radke *et al.*, (2004) showed that EGFR co-localises with Annexin A1 at the plasma membrane and at vesicular structures during internalisation in HeLa cells. This co-localisation suggests that the inward vesiculation of EGFR and its degradation via the lysosomal pathway may be a consequence of EGFR mediated phosphorylation of Annexin A1 (Futter *et al.*, 1993; Radke *et al.*, 2004). Interestingly, Annexin A1 binds S100A11 (Annexin A2 binds S100A10 and Annexin A11 binds S100A6) (Gerke and Moss, 2002). Another S100 family member, S100A9 (Calgranulin B), was also identified as differentiation-associated in the microarray analysis and it could be speculated that it too may bind a member of the Annexin family. Annexin A1 has extracellular activity as a regulator of leukocyte extravasation via the finding of Annexin A1 specific receptors on neutrophils and monocytes that are members of the formyl peptide receptor (FPR) family of chemoattractant receptors (Rescher and Gerke, 2004). An anti-inflammatory action for Annexin A1 has been suggested and supported by experiments in knock-out mice revealing it to be an important regulator of leukocyte migration (Rescher and Gerke, 2004). Annexins A1 and A3 have been linked to exocytosis, specifically the post-*trans*-Golgi network, and stimulate Ca^{2+}-dependent secretion in neutrophils (Gerke and Moss, 2002). Annexin A3 is expressed in differentiated cells of the myeloid cell lineage and is found highly expressed in neutrophils where it is an abundant phospholipid in the neutrophil specific granules (Perron *et al.*, 1997). Annexin A9 is a unique member of this family that does not bind acidic phospholipids in the presence of submillimolar Ca^{2+} concentrations, unlike the other family members (Goebeler *et al.*, 2003). This suggests that Annexin A9 is not regulated by intracellular Ca^{2+}, but it remains to be determined whether this protein is capable of interacting with certain types of cellular membranes via a Ca^{2+} independent mechanism (Goebeler *et al.*, 2003). *Annexin A9* mRNA expression is observed very early in mouse embryonic development and ceases by day 19 post conception (Markoff *et al.*, 2002). Interestingly, the *Annexin A9* promoter contains sites for the transcription activators GATA1 and PU.1 as well as KLF1 and KLF3 (Chlystun *et al.*, 2004). Over-expression of PU.1 blocks erythroid

differentiation, whereas enforced expression of GATA1 inhibits differentiation of myeloid cells (Rekhtman *et al.*, 2003).

Two genes, *yippee-like 3* (*Ypel3*) and *yippee-like 5* (*Ypel5*), from a novel gene family called YPEL were also found in the differentiation-associated list. *Drosophila yippee* was the first member of YPEL family identified, based on its interaction with hemolin, an immunoglobulin superfamily member from a species of moth (Roxstrom-Lindquist and Faye, 2001). Since then, five family members YPEL1-5 and Ypel1-5 have been identified in human and mouse respectively (Hosono *et al.*, 2004). The YPEL family proteins have been found in essentially all eukaryotes and were thus reasoned to have a fundamental function (Hosono *et al.*, 2004). *Ypel3* and *Ypel5* were found to be ubiquitously expressed in a multi-tissue cDNA panel. Indirect immunofluorescent staining of COS-7 cells revealed that these proteins are localised to centrosome and nucleolus during interphase but have different staining around or at the mitotic spindle during mitosis (Hosono *et al.*, 2004). It has been suggested that based on their structure YPEL proteins are capable of binding zinc, or another metal ion, while their protein localisation suggests a role involved in cell division (Roxstrom-Lindquist and Faye, 2001; Hosono *et al.*, 2004). The gene expression profiles of *Ypel3* and *Ypel5* in the microarray study reported in this chapter suggests a role for these genes in the differentiation of FIΔ expressing FDB-1 cells, and perhaps in the differentiation of myeloid cells.

Four haemoglobin genes (*Hba-a1*, *Hbb-b1*, *Hbb-bh1* and *Hbb-γ*) were also differentially expressed and were associated with a decrease in FIΔ signalling. This would be expected as the differentiating FIΔ expressing FDB-1 cells do not form mature erythrocytes and thus would not be expected to express these genes. The expression of these genes in the proliferating IL-3 stimulated FDB-1 cells could be indicative of a multi-lineage priming event within the FDB-1 cells (even though they are not known to be capable of differentiating to erythrocytes) similar to FDCP-mix cells investigated by Bruno *et al*., (2004). Multi-lineage priming is also shown in the results of Miyamoto *et al*., (2002) who demonstrated that the earliest myeloid progenitors (common myeloid progenitors, CMPs) coexpress both granulocyte/monocyte-affiliated and megakaryocyte/erythrocyte-affiliated genes. Thus, IL-3 stimulated FDB-1 cells may express markers for multiple cell lineages until they commit to a specific differentiation pathway. Furthermore, expression of V449E (but not FIΔ or I374N) in primary

haematopoietic cells leads to the production of factor independent erythroid cells (McCormack and Gonda, 1997; McCormack, 1998). Thus, expression of haemoglobin genes may be specifically driven by V449E mediated signalling.

It was also interesting to note that, according to the array data, some genes (e.g. *Ltf* and *S100a9*) associated with the differentiation phenotype are already significantly different at time zero (i.e. in IL-3) between FIΔ and V449E and increase in expression following IL-3 withdrawal (Table 13). This may suggest that in the presence of IL-3, FIΔ signalling results in the activation of several genes associated with differentiation. However, the expression of these genes is clearly not sufficient to cause differentiation in the presence of the dominant IL-3 signal. FIΔ induced differentiation may require further gene expression changes that are associated with IL-3 withdrawal.

4.3.2.5 Genes with a known association with transcription

GATA binding protein 2 (Gata2) was found to be over-expressed in the proliferation-associated list (Table 15D) and is another example of a lineage-affiliated transcription factor along with the PU.1 and C/EBP family. GATA2 is expressed in haematopoietic stem and progenitor cells as well as endothelial cells and in a diverse range of tissues (Grass *et al.,* 2003). GATA2 controls early stages of haematopoiesis, whereas GATA1 regulates terminal differentiation and function of erythroid and megakaryocytic cells and early stages of eosinophil differentiation (Grass *et al.*, 2003). In human myeloid KG1 cells, all-trans-retinoic acid (RA) receptor alpha (RARα) associates with GATA2 and makes the transcriptional activity of GATA2 RA responsive (Tsuzuki *et al.,* 2004). Over-expression of GATA2 in murine primitive haematopoietic cells suppresses haematopoiesis (Briegel *et al.*, 1993; Heyworth *et al.*, 1999; Persons *et al.*, 1999) and when expressed in ES cells it increases primitive haematopoietic colony formation (Kitajima *et al.,* 2002). This reveals a common theme whereby enhanced GATA2 expression in progenitors affects differentiation by promoting self renewal over differentiation (Grass *et al.,* 2003). This would be in agreement with the increased *Gata2* mRNA expression levels seen in the proliferating V449E cells over the differentiating FIΔ cells.

Other genes in the transcription factor category are *Helios, Emx-2* (Table 15C) and *Dach1* (Table 15A) which could have (novel) roles in either promoting proliferation and self-renewal or blocking differentiation. *Empty spiracles homolog 2* (*Emx2*) is a homeobox-containing gene, a homologue of the *Drosophila empty spiracles* gene (*ems*),

that codes for a transcription factor involved in mouse cerebral cortex development (Noonan *et al.*, 2001; Cecchi, 2002). Data from Noonan *et al*., (2001) support a role of Emx2 functioning as a negative regulator of cellular proliferation and as a tumour suppressor in the uterine endometrium. Emx2 is involved in the regionalization and specification of cortical areas as well as having functions involved in neuroblast proliferation, migration and differentiation (Cecchi, 2002). In mice, w*ingless-related MMTV integration site 1* (*Wnt1*) gene has been identified as an in-vitro downstream target gene of *Emx2*, which negatively controls its transcription (Iler *et al.*, 1995; Cecchi, 2002).

Helios is a member of the Ikaros gene family that belongs to the Krupel transcription factor family (Rebollo and Schmitt, 2003). The Ikaros gene family is a primary regulatory of lymphocyte differentiation (Takanashi *et al.,* 2002). Helios can dimerise with itself or heterodimerise with the other Ikaros family members, Aiolos and Ikaros. Helios is expressed in the earliest haematopoietic sites of the embryo, in adult haematopoietic stem cells and, with decreasing expression, in maturing erythroid, macrophage, T and B lineage cells (Rebollo and Schmitt, 2003). Helios has been implicated in the development of acute lymphoblastic leukaemia, where its expression is dysregulated (Sun *et al.*, 2002; Takanashi *et al.*, 2002).

Dachshund 1 (*Dach1*) is a mammalian homologue of the *Drosophila dachshund* (*dac*) gene. Dach1 is a nuclear factor that is expressed during development in restricted areas of the central nervous system, neural crest and limb buds (Backman *et al.,* 2003). Targeted disruption of *Dach1* results in postnatal lethality associated with failure to suckle, cyanosis and respiratory distress. However, growth kinetics of the neural stem cells, histological examination, blood chemistry and mRNA expression studies all failed to reveal a cellular basis for this lethality (Davis *et al.,* 2001; Backman *et al.,* 2003). *Dac* and its mammalian homologues were found to be related to the Ski proto-oncogene, which functions during myogenesis and neurulation, and Sno, a Ski-related protein (Hammond *et al.,* 1998). Interaction of Ski and Sno with Smad3 represses transcription of transforming growth factor-β (TGF-β)-responsive genes and recently Dach1 has been shown to inhibit TGF-β signalling through binding of Smad4 and repression of Smad signalling (Wu *et al.,* 2003). Interestingly, Ski functions as a co-repressor with c-Myb (Nomura *et al.,* 2004) and it would be interesting to investigate whether Dach1 would also negatively regulate c-Myb activity.

Several genes displaying differential expression in the microarray analysis are known targets of the granulocytic lineage transcription factor, C/EBPα; e.g. *Calgranulin-b* (S100A9) (Cammenga *et al.,* 2003), *KLF5* (Cammenga *et al.,* 2003), *ANXA1* (Iwama *et al.*, 1998; Johansen *et al.*, 2001; Cammenga *et al.*, 2003), *LTF* (Iwama *et al.*, 1998; Khanna-Gupta *et al.*, 2000; Khanna-Gupta *et al.*, 2003), c-*Myc* (Johansen *et al.*, 2001; Cammenga *et al.*, 2003) and *CSF3R* (G-CSF receptor) (Smith *et al.*, 1996; Zhang *et al.*, 1998). Kruppel-like factor 5 (*KLF5*) is particularly interesting because not only is it a transcription factor that interacts with the retinoic acid receptor (Shindo *et al.,* 2002), but it also has known tumour suppressor activity (Chen *et al.*, 2002; Chen *et al.*, 2003) and has been reported to regulate the *Ltf* promoter (Shi *et al.,* 1999). The regulation of the *Ltf* promoter by *Klf5* may explain why expression of *Ltf* is induced in the FIΔ expressing FDB-1 cells (greater than 10 fold F/V at time 24 hours – see Table 13). KLF5 is a downstream target of Wnt1 signalling and its induction occurs via a β-catenin/T-cell independent mechanism that may involve activation of protein kinase C (another differentially expressed gene in the Top 200 list – associated with differentiation (Table 12A) (Bieker, 2001). *KLF5* has also been shown to be a direct target of oncogenic H-Ras and that it acts downstream of MAPK, as determined by the use of MEK inhibitors (Nandan *et al.,* 2004). KLF5 is also a positive regulator of cellular proliferation (Sun *et al.,* 2001) and cyclin D1 has been shown to be a target of KLF5, which may mediate KLF5's effect on cell proliferation (Nandan *et al.,* 2004).

Other Kruppel-like zinc finger transcription factor family members that were also found to be differentially expressed were KLF6 and KLF7. Aliases of KLF6 include Copeb, Bcd, Zf9 and CPBP. Known KLF6 target genes include a placental glycoprotein, collagen α1(I), the collagen-specific molecular chaperone HSP47, leukotriene C(4) synthase, keratin 4 and 12, transforming growth factor β1 (*TGF*β1), types I and II *TGFβ* receptors, urokinase type plasminogen activator (*uPA*), p21, endoglin, nitro-oxide synthase, insulin-like growth factor 1 receptor and the human immunodeficiency virus long terminal repeat (*HIV-1 LTR*) (Koritschoner *et al.*, 1997; Kim *et al.*, 1998a; Ratziu *et al.*, 1998; Suzuki *et al.*, 1998; Kojima *et al.*, 2000; Okano *et al.*, 2000; Zhao *et al.*, 2000; Narla *et al.*, 2001; Botella *et al.*, 2002; Chiambaretta *et al.*, 2002; Wang *et al.*, 2002; Yasuda *et al.*, 2002; Warke *et al.*, 2003; Rubinstein *et al.*, 2004). Narla *et al.*, (2001) identified *KLF6* as a candidate tumour suppressor gene in prostate cancer and based on its ubiquitous expression and ability to suppress growth,

suggested that it may play a more general role in the development or progression of human cancers, particularly those associated with LOH at chromosome 10p15. KLF7 has been suggested to play a critical role in the decision between proliferation and cell cycle arrest/differentiation, particularly within the developing and adult nervous system (Bieker, 2001).

The receptor for uPA is the urokinase-type plasminogen activator receptor, uPAR, (gene name *Plaur*), which was found to be differentially expressed in the microarray analysis (differentiation-associated). uPAR (CD87) has been shown to act as a marker for terminal granulocytic differentiation but is not expressed on $CD34^+$ progenitor cells (Elghetany *et al.,* 2003). It is interesting to note that uPA and phosphoglycerate kinase can up-regulate the expression of uPAR receptor post-transciptionally (Montuori *et al.*, 2001; Shetty *et al.*, 2004). TGF-β strongly up-regulates uPAR expression (Lund *et al.,* 1991) and the transcription factor KLF4 also regulates uPAR in epithelial cells (Wang *et al.*, 2004). Hence not only is *KLF6* induced upon differentiation in the microarray analysis, but two of its target genes (*TGFβ1* and *uPA*) up-regulate *uPAR* that was also found to be differentially expressed.

4.3.2.6 Genes with a known association with proliferation

The "proliferation-associated" list includes genes such as *Cd34*, a marker of immature haematopoietic stem cells with unknown function, as well as *c-Myc* a proto-oncogene. *c-Myc* is expressed in proliferating cells and is down-regulated in terminally differentiated cells. The down-regulation of *c-Myc* expression is a critical event for cell commitment to a differentiation pathway, particularly in myeloid cells (Johansen *et al.*, 2001). The gene encoding *Leukemia inhibitory factor* (*Lif*), a pleitropic cytokine that belongs to the IL-6 related cytokine family, was found to be differentially expressed and down-regulated upon FIΔ differentiation (see Table 15D). Lif is able to enhance stem cell proliferation and stimulate proliferation of factor dependent haematopoietic cell lines (Metcalf, 2003), which may explain its increased expression in the proliferating V499E cells. Another gene of note is *Spi2-1* (*Serpina3g*), a serine protease inhibitor, which decreased 3.9-4.3 fold (see Table 13 and Table 15C) upon FIΔ differentiation and has also been reported in other myeloid differentiation models (Hampson *et al.*, 1997; Bruno *et al.*, 2004). *Spi2-1* was shown by Hampson *et al.*, (1997) to be expressed in undifferentiated FDCP-mix cells, a multipotent haematopoietic cell line, but showed a marked down-regulation upon differentiation in this cell line and in bone marrow

derived bipotent granulocyte-macrophage progenitor cells. Over-expression of Spi2-1 protein caused a delay in maturation and an increase in the clonogenic potential of the FDCP-mix cells (Hampson *et al.*, 1997). It is thus likely that *Spi2-1* has a role in myeloid cell development that is also recapitulated in FDB-1 cell line via interaction with non-protease intracellular targets (Morris *et al.*, 2003).

4.3.2.7 Genes with a known association with cancer

Several of the genes predicted to be associated with proliferation or blocked differentiation are represented as proviral integration sites in mouse haematopoietic tumours generated in screens for oncogenic activation (Retroviral Tagged Cancer Gene Database, RTCGD) (Suzuki *et al.*, 2002; Akagi *et al.*, 2004). This information has been included in Table 17 and suggests a potential role of such genes in oncogenesis. In addition, several of the genes found to be differentially expressed in the microarray study presented in this chapter are over-expressed in various cancers or suggested to be tumour suppressors. It is possible that one or more of these differentially expressed genes may be partially responsible for the leukaemogenic properties of the V449E mutant observed *in vivo* (McCormack and Gonda, 1999). For example, the *BTG1* locus was discovered by mapping a *c-Myc* chromosomal translocation [t(8:12) (q24, q22)] in a B-cell chronic lymphocytic leukaemia (Rimokh *et al.*, 1991) and *BTG1* has been found to be a potential tumour suppressor gene (Rouault *et al.*, 1992). *c-Myc* is frequently over-expressed in human neoplasia (Marcu *et al.*, 1992), induces tumours in mice and can lead to the formation of myeloid leukaemias (Adams *et al.*, 1985; Felsher and Bishop, 1999). *uPAR* expression has been implicated in tumour growth and progression (Wang *et al.*, 2004) as it is expressed in several tumours and plays a fundamental role in tumour cell invasion and metastasis (Grondahl-Hansen *et al.*, 1995; Miyake *et al.*, 1999; Mustjoki *et al.*, 1999; Pappot, 1999; Stephens *et al.*, 1999; Montuori *et al.*, 2001; Memarzadeh *et al.*, 2002; Dano *et al.*, 2005). *uPAR* has also been found to be expressed on blasts from patients with acute myeloid leukaemia, particularly monocytic leukaemia, implicating the involvement of this gene in myeloid leukaemia (Plesner *et al.*, 1994; Lanza *et al.*, 1998; Mustjoki *et al.*, 1999). *Calgranulin-b* (S100a9) has also been found to be over-expressed in certain subtypes of AML (Golub *et al.*, 1999; Schoch *et al.*, 2002) but this gene's expression may be more indicative of a role as a marker gene rather than having a causative role in leukaemia. Similarly, *Lipocalin 2* has also been identified as a candidate marker gene for pancreatic

tumorigenesis and metastasis (Missiaglia *et al.*, 2004) and for ovarian cancer diagnosis and thereapy (Santin *et al.*, 2004).

Acute promyelocytic leukaemia (APL) is a common form of AML characterised by an accumulation of myeloid precursors blocked at the promyelocytic stage of maturation. In APL, four types of chromosomal translocations have been described that express a fusion protein involving the RARα coding region and another nuclear protein that contains a homodimerization domain (Lin *et al.*, 1999). One of these proteins is nucleophosmin (NPM), a proto-oncogene, which is down-regulated upon FIΔ differentiation (see Table 13 and Table 17). NPM is a nucleolar phosphoprotein that is more abundant in cancer cells than in normal resting cells (Kondo *et al.*, 1997; Wu and Yung, 2002) and over-expression of NPM in NIH3T3 cells leads to malignant transformation (Kondo *et al.*, 1997). *NPM* is an immediate-early gene induced by UV-damaged DNA (Wu and Yung, 2002). Its expression is increased in associated with commitment for mitogenesis and plays a potential role as a positive regulator of cell proliferation (Feuerstein and Mond, 1987; Feuerstein *et al.*, 1988). NPM is a target of CDK2-cyclin E and its phosphorylation by CDK2-cyclin E is involved in the initiation of centrosome duplication (Okuda *et al.*, 2000; Tokuyama *et al.*, 2001). These last findings fit well with its increased expression in the proliferating V449E cells (as compared to the differentiating FIΔ cells).

The order of the genes *ALK-YPEL5-MST014* on human chromosome 2 and mouse chromosome 17 are identical (Hosono *et al.*, 2004). *Ypel5* was differentially expressed in the microarray, but has not been associated with cancer; however the adjacent gene, anaplastic lymphoma kinase (*ALK),* is involved in a specific translocation in lymphoma. *ALK* encodes a neural tyrosine kinase that is involved in the development and function of the nervous system (Morris *et al.*, 1997). NPM occurs as a fusion protein with ALK in non-Hodgkin's lymphomas (Morris *et al.*, 1994). Another chromosomal translocation involving *NPM* is the t(3;5)(q25.1;q34) of myelodysplastic syndrome and AML, which produces the fusion gene, *NPM-MLF1* (Yoneda-Kato *et al.*, 1996). These findings involving NPM translocation clearly suggest a role for NPM in the development of lymphomas and leukaemia.

Genes associated with PML oncogenic domains.

The genes *Sp100*, *Cbx5* and *Hipk1*, found to be differentially expressed in the microarray, have been discovered to be associated with PML oncogenic domains or

PODs. *Sp100* and *Hipk1* were both found to be up-regulated with differentiation of FIΔ cells (see Table 13 and Table 17). The PML-SP100 nuclear bodies, also referred to as PODs, were first described as discrete subnuclear structures containing the SP100 autoantigen (Seeler and Dejean, 1999). Subsequently they were shown to contain the promyelocytic leukaemia protein (PML) that is one part of the leukaemogenic PML-RARa fusion protein associated with APL (for review see Lin *et al*., 1999). Cell division causes the disaggregation of PODs and this structure is also disrupted in APL, but is restored by retinoic acid which is used in the therapeutic treatment of APL (Seeler and Dejean, 1999). Interferons, inducers of cellular antiproliferative and anti-viral responses, cause the up-regulation of PML and SP100 (Guldner *et al.*, 1992; Lavau *et al.*, 1995; Stadler *et al.*, 1995).

Seeler *et al*., (2001) suggest a link between POD components and transcriptional control at the chromatin level. They show that SP100, like PML, is posttranslationally modified by SUMO and this modification enhances the stability of the interaction of SP100 with HP1α, a member of the heterochromatin protein 1 (HP1) family of non-histone chromosomal proteins previously identified to bind to SP100 (Lehming *et al.*, 1998; Seeler *et al.*, 1998). Interestingly *Cbx5* (or HP1α as it is also known) was found to be differentially expressed between V499E and FIΔ at 24 hours and showed selective repression during differentiation of FIΔ (see Table 13 and Table 17B). Both SP100 and HP1α proteins act as transcriptional repressors when bound to DNA (Seeler and Dejean, 1999). Lehming *et al*., (1998) showed that SP100 also binds high-mobility group 1/2 (HMG1/2) type proteins. Previously, Seeler *et al*., (1998) identified a novel splice variant, SP100-HMG that contains a high mobility group (HMG1/2) protein domain that may posses DNA-binding activities and/or act as a transcription factor.

Hipk1 is a member of the HIPK family nuclear protein kinases that act as co-repressors of homeodomain transcription factors (Kim *et al.*, 1998b). It is ubiquitously expressed in mice and humans and localises predominantly to nuclear speckles (the majority of which were distinct from PODs) with diffuse staining through-out the nucleus and to a lesser extent the cytoplasm (Ecsedy *et al.*, 2003). Hipk1 has been shown to interact with Daxx and phosphorylates Daxx on Ser 669, which is important for modulating Daxx's transcription-repressive activity (Ecsedy *et al.*, 2003). Daxx, a nuclear protein, interacts with PML and localises to PODs, but the interaction of PML with Daxx depends on the SUMO modification of PML (Li *et al.*, 2000). Hipk1

relocalizes Daxx from PODs disrupting its interaction with PML and augmenting its association with histone deacetylase (HDAC1) (Ecsedy *et al.*, 2003). Given that PODs and the proteins that are associated with them are involved in the regulation of transcription and that some of these genes were found to be differentially expressed in this study, it therefore warrants further investigation to gain a more complete understanding of the role these proteins play in cell growth, differentiation and neoplasia.

4.3.3 Summary

In this chapter, microarray analysis was used to identify genes that were selectively expressed during differentiation of FIΔ expressing FDB-1 cells compared to proliferating V449E expressing FDB-1 cells. Among these selectively expressed genes were *Btg1*, *Calgranulin B* (*s100a9*), *Cd24* and *Ltf* found to be differentiation-associated; *Hipk1*, *Klf6*, *Sp100*, *Sfrs3* found to be transcription-associated; and *Bnip3*, *Cd34*, *Myc*, *Nucleophosmin* and *Nucleostemin* found to be proliferation-associated (Table 17). In order to confirm differential expression of these 13 genes it was decided to examine specific mRNA levels at different times during the differentiation process (see Chapter 5). This was performed in order to confirm a kinetic relationship between the regulation of these genes and the differentiation or proliferation of FDB-1 cells.

Chapter 5

Confirmation of Microarray Data by Expression Analysis and Functional Studies in Myeloid Cells

5.1 Introduction

5.1.1 Aims

The work described in this chapter had two aims. The first was to confirm differences in gene expression identified from the microarray experiment (see Chapter 4) using another method of measuring gene expression, Northern blot analysis. The second aim was to over-express a subset of these genes in FDB-1 cells in order to investigate the role of these genes in FDB-1 cell proliferation and differentiation. To address the first aim, gene expression differences between FDB-1 cells expressing the extracellular mutant (FIΔ) and the transmembrane mutant (V449E) were assessed after IL-3 withdrawal. This was performed at time zero (0), where cells were grown in media containing IL-3, and 24 hours after growth of the cells in factor-free media. Northern blots were performed for 13 genes identified in Table 17 (see Chapter 4). Six of the 13 genes were investigated further by examining a time course of FDB-1 mRNA expression over three days of mutant hβc signalling in the FDB-1 cells. To investigate whether changes in gene expression levels seen at the RNA level also extended to the protein level, protein levels were compared for one of the selected gene-products. To investigate the second aim, retroviral vectors were used to enable over-expression following infection of FDB-1 cells. The goal from the over-expression studies was to gain insights into the possible function of these genes in FDB-1 cells and enable prioritisation of genes for further functional analysis primary cell systems.

5.1.2 Selection of specific genes

Using both the Top 200 ranked genes and the gene lists generated from the F-stat analysis, a list of specific genes of interest were selected based on whether published data had shown an association with proliferation, differentiation, transcription or cancer (see Chapter 4, Section 4.3.2.3 and Table 17). This group of genes spanned a variety of biological processes that could be involved in aspects of myeloid differentiation (FIΔ

expressing cells) or involved in the proliferation and/or leukaemic potential of the V449E cells. The expression levels of thirteen genes selected from the genes of interest list were investigated in Section 5.2.1, below.

Following the analysis of gene expression for 13 genes at time 0 and 24 hours, a subset of six genes was selected for a more detailed analysis over a 72 hour period following factor independent growth of FIΔ or V449E expressing FDB-1 cells (see Section 5.2.1.2.2). Four genes (*Btg1*, *Cd24*, *Klf6* and *Nucleostemin* (also known as *Gnl3*, but is referred to as *Nucleostemin* throughout this thesis)) were then selected based on their confirmed differential expression profile in the Northern blots, and their known function(s) identified from the literature (summarised below), for functional analysis in FDB-1 cells (see Section 5.2.2). Of the four genes chosen for the over-expression studies, *Btg1*, *Cd24* and *Klf6* showed increased expression in FIΔ cells over time in the microarray data, which suggested a role in differentiation.

5.1.2.1 Genes up-regulated with differentiation

Btg1

BTG1's anti-proliferative role, potential tumour suppressor activity and its involvement as a translocation product in B-cell chronic lymphocytic leukaemia (B-CLL) (Rimokh *et al.*, 1991; Rouault *et al.*, 1992; Matsuda *et al.*, 2001) justified an examination of the role of BTG1 in FDB-1 cell differentiation. Over-expression of BTG1 has also been shown to inhibit proliferation of NIH 3T3 (Rouault *et al.*, 1993), rat aortic smooth-muscle cell (RSMC) (Wilcox *et al.*, 1995) and induce myoblast differentiation (Rodier *et al.*, 2001). While its anti-proliferative role is interesting in terms of the study presented in this thesis, the spatial and temporal expression of BTG1 has suggested an involvement in differentiation processes such as embryogenesis, spermatogenesis and angiogenesis (Raburn *et al.*, 1995; Sakaguchi *et al.*, 2001; Iwai *et al.*, 2004). BTG1 expression has also been linked to erythroid progenitor differentiation (Kolbus *et al.*, 2003; Bakker *et al.*, 2004) and also to myeloid differentiation (Cho *et al.*, 2004). When over-expressed in murine bone marrow cells, BTG1 expression was shown to inhibit the proliferation of erythroid progenitor cells – decreasing the size and number of erythroid colonies that formed in colony assays (Bakker *et al.*, 2004). A study by Cho *et al*., (2004) showed that *BTG1* mRNA was over-expressed in the complete remission state of all-trans-retinoic acid (ATRA)-treated AML-M3 and ATRA-treated HL-60 cells. They found that expression of BTG1 protein was up-regulated in bone marrow mononuclear cells

from AML-M2 and M3 patients in complete remission and in normal bone marrow mononuclear cells, but not in the non-remission state of AML-M2 patients. Cho *et al*., (2004) concluded that BTG1 may play a role in the differentiation of myeloid cells and as a potential treatment related biomarker for monitoring AML-M2 and M3 patients after therapy. Furthermore, BTG1 has been identified as a target gene of the forkhead transcription factor (FoxO3a), and a PI3 kinase target (Kolbus *et al.*, 2003; Bakker *et al.*, 2004).

BTG1 acts as an activator of protein-arginine *N*-methyltransferase 1 (PRMT1) via its binding to the BoxC region of BTG1 that is conserved in BTG1 and BTG2 (Lin *et al.*, 1996; Berthet *et al.*, 2002). This interaction with PRMT1 suggests that BTG1 may direct the associated methyl transferase activity towards substrates that bind to protein interaction domains contained within BTG1 (e.g. BoxA and B conserved in all six family members) (Matsuda *et al.*, 2001; Bakker *et al.*, 2004). BTG1 also interacts with the carbon catabolite repressor (CCR4)-associated factor 1, CAF1 and the CCR4-CAF1 complex is involved with heterochromatin formation, gene silencing and the negative regulation of mRNA stability (Bogdan *et al.*, 1998; Tucker *et al.*, 2001). Together the interaction and activation of PRMT1 and CAF1 by BTG1 could contribute to epigenetic gene regulation via the generation of a chromatic structure that facilitates promoter accessibility or the displacement of negative regulatory transcription factors (Prevot *et al.*, 2000; Bakker *et al.*, 2004).

Klf6

The Kruppel-like transcription factor, *Klf6* (also known as *COPEB, CPBP, Zf9* and *Bcd*) was isolated as a proto-oncogene from the peripheral blood of a B-CLL patient and has been shown to be involved in B-cell growth and development (El Rouby and Newcomb, 1996; Gehrau *et al.*, 2005). The *KLF6* gene has been shown to be mutated in a subset of human prostate cancers, where it may act as a tumour suppressor gene (Narla *et al.*, 2001; Narla *et al.*, 2005a; Narla *et al.*, 2005b). Consistent with this, Narla *et al*., (2001) showed that up-regulation of the *KLF6* gene suppresses cell growth in a p53-independent manner. This growth suppression was due to the up-regulated expression of p21 (WAF1/CIP1), an inhibitor of several cyclin-dependent kinases (cdk) and a key regulator of the G_1/S transition (Sherr and Roberts, 1999), which was revealed to be a direct target of KLF6. Benzeno *et al*., (2004) have suggested that KLF6 converges with the retinoblastoma protein (Rb) pathway resulting in KLF6's effects on growth

suppression. KLF6 mediates growth inhibition by reducing the phosphorylation of Rb through disruption of cyclin D1/cdk4 complexes and forcing the redistribution of p21 onto cdk2, which promotes G_1 cell cycle arrest (Benzeno *et al.*, 2004). KLF6 has also been suggested to be a tumour suppressor gene product in non-small cell lung cancers (NSCLC) where it was shown to be down-regulated in NSCLC and to suppresses tumour growth by induction of apoptosis (Ito *et al.*, 2004). Further support for KLF6 as a potential tumour suppressor gene product and as an inhibitor of cell proliferation was revealed by its functional antagonism of the c-Jun proto-oncoprotein, whereby KLF6 was shown to enhance the degradation of c-Jun protein by the proteasome-dependent pathway (Slavin *et al.*, 2004).

Cd24

Published studies describe a role for *Cd24* gene expression in FDB-1 cell differentiation. CD24 protein is a ligand of an adhesion receptor, P-selectin, expressed on activated endothelial cells and platelets (Sammar *et al.*, 1994; Aigner *et al.*, 1995; Aigner *et al.*, 1997). It is a glycosylphosphatidylinositol (GPI)-anchored cell surface protein that is heavily glycosylated and has been previously implicated in T- and B-cell differentiation (Pierres *et al.*, 1987; Alterman *et al.*, 1990; Kay *et al.*, 1990; Kay *et al.*, 1991). CD24 protein is expressed on most immature haematopoietic lineages including myeloid precursors, as well as developing neurons, regenerating muscle, dendritic cells and epithelial cells (Nielsen *et al.*, 1997 and references therein). CD24 protein expression is regulated during cellular maturation of most haematopoietic lineages but, with the exception of erythrocytes, its expression is absent in terminally differentiated lymphocytes and most myeloid cells (Nielsen *et al.*, 1997). CD24 protein is a known marker of granulocytic differentiation (Elghetany and Patel, 2002; Elghetany *et al.*, 2004) and hence its up-regulation at the mRNA level is consistent with granulocytic differentiation of FIΔ expressing FDB-1 cells. CD24 is also involved in apoptosis as demonstrated by Jung *et al.*, (2004) who showed that cross-linking of CD24 with mT-20 monoclonal antibody induced apoptosis of immature thymocytes in a T-cell receptor (TCR)- and caspase-independent manner. Expression of CD24 protein is predictive of monocytic AML (FAB subtype M4 or M5) (Raife *et al.*, 1994) and it is also expressed in B-cell lineage leukaemias including ALL, CLL and non-Hodgkin's B-cell leukaemia (Perri *et al.*, 1983). CD24 has been found to be expressed on solid tumours including small cell lung carcinoma, neuroblastoma, rhabdomyosarcoma, and renal cell carcinoma

(Droz *et al.*, 1990; Akashi *et al.*, 1994). CD24 protein expression is not only important in haematological malignancy but it is also an important prognostic tissue marker in prostate cancer (Kristiansen *et al.*, 2004), ovarian cancer (Kristiansen *et al.*, 2002), breast cancer (Kristiansen *et al.*, 2003b), colorectal cancer (Weichert *et al.*, 2005) and non-small cell lung cancer (Kristiansen *et al.*, 2003a) whereby increased expression was linked to earlier disease progression and shortened patient survival.

5.1.2.2 Down-regulated with differentiation

Nucleostemin

Nucleostemin (*Gnl3*) displayed an increased gene expression in V449E cells over time (decreased in FIΔ) in the microarray data suggesting a possible role in the sustained proliferation of the V449E FDB-1 cells in the absence of mIL-3. *Nucleostemin*'s reported involvement in stem cell maintenance and expression in several cancer cell lines and malignant tumour tissues made it an exciting candidate to investigate for its possible role in V449E mediated proliferation of FDB-1 cells (Tsai and McKay, 2002; Liu *et al.*, 2004; Sijin *et al.*, 2004; Xu *et al.*, 2004; Han *et al.*, 2005; Yaghoobi *et al.*, 2005; Yang *et al.*, 2005). Nucleostemin protein was identified by Tsai and McKay (2002) to be expressed in central nervous system stem cells, embryonic stem cells and several human cancer cell lines. Hence, it was suggested that Nucleostemin regulates the self-renewal of stem cells (Tsai and McKay, 2002; Yaghoobi *et al.*, 2005). Yang *et al*., (2005) have recently identified the alpha isoform of human protein phosphatase 2 regulatory subunit B (B56) (PPP2R5A) as a protein that interacts with Nucleostemin. PPP2R5A plays an important role in cell cycle progression and the negative control of cell growth and division by dephosphorylating and inactivating protein kinases that regulate critical signalling cascades (Millward *et al.*, 1999). The interaction of PPP2RSA with Nucleostemin implicates the involvement of Nucleostemin in these processes. Furthermore, Nucleostemin's expression in cancer cell lines and tumour tissues implies an important role in tumourigenesis and the growth regulation of cancers (Tsai and McKay, 2002; Liu *et al.*, 2004; Sijin *et al.*, 2004; Xu *et al.*, 2004; Han *et al.*, 2005).

5.1.3 Functional studies: retroviral over-expression in target cell lines

A common strategy used to investigate the function of a gene of interest in a given cell type is to over-express that gene using a powerful viral promoter. As murine and human

haematopoietic cells and cell lines can be difficult to transfect via standard approaches, retroviral vectors have been used in this chapter to transfect haematopoietic cells. Retroviral vectors can infect many commonly used haematopoietic cell lines and primary cells. One way of producing an infectious retrovirus containing a gene of interest is to transiently transfect a layer of HEK 293T cells with a retroviral plasmid construct and a helper ("packaging") plasmid(s) that encodes the required viral proteins (Gag, Pol and Env proteins) needed to produce infectious retroviral particles (see Pear *et al.*, 1993). The resulting retroviral supernatant is used to deliver the gene of interest to a target cell by infection of the cell. Following virus entry into the cell, the viral RNA molecule is reverse transcribed and the cDNA of interest, flanked by long terminal repeats (LTRs), is integrated into the target cell where expression of the gene of interest is driven by viral promoters within the LTRs. Retroviruses can carry selectable markers driven by separate promoters or linked by an internal ribosomal entry site (IRES) to the gene of interest. The use of selectable markers allows selection of the infected population containing the desired gene of interest for analysis.

The FDB-1 system is amenable to over-expression analysis and enables analysis of the effects of gene over-expression on growth, differentiation and survival in a myeloid system. It is likely that some of the genes identified as differentially expressed in FDB-1 cells are altered as a downstream consequence of other changes, and do not directly affect growth or differentiation. However, it is anticipated that some of the genes that were identified as differentially expressed in the study presented in Chapter 4 will be found to have a direct role in the growth, differentiation and/or survival of the FDB-1 cells.

It is predicted that genes that are more highly expressed in a proliferating population of FDB-1 cells (like the V449E expressing FDB-1 cells), when over-expressed in another population of cells that normally differentiate (e.g. the FIΔ expressing FDB-1 cells or parental FDB-1 cells grown in medium containing mGM-CSF), may lead to an enforced proliferation and/or a block in differentiation. Similarly, a differentiation-associated gene when over-expressed in conditions that normally support proliferation may cause differentiation of the FDB-1 cells. Functional experiments in the FDB-1 model therefore provide a valuable initial assessment of a gene's role in myeloid cell growth and differentiation.

5.2 Results

5.2.1 Expression Analysis

5.2.1.1 Rationale

To confirm the expression differences for several of the genes identified by the microarray analysis described in Chapter 4, Northern analysis was used. Analysis by Northern blots was chosen because this technique provides a method of measuring both gene expression differences and mRNA size (the latter of which is indicative of the specificity of hybridised target).

5.2.1.2 Northern blotting

5.2.1.2.1 Initial Northern Blots to confirm microarray data

Northern blots were performed using RNA that was prepared from the same batch of FDB-1 cells expressing V449E or FIΔ as was used for the microarray analysis. Cells were harvested at time 0 (i.e. growing in the presence of mIL-3) and 24 hours after mIL-3 withdrawal (mutant hβc signalling). Poly-A^+ RNA was isolated from these cell populations, separated by agarose gel electrophoresis, and transferred to a nitrocellulose membrane (see Section 2.3.5). Gene specific probes were obtained from clones produced by the National Institute of Aging (NIA). These clones contained the cDNA fragments represented on the microarray. The NIA clones (see Section 2.1.9) were restriction enzyme digested to isolate a specific probe fragment for each gene. Gene specific DNA sequences were radiolabelled with ^{32}P and used to probe the immobilised FDB-1 RNA on the Northern filter. This enabled the mRNA expression levels for each of the genes in the FDB-1 samples to be determined. The results of these Northern blots are shown in Figure 29. Two genes (*Sfrs3* and *Bnip3*) did not show obvious regulation on the Northern blots when compared to a control gene, Cyclophilin A (Cyc A) (*Ppia*), blotted on the same Northern filters. The results of the quantitation are presented in Figure 30 and summarised in Table 18. Figure 29 and Figure 30 and Table 18 show that the vast majority of the genes tested (with the exceptions of *Bnip3* and *Sfrs3*) were in fact differentially expressed between FIΔ and V449E expressing FDB-1 cells, and confirm the differences in gene expression identified by the microarray analysis. Northern blots of *Cd34, Sp100, Klf6* and *Nucleostemin* showed increases in expression over time (24 hours) in V449E cells (greater than 1.5 fold) whereas *Ltf*, *Hipk1*, *Cd24*,

S100a9, *Btg1*, and *Klf6* increased in expression over time in the FIΔ cells (greater than 1.5 fold) (see Figure 29, Figure 30 and Table 18).

