UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH & LIFE SCIENCES

School of Medicine

Abnormalities affecting tyrosine kinase signalling in atypical myeloproliferative disorders

by

Claire Hidalgo-Curtis

Thesis for the degree of Doctor of Philosophy

May 2009

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH & LIFE SCIENCES
SCHOOL OF MEDICINE

Doctor of Philosophy

ABNORMALITIES AFFECTING TYROSINE KINASE SIGNALLING IN ATYPICAL MYELOPROLIFERATIVE DISORDERS

By Claire Hidalgo-Curtis

The myeloproliferative disorders (MPDs) are a group of haematopoietic stem cell diseases, characterised by proliferation of one or more cells of the myeloid lineage. Several lines of evidence have highlighted the importance of aberrant tyrosine kinase signalling in the pathogenesis of these disorders. Cloning of rare chromosomal translocations and point mutation analysis in the MPDs has identified diverse deregulated tyrosine kinase genes, notably *PDGFRA*, *PDGFRB*, *FGFR1* and *JAK2*. However the majority of atypical MPDs still remain to be characterised and identification of patients harbouring fusions, particularly those involving the PDGF receptors is of increasing importance, as they are likely to be responsive to targeted therapy with imatinib.

I am investigating MPD patients for abnormalities affecting tyrosine kinase signalling, and have used three approaches, translocation cloning, expression analysis and SNP array analysis to detect regions of loss of heterozygosity (LOH). Thus far, by translocation cloning I have identified a previously undescribed partner gene fused to *PDGFRB* and two new *PDGFRA* fusion genes. I have also designed two reverse transcriptase PCR (RT-PCR) assays and a cDNA MLPA assay to detect over-expression of specific tyrosine kinases screening approximately 200 patients. Each assay identified all patients previously diagnosed with known fusions. Additionally, two patients identified with overexpression of *PDGFRB* have been found to have cryptic *ETV6-PDGFRB* fusions and overexpression of *PDGFRA* in one patient lead to the discovery of a previously undescribed fusion involving a novel partner gene (*KIF5B*).

Recent evidence has indicated that acquired isodisomy is a novel mechanism by which mutations in cancer may be reduced to homozygosity. Typically, acquired isodisomy is associated with oncogenic changes rather than tumour suppressor genes, eg. the activating *JAK2* V617F mutation and 9p aUPD. I have undertaken a screen using Affymetrix 50K SNP arrays for regions of acquired isodisomy as a means to identify genomic regions that may harbour novel oncogenes in different subgroups of MPD patients. Large tracts of homozygosity (defined as >20Mb running to a telomere), strongly suggesting acquired isodisomy, were seen in 40% aMPD patients. The homozygous tracts encompassed diverse genomic regions in aMPD, but two common regions (3 cases for each region) were identified at 7q and 11q. Mutations in the CBL ubiquitin ligase gene were discovered in all three aCML patients with 11q aUPD as well as in an additional 23 MPD patients following further screening.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION OF AUTHORSHIP

I, Claire Elizabeth Hidalgo-Curtis, declare that the thesis entitled:

Abnormalities affecting tyrosine kinase signalling in atypical myeloproliferative disorders

And the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:

    - Two novel imatinib-responsive PDGFRA fusion genes in chronic eosinophilic leukaemia. Curtis CE, Grand FH, Musto P, Clark A, Murphy J, Perla G, Minervini MM, Stewart J, Reiter A, Cross NC. Br J Haematol. 2007 Jul;138(1):77-81.

    - A novel ETV6-PDGFRB fusion transcript missed by standard screening in a patient with an imatinib responsive chronic myeloproliferative disease.

Curtis CE, Grand FH, Waghorn K, Sahoo TP, George J, Cross NC. Leukemia. 2007 Aug;21(8):1839-41. Epub 2007 May 17.

- The t(1;9)(p34;q34) and t(8;12)(p11;q15) fuse pre-mRNA processing proteins SFPQ (PSF) and CPSF6 to ABL and FGFR1. Hidalgo-Curtis C, Chase A, Drachenberg M, Roberts MW, Finkelstein JZ, Mould S, Oscier D, Cross NC, Grand FH. Genes Chromosomes Cancer. 2008 May;47(5):379-85.

- Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms. Grand FH, Hidalgo-Curtis CE, Ernst T, Zoi K, Zoi C, McGuire C, Kreil S, Jones A, Score J, Metzgeroth G, Oscier D, Hall A, Brandts C, Serve H, Reiter A, Chase AJ, Cross NC. Blood. 2009 Apr 22. [Epub ahead of print]

- Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia. Score J, Curtis C, Waghorn K, Stalder M, Jotterand M, Grand FH, Cross NC. Leukemia. 2006 May;20(5):827-32.

- Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. Jones AV, Kreil S, Zoi K, Waghorn K, Curtis C, Zhang L, Score J, Seear R, Chase AJ, Grand FH, White H, Zoi C, Loukopoulos D, Terpos E, Vervessou EC, Schultheis B, Emig M, Ernst T, Lengfelder E, Hehlmann R, Hochhaus A, Oscier D, Silver RT, Reiter A, Cross NC.Blood. 2005 Sep 15;106(6):2162-8. Epub 2005 May 26.

Signed: ……..

Date:…….

Acknowledgements

I would like to gratefully acknowledge the invaluable supervision, advice and support of Professor Nick Cross and Dr Francis Grand whilst completing this work. I would also like to acknowledge the help of Dr Andy Chase for all his additional technical advice and guidance. I am also thankful to have such tolerant and supportive laboratory colleagues at the Wessex Regional Genetics Laboratory, to name but a few; Amy Jones, Joannah Score, Kathy Waghorn, Thomas Ernst, Victoria Hall and Gemma Watkins. Their encouragement and friendship in all areas has been a great source of strength and comfort. Finally, I want to thank all those behind the scenes. My parents have been an endless source of love and support, believing that I can achieve anything I set my mind to. My brother, Dr Tim Curtis, whose footsteps I have followed (rather more slowly), has always been a great role model and someone to whom I can only aspire. A special thanks for understanding what a PhD entails, and for all your additional advice. Thanks must also go to Sue and Damian for their never-ending support. Last, but most definitely not least, my husband Marcos deserves exceptional gratitude for encouraging me throughout the whole course of my work whilst also putting up with my many crisis and mood swings! He is an endless source of joy and love, and has given me the most precious gift of all, our beautiful baby son Rafë.

LIST OF ABBREVIATIONS

ABI	Applied Biosystems
ABL	Abelson gene
aCML	atypical CML
AML	Acute Myeloid Leukaemia
ARG	Abelson related gene
ATP	Adenosine Triphosphate
BAC	Bacterial Artificial Chromosome
BAR domains	Bin/Amphiphysin/Rvs domains
BCR	Breakpoint Cluster Region
BIN2	Bridging integrator 2
BLAST	Basic local alignment search tool
BLAT	The BLAST-like alignment tool
bp	base pair
53BP1	53-Binding Protein 1
BRLMM	Bayesian Robust Linear Model with Mahalanobis
BTK	Bruton agammaglobulinemia tyrosine kinase
cDNA	complementary DNA
CEL	Chronic Eosinophilic Leukaemia
CEP110	Centriole associated protein 110
c-FMS	Colony stimulating factor 1 receptor
CFU-GM	Colony forming unit – granulocyte macrophage
CIP	Calf Intestinal phosphatase
CML	Chronic Myeloid Leukaemia
CML-CP	Chronic Myeloid Leukaemia – Chronic Phase
CMML	Chronic Myelomonocytic Leukaemia
CMPD	Chronic Myeloproliferative disorder
CNL	Chronic Neutrophilic Leukaemia
CNAT	GeneChip® Copy Number Analysis Tool
dATP	Deoxyadenosine 5'-triphosphate