Many of the genes showed differences greater than 1.5 fold between FIΔ and V449E even at time 0 on the Northern blots (e.g. *Ltf*, *Hipk1*, *Cd24*, *Cd34*, *Sp100*, *Klf6*, *Myc* and *Nucleostemin*) (see Figure 29 and Figure 30 and Table 18). For *Ltf*, *Hipk1* and *Nucleostemin* these time 0 differences in gene expression between V449E and FIΔ agreed with the findings from the microarray analysis (based on Adelaide scan data presented in Chapter 4, Table 13). However, for *Cd24*, *Cd34* and *Klf6* the differences in gene expression were opposite –i.e. all these genes were greater in V449E at time 0 on the microarray, but were greater in FIΔ on the Northern blots. *Sp100* showed a large time 0 difference in gene expression (higher in FIΔ than V449E) on the Northern blot that was not evident from the microarray analysis.

Table 18 summarises the Northern blot and the microarray data. The Northern blot expression analysis revealed that 11 of the 13 genes examined had differences in gene expression between V449E and FIΔ.

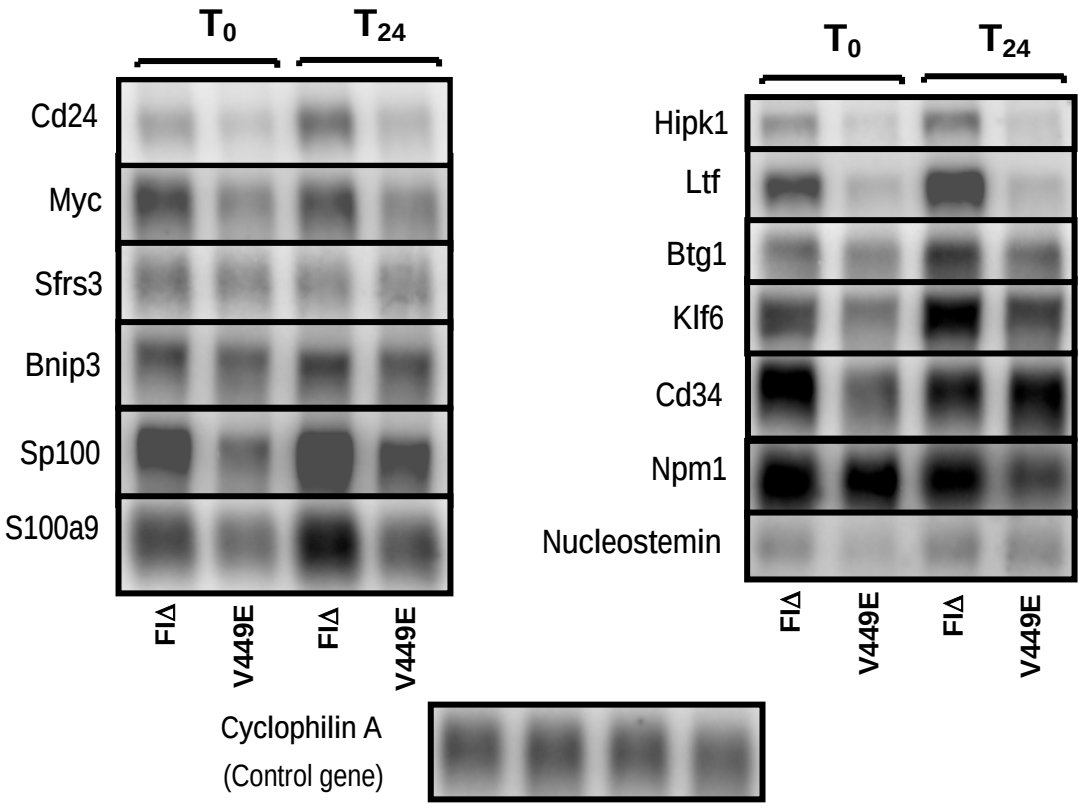


Figure 29: Initial Northern blots. Gene specific sequences were radiolabelled and used to probe Northern blots prepared from Poly-A^+ RNA purified from FDB-1 cells expressing either FIΔ or V449E hβc after 0 or 24 hours following mIL-3 withdrawal (mutant βc signalling). A representative Cyclophilin A probed Northern blot is shown for comparison. The results from these Northern blots are quantified and plotted in Figure 30.

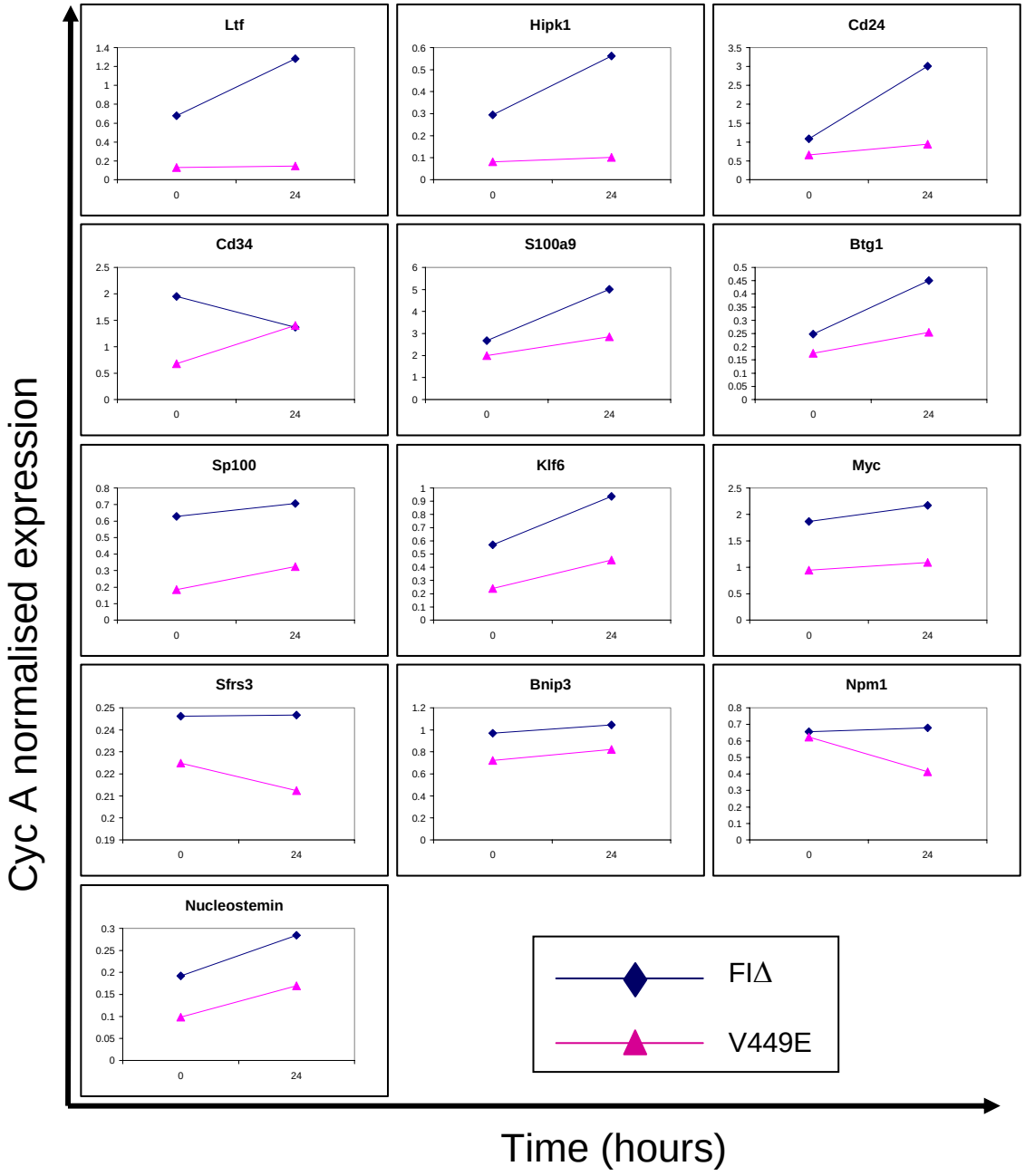


Figure 30: Normalised gene expression ratios. Initial Northern blots probed with the above genes were quantitated by comparison to the Cyclophilin A (Cyc A) control blot. The quantitation was performed using ImageQuant software and normalised to the expression of Cyclophilin A on the same Northern filter. The data is presented after time 0 and 24 hours of FIΔ or V449E expression in FDB-1 cells following mIL-3 withdrawal.

	Initial Northern blots				Microarray			
	F/V time 0	**V time 24/0**	**F time 24/0**	**F/V time 24**	**F/V time 0**	**V time 24/0**	**F time 24/0**	**F/V time 24**
Ltf	5.18	1.12	1.90	8.74	2.46	0.92	5.15	12.46
Hipk1	3.59	1.24	1.91	5.51	1.20	1.13	2.35	1.97
Cd24	1.66	1.44	2.77	3.20	0.91	0.92	2.88	2.05
Cd34	2.88	2.07	0.70	0.98	0.72	0.89	0.61	0.37
S100a9	1.34	1.43	1.87	1.76	2.81	0.98	2.66	7.49
Btg1	1.41	1.45	1.82	1.77	0.60	1.11	1.91	1.06
Sp100	3.42	1.76	1.12	2.18	0.93	1.10	1.71	1.76
Klf6	2.37	1.89	1.64	2.06	0.65	1.06	1.53	1.24
Myc	1.97	1.15	1.17	1.99	1.28	1.39	1.00	0.86
Sfrs3	1.09	0.94	1.00	1.16	1.14	1.05	0.79	0.78
Bnip3	1.34	1.14	1.08	1.27	1.07	1.12	0.70	0.54
Npm1	1.05	0.66	1.04	1.64	1.26	1.10	0.58	0.80
Ns	1.96	1.73	1.48	1.68	1.35	1.10	0.61	0.73

Table 18: Ratios obtained from the quantitated initial Northern blots and Microarray experiment. Time 0 and 24 hour FIΔ over V449E ratios are presented to allow comparison to microarray results presented in Chapter 4. The V449E and FIΔ time 24/0 ratios are presented to show clearly those genes that increase or decrease over time in these cell lines. The microarray data presented for this comparison uses the Adelaide scan (mean) values. Nucleostemin has been abbreviated to Ns.

5.2.1.2.2 Time course analysis - examining expression over 72 hours

Six of the genes that showed regulation on the above Northern blots were examined in more detail to determine their expression profile over a longer time course. This was achieved by preparing RNA from the two populations of FDB-1 cells at time 0, 6, 12, 24, 48 and 72 hours following mIL-3 withdrawal. The six genes selected were *Btg1*, *Hipk1*, *Cd24*, *Cd34*, *Klf6* and *Nucleostemin*. These genes were selected firstly because their differential expression was confirmed on the initial Northern blots, and secondly to investigate their expression at time points between 0 and 24 hours and to monitor their expression over later time points. This could reveal possible involvement in early (6 and 12 hours) or late (48 and 72 hours) events in differentiation or in early events associated with switching from mIL-3 dependent growth.

Quantitation was performed via comparison to the Cyc A probed Northern blot (as described for the data presented in Figure 30) and the results from the time course analysis are shown in Figure 31 and Figure 32. Four of the genes showed increased expression in the FIΔ cell line over time (Figure 32). The two exceptions were *Cd34* and *Nucleostemin* where there is a dramatic decrease in expression of both of these genes in FIΔ over the 72 hours. For *Cd34*, this is matched by a similar decrease in its expression in V449E over time, but for *Nucleostemin* its expression increases over time in V449E cells. *Hipk1* and *Klf6* also show moderate increases in their expression in the V449E cells over the 72 hours while *Cd24* and *Btg1* show constant expression within the V449E cells over this time period (see Figure 31 and Figure 32). There was a noticeable decrease in expression at the 12 hour time point in both the FIΔ and V449E expressing cells for all the genes studied.It is unclear whether this is a real response in the cells to FIΔ or V449E signalling, or an artefact the experimental conditions.

5.2.1.2.3 Concordance between initial and time course Northern blots

Previous data showed that there were differences between the two cell lines at time 0 (see Section 5.2.1.2.1). In the time course analysis these differences at time 0 are still apparent for *Nucleostemin*, *Klf6* and *Cd24*. *Hipk1* was also different but only by 1.5 fold on the time course analysis at time 0, compared to 3.6 fold on the initial Northern blots (Table 18 and Table 19). The time zero difference for *Cd24* was however increased on the time-course Northern compared to initial Northern blot (4.9 fold versus 1.7 fold respectively) (Table 18 and Table 19). Examination of the time 0 and time 24 FIΔ over V449E ratios found that there was generally good agreement between the initial

Northern blots and the time course analysis, with two exceptions – *Cd34* and *Nucleostemin*. *Cd34* at time 0 was found to have ratios of 2.9 and 0.9 fold for the initial Northern blot and time course analysis respectively, while *Nucleostemin* at time 24 had ratios of 1.7 and 1.0 for the initial Northern blot and time course analysis respectively. Interestingly, while there was no significant difference at 24 hours, later time points (48 and 72 hours) showed an increasing difference between V449E and FIΔ such that at 72 hours V449E expression was three fold greater than FIΔ. Also some of the genes showed a different magnitude of change but that change was in the same direction for the initial Northern blots and time course analysis (e.g. *Cd24* time 24 ratio was 3.2 and 7.2 fold greater in FIΔ than V449E for the initial Northern blots and time course analysis respectively).

In summary, the time course Northern blots showed that the time 24 hour differences in gene expression between V449E and FIΔ that were revealed by the microarray analysis and initial northern blots were also evident at earlier (0, 6, 12 hour) and later (48 and 72 hour) time points.

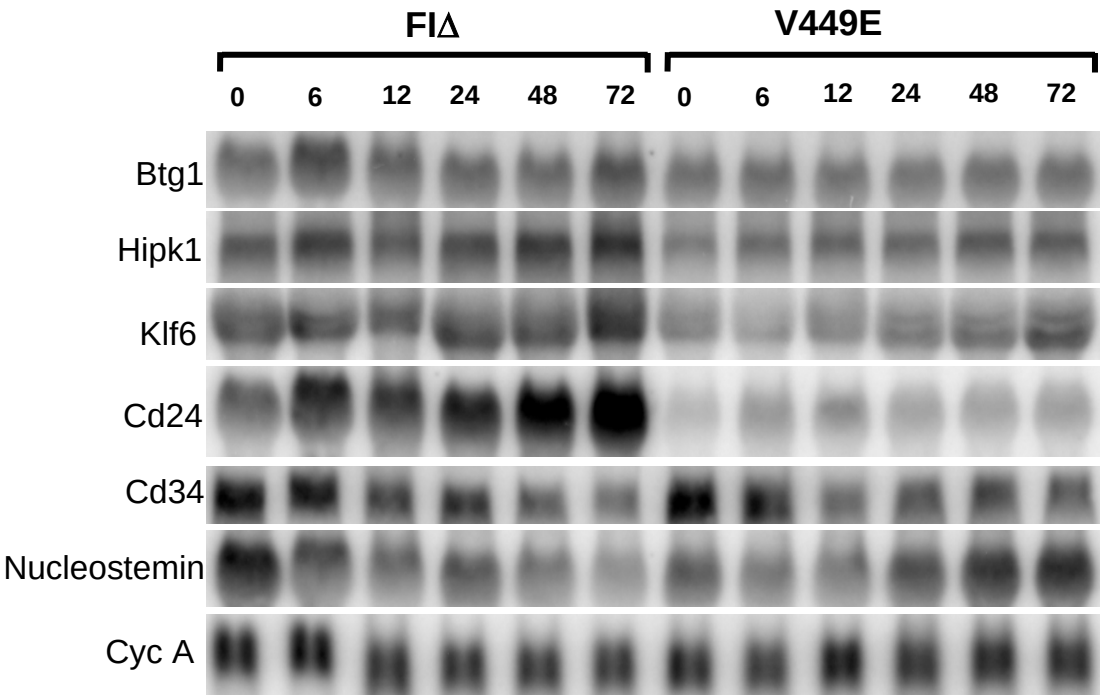


Figure 31: Time course Northern blots. Genes specific sequences were radiolabelled and used to probe a Northern Blot prepared from total RNA from FDB-1 cells expressing either FIΔ or V449E hβc after 0, 6,12, 24, 48, and 72 hours following mIL-3 withdrawal (mutant βc signalling). The Cyc A probed Northern blot is shown as a loading control and was used to normalise the expression levels of each gene.

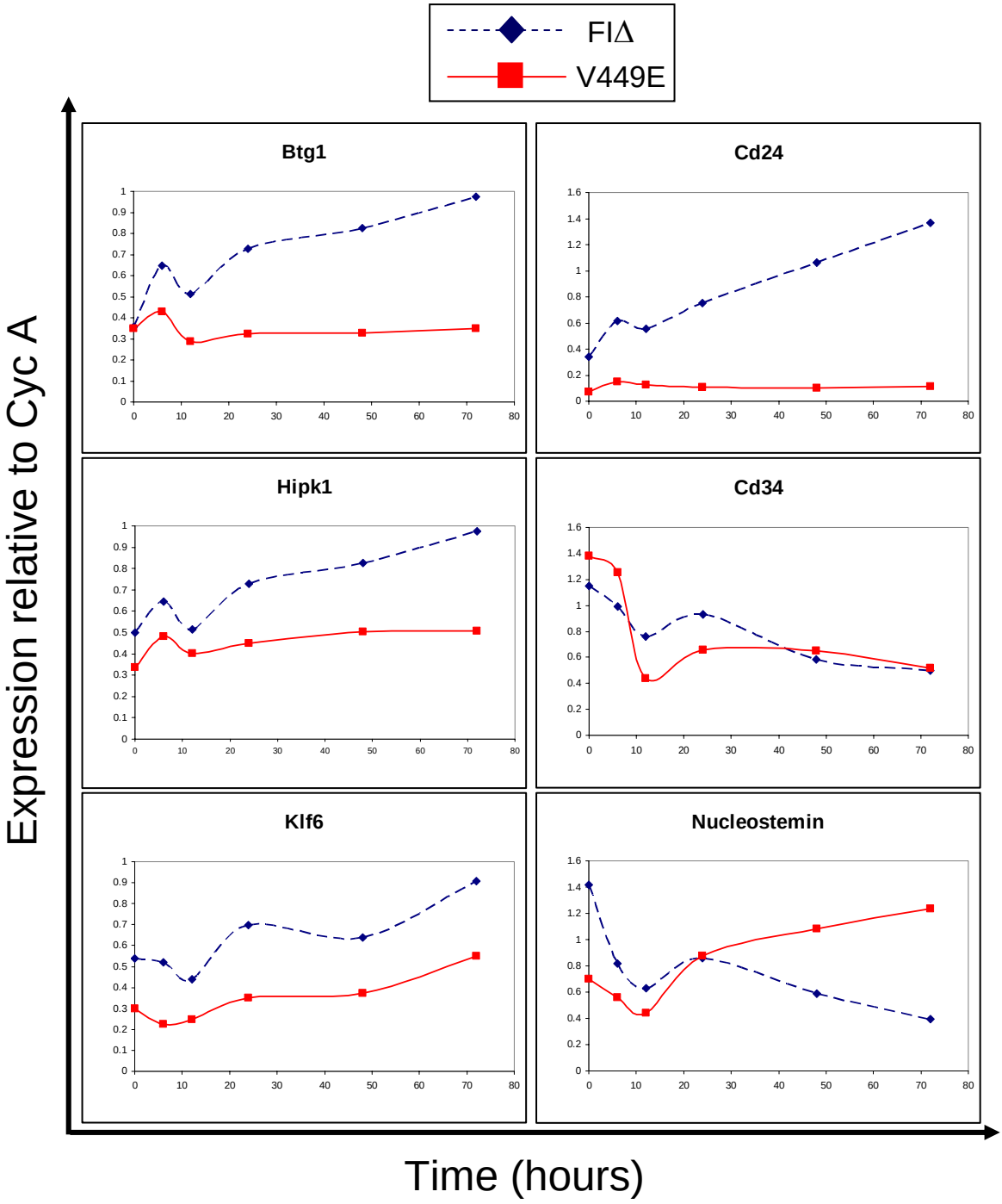


Figure 32: Time course Northern normalised expression data. Expression profiles of the six selected genes are shown after normalisation to the expression level of Cyclophilin A (Cyc A). The quantitation shows the changes in gene expression in FIΔ and V449E expressing FDB-1 cells over 72 hours following mIL-3 withdrawal.

	Time course Northern blot			
	F/V time 0	**V time 24/0**	**F time 24/0**	**F/V time 24**
Btg1	1.02	0.93	2.04	2.25
Hipk1	1.49	1.34	1.45	1.62
Klf6	1.80	1.17	1.29	1.99
Cd24	4.93	1.52	2.21	7.17
Ns	2.03	1.25	0.60	0.98
Cd34	0.85	0.44	0.77	1.50

Table 19: Ratios obtained from the quantitated Time course Northern blots. Time 0 and 24 hour FIΔ over V449E ratios are present to allow comparison to microarray results presented in Chapter 4 and those from the initial Northern blots (see Table 18). The V449E and FIΔ time 24/0 ratios are also presented for comparison to previous results but a better representation of those genes that increase or decrease over time in these cell lines is obtained from the graphs above (Figure 32). Nucleostemin has been abbreviated to Ns.

5.2.1.3 CD24 protein expression at the cell surface of FDB-1 cells.

Cd24 mRNA displayed a striking change in expression on Northern analysis between V449E and FIΔ. To establish whether the change in mRNA levels resulted in changes in CD24 protein levels, the cell surface levels of CD24 in the parental FDB-1 cells and in FDB-1 cells expressing the V449E and FIΔ mutants were examined. It was expected that cell surface expression of CD24 would correlate with total protein levels as CD24 is a cell surface receptor (Pierres *et al.*, 1987; Alterman *et al.*, 1990; Kay *et al.*, 1990; Kay *et al.*, 1991).

5.2.1.3.1 CD24 protein expression in parental FDB-1 cells

Levels of CD24 protein were measured via flow cytometry in the FDB-1 cell population using an anti-CD24 monoclonal antibody conjugated to phycoerythrin (PE) and compared to the staining observed with an irrelevant isotype matched control antibody. The cell surface expression of CD24 protein was examined on parental FDB-1 cells which were washed and placed into culture medium containing either mIL-3 (to induce proliferation) or mGM-CSF (to induce differentiation). The results of these experiments on the parental FDB-1 cells showed that in the presence of mIL-3 on day 3 and 5, 8-11% of the parental FDB-1 cells grown in mIL-3 showed binding of the anti-CD24 antibody beyond the level seen with the irrelevant isotype matched control antibody (Figure 33 and data not shown). This binding indicates that 8-11% of the parental FDB-

1 cells grown in mIL3 express CD24 protein at the cell surface. When parental FDB-1 cells were placed in culture medium containing mGM-CSF, there was an increase in CD24 protein expression seen on day 3 and day 5 (see Figure 33). This is indicative of there being an increase in CD24 protein expression upon differentiation, and the gene expression analysis suggesting that a significant component of this is due to up-regulation of *Cd24* transcript.

5.2.1.3.2 CD24 protein expression in V449E and FIΔ expressing FDB-1 cells

In light of the increased protein expression seen in differentiating FDB-1 cells, the next step was to confirm whether FDB-1 cells expressing the two hβc mutants (V449E and FIΔ) also express differing levels of CD24 protein. Differences in the CD24 protein expression were predicted in the two hβc mutants given the findings from the microarray analysis and Northern blots above. Parental FDB-1 cells, V449E and FIΔ expressing FDB-1 cells were maintained in medium containing mIL-3 and their levels of CD24 protein expressed at the cell surface were determined using an anti-CD24 antibody as described in Section 5.2.1.3.1. In Figure 34 it can be seen there is a population of cells expressing CD24 in both the V449E and parental FDB-1 cell populations (10% and 9% respectively). This level of expression of CD24 protein is comparable to the CD24 expression seen in the parental FDB-1 population proliferating in mIL3 for 3 days shown in the previous experiment (11% see Figure 33). In contrast, the FIΔ expressing FDB-1 cells maintained in mIL-3 have a larger proportion of cells expressing CD24 (25%). The fact that V449E expression of CD24 is comparable to wild type FDB-1 cells indicates that there is an increased expression of CD24 protein that is specifically associated with FIΔ expression (rather than a decreased expression of CD24 protein in V449E cells) (Figure 34). This finding agrees with the results of both the microarray and Northern analysis and confirms that there is an increased expression of CD24 at the RNA and protein level when the FIΔ mutant is expressed in the FDB-1 cells in the presence of IL-3. Differentiation of FIΔ cells would be expect to lead to an increased expression of CD24 protein, given that the mRNA levels of *Cd24* increase over 72 hours, and that differentiation of the parental FDB-1 cells is also associated with an increase in CD24 protein expression at the cell surface.

In summary, it has been demonstrated that differentiation of FDB-1 cells with mGM-CSF and expression of FIΔ in FDB-1 cells leads to an increases in CD24 protein

expression at the cell surface. This increase in CD24 protein levels reflects a difference in *Cd24* mRNA levels in FIΔ cells shown in the Northern blots and microarray analysis.

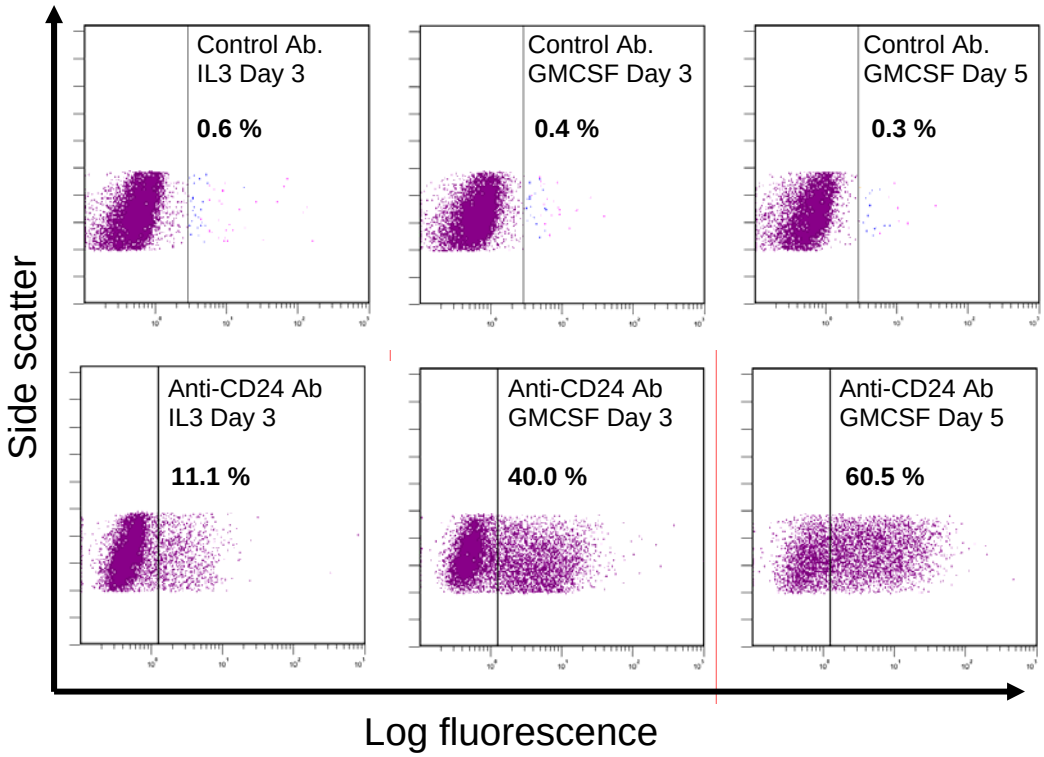


Figure 33: CD24 protein expression on the cell surface of FDB-1 cells. FDB-1 cells were washed then grown in medium containing either mIL-3 (IL3) (500 U/mL) or mGM-CSF (GM) (500 U/mL) for up to 5 days. After 3 and 5 days of culture, cells were collected for staining with either an isotype matched control antibody or anti-CD24 antibody prior to analysis via flow cytometry. Flow cytometry profiles of the isotype matched control antibody ("Control Ab.") and anti-CD24 antibody ("Anti-CD24 Ab.") stained FDB-1 cells grown under the stated conditions are presented. The number of cells positive for CD24 staining are shown on the profiles as a percentage of the total live cells analysed for a representative experiment. Note the FDB-1 cells grown in mIL-3 on Day 5 showed 8% CD24 positive cells (data not shown).

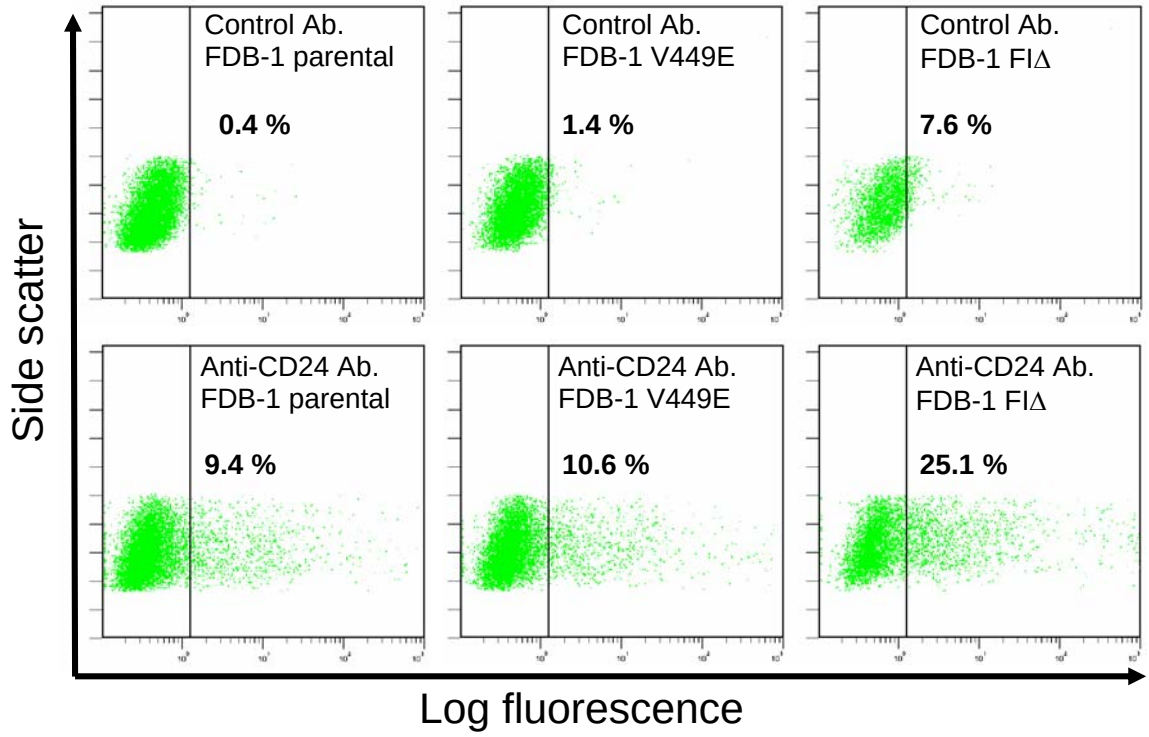


Figure 34: CD24 protein expression on the cell surface of FDB-1 cells expressing FIΔ and V449E. FDB-1 cells ("FDB-1 parental") and FDB-1 cells expressing either FIΔ ("FIΔ"), or V449E ("V449E") were grown continuously in the presence of mIL-3 (500 U/mL) and stained with either an isotype matched control antibody or anti-CD24 antibody prior to analysis via flow cytometry. Flow cytometry profiles obtained with the isotype matched control antibody ("Control Ab.") and anti-CD24 antibody ("Anti-CD24 Ab.") on the FDB-1 cells are presented. The number of cells positive for CD24 staining are shown on the profiles as a percentage of the total live cells analysed for a representative experiment.

5.2.2 Functional Studies

The approach used to investigate the function of four genes (*Btg1*, *Cd24*, *Klf6* and *Nucleostemin*) was over-expression using retroviral vectors in the FDB-1 cell line. The effect of over-expression of *Btg1* was examined in the FDC-P1 cell line. These two cell lines were used for a preliminary investigation of gene function that could be used to prioritise genes for further study.

5.2.2.1 Cloning strategy for the target genes

The coding regions of the four genes of interest (*Btg1*, *Cd24*, *Klf6* and *Nucleostemin*) were PCR-amplified from FDB-1 cDNA (see Figure 35 and Table 20), cloned into MSCV-IRES-GFP vector and sequenced for verification of their identity. In addition, these four genes were also PCR amplified with an in frame c-terminal haemagglutinin (HA) tag to enable detection of the exogenous expressed proteins.

5.2.2.2 Retroviral infection of FDC-P1 and FDB-1 cells

Many haematopoietic cells are resistant to standard transfection methods used to over-express a gene of interest in the target cells. An efficient alternative involves the infection of target cells (e.g. FDB-1 cells) with a replication incompetent retrovirus. This infection generates a pool of infected cells that express a gene of interest, which can be isolated and used in functional assays. When efficient infection occurs, a large pool of cells can be analysed. Efficient infection of haematopoietic cells is important as it allows analysis of many different independent clones and would avoid the consequences of the analysis of clonal artefacts.

A standard technique for infecting FDB-1 and FDC-P1 cell lines was used in these studies (see Section 2.4.4.2 and Figure 36). Briefly, this involved the transient co-transfection of HEK 293T cells with the ecotropic packaging plasmid pEQ (Persons *et al.*, 1998), which provided the necessary viral packaging products, and a retroviral plasmid containing the gene of interest. Twenty-four hours post transfection, FDB-1 or FDC-P1 cells were co-cultured with the HEK 293T cells for two days, before the FDB-1 or FDC-P1 cells were harvested. The infected cells used in this study were selected based on their expression of the fluorescent protein, Green Fluorescent Protein (GFP), which was encoded by the retroviral vector and linked to the gene of interest by an internal ribosomal entry site (IRES). Flow cytometry was used to select and sort infected cells based on their expression of GFP (see Figure 36). This approach to use

GFP as a selectable marker was undertaken as antibiotic selection was predicted to be problematic for cells expressing a gene that negatively affects growth. The approach used in this chapter (see Figure 36) can be adopted for use on other haematopoietic cells to compare with the results obtained in the FDB-1 cell system.

The retroviral vector that was used for the over-expression studies was the MSCV-IRES-GFP vector (Van Parijs *et al.*, 1999). The Murine Stem Cell Virus (MSCV) vectors were derived from the Murine Embryonic Stem Cell Virus (MESV) and the LN retroviral vectors (Miller and Rosman, 1989; Grez *et al.*, 1990). The MSCV vector is able to achieve high-level gene expression in haematopoietic and embryonic stem cells via a mutated 5' LTR (Grez *et al.*, 1990; Cherry *et al.*, 2000). This mutated 5' LTR is preferred to the LTR from Moloney Murine Leukaemia Virus (MMLV) which has very low activity in stem cells (Hawley *et al.*, 1994). The IRES present in this vector enables a gene cloned into the multiple cloning site (MCS) upstream of GFP to be translated from a single bicistronic mRNA (Jang *et al.*, 1988; Jackson *et al.*, 1990). This expression of both GFP and the cloned gene allows enrichment and analysis of infected cell populations via flow cytometry as described above.

Gene name	ORF Length (bp)	Expected PCR product size (bp)
Btg1	516	541
Btg1-HA		568
Cd24	231	256
Cd24-HA		283
Klf6	855	880
Klf6-HA		907
Nucleostemin	1617	1643
Ns-HA		1670

Table 20: PCR amplification of the genes for functional analysis. The length of the open reading frame (ORF) for each of the genes is stated. The expected lengths of the PCR products are shown after the incorporation of the primers and also the haemagglutinin (HA) tag to the amplified ORF. Primers used to amplify the eight PCR products are stated in Section 2.2.10.4. The size determined for each gene's ORF was based on the following sequences obtained from the GenBank database: Btg1 = NM_007569, Cd24= NM_009846, Klf6= NM_011803 and Nucleostemin = NM_153547 (long isoform).

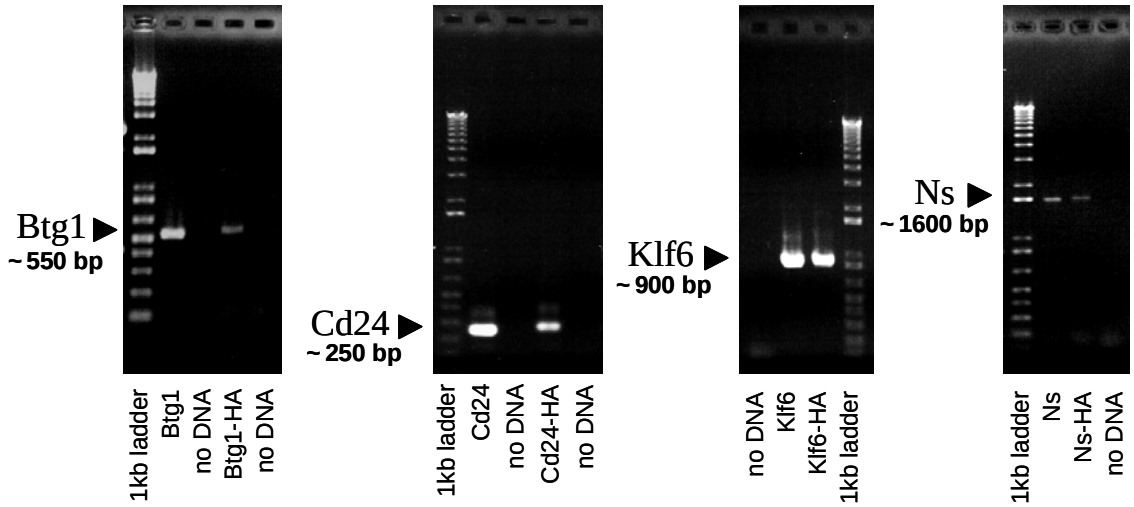


Figure 35: RT-PCR amplification of genes for functional studies. FDB-1 RNA was reverse transcribed and used as a template to amplify the open reading frames of Btg1, Cd24, Klf6 and Nucleostemin (Ns) via PCR. Primers used to amplify the eight PCR products are stated in Section 2.2.10.4. Primers containing the haemagglutinin tag (HA) were used to amplify the ORF of each gene together with the sequence for the HA tag. No template PCR were run as controls (no DNA). PCR products were cloned and sequenced to confirm their identity.

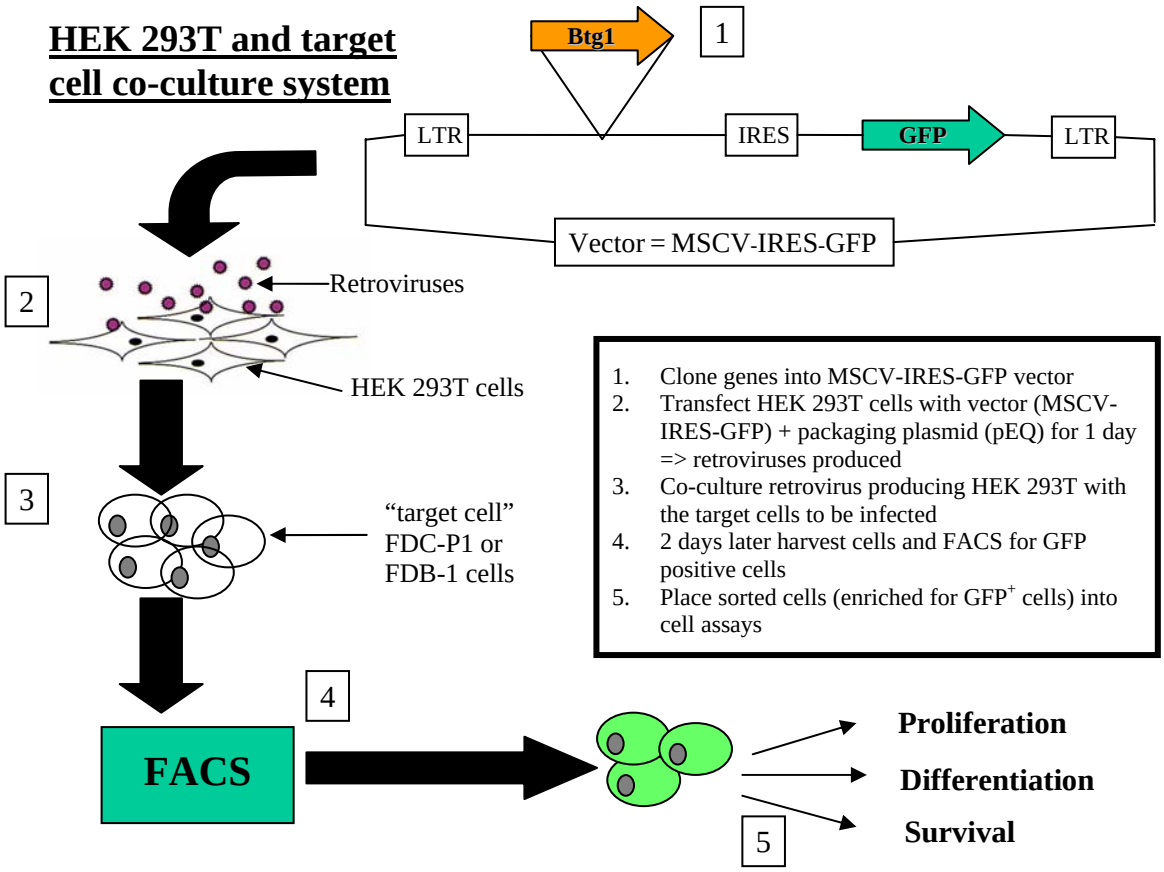


Figure 36: The HEK 293T and target cell co-culture infection system. 1. Genes of interest are cloned into the MSCV-IRES-GFP vector. 2. HEK 293T cells are transfected with the gene containing MSCV vector from step 1 and the necessary packaging plasmid (e.g. pEQ). 3. Either FDC-P1 or FDB-1 cells are placed on top of the adherent HEK 293T cells in co-culture. 4. After 2 days the FDC-P1/FDB-1 cells are harvested and then analysed and sorted via fluorescence activated cell sorting (FACS). 5. The sorted GFP-positive (GFP^+) cells can then be placed in cell assays to assess a given gene's function in cell proliferation, differentiation and survival.

5.2.2.3 Analysis of BTG1 function in FDC-P1 cells

The first functional experiments were performed with BTG1 (Figure 36) in the FDC-P1 cell line. The *BTG1* gene is established in the literature as an anti-proliferative gene (Rouault *et al.*, 1992) and it was reasoned that its over-expression in FDC-P1 cells may alter their proliferation rate. The FDC-P1 cell line was chosen as it could be readily infected with retroviruses and used to investigate the effects of BTG1 on proliferation in response to GM-CSF and IL-3. FDC-P1 cells were co-cultured with HEK 293T cells producing MSCV-BTG1-IRES-GFP retroviruses and after infection, GFP-positive (GFP^+) FDC-P1 cells were sorted by FACS. The population of infected cells expressing GFP and BTG1 was obtained and proliferation assays were performed to determine whether BTG1 affects FDC-P1 cell proliferation rate. The result of one experiment is shown below (Figure 37 and Figure 38). In this experiment, 18% of cells were infected with the BTG1 retroviral vector after co-culture of the FDC-P1 cells with the retrovirus-producing HEK 293T cells, as measured by GFP expression (Figure 37). These infected cells were sorted via FACS to obtain an enriched population of GFP^+/BTG1 expressing cells (Figure 37), which were analysed in a proliferation assay to determine the effect of unregulated BTG1 expression on FDC-P1 growth and compared to a control population infected with empty vector. In Figure 38, a minor effect on the proliferation of FDC-P1 cells could be seen at day 4. The BTG1 expressing cells showed slightly reduced proliferation compared to the control population of cells expressing the empty MSCV-vector. This effect was significantly different ($p < 0.05$) between the control and BTG1 infected cell lines grown in mIL3 at day 4. The factor dependence of the infected FDC-P1 cells was confirmed by their lack of growth in the absence of mIL-3 (Figure 38). The experiment was repeated to confirm this reduced proliferation, however, it was decided that sorting the cells into two distinct populations expressing different levels of BTG1 might reveal the effects of BTG1 more clearly.

In this experiment the infected FDC-P1 cells were sorted into GFP^{HI} and GFP^{LOW} positive populations and assayed for proliferation (Figure 39 and Figure 40). In this experiment, 14% of the FDC-P1 cells were infected with the BTG1 retroviral vector as measured by GFP expression and sorted into GFP^{HI} and GFP^{LOW} (Figure 39). Figure 39 shows the percentage of infected cells collected for the GFP^{HI} and GFP^{LOW} populations and revealed that the purity of the sorted populations were 87% and 92% for the GFP^{HI}/BTG1 and GFP^{LOW}/BTG1 populations respectively. GFP^{HI} and GFP^{LOW}

populations were sorted for the empty vector infected FDC-P1 cells and all four populations of cells were compared in a proliferation assay (Figure 40). This experiment tested whether increased expression of BTG1 in the GFP^{HI} population leads to a decrease in the proliferation rate of FDC-P1 cells. In this experiment, there was no significant difference between the GFP^{HI} and GFP^{LOW} expressing BTG1 FDC-P1 populations and control populations in terms of their proliferation rates (Figure 40 and data not shown). In summary, the use of a co-culture infection system enabled the infection of FDC-P1 cells with a BTG1 expressing retroviral vector, which indicated that BTG1 over-expression does not profoundly affect the proliferation rate of FDC-P1 cells.

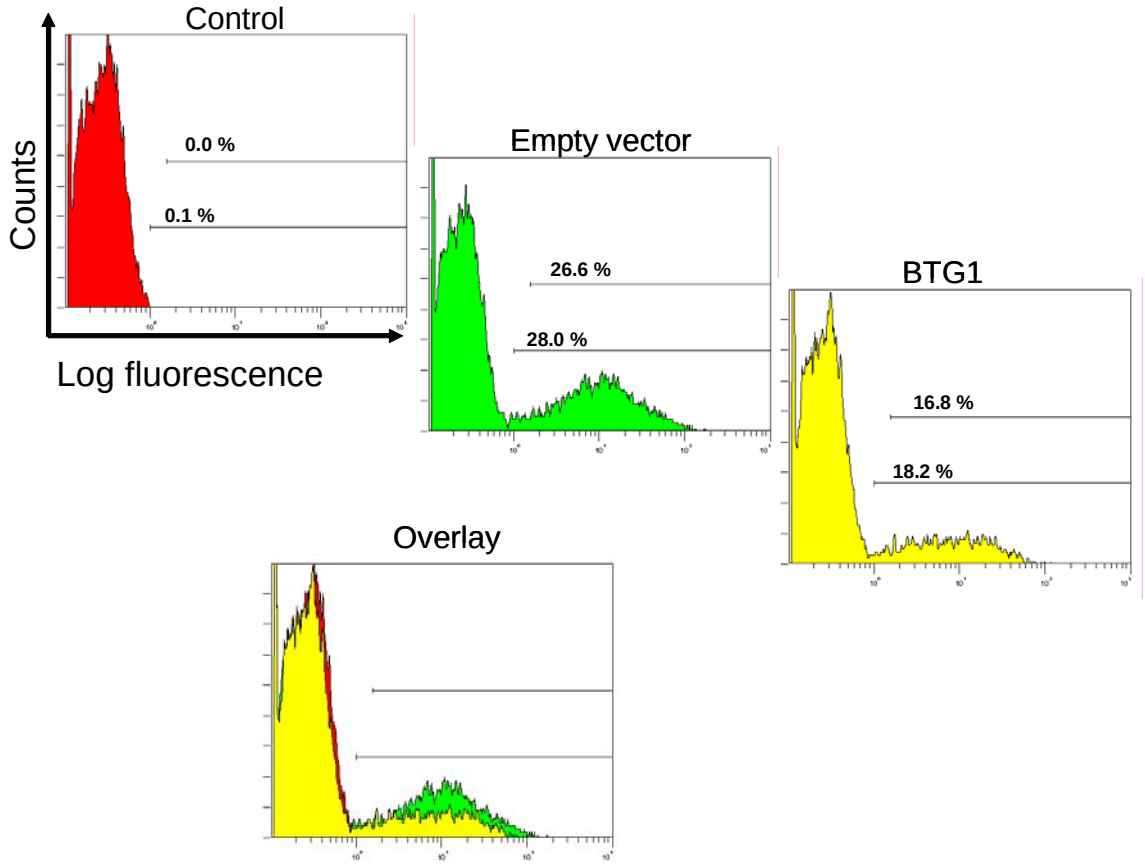


Figure 37: Fluorescence activated cell sorting (FACS) of the BTG1 and empty vector infected FDC-P1 cells. Retrovirally infected FDC-P1 cells were analysed and sorted based on the expression of GFP. "BTG1 infected" FDC-P1 cells were infected with retroviral vector containing the BTG1 orf linked via an IRES to GFP (BTG1-MSCV-IRES-GFP), whereas the "empty vector infected" FDC-P1 cells were infected with the empty retroviral vector containing GFP (MSCV-IRES-GFP). The largest gate indicates the cells positive for expression GFP and the smaller gate shows the expression of the cells that were collected. Uninfected FDC-P1 cells were used as the "control" population to define the non GFP expressing cells. In each panel, the x-axis scale for the major graduations is 10^0, 10^1, 10^2 and 10^3.

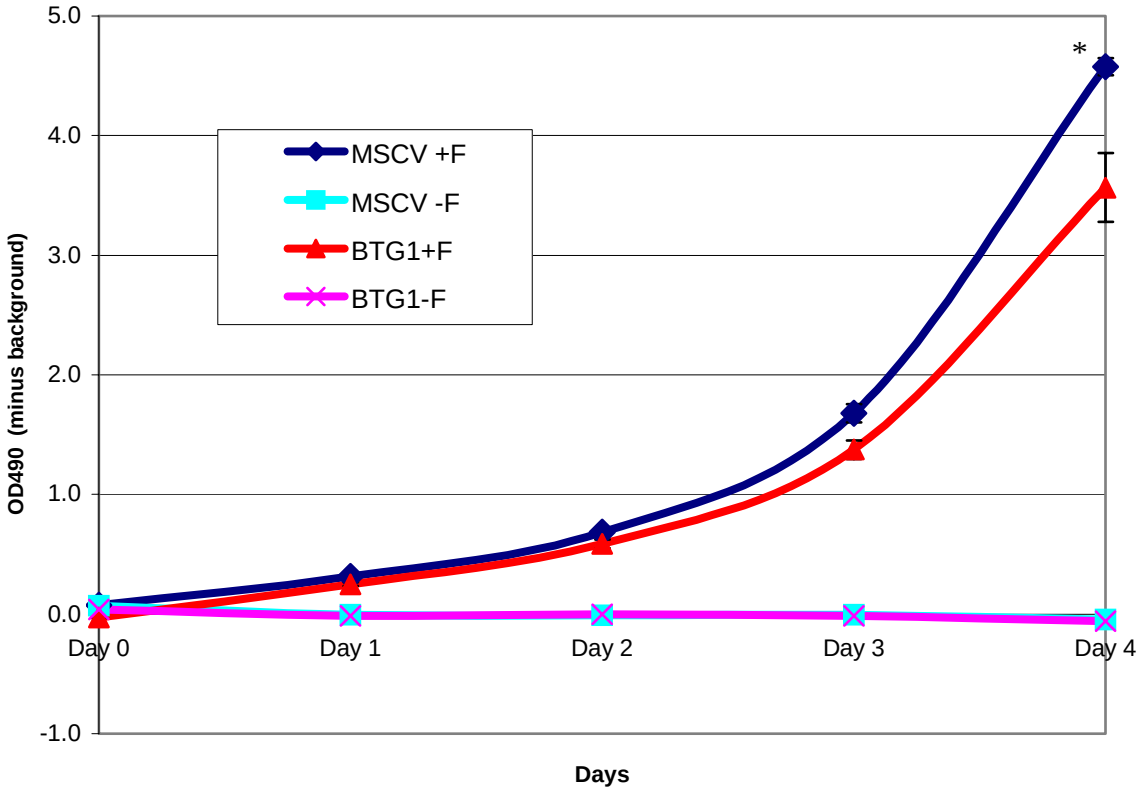


Figure 38: Proliferation of FDC-P1 cells infected with MSCV-BTG1. FDC-P1 cells were infected with either the empty or BTG-1 containing MSCV-IRES-GFP vector and after sorting were placed into a four day proliferation assay in the presence (+F) and absence (-F) of growth factor (mIL-3 used at 100 U/mL). Proliferation was determined using an aqueous proliferation assay dye and measuring the coloured product formed after 2 hours by viable cells at OD490 which is directly proportional to the number of cells. * = p value less than 0.05 via Student's t-Test

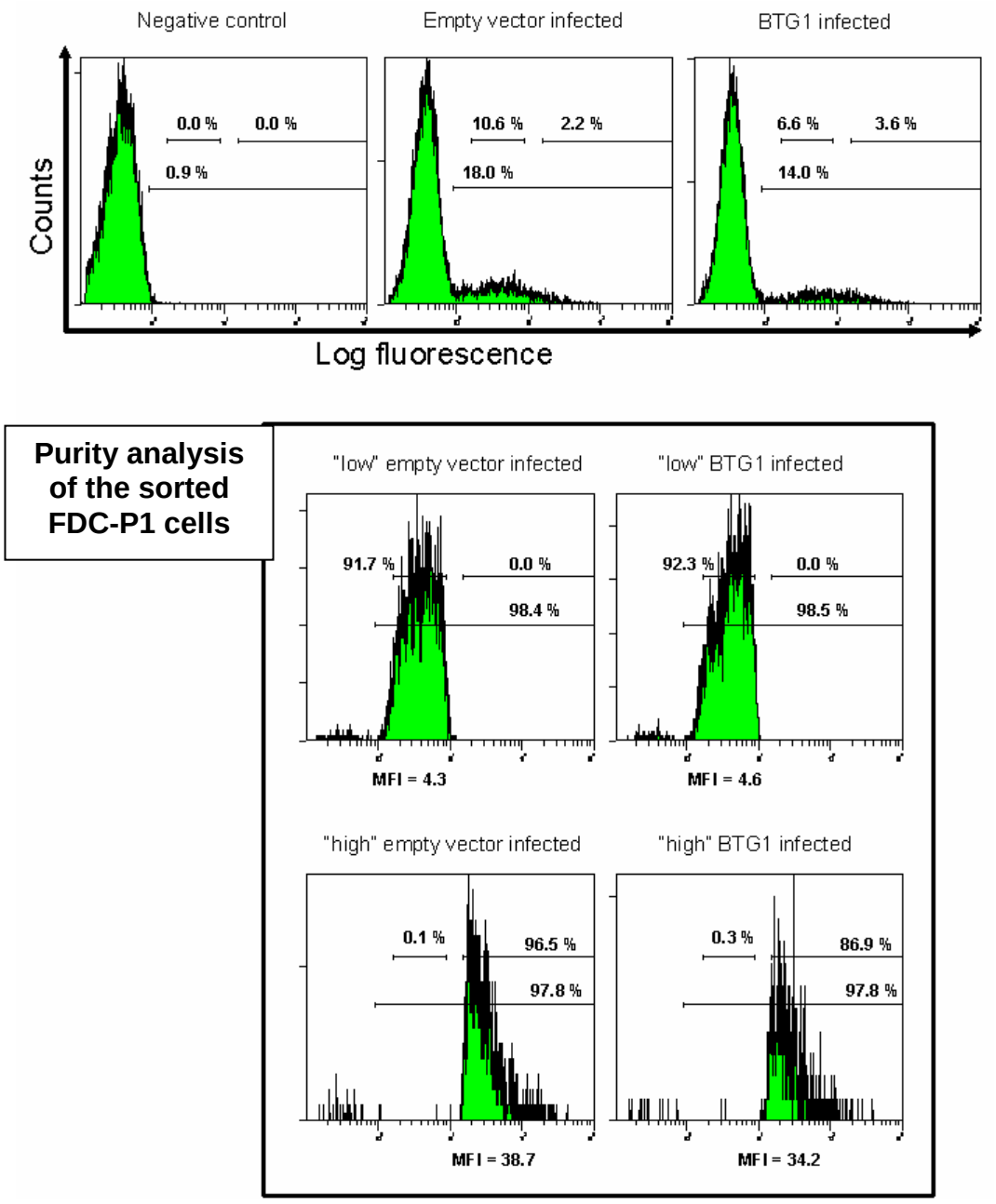


Figure 39: Fluorescence activated cell sorting (FACS) of the BTG1 and empty vector infected FDC-P1 cells into high and low expressing populations. Retrovirally infected FDC-P1 cells were analysed and sorted into high and low expressing cells based on the expression of GFP. "BTG1 infected" FDC-P1 cells were infected with retroviral vector containing the BTG1 open reading frame linked via an IRES to GFP (MSCV-BTG1-IRES-GFP), whereas the "empty vector infected" FDC-P1 cells were infected with the empty retroviral vector containing GFP (MSCV-IRES-GFP). The largest gate indicates the cells positive for expression GFP and the two smaller gates show the expression of the cells that were collected as the GFP^{HI} ("high") and GFP^{LOW} ("low") expressing cell populations. A sample of each batch of sorted cells was reanalysed to determine the purity of the sorted population (lower panel). Mean fluorescence intensity (MFI) is indicated on the x-axis to indicate the intensity of the GFP^{HI} and GFP^{LOW} populations. Uninfected FDC-P1 cells were used as the "control" population to define the non GFP expressing cells.

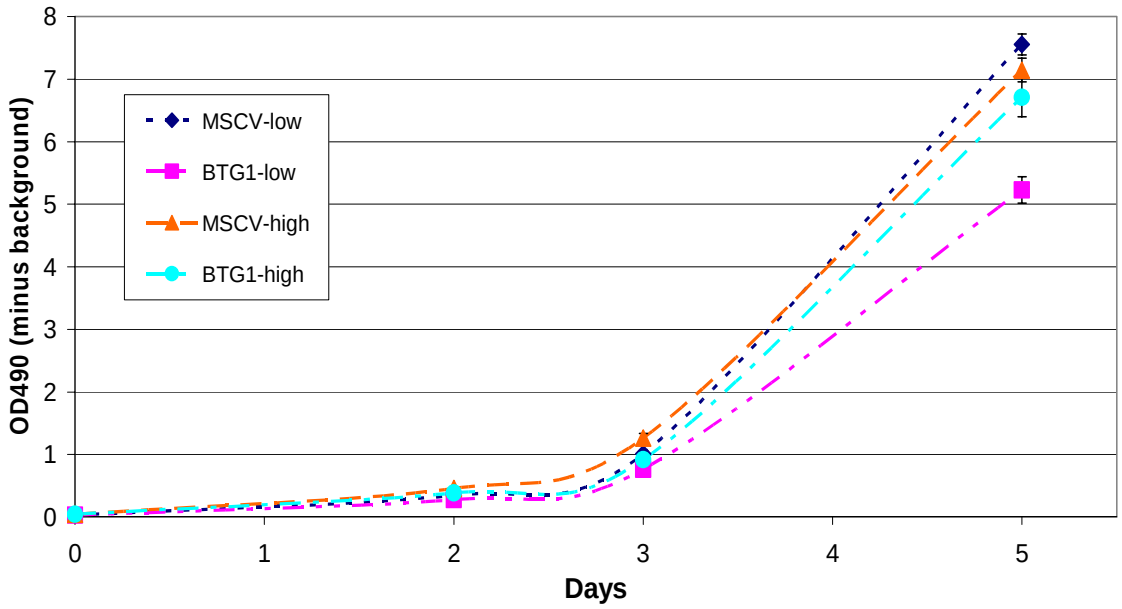


Figure 40: Proliferation of FDC-P1 cells infected with high and low levels of MSCV-BTG1. FDC-P1 cells were infected with either the empty or BTG-1 containing MSCV-IRES-GFP vector and were sorted into high and low GFP expression for both constructs. Proliferation rates for the four cells populations were compared in a five day assay. Proliferation was determined by the addition of an aqueous proliferation assay dye and measuring the coloured product formed after two hours by the viable cells at OD490. The formation of the coloured product is directly proportional to the number of viable cells.

5.2.2.4 Analysis of gene function in FDB-1 cells

To explore the role of genes identified by the microarray analysis a major aim of work described in this chapter was to introduce retroviral constructs into the FDB-1 cell line. The ability to introduce retroviral constructs in the FDB-1 cell line would permit over-expression and gene knockdown approaches in an experimental model allowing prioritisation of genes for study in primary cell systems. Initial experiments with FDB-1 cells were problematic because retroviral infection routinely led to the infection of one to two percent of the FDB-1 cells, which after FACS purification, led to the recovery of a small number of GFP-positive cells (data not shown). The sorted population required expansion of the number of cells prior to a second round of sorting to obtain a sufficiently enriched population that could be used in proliferation or differentiation assays. The sorting procedure was further complicated by significant contamination of the sorted FDB-1 cell population with more rapidly proliferating HEK 293T cells from the co-culture infection procedure. The contamination of the sorted cells with HEK 293T cells prevented the accurate assessment of the proliferation of the FDB-1 cells after retroviral infection with the genes of interest. The aim of the work described in Section 5.2.2.5 was to optimise or modify the infection method to obtain increased numbers of infected cells in order to perform experiments looking at proliferation and differentiation responses of FDB-1 cells.

5.2.2.5 FDB-1 infections – optimisation of retroviral infection

Due to the low percentage infection obtained by the standard approach (see Section 5.2.2.4), it was necessary to optimise the infection procedure to increase the number of infected FDB-1 cells obtained. Two approaches were examined; (1) the use of a pantropic retroviral packaging vector that uses the G glycoprotein of vesicular somatitis virus (VSV-G) and (2) the improvement of the co-culture infection method.

In contrast to the co-culture system outlined in Section 5.2.2.2, which described the use of the ecotropic (pEQ) packaging system, the VSV-G packaging system required co-transfection of three plasmids. The first plasmid (pVPackVSV-G) contained the *env*-expressing gene, which encoded the viral envelope protein, VSV-G. The second plasmid (pVPack-GP) contained the *gag-pol* genes required for virus production (e.g. structural proteins, reverse transcriptase and integrase). The third plasmid was the replication deficient retroviral vector, which in these experiments was the MSCV-IRES-GFP vector that contained the gene of interest. Unlike other MMLV–derived envelope

proteins, which recognize cell surface receptors to gain viral entry into the cell (such as the ecotropic receptor used by the pEQ retroviral packaging system), VSV-G recognizes a phospholipid (G glycoprotein) that is present on mammalian and non-mammalian cell types and allows the infection of broad range of mitotic cell types (Emi *et al.*, 1991; Burns *et al.*, 1993; Yee *et al.*, 1994a; Yee *et al.*, 1994b).

5.2.2.5.1 Testing of the VSV-G retroviral packaging system

A pantropic retroviral packaging system (consisting of the pVPack-VSV-G and pVPack-GP plasmid vectors) was tested as it was reasoned it would have a better infection rate of FDB-1 cells because virus entry would not be limited to expression of the mouse ecotropic receptor (unlike the situation with the pEQ retroviral packaging system). When the VSV-G packaging system was tested in the co-culture system it was found that an almost identical level of infection of the FDB-1 cells was obtained (data not shown). This finding suggested that it was not the particular retroviral packaging system that was responsible for the poor percentage infection of the FDB-1 cells and pointed to the fact that the co-culture system may need optimisation (see Section 5.2.2.5.3).

5.2.2.5.2 Determination of retroviral titres

To determine the concentration of retroviruses produced by these two retroviral packaging systems, their viral titres were determined following their transfection into HEK 293T cells and the subsequent infection of HEK 293T and NIH 3T3 cells with the retroviral supernatant. The MSCV-IRES-GFP vector was used with both these systems to determine infected cells by expression of GFP by flow cytometry. Viral titres were determined by calculating the number of GFP-positive infected target cells obtained from a limiting dilution of the viral supernatant. As the number of virus particles is inferred by the infection of a target cell, the viral titre is expressed as transducing units per mL. The actual viral titre may be higher than the calculated transducing units per mL due to multiple virus particles infecting the same cell (which occurs with high concentrations of the retroviral supernatant). The NIH 3T3 cells were selected as a target cell line for two reasons. Firstly, these cells have been widely used in the past for determining viral titres. Secondly this enables the viral titre from the pEQ system to be determined as this system can only infect mouse cell lines and would be unable to infect the human HEK 293T cells. The results of this experiment are shown in Figure 41. As expected, poor infection of (human) HEK 293T cells was achieved with the ecotropic

packaging system (pEQ) due to an absence of expression of the mouse ecotropic receptor on these cells. In Figure 41 it can be observed that the ecotropic packaging system produced a higher titre than the amphotropic packaging system when titred on NIH 3T3 cells. It was calculated that 8.5×10^5 transducing units per mL of supernatant was obtained for the ecotropic packaged virus compared to 5.4×10^4 for the amphotropic packaged virus using NIH 3T3 cells (Figure 41). However, the limiting dilution series used may not have been diluted sufficiently to determine the maximum viral titre of ecotropic system on the NIH 3T3 cells because at the highest dilution used the data had not yet reached plateau. The results also suggest that NIH 3T3 cells were more infectable than the HEK 293T cells based on the higher infection obtained using the VSV-G retroviral supernatant on both of these cell lines (5.4×10^4 versus 2.7×10^4 transducing units per mL respectively) (Figure 41). However, the observed increase in infection could also be explained by a greater proliferation of NIH 3T3 cells compared to HEK 293T cells as the progeny of an infected cell will also be GFP-positive.

In summary, the ecotropic packaging system (pEQ) produced a superior retroviral titre compared to the VSV-G packaging system when examined on NIH 3T3 cells. Consequently, the ecotropic packaging system was the system chosen for optimisation experiments performed on FDB-1 cells described in the next section. The superior titre produced by the ecotropic packaging system was an interesting finding considering that both packaging systems infect FDB-1 cells in the co-culture system to similar levels (see Sections 5.2.2.4 and 5.2.2.5.1). This may suggest that while the ecotropic packaging system produces more virus particles (as determined by the higher transducing units per mL), the amphotropic viruses may be more efficient at gaining entry into the FDB-1 cells and would therefore require fewer viral particles per mL to obtain the same level of infection. Confirmation of this would require measurement of the relative number of viruses required by the ecotropic (pEQ) and the amphotropic (VSV-G) packaging systems to infect FDB-1 cells.

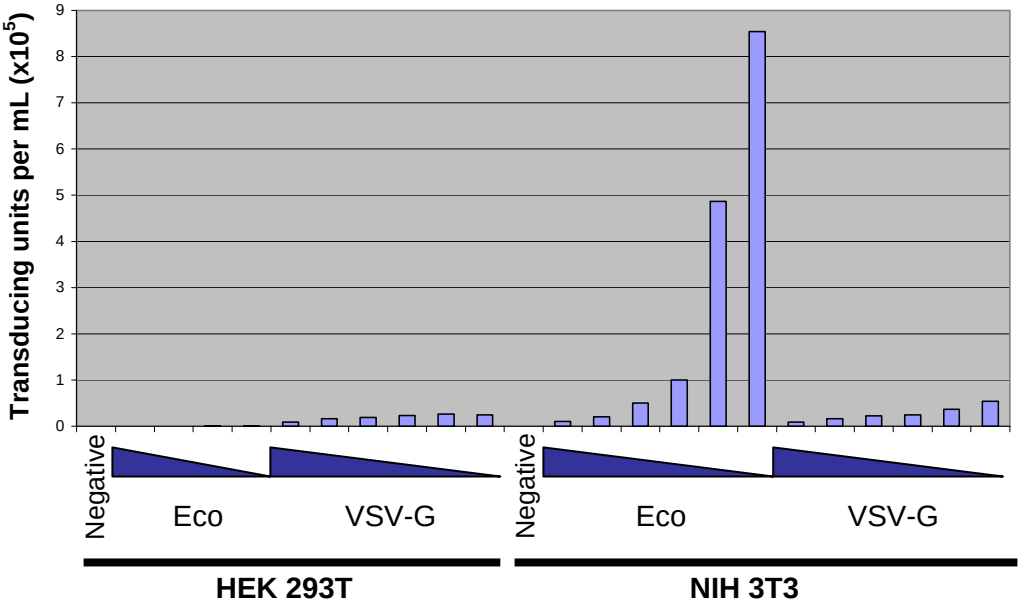


Figure 41: Titering virus production from ecotropic and VSV-G packaging systems. HEK 293T cells were transiently transfected with either pEQ (Eco) or pVPack-VSV-G (VSV-G) and pVPack-GP plasmids together with the MSCV-IRES-GFP plasmid (as a reporter plasmid) to produce retroviral supernatants. NIH 3T3 and HEK 293T cells were infected with retrovirus containing supernatant for 24 hours and then this media was replaced with fresh media for a further 48h before harvesting and assaying for GFP expression via flow cytometry. The cell line used and the dilution of the retroviral supernatant used for infection is stated on the x-axis with the calculated number of virus particles per mL stated on the y-axis as transducing units per mL. “Negative” is either HEK 293T or NIH 3T3 cells cultured without any virus supernatant added. The dilution series used for the pEQ on HEK 293T cells was 1:1, 1:2, 1:5 and 1:10 and for all the other treatments 1:1, 1:2, 1:5, 1:10, 1:50, 1:100 was used.

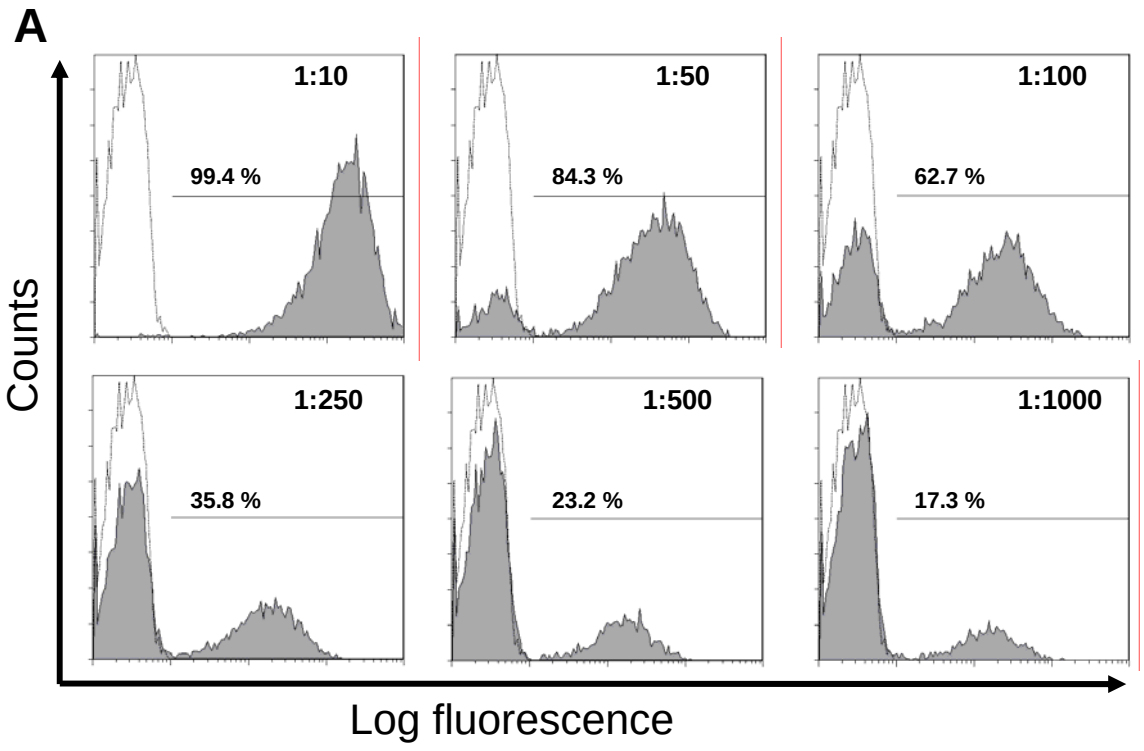
5.2.2.5.3 Comparing co-culture and supernatant infection methods for FDB-1 cell infection

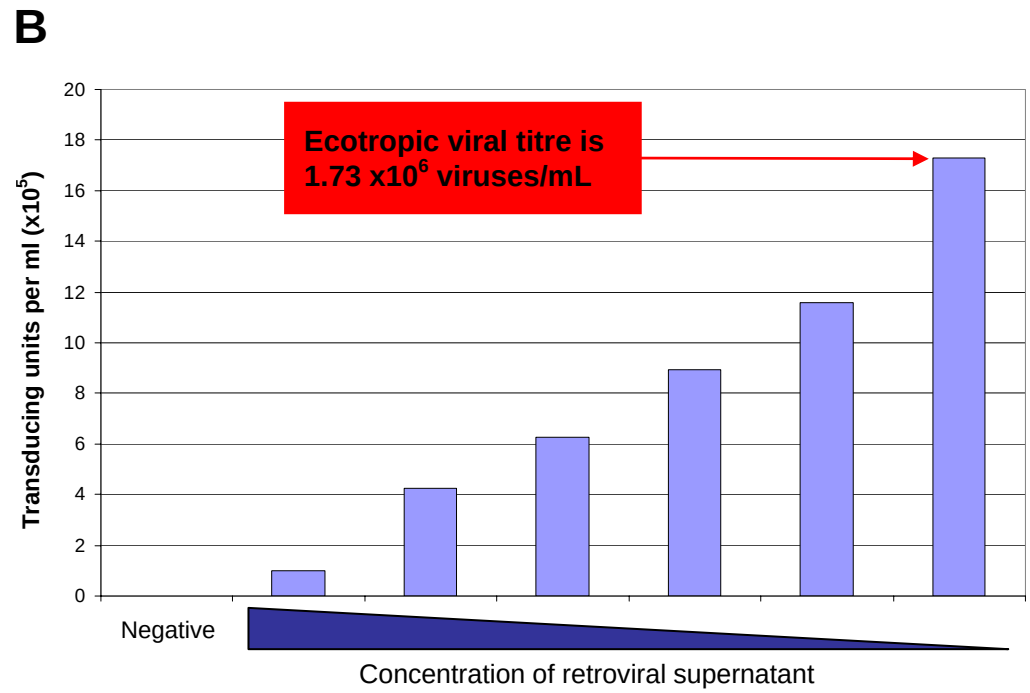
The question now arises as to why the ecotropic packaging system was producing significant levels of retrovirus, while only achieving very low levels of infection of the FDB-1 cells in co-culture with retroviral producing HEK 293T cells (1% data not shown). With a view to improving infection levels it was examined whether the FDB-1 cells could be infected directly using a retroviral supernatant (without co-culturing the cells with HEK 293T cells) as described in the NIH 3T3 titering experiment above. It was suspected that the co-culture of HEK 293T cells and FDB-1 cells was leading to a slower proliferation of the FDB-1 cells and in turn a reduced level of infection. The direct infection of FDB-1 cells with retrovirus containing supernatant would overcome any inhibitory effect that the HEK 293T cells may have on the growth of FDB-1 cells in co-culture. An experiment was designed to directly compare the performance of the two methods of infecting FDB-1 cells – co-culture and direct supernatant infection. This consisted of parallel transfections of HEK 293T cells where one population was used to produce retroviral supernatant for direct infection of FDB-1 cells, and the second population of HEK 293T cells were setup in co-culture with FDB-1 cells. The viral titre of the retroviral supernatant used to directly infect the FDB-1 cells was determined to be 1.7×10^6 viruses per mL based on infection of NIH 3T3 cells (Figure 42). FDB-1 cells were assayed under co-culture conditions with HEK 293T cells or cultured with the retroviral supernatant and the number of GFP-positive cells following infection was determined by flow cytometry. Co-culture resulted in a low level of FDB-1 infection (0.35% GFP-positive cells, see Figure 43), comparable to infection levels obtained previously. Direct infection of FDB-1 cells with retroviral supernatant showed a vastly increased percentage of infection (20.8%, see Figure 43). The superior infection obtained with supernatant infection method demonstrated that it is possible to infect FDB-1 cells via the use of a retroviral supernatant and that this is a significant improvement to the co-culture method.

In conclusion, the co-culture of FDB-1 cells with HEK 293T cells results in a very low percentage infection of FDB-1 cells, while the addition of retrovirus containing supernatant to FDB-1 cultures results in increased infection of the FDB-1 cells by at least 50 fold (20.8% GFP-positive cells via supernatant infection compared to 0.35% via the co-culture infection). The direct addition of retroviral supernatant

provides a suitable method for infecting FDB-1 cells with retroviral constructs in order to perform functional experiments. This modified protocol would greatly increase the number of FDB-1 cells infected with a retroviral vector above the levels seen previously using the co-culture method. FDB-1 cells could then be obtained over-expressing any gene of interest in sufficient numbers to permit FACS and their immediate use in functional assays (e.g. the proliferation assays reported for the FDC-P1 cells infected with BTG1 in Section 5.2.2.3) without requiring further expansion of the cell population. From such studies it was reasoned that insights into the role of particular genes in the processes of FDB-1 cell proliferation and differentiation could be obtained.

Figure 42: Titering of the retroviral stock used for FDB-1 infection. The titre of the viral supernatant produced from transient transfection of HEK 293T cells with the ecotropic (pEQ) virus vector and the MSCV-IRES-GFP vector on NIH 3T3 cells was determined. Virus containing supernatant was serially diluted and added to the media of NIH 3T3 cells for 24 hours. Following the infection, media was replaced with fresh media and the cells were given 48 hours recovery (to allow for expression of GFP) before being harvested and analysed via flow cytometry. In (A) the panels show the flow cytometry profiles of the infected NIH 3T3 cells (grey) compared to uninfected cells (white) obtained with the serial diluted viral stock, with the percentage GFP-positive cells stated. In each panel in (A), the x-axis scale for the major graduations is 10^0, 10^1, 10^2 and 10^3. From the flow cytometry data, the number of GFP-positive cells was determined and used to derive the number of infected cells which is plotted in (B) as transducing units per mL for each dilution tested. The number of transducing units per ml gives an indication of the minimal viral titre, which from the highest dilution of stock supernatant (1:1000) gave 1.73 x10^6 viruses per mL of supernatant.


A
1:10
99.4 %
1:50
84.3 %
1:100
62.7 %
1:250
35.8 %
1:500
23.2 %
1:1000
17.3 %
Counts
Log fluorescence


B
Ecotropic viral titre is
1.73 x10^6 viruses/mL
Transducing units per ml (x10^5)
20
18
16
14
12
10
8
6
4
2
0
Negative
Concentration of retroviral supernatant

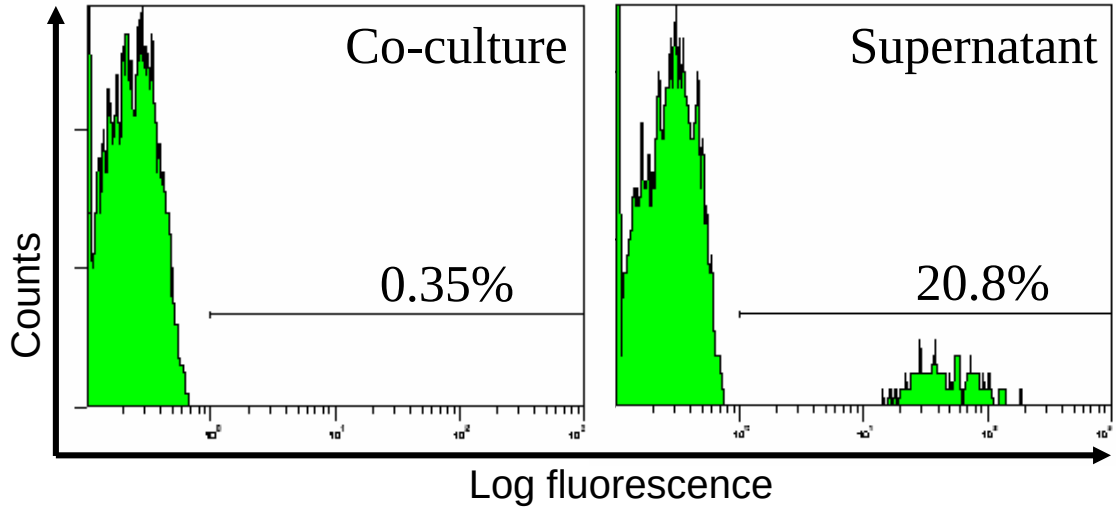


Figure 43: Comparison between co-culture versus direct supernatant infection methods. HEK 293T cells were transfected with the ecotropic (pEQ) viral vector and the MSVC-IRES-GFP vector and used to infect FDB-1 cells via either direct contact with the HEK 293T cells ("co-culture") or via culture of the FDB-1 cells with the retrovirus containing supernatant ("supernatant") collected from transfected HEK 293T cells. FDB-1 cells were infected via the two methods and then harvested and assessed via flow cytometry to determine the percentage of GFP-positive cells. This allowed direct comparisons between the two methods of infecting of FDB-1 cells – co-culture and supernatant infection – to be performed. This indicated that the supernatant infection method gave a superior infection level of FDB-1 cells in this experiment.