dCTP	Deoxycytidine 5'-triphosphate
dGTP	Deoxyguanosine 5'-triphosphate
dHPLC	denaturing High Performance Liquid Chromatography
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide Triphosphates
DTT	Dithiothreitol
dTTP	Deoxythymidine 5'-triphosphate
EEN	SH3-domain GRB2-like
EGFR	Epidermal Growth Factor Receptor
EMS	Eight 'p11' myeloproliferative syndrome
EPHB4	Ephrin receptor B4
ERK	Extracellular signal-regulated kinase
ET	Essential Thrombocythaemia
ETV6	ets variant gene 6
FCS	Foetal Calf Serum
FGFR1	Fibroblast growth factor receptor 1
FGFR1OP2	FGFR1 oncogene partner 2
FGFR3	Fibroblast growth factor receptor 3
FIG	Golgi associated PDZ and coiled-coil motif
FIP1L1	FIP-1 Like-1 gene
FISH	Fluorescence *in situ* Hybridisation
FLT3	FMS-related tyrosine kinase 3
FOP	FGFR1 oncogene partner
G-CSF	Granulocyte-Colony stimulating factor
GOLGA4	Golgi autoantigen, golgin subfamily a4
GTC	Guanidium Thiocyanate
GTYPE	GeneChip® Genotyping Analysis Software
HB	Hybridisation Buffer
HBSS	Hanks Buffered Saline Solution
HES	Hypereosinophilic Syndrome
HIP1	Huntingtin interacting protein 1

HMM	Hidden Markov Model
IFNα	Interferon alpha
IL-3	Interleukin 3
IMF	Idiopathic Myelofibrosis
JAK2	Janus kinase 2
JMML	Juvenile Myelomonocytic Leukaemia
KI	Kinase Insert
KIF5B	Kinesin family member 5B
LOH	Loss of heterozygosity
LREC	Local Research Ethics Committee
Mb	Megabase
MDS	Myelodysplastic
MEK	MAPK [mitogen-activated protein kinase]/ERK kinase
MET	Hepatocyte growth factor receptor
MLL	Myeloid/lymphoid or mixed-lineage leukemia
MLPA	Multiplex Ligation-dependent Probe Amplification
MMLV	Moloney Murine Leukaemia Virus
MNCs	Mononuclear Cells
MPDs	Myeloproliferative disorders
MPL	Myeloproliferative leukemia virus oncogene
mRNA	messenger RNA
MYO18A	Myosin 18A
NCBI	National Centre for Biotechnology Information
NIN	Ninein
NTRKs	Non-receptor tyrosine kinases
PBS	Phosphate Buffered saline
PCR	Polymerase Chain Reaction
PDE4DIP	Phosphodiesterase 4D interacting protein (myomegalin),
PDGF	Platelet derived growth factor
PDGFR	Platelet derived growth factor receptor
PDGFRA	Platelet derived growth factor receptor alpha

PDGFRB	Platelet derived growth factor receptor beta
PE	Perkin Elmer
PI3K	Phosphatidylinositol-3-kinases
PLCγ	Phospholipase C gamma
PV	Polycythaemia Vera
5'RACE	5'Rapid Amplification of cDNA Ends
RCLB	Red Cell Lysis Buffer
RET	Ret proto-oncogene
RNA	Ribonucleic acid
RTKs	Receptor tyrosine kinases
RT-PCR	Reverse Transcriptase – PCR
SCF	Stem Cell Factor
SCLL	Stem cell leukaemia-lymphoma
siRNA	Short interfering RNA
SNP	Single nucleotide polymorphism
STATs	Signal transducer and activator of transcription
SYK	Spleen tyrosine kinase
TAP	Tobacco alkaline phosphatase
TGFβ	Transforming Growth Factor beta
TGN	*Trans* Golgi Network
TK	Tyrosine Kinase
TRKB	Neurotrophic tyrosine kinase receptor, type 2
UPD	Uniparental disomy
WBC	White Blood cells
WHO	World Health Organisation
ZNF198	Zinc Finger protein 198

1 INTRODUCTION

The focus of the Leukaemia Research group based at the Wessex Regional Genetics Laboratory in Salisbury is the study of a rare subset of neoplastic conditions known as the myeloproliferative disorders (MPDs) and in particular, the association between MPDs and deregulated tyrosine kinases. This introduction discusses normal haematopoiesis and its subversion leading to leukaemia, and specifically the involvement of mutations targeting tyrosine kinase genes.

1.1 Normal Haematopoiesis

Haematopoiesis is the dynamic process of blood cell formation from a common haematopoietic progenitor or stem cell, resulting in the production of a diverse range of cell types with very different functions, ranging from oxygen transport and clotting to the production of antibodies.