5.3 Discussion

5.3.1 The use of Northern blots to confirm differential expression

The interpretation of the microarray results is subject to two considerations which would be relevant to every expression profiling study. The first is whether the magnitude of the differential expression observed in the microarray is likely to be indicative of the actual changes in gene expression, and the second is whether the differentially expressed genes identified are significant. To identify significant genes a statistical analysis of the microarray data was performed (Chapter 4). It was also observed that on the microarray there were several clones (or cDNA sequences) that corresponded to the same gene, and that these clones had similar gene expression levels to one another, indicating reproducibility and consistency in the data (see Table 13 and Figure 26). Furthermore, 13 of the genes identified as statistically significant by the microarray analysis were examined using Northern blots to confirm their differential expression in FIΔ and V449E expressing FDB-1 cell lines (see Figure 29). The Northern blots were also used to determine the magnitude of the gene expression differences. The comparisons between the Northern blots and the microarray data were the focus of the work described in this chapter and are discussed below.

The initial Northern blots confirmed that 11 out of 13 genes tested from a list of specific genes of interest identified from the microarray analysis (defined in Chapter 4, Section 4.3.2.3 and Table 17) displayed differential expression between FIΔ and V449E expressing FDB-1 cell lines. *Sfrs3* and *Bnip3* showed less than a 1.4 fold difference between FIΔ and V449E at either time 0 or 24 hours on the Northern blots. This compares to results from the microarray, which showed that *Sfrs3* and *Bnip3* were 1.3 and 1.8 fold down-regulated in FIΔ compared to V449E at 24 hours respectively (see Table 18). With the exception of only a few genes (such as *Ltf* and *S100a9*) the fold changes seen on the microarray or Northern blots was in the range 1.5 - 2 fold. Nevertheless, such small differences can be important as it has been reported that a small change in the expression level of key genes (like transcription factors) can have dramatic consequences on the expression of other genes and the properties of the cells. An example of this was reported by Rosenbauer *et al*., (2004), where it was found that a reduction in expression of the lineage-specific transcription factor, PU.1, to 20% of normal levels in mice caused AML, unlike a complete or 50% loss in PU.1 expression.

Hence, genes with a relatively small change in gene expression are worthy of further investigation to examine whether the change alters the protein level in the cell, or causes any functional properties of the cell to change, such as altered an proliferation rate or differentiation.

It has been reported that microarrays frequently underestimate the fold change (Taniguchi *et al.,* 2001) however, the genes analysed here via Northern blots generally had fold changes of similar magnitude as that seen on the microarray (Table 18). There were some exceptions, including the gene *Cd34* for which the 24 hour difference seen between FIΔ and V449E on the microarray was not confirmed on either the initial Northern blot or time course analysis. Another exception was the *Cd24* gene where the large difference (4.9 fold) between FIΔ and V449E in mIL-3 (time 0) on the time course analysis was not observed on either the initial Northern blot or from the microarray analysis. Yet the difference in *Cd24* expression between FIΔ and V449E at 24 hours was comparable between the microarray and the Northern blots (initial blots and time course analysis). Some of the discrepancies between the Northern blots and microarray may reflect variation between culture conditions, suggesting that some genes are regulated by subtle changes in growth conditions that occurred between experiments.

A striking feature of a number of the genes studied was that they showed a significant difference at time 0 (i.e. in the presence of mIL-3) between FIΔ and V449E (see Sections 5.2.1.2.1 and 5.2.1.2.2 and Chapter 4). This was a surprising finding given that there is little morphological difference between the FIΔ and V449E expressing FDB-1 cells in mIL-3, and as the cells show no differences in proliferation rate or survival in presence of mIL-3. The differential expression of several genes at time 0 suggests that either the differentiation machinery is already activated in FIΔ cells in the presence of mIL-3, and/or alternatively V449E activates an altered gene profile. A suggestion that the former is correct was hinted at from studies examining CD24 protein levels (discussed in Section 5.3.2). The studies examining CD24 protein levels revealed that expression of FIΔ was associated with increased levels of CD24 protein expression at the cell surface compared to both V449E and parental FDB-1 cells grown in the presence of mIL-3. This is consistent with the differentiation machinery being already active in FIΔ expressing FDB-1 cells.

5.3.2 CD24 protein expression at the cell surface of FDB-1 cells

The one concern frequently raised from microarray analysis is that small changes at the mRNA level may not be reflected at the protein level. To assess the correlation between gene expression changes and levels of protein expression, the expression of one of the selected genes, *Cd24*, at the cell surface was investigated via flow cytometry. Although this technique does not measure total protein levels in the cell, it was reasoned that this method would be appropriate to investigate expression of this protein, given that CD24 is a GPI-anchored cell surface receptor (Pierres *et al.*, 1987; Alterman *et al.*, 1990; Kay *et al.*, 1990; Kay *et al.*, 1991) and that its surface expression has been found to correlate with mRNA levels (Wenger *et al.*, 1991). It was observed that CD24 protein expression is increased in FIΔ cells compared to V449E cells or parental FDB-1 cells when grown in mIL-3, and that when the parental cells are grown in mGM-CSF there is also increased expression of CD24 protein. It was thus concluded that FDB-1 cell differentiation induced by both FIΔ expression and growth in mGM-CSF induces the gene expression of *Cd24* which leads to increased levels of CD24 protein. While the specific role of *Cd24* in the process of FDB-1 cell differentiation is unclear, there is emerging evidence in the literature that this gene has roles in cell signalling, adhesion and differentiation (discussed below) (Alterman *et al.*, 1990; Kay *et al.*, 1991; Kadmon *et al.*, 1992; Wenger *et al.*, 1993; Aigner *et al.*, 1995; Aigner *et al.*, 1997). Further functional experiments are required to elucidate the role of this gene/protein in FDB-1 cell differentiation. In particular, it would be of interest to determine whether over-expression of Cd24 is able to reduce proliferation or induce differentiation of parental FDB-1 cells in mIL-3.

Investigation of CD24 function in myeloid cells may shed light on the significance of CD24 expression in a range of cancers and how its expression might be advantageous to tumour cells. In hepatocellular carcinoma, *CD24* is a potential early tumour marker gene that correlates with p53 mutation (Huang and Hsu, 1995). In human tumour cell lines, SW2 (small cell lung carcinoma) and K562 (CML), CD24 has also been found to associate with the protein tyrosine kinases c-fgr and lyn (Zarn *et al.*, 1996). The ability of CD24 to bind to P-selectin enables tumour cells or myeloid cells expressing CD24 to bind to endothelial cells and platelets (Aigner *et al.*, 1995; Aigner *et al.*, 1997). The role of both the CD24/P-selectin binding pathway as well as the c-src kinase signalling pathway in promoting cancer cell survival or tumour metastasis would justify further investigation of this protein. Furthermore, given that lyn associates with

GM-CSF receptor alpha subunit (Michelle Peters, unpublished observation) and that this subunit is required for the factor independent growth of FIΔ, suggests that lyn, CD24, together with both the alpha and beta subunits of the GMCSF receptor, may all be involved in a multi-receptor signalling complex that may have a specific role in the differentiation of myeloid cells that express these receptors and proteins.

5.3.3 Functional studies with the genes of interest

Four candidate genes (*Btg1*, *Cd24*, *Klf6* and *Nucleostemin*) were cloned into the retroviral vector MSCV-IRES-GFP to permit functional characterisation studies. The selection of these genes was based on their differential expression profile from the microarray and Northern blots together with their known function(s) (see Section 5.1.2).

5.3.3.1 Functional experiments performed with BTG1 in FDC-P1 cells

Infection of FDC-P1 cells with a retroviral vector containing the *Btg1* gene (MSCV-BTG1-IRES-GFP, see Figure 37) was performed to examine the consequences of over-expression of this protein on the proliferation of myeloid cells. Over-expression of BTG showed no effect on the proliferation of FDC-P1 cells (Figure 38 and Figure 40), as might have been expected from the anti-proliferative role associated with this protein. It is possible that the negative effects on proliferation of this protein are not observed in FDC-P1 cells, and perhaps would only be evident in primary myeloid cells. This would agree with the findings from Bakker *et al*., (2004) who reported a reduced expansion of myeloid progenitor cells in colony assays caused by the over-expression of BTG1.

A recent study by Bakker *et al*., (2004) showed that BTG1 over-expression only marginally affected the proliferation of infected myeloid progenitors, while a profound negative effect on proliferation was observed for erythroid progenitors. These findings were achieved by transducing primary murine bone marrow with a retroviral expression vector containing BTG1 and scoring the number of colonies that were obtained from serum free semi-solid media supplemented with either GM-CSF (giving myeloid colonies), or a cocktail of Epo, SCF and dexamethasone (producing exclusively erythroid colonies). While the expression of BTG1 reduced the size and number of erythroid colonies, there was only a slight reduction in the number of myeloid colonies (Bakker *et al.*, 2004). Taken together, the findings from Bakker *et al*., (2004) and the results from this chapter suggest that BTG1 has only a minor effect on the proliferation of myeloid cells.

While over-expression of BTG1 may not effect proliferation of FDC-P1 cells, there is the chance that additional mutations or consequences of the expression of the mutant hβc in FDB-1 cells may render this protein to be a potent suppressor of proliferation in FIΔ but not V449E expressing FDB-1 cells. Given that FDC-P1 cells are unable to differentiate, it would warrant investigation of the effect of BTG1 over-expression on the differentiation of FDB-1 cells.

The work of Bakker *et al*., (2004) demonstrates the importance of BTG1 in erythrocyte differentiation and along with Cho *et al*., (2004) suggests that BTG1 is involved in myeloid cell differentiation. BTG2, a homologous family member is suppressed by SCF (similar to BTG1) and induced during differentiation of erythroid progenitors (Kolbus *et al.,* 2003). A recent microarray study, similar to the one presented in this thesis, has found that the gene *Btg2* is also up-regulated with GM-CSF induced differentiation of FDB-1 cells and with FIΔ expression in FDB-1 cells (Brown *et al.,* 2006). Given that BTG2, a closely related protein to BTG1, is also up-regulated in neuronal differentiation (Bradbury *et al.,* 1991; Corrente *et al.*, 2002) and shown to be involved in germ cell and muscle cell differentiation (for a review see Tirone, 2001) it points to a general role for BTG proteins in cell differentiation. Under-expression studies involving BTG1 in myeloid cells would provide more conclusive evidence of its role in myeloid cells e.g. by examining its role in the differentiation of primary myeloid cells. However, it is highly likely that BTG2 could substitute for the effects of BTG1 given that it is also expressed in myeloid cells (Lim *et al.,* 1995). This compensation by BTG2 may preclude seeing a direct effect of under-expression of BTG1 via siRNA directed against BTG1, or from a transgenic gene knockout animal.

5.3.3.2 Optimising the retroviral infection of FDB-1 cells

In order to perform functional experiments in FDB-1 cells an efficient retroviral delivery method was required. Via optimisation of an existing protocol in the laboratory and separating the stages of virus production and viral infection, the conditions were improved such that a 20% infection of FDB-1 cells was obtained (Figure 43). The major change to the protocol was associated with the collection of a high titre supernatant for infection of the FDB-1 cells. The addition of the viral supernatant to the culture medium of the FDB-1 cells proved much more efficient than direct contact with the virus producer cells in co-culture, which gave infection levels of 1-2%. The superiority of retroviral supernatant when used alone is likely to be due to increased growth rate (and

hence mitotic index) of the FDB-1 cells in the absence of HEK 293T cells. This is because the cells must be cycling for integration of the virally encoded genes to occur (Van Parijs *et al.*, 1999). The increased growth rate leads to increased uptake of the virus and incorporation into the FDB-1 cell genome resulting in expression of the retroviral vector containing genes (e.g. GFP). The increased level of infection enables sufficiently large numbers of infected FDB-1 cells to be isolated so that the original functional experiments planned with the four genes of interest can now be performed.

The optimised retroviral infection procedure used the ecotropic packaging vector pEQ rather than the amphotropic packaging plasmids pVPack-VSV-G and pVPack-GP (referred to as "VSV-G"). The ecotropic packaging system was chosen as it gave a higher retroviral titre than the amphotropic packaging system when tested on NIH 3T3 cells (8.5×10^5 versus 5.4×10^4 viruses/mL, see Figure 41). It is stated that up to 1×10^7 viruses/mL is obtainable with the VSV-G retroviral packaging system and that a good viral titre is considered greater than 1×10^6 viruses/mL (Hopkins, 1993; Clontech, 2004). However, such levels were not observed under this study's culture conditions. Virus containing supernatant produced from the pEQ packaging plasmid is able to be stored in the fridge but cannot be concentrated by high speed centrifugation. In contrast, pseudotyped VSV-G virus particles are able to be concentrated by ultra-centrifugation and stored as a frozen concentrate that can be thawed when required (Yee *et al.*, 1994a; Yee *et al.*, 1994b). This concentration process would enable the retroviral supernatant prepared from pVPack-VSV-G transfected cells to be concentrated to much higher levels (between 100 and 300 fold or to titres $>10^9$ cfu/mL (Burns *et al.*, 1993; Yee *et al.*, 1994a; Yee *et al.*, 1994b). Viral titres in the order of 10^9 viruses/mL may justify the use of this concentrated supernatant to achieve even higher infection levels of FDB-1 cells. As an approach to further optimising this procedure it would be interesting to compare the infection levels obtained with ecotropic (pEQ) supernatant and concentrated amphotropic (VSV-G) supernatant.

5.3.3.3 Problems with retroviral infection and future directions

The difficulties infecting FDB-1 cells for the functional experiments were resolved by optimising viral production and the infection procedure (as discussed above). Another potential problem with the retroviral over-expression approach is that if the over-expressed gene negatively affects growth (i.e. causes differentiation of the cells) then it can be difficult to isolate and expand a population of infected cells. This could be the case for the genes Btg1 and Cd24 which increased with FIΔ differentiation (see Figure

31, Figure 32, Figure 33 and Figure 34). On the other hand, cells expressing genes that positively effect growth (suppress differentiation or promote growth) can be expanded readily from the infected cell population. One possible solution for this problem would be to use an inducible system (such as those based on the tetracycline (Tet)-mediated expression systems – see Baron and Bujard, 2000 for review) in which the gene containing plasmid could be selected for prior to inducing the expression of the target gene in the cells of interest. The use of an inducible expression system would allow expansion of a cell population prior to the induction of an over-expressed gene that may negatively affect growth or cause differentiation. An additional approach for the investigation of gene function is offered by small interfering RNA (siRNA), which would permit studies examining the effect of these genes when under-expressed in cells or tissues (for review see Tuschl and Borkhardt, 2002, Mello and Conte, 2004 and Hannon and Rossi, 2004). These approaches are discussed further in Chapter 6, Section 6.1.4.

5.3.4 Summary

In the results described in this chapter, the differential expression of eleven genes from the microarray analysis were confirmed via the use of Northern blots. Time course analysis was used to show gene expression differences between FIΔ expressing FDB-1 cells and V449E expressing FDB-1 cells over 72 hours for *Btg1*, *Hipk1*, *Cd24*, *Cd34*, *Klf6* and *Nucleostemin*. Flow cytometry was used to investigate whether increased mRNA levels of *Cd24* would be reflected by increased CD24 protein expressed at the cell surface. Increased expression of CD24 protein was shown to be induced by FIΔ expression and mGM-CSF induced differentiation of FDB-1 cells, in line with elevated levels of *Cd24* mRNA in FIΔ cells over time. Preliminary functional studies with Btg1 in FDC-P1 cells showed no significant effect on proliferation due to over-expression of Btg1.A retroviral infection procedure for FDB-1 cells was optimised making it possible to investigate the functional role of Btg1, Cd24, Klf6 and *Nucleostemin* in the growth, differentiation and survival of FDB-1 cells using a number of standard growth, survival and differentiation assays.

Chapter 6

Concluding Discussion and Future Perspectives

6.1 Overview

The cytokines IL-3, IL-5 and GM-CSF affect the survival, proliferation, differentiation and functional activation of cells within the haematopoietic system (Metcalf and Nicola, 1995; Martinez-Moczygemba and Huston, 2003). Ligand specific α subunits in conjunction with hβc comprise the receptors for human IL-3, IL-5 and GM-CSF. A series of activating mutations in hβc have been isolated based on their ability to induce cytokine independent proliferation of FDC-P1 cells (D'Andrea *et al.*, 1994; Jenkins *et al.*, 1995; D'Andrea *et al.*, 1996; Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000). The mutants can be classified based on the transmembrane or extracellular position of the mutation. Signalling and biochemical differences exist between the extracellular hβc mutants (I374N and FIΔ) and the transmembrane mutant V449E (reviewed in D'Andrea and Gonda, 2000) (see Introduction, Section 1.5.3.3). These differences are predicted to result in the different biological effects of the two classes of hβc mutants seen *in vitro* and *in vivo* (Jenkins *et al.*, 1995; McCormack and Gonda, 1997; D'Andrea *et al.*, 1998; McCormack and Gonda, 1999) (see Sections 1.5.3.2 and 1.5.3.4). The identification of genes that were differentially expressed between myeloid cell lines expressing the transmembrane mutant, V449E, and one of the extracellular mutants, I374N or FIΔ, was predicted to provide insights into the events involved in hβc signalling that are associated with myeloid proliferation, differentiation and the leukaemogenic properties of the V449E mutant.

6.2 Summary of data

The primary aim of this work was to identify genes that were differentially expressed in cell lines expressing either a transmembrane hβc mutant or an extracellular hβc mutant. To achieve this, two screening approaches were used, the first involved using membrane arrays and the second used microarrays to examine the expression of thousands of genes at once and identify differentially expressed genes between the samples.

In the first experiment a membrane array analysis was performed to identify the gene expression differences between the extracellular mutant, I374N, and the transmembrane mutant, V449E, expressed in FDC-P1 cells (see Chapter 3). This screen was predicted to reveal gene expression changes associated with the different signalling and biological properties of the activated hβc mutants. Using a novel approach consisting of multiple image analysis programs and statistical methods, several gene lists of differentially expressed genes were identified. *Ptpmt1* was confirmed to be differentially expressed between V449E and I374N using Northern blots, which suggested that changes in the level of this phosphatase may be involved in the different signalling properties of these two hβc mutants. The Northern blots revealed increased expression of *Ptpmt1* in V449E expressing FDC-P1 cells, compared to FDC-P1 cells expressing hGM-CSFα and hβc (αβ) grown in the presence or absence of hGM-CSF and I374N expressing FDC-P1 cells. Northern blots also revealed *Nab1* and *Ddx26b* to be regulated in response to hGM-CSF in FDC-P1 αβ cells, however, differential regulation between V449E and I374N was not confirmed. Stimulation of FDC-P1 αβ cells with hGM-CSF led to an up-regulation of *Ptpmt1* and *Nab1* mRNA, and a down-regulation of *Ddx26b* mRNA. The differential expression of these three genes in response to hGM-CSF in FDC-P1 cells indicates they may be downstream targets of GM-CSF signalling through hβc. It was expected that more of the genes identified from the membrane array analysis would be confirmed to be differentially expressed between V449E and I374N by Northern blots, given that these two cell lines show biochemical and signalling differences (Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000). The confirmation of only one gene from the 15 examined by Northern blots suggests that many of the genes identified as differentially expression from the membrane array data are false positives. The results may indicate that membrane array method was not sensitive or robust enough to detect the differentially expressed genes. Alternatively, the large no of false positives may be due to the choice of cell line used. In the FDC-P1 cells, there are no morphological differences or alterations in their rate of proliferation induced by these two types of hβc mutants. It has been shown however, that these hβc mutants differ in the requirement for the GM-CSF alpha subunit in these cells and in the level of phosphorylation of the hβc and the adapter protein SHC (Jenkins *et al.*, 1998; D'Andrea and Gonda, 2000, T Blake personal communication). Perhaps, despite upstream differences between these mutants, it may be that these two classes of hβc

mutants do not induce differential changes in gene expression within the FDC-P1 cell line.

The second screening approach used a microarray analysis to identify gene expression differences between the extracellular mutant, FIΔ, and the transmembrane mutant, V449E, expressed in FDB-1 cells (see Chapter 4). FDB-1 cells were chosen as they have the ability to proliferate in response to mIL-3 and differentiate in response to mGM-CSF and most importantly, respond differently to the two hβc mutants. Thus, the use of the FDB-1 cell line enabled the examination of genes involved in hβc signalling that are associated with the alternative biological processes of proliferation and differentiation. The microarray experiment was performed as a 2×2 factorial design, which was superior to the commonly used reference design (Glonek and Solomon, 2004). This design allowed the determination of genes that were differentially expressed over time, or differentially expressed between samples. The analysis focused on those genes that showed differential expression over time (24 hours) and between the samples, which was defined as the interaction parameter. A comprehensive approach was adopted for the analysis of the microarray data, which was acquired from two separate scans of each microarray slide, using two measures of gene intensity. The data was analysed using three statistical tests (T-test, B-stat and F-stat) to identify significant gene expression changes. This novel and robust analysis method enabled the identification of genes that performed consistently well across all approaches, or those genes that ranked highly based on several different approaches. The use of multiple scans of a microarray slide to improve the detection of differentially expressed genes has since been reported by Romualdi *et al*., (2003). Romualdi *et al*., (2003) have developed a software program to specifically manage multiple scan array images, which they show minimises scanning artefacts and provides increased confidence in the differentially expressed genes observed. An extension on the idea of using multiple scans would be to perform both a low and high intensity scan of each slide as suggested by Lyng *et al*., (2004). Scanning each slide at high and low intensities would enable the integration of data obtained from both high and low abundance genes found on the microarray, which may otherwise be overlooked using just a single moderate intensity scan.

The outcome of the microarray analysis of FDB-1 cells expressing the activated hβc mutants was a list of 34 differentially expressed genes stated in Table 17. Thirteen

of the genes listed in Table 8 were examined by Northern blotting to confirm the differential expression these genes identified as differentially expressed from the microarray analysis. Northern blots confirmed the differential expression of 11 of 13 genes between V449E and FIΔ expressing FDB-1 cells over 24 hours (see Section 5.2.1.2.1). Several genes showed differential expression on Northern blots ranging from 2-8 fold at 24 hours when comparing the gene expression in FIΔ cells over the gene expression in V449E cells. The Northern blot results provided evidence that the microarray technique and the statistical analysis approaches adopted were capable of identifying differentially expressed genes in FDB-1 cells. Furthermore, time course mRNA expression profiles of six of the regulated genes were examined using Northern blots to reveal the expression pattern of these genes over 72 hours. The time course analysis demonstrated the differential expression of these genes at times prior to, and after, 24 hours (see Chapter 5, Section 5.2.1.2.2). This expression analysis, together with the literature that was available at the time, allowed selection of these genes for further studies.

6.2.1 Functional experiments using retroviral infection of FDC-P1 and FDB-1 cell lines

The secondary aim of this thesis was to functionally characterise genes that were identified to be differentially expressed between the extracellular and transmembrane hβc mutants. Four genes (*Btg1*, *Cd24*, *Klf6* and *Nucleostemin*) were cloned into the MSCV-IRES-GFP retroviral expression vector to enable over-expression studies by retroviral infection of target myeloid cell lines to be performed. The protein motifs and known function of these four genes reported in other systems suggested roles in proliferation, differentiation or in the regulation of gene expression (see Sections 5.1.2.1 and 5.1.2.2), and were thus of interest to the study presented in this thesis. Initial functional studies were performed in the FDC-P1 cell line, prior to optimisation of the FDB-1 system for over-expression analysis. The results of the preliminary functional studies in FDC-P1 and FDB-1 cells were predicted to provide important leads allowing prioritisation of genes for more sophisticated studies in primary cells.

The effect of the gene *Btg1* on the proliferation of cells was investigated by infection of FDC-P1 cells with MSCV-BTG1-IRES-GFP in two independent experiments. Over-expression of BTG1 (based on co-expression of GFP) did not

significantly alter the proliferation rate of FDC-P1 cells (Figure 38 and Figure 40). However, while the cells were shown to express GFP, the expression of BTG1 protein in these cells was not confirmed. Expression of BTG1 protein could now be confirmed by Western blotting approaches using commercially available antibodies. Results published by Bakker *et al*., (2004) revealed a small but significant effect on the proliferation of myeloid progenitor cells over-expressing BTG1. Furthermore, Bakker *et al*., (2004) showed that BTG1 expression had a profound negative effect on the proliferation of erythroid progenitors indicating that BTG1 negatively effects growth of erythroid cells. Cho *et al*., (2004) have suggested that BTG1 has a role in the differentiation of myeloid cells based on its regulated expression in primary AML cells. It is likely that FDC-P1 cells are unresponsive to the anti-proliferative effects of BTG1, perhaps as a consequence of accumulated genetic lesions. Therefore, further studies investigating the role of over-expression of BTG1 in FDB-1 cells and primary cells were of interest to reveal the functional consequences of increased *Btg1* mRNA levels seen in FIΔ expressing FDB-1 cells. (Bakker *et al.,* 2004)

Optimisation of FDB-1 infection was required as preliminary experiments performed to infect FDB-1 cells with MSCV-BTG1-IRES-GFP vector revealed that the FDB-1 cells were poorly infectable under standard co-culture conditions. The low level of retroviral infection of the FDB-1 cells was a major limitation to the use of these cells in functional assays. An additional problem with the infection of FDB-1 cells was that two of the four genes first chosen for investigation (*Btg1* and *Klf6*) were predicted to negatively effect the growth rate of the FDB-1 cells (i.e. cause differentiation or growth arrest) when over-expressed. With only small numbers of cells infected, it was not be possible to expand the initially infected population to sufficient numbers to enable an investigation of that gene's function in FDB-1 cells. This was likely to be an ongoing limitation, as several genes from the microarray study were also differentiation-associated (and likely to negatively affect the growth rate of cells) thereby making future studies by this method difficult.

To achieve higher infection levels of FDB-1 cells an optimisation of the retroviral infection process was performed. This led to the discovery that by separating the virus production steps from the infection of the FDB-1 cells, it was possible to increase the infection of FDB-1 cells from levels of 1-2% to levels up to 20% (see Chapter 5, Section 5.2.2.5). The optimised infection protocol would enable isolation of the infected FDB-1 cells in sufficient numbers to allow an examination of the effects of

over-expression of a gene of interest. The examination of the consequences of the over-expression of the four genes in FDB-1 cells would be the next logical step for further studies (see Section 6.4.2).

6.3 Comparison of the results with published microarray studies

6.3.1 Further microarray experiments

As mention earlier (see Section 4.1.3) the choice of time point used in the microarray study was an important issue, in order to identify genes that were a direct consequence of the differential signalling associated with the hβc mutants, rather than secondary events from the differentiation of the FIΔ expressing FDB-1 cells. The decision to use 24 hours was in line with other differentiation studies using similar cell lines e.g. FDCP-mix cells used by Bruno *et al*., (2004) (who examined 4, 8, 16, 24, 48, 72 and 168 hours), THP-1 cell line used by Tuomisto *et al*., (2005) (who studied 3, 12, 24 and 72 hours), MPRO cell line used by Lian *et al*., (2001) (who examined 24, 48 and 72 hours), HL-60 used by Zheng *et al*., (2002) (who studied 0, ¼, ½ , 1, 2, 3, 6, 12, 24, and 48 hours), NB4 cells used by Zheng *et al*., (2005) (who examined 6, 12, 24, 28 and 72 hours) and both NB4 and HL-60 cells used by Lee *et al*., (2002) (who studied 3, 8, 12, 24, and 48 hours). The examination of multiple time points in the microarray experiment reported in this thesis would have been possible with the use of additional slides. Examination of multiple time points should reveal co-ordinately regulated genes by comparison of their time course profiles. The identification of co-ordinately regulated genes by examining multiple time points has been demonstrated in a study by Brown *et al*., (2006), which is discussed below.

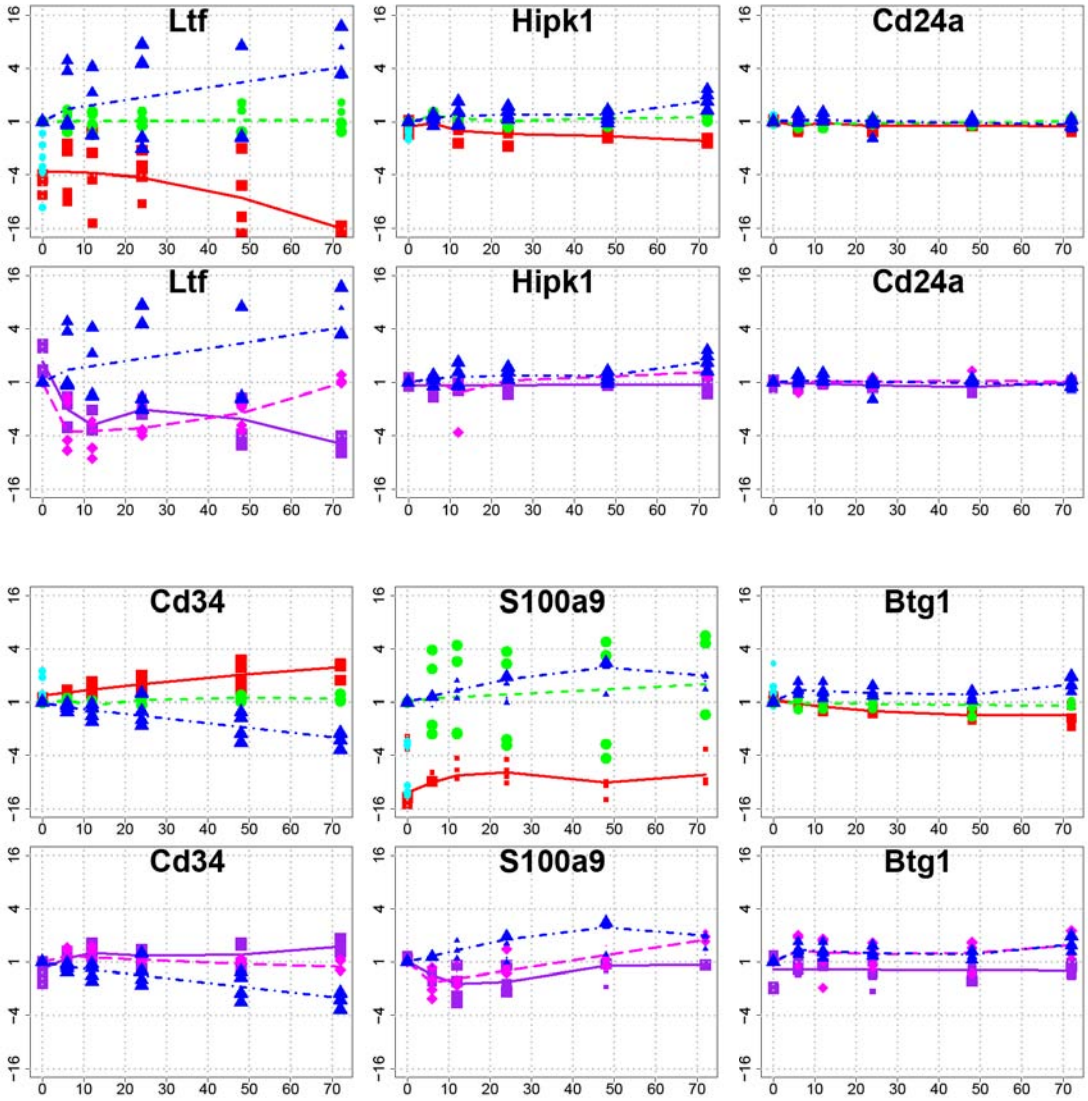
In parallel to the studies with the activated hβc mutants was the interest in examining the parental FDB-1 cells grown in the presence of either mIL-3 or mGM-CSF. The comparison of the proliferative and differentiation phenotypes of the FDB-1 cells grown in IL-3 and GM-CSF, respectively, to the phenotypes produced by V449E and FIΔ in the absence of these exogenously added growth factors was expected to identify overlapping gene-sets. By examining the time course profiles of genes found to be differentially expressed it was predicted that common gene expression changes that are associated with differentiation (FIΔ expression or FDB-1 cells +GM-CSF) or

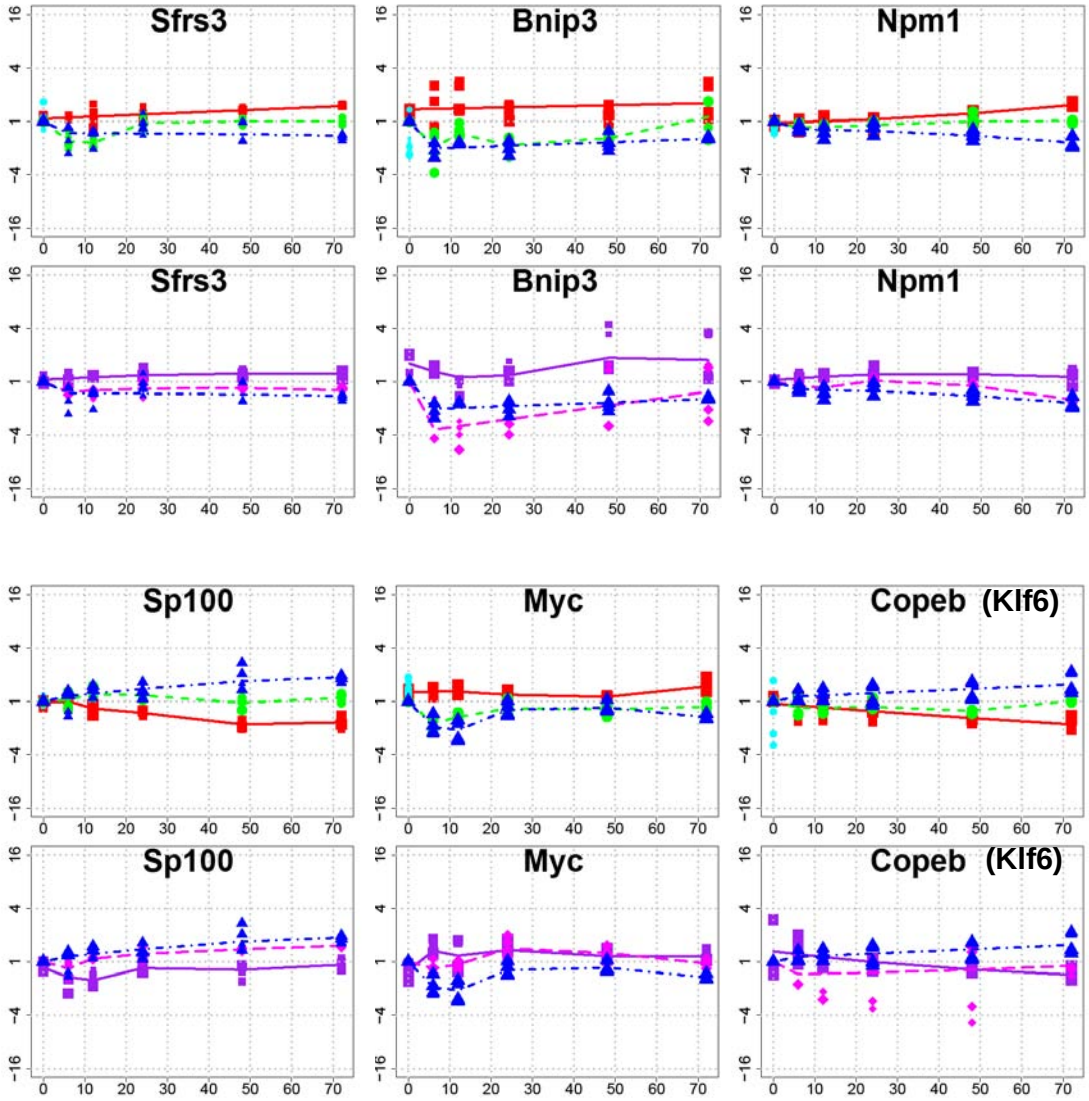
proliferation (V449E expression or FDB-1 cells +IL-3) would be identified. Recently the findings from a large microarray study performed in the laboratory of Associate Professor R. D'Andrea on the FDB-1 cell lines (parental +GM-CSF or IL-3, FIΔ and V449E) has been published (Brown *et al.*, 2006). This study by Brown *et al*., (2006) examined the gene expression changes associated with proliferation and differentiation of the FDB-1 cell lines at multiple time points over 72 hours. The study was successful in identifying genes that were co-ordinately regulated over time and potentially important regulators of proliferation, survival and differentiation. From the analysis of differentially expressed genes identified between the two proliferating populations (FDB-1 +IL3 or V449E), Brown *et al*., (2006) were able to identify a set of V449E-specific genes that were associated with the leukaemic V449E mutation. Interestingly, this list includes the genes for *S100a9*, *Ltf* and *Lcn2* which were also identified as differentially expressed in the microarray study reported in Chapter 4.

The microarrays used in this thesis were printed with the NIA 15K cDNA clone set, whereas the microarrays used by Brown *et al*., (2006) contained the CompuGen Mouse OligoLibrary consisting of 21587 unique genes. There were 8315 genes that were represented on both of the microarray slides, which included, 102 of the 177 genes identified in the F-stat list (Table 15) and 146 of the Top 200 genes (Table 12). 69 genes of the 102 overlapping genes in the F-stat gene list were identified as significantly differentially expressed in the study reported by Brown *et al*., (2006). Similarly, 99 genes of the 146 overlapping genes in the Top 200 gene list were also significant, which equates to 68% in both cases. Thus, each of the significant differentially expressed gene lists, identified in Chapter 4, reveal genes that were also found to be significant by Brown *et al*., (2006). Hence, there were a number of genes identified in the analysis presented in Chapter 4, which have also shown interesting time course profiles over 72 hours in the microarray analysis reported by Brown *et al*., (2006). For the 13 genes examined by Northern blots in Chapter 5 (Figure 30) their expression profiles in the microarray study of Brown *et al*., (2006) are shown Figure 44, with the exception of *Nucleostemin* as this gene was not present in the CompuGen Mouse Oligo Library. The expression profiles for the remaining 12 genes was similar in the time course microarray analysis of Brown *et al*., (2006) to the Northern results and microarray results reported in this thesis. The notable exception was the gene *Cd24* which showed no differential expression in any of the FDB-1 cell lines examined. The reason for this discrepancy at present remains unclear, but could be due to a failure of the oligo on the CompuGen

microarray slides to hybridise to its *Cd24* mRNA target. The expression profile of *S100a9* in Figure 44 revealed opposing fold changes in V449E expressing FDB-1 cells over the 72 hours in replicate samples analysed on the microarray. While the large fold change at 24 hours agrees with the 7.5 fold change seen in FIΔ over V449E in the microarray experiments performed in this thesis, given the variability of the V449E expression these results must be regarded cautiously, as the Northern blots (see Table 18) did not confirm these gene expression differences. The profiles shown in Figure 44 also revealed the expression of these twelve genes in parental FDB-1 induced to differentiate by treatment with GM-CSF. Many of the genes showed co-ordinated expression of genes induced by FIΔ or GM-CSF treatment, although in several genes there is a delay before expression of the gene increases in response to GM-CSF (e.g. *Ltf*). Examination of genes expressed in common between FIΔ and GM-CSF stimulation of FDB-1 cells, may reveal a common myeloid differentiation pathway. Ongoing studies by other researchers in the laboratory are examining selected genes from the microarray analysis, to reveal their role in myeloid cell proliferation and differentiation.

Figure 44: Expression profiles of genes studied by Brown *et al*., (2006) that were examined by Northern blots. The gene expression profiles of twelve genes examined by Northern blots shown in Chapter 5 were isolated from the experimental data published by Brown *et al*., (2006). Note: the profile of *Nucleostemin* could not be obtained as this gene was not contained in the CompuGen Mouse Oligo Library. In each plot, points representing the normalised microarray fold-change values (one point per array), and the lines representing weighted, smoothed spline estimates of the change over 72 hours. Upper plots: green circles and dashed line, V449E profile; blue triangles and a dot-dashed line, FIΔ profile; red squares and a solid line, V449E versus FIΔ direct comparisons; Cyan points, V449E compared with parental FDB-1 at time zero (i.e. in mIL3). Lower plots: blue triangles and a dot-dashed line, FIΔ; pink diamonds and long dashed line, GM-CSF; purple squares and a solid line, FIΔ versus GM-CSF direct comparisons.


Ltf
Hipk1
Cd24a
Ltf
Hipk1
Cd24a
Cd34
S100a9
Btg1
Cd34
S100a9
Btg1
16
4
1
−4
−16
0
10
20
30
40
50
60
70


Sfrs3
Bnip3
Npm1
Sfrs3
Bnip3
Npm1
Sp100
Myc
Copeb (Klf6)
Sp100
Myc
Copeb (Klf6)

6.3.1.1 Comparison of the microarray results with published studies of cell differentiation

A number of studies by other researchers have investigated the molecular profiles of various classes of haematopoietic cells using techniques including differential library construction, differential display, suppression PCR, membrane arrays, microarrays and oligo-based Affymetrix GeneChips. A number of the microarray studies (either cDNA or oligo arrays) have reported genes that are in common to the genes identified to be differential expressed in the microarray experiments presented in this thesis (see Table 21).

Akashi *et al*., (2003) examined different classes of prospectively isolated haematopoietic progenitors and found *Gata2* (up-regulated in V449E) as one of the representative genes that are predominantly expressed in haematopoietic stem cells (HSC) and down-regulated in multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs). G-CSFR (*Csf3r*) was found to be highly expressed in CMPs while lysozyme (*LYZS*) was under-expressed in CMPs and CLPs. *Lysz* was 3.9 fold up-regulated with FIΔ differentiation time (Chapter 4, Table 15e) while G-CSFR was 2.0 fold up-regulated with FIΔ differentiation (Chapter 4, Table 15a).

Bruno *et al*., (2004) used FDCP-mix cells, a multipotent haematopoietic cell line, differentiated over 4 hours to 7 days to identify *Ltf*, *Spi2-1*, *Hbb-y*, *Hbb-bh1*, *c-myc*, *Lyzs* (lysozyme), *Gata2*, *Csf3r* (G-CSF receptor), *Bnip3*, *S100a6* and *Daf1* as differentially expressed in common with the studies reported in this thesis. They found that *Bnip3* was one of the genes that was commonly down-regulated in all four differentiation pathways and this was found to be down-regulated upon FIΔ differentiation in this study (1.4-1.7 fold, see Table 12b, Table 13 and Table 15c). *Spi2-1* was found to be over-expressed after 16 hours of differentiation down the neutrophil pathway and under-expressed at 72 and 168 hours. In the study presented in this thesis, *Spi2-1 (Serpina3g)* decreased 3.9-4.3 fold upon FIΔ differentiation consistent with the later time points (see Table 12b, Table 13 and Table 15c).

Spi2-1 was shown by Hampson *et al*., (1997) to be expressed in undifferentiated FDCP-mix cells, but showed a marked down-regulation upon granulocyte macrophage differentiation in this cell line and in bone marrow derived bipotent granulocyte macrophage progenitor cells. Over-expression of *Spi2-1* caused a delay in maturation

and an increase in the clonogenic potential of the FDCP-mix cells (Hampson *et al.*, 1997).

Using the MPRO cell line as a model of myeloid differentiation, Lian *et al*., (2001) identified *Lcn2* as being up-regulated with retinoic acid induced granulocyte differentiation of MPRO after 24 hours and consistent with the observed regulation with FIΔ differentiation (3.3 fold, see Table 12a and Table 15b). *S100a9* (Calgranulin B) was found to be up-regulated late in differentiation (72 hours), *Plaur* (urokinase plasminogen activator receptor) was up-regulated between 2-3 fold increasing from 0–72 hours and *Ltf* showed a transient increase, peaking at 48 hours (Lian *et al.*, 2001). This is consistent with the studies present in this thesis (*S100a9*, *Plaur* and *Ltf* were up-regulated with FIΔ differentiation 2.6-2.8, 2.1-2.5 and 4.8-5.6 fold respectively - see Table 12a, Table 13 and Table 15b). Also consistent with the expression analysis in Chapter 4, Lian *et al*., (2001) found *Spi2-1* (*Serpina3g*), *Npm1*, and *Myc* to be down-regulated early (within 24 hours) of MPRO differentiation.

Tuomisto *et al*., (2005) used human monocytic THP-1 cells stimulated with PMA (phorbol 12-myristate 13-acetate) to induce differentiation into macrophages. They showed that both *ANXA9* and *SP100* increased expression over the differentiation time course (72 hours), while *PLAUR* showed an early peak in expression at 3 hours with increased expression at 24 and 72 hours.

As mentioned in Chapter 4 (Section 4.3.2.5), several genes displaying differential expression in the microarray analysis presented in this thesis are known targets of the granulocytic lineage transcription factor, C/EBPα; e.g. *Calgranulin-b* (*S100a*9), *KLF5*, *Anxa1*, *Ltf*, *c-myc*, and *Csf3r* and thus, have been reported in studies investigating C/EBP function in the differentiation of haematopoietic cells (Smith *et al.*, 1996; Iwama *et al.*, 1998; Zhang *et al.*, 1998; Khanna-Gupta *et al.*, 2000; Johansen *et al.*, 2001; Cammenga *et al.*, 2003; Khanna-Gupta *et al.*, 2003). Furthermore, the set of 44 down-regulated V449E-specific genes that were identified by Brown *et al*., (2006) also contains ten known C/EBPα and/or C/EBPε target genes and includes the C/EBPα target genes, *S100a9*, *Ltf* and *Lcn2*, which were identified as differentially expressed in the microarray study reported in Chapter 4. These findings are consistent with the key role of C/EBPα in granulocyte differentiation (Radomska *et al.*, 1998; Friedman, 2002; Rosmarin *et al.*, 2005) and suggests that V449E signalling acts in part to affect activity of C/EBPα. It will be of interest to determine the nature of this pathway.

In summary, the above studies provide further evidence for genes identified as differentially expressed between FIΔ and V449E in FDB-1 cells to be involved in myeloid differentiation.

Table 21: Comparison of microarray results to differentiation studies. Table stating the overlapping genes identified to be differentially expressed in the microarray analysis presented in this thesis and other studies of cell differentiation.

Differentially expressed gene identified	Cell line, cell type or leukaemia investigated	Referenced study	Comments and findings
Anxa9	THP-1	Tuomisto *et al*., 2005	increased expression over 72 h after PMA stimulation to induce differentiation to macrophages
Bnip3	FDCP-mix	Bruno *et al*., 2004	down-regulated in all four differentiation pathways (Erythroid, Neutrophil, Neutrophil/Monocyte, Megakaryocyte)
c-myc	FDCP-mix	Bruno *et al*., 2004	Erythroid pathway - overexpressed 4 & 8h and underexpressed at 72 & 168h. Neutrophil pathway - underexpressed 4,24, 168 and overexpressed 0 and 72h
Csf3r (G-CSFR)	haematopoietic progenitors	Akashi *et al*., 2003	highly expressed in CMPs
Csf3r (G-CSFR)	FDCP-mix	Bruno *et al*., 2004	Neutrophil pathway - overexpressed at 168h
Daf1	FDCP-mix	Bruno *et al*., 2004	down-regulated in all four differentiation pathways (Erythroid, Neutrophil, Neutrophil/Macrophage, Megakaryocyte)
Gata2	haematopoietic progenitors	Akashi *et al*., 2003	predominatly overexpressed in HSC and down regulated in MPPs, CLPs and CMPs
Gata2	FDCP-mix	Bruno *et al*., 2004	Erythroid pathway - underexpressed 4 & 168h and overexpressed at 48h & 72h
Hbb-bh1	FDCP-mix	Bruno *et al*., 2004	Erythroid pathway - underexpressed at 0 & 16h and overexpressed 168h. Neutrophil pathway - underexpressed 72 & 168 and overexpressed 8 & 16h
Hbb-y	FDCP-mix	Bruno *et al*., 2004	Erythroid pathway - underexpressed 0,4 & 8 h and overexpressed at 72 & 168h
Lcn2	MPRO	Lian *et al*., 2001	up regulated upon RA induced differentiation after 24h
Ltf	FDCP-mix	Bruno *et al*., 2004	Erythroid pathway - overexpressed at 0h and underexpressed at 72h
Ltf	MPRO	Lian *et al*., 2001	transient increase in expression peaking at 48h after RA induced differentiation
Lyzs	haematopoietic progenitors	Akashi *et al*., 2003	underexpressed in CMPs and CLPs
Lyzs	FDCP-mix	Bruno *et al*., 2004	Erythroid pathway - underexpressed 0, 4 & 168h and overexpressed at 24h
Myc	MPRO	Lian *et al*., 2001	down regulated within 24 hours after RA induced differentiation
Npm1	MPRO	Lian *et al*., 2001	down regulated within 24 hours after RA induced differentiation
Plaur	MPRO	Lian *et al*., 2001	up regulated 2-3 fold over 0-72h after RA induced differentiation
Plaur	THP-1	Tuomisto *et al*., 2005	increased expression at 24 & 72 h with an early peak at 3 h after PMA stimulation to induce differentiation to macrophages
S100a6	FDCP-mix	Bruno *et al*., 2004	down-regulated in all four differentiation pathways (Erythroid, Neutrophil, Neutrophil/Macrophage, Megakaryocyte)
S100a9 (Calgranulin B)	MPRO	Lian *et al*., 2001	up regulated late in RA induced differentiation (72h)
SP100	THP-1	Tuomisto *et al*., 2005	increased expression over 72 h after PMA stimulation to induce differentiation to macrophages
Spi2-1	FDCP-mix	Bruno *et al*., 2004	Neutrophil pathway - overexpressed after 16h and underexpressed at 72h and 168h
Spi2-1	FDCP-mix	Hampson *et al*., 1997	expressed in undifferentiated FDCP-mix cells and downregulated upon granulocye-macrophage differentiation. Overexpression delays maturation and increases clonogenic potential
Spi2-1	bipotent GM progenitor cells	Hampson *et al*., 1997	downregulated upon granulocye-macrophage differentiation
Spi2-1	MPRO	Lian *et al*., 2001	down regulated within 24 hours after RA induced differentiation

6.3.1.2 Comparison of the microarray results with published studies of leukaemic cell lines and leukaemia samples

As described for the differentiation studies mentioned above, other researchers have performed array studies with leukaemic cell lines and leukaemia samples, and have revealed genes in common to the genes displaying differential expression in this thesis (see Table 22). *Cd24* expression was found to be up-regulated upon FIΔ differentiation (2.5-2.9 fold, see Table 12a, Table 13 and Table 15a) and Armstrong *et al*., (2002) found it to be differentially regulated between ALL, MLL and AML samples, with higher expression in the ALL samples compared to the MLL or AML samples. In addition *Cd44* was up-regulated upon FIΔ differentiation (2.0 fold, see Table 15b) and *CD44* was also differentially regulated between ALL, MLL and AML samples with higher expression in MLL samples than ALL samples (Armstrong *et al.*, 2002). The regulation of *Cd44* and *Cd24* in this study indicates these genes could be used as molecular markers to distinguish ALL, MLL and AML leukaemias and is suggestive of a role for these genes in these diseases.

Calgranulin B (*S100a9*) was up-regulated in FIΔ (2.6-2.8 fold, see Table 12a, Table 13 and Table 15b) and reported by Schoch *et al*., (2002) to be one of 13 genes that were sufficient to accurately separate three distinct cytogenetic AML subtypes. In the study reported by Schoch *et al*., (2002), Calgranulin B was found to be expressed at lower levels in the M3/M3v subtype of AML (t(15;17)(q22;q12)) compared to the AML M4eo (inv(16)(p13q22)) and AML M2 (t(8;21)(q22;q22)) subtypes.

Virtaneva *et al*., (2001) reported *BTG1*, *COPEB* (*KLF6*) and *MYC* as differentially expressed between AML and $CD34^+$ cells (see Table 22). *Btg1* and *Copeb* showed increased expression in FIΔ cells as they differentiated (*Btg1* 2.03 fold, Table 15b and *Copeb* 1.5-1.7 fold, Table 12a, Table 13 and Table 15a) and were up-regulated in AML patient samples compared to $CD34^+$ cells. When examining the two AML groups separately, Virtaneva *et al*., (2001) found that *COPEB* was up-regulated in the AML samples with trisomy 8 (+8) and the AML samples with normal cytogenetics (CN), but *BTG1* was up-regulated in only the AML+8 samples. The increased expression of *BTG1* in AML+8 is not due to a gene dosage effect from trisomy of chromosome 8, as the human *BTG1* gene is located on chromosome 12. In contrast, *Myc* expression was down-regulated upon differentiation in FIΔ cells, up-regulated in V449E cells (see Table 12b and Table 13) and was down-regulated in AML+8 (Virtaneva *et al.*,

2001). *COPEB* was also identified in the study by Kohlmann *et al*., (2003) as part of their minimal set of genes required for subclassification of AML subtypes.

Haferlach *et al*., (2005) analysed bone marrow and peripheral blood samples from leukaemia patients in an attempt to demonstrate the use microarrays for the diagnosis and classification of leukaemias. In their analysis they identified the top 100 differentially expressed genes from the twelve subtypes of leukaemia and using support vector machines were able to confirm and reproduce the twelve predefined leukaemia subtypes. *CD24* occurred in numerous top 100 differentially expressed lists suggesting its down-regulation may be important for leukaemia generally. *SCAP2*, *RAB32* and *LCN2* were found in two lists identified by Haferlach *et al*., (2005) (see Table 22). The *Rab32* was 1.8-1.9 fold up-regulated with FIΔ differentiation (Table 12a and Table 15a), *Scap2* was 1.7-1.8 fold up-regulated (Table 12a and Table 15b) and *Lcn2* was up-regulated 3.3 fold (Table 12a and Table 15b). Other genes discussed in this thesis were also stated in the twelve top 100 lists reported by Haferlach *et al*., (2005) included *ADAM8*, *ANXA3*, *CSF3R*, *MYBBP1A* (P160), *S100A9* and *SAMHD1*.

Lee *et al*., (2002) also found *S100A9* to be up-regulated in the acute promyelocytic cell lines, NB4 and HL-60, when induced to differentiate with retinoic acid (RA). In a similar study, Zheng *et al*., (2002) found *PLAUR* to be up-regulated within 2-3 hours of 12-*O*-tetradecanoylphorbol-13-acetate (TPA) treatment of HL-60 cells. Another study using HL-60 cells induced to differentiate with RA and 1,25-dihydroxyvitamin D3, *BTG1* was found to be up-regulated while *MYC* was down-regulated upon differentiation (Song *et al.*, 2003). Zheng *et al*., (2005) examined NB4 differentiation with RA and arsenic trioxide and found *CD44*, *S100A9* and *SP100* to be up-regulated and *MYC* to be down-regulated upon differentiation. *BTG1*, *FCERG1* and *GATA2* were found to be up-regulated by over four fold upon HL-60 differentiation into macrophages with TPA after 48 hours (Juan *et al.*, 2002).

In summary, the above studies provide evidence that genes identified to be differentially expressed between the FIΔ mutant and the leukaemic hβc mutant V449E, have been associated with leukaemic cell lines and myeloid leukaemias. The identification of these genes in the published studies summarised above suggests that they have been identified as a consequence of their roles in the proliferation and differentiation of myeloid cells and could indicate specific roles in the pathogenesis of human myeloid leukeamias.

Table 22: Comparison of microarray results to leukaemia studies. Table stating the overlapping genes identified to be differentially expressed in the microarray analysis presented in this thesis, and identified in common to other studies of leukaemic cell lines and leukaemia patient samples.

Differentially expressed gene identified	Cell line, cell type or leukaemia investigated	Referenced study	Comments / findings
Adam8	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in Pro-B-ALL/t(11q23) vs rest
Anxa3	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in CML vs rest
Btg1	AML & $CD34^+$	Virtaneva *et al*., 2001	up regulated in AML compared to $CD34^+$ cells. Up regulated in AML+8
Btg1	HL-60	Song *et al*., 2003	up regulated upon differentiation with RA and 1,25-dihydroxyvitamin D3
Btg1	HL-60	Juan *et al*., 2002	up regulated over 4 fold upon differentiation with TPA
Cd24	ALL, MLL, AML samples	Armstrong *et al*., 2002	higher expression in ALL compared to the MLL and AML samples
Cd24	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in AML with t(15;17) vs rest, AML with inv(16) vs rest, AML with 11q23/MLL rearrangement vs rest and CML vs rest
Cd44	ALL, MLL, AML samples	Armstrong *et al*., 2002	higher expression in MLL compared to the AML samples
Cd44	NB4	Zheng *et al*., 2005	up regulated upon differentiation with RA and arsenic trioxid
Copeb (Klf6)	AML & $CD34^+$	Virtaneva *et al*., 2001	up regulated in AML compared to CD34+ cells. Up regulated in AML+8 and AML samples with normal cytogenetics
Copeb (Klf6)	ALL & AML	Kohlmann *et al*., 2003	part of minimal set of genes required for subclassification of AML subtypes
Csf3r	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in Pro-B-ALL/t(11q23) vs rest
Fcerg1	HL-60	Juan *et al*., 2002	up regulated over 4 fold upon differentiation with TPA
Gata2	HL-60	Juan *et al*., 2002	up regulated over 4 fold upon differentiation with TPA
Lcn2	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in AML with normal karyotype or so called "other" cytogenetic abnormalities vs rest and CML vs rest
Mybbp1a (P160)	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in CLL vs rest
Myc	AML & $CD34^+$	Virtaneva *et al*., 2001	differentially expressed between AML and $CD34^+$ cells. Down regulated in AML+8
Myc	HL-60	Song *et al*., 2003	down regulated upon differentiation with RA and 1,25-dihydroxyvitamin D4
Myc	NB4	Zheng *et al*., 2005	down regulated upon differentiation with RA and arsenic trioxide
Plaur	HL-60	Zheng *et al*., 2002	up regulated upon induction of differentiation with TPA with 2-3 h
Rab32	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in Pro-B-ALL/t(11q23) vs rest and c-ALL/Pre-B-ALL with or without t(9;22) vs rest
S100a9	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in CML vs rest
S100a9 (Calgranulin B)	AML	Schoch *et al*., 2002	1 of 13 genes used to separate 3 distinct AML subtypes -expressed lower in M3/M3v compared to M4eo and M2
S100a9 (Calgranulin B)	NB4 and HL-60	Lee *et al*., 2002	up regulated upon induction of differentiation with RA
S100a9 (Calgranulin B)	NB4	Zheng *et al*., 2005	up regulated upon differentiation with RA and arsenic trioxide
Samhd1	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in T-ALL vs rest
Scap2	AML, ALL, CML & CLL samples	Haferlach *et al*., 2005	Differentially expressed in AML with t(15;17) vs rest and c-ALL/Pre-B-ALL with or without t(9;22) vs rest
SP100	NB4	Zheng *et al*., 2005	up regulated upon differentiation with RA and arsenic trioxide

6.4 Project overview and future studies

6.4.1 Overall success of the approach

Overall the broad aim of identifying genes that display differential expression in cells responding to the extracellular and transmembrane hβc was successful. A phosphatase encoded by the gene *Ptpmt1* was identified as differentially expressed between I374N and V449E expressing FDC-P1 cells. The differential expression of *Ptpmt1* could not be established in FDB-1 cells, as it was not present within the NIA 15K cDNA clone set. In FDB-1 cells, changes in gene expression for *Cd34*, *Gata2*, *Ltf*, *Cd24* and *Lcn2* were consistent with the differentiation of the FIΔ expressing FDB-1 cells as these cells cease their proliferation and switch to a differentiation program. This differentiation switch is clearly associated with expression of markers of mature granulocytes and neutrophils (e.g. CD24, (Elghetany and Patel, 2002; Elghetany *et al.*, 2004) and inducing the known neutrophil granule proteins Lcn2 (Kjeldsen *et al.*, 2000) and Ltf (Khanna-Gupta *et al.*, 2003)). CD34 expression is also known to decrease as myeloid progenitors progress through differentiation (Krause *et al.*, 1996). However several of the genes listed in Chapter 4, Table 17 (for example the three KLF family members, *KLF5*-*KLF7*) are genes which have not been associated with myeloid differentiation. The KLF genes as well as several other transcription factors or transcription associated genes identified in Table 17 reveal a potential role for these proteins in the transcriptional control of the differentiation of FDB-1 cells and perhaps in myeloid cells in general. This demonstrates that the microarrays can be successfully used to identify genes with regulatory roles in key biological processes (e.g. myeloid differentiation) without prior knowledge of the genes involved.

6.4.2 Future studies

6.4.2.1 Confirmation of genes showing an effect in the functional experiments

The infection optimisation experiments (described in Section 5.2.2.5) enable increased levels of FDB-1 cells to be infected so that the function experiments originally planned in Chapter 5 can now be performed. The next step would be to investigate the consequence of the over-expression of the genes, *Btg1*, *Cd24*, *Klf5* and *Nucleostemin* in FDB-1 cells. The expression of these proteins would be examied by flow cytometry

analysis for GFP expression, as was performed in the BTG1 over-expression studies performed in FDC-P1 cells. Any genes that showed an effect on proliferation and differentiation in the functional experiments would have been examined by Western blotting to confirm that the gene's effect was due to over-expression of that protein in FDB-1 cells. The HA tagged versions of the above genes (prepared in Section 5.2.2.1) could be used where suitable antibodies are not available for measuring the desired protein using Western blotting. If the C-terminal HA tag interferes with the target gene's function, the use of an N-terminal tag could be utilised or alternative tags could be investigated.

6.4.2.2 Use of differentiation and proliferation assays to assess gene function

If the differentiation associated genes examined in the FDB-1 cells behave as predicted i.e. they negatively effect the growth of the FDB-1 cells or lead to their differentiation, this will be revealed by mitogenic assays (e.g. CellTiter96^{TM} Aqueous One Solution Cell Proliferation Assay dye), examination of cell morphology on cytospin slide preparations (differential cell counts) and staining for differentiation antigens (flow cytometry). Viability assays, such as trypan blue exclusion or Annexin-V staining, would also be used to monitor changes in cell survival as a result of over-expression by comparison to control cells infected with empty vector. If the inhibition of cell proliferation by expression of these genes prevents isolation of an adequate number of infected cells, then inducible methods of gene expression may have to be investigated (see Section 6.4.2.3). An alternative technique would be to investigate the effect of the over-expression of these genes in colony assays using primary cells. In these assays a gene that induces differentiation would be measured by a decrease in colony formation compared to control cells infected with empty vector. Similarly, a gene that induces proliferation (or overcomes a block in differentiation) would be measured by an increase in colony formation and/or size compared to controls.

It would be interesting to examine the functional effects of these genes in the parental FDB-1 cells with the addition of either IL-3 or GM-CSF, as well as in FDB-1 expressing the hβc mutants. For example, it has been shown that enforced over-expression of c-Myb blocks the differentiation of FDB-1 cells in the presence of GM-CSF (Brown *et al.*, 2006). This is consistent with the known functions of Myb as an oncogene and modulator of the cell cycle (Gonda, 1998) and confirms that FDB-1 cells are responsive to enforced expression of oncogenes. It would be of interest to examine

whether the predicted differentiation-associated genes are able to overcome the proliferative signal induced by either IL3 or V449E expression in FDB-1 cells, and lead to growth arrest and other gene expression changes associated with differentiation.

6.4.2.3 Inducible systems to control expression of growth inhibitory genes

For genes that induce growth arrest, it may still be necessary to consider inducible systems in order to examine function. For example, with the use of either the Tet On or Tet Off systems (see Baron and Bujard, (2000) for review) an infected FDB-1 population could be isolated and expanded prior to the induction of expression of the desired gene. If coupled with the use of selectable markers (such as GFP) the inducible vector containing cells could be selected and enriched (e.g. by FACS) so that the functional experiments were being performed in the presence or absence of induction and only with infected cells harbouring the desired gene. Using a highly enriched population of infected cells with induced expression of the gene of interest for the functional assays, would provide an improved method for detecting small effects on differentiation induced by these genes (without the effects of uninfected cells to obscure the results).

6.4.2.4 Short interfering RNA (siRNA)

As a complementary approach to the over-expression studies suggested above, it would be of considerable interest to examine the functional consequences of gene ablation. These types of experiments can be accomplished via small interfering RNA (siRNA), which has rapidly become an accepted approach to assess gene function (Dykxhoorn *et al.,* 2003). Frequently, many oligonucleotides need to be designed in order to induce a significant siRNA effect and results would then be obtained from at least two independent functional siRNA sequences. These siRNA sequences can then be expressed from retroviral or lentiviral vectors (such as those based on the vectors described by Wiznerowicz & Trono (2003) and would be introduced into FDB-1 cells or other cell lines to examine the functional consequences of gene “knockdown”. At the time the over-expression studies were performed, no efficient method of transfecting FDB-1 cells with siRNA sequences existed. Now that the infection protocol has been improved, FDB-1 cells could be infected with retroviral vectors that express siRNA sequences. Studies using siRNA would be useful in confirming the role of any proliferation-associated genes (such as *Nucleostemin*) that were shown to increase in the V449E cell population and regulated siRNA vectors are now available that would

greatly facilitate such a study. Using siRNA directed against a proliferation-associated gene may induce either growth arrest or differentiation in the absence of the proliferative gene signal in the proliferating cell population. SiRNA could also be useful for confirming the role of genes predicted to induce growth arrest, where it would be predicted that gene knockdown may remove the growth arrest and allow the continual proliferation of FDB-1 cells growing in mGM-CSF, or FIΔ expressing FDB-1 cells growing in the absence of mIL-3.

6.4.2.5 Examination of gene function in other cell lines and primary cells

The expression of the candidate genes from the microarray study would also be interesting in other murine and human myeloid/leukaemic cell lines (for example, 32D, KG1a, HL60, NB4, THP-1 and NKM0-1 cells) and primary cells, to determine if their effects are cell type specific or of more general significance to myeloid cells. For example, mouse foetal liver cells or bone marrow cells would be harvested and infected with a retroviral over-expression construct. The properties of the infected cells would be assessed in liquid and semi-solid media to monitor proliferation and differentiation by analysing cell numbers, cell surface markers and Jenna-Giemsa staining of cytospin cell preparations. Genes that display significant effects in these systems, consistent with a regulatory role, could be assessed further in animal models (see below).

6.4.2.6 Examination of gene function in vivo

Finally, after the examination of over-expression and knockdown in cell lines and primary cells, in-vivo models should be investigated. One method is to use retrovirally infected bone marrow populations over-expressing a gene of interest and/or a dominant negative form thereof. Retrovirally infected bone marrow would be introduced into lethally irradiated recipient mice to monitor the haematopoietic reconstitution of the mice and the effect of the transgene. This technique was used by (McCormack and Gonda, 1999) to investigate the effect of FIΔ expression *in vivo*. There are also a number of transgenic models in which transgene expression is inducible in the haematopoietic system and these would be appropriate for this analysis. One such model being developed by R. D'Andrea and others (personal communication) uses the haematopoietic-specific promoter from the *Vav* gene (Ogilvy *et al.*, 1999) coupled to the reverse Tet transactivator (rtTA) to generate mice where haematopoietic over-expression of genes can be obtained. To examine the function of selected genes of interest *in vivo*, an inducible transgene would be constructed to combine with the Vav-

rtTA line of mice. For example, the transgene could be constructed using the Clontech pBI-EGFP vector which contains the Tet-operator sequences and leads to the expression of EGFP induced with the expression of the gene of interest. Inducible haematopoietic-specific gene ablation could be obtained by using these mice to express dominant-negative constructs, or by engineering lox P-recombination sites in specific genes of interest and crossing these mice with LC-1 mice expressing Cre-recombinaese regulated by a bi-directional tetracycline-inducible promoter (Schonig *et al.*, 2002).

Another inducible system for over-expression studies utilises the estrogen receptor ligand binding domain (ER) fused to a target gene. In this system, used by Cammenga *et al*., (2003) and Hogg *et al*., (1997) for the infection of human $CD34^+$ cells and primary murine foetal haematopoietic cells respectively, the addition of β-estradiol leads to activation of a fusion protein containing the target gene and the ER ligand binding domain (e.g. C/EBPα-ER). Other alternatives to this system use a modified ER ligand binding domain such that it is responsive to 4-hydroxytamoxifen (Danielian *et al.*, 1993; Littlewood *et al.*, 1995; Jager *et al.*, 2004). The 4-hydroxytamoxifen responsive system could be adapted for *in vivo* use through the use of either a transgenic mouse expressing the desired ER-target gene construct, or by retrovirally infecting haematopoietic cells with the ER-target gene and transplanting these cells into lethally irradiated recipients. Expression of the target gene in this model is induced by administration of tamoxifen to the mice. Alternatively, mice expression an ER-Cre fusion protein can be used to induce expression of Cre recombinase *in vivo*. When this is utilised in a double transgenic setting, the expression of Cre protein can be used to delete a lox-P flanked gene in a transgenic mouse, or to delete a stop cassette from a transgene *in vivo*, for loss of function or gain of function studies respectively (DeMayo and Tsai, 2001; Lewandoski, 2001; Schonig *et al.*, 2002; Jager *et al.*, 2004; Beppu *et al.*, 2005; Yu *et al.*, 2005).

Other haematopoietic specific transgenic models include the use of regulatory sequences from the human *CD34* gene (Okuno *et al.*, 2002; Radomska *et al.*, 2002). Radomska *et al*., (2002) demonstrated that they could use the regulatory elements from human *CD34* gene to drive expression of the tetracycline-responsive transactivator protein tTA *in vivo*. Furthermore, by cross breeding these CD34-tTA mice with another strain carrying Cre recombinase (under the control of tetracycline-response elements) mice would be created in which Cre expression occurred in a pattern consistent with human *CD34* transgene expression and was inducible following the withdrawal of

tetracycline administration. Adapting this system to produce Cre expression under the control of *CD34* regulatory elements would enable cell type specific recombination of genes in haematopoietic stem cells and progenitor cells (Radomska *et al.*, 2002). The *CD34* transgene model would be useful for the establishment of leukaemia models in mice in which transgene expression can be directed to stem and progenitor cells. This has been demonstrated by Huettner *et al*., (2003) for the expression of an inducible BCR/ABL disease via *CD34* regulatory elements in transgenic mice causing a megarkaryocytic myeloproliferative syndrome. Recently, Koschemieder *et al*., (2005) have used a similar system with the tTA placed under the control of the murine stem cell leukaemia gene 3' enhancer to induce BCR-ABL expression and cause chronic neutrophilia and leukocytosis in mice.

The above mentioned inducible and haematopoietic specific expression models could be adapted to investigate the gene function *in vivo* for any of the genes identified from the microarray study in which further studies are warranted. Such studies will be of particular interest in investigating the role of a gene of interest in the initiation of leukaemia. Also of interest would be any genes that caused disruption of the normal haematopoietic differentiation pathway *in vivo*, as this may provide further insights into what are the critical pathways and genes involved in haematopoiesis.

6.4.2.7 Future trends

A number of issues still limit microarray studies such as the one presented in this thesis. These relate to how to (i) prioritise, (ii) verify, (iii) analyse the function of the large number of genes identified as differentially expressed, and (iv) how to distinguish cause from effect. In this thesis methods have been presented to prioritise (bioinformatics), verify (Northern blots), and to provide a preliminary assessment of the functional consequences of the genes under investigation (over- and under-expression studies). However, the methods utilised in this thesis were not easily scaleable for the analysis of large numbers of genes. For example, in the study reported in Chapter 4, 254 genes had significant interaction parameters over the four analysis methods at a significance level of 0.001 via the F test statistic. Hence, there is a clear need for high throughput methods of real-time PCR and functional analysis so that large number of genes identified in a microarray experiment can be examined further. With the availability of vast amounts of sequence information and annotation on the web, approaches using bioinformatics that facilitate interrogation and integration of this information will be critical and will significantly facilitate the gene prioritisation process. Programs to perform such tasks

are already being developed (including DAVID (Dennis *et al.,* 2003) described in Chapter 4, Section 4.2.6.3) and several websites such as GeneCards and LocusLink which make finding information about given genes a much simpler task than it was in the past. Unlike Northern blots, real-time PCR approaches have the best potential for high throughput analysis and already companies are releasing products which allow analysis of up to 384 genes by real-time PCR (e.g microfluidic card system from Applied Biosytems). Using the TaqMan® Low Density Arrays from Applied Biosystems up to 384 primer and probe sets can be tested at once and run on the ABI PRISM® 7900HT real time PCR instrument to determine gene expression changes in a high throughput manner. However, adapting functional assays to methods suitable for high throughput techniques is more challenging, but work in this field is also progressing (Brummelkamp and Bernards, 2003). A method under development by T. Gonda and others (personal communication) is an unbiased approach for the functional analysis of large gene-sets using lentiviral over-expression vectors for high throughput functional screening. The use of lentiviral vectors offers several advantages over retroviral vectors such as their ability to transduce slowly dividing cells (e.g. haematopoietic stem cells), resistance to transcriptional silencing of the integrated transgene, and ability to use various transcriptional promoters (Yu *et al.,* 2003). Also self-inactivating lentiviral vectors enable transgene expression to be driven by internal promoters within the targeted cell - upon disabling of the viral promoter within the long terminal repeat after integration (Yu *et al.,* 2003). Access to such high throughput methods would be a significant aid in the analysis of candidate genes by enabling a more rapid functional screening of the myriad genes arising from a microarray analysis.

6.5 Concluding remarks

Membrane arrays and microarrays offer the potential of interrogating thousands of genes in single experiment. This can provide insights into genes that are involved in key biological process of interest without needing prior knowledge of the genes involved.

The studies presented here identified differentially expressed genes between FDB-1 cells expressing either an extracellular or transmembrane hβc mutation. These genes could be grouped into three categories: those genes which were differentiation-associated, those which were proliferation-associated and those which were transcription-associated (see Chapter 4, Table 14). The genes in transcription-associated list revealed a number of transcription factors with a potential role in granulocyte/macrophage growth and differentiation (including *Gata2*, see Section 4.3.2.5). Many of the genes from these three groups have been identified in other studies as important for granulocyte/macrophage proliferation and differentiation, or as having an association with myeloid leukaemia. Investigation of these genes using *in vitro* and *in vivo* models will provide insights into hβc signalling and the biological processes of proliferation and differentiation of myeloid cells. Identification of those genes that give the V449E mutant its leukaemic properties *in vivo* should reveal a pathway that when deregulated gives rise to acute myeloid leukaemia. The differential expression of C/EBPα target genes identified in this thesis is consistent with the idea reported by Brown *et al*., (2006), which suggests that V449E signalling may affect C/EBPα activity.

Bibliography

Adams, J. M., Harris, A. W., Pinkert, C. A., Corcoran, L. M., Alexander, W. S., Cory, S., Palmiter, R. D., and Brinster, R. L. (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature.* 318(6046): 533-538.

Aigner, S., Ruppert, M., Hubbe, M., Sammar, M., Sthoeger, Z., Butcher, E. C., Vestweber, D., and Altevogt, P. (1995). Heat stable antigen (mouse CD24) supports myeloid cell binding to endothelial and platelet P-selectin. *Int. Immunol.* 7(10): 1557-1565.

Aigner, S., Sthoeger, Z. M., Fogel, M., Weber, E., Zarn, J., Ruppert, M., Zeller, Y., Vestweber, D., Stahel, R., Sammar, M., and Altevogt, P. (1997). CD24, a mucin-type glycoprotein, is a ligand for P-selectin on human tumor cells. *Blood.* 89(9): 3385-3395.

Akagi, K., Suzuki, T., Stephens, R. M., Jenkins, N. A., and Copeland, N. G. (2004). RTCGD: retroviral tagged cancer gene database. *Nucleic Acids Res.* 32(Database issue): D523-527.

Akashi, K., He, X., Chen, J., Iwasaki, H., Niu, C., Steenhard, B., Zhang, J., Haug, J., and Li, L. (2003). Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis. *Blood.* 101(2): 383-389.

Akashi, T., Shirasawa, T., and Hirokawa, K. (1994). Gene expression of CD24 core polypeptide molecule in normal rat tissues and human tumor cell lines. *Virchows Arch.* 425(4): 399-406.

Alexander, W. S., Metcalf, D., and Dunn, A. R. (1995). Point mutations within a dimer interface homology domain of c-Mpl induce constitutive receptor activity and tumorigenicity. *EMBO J.* 14: 5569-5578.

Alexander, W. S., Roberts, A. W., Nicola, N. A., Li, R., and Metcalf, D. (1996). Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl. *Blood.* 87(6): 2162-2170.

Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grever, M. R., Byrd, J. C., Botstein, D., Brown, P. O., and Staudt, L. M. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 403(6769): 503-511.

Alterman, L. A., Crispe, I. N., and Kinnon, C. (1990). Characterization of the murine heat-stable antigen: an hematolymphoid differentiation antigen defined by the J11d, M1/69 and B2A2 antibodies. *Eur. J. Immunol.* 20(7): 1597-1602.

Appel, K., Buttini, M., Sauter, A., and Gebicke-Haerter, P. J. (1995). Cloning of rat interleukin-3 receptor beta-subunit from cultured microglia and its mRNA expression in vivo. *J. Neurosci.* 15(8): 5800-5809.

Aricò, M., Biondi, A., and Pui, C. H. (1997). Juvenile myelomonocytic leukemia. *Blood.* 90(2): 479-488.

Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., Minden, M. D., Sallan, S. E., Lander, E. S., Golub, T. R., and Korsmeyer, S. J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. Genet.* 30(1): 41-47.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25(1): 25-29.

Backman, M., Machon, O., Van Den Bout, C. J., and Krauss, S. (2003). Targeted disruption of mouse Dach1 results in postnatal lethality. *Dev. Dyn.* 226(1): 139-144.

Bagley, C. J., Woodcock, J. M., Stomski, F. C., and Lopez, A. F. (1997). The structural and functional basis of cytokine receptor activation: lessons from the common β **s**ubunit of the granulocyte-macrophage colony-stimulating factor, interleukin-3 (IL-3) and IL-5 receptors. *Blood.* 89: 1471-1482.

Baird, P. N., D'Andrea, R. J., and Goodall, G. J. (1995). Cytokine receptor genes: structure, chromosomal location, and involvement in human disease. *Leuk. Lymphoma.* 18(5-6): 373-383.

Baker, B. W., Boettiger, D., Spooncer, E., and Norton, J. D. (1992). Efficient retroviral-mediated gene transfer into human B lymphoblastoid cells expressing mouse ecotropic viral receptor. *Nucleic Acids Res.* 20(19): 5234.

Bakker, W. J., Blazquez-Domingo, M., Kolbus, A., Besooyen, J., Steinlein, P., Beug, H., Coffer, P. J., Lowenberg, B., von Lindern, M., and van Dijk, T. B. (2004). FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1. *J. Cell Biol.* 164(2): 175-184.

Bargmann, C. I., Hung, M. C., and Weinberg, R. A. (1986). Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell.* 45(5): 649-657.

Baron, U., and Bujard, H. (2000). Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol.* 327: 401-421.

Bartley, P. A., Lutwyche, J. K., and Gonda, T. J. (2001). Identification and validation of candidate Myb target genes. *Blood Cells. Mol. Dis.* 27(2): 409-415.

Basu, S., Dunn, A. R., Marino, M. W., Savoia, H., Hodgson, G., Lieschke, G. J., and Cebon, J. (1997). Increased tolerance to endotoxin by granulocyte-macrophage colony-stimulating factor-deficient mice. *J. Immunol.* 159(3): 1412-1417.

Basu, S., Hodgson, G., Zhang, H. H., Katz, M., Quilici, C., and Dunn, A. R. (2000). "Emergency" granulopoiesis in G-CSF-deficient mice in response to Candida albicans infection. *Blood.* 95(12): 3725-3733.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practicle and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological.* 57(1): 289-300.

Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R., and Sultan, C. (1985). Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann. Intern. Med.* 103(4): 620-625.

Benzeno, S., Narla, G., Allina, J., Cheng, G. Z., Reeves, H. L., Banck, M. S., Odin, J. A., Diehl, J. A., Germain, D., and Friedman, S. L. (2004). Cyclin-Dependent Kinase Inhibition by the KLF6 Tumor Suppressor Protein through Interaction with Cyclin D1. *Cancer Res.* 64(11): 3885-3891.

Beppu, H., Lei, H., Bloch, K. D., and Li, E. (2005). Generation of a floxed allele of the mouse BMP type II receptor gene. *Genesis.* 41(3): 133-137.

Berthet, C., Guehenneux, F., Revol, V., Samarut, C., Lukaszewicz, A., Dehay, C., Dumontet, C., Magaud, J. P., and Rouault, J. P. (2002). Interaction of PRMT1 with BTG/TOB proteins in cell signalling: molecular analysis and functional aspects. *Genes Cells.* 7(1): 29-39.