Haematopoietic cells first appear in the yolk sac in the second week of embryogenesis. By 8 weeks, haematopoiesis has become established in the liver and spleen of the embryo and by 12-16 weeks these are the major sites of haematopoietic cell formation and they remain active until birth. At 20 weeks gestation the bone marrow takes over as the active site for haematopoiesis, gradually increasing its activity until it then becomes the major site of production. In infants, haematopoiesis occurs in the bone marrow of all bones, whilst in adults haematopoiesis is confined to the flat bones of the sternum, ribs, cranium, pelvis, and the proximal ends of the femur and humerus.

The haematopoietic system itself consists of the haematopoietic stem cell and its progeny, and the supporting bone marrow stroma composed of endothelial cells, fat cells, fibroblasts, macrophages and an extracellular matrix. The stem cells are a self-

renewing population that have the potential to differentiate and become committed to a particular blood cell lineage. Initial differentiation of haematopoietic stem cells is along one of two major pathways, the myeloid (erythrocytes, neutrophils, monocytes/macrophages, eosinophils, basophils and megakaryocytes/platelets) or lymphoid (T-cells, B-cells and natural killer-cells) lineages (Figure 1.1). At this stage the stem cells have become progenitor cells for each type of blood cell, and have lost the capacity for self-renewal and are thus committed to a given cell lineage. These progenitor cells then divide and generate mature blood cells. Commitment of a progenitor cell to a given cell lineage depends upon the acquisition of responsiveness to certain growth factors. The particular microenvironment within which the progenitor cell resides controls differentiation. The bone marrow stroma is where haematopoiesis is controlled providing an inductive microenvironment consisting of an extracellular matrix and either membrane bound or diffusible growth factors (Dexter and Spooncer 1987). Evidence suggests that the dynamics of transcription factor networks within a cell, in addition to the combination of growth factors and cytokines in the environment are associated with commitment of a progenitor cell and hence cell diversity (Hoang 2004).

Haematopoiesis is a continuous process throughout adulthood, and indeed the numbers and distribution of the different cell types in the peripheral blood exhibit little variation under steady-state conditions. This process is regulated by the appropriate balance of cell division, differentiation and apoptosis to maintain a steady state. Apoptosis (also known as programmed cell death) is the process in which unwanted cells are removed. A great number of cells produced by haematopoiesis are destined to die by apoptosis before becoming fully differentiated (Williams, *et al* 1990). An imbalance between cell proliferation and cell death as a result of genetic mutations are observed in a number of haematological malignancies (Adjei 2001, Cucuianu 2001, Warmuth, *et al* 1999).