Bieker, J. J. (2001). Kruppel-like factors: three fingers in many pies. *J. Biol. Chem.* 276(37): 34355-34358.

Birnbaum, R. A., O'Marcaigh, A., Wardak, Z., Zhang, Y. Y., Dranoff, G., Jacks, T., Clapp, D. W., and Shannon, K. M. (2000). Nf1 and Gmcsf interact in myeloid leukemogenesis. *Mol. Cell.* 5(1): 189-195.

Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., and Sondak, V. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature.* 406(6795): 536-540.

Bogdan, J. A., Adams-Burton, C., Pedicord, D. L., Sukovich, D. A., Benfield, P. A., Corjay, M. H., Stoltenborg, J. K., and Dicker, I. B. (1998). Human carbon catabolite repressor protein (CCR4)-associative factor 1: cloning, expression and characterization of its interaction with the B-cell translocation protein BTG1. *Biochem. J.* 336 (Pt 2): 471-481.

Boisclair, Y., and Tremblay, M. L. (2005). Firing up mitochondrial activities with PTPMT1. *Mol. Cell.* 19(3): 291-292.

Botella, L. M., Sanchez-Elsner, T., Sanz-Rodriguez, F., Kojima, S., Shimada, J., Guerrero-Esteo, M., Cooreman, M. P., Ratziu, V., Langa, C., Vary, C. P., Ramirez, J. R., Friedman, S., and Bernabeu, C. (2002). Transcriptional activation of endoglin and transforming growth factor-beta signaling components by cooperative interaction between Sp1 and KLF6: their potential role in the response to vascular injury. *Blood.* 100(12): 4001-4010.

Bradbury, A., Possenti, R., Shooter, E. M., and Tirone, F. (1991). Molecular cloning of PC3, a putatively secreted protein whose mRNA is induced by nerve growth factor and depolarization. *Proc. Natl. Acad. Sci. U. S. A.* 88(8): 3353-3357.

Brejning, J., Jespersen, L., and Arneborg, N. (2003). Genome-wide transcriptional changes during the lag phase of Saccharomyces cerevisiae. *Arch. Microbiol.* 179(4): 278-294.

Brendel, C., and Neubauer, A. (2000). Characteristics and analysis of normal and leukemic stem cells: current concepts and future directions. *Leukemia.* 14(10): 1711-1717.

Briegel, K., Lim, K. C., Plank, C., Beug, H., Engel, J. D., and Zenke, M. (1993). Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. *Genes Dev.* 7(6): 1097-1109.

Brown, A. L., Peters, M., D'Andrea, R. J., and Gonda, T. J. (2004). Constitutive mutants of the GM-CSF receptor reveal multiple pathways leading to myeloid cell survival, proliferation, and granulocyte-macrophage differentiation. *Blood.* 103(2): 507-516.

Brown, A. L., Wilkinson, C. R., Waterman, S. R., Kok, C. H., Salerno, D. G., Diakiw, S. M., Reynolds, B., Scott, H. S., Tsykin, A., Glonek, G. F., Goodall, G. J., Solomon, P. J., Gonda, T. J., and D'Andrea, R. J. (2006). Genetic regulators of myelopoiesis and leukemic signaling identified by gene profiling and linear modeling. *J. Leukoc. Biol.* 80(2): 433-447.

Brummelkamp, T. R., and Bernards, R. (2003). New tools for functional mammalian cancer genetics. *Nat. Rev. Cancer.* 3(10): 781-789.

Bruno, L., Hoffmann, R., McBlane, F., Brown, J., Gupta, R., Joshi, C., Pearson, S., Seidl, T., Heyworth, C., and Enver, T. (2004). Molecular signatures of self-renewal, differentiation, and lineage choice in multipotential hemopoietic progenitor cells *in vitro*. *Mol. Cell. Biol.* 24(2): 741-756.

Buckley, M. J. (2000). The Spot user's guide. CSIRO Mathematical and Information Sciences, North Ryde, NSW 1670, Australia. http://www.cmis.csiro.au/iap/Spot/spotmanual.htm.

Buettner, R., Mora, L. B., and Jove, R. (2002). Activated STAT signaling in human tumors provides novel molecular targets for therapeutic intervention. *Clin. Cancer Res.* 8(4): 945-954.

Burns, J. C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J. K. (1993). Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 90(17): 8033-8037.

Cammenga, J., Mulloy, J. C., Berguido, F. J., MacGrogan, D., Viale, A., and Nimer, S. D. (2003). Induction of C/EBPalpha activity alters gene expression and differentiation of human CD34+ cells. *Blood.* 101(6): 2206-2214.

Carricaburu, V., Lamia, K. A., Lo, E., Favereaux, L., Payrastre, B., Cantley, L. C., and Rameh, L. E. (2003). The phosphatidylinositol (PI)-5-phosphate 4-kinase type II enzyme controls insulin signaling by regulating PI-3,4,5-trisphosphate degradation. *Proc. Natl. Acad. Sci. U. S. A.* 100(17): 9867-9872.

Caudill, S. P., and Hill, R. H. (1995). Multiple hypothesis tests in multiple investigations. *Stat. Med.* 14(5-7): 585-589.

Cecchi, C. (2002). Emx2: a gene responsible for cortical development, regionalization and area specification. *Gene.* 291(1-2): 1-9.

Cerottini, J. C., Engers, H. D., Macdonald, H. R., and Brunner, T. (1974). Generation of cytotoxic T lymphocytes in vitro. I. Response of normal and immune mouse spleen cells in mixed leukocyte cultures. *J. Exp. Med.* 140(3): 703-717.

Chambon, P. (1996). A decade of molecular biology of retinoic acid receptors. *FASEB J.* 10(9): 940-954.

Chang, J. M., Metcalf, D., Lang, R. A., Gonda, T. J., and Johnson, G. R. (1989). Nonneoplastic hematopoietic myeloproliferative syndrome induced by dysregulated multi-CSF (IL-3) expression. *Blood.* 73(6): 1487-1497.

Chen, C., Bhalala, H. V., Qiao, H., and Dong, J. T. (2002). A possible tumor suppressor role of the KLF5 transcription factor in human breast cancer. *Oncogene.* 21(43): 6567-6572.

Chen, C., Bhalala, H. V., Vessella, R. L., and Dong, J. T. (2003). KLF5 is frequently deleted and down-regulated but rarely mutated in prostate cancer. *Prostate.* 55(2): 81-88.

Chen, Y., Dougherty, E. R., and Bittner, M. L. (1997). Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Opt.* 2: 364-374.

Cherry, S. R., Biniszkiewicz, D., van Parijs, L., Baltimore, D., and Jaenisch, R. (2000). Retroviral expression in embryonic stem cells and hematopoietic stem cells. *Mol. Cell. Biol.* 20(20): 7419-7426.

Cheson, B. D., Cassileth, P. A., Head, D. R., Schiffer, C. A., Bennett, J. M., Bloomfield, C. D., Brunning, R., Gale, R. P., Grever, M. R., Keating, M. J., and *et al*. (1990). Report of the National Cancer Institute-sponsored workshop on definitions of diagnosis and response in acute myeloid leukemia. *J. Clin. Oncol.* 8(5): 813-819.

Cheung, V. G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R., and Childs, G. (1999). Making and reading microarrays. *Nat. Genet.* 21(1 Suppl): 15-19.

Chiambaretta, F., Blanchon, L., Rabier, B., Kao, W. W., Liu, J. J., Dastugue, B., Rigal, D., and Sapin, V. (2002). Regulation of corneal keratin-12 gene expression by the human Kruppel-like transcription factor 6. *Invest. Ophthalmol. Vis. Sci.* 43(11): 3422-3429.

Chlystun, M., Markoff, A., and Gerke, V. (2004). Structural and functional characterisation of the mouse annexin A9 promoter. *Biochim. Biophys. Acta.* 1742(1-3): 141-149.

Cho, J. W., Kim, J. J., Park, S. G., Lee do, H., Lee, S. C., Kim, H. J., Park, B. C., and Cho, S. (2004). Identification of B-cell translocation gene 1 as a biomarker for monitoring the remission of acute myeloid leukemia. *Proteomics.* 4(11): 3456-3463.

Chretien, S., Moreau-Gachelin, F., Apiou, F., Courtois, G., Mayeux, P., Dutrillaux, B., Cartron, J. P., Gisselbrecht, S., and Lacombe, C. (1994). Putative oncogenic role of the erythropoietin receptor in murine and human erythroleukemia cells. *Blood.* 83(7): 1813-1821.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O., and Herskowitz, I. (1998). The transcriptional program of sporulation in budding yeast. *Science.* 282(5389): 699-705.

Chu, S., Xu, H., Shah, N. P., Snyder, D. S., Forman, S. J., Sawyers, C. L., and Bhatia, R. (2005). Detection of BCR-ABL kinase mutations in CD34+ cells from chronic myelogenous leukemia patients in complete cytogenetic remission on imatinib mesylate treatment. *Blood.* 105(5): 2093-2098.

Churchill, G. A. (2002). Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* 32 Suppl: 490-495.

Clontech. (2004). Retroviral Gene Transfer and Expression User Manual. PT3132-1 (PR43789), BD Biosciences Clontech. http://www.clontech.com/clontech/techinfo/manuals/PDF/PT3132-1.pdf.

Constantinescu, S. N., Liu, X., Beyer, W., Fallon, A., Shekar, S., Henis, Y. I., Smith, S. O., and Lodish, H. F. (1999). Activation of the erythropoietin receptor by the gp55-P viral envelope protein is determined by a single amino acid in its transmembrane domain. *EMBO J.* 18(12): 3334-3347.

Coombes, K. R., Highsmith, W. E., Krogmann, T. A., Baggerly, K. A., Stivers, D. N., and Abruzzo, L. V. (2002). Identifying and quantifying sources of variation in microarray data using high-density cDNA membrane arrays. *J. Comput. Biol.* 9(4): 655-669.

Corrente, G., Guardavaccaro, D., and Tirone, F. (2002). PC3 potentiates NGF-induced differentiation and protects neurons from apoptosis. *Neuroreport.* 13(4): 417-422.

Dahl, M. E., Arai, K. I., and Watanabe, S. (2000). Association of Lyn tyrosine kinase to the GM-CSF and IL-3 receptor common betac subunit and role of Src tyrosine kinases in DNA synthesis and anti-apoptosis. *Genes. Cells.* 5(2): 143-153.

D'Andrea, R., Rayner, J., Moretti, P., Lopez, A., Goodall, G. J., Gonda, T. J., and Vadas, M. (1994). A mutation of the common receptor subunit for interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor, and IL-5 that leads to ligand independence and tumorigenicity. *Blood.* 83(10): 2802-2808.

D'Andrea, R. J., Barry, S. C., Moretti, P. A., Jones, K., Ellis, S., Vadas, M. A., and Goodall, G. J. (1996). Extracellular truncations of h beta c, the common signaling subunit for interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor (GM-CSF), and IL-5, lead to ligand-independent activation. *Blood.* 87(7): 2641-2648.

D'Andrea, R. J., and Gonda, T. J. (2000). A model for assembly and activation of the GM-CSF, IL-3 and IL-5 receptors: insights from activated mutants of the common beta subunit. *Exp. Hematol.* 28(3): 231-243.

D'Andrea, R. J., Harrison-Findik, D., Butcher, C. M., Finnie, J., Blumbergs, P., Bartley, P., McCormack, M., Jones, K., Rowland, R., Gonda, T. J., and Vadas, M. A. (1998). Dysregulated hematopoiesis and a progressive neurological disorder induced by expression of an activated form of the human common beta chain in transgenic mice. *J Clin. Invest.* 102(11): 1951-1960.

Danielian, P. S., White, R., Hoare, S. A., Fawell, S. E., and Parker, M. G. (1993). Identification of residues in the estrogen receptor that confer differential sensitivity to estrogen and hydroxytamoxifen. *Mol. Endocrinol.* 7(2): 232-240.

Dano, K., Behrendt, N., Hoyer-Hansen, G., Johnsen, M., Lund, L. R., Ploug, M., and Romer, J. (2005). Plasminogen activation and cancer. *Thromb. Haemost.* 93(4): 676-681.

Davis, R. J., Shen, W., Sandler, Y. I., Amoui, M., Purcell, P., Maas, R., Ou, C. N., Vogel, H., Beaudet, A. L., and Mardon, G. (2001). Dach1 mutant mice bear no gross abnormalities in eye, limb, and brain development and exhibit postnatal lethality. *Mol. Cell. Biol.* 21(5): 1484-1490.

de Groot, R. P., Coffer, P. J., and Koenderman, L. (1998). Regulation of proliferation, differentiation and survival by the IL-3/IL-5/GM-CSF receptor family. *Cell. Signal.* 10(9): 619-628.

de Haan, G., Bystrykh, L. V., Weersing, E., Dontje, B., Geiger, H., Ivanova, N., Lemischka, I. R., Vellenga, E., and Van Zant, G. (2002). A genetic and genomic analysis identifies a cluster of genes associated with hematopoietic cell turnover. *Blood.* 100(6): 2056-2062.

de la Cruz, J., Kressler, D., and Linder, P. (1999). Unwinding RNA in Saccharomyces cerevisiae: DEAD-box proteins and related families. *Trends Biochem. Sci.* 24(5): 192-198.

de Sauvage, F. J., Carver-Moore, K., Luoh, S. M., Ryan, A., Dowd, M., Eaton, D. L., and Moore, M. W. (1996). Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. *J. Exp. Med.* 183(2): 651-656.

DeMayo, F. J., and Tsai, S. Y. (2001). Targeted gene regulation and gene ablation. *Trends Endocrinol Metab.* 12(8): 348-353.

Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4(5): P3.

DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. (1996). Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* 14(4): 457-460.

Dexter, T. M., Garland, J., Scott, D., Scolnick, E., and Metcalf, D. (1980). Growth of factor-dependent haemopoietic precursor cell lines. *J. Exp. Med.* 152: 1036-1047.

Dilworth, F. J., Fromental-Ramain, C., Remboutsika, E., Benecke, A., and Chambon, P. (1999). Ligand-dependent activation of transcription *in vitro* by retinoic acid receptor alpha/retinoid X receptor alpha heterodimers that mimics transactivation by retinoids *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* 96(5): 1995-2000.

Dong, F., Brynes, R. K., Tidow, N., Welte, K., Lowenberg, B., and Touw, I. P. (1995a). Mutations in the gene for the granulocyte colony-stimulating-factor receptor in patients with acute myeloid leukemia preceded by severe congenital neutropenia. *N. Engl. J. Med.* 333(8): 487-493.

Dong, F., Dale, D. C., Bonilla, M. A., Freedman, M., Fasth, A., Neijens, H. J., Palmblad, J., Briars, G. L., Carlsson, G., Veerman, A. J., Welte, K., Lowenberg, B., and Touw, I. P. (1997). Mutations in the granulocyte colony-stimulating factor receptor gene in patients with severe congenital neutropenia. *Leukemia.* 11(1): 120-125.

Dong, F., van Paassen, M., van Buitenen, C., Hoefsloot, L. H., Lowenberg, B., and Touw, I. P. (1995b). A point mutation in the granulocyte colony-stimulating factor receptor (G-CSF-R) gene in a case of acute myeloid leukemia results in the overexpression of a novel G-CSF-R isoform. *Blood.* 85(4): 902-911.

Doyle, S. E., and Gasson, J. C. (1998). Characterization of the role of the human granulocyte-macrophage colony- stimulating factor receptor alpha subunit in the activation of JAK2 and STAT5. *Blood.* 92(3): 867-876.

Dranoff, G., Crawford, A. D., Sadelain, M., Ream, B., Rashid, A., Bronson, R. T., Dickersin, G. R., Bachurski, C. J., Mark, E. L., Whitsett, J. A., and *et al.* (1994). Involvement of granulocyte-macrophage colony-stimulating factor in pulmonary homeostasis. *Science.* 264(5159): 713-716.

Drexler, H. G. (1996). Expression of FLT3 receptor and response to FLT3 ligand by leukemic cells. *Leukemia.* 10(4): 588-599.

Droz, D., Rousseau-Merck, M. F., Jaubert, F., Diebold, N., Nezelof, C., Adafer, E., and Mouly, H. (1990). Cell differentiation in Wilms' tumor (nephroblastoma): an immunohistochemical study. *Hum. Pathol.* 21(5): 536-544.

Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., Lydon, N. B., Kantarjian, H., Capdeville, R., Ohno-Jones, S., and Sawyers, C. L. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* 344(14): 1031-1037.

Dudley, A. M., Aach, J., Steffen, M. A., and Church, G. M. (2002). Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. *Proc. Natl. Acad. Sci. U. S. A.* 99(11): 7554-7559.

Dudoit, S., Yang, Y. H., Callow, M. J., and Speed, T. P. (2000). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. Technical Report 578, Department of Statistics, University of California, Berkeley. http://www.stat.berkeley.edu/users/terry/zarray/Html/papersindex.html.

Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J. M. (1999). Expression profiling using cDNA microarrays. *Nat. Genet.* 21(1 Suppl): 10-14.

Dykxhoorn, D. M., Novina, C. D., and Sharp, P. A. (2003). Killing the messenger: short RNAs that silence gene expression. *Nat. Rev. Mol. Cell Biol.* 4(6): 457-467.

Ecsedy, J. A., Michaelson, J. S., and Leder, P. (2003). Homeodomain-interacting protein kinase 1 modulates Daxx localization, phosphorylation, and transcriptional activity. *Mol. Cell. Biol.* 23(3): 950-960.

Efron, B. (2003). Robbins, Empirical Bayes and Microarrays. *The Annals of Statistics.* 31(2): 366-378.

Efron, B., Tibshirani, R., Storey, J. D., and Tusher, V. (2001). Empirical Bayes analysis of a microarray experiment. *Journal of the American Statisitical Association.* 96(456): 1151-1160.

Eisen, M. B., and Brown, P. O. (1999). DNA arrays for analysis of gene expression. *Methods Enzymol.* 303: 179-205.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* 95(25): 14863-14868.

El Rouby, S., and Newcomb, E. W. (1996). Identification of Bcd, a novel proto-oncogene expressed in B-cells. *Oncogene.* 13(12): 2623-2630.

Elghetany, M. T., Ge, Y., Patel, J., Martinez, J., and Uhrova, H. (2004). Flow cytometric study of neutrophilic granulopoiesis in normal bone marrow using an expanded panel of antibodies: correlation with morphologic assessments. *J. Clin. Lab. Anal.* 18(1): 36-41.

Elghetany, M. T., and Patel, J. (2002). Assessment of CD24 expression on bone marrow neutrophilic granulocytes: CD24 is a marker for the myelocytic stage of development. *Am. J. Hematol.* 71(4): 348-349.

Elghetany, M. T., Patel, J., Martinez, J., and Schwab, H. (2003). CD87 as a marker for terminal granulocytic maturation: assessment of its expression during granulopoiesis. *Cytometry B Clin Cytom.* 51(1): 9-13.

Emanuel, P. D., Bates, L. J., Castleberry, R. P., Gualtieri, R. J., and Zuckerman, K. S. (1991). Selective hypersensitivity to granulocyte-macrophage colony-stimulating factor by juvenile chronic myeloid leukemia hematopoietic progenitors. *Blood.* 77(5): 925-929.

Emanuel, P. D., Shannon, K. M., and Castleberry, R. P. (1996). Juvenile myelomonocytic leukemia: molecular understanding and prospects for therapy. *Mol. Med. Today.* 2(11): 468-475.

Emi, N., Friedmann, T., and Yee, J. K. (1991). Pseudotype formation of murine leukemia virus with the G protein of vesicular stomatitis virus. *J. Virol.* 65(3): 1202-1207.

Enzler, T., Gillessen, S., Manis, J. P., Ferguson, D., Fleming, J., Alt, F. W., Mihm, M., and Dranoff, G. (2003). Deficiencies of GM-CSF and interferon gamma link inflammation and cancer. *J. Exp. Med.* 197(9): 1213-1219.

Evans, V. J. (1998). Acute Myeloproliferative Leukemia (Acute Nonlymphocytic Leukemia). In *Clinical Hematology : principles, procedures, correlations.* (Stiene-Martin, E. A., Lotspeich-Steininger, C. A., and Koepke, J. A., eds), 2nd Ed., pp. 442-454, Lippincott-Raven Publishers, Philadelphia, USA.

Felsher, D. W., and Bishop, J. M. (1999). Reversible tumorigenesis by MYC in hematopoietic lineages. *Mol. Cell.* 4(2): 199-207.

Fenski, R., Flesch, K., Serve, S., Mizuki, M., Oelmann, E., Kratz-Albers, K., Kienast, J., Leo, R., Schwartz, S., Berdel, W. E., and Serve, H. (2000). Constitutive activation of FLT3 in acute myeloid leukaemia and its consequences for growth of 32D cells. *Br. J. Haematol.* 108(2): 322-330.

Feuerstein, N., Chan, P. K., and Mond, J. J. (1988). Identification of numatrin, the nuclear matrix protein associated with induction of mitogenesis, as the nucleolar protein B23. Implication for the role of the nucleolus in early transduction of mitogenic signals. *J. Biol. Chem.* 263(22): 10608-10612.

Feuerstein, N., and Mond, J. J. (1987). "Numatrin," a nuclear matrix protein associated with induction of proliferation in B lymphocytes. *J. Biol. Chem.* 262(23): 11389-11397.

Forbes, L. V., Gale, R. E., Pizzey, A., Pouwels, K., Nathwani, A., and Linch, D. C. (2002). An activating mutation in the transmembrane domain of the granulocyte colony-stimulating factor receptor in patients with acute myeloid leukemia. *Oncogene.* 21(39): 5981-5989.

Franke, T. F., Kaplan, D. R., and Cantley, L. C. (1997). PI3K: downstream AKTion blocks apoptosis. *Cell.* 88(4): 435-437.

Freeburn, R. W., Gale, R. E., and Linch, D. C. (1998). Activating point mutations in the betaC chain of the GM-CSF, IL-3 and IL-5 receptors are not a major contributory factor in the pathogenesis of acute myeloid leukaemia. *Br. J. Haematol.* 103(1): 66-71.

Freedman, M. H., Grunberger, T., Correa, P., Axelrad, A. A., Dube, I. D., and Cohen, A. (1993). Autocrine and paracrine growth control by granulocyte-monocyte colony-stimulating factor of acute lymphoblastic leukemia cells. *Blood.* 81(11): 3068-3075.

Friedman, A. D. (2002). Transcriptional regulation of granulocyte and monocyte development. *Oncogene.* 21(21): 3377-3390.

Futter, C. E., Felder, S., Schlessinger, J., Ullrich, A., and Hopkins, C. R. (1993). Annexin I is phosphorylated in the multivesicular body during the processing of the epidermal growth factor receptor. *J. Cell Biol.* 120(1): 77-83.

Gallicchio, V. S. (1998). Hematopoiesis and Review of Genetics. In *Clinical Hematology : principles, procedures, correlations.* (Stiene-Martin, E. A., Lotspeich-Steininger, C. A., and Koepke, J. A., eds), 2nd Ed., pp. 46-56, Lippincott-Raven Publishers, Philadelphia, USA.

Gaur, M., Murphy, G. J., deSauvage, F. J., and Leavitt, A. D. (2001). Characterization of Mpl mutants using primary megakaryocyte-lineage cells from mpl(-/-) mice: a new system for Mpl structure-function studies. *Blood.* 97(6): 1653-1661.

Gehrau, R. C., D'Astolfo, D. S., Prieto, C., Bocco, J. L., and Koritschoner, N. P. (2005). Genomic organization and functional analysis of the gene encoding the Kruppel-like transcription factor KLF6. *Biochim. Biophys. Acta.* 1730(2): 137-146.

Geissler, K., Ohler, L., Fodinger, M., Virgolini, I., Leimer, M., Kabrna, E., Kollars, M., Skoupy, S., Bohle, B., Rogy, M., and Lechner, K. (1996). Interleukin 10 inhibits growth and granulocyte/macrophage colony- stimulating factor production in chronic myelomonocytic leukemia cells. *J. Exp. Med.* 184(4): 1377-1384.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y., and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5(10): R80.

Gerke, V., and Moss, S. E. (2002). Annexins: from structure to function. *Physiol. Rev.* 82(2): 331-371.

Gilliland, D. G., and Griffin, J. D. (2002). The roles of FLT3 in hematopoiesis and leukemia. *Blood.* 100(5): 1532-1542.

Gilliland, D. G., Jordan, C. T., and Felix, C. A. (2004). The molecular basis of leukemia. *Hematology Am Soc Hematol Educ Program.* 80-97.

Glonek, G. F., and Solomon, P. J. (2004). Factorial and time course designs for cDNA microarray experiments. *Biostatistics.* 5(1): 89-111.

Goebeler, V., Ruhe, D., Gerke, V., and Rescher, U. (2003). Atypical properties displayed by annexin A9, a novel member of the annexin family of Ca(2+) and lipid binding proteins. *FEBS Lett.* 546(2-3): 359-364.

Golub, T. R. (1999). The Genetics of AML: An Update. *Hematology Am Soc Hematol Educ Program.* 102-111.

Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science.* 286(5439): 531-537.

Gonda, T. J. (1998). The c-Myb oncoprotein. *Int. J. Biochem. Cell Biol.* 30(5): 547-551.

Gonda, T. J., and D'Andrea, R. J. (1997). Activating mutations in cytokine receptors: implications for receptor function and role in disease. *Blood.* 89(2): 355-369.

Gonda, T. J., Ramsay, R. G., and Johnson, G. R. (1989). Murine myeloid cell lines derived by *in vitro* infection with recombinant c-myb retroviruses express myb from rearranged vector proviruses. *EMBO J.* 8(6): 1767-1775.

Gorman, D. M., Itoh, N., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., and Miyajima, A. (1992). Chromosomal localization and organization of the murine genes encoding the beta subunits (AIC2A and AIC2B) of the interleukin 3, granulocyte/macrophage colony-stimulating factor, and interleukin 5 receptors. *J. Biol. Chem.* 267(22): 15842-15848.

Gorman, D. M., Itoh, N., Kitamura, T., Schreurs, J., Yonehara, S., Yahara, I., Arai, K., and Miyajima, A. (1990). Cloning and expression of a gene encoding an interleukin 3 receptor-like protein: identification of another member of the cytokine receptor gene family. *Proc. Natl. Acad. Sci. U. S. A.* 87(14): 5459-5463.

Gouilleux-Gruart, V., Debierre-Grockiego, F., Gouilleux, F., Capiod, J. C., Claisse, J. F., Delobel, J., and Prin, L. (1997). Activated Stat related transcription factors in acute leukemia. *Leuk. Lymphoma.* 28(1-2): 83-88.

Goyal, R. K., and Longmore, G. D. (1999). Abnormalities of cytokine receptor signalling contributing to diseases of red blood cell production. *Ann. Med.* 31(3): 208-216.

Grass, J. A., Boyer, M. E., Pal, S., Wu, J., Weiss, M. J., and Bresnick, E. H. (2003). GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc. Natl. Acad. Sci. U. S. A.* 100(15): 8811-8816.

Grez, M., Akgun, E., Hilberg, F., and Ostertag, W. (1990). Embryonic stem cell virus, a recombinant murine retrovirus with expression in embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 87(23): 9202-9206.

Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., Fanelli, M., Ruthardt, M., Ferrara, F. F., Zamir, I., Seiser, C., Lazar, M. A., Minucci, S., and Pelicci, P. G. (1998). Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. *Nature.* 391(6669): 815-818.

Grondahl-Hansen, J., Peters, H. A., van Putten, W. L., Look, M. P., Pappot, H., Ronne, E., Dano, K., Klijn, J. G., Brunner, N., and Foekens, J. A. (1995). Prognostic significance of the receptor for urokinase plasminogen activator in breast cancer. *Clin. Cancer Res.* 1(10): 1079-1087.

Gualtieri, R. J., Emanuel, P. D., Zuckerman, K. S., Martin, G., Clark, S. C., Shadduck, R. K., Dracker, R. A., Akabutu, J., Nitschke, R., Hetherington, M. L., and *et al*. (1989). Granulocyte-macrophage colony-stimulating factor is an endogenous regulator of cell proliferation in juvenile chronic myelogenous leukemia. *Blood.* 74(7): 2360-2367.

Guidez, F., Ivins, S., Zhu, J., Soderstrom, M., Waxman, S., and Zelent, A. (1998). Reduced retinoic acid-sensitivities of nuclear receptor corepressor binding to PML- and PLZF-RARalpha underlie molecular pathogenesis and treatment of acute promyelocytic leukemia. *Blood.* 91(8): 2634-2642.

Guldner, H. H., Szostecki, C., Grotzinger, T., and Will, H. (1992). IFN enhance expression of Sp100, an autoantigen in primary biliary cirrhosis. *J. Immunol.* 149(12): 4067-4073.

Gurney, A. L., Carver-Moore, K., de Sauvage, F. J., and Moore, M. W. (1994). Thrombocytopenia in c-mpl-deficient mice. *Science.* 265(5177): 1445-1447.

Guthridge, M. A., Stomski, F. C., Barry, E. F., Winnall, W., Woodcock, J. M., McClure, B. J., Dottore, M., Berndt, M. C., and Lopez, A. F. (2000). Site-specific serine phosphorylation of the IL-3 receptor is required for hemopoietic cell survival. *Mol. Cell.* 6(1): 99-108.

Guthridge, M. A., Stomski, F. C., Thomas, D., Woodcock, J. M., Bagley, C. J., Berndt, M. C., and Lopez, A. F. (1998). Mechanism of activation of the GM-CSF, IL-3, and IL-5 family of receptors. *Stem Cells.* 16(5): 301-313.

Haferlach, T., Kohlmann, A., Schnittger, S., Dugas, M., Hiddemann, W., Kern, W., and Schoch, C. (2005). Global approach to the diagnosis of leukemia using gene expression profiling. *Blood.* 106(4): 1189-1198.

Hammond, K. L., Hanson, I. M., Brown, A. G., Lettice, L. A., and Hill, R. E. (1998). Mammalian and Drosophila dachshund genes are related to the Ski proto-oncogene and are expressed in eye and limb. *Mech. Dev.* 74(1-2): 121-131.

Hampson, I. N., Hampson, L., Pinkoski, M., Cross, M., Heyworth, C. M., Bleackley, R. C., Atkinson, E., and Dexter, T. M. (1997). Identification of a serpin specifically expressed in multipotent and bipotent hematopoietic progenitor cells and in activated T cells. *Blood.* 89(1): 108-118.

Han, C., Zhang, X., Xu, W., Wang, W., Qian, H., and Chen, Y. (2005). Cloning of the nucleostemin gene and its function in transforming human embryonic bone marrow mesenchymal stem cells into F6 tumor cells. *Int. J. Mol. Med.* 16(2): 205-213.

Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell.* 100(1): 57-70.

Hann, B., and Balmain, A. (2001). Building 'validated' mouse models of human cancer. *Curr. Opin. Cell Biol.* 13(6): 778-784.

Hannon, G. J., and Rossi, J. J. (2004). Unlocking the potential of the human genome with RNA interference. *Nature.* 431(7006): 371-378.

Hapel, A. J., Vande Woude, G., Campbell, H. D., Young, I. G., and Robins, T. (1986). Generation of an autocrine leukaemia using a retroviral expression vector carrying the interleukin-3 gene. *Lymphokine Res.* 5(4): 249-254.

Hara, T., and Miyajima, A. (1992). Two distinct functional high affinity receptors for mouse interleukin-3 (IL-3). *EMBO J.* 11(5): 1875-1884.

Hawley, R. G., Lieu, F. H., Fong, A. Z., and Hawley, T. S. (1994). Versatile retroviral vectors for potential use in gene therapy. *Gene Ther.* 1(2): 136-138.

He, L. Z., Guidez, F., Tribioli, C., Peruzzi, D., Ruthardt, M., Zelent, A., and Pandolfi, P. P. (1998). Distinct interactions of PML-RARalpha and PLZF-RARalpha with co-repressors determine differential responses to RA in APL. *Nat. Genet.* 18(2): 126-135.

He, L. Z., Merghoub, T., and Pandolfi, P. P. (1999). *In vivo* analysis of the molecular pathogenesis of acute promyelocytic leukemia in the mouse and its therapeutic implications. *Oncogene.* 18(38): 5278-5292.

Heisterkamp, N., Stephenson, J. R., Groffen, J., Hansen, P. F., de Klein, A., Bartram, C. R., and Grosveld, G. (1983). Localization of the c-ab1 oncogene adjacent to a translocation break point in chronic myelocytic leukaemia. *Nature.* 306(5940): 239-242.

Heyworth, C., Gale, K., Dexter, M., May, G., and Enver, T. (1999). A GATA-2/estrogen receptor chimera functions as a ligand-dependent negative regulator of self-renewal. *Genes Dev.* 13(14): 1847-1860.

Hill, A. A., Brown, E. L., Whitley, M. Z., Tucker-Kellogg, G., Hunter, C. P., and Slonim, D. K. (2001). Evaluation of normalization procedures for oligonucleotide array data based on spiked cRNA controls. *Genome Biol.* 2(12): RESEARCH0055.

Hogg, A., Schirm, S., Nakagoshi, H., Bartley, P., Ishii, S., Bishop, J. M., and Gonda, T. J. (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(24): 2885-2898.

Holloway, A. J., van Laar, R. K., Tothill, R. W., and Bowtell, D. D. (2002). Options available--from start to finish--for obtaining data from DNA microarrays II. *Nat. Genet.* 32 Suppl: 481-489.

Hopkins, N. (1993). High titers of retrovirus (vesicular stomatitis virus) pseudotypes, at last. *Proc. Natl. Acad. Sci. U. S. A.* 90(19): 8759-8760.

Hosack, D. A., Dennis, G., Jr., Sherman, B. T., Lane, H. C., and Lempicki, R. A. (2003). Identifying biological themes within lists of genes with EASE. *Genome Biol.* 4(10): R70.

Hosono, K., Sasaki, T., Minoshima, S., and Shimizu, N. (2004). Identification and characterization of a novel gene family YPEL in a wide spectrum of eukaryotic species. *Gene.* 340(1): 31-43.

Huang, L. R., and Hsu, H. C. (1995). Cloning and expression of CD24 gene in human hepatocellular carcinoma: a potential early tumor marker gene correlates with p53 mutation and tumor differentiation. *Cancer Res.* 55(20): 4717-4721.

Huettner, C. S., Koschmieder, S., Iwasaki, H., Iwasaki-Arai, J., Radomska, H. S., Akashi, K., and Tenen, D. G. (2003). Inducible expression of BCR/ABL using human CD34 regulatory elements results in a megakaryocytic myeloproliferative syndrome. *Blood.* 102(9): 3363-3370.

Huffman, J. A., Hull, W. M., Dranoff, G., Mulligan, R. C., and Whitsett, J. A. (1996). Pulmonary epithelial cell expression of GM-CSF corrects the alveolar proteinosis in GM-CSF-deficient mice. *J. Clin. Invest.* 97(3): 649-655.

Hunter, T. (2000). Signaling--2000 and beyond. *Cell.* 100(1): 113-127.

Hwang, H. C., Martins, C. P., Bronkhorst, Y., Randel, E., Berns, A., Fero, M., and Clurman, B. E. (2002). Identification of oncogenes collaborating with p27Kip1 loss by insertional mutagenesis and high-throughput insertion site analysis. *Proc. Natl. Acad. Sci. U. S. A.* 99(17): 11293-11298.

Ihaka, R., and Gentleman, R. (1996). R: a language for data analysis and graphics. *J. Comput. Graph. Statist.* 5: 299-314.

Ihle, J. N. (1996). STATs: signal transducers and activators of transcription. *Cell.* 84(3): 331-334.

Ikegami, M., Ueda, T., Hull, W., Whitsett, J. A., Mulligan, R. C., Dranoff, G., and Jobe, A. H. (1996). Surfactant metabolism in transgenic mice after granulocyte macrophage-colony stimulating factor ablation. *Am. J. Physiol.* 270(4 Pt 1): L650-658.

Ilaria, R. L., Jr., Hawley, R. G., and Van Etten, R. A. (1999). Dominant negative mutants implicate STAT5 in myeloid cell proliferation and neutrophil differentiation. *Blood.* 93(12): 4154-4166.

Iler, N., Rowitch, D. H., Echelard, Y., McMahon, A. P., and Abate-Shen, C. (1995). A single homeodomain binding site restricts spatial expression of Wnt-1 in the developing brain. *Mech. Dev.* 53(1): 87-96.

Insinga, A., Monestiroli, S., Ronzoni, S., Gelmetti, V., Marchesi, F., Viale, A., Altucci, L., Nervi, C., Minucci, S., and Pelicci, P. G. (2005). Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. *Nat. Med.* 11(1): 71-76.

Ito, G., Uchiyama, M., Kondo, M., Mori, S., Usami, N., Maeda, O., Kawabe, T., Hasegawa, Y., Shimokata, K., and Sekido, Y. (2004). Kruppel-Like Factor 6 Is Frequently Down-Regulated and Induces Apoptosis in Non-Small Cell Lung Cancer Cells. *Cancer Res.* 64(11): 3838-3843.

Iversen, P. O., Lewis, I. D., Turczynowicz, S., Hasle, H., Niemeyer, C., Schmiegelow, K., Bastiras, S., Biondi, A., Hughes, T. P., and Lopez, A. F. (1997). Inhibition of granulocyte-macrophage colony-stimulating factor prevents dissemination and induces remission of juvenile myelomonocytic leukemia in engrafted immunodeficient mice. *Blood.* 90(12): 4910-4917.

Iwai, K., Hirata, K., Ishida, T., Takeuchi, S., Hirase, T., Rikitake, Y., Kojima, Y., Inoue, N., Kawashima, S., and Yokoyama, M. (2004). An anti-proliferative gene BTG1 regulates angiogenesis *in vitro*. *Biochem. Biophys. Res. Commun.* 316(3): 628-635.

Iwama, A., Zhang, P., Darlington, G. J., McKercher, S. R., Maki, R., and Tenen, D. G. (1998). Use of RDA analysis of knockout mice to identify myeloid genes regulated *in vivo* by PU.1 and C/EBPalpha. *Nucleic Acids Res.* 26(12): 3034-3043.

Jackson, R. J., Howell, M. T., and Kaminski, A. (1990). The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem. Sci.* 15(12): 477-483.

Jafari, P., and Azuaje, F. (2006). An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors. *BMC Med Inform Decis Mak.* 6: 27.

Jager, R., Maurer, J., Jacob, A., and Schorle, H. (2004). Cell type-specific conditional regulation of the c-myc proto-oncogene by combining Cre/loxP recombination and tamoxifen-mediated activation. *Genesis.* 38(3): 145-150.

James, C., Ugo, V., Le Couedic, J. P., Staerk, J., Delhommeau, F., Lacout, C., Garcon, L., Raslova, H., Berger, R., Bennaceur-Griscelli, A., Villeval, J. L., Constantinescu, S. N., Casadevall, N., and Vainchenker, W. (2005). A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature.* 434(7037): 1144-1148.

Jang, S. K., Krausslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.* 62(8): 2636-2643.

Jenkins, B. J., Blake, T. J., and Gonda, T. J. (1998). Saturation mutagenesis of the beta subunit of the human granulocyte- macrophage colony-stimulating factor receptor shows clustering of constitutive mutations, activation of ERK MAP kinase and STAT pathways, and differential beta subunit tyrosine phosphorylation. *Blood.* 92(6): 1989-2002.

Jenkins, B. J., D'Andrea, R., and Gonda, T. J. (1995). Activating point mutations in the common beta subunit of the human GM- CSF, IL-3 and IL-5 receptors suggest the involvement of beta subunit dimerization and cell type-specific molecules in signalling. *EMBO J.* 14(17): 4276-4287.

Jenkins, B. J., Le, F., and Gonda, T. J. (1999). A cell type-specific constitutive point mutant of the common beta- subunit of the human granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin (IL)-3, and IL-5 receptors requires the GM-CSF receptor alpha-subunit for activation. *J. Biol. Chem.* 274(13): 8669-8677.

Jiang, X., Fujisaki, T., Nicolini, F., Berger, M., Holyoake, T., Eisterer, W., Eaves, C., and Eaves, A. (2000). Autonomous multi-lineage differentiation in vitro of primitive CD34+ cells from patients with chronic myeloid leukemia. *Leukemia.* 14(6): 1112-1121.

Jiang, X., Lopez, A., Holyoake, T., Eaves, A., and Eaves, C. (1999). Autocrine production and action of IL-3 and granulocyte colony-stimulating factor in chronic myeloid leukemia. *Proc. Natl. Acad. Sci. U. S. A.* 96(22): 12804-12809.

Johansen, L. M., Iwama, A., Lodie, T. A., Sasaki, K., Felsher, D. W., Golub, T. R., and Tenen, D. G. (2001). c-Myc is a critical target for c/EBPalpha in granulopoiesis. *Mol. Cell. Biol.* 21(11): 3789-3806.

Johnson, G. R., Gonda, T. J., Metcalf, D., Hariharan, I. K., and Cory, S. (1989). A lethal myeloproliferative syndrome in mice transplanted with bone marrow cells infected with a retrovirus expressing granulocyte-macrophage colony stimulating factor. *EMBO J.* 8(2): 441-448.

Jones, K. L., Bagley, C. J., Butcher, C., Barry, S. C., Vadas, M. A., and D'Andrea, R. J. (2001). Peptide insertions in domain 4 of hbeta(c), the shared signalling receptor subunit for GM-CSF, IL3 and IL5, induce ligand-independent activation. *Cytokine.* 14(6): 303-315.

Juan, H. F., Lin, J. Y., Chang, W. H., Wu, C. Y., Pan, T. L., Tseng, M. J., Khoo, K. H., and Chen, S. T. (2002). Biomic study of human myeloid leukemia cells differentiation to macrophages using DNA array, proteomic, and bioinformatic analytical methods. *Electrophoresis.* 23(15): 2490-2504.

Jung, K. C., Park, W. S., Kim, H. J., Choi, E. Y., Kook, M. C., Lee, H. W., and Bae, Y. (2004). TCR-independent and caspase-independent apoptosis of murine thymocytes by CD24 cross-linking. *J. Immunol.* 172(2): 795-802.

Just, U., Katsuno, M., Stocking, C., Spooncer, E., and Dexter, M. (1993). Targeted *in vivo* infection with a retroviral vector carrying the interleukin-3 (multi-CSF) gene leads to immortalization and leukemic transformation of primitive hematopoietic progenitor cells. *Growth Factors.* 9(1): 41-55.

Kadmon, G., Eckert, M., Sammar, M., Schachner, M., and Altevogt, P. (1992). Nectadrin, the heat-stable antigen, is a cell adhesion molecule. *J. Cell Biol.* 118(5): 1245-1258.

Kanei-Ishii, C., Nomura, T., Tanikawa, J., Ichikawa-Iwata, E., and Ishii, S. (2004). Differential sensitivity of v-Myb and c-Myb to Wnt-1-induced protein degradation. *J. Biol. Chem.* 279(43): 44582-44589.

Katsumoto, T. R., Duda, J., Kim, A., Wardak, Z., Dranoff, G., Clapp, D. W., and Shannon, K. (2005). Granulocyte/macrophage colony-stimulating factor and accessory cells modulate radioprotection by purified hematopoietic cells. *J. Exp. Med.* 201(6): 853-858.

Kaushansky, K., Lin, N., Grossmann, A., Humes, J., Sprugel, K. H., and Broudy, V. C. (1996). Thrombopoietin expands erythroid, granulocyte-macrophage, and megakaryocytic progenitor cells in normal and myelosuppressed mice. *Exp. Hematol.* 24(2): 265-269.

Kay, R., Rosten, P. M., and Humphries, R. K. (1991). CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. *J. Immunol.* 147(4): 1412-1416.

Kay, R., Takei, F., and Humphries, R. K. (1990). Expression cloning of a cDNA encoding M1/69-J11d heat-stable antigens. *J. Immunol.* 145(6): 1952-1959.

Kelly, J. A., Spolski, R., Kovanen, P. E., Suzuki, T., Bollenbacher, J., Pise-Masison, C. A., Radonovich, M. F., Lee, S., Jenkins, N. A., Copeland, N. G., Morse, H. C., 3rd, and Leonard, W. J. (2003). Stat5 synergizes with T cell receptor/antigen stimulation in the development of lymphoblastic lymphoma. *J. Exp. Med.* 198(1): 79-89.

Kelly, L. M., Liu, Q., Kutok, J. L., Williams, I. R., Boulton, C. L., and Gilliland, D. G. (2002). FLT3 internal tandem duplication mutations associated with human acute myeloid leukemias induce myeloproliferative disease in a murine bone marrow transplant model. *Blood.* 99(1): 310-318.

Khanna-Gupta, A., Zibello, T., Simkevich, C., Rosmarin, A. G., and Berliner, N. (2000). Sp1 and C/EBP are necessary to activate the lactoferrin gene promoter during myeloid differentiation. *Blood.* 95(12): 3734-3741.

Khanna-Gupta, A., Zibello, T., Sun, H., Gaines, P., and Berliner, N. (2003). Chromatin immunoprecipitation (ChIP) studies indicate a role for CCAAT enhancer binding proteins alpha and epsilon (C/EBP alpha and C/EBP epsilon) and CDP/cut in myeloid maturation-induced lactoferrin gene expression. *Blood.* 101(9): 3460-3468.

Kim, Y., Ratziu, V., Choi, S. G., Lalazar, A., Theiss, G., Dang, Q., Kim, S. J., and Friedman, S. L. (1998a). Transcriptional activation of transforming growth factor beta1 and its receptors by the Kruppel-like factor Zf9/core promoter-binding protein and Sp1. Potential mechanisms for autocrine fibrogenesis in response to injury. *J. Biol. Chem.* 273(50): 33750-33758.

Kim, Y. H., Choi, C. Y., Lee, S. J., Conti, M. A., and Kim, Y. (1998b). Homeodomain-interacting protein kinases, a novel family of co-repressors for homeodomain transcription factors. *J. Biol. Chem.* 273(40): 25875-25879.

Kimura, A., and Sultana, T. A. (2004). Granulocyte colony-stimulating factor receptors on CD34++ cells in patients with myelodysplastic syndrome (MDS) and MDS-acute myeloid leukemia. *Leuk. Lymphoma.* 45(10): 1995-2000.

Kitajima, K., Masuhara, M., Era, T., Enver, T., and Nakano, T. (2002). GATA-2 and GATA-2/ER display opposing activities in the development and differentiation of blood progenitors. *EMBO J.* 21(12): 3060-3069.

Kjeldsen, L., Cowland, J. B., and Borregaard, N. (2000). Human neutrophil gelatinase-associated lipocalin and homologous proteins in rat and mouse. *Biochim. Biophys. Acta.* 1482(1-2): 272-283.

Klempnauer, K. H., Gonda, T. J., and Bishop, J. M. (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell.* 31(2 Pt 1): 453-463.

Kohlmann, A., Schoch, C., Schnittger, S., Dugas, M., Hiddemann, W., Kern, W., and Haferlach, T. (2003). Molecular characterization of acute leukemias by use of microarray technology. *Genes. Chromosomes Cancer.* 37(4): 396-405.

Kojima, S., Hayashi, S., Shimokado, K., Suzuki, Y., Shimada, J., Crippa, M. P., and Friedman, S. L. (2000). Transcriptional activation of urokinase by the Kruppel-like factor Zf9/COPEB activates latent TGF-beta1 in vascular endothelial cells. *Blood.* 95(4): 1309-1316.

Kolbus, A., Blazquez-Domingo, M., Carotta, S., Bakker, W., Luedemann, S., von Lindern, M., Steinlein, P., and Beug, H. (2003). Cooperative signaling between cytokine receptors and the glucocorticoid receptor in the expansion of erythroid progenitors: molecular analysis by expression profiling. *Blood.* 102(9): 3136-3146.

Kondo, T., Minamino, N., Nagamura-Inoue, T., Matsumoto, M., Taniguchi, T., and Tanaka, N. (1997). Identification and characterization of nucleophosmin/B23/numatrin which binds the anti-oncogenic transcription factor IRF-1 and manifests oncogenic activity. *Oncogene.* 15(11): 1275-1281.

Kopf, M., Brombacher, F., Hodgkin, P. D., Ramsay, A. J., Milbourne, E. A., Dai, W. J., Ovington, K. S., Behm, C. A., Kohler, G., Young, I. G., and Matthaei, K. I. (1996). IL-5-deficient mice have a developmental defect in CD5+ B-1 cells and lack eosinophilia but have normal antibody and cytotoxic T cell responses. *Immunity.* 4(1): 15-24.

Koritschoner, N. P., Bocco, J. L., Panzetta-Dutari, G. M., Dumur, C. I., Flury, A., and Patrito, L. C. (1997). A novel human zinc finger protein that interacts with the core promoter element of a TATA box-less gene. *J. Biol. Chem.* 272(14): 9573-9580.

Koschmieder, S., Gottgens, B., Zhang, P., Iwasaki-Arai, J., Akashi, K., Kutok, J. L., Dayaram, T., Geary, K., Green, A. R., Tenen, D. G., and Huettner, C. S. (2005). Inducible chronic phase of myeloid leukemia with expansion of hematopoietic stem cells in a transgenic model of BCR-ABL leukemogenesis. *Blood.* 105(1): 324-334.

Kralovics, R., Passamonti, F., Buser, A. S., Teo, S. S., Tiedt, R., Passweg, J. R., Tichelli, A., Cazzola, M., and Skoda, R. C. (2005a). A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N. Engl. J. Med.* 352(17): 1779-1790.

Kralovics, R., Teo, S. S., Buser, A. S., Brutsche, M., Tiedt, R., Tichelli, A., Passamonti, F., Pietra, D., Cazzola, M., and Skoda, R. C. (2005b). Altered gene expression in myeloproliferative disorders correlates with activation of signaling by the V617F mutation of Jak2. *Blood.*

Krause, D. S., Fackler, M. J., Civin, C. I., and May, W. S. (1996). CD34: structure, biology, and clinical utility. *Blood.* 87(1): 1-13.

Kristiansen, G., Denkert, C., Schluns, K., Dahl, E., Pilarsky, C., and Hauptmann, S. (2002). CD24 is expressed in ovarian cancer and is a new independent prognostic marker of patient survival. *Am. J. Pathol.* 161(4): 1215-1221.

Kristiansen, G., Pilarsky, C., Pervan, J., Sturzebecher, B., Stephan, C., Jung, K., Loening, S., Rosenthal, A., and Dietel, M. (2004). CD24 expression is a significant predictor of PSA relapse and poor prognosis in low grade or organ confined prostate cancer. *Prostate.* 58(2): 183-192.

Kristiansen, G., Schluns, K., Yongwei, Y., Denkert, C., Dietel, M., and Petersen, I. (2003a). CD24 is an independent prognostic marker of survival in nonsmall cell lung cancer patients. *Br. J. Cancer.* 88(2): 231-236.

Kristiansen, G., Winzer, K. J., Mayordomo, E., Bellach, J., Schluns, K., Denkert, C., Dahl, E., Pilarsky, C., Altevogt, P., Guski, H., and Dietel, M. (2003b). CD24 expression is a new prognostic marker in breast cancer. *Clin. Cancer Res.* 9(13): 4906-4913.

Kuby, J. (1997a). Cancer and the Immune System. In *Immunolgy* (Allen, D., ed), 3rd Ed., pp. 573-596, W. H. Freeman and Company, New York, USA.

Kuby, J. (1997b). Cytokines. In *Immunolgy* (Allen, D., ed), 3rd Ed., pp. 313-334, W. H. Freeman and Company, New York, USA.

Lang, R. A., Metcalf, D., Cuthbertson, R. A., Lyons, I., Stanley, E., Kelso, A., Kannourakis, G., Williamson, D. J., Klintworth, G. K., Gonda, T. J., and Dunn, A. R. (1987). Transgenic mice expressing a hemopoietic growth factor gene (GM-CSF) develop accumulations of macrophages, blindness, and a fatal syndrome of tissue damage. *Cell.* 51: 675-686.

Lang, R. A., Metcalf, D., Gough, N. M., Dunn, A. R., and Gonda, T. J. (1985). Expression of a hemopoietic growth factor cDNA in a factor-dependent cell line results in autonomous growth and tumorigenicity. *Cell.* 43(2 Pt 1): 531-542.

Lantz, C. S., Boesiger, J., Song, C. H., Mach, N., Kobayashi, T., Mulligan, R. C., Nawa, Y., Dranoff, G., and Galli, S. J. (1998). Role for interleukin-3 in mast-cell and basophil development and in immunity to parasites. *Nature.* 392(6671): 90-93.

Lanza, F., Castoldi, G. L., Castagnari, B., Todd, R. F., 3rd, Moretti, S., Spisani, S., Latorraca, A., Focarile, E., Roberti, M. G., and Traniello, S. (1998). Expression and functional role of urokinase-type plasminogen activator receptor in normal and acute leukaemic cells. *Br. J. Haematol.* 103(1): 110-123.

Lavau, C., Marchio, A., Fagioli, M., Jansen, J., Falini, B., Lebon, P., Grosveld, F., Pandolfi, P. P., Pelicci, P. G., and Dejean, A. (1995). The acute promyelocytic leukaemia-associated PML gene is induced by interferon. *Oncogene.* 11(5): 871-876.

Lee, K. H., Chang, M. Y., Ahn, J. I., Yu, D. H., Jung, S. S., Choi, J. H., Noh, Y. H., Lee, Y. S., and Ahn, M. J. (2002). Differential gene expression in retinoic acid-induced differentiation of acute promyelocytic leukemia cells, NB4 and HL-60 cells. *Biochem. Biophys. Res. Commun.* 296(5): 1125-1133.

Lehming, N., Le Saux, A., Schuller, J., and Ptashne, M. (1998). Chromatin components as part of a putative transcriptional repressing complex. *Proc. Natl. Acad. Sci. U. S. A.* 95(13): 7322-7326.

LeVine, A. M., Reed, J. A., Kurak, K. E., Cianciolo, E., and Whitsett, J. A. (1999). GM-CSF-deficient mice are susceptible to pulmonary group B streptococcal infection. *J. Clin. Invest.* 103(4): 563-569.

Levine, R. L., Wadleigh, M., Cools, J., Ebert, B. L., Wernig, G., Huntly, B. J., Boggon, T. J., Wlodarska, I., Clark, J. J., Moore, S., Adelsperger, J., Koo, S., Lee, J. C., Gabriel, S., Mercher, T., D'Andrea, A., Frohling, S., Dohner, K., Marynen, P., Vandenberghe, P., Mesa, R. A., Tefferi, A., Griffin, J. D., Eck, M. J., Sellers, W. R., Meyerson, M., Golub, T. R., Lee, S. J., and Gilliland, D. G. (2005). Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell.* 7(4): 387-397.

Lewandoski, M. (2001). Conditional control of gene expression in the mouse. *Nat. Rev. Genet.* 2(10): 743-755.

Li, H., Leo, C., Zhu, J., Wu, X., O'Neil, J., Park, E. J., and Chen, J. D. (2000). Sequestration and inhibition of Daxx-mediated transcriptional repression by PML. *Mol. Cell. Biol.* 20(5): 1784-1796.

Lian, Z., Wang, L., Yamaga, S., Bonds, W., Beazer-Barclay, Y., Kluger, Y., Gerstein, M., Newburger, P. E., Berliner, N., and Weissman, S. M. (2001). Genomic and proteomic analysis of the myeloid differentiation program. *Blood.* 98(3): 513-524.

Lieschke, G. J., Grail, D., Hodgson, G., Metcalf, D., Stanley, E., Cheers, C., Fowler, K. J., Basu, S., Zhan, Y. F., and Dunn, A. R. (1994). Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. *Blood.* 84(6): 1737-1746.

Lim, I. K., Lee, M. S., Lee, S. H., Kim, N. K., Jou, I., Seo, J. S., and Park, S. C. (1995). Differential expression of TIS21 and TIS1 genes in the various organs of Balb/c mice, thymic carcinoma tissues and human cancer cell lines. *J. Cancer Res. Clin. Oncol.* 121(5): 279-284.

Lin, R. J., Egan, D. A., and Evans, R. M. (1999). Molecular genetics of acute promyelocytic leukemia. *Trends Genet.* 15(5): 179-184.

Lin, R. J., Nagy, L., Inoue, S., Shao, W., Miller, W. H., Jr., and Evans, R. M. (1998). Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature.* 391(6669): 811-814.

Lin, W. J., Gary, J. D., Yang, M. C., Clarke, S., and Herschman, H. R. (1996). The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J. Biol. Chem.* 271(25): 15034-15044.

Lipshutz, R. J., Fodor, S. P., Gingeras, T. R., and Lockhart, D. J. (1999). High density synthetic oligonucleotide arrays. *Nat. Genet.* 21(1 Suppl): 20-24.

Littlewood, T. D., Hancock, D. C., Danielian, P. S., Parker, M. G., and Evan, G. I. (1995). A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res.* 23(10): 1686-1690.

Liu, F., Wu, H. Y., Wesselschmidt, R., Kornaga, T., and Link, D. C. (1996). Impaired production and increased apoptosis of neutrophils in granulocyte colony-stimulating factor receptor-deficient mice. *Immunity.* 5(5): 491-501.

Liu, S. J., Cai, Z. W., Liu, Y. J., Dong, M. Y., Sun, L. Q., Hu, G. F., Wei, Y. Y., and Lao, W. D. (2004). Role of nucleostemin in growth regulation of gastric cancer, liver cancer and other malignancies. *World J. Gastroenterol.* 10(9): 1246-1249.

Loeb, L. A. (1998). Cancer cells exhibit a mutator phenotype. *Adv. Cancer Res.* 72: 25-56.

Longmore, G. D., and Lodish, H. F. (1991). An activating mutation in the murine erythropoietin receptor induces erythroleukemia in mice: a cytokine receptor superfamily oncogene. *Cell.* 67(6): 1089-1102.

Longmore, G. D., Pharr, P., and Lodish, H. F. (1992). Mutation in murine erythropoietin receptor induces erythropoietin-independent erythroid proliferation *in vitro*, polycythemia *in vivo*. *Leukemia.* 6 Suppl 3: 130S-134S.

Longmore, G. D., Pharr, P., Neumann, D., and Lodish, H. F. (1993). Both megakaryocytopoiesis and erythropoiesis are induced in mice infected with a retrovirus expressing an oncogenic erythropoietin receptor. *Blood.* 82(8): 2386-2395.

Longmore, G. D., You, Y., Molden, J., Liu, K. D., Mikami, A., Lai, S. Y., Pharr, P., and Goldsmith, M. A. (1998). Redundant and selective roles for erythropoietin receptor tyrosines in erythropoiesis *in vivo*. *Blood.* 91(3): 870-878.

Lonnstedt, I., and Speed, T. P. (2002). Replicated Microarray Data. *Statistical Sinica.* 12(1): 31-46.

Look, A. T. (1997). Oncogenic transcription factors in the human acute leukemias. *Science.* 278(5340): 1059-1064.

Lowenberg, B., Downing, J. R., and Burnett, A. (1999). Acute myeloid leukemia. *N. Engl. J. Med.* 341(14): 1051-1062.

Lund, L. R., Romer, J., Ronne, E., Ellis, V., Blasi, F., and Dano, K. (1991). Urokinase-receptor biosynthesis, mRNA level and gene transcription are increased by transforming growth factor beta 1 in human A549 lung carcinoma cells. *EMBO J.* 10(11): 3399-3407.

Lyng, H., Badiee, A., Svendsrud, D. H., Hovig, E., Myklebost, O., and Stokke, T. (2004). Profound influence of microarray scanner characteristics on gene expression ratios: analysis and procedure for correction. *BMC Genomics.* 5(1): 10.

Mackarehtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P., and Lemischka, I. R. (1995). Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity.* 3(1): 147-161.

Magee, T., and Marshall, C. (1999). New insights into the interaction of Ras with the plasma membrane. *Cell.* 98(1): 9-12.

Mahon, F. X. (2005). JAK the trigger. *Oncogene.*

Marcu, K. B., Bossone, S. A., and Patel, A. J. (1992). myc function and regulation. *Annu. Rev. Biochem.* 61: 809-860.

Markoff, A., Kuryshev, V., Vorobyov, E., Bogdanova, N., Rescher, U., Goebeler, V., Kondrashov, A., and Gerke, V. (2002). Structure and Expression of the Murine Annexin A9 Gene. *Genome Letters.* 1(4): 189-197.

Martinez-Moczygemba, M., and Huston, D. P. (2003). Biology of common beta receptor-signaling cytokines: IL-3, IL-5, and GM-CSF. *J. Allergy Clin. Immunol.* 112(4): 653-665; quiz 666.

Matsuda, S., Rouault, J., Magaud, J., and Berthet, C. (2001). In search of a function for the TIS21/PC3/BTG1/TOB family. *FEBS Lett.* 497(2-3): 67-72.

Matsumura, I., Kanakura, Y., Kato, T., Ikeda, H., Ishikawa, J., Horikawa, Y., Hashimoto, K., Moriyama, Y., Tsujimura, T., Nishiura, T., and *et al*. (1995). Growth response of acute myeloblastic leukemia cells to recombinant human thrombopoietin. *Blood.* 86(2): 703-709.

McCormack, M. P. (1998). The biological effects of constitutively active mutants of the common beta subunit of the human IL-3, IL-5 and GM-CSF receptors. University of Adelaide.

McCormack, M. P., and Gonda, T. J. (1997). Expression of activated mutants of the human interleukin-3/interleukin- 5/granulocyte-macrophage colony-stimulating factor receptor common beta subunit in primary hematopoietic cells induces factor-independent proliferation and differentiation. *Blood.* 90(4): 1471-1481.

McCormack, M. P., and Gonda, T. J. (1999). Myeloproliferative disorder and leukaemia in mice induced by different classes of constitutive mutants of the human IL-3/IL-5/GM-CSF receptor common beta subunit. *Oncogene.* 18(51): 7190-7199.

McCormack, M. P., and Gonda, T. J. (2000). Novel murine myeloid cell lines that exhibit a differentiation switch in response to IL-3 or GM-CSF, or to different constitutively active mutants of the GM-CSF receptor beta subunit. *Blood.* 95(1): 120-127.

McKinstry, W. J., Li, C. L., Rasko, J. E., Nicola, N. A., Johnson, G. R., and Metcalf, D. (1997). Cytokine receptor expression on hematopoietic stem and progenitor cells. *Blood.* 89(1): 65-71.

Medina-Acosta, E., and Cross, G. A. (1993). Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. *Mol. Biochem. Parasitol.* 59(2): 327-329.

Mello, C. C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. *Nature.* 431(7006): 338-342.

Melnick, A., and Licht, J. D. (1999). Deconstructing a disease: RARalpha, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia. *Blood.* 93(10): 3167-3215.

Memarzadeh, S., Kozak, K. R., Chang, L., Natarajan, S., Shintaku, P., Reddy, S. T., and Farias-Eisner, R. (2002). Urokinase plasminogen activator receptor: Prognostic biomarker for endometrial cancer. *Proc. Natl. Acad. Sci. U. S. A.* 99(16): 10647-10652.

Metcalf, D. (2003). The unsolved enigmas of leukemia inhibitory factor. *Stem Cells.* 21(1): 5-14.

Metcalf, D., and Nicola, N. A. (1995). *The Hemopoietic Colony-Stimulating Factors: From Biology to Clinical Applications.* Cambridge University Press, Cambridge.

Mikkers, H., and Berns, A. (2003). Retroviral insertional mutagenesis: tagging cancer pathways. *Adv. Cancer Res.* 88: 53-99.

Miller, A. D., and Rosman, G. J. (1989). Improved retroviral vectors for gene transfer and expression. *Biotechniques.* 7(9): 980-982, 984-986, 989-990.

Millward, T. A., Zolnierowicz, S., and Hemmings, B. A. (1999). Regulation of protein kinase cascades by protein phosphatase 2A. *Trends Biochem. Sci.* 24(5): 186-191.

Minden, M. (1995). Growth factor requirements for normal and leukemic cells. *Semin. Hematol.* 32(3): 162-182.

Missiaglia, E., Blaveri, E., Terris, B., Wang, Y. H., Costello, E., Neoptolemos, J. P., Crnogorac-Jurcevic, T., and Lemoine, N. R. (2004). Analysis of gene expression in cancer cell lines identifies candidate markers for pancreatic tumorigenesis and metastasis. *Int. J. Cancer.* 112(1): 100-112.

Miyajima, A., Mui, A. L., Ogorochi, T., and Sakamaki, K. (1993). Receptors for granulocyte-macrophage colony-stimulating factor, interleukin-3, and interleukin-5. *Blood.* 82(7): 1960-1974.

Miyake, H., Hara, I., Yamanaka, K., Gohji, K., Arakawa, S., and Kamidono, S. (1999). Elevation of serum levels of urokinase-type plasminogen activator and its receptor is associated with disease progression and prognosis in patients with prostate cancer. *Prostate.* 39(2): 123-129.

Miyamoto, T., Iwasaki, H., Reizis, B., Ye, M., Graf, T., Weissman, I. L., and Akashi, K. (2002). Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev. Cell.* 3(1): 137-147.

Model, F., Adorjan, P., Olek, A., and Piepenbrock, C. (2001). Feature selection for DNA methylation based cancer classification. *Bioinformatics.* 17 Suppl 1: S157-164.

Montuori, N., Mattiello, A., Mancini, A., Santoli, M., Taglialatela, P., Caputi, M., Rossi, G., and Ragno, P. (2001). Urokinase-type plasminogen activator up-regulates the expression of its cellular receptor through a post-transcriptional mechanism. *FEBS Lett.* 508(3): 379-384.

Moodie, S. A., and Wolfman, A. (1994). The 3Rs of life: Ras, Raf and growth regulation. *Trends Genet.* 10(2): 44-48.

Moore, D. S., and McCabe, G. P. (1998). *Introduction to the practice of statistics*, 3rd Ed. W.H. Freeman and Company, New York, USA.

Moore, M. A. (2005). Converging pathways in leukemogenesis and stem cell self-renewal. *Exp. Hematol.* 33(7): 719-737.

Morris, E. C., Dafforn, T. R., Forsyth, S. L., Missen, M. A., Horvath, A. J., Hampson, L., Hampson, I. N., Currie, G., Carrell, R. W., and Coughlin, P. B. (2003). Murine serpin 2A is a redox-sensitive intracellular protein. *Biochem. J.* 371(Pt 1): 165-173.

Morris, S. W., Kirstein, M. N., Valentine, M. B., Dittmer, K. G., Shapiro, D. N., Saltman, D. L., and Look, A. T. (1994). Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science.* 263(5151): 1281-1284.

Morris, S. W., Naeve, C., Mathew, P., James, P. L., Kirstein, M. N., Cui, X., and Witte, D. P. (1997). ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). *Oncogene.* 14(18): 2175-2188.

Mui, A. L., Wakao, H., Kinoshita, T., Kitamura, T., and Miyajima, A. (1996). Suppression of interleukin-3-induced gene expression by a C-terminal truncated Stat5: role of Stat5 in proliferation. *EMBO J.* 15(10): 2425-2433.

Mui, A. L., Wakao, H., O'Farrell, A. M., Harada, N., and Miyajima, A. (1995). Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin-5 transduce signals through two STAT5 homologs. *EMBO J.* 14(6): 1166-1175.

Murakami, M., Narazaki, M., Hibi, M., Yawata, H., Yasukawa, K., Hamaguchi, M., Taga, T., and Kishimoto, T. (1991). Critical cytoplasmic region of the interleukin 6 signal transducer gp130 is conserved in the cytokine receptor family. *Proc. Natl. Acad. Sci. U. S. A.* 88(24): 11349-11353.

Mustjoki, S., Alitalo, R., Stephens, R. W., and Vaheri, A. (1999). Blast cell-surface and plasma soluble urokinase receptor in acute leukemia patients: relationship to classification and response to therapy. *Thromb. Haemost.* 81(5): 705-710.

Nakao, M., Yokota, S., Iwai, T., Kaneko, H., Horiike, S., Kashima, K., Sonoda, Y., Fujimoto, T., and Misawa, S. (1996). Internal tandem duplication of the flt3 gene found in acute myeloid leukemia. *Leukemia.* 10(12): 1911-1918.

Nandan, M. O., Yoon, H. S., Zhao, W., Ouko, L. A., Chanchevalap, S., and Yang, V. W. (2004). Kruppel-like factor 5 mediates the transforming activity of oncogenic H-Ras. *Oncogene.* 23(19): 3404-3413.

Narla, G., Difeo, A., Reeves, H. L., Schaid, D. J., Hirshfeld, J., Hod, E., Katz, A., Isaacs, W. B., Hebbring, S., Komiya, A., McDonnell, S. K., Wiley, K. E., Jacobsen, S. J., Isaacs, S. D., Walsh, P. C., Zheng, S. L., Chang, B. L., Friedrichsen, D. M., Stanford, J. L., Ostrander, E. A., Chinnaiyan, A. M., Rubin, M. A., Xu, J., Thibodeau, S. N., Friedman, S. L., and Martignetti, J. A. (2005a). A germline DNA polymorphism enhances alternative splicing of the KLF6 tumor suppressor gene and is associated with increased prostate cancer risk. *Cancer Res.* 65(4): 1213-1222.

Narla, G., DiFeo, A., Yao, S., Banno, A., Hod, E., Reeves, H. L., Qiao, R. F., Camacho-Vanegas, O., Levine, A., Kirschenbaum, A., Chan, A. M., Friedman, S. L., and Martignetti, J. A. (2005b). Targeted inhibition of the KLF6 splice variant, KLF6 SV1, suppresses prostate cancer cell growth and spread. *Cancer Res.* 65(13): 5761-5768.

Narla, G., Heath, K. E., Reeves, H. L., Li, D., Giono, L. E., Kimmelman, A. C., Glucksman, M. J., Narla, J., Eng, F. J., Chan, A. M., Ferrari, A. C., Martignetti, J. A., and Friedman, S. L. (2001). KLF6, a candidate tumor suppressor gene mutated in prostate cancer. *Science.* 294(5551): 2563-2566.

Natarajan, K., Meyer, M. R., Jackson, B. M., Slade, D., Roberts, C., Hinnebusch, A. G., and Marton, M. J. (2001). Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. *Mol. Cell. Biol.* 21(13): 4347-4368.

Neil, J. C., and Cameron, E. R. (2002). Retroviral insertion sites and cancer: fountain of all knowledge? *Cancer Cell.* 2(4): 253-255.

Newton, M. A., Kendziorski, C. M., Richmond, C. S., Blattner, F. R., and Tsui, K. W. (2001). On differential variability of expression ratios: improving statistical inference about gene expression changes from microarray data. *J. Comput. Biol.* 8(1): 37-52.

Nguyen, H. Q., Hoffman-Liebermann, B., and Liebermann, D. A. (1993). The zinc finger transcription factor Egr-1 is essential for and restricts differentiation along the macrophage lineage. *Cell.* 72(2): 197-209.

Nicola, N. A. (1994). *Guidebook to Cytokines and Their Receptors.* Oxford University Press, Oxford, UK.

Nicola, N. A., and Metcalf, D. (1985). Binding of 125I-labeled granulocyte colony-stimulating factor to normal murine hemopoietic cells. *J. Cell. Physiol.* 124(2): 313-321.

Nicola, N. A., Robb, L., Metcalf, D., Cary, D., Drinkwater, C. C., and Begley, C. G. (1996). Functional inactivation in mice of the gene for the interleukin-3 (IL-3)-specific receptor beta-chain: implications for IL-3 function and the mechanism of receptor transmodulation in hematopoietic cells. *Blood.* 87(7): 2665-2674.

Nielsen, P. J., Lorenz, B., Muller, A. M., Wenger, R. H., Brombacher, F., Simon, M., von der Weid, T., Langhorne, W. J., Mossmann, H., and Kohler, G. (1997). Altered erythrocytes and a leaky block in B-cell development in CD24/HSA-deficient mice. *Blood.* 89(3): 1058-1067.

Nishinakamura, R., Miyajima, A., Mee, P. J., Tybulewicz, V. L., and Murray, R. (1996a). Hematopoiesis in mice lacking the entire granulocyte-macrophage colony-stimulating factor/interleukin-3/interleukin-5 functions. *Blood.* 88(7): 2458-2464.

Nishinakamura, R., Nakayama, N., Hirabayashi, Y., Inoue, T., Aud, D., McNeil, T., Azuma, S., Yoshida, S., Toyoda, Y., Arai, K., and *et al*. (1995). Mice deficient for the IL-3/GM-CSF/IL-5 beta c receptor exhibit lung pathology and impaired immune response, while beta IL3 receptor-deficient mice are normal. *Immunity.* 2(3): 211-222.

Nishinakamura, R., Wiler, R., Dirksen, U., Morikawa, Y., Arai, K., Miyajima, A., Burdach, S., and Murray, R. (1996b). The pulmonary alveolar proteinosis in granulocyte macrophage colony-stimulating factor/interleukins 3/5 beta c receptor-deficient mice is reversed by bone marrow transplantation. *J. Exp. Med.* 183(6): 2657-2662.

Noguchi, Y., Wada, H., Marino, M. W., and Old, L. J. (1998). Regulation of IFN-gamma production in granulocyte-macrophage colony-stimulating factor-deficient mice. *Eur. J. Immunol.* 28(12): 3980-3988.

Nomura, T., Tanikawa, J., Akimaru, H., Kanei-Ishii, C., Ichikawa-Iwata, E., Khan, M. M., Ito, H., and Ishii, S. (2004). Oncogenic activation of c-Myb correlates with a loss of negative regulation by TIF1beta and Ski. *J. Biol. Chem.* 279(16): 16715-16726.

Noonan, F. C., Mutch, D. G., Ann Mallon, M., and Goodfellow, P. J. (2001). Characterization of the homeodomain gene EMX2: sequence conservation, expression analysis, and a search for mutations in endometrial cancers. *Genomics.* 76(1-3): 37-44.

Nyga, R., Pecquet, C., Harir, N., Gu, H., Dhennin-Duthille, I., Regnier, A., Gouilleux-Gruart, V., Lassoued, K., and Gouilleux, F. (2005). Activated STAT5 proteins induce activation of the PI 3-kinase/Akt and Ras/MAPK pathways via the Gab2 scaffolding adapter. *Biochem. J.* 390(Pt 1): 359-366.

Ogilvy, S., Metcalf, D., Gibson, L., Bath, M. L., Harris, A. W., and Adams, J. M. (1999). Promoter elements of vav drive transgene expression *in vivo* throughout the hematopoietic compartment. *Blood.* 94(6): 1855-1863.

Okano, J., Opitz, O. G., Nakagawa, H., Jenkins, T. D., Friedman, S. L., and Rustgi, A. K. (2000). The Kruppel-like transcriptional factors Zf9 and GKLF coactivate the human keratin 4 promoter and physically interact. *FEBS Lett.* 473(1): 95-100.

Okuda, K., Foster, R., and Griffin, J. D. (1999). Signaling domains of the beta c chain of the GM-CSF/IL-3/IL-5 receptor. *Ann. N. Y. Acad. Sci.* 872: 305-312.

Okuda, K., Smith, L., Griffin, J. D., and Foster, R. (1997). Signaling functions of the tyrosine residues in the betac chain of the granulocyte-macrophage colony-stimulating factor receptor. *Blood.* 90(12): 4759-4766.

Okuda, M., Horn, H. F., Tarapore, P., Tokuyama, Y., Smulian, A. G., Chan, P. K., Knudsen, E. S., Hofmann, I. A., Snyder, J. D., Bove, K. E., and Fukasawa, K. (2000). Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell.* 103(1): 127-140.

Okuno, Y., Huettner, C. S., Radomska, H. S., Petkova, V., Iwasaki, H., Akashi, K., and Tenen, D. G. (2002). Distal elements are critical for human CD34 expression *in vivo*. *Blood.* 100(13): 4420-4426.

Ooi, J., Tojo, A., Asano, S., Sato, Y., and Oka, Y. (1998). Thrombopoietin induces tyrosine phosphorylation of a common beta subunit of GM-CSF receptor and its association with Stat5 in TF-1/TPO cells. *Biochem. Biophys. Res. Commun.* 246(1): 132-136.

Orkin, S. H. (2000). Diversification of haematopoietic stem cells to specific lineages. *Nat. Rev. Genet.* 1(1): 57-64.

O'Shea, J. J. (1997). Jaks, STATs, cytokine signal transduction, and immunoregulation: are we there yet? *Immunity.* 7(1): 1-11.

Pagliarini, D. J., Wiley, S. E., Kimple, M. E., Dixon, J. R., Kelly, P., Worby, C. A., Casey, P. J., and Dixon, J. E. (2005). Involvement of a mitochondrial phosphatase in the regulation of ATP production and insulin secretion in pancreatic beta cells. *Mol. Cell.* 19(2): 197-207.

Pagliarini, D. J., Worby, C. A., and Dixon, J. E. (2004). A PTEN-like phosphatase with a novel substrate specificity. *J. Biol. Chem.* 279(37): 38590-38596.

Paine, R., 3rd, Preston, A. M., Wilcoxen, S., Jin, H., Siu, B. B., Morris, S. B., Reed, J. A., Ross, G., Whitsett, J. A., and Beck, J. M. (2000). Granulocyte-macrophage colony-stimulating factor in the innate immune response to Pneumocystis carinii pneumonia in mice. *J. Immunol.* 164(5): 2602-2609.

Pandolfi, P. P. (2001). Oncogenes and tumor suppressors in the molecular pathogenesis of acute promyelocytic leukemia. *Hum. Mol. Genet.* 10(7): 769-775.

Pang, A. L., Taylor, H. C., Johnson, W., Alexander, S., Chen, Y., Su, Y. A., Li, X., Ravindranath, N., Dym, M., Rennert, O. M., and Chan, W. Y. (2003). Identification of differentially expressed genes in mouse spermatogenesis. *J. Androl.* 24(6): 899-911.

Pappot, H. (1999). The plasminogen activation system in lung cancer--with special reference to the prognostic role in "non-small cell lung cancer". *APMIS. Suppl.* 92: 1-29.

Parganas, E., Wang, D., Stravopodis, D., Topham, D. J., Marine, J. C., Teglund, S., Vanin, E. F., Bodner, S., Colamonici, O. R., van Deursen, J. M., Grosveld, G., and Ihle, J. N. (1998). Jak2 is essential for signaling through a variety of cytokine receptors. *Cell.* 93(3): 385-395.

Pawson, T., and Scott, J. D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. *Science.* 278(5346): 2075-2080.

Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993). Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci. U. S. A.* 90(18): 8392-8396.

Penciolelli, J. F., Wendling, F., Robert-Lezenes, J., Barque, J. P., Tambourin, P., and Gisselbrecht, S. (1987). Genetic analysis of myeloproliferative leukemia virus, a novel acute leukemogenic replication-defective retrovirus. *J. Virol.* 61(2): 579-583.

Perri, R. T., Royston, I., LeBien, T. W., and Kay, N. E. (1983). Chronic lymphocytic leukemia progenitor cells carry the antigens T65, BA-1, and Ia. *Blood.* 61(5): 871-875.

Perron, B., Lewit-Bentley, A., Geny, B., and Russo-Marie, F. (1997). Can enzymatic activity, or otherwise, be inferred from structural studies of annexin III? *J. Biol. Chem.* 272(17): 11321-11326.

Persons, D. A., Allay, J. A., Allay, E. R., Ashmun, R. A., Orlic, D., Jane, S. M., Cunningham, J. M., and Nienhuis, A. W. (1999). Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood.* 93(2): 488-499.

Persons, D. A., Mehaffey, M. G., Kaleko, M., Nienhuis, A. W., and Vanin, E. F. (1998). An improved method for generating retroviral producer clones for vectors lacking a selectable marker gene. *Blood Cells. Mol. Dis.* 24(2): 167-182.

Pierres, M., Naquet, P., Barbet, J., Marchetto, S., Marics, I., Devaux, C., Barad, M., Hyman, R., and Rougon, G. (1987). Evidence that murine hematopoietic cell subset marker J11d is attached to a glycosyl-phosphatidylinositol membrane anchor. *Eur. J. Immunol.* 17(12): 1781-1785.

Plesner, T., Ralfkiaer, E., Wittrup, M., Johnsen, H., Pyke, C., Pedersen, T. L., Hansen, N. E., and Dano, K. (1994). Expression of the receptor for urokinase-type plasminogen activator in normal and neoplastic blood cells and hematopoietic tissue. *Am. J. Clin. Pathol.* 102(6): 835-841.

Prevot, D., Voeltzel, T., Birot, A. M., Morel, A. P., Rostan, M. C., Magaud, J. P., and Corbo, L. (2000). The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. *J. Biol. Chem.* 275(1): 147-153.

Quackenbush, J. (2001). Computational analysis of microarray data. *Nat. Rev. Genet.* 2(6): 418-427.

Quackenbush, J. (2002). Microarray data normalization and transformation. *Nat. Genet.* 32 Suppl: 496-501.