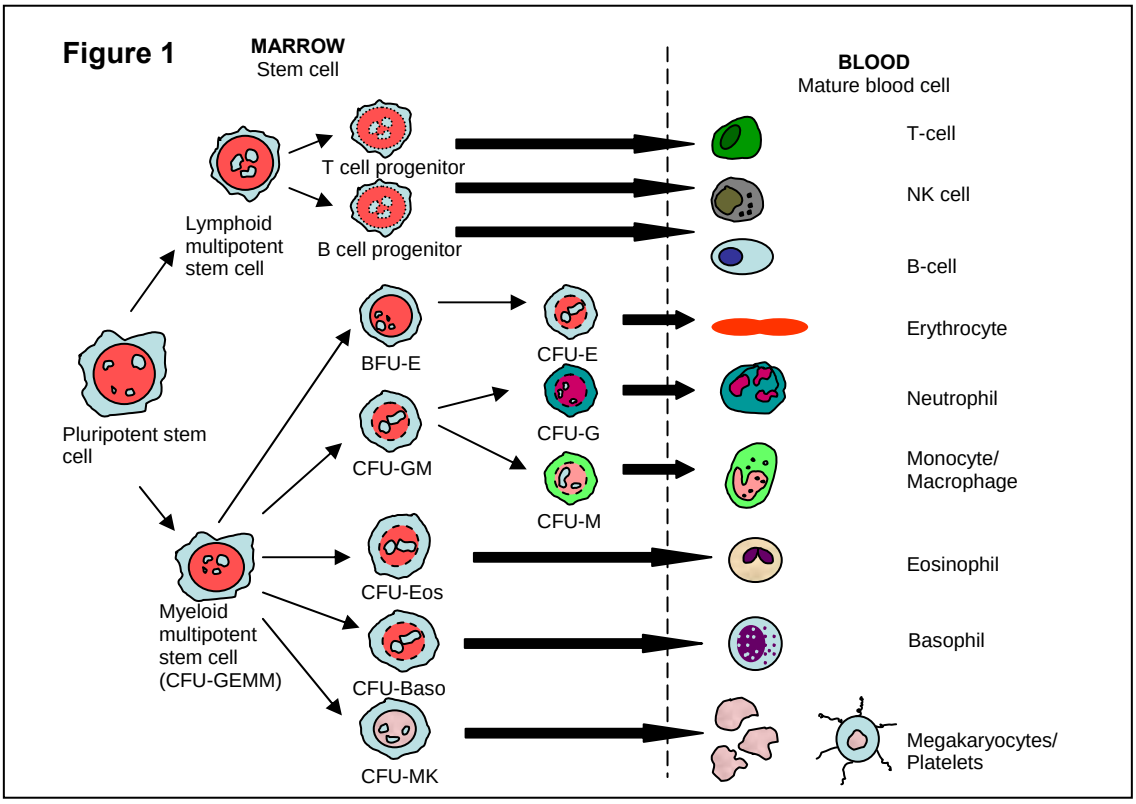


Figure 1. 1: A hierarchical model of haematopoiesis

1.2 Leukaemia

The term leukaemia comes from the Greek word meaning literally ‘white blood’ and was first described by Velpeau in 1827 (Velpeau 1827) but then took over a decade before the disease became recognised by other physicians (Bennett 1845, Cragie 1845, Virchow 1845). Leukaemia is defined as the clonal proliferation and accumulation of blood cells as a result of disruption to normal control mechanisms. The outcome of this is the bone marrow produces large numbers of abnormal cells, assumed to be derived from a single haematopoietic stem cell, thus disrupting the normal production of blood cells. The resultant release of immature cells into the peripheral blood and infiltration of organs such as the liver and spleen causing enlargement are common consequences of the disruption of normal haematopoiesis. Frequently other co-morbidities such as anaemia, thrombocytopenia and immunological dysfunction are also observed.

1.2.1 Clonality

There is a large amount of evidence to suggest that haematological malignancies arise from a single ancestral cell that has acquired a selective growth advantage due to acquired somatic mutations, and are thus clonal. In normal individuals haematopoiesis is thought to be polyclonal and this has been shown by clonality assays based on Lyon's hypothesis of X-inactivation (Lyon 1988). This states that there is random inactivation in females of one X-chromosome and the genes it carries during early embryogenesis. Therefore normal female tissue will be mosaic of cells expressing genes from one or other X-chromosomes whereas clonal tissue will express genes from one X-chromosome only. The assessment of clonality is now an essential element of the characterisation and definition of haematological malignancy. There are a number of different assays used to detect clonal derivation, most dependent on differential methylation patterns between active and inactive X-chromosomes at various polymorphic loci.

Clonality in leukaemia is now well established and is a basic concept for the definition of myeloid leukaemia (Fialkow*, et al* 1967, Raskind and Fialkow 1987, Wainscoat and Fey 1990). Most forms of leukaemia have demonstrated one variant of a polymorphic allele in all their leukaemic cells (Raskind*, et al* 1998). Another important step in defining clonality and leukaemia was demonstrated by the Philadelphia chromosome in CML patients, which is found in multiple cell types (neutrophils, basophils, monocytes and B-cells) suggesting that the disease derives from a stem cell rather than a myeloid progenitor (Fialkow*, et al* 1977).

1.2.2 Classification

Leukaemia is broadly classified as either acute or chronic, referring to the speed at which the leukaemia progresses. In chronic leukaemia the disease is both slow to

develop and progress and is characterised by the over-accumulation of mature cells. In acute leukaemia an amassing of immature haematopoietic blast or progenitor cells is seen due to the rapid progression of the disease. If left untreated, acute leukaemia is almost always fatal. Both classifications can be further subdivided into myeloid and lymphoid which refers to the phenotype of the affected cells.

Classification of the myeloid neoplasms has been defined by the World Health Organisation (WHO) (Vardiman, *et al* 2002, Vardiman, *et al* 2009). The focus of my work encompasses two of these categories: the chronic myeloproliferative disorders (CMPD) and the myelodysplastic/myeloproliferative disorders (MDS/MPD). Although these disorders are classified separately, the entities comprising each category are similar in that they are clonal stem cell proliferations with excess myeloproliferation. They exhibit increased marrow cellularity and organomegaly, and are susceptible to undergo genetic evolution associated with bone marrow failure or transformation to blast crisis (acute leukaemia). Despite these similarities, there are differences in the pathogenesis of each specific disease subtype that lead to clinical and morphological disparity.

1.3 The chronic myeloproliferative disorders (CMPD) and the myelodysplastic/myeloproliferative disorders (MDS/MPD)

The CMPD are haematological diseases that arise in a primitive bone marrow progenitor cell (Adamson, *et al* 1976, Fialkow, *et al* 1981, Fialkow, *et al* 1980, Jacobson, *et al* 1978). They are characterised by proliferation of one or more cells of the myeloid lineages (i.e., granulocytic, erythroid and megakaryocytic), believed to derive from a single neoplastic clone. The neoplastic haematopoiesis shows maturation followed by a corresponding increase in the number of one or more of the mature myeloid lineages in the peripheral blood (Figure 1.1). Splenomegaly and hepatomegaly are common features. These disorders are also prone to undergo

genetic evolution associated with bone marrow failure or transformation to blast phase. A finding of 10-19% blasts in the blood or bone marrow signifies disease acceleration and 20% or above indicates blast phase. There are seven entities identified within the CMPD classification: Chronic myelogenous leukaemia (CML), polycythaemia vera (PV), chronic idiopathic myelofibrosis (IMF), essential thrombocythaemia (ET), chronic neutrophilic leukaemia (CNL), chronic eosinophilic leukaemia (CEL) and unclassifiable CMPD.

The MDS/MPD group was created to accommodate disorders that have both MDS-like and MPD-like features, and as a consequence are difficult to assign to separate categories. They are also clonal haematopoietic neoplasms, however these disorders exhibit dysplastic and proliferative clinical, laboratory and morphological features at the time of initial presentation. They are characterised by hypercellular bone marrow with fibrosis, organomegaly and leukocytosis, commonly neutrophilic proliferation with or without monocytosis, similar to the CMPD. However, the cells are often morphologically dysplastic. Occasionally, simultaneous cytopenia of one or more of the other cell lineages is also observed due to ineffective proliferation or a decrease in the number of precursors. There are three main entities acknowledged in this classification: chronic myelomonocytic leukaemia (CMML), atypical chronic myeloid leukaemia (aCML) juvenile myelomonocytic leukaemia (JMML), in addition to unclassified MDS/MPD.

With the overlap of clinical, laboratory and morphological findings among the individual CMPD and MDS/MPD entities, distinction between the diseases in each category is made more difficult. For this reason, and the lack of knowledge surrounding the underlying molecular pathogenesis, the diagnosis of these disorders is based mainly upon the lineage of the predominant proliferating cells and the prominence of marrow fibrosis, combined with clinical, morphological and laboratory features. Due to the considerable overlap in the disease spectrum of these disorders I shall refer broadly to both groups together as myeloproliferative disorders (MPDs).

To date, there is no single molecular abnormality associated with this subset of disorders, although a pathognomonic molecular marker is associated with CML, and the V617F *JAK2* mutation is seen in the great majority of cases of PV as well as other MPDs.

1.4 The paradigm for deregulated tyrosine kinases - Chronic Myelogenous Leukaemia (CML)

CML is perhaps the best characterised MPD it has an incidence of 1-1.5/100,000 population annually (Faderl*, et al* 1999, Gunz 1977, Young*, et al* 1981). It is a bi- or tri-phasic disorder with an indolent chronic phase (CML-CP), followed by one or both of the aggressive transformed stages; accelerated phase and/or blast phase. CML usually presents in chronic phase, which is broadly defined by the absence of features of accelerated or blast phase. The haematological indications of CML-CP, as observed on blood smears, show leukocytosis due to the large number and variable maturation stages of neutrophils. Marrow blasts usually account for less than 2% of the white blood cells in the blood and absolute basophilia, a defining feature, is commonly present. The natural course of the disease progression from CML-CP to a more aggressive phase is generally within 3-5 years, and is thought to be due to genetic instability (Mitelman 1993).

1.4.1 Molecular pathogenesis of CML

A molecular marker was first described in 1960, when Nowell and Hungerford discovered a 'minute chromosome' associated with the condition, a shortened chromosome 22 (Nowell and Hungerford 1960). It was not until much later in 1985 that the underlying molecular lesion was defined (Shtivelman*, et al* 1985). The hallmark of CML is the Philadelphia (Ph) chromosome, a reciprocal translocation,

t(9;22)(q34;q11), that fuses *BCR* on chromosome 22 with the *ABL* tyrosine kinase on chromosome 9 (de Klein, *et al* 1982, Faderl, *et al* 1999, Gunz 1977, Rowley 1973, Shtivelman, *et al* 1985, Young, *et al* 1981). The *ABL* gene on chromosome 9 has homology with a known viral oncogene v-*abl*, which was identified as the cause of some murine leukaemias. The normal *ABL* gene encodes a non-receptor tyrosine kinase that is active in signal transduction. The chimaeric *BCR-ABL* gene however encodes an aberrant tyrosine kinase of higher molecular weight and depending on the location of the breakpoint within *BCR* (Quackenbush, *et al* 2000), three oncogenic BCR-ABL fusion proteins: p190, p210 and p230 can be formed. They are all found to be sufficient to generate leukaemia in murine models (Daley, *et al* 1990, Huettner, *et al* 2000). Fusion to BCR constitutively activates the ABL tyrosine kinase moiety, and this activation depends on the BCR N-terminal coiled-coil domain (McWhirter and Wang 1997). The effect of this is autophosphorylation of the protein and activation of a number of cellular pathways important for transducing proliferative and anti-apoptotic signals (McGahon, *et al* 1994, Puil, *et al* 1994). It has been suggested that the strong resistance to apoptosis along with DNA repair deficiency induced by *BCR-ABL* allows the accumulation of secondary genetic abnormalities leading to clonal evolution toward blast crisis. Cytogenetically, such abnormalities typically include +8, +Ph or i(17q).(Barnes and Melo 2002, Faderl, *et al* 1999, Mitelman 1993, Mitelman, *et al* 1976, Sokal, *et al* 1988). Abrogation of apoptosis suggests that the expansion of the neoplastic clone may be due not only to a proliferative advantage but also to prolonged survival.

The Ph chromosome is now detected by routine cytogenetic techniques in 85-90% of CML patients at diagnosis. The remaining patients have variant translocations or a submicroscopic translocation of *BCR-ABL* that can only be detected by molecular techniques or fluorescence *in situ* hybridisation (FISH) (Hooberman, *et al* 1989, Sinclair, *et al* 1997, Werner, *et al* 1997, Westbrook and Keinanen 1992).

1.4.2 Therapy in CML and the development of Imatinib

Allogeneic stem cell transplantation is currently believed to be the only curative therapy for CML, however only a minority of patients is eligible for this procedure due to a relatively limited number of HLA-matched donors. In the past, the therapy most commonly used to achieve haematological remission in CML-CP was interferon-alpha (IFNα). However, an important advance in the therapy of CML came with the development of imatinib mesylate, a small molecule inhibitor that competes with ATP for occupancy of its specific binding site within the kinase domain of BCR-ABL. Inhibition of tyrosine kinase activity occurs, as imatinib lacks the essential phosphate groups usually provided by ATP. Thus, phosphorylation of substrate proteins cannot occur and they therefore fail to adopt the conformational changes necessary for binding of effector proteins. Ultimately, downstream signalling is prevented (Druker*, et al* 1996). The initial clinical trials of imatinib included 83 Ph(+) CML patients who had failed IFNα-based treatment and were in CML-CP. A complete haematological response was seen in 98% of patients, with complete cytogenetic responses in 13%. These initial studies were further expanded to include 260 CML patients in myeloid blast crisis. The overall response rate was 64%, with 11% achieving complete remission (<5% marrow blasts). However, 38% of patients either returned