Quelle, F. W., Sato, N., Witthuhn, B. A., Inhorn, R. C., Eder, M., Miyajima, A., Griffin, J. D., and Ihle, J. N. (1994). JAK2 associates with the beta c chain of the receptor for granulocyte-macrophage colony-stimulating factor, and its activation requires the membrane-proximal region. *Mol. Cell. Biol.* 14(7): 4335-4341.

Raburn, D. J., Hamil, K. G., Tsuruta, J. K., O'Brien, D. A., and Hall, S. H. (1995). Stage-specific expression of B cell translocation gene 1 in rat testis. *Endocrinology.* 136(12): 5769-5777.

Radke, S., Austermann, J., Russo-Marie, F., Gerke, V., and Rescher, U. (2004). Specific association of annexin 1 with plasma membrane-resident and internalized EGF receptors mediated through the protein core domain. *FEBS Lett.* 578(1-2): 95-98.

Radomska, H. S., Gonzalez, D. A., Okuno, Y., Iwasaki, H., Nagy, A., Akashi, K., Tenen, D. G., and Huettner, C. S. (2002). Transgenic targeting with regulatory elements of the human CD34 gene. *Blood.* 100(13): 4410-4419.

Radomska, H. S., Huettner, C. S., Zhang, P., Cheng, T., Scadden, D. T., and Tenen, D. G. (1998). CCAAT/enhancer binding protein alpha is a regulatory switch sufficient for induction of granulocytic development from bipotential myeloid progenitors. *Mol. Cell. Biol.* 18(7): 4301-4314.

Raife, T. J., Lager, D. J., Kemp, J. D., and Dick, F. R. (1994). Expression of CD24 (BA-1) predicts monocytic lineage in acute myeloid leukemia. *Am. J. Clin. Pathol.* 101(3): 296-299.

Ramshaw, H. S., Bardy, P. G., Lee, M. A., and Lopez, A. F. (2002). Chronic myelomonocytic leukemia requires granulocyte-macrophage colony-stimulating factor for growth in vitro and in vivo. *Exp. Hematol.* 30(10): 1124-1131.

Randolph, T. R. (2005a). Chronic myelocytic leukemia--Part I: History, clinical presentation, and molecular biology. *Clin. Lab. Sci.* 18(1): 38-48.

Randolph, T. R. (2005b). Chronic myelocytic leukemia--Part II: Approaches to and molecular monitoring of therapy. *Clin. Lab. Sci.* 18(1): 49-56.

Rasko, J. E., Metcalf, D., Rossner, M. T., Begley, C. G., and Nicola, N. A. (1995). The flt3/flk-2 ligand: receptor distribution and action on murine haemopoietic cell survival and proliferation. *Leukemia.* 9(12): 2058-2066.

Ratziu, V., Lalazar, A., Wong, L., Dang, Q., Collins, C., Shaulian, E., Jensen, S., and Friedman, S. L. (1998). Zf9, a Kruppel-like transcription factor up-regulated *in vivo* during early hepatic fibrosis. *Proc. Natl. Acad. Sci. U. S. A.* 95(16): 9500-9505.

Rayner, J. R., and Gonda, T. J. (1994). A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. *Mol. Cell. Biol.* 14(2): 880-887.

Rebollo, A., and Schmitt, C. (2003). Ikaros, Aiolos and Helios: transcription regulators and lymphoid malignancies. *Immunol. Cell Biol.* 81(3): 171-175.

Rego, E. M., Wang, Z. G., Peruzzi, D., He, L. Z., Cordon-Cardo, C., and Pandolfi, P. P. (2001). Role of promyelocytic leukemia (PML) protein in tumor suppression. *J. Exp. Med.* 193(4): 521-529.

Reilly, J. T. (2003). Receptor tyrosine kinases in normal and malignant haematopoiesis. *Blood Rev.* 17(4): 241-248.

Rekhtman, N., Choe, K. S., Matushansky, I., Murray, S., Stopka, T., and Skoultchi, A. I. (2003). PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation. *Mol. Cell. Biol.* 23(21): 7460-7474.

Ren, R. (2005). Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. *Nat. Rev. Cancer.* 5(3): 172-183.

Rescher, U., and Gerke, V. (2004). Annexins--unique membrane binding proteins with diverse functions. *J. Cell Sci.* 117(Pt 13): 2631-2639.

Rimokh, R., Rouault, J. P., Wahbi, K., Gadoux, M., Lafage, M., Archimbaud, E., Charrin, C., Gentilhomme, O., Germain, D., Samarut, J., and *et al*. (1991). A chromosome 12 coding region is juxtaposed to the MYC protooncogene locus in a t(8;12)(q24;q22) translocation in a case of B-cell chronic lymphocytic leukemia. *Genes. Chromosomes Cancer.* 3(1): 24-36.

Robb, L., Drinkwater, C. C., Metcalf, D., Li, R., Kontgen, F., Nicola, N. A., and Begley, C. G. (1995). Hematopoietic and lung abnormalities in mice with a null mutation of the common beta subunit of the receptors for granulocyte-macrophage colony-stimulating factor and interleukins 3 and 5. *Proc. Natl. Acad. Sci. U. S. A.* 92(21): 9565-9569.

Roberts, A. W. (2005). G-CSF: a key regulator of neutrophil production, but that's not all! *Growth Factors.* 23(1): 33-41.

Roberts, C. J., Nelson, B., Marton, M. J., Stoughton, R., Meyer, M. R., Bennett, H. A., He, Y. D., Dai, H., Walker, W. L., Hughes, T. R., Tyers, M., Boone, C., and Friend, S. H. (2000). Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science.* 287(5454): 873-880. Web supplement.

Rodier, A., Rochard, P., Berthet, C., Rouault, J. P., Casas, F., Daury, L., Busson, M., Magaud, J. P., Wrutniak-Cabello, C., and Cabello, G. (2001). Identification of functional domains involved in BTG1 cell localization. *Oncogene.* 20(21): 2691-2703.

Romualdi, C., Trevisan, S., Celegato, B., Costa, G., and Lanfranchi, G. (2003). Improved detection of differentially expressed genes in microarray experiments through multiple scanning and image integration. *Nucleic Acids Res.* 31(23): e149.

Rosenbauer, F., Koschmieder, S., Steidl, U., and Tenen, D. G. (2005). Effect of transcription-factor concentrations on leukemic stem cells. *Blood.* 106(5): 1519-1524.

Rosenbauer, F., Wagner, K., Kutok, J. L., Iwasaki, H., Le Beau, M. M., Okuno, Y., Akashi, K., Fiering, S., and Tenen, D. G. (2004). Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat. Genet.* 36(6): 624-630.

Rosmarin, A. G., Yang, Z., and Resendes, K. K. (2005). Transcriptional regulation in myelopoiesis: Hematopoietic fate choice, myeloid differentiation, and leukemogenesis. *Exp. Hematol.* 33(2): 131-143.

Rosnet, O., Buhring, H. J., Marchetto, S., Rappold, I., Lavagna, C., Sainty, D., Arnoulet, C., Chabannon, C., Kanz, L., Hannum, C., and Birnbaum, D. (1996). Human FLT3/FLK2 receptor tyrosine kinase is expressed at the surface of normal and malignant hematopoietic cells. *Leukemia.* 10(2): 238-248.

Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J. C., Lashkari, D., Shalon, D., Myers, T. G., Weinstein, J. N., Botstein, D., and Brown, P. O. (2000). Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* 24(3): 227-235.

Rouault, J. P., Rimokh, R., Tessa, C., Paranhos, G., Ffrench, M., Duret, L., Garoccio, M., Germain, D., Samarut, J., and Magaud, J. P. (1992). BTG1, a member of a new family of antiproliferative genes. *EMBO J.* 11(4): 1663-1670.

Rouault, J. P., Samarut, C., Duret, L., Tessa, C., Samarut, J., and Magaud, J. P. (1993). Sequence analysis reveals that the BTG1 anti-proliferative gene is conserved throughout evolution in its coding and 3' non-coding regions. *Gene.* 129(2): 303-306.

Rowley, J. D. (1998). The critical role of chromosome translocations in human leukemias. *Annu. Rev. Genet.* 32: 495-519.

Roxstrom-Lindquist, K., and Faye, I. (2001). The Drosophila gene Yippee reveals a novel family of putative zinc binding proteins highly conserved among eukaryotes. *Insect Mol. Biol.* 10(1): 77-86.

Rubinstein, M., Idelman, G., Plymate, S. R., Narla, G., Friedman, S. L., and Werner, H. (2004). Transcriptional activation of the insulin-like growth factor I receptor gene by the Kruppel-like factor 6 (KLF6) tumor suppressor protein: potential interactions between KLF6 and p53. *Endocrinology.* 145(8): 3769-3777.

Russo, M. W., Sevetson, B. R., and Milbrandt, J. (1995). Identification of NAB1, a repressor of NGFI-A- and Krox20-mediated transcription. *Proc. Natl. Acad. Sci. U. S. A.* 92(15): 6873-6877.

Saglio, G., Cilloni, D., Rancati, F., and Boano, L. (2004). Glivec and CML: a lucky date. *J. Biol. Regul. Homeost. Agents.* 18(2): 246-251.

Saito, M., Smiley, J., Toth, R., and Vadasz, C. (2002). Microarray analysis of gene expression in rat hippocampus after chronic ethanol treatment. *Neurochem. Res.* 27(10): 1221-1229.

Sakaguchi, T., Kuroiwa, A., and Takeda, H. (2001). Expression of zebrafish btg-b, an anti-proliferative cofactor, during early embryogenesis. *Mech. Dev.* 104(1-2): 113-115.

Sammar, M., Aigner, S., Hubbe, M., Schirrmacher, V., Schachner, M., Vestweber, D., and Altevogt, P. (1994). Heat-stable antigen (CD24) as ligand for mouse P-selectin. *Int. Immunol.* 6(7): 1027-1036.

Santin, A. D., Zhan, F., Bellone, S., Palmieri, M., Cane, S., Bignotti, E., Anfossi, S., Gokden, M., Dunn, D., Roman, J. J., O'Brien, T. J., Tian, E., Cannon, M. J., Shaughnessy, J., Jr., and Pecorelli, S. (2004). Gene expression profiles in primary ovarian serous papillary tumors and normal ovarian epithelium: identification of candidate molecular markers for ovarian cancer diagnosis and therapy. *Int. J. Cancer.* 112(1): 14-25.

Sapir, M., and Churchill, G. A. (2000). Estimating the posterior probability of differential gene expression from microarray data. Poster, The Jackson Laboratory. http://www.jax.org/staff/churchill/labsite/pubs/index.html.

Sasajima, H., Nakagawa, K., and Yokosawa, H. (2002). Antiproliferative proteins of the BTG/Tob family are degraded by the ubiquitin-proteasome system. *Eur. J. Biochem.* 269(14): 3596-3604.

Sawyers, C. L. (1997). Signal transduction pathways involved in BCR-ABL transformation. *Baillieres Clin. Haematol.* 10(2): 223-231.

Scheid, M. P., Woodgett, J. R., and Woodgett, J. R. (2001). Pkb/akt: functional insights from genetic models. *Nat. Rev. Mol. Cell Biol.* 2(10): 760-768.

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* 270(5235): 467-470.

Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P. O., and Davis, R. W. (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. U. S. A.* 93(20): 10614-10619.

Schoch, C., Kohlmann, A., Schnittger, S., Brors, B., Dugas, M., Mergenthaler, S., Kern, W., Hiddemann, W., Eils, R., and Haferlach, T. (2002). Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 99(15): 10008-10013.

Schonig, K., Schwenk, F., Rajewsky, K., and Bujard, H. (2002). Stringent doxycycline dependent control of CRE recombinase *in vivo*. *Nucleic Acids Res.* 30(23): e134.

Schoondermark-Stolk, S. A., ter Schure, E. G., Verrips, C. T., Verkleij, A. J., and Boonstra, J. (2002). Identification of salt-induced genes of Zygosaccharomyces rouxii by using Saccharomyces cerevisiae GeneFilters. *FEMS Yeast Res.* 2(4): 525-532.

Scott, C. L., Hughes, D. A., Cary, D., Nicola, N. A., Begley, C. G., and Robb, L. (1998). Functional analysis of mature hematopoietic cells from mice lacking the betac chain of the granulocyte-macrophage colony-stimulating factor receptor. *Blood.* 92(11): 4119-4127.

Seeler, J. S., and Dejean, A. (1999). The PML nuclear bodies: actors or extras? *Curr. Opin. Genet. Dev.* 9(3): 362-367.

Seeler, J. S., Marchio, A., Losson, R., Desterro, J. M., Hay, R. T., Chambon, P., and Dejean, A. (2001). Common properties of nuclear body protein SP100 and TIF1alpha chromatin factor: role of SUMO modification. *Mol. Cell. Biol.* 21(10): 3314-3324.

Seeler, J. S., Marchio, A., Sitterlin, D., Transy, C., and Dejean, A. (1998). Interaction of SP100 with HP1 proteins: a link between the promyelocytic leukemia-associated nuclear bodies and the chromatin compartment. *Proc. Natl. Acad. Sci. U. S. A.* 95(13): 7316-7321.

Semenza, G. L., Dureza, R. C., Traystman, M. D., Gearhart, J. D., and Antonarakis, S. E. (1990). Human erythropoietin gene expression in transgenic mice: multiple transcription initiation sites and cis-acting regulatory elements. *Mol.Cell. Biol.* 10(3): 930-938.

Seymour, J. F., Lieschke, G. J., Grail, D., Quilici, C., Hodgson, G., and Dunn, A. R. (1997). Mice lacking both granulocyte colony-stimulating factor (CSF) and granulocyte-macrophage CSF have impaired reproductive capacity, perturbed neonatal granulopoiesis, lung disease, amyloidosis, and reduced long-term survival. *Blood.* 90(8): 3037-3049.

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annu. Rev. Psychol.* 46: 561-584.

Sherr, C. J., and Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev.* 13(12): 1501-1512.

Shetty, S., Muniyappa, H., Halady, P. K., and Idell, S. (2004). Regulation of urokinase receptor expression by phosphoglycerate kinase. *Am. J. Respir. Cell Mol. Biol.* 31(1): 100-106.

Shi, H., Zhang, Z., Wang, X., Liu, S., and Teng, C. T. (1999). Isolation and characterization of a gene encoding human Kruppel-like factor 5 (IKLF): binding to the CAAT/GT box of the mouse lactoferrin gene promoter. *Nucleic Acids Res.* 27(24): 4807-4815.

Shindo, T., Manabe, I., Fukushima, Y., Tobe, K., Aizawa, K., Miyamoto, S., Kawai-Kowase, K., Moriyama, N., Imai, Y., Kawakami, H., Nishimatsu, H., Ishikawa, T., Suzuki, T., Morita, H., Maemura, K., Sata, M., Hirata, Y., Komukai, M., Kagechika, H., Kadowaki, T., Kurabayashi, M., and Nagai, R. (2002). Kruppel-like zinc-finger transcription factor KLF5/BTEB2 is a target for angiotensin II signaling and an essential regulator of cardiovascular remodeling. *Nat. Med.* 8(8): 856-863.

Shinjo, K., Takeshita, A., Ohnishi, K., and Ohno, R. (1997). Granulocyte colony-stimulating factor receptor at various differentiation stages of normal and leukemic hematopoietic cells. *Leuk. Lymphoma.* 25(1-2): 37-46.

Sijin, L., Ziwei, C., Yajun, L., Meiyu, D., Hongwei, Z., Guofa, H., Siguo, L., Hong, G., Zhihong, Z., Xiaolei, L., Yingyun, W., Yan, X., and Weide, L. (2004). The effect of knocking-down nucleostemin gene expression on the *in vitro* proliferation and *in vivo* tumorigenesis of HeLa cells. *J. Exp. Clin. Cancer Res.* 23(3): 529-538.

Slavin, D. A., Koritschoner, N. P., Prieto, C. C., Lopez-Diaz, F. J., Chatton, B., and Bocco, J. L. (2004). A new role for the Kruppel-like transcription factor KLF6 as an inhibitor of c-Jun proto-oncoprotein function. *Oncogene.* 23(50): 8196-8205.

Slonim, D. K. (2002). From patterns to pathways: gene expression data analysis comes of age. *Nat. Genet.* 32 Suppl: 502-508.

Smith, L. T., Hohaus, S., Gonzalez, D. A., Dziennis, S. E., and Tenen, D. G. (1996). PU.1 (Spi-1) and C/EBP alpha regulate the granulocyte colony-stimulating factor receptor promoter in myeloid cells. *Blood.* 88(4): 1234-1247.

Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3: Article 3.

Smyth, G. K., and Speed, T. (2003). Normalization of cDNA microarray data. *Methods.* 31(4): 265-273.

Smyth, G. K., Yang, Y. H., and Speed, T. (2003). Statistical issues in cDNA microarray data analysis. *Methods Mol. Biol.* 224: 111-136.

Soille, P. (1999). *Morphological image analysis: principles and applications* Springer Verlag, New York, USA.

Song, J. H., Kim, J. M., Kim, S. H., Kim, H. J., Lee, J. J., Sung, M. H., Hwang, S. Y., and Kim, T. S. (2003). Comparison of the gene expression profiles of monocytic versus granulocytic lineages of HL-60 leukemia cell differentiation by DNA microarray analysis. *Life Sci.* 73(13): 1705-1719.

Southern, E., Mir, K., and Shchepinov, M. (1999). Molecular interactions on microarrays. *Nat. Genet.* 21(1 Suppl): 5-9.

Souyri, M., Vigon, I., Penciolelli, J. F., Heard, J. M., Tambourin, P., and Wendling, F. (1990). A putative truncated cytokine receptor gene transduced by the myeloproliferative leukemia virus immortalizes hematopoietic progenitors. *Cell.* 63(6): 1137-1147.

Spangrude, G., Smith, L., Uchida, N., Ikuta, K., Heimfeld, S., Friedman, J., and Weissman, I. (1991). Mouse hematopoietic stem cells. *Blood.* 78(6): 1395-1402.

Spangrude, G. J., Heimfeld, S., and Weissman, I. L. (1988). Purification and characterization of mouse hematopoietic stem cells. *Science.* 241(4861): 58-62.

Speed, T. P. (2001). Gene expression. Wald Lecture III, Joint Statistical Meetings, August 2001, http://www.stat.berkeley.edu/users/terry/Wald/waldlec.html

Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol. Biol. Cell.* 9(12): 3273-3297.

Stadler, M., Chelbi-Alix, M. K., Koken, M. H., Venturini, L., Lee, C., Saib, A., Quignon, F., Pelicano, L., Guillemin, M. C., Schindler, C., and *et al*. (1995). Transcriptional induction of the PML growth suppressor gene by interferons is mediated through an ISRE and a GAS element. *Oncogene.* 11(12): 2565-2573.

Stanley, E., Lieschke, G. J., Grail, D., Metcalf, D., Hodgson, G., Gall, J. A., Maher, D. W., Cebon, J., Sinickas, V., and Dunn, A. R. (1994). Granulocyte/macrophage colony-stimulating factor-deficient mice show no major perturbation of hematopoiesis but develop a characteristic pulmonary pathology. *Proc. Natl. Acad. Sci. U. S. A.* 91(12): 5592-5596.

Stephens, R. W., Nielsen, H. J., Christensen, I. J., Thorlacius-Ussing, O., Sorensen, S., Dano, K., and Brunner, N. (1999). Plasma urokinase receptor levels in patients with colorectal cancer: relationship to prognosis. *J. Natl. Cancer Inst.* 91(10): 869-874.

Sternberg, M. J., and Gullick, W. J. (1989). Neu receptor dimerization. *Nature.* 339(6226): 587.

Sun, L., Kerawalla, H., Wu, X., Lehnert, M. S., and Uckun, F. M. (2002). Expression of a unique helios isoform in human leukemia cells. *Leuk. Lymphoma.* 43(4): 841-849.

Sun, R., Chen, X., and Yang, V. W. (2001). Intestinal-enriched Kruppel-like factor (Kruppel-like factor 5) is a positive regulator of cellular proliferation. *J. Biol. Chem.* 276(10): 6897-6900.

Suzuki, T., Shen, H., Akagi, K., Morse, H. C., Malley, J. D., Naiman, D. Q., Jenkins, N. A., and Copeland, N. G. (2002). New genes involved in cancer identified by retroviral tagging. *Nat. Genet.* 32(1): 166-174.

Suzuki, T., Yamamoto, T., Kurabayashi, M., Nagai, R., Yazaki, Y., and Horikoshi, M. (1998). Isolation and initial characterization of GBF, a novel DNA-binding zinc finger protein that binds to the GC-rich binding sites of the HIV-1 promoter. *J. Biochem. (Tokyo).* 124(2): 389-395.

Takanashi, M., Yagi, T., Imamura, T., Tabata, Y., Morimoto, A., Hibi, S., Ishii, E., and Imashuku, S. (2002). Expression of the Ikaros gene family in childhood acute lymphoblastic leukaemia. *Br. J. Haematol.* 117(3): 525-530.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E. S., and Golub, T. R. (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 96(6): 2907-2912.

Tambourin, P. E., Wendling, F., Jasmin, C., and Smadja-Joffe, F. (1979). The physiopathology of Friend leukemia. *Leuk. Res.* 3(3): 117-129.

Tanaka, T. S., Jaradat, S. A., Lim, M. K., Kargul, G. J., Wang, X., Grahovac, M. J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W. H., 3rd, Becker, K. G., and Ko, M. S. (2000). Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci. U. S. A.* 97(16): 9127-9132.

Taniguchi, M., Miura, K., Iwao, H., and Yamanaka, S. (2001). Quantitative assessment of DNA microarrays--comparison with Northern blot analyses. *Genomics.* 71(1): 34-39.

Tartaglia, L. A., Dembski, M., Weng, X., Deng, N., Culpepper, J., Devos, R., Richards, G. J., Campfield, L. A., Clark, F. T., Deeds, J., and *et al*. (1995). Identification and expression cloning of a leptin receptor, OB-R. *Cell.* 83(7): 1263-1271.

Tefferi, A., and Gilliland, D. G. (2005). The JAK2V617F tyrosine kinase mutation in myeloproliferative disorders: status report and immediate implications for disease classification and diagnosis. *Mayo Clin. Proc.* 80(7): 947-958.

Teglund, S., McKay, C., Schuetz, E., van Deursen, J. M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J. N. (1998). Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. *Cell.* 93(5): 841-850.

Tenen, D. G. (2003). Disruption of differentiation in human cancer: AML shows the way. *Nat. Rev. Cancer.* 3(2): 89-101.

Testa, U., Riccioni, R., Diverio, D., Rossini, A., Lo Coco, F., and Peschle, C. (2004). Interleukin-3 receptor in acute leukemia. *Leukemia.* 18(2): 219-226.

Thomson, A. W. (1998). *The cytokine handbook*, 3rd Ed. Academic Press, San Diego, California, USA.

Tirone, F. (2001). The gene PC3(TIS21/BTG2), prototype member of the PC3/BTG/TOB family: regulator in control of cell growth, differentiation, and DNA repair? *J. Cell. Physiol.* 187(2): 155-165.

Tokuyama, Y., Horn, H. F., Kawamura, K., Tarapore, P., and Fukasawa, K. (2001). Specific phosphorylation of nucleophosmin on Thr(199) by cyclin-dependent kinase 2-cyclin E and its role in centrosome duplication. *J. Biol. Chem.* 276(24): 21529-21537.

Tomasson, M. H., Williams, I. R., Li, S., Kutok, J., Cain, D., Gillessen, S., Dranoff, G., Van Etten, R. A., and Gilliland, D. G. (2001). Induction of myeloproliferative disease in mice by tyrosine kinase fusion oncogenes does not require granulocyte-macrophage colony-stimulating factor or interleukin-3. *Blood.* 97(5): 1435-1441.

Tsai, R. Y., and McKay, R. D. (2002). A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. *Genes Dev.* 16(23): 2991-3003.

Tseng, G. C., Oh, M. K., Rohlin, L., Liao, J. C., and Wong, W. H. (2001). Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.* 29(12): 2549-2557.

Tsuzuki, S., Kitajima, K., Nakano, T., Glasow, A., Zelent, A., and Enver, T. (2004). Cross talk between retinoic acid signaling and transcription factor GATA-2. *Mol. Cell. Biol.* 24(15): 6824-6836.

Tucker, M., Valencia-Sanchez, M. A., Staples, R. R., Chen, J., Denis, C. L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. *Cell.* 104(3): 377-386.

Tuomisto, T. T., Riekkinen, M. S., Viita, H., Levonen, A. L., and Yla-Herttuala, S. (2005). Analysis of gene and protein expression during monocyte-macrophage differentiation and cholesterol loading--cDNA and protein array study. *Atherosclerosis.* 180(2): 283-291.

Tuschl, T., and Borkhardt, A. (2002). Small interfering RNAs: a revolutionary tool for the analysis of gene function and gene therapy. *Mol. Interv.* 2(3): 158-167.

Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A.* 98(9): 5116-5121.

Ullrich, A., and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. *Cell.* 61(2): 203-212.

Usui, N. (2005). [Imatinib therapy for patients with chronic myelogenous leukemia]. *Gan To. Kagaku Ryoho.* 32(3): 297-303.

Van Parijs, L., Refaeli, Y., Lord, J. D., Nelson, B. H., Abbas, A. K., and Baltimore, D. (1999). Uncoupling IL-2 signals that regulate T cell proliferation, survival, and Fas-mediated activation-induced cell death. *Immunity.* 11(3): 281-288.

Vanhaesebroeck, B., Leevers, S. J., Ahmadi, K., Timms, J., Katso, R., Driscoll, P. C., Woscholski, R., Parker, P. J., and Waterfield, M. D. (2001). Synthesis and function of 3-phosphorylated inositol lipids. *Annu. Rev. Biochem.* 70: 535-602.

Virtaneva, K., Wright, F. A., Tanner, S. M., Yuan, B., Lemon, W. J., Caligiuri, M. A., Bloomfield, C. D., de La Chapelle, A., and Krahe, R. (2001). Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proc. Natl. Acad. Sci. U. S. A.* 98(3): 1124-1129.

Wada, H., Noguchi, Y., Marino, M. W., Dunn, A. R., and Old, L. J. (1997). T cell functions in granulocyte/macrophage colony-stimulating factor deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* 94(23): 12557-12561.

Wang, H., Yang, L., Jamaluddin, M. S., and Boyd, D. D. (2004). The Kruppel-like KLF4 transcription factor, a novel regulator of urokinase receptor expression, drives synthesis of this binding site in colonic crypt luminal surface epithelial cells. *J. Biol. Chem.* 279(21): 22674-22683.

Wang, I. J., Carlson, E. C., Liu, C. Y., Kao, C. W., Hu, F. R., and Kao, W. W. (2002). Cis-regulatory elements of the mouse Krt1.12 gene. *Mol. Vis.* 8: 94-101.

Ward, A. C., Touw, I., and Yoshimura, A. (2000). The Jak-Stat pathway in normal and perturbed hematopoiesis. *Blood.* 95(1): 19-29.

Warke, V. G., Nambiar, M. P., Krishnan, S., Tenbrock, K., Geller, D. A., Koritschoner, N. P., Atkins, J. L., Farber, D. L., and Tsokos, G. C. (2003). Transcriptional activation of the human inducible nitric-oxide synthase promoter by Kruppel-like factor 6. *J. Biol. Chem.* 278(17): 14812-14819.

Warner, J. K., Wang, J. C., Hope, K. J., Jin, L., and Dick, J. E. (2004). Concepts of human leukemic development. *Oncogene.* 23(43): 7164-7177.

Watanabe, S., Itoh, T., and Arai, K. (1997). Roles of JAK kinase in human GM-CSF receptor signals. *Leukemia.* 11 Suppl 3: 76-78.

Watowich, S. S., Hilton, D. J., and Lodish, H. F. (1994). Activation and inhibition of erythropoietin receptor function: role of receptor dimerization. *Mol. Cell. Biol.* 14(6): 3535-3549.

Watowich, S. S., Wu, H., Socolovsky, M., Klingmuller, U., Constantinescu, S. N., and Lodish, H. F. (1996). Cytokine receptor signal transduction and the control of hematopoietic cell development. *Annu. Rev. Cell Dev. Biol.* 12: 91-128.

Weichert, W., Denkert, C., Burkhardt, M., Gansukh, T., Bellach, J., Altevogt, P., Dietel, M., and Kristiansen, G. (2005). Cytoplasmic CD24 expression in colorectal cancer independently correlates with shortened patient survival. *Clin. Cancer Res.* 11(18): 6574-6581.

Weiner, D. B., Liu, J., Cohen, J. A., Williams, W. V., and Greene, M. I. (1989). A point mutation in the neu oncogene mimics ligand induction of receptor aggregation. *Nature.* 339(6221): 230-231.

Wendling, F., Penciolelli, J. F., Charon, M., and Tambourin, P. (1989a). Factor-independent erythropoietic progenitor cells in leukemia induced by the myeloproliferative leukemia virus. *Blood.* 73(5): 1161-1167.

Wendling, F., Varlet, P., Charon, M., and Tambourin, P. (1986). MPLV: a retrovirus complex inducing an acute myeloproliferative leukemic disorder in adult mice. *Virology.* 149(2): 242-246.

Wendling, F., Vigon, I., Souyri, M., and Tambourin, P. (1989b). Myeloid progenitor cells transformed by the myeloproliferative leukemia virus proliferate and differentiate *in vitro* without the addition of growth factors. *Leukemia.* 3(7): 475-480.

Wenger, R. H., Ayane, M., Bose, R., Kohler, G., and Nielsen, P. J. (1991). The genes for a mouse hematopoietic differentiation marker called the heat-stable antigen. *Eur. J. Immunol.* 21(4): 1039-1046.

Wenger, R. H., Rochelle, J. M., Seldin, M. F., Kohler, G., and Nielsen, P. J. (1993). The heat stable antigen (mouse CD24) gene is differentially regulated but has a housekeeping promoter. *J. Biol. Chem.* 268(31): 23345-23352.

Wieland, I., Ropke, A., Stumm, M., Sell, C., Weidle, U. H., and Wieacker, P. F. (2001). Molecular characterization of the DICE1 (DDX26) tumor suppressor gene in lung carcinoma cells. *Oncol. Res.* 12(11-12): 491-500.

Wieland, I., Sell, C., Weidle, U. H., and Wieacker, P. (2004). Ectopic expression of DICE1 suppresses tumor cell growth. *Oncol. Rep.* 12(2): 207-211.

Wilcox, J. N., Scott, N. A., Subramanian, R. R., Ross, C. E., Adams-Burton, C. R., Stoltenborg, J. K., and Corjay, M. H. (1995). *Circulation.* 92: 134-135.

Winkelmann, J. C., Ward, J., Mayeux, P., Lacombe, C., Schimmenti, L., and Jenkins, R. B. (1995). A translocated erythropoietin receptor gene in a human erythroleukemia cell line (TF-1) expresses an abnormal transcript and a truncated protein. *Blood.* 85(1): 179-185.

Wittinghofer, F. (1998). Ras signalling. Caught in the act of the switch-on. *Nature.* 394(6691): 317, 319-317, 320.

Wiznerowicz, M., and Trono, D. (2003). Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference. *J. Virol.* 77(16): 8957-8961.

Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., and Lockhart, D. J. (1997). Genome-wide expression monitoring in Saccharomyces cerevisiae. *Nat. Biotechnol.* 15(13): 1359-1367.

Wolff, L. (1997). Contribution of oncogenes and tumor suppressor genes to myeloid leukemia. *Biochim. Biophys. Acta.* 1332(3): F67-104.

Wong, P. M., Chung, S. W., Dunbar, C. E., Bodine, D. M., Ruscetti, S., and Nienhuis, A. W. (1989). Retrovirus-mediated transfer and expression of the interleukin-3 gene in mouse hematopoietic cells result in a myeloproliferative disorder. *Mol. Cell. Biol.* 9(2): 798-808.

Wride, M. A., Mansergh, F. C., Adams, S., Everitt, R., Minnema, S. E., Rancourt, D. E., and Evans, M. J. (2003). Expression profiling and gene discovery in the mouse lens. *Mol. Vis.* 9: 360-396.

Wu, K., Yang, Y., Wang, C., Davoli, M. A., D'Amico, M., Li, A., Cveklova, K., Kozmik, Z., Lisanti, M. P., Russell, R. G., Cvekl, A., and Pestell, R. G. (2003). DACH1 inhibits transforming growth factor-beta signaling through binding Smad4. *J. Biol. Chem.* 278(51): 51673-51684.

Wu, M. H., and Yung, B. Y. (2002). UV stimulation of nucleophosmin/B23 expression is an immediate-early gene response induced by damaged DNA. *J. Biol. Chem.* 277(50): 48234-48240.

Xu, W., Qian, H., Zhu, W., Chen, Y., Shao, Q., Sun, X., Hu, J., Han, C., and Zhang, X. (2004). A novel tumor cell line cloned from mutated human embryonic bone marrow mesenchymal stem cells. *Oncol. Rep.* 12(3): 501-508.

Yaghoobi, M. M., Mowla, S. J., and Tiraihi, T. (2005). Nucleostemin, a coordinator of self-renewal, is expressed in rat marrow stromal cells and turns off after induction of neural differentiation. *Neurosci. Lett.*

Yale, J., and Bohnert, H. J. (2001). Transcript expression in Saccharomyces cerevisiae at high salinity. *J. Biol. Chem.* 276(19): 15996-16007.

Yamada, T., Kaneko, H., Iizuka, K., Matsubayashi, Y., Kokai, Y., and Fujimoto, J. (1996). Elevation of lymphocyte and hematopoietic stem cell numbers in mice transgenic for human granulocyte CSF. *Lab. Invest.* 74(2): 384-394.

Yamamoto, Y., Kiyoi, H., Nakano, Y., Suzuki, R., Kodera, Y., Miyawaki, S., Asou, N., Kuriyama, K., Yagasaki, F., Shimazaki, C., Akiyama, H., Saito, K., Nishimura, M., Motoji, T., Shinagawa, K., Takeshita, A., Saito, H., Ueda, R., Ohno, R., and Naoe, T. (2001). Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood.* 97(8): 2434-2439.

Yang, H. X., Jin, G. L., Meng, L., Zhang, J. Z., Liu, W. B., and Shou, C. C. (2005). Screening and identification of proteins interacting with nucleostemin. *World J. Gastroenterol.* 11(31): 4812-4814.

Yang, Y. H., Buckley, M. J., Dudoit, S., and Speed, T. P. (2001a). Comparison of methods for image analysis on cDNA microarray data. Technical Report 584, Department of Statistics, University of California, Berkeley. http://www.stat.berkeley.edu/users/terry/zarray/Html/papersindex.html.

Yang, Y. H., Buckley, M. J., and Speed, T. P. (2001b). Analysis of cDNA microarray images. *Brief. Bioinform.* 2(4): 341-349.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 30(4): e15.

Yang, Y. H., Dudoit, S., Luu, P., and Speed, T. P. (2001c). Normalisation for cDNA microarray data. In *Microarrays: Optical Technologies and Informatics* (Bittner, M. L., Chen, Y., Dorsel, A. N., and Dougherty, E. R., eds), SPIE, Society for Optical Engineering, San Jose, CA, USA.

Yasuda, K., Hirayoshi, K., Hirata, H., Kubota, H., Hosokawa, N., and Nagata, K. (2002). The Kruppel-like factor Zf9 and proteins in the Sp1 family regulate the expression of HSP47, a collagen-specific molecular chaperone. *J. Biol. Chem.* 277(47): 44613-44622.

Yee, J. K., Friedmann, T., and Burns, J. C. (1994a). Generation of high-titer pseudotyped retroviral vectors with very broad host range. *Methods Cell Biol.* 43 Pt A: 99-112.

Yee, J. K., Miyanohara, A., LaPorte, P., Bouic, K., Burns, J. C., and Friedmann, T. (1994b). A general method for the generation of high-titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. *Proc. Natl. Acad. Sci. U. S. A.* 91(20): 9564-9568.

Yoneda-Kato, N., Look, A. T., Kirstein, M. N., Valentine, M. B., Raimondi, S. C., Cohen, K. J., Carroll, A. J., and Morris, S. W. (1996). The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid leukemia produces a novel fusion gene, NPM-MLF1. *Oncogene.* 12(2): 265-275.

Yoshida, T., Ikuta, K., Sugaya, H., Maki, K., Takagi, M., Kanazawa, H., Sunaga, S., Kinashi, T., Yoshimura, K., Miyazaki, J., Takaki, S., and Takatsu, K. (1996). Defective B-1 cell development and impaired immunity against Angiostrongylus cantonensis in IL-5R alpha-deficient mice. *Immunity.* 4(5): 483-494.

Yoshimura, A., D'Andrea, A. D., and Lodish, H. F. (1990). Friend spleen focus-forming virus glycoprotein gp55 interacts with the erythropoietin receptor in the endoplasmic reticulum and affects receptor metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 87(11): 4139-4143.

Young, D. C., and Griffin, J. D. (1986). Autocrine secretion of GM-CSF in acute myeloblastic leukemia. *Blood.* 68(5): 1178-1181.

Yu, H. M., Liu, B., Chiu, S. Y., Costantini, F., and Hsu, W. (2005). Development of a unique system for spatiotemporal and lineage-specific gene expression in mice. *Proc. Natl. Acad. Sci. U. S. A.* 102(24): 8615-8620.

Yu, X., Zhan, X., D'Costa, J., Tanavde, V. M., Ye, Z., Peng, T., Malehorn, M. T., Yang, X., Civin, C. I., and Cheng, L. (2003). Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Mol. Ther.* 7(6): 827-838.

Zarn, J. A., Zimmermann, S. M., Pass, M. K., Waibel, R., and Stahel, R. A. (1996). Association of CD24 with the kinase c-fgr in a small cell lung cancer cell line and with the kinase lyn in an erythroleukemia cell line. *Biochem. Biophys. Res. Commun.* 225(2): 384-391.

Zhan, Y., Basu, S., Lieschke, G. J., Grail, D., Dunn, A. R., and Cheers, C. (1999). Functional deficiencies of peritoneal cells from gene-targeted mice lacking G-CSF or GM-CSF. *J. Leukoc. Biol.* 65(2): 256-264.

Zhan, Y., Lieschke, G. J., Grail, D., Dunn, A. R., and Cheers, C. (1998). Essential roles for granulocyte-macrophage colony-stimulating factor (GM-CSF) and G-CSF in the sustained hematopoietic response of Listeria monocytogenes-infected mice. *Blood.* 91(3): 863-869.

Zhang, P., Iwama, A., Datta, M. W., Darlington, G. J., Link, D. C., and Tenen, D. G. (1998). Upregulation of interleukin 6 and granulocyte colony-stimulating factor receptors by transcription factor CCAAT enhancer binding protein alpha (C/EBP alpha) is critical for granulopoiesis. *J. Exp. Med.* 188(6): 1173-1184.

Zhao, J. L., Austen, K. F., and Lam, B. K. (2000). Cell-specific transcription of leukotriene C(4) synthase involves a Kruppel-like transcription factor and Sp1. *J. Biol. Chem.* 275(12): 8903-8910.

Zhao, Y., Wagner, F., Frank, S. J., and Kraft, A. S. (1995). The amino-terminal portion of the JAK2 protein kinase is necessary for binding and phosphorylation of the granulocyte-macrophage colony-stimulating factor receptor beta c chain. *J. Biol. Chem.* 270(23): 13814-13818.

Zheng, P. Z., Wang, K. K., Zhang, Q. Y., Huang, Q. H., Du, Y. Z., Zhang, Q. H., Xiao, D. K., Shen, S. H., Imbeaud, S., Eveno, E., Zhao, C. J., Chen, Y. L., Fan, H. Y., Waxman, S., Auffray, C., Jin, G., Chen, S. J., Chen, Z., and Zhang, J. (2005). Systems analysis of transcriptome and proteome in retinoic acid/arsenic trioxide-induced cell differentiation/apoptosis of promyelocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* 102(21): 7653-7658.

Zheng, X., Ravatn, R., Lin, Y., Shih, W. C., Rabson, A., Strair, R., Huberman, E., Conney, A., and Chin, K. V. (2002). Gene expression of TPA induced differentiation in HL-60 cells by DNA microarray analysis. *Nucleic Acids Res.* 30(20): 4489-4499.

Zhu, J., and Emerson, S. G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene.* 21(21): 3295-3313.

Zon, L. I., Moreau, J. F., Koo, J. W., Mathey-Prevot, B., and D'Andrea, A. D. (1992). The erythropoietin receptor transmembrane region is necessary for activation by the Friend spleen focus-forming virus gp55 glycoprotein. *Mol. Cell. Biol.* 12(7): 2949-2957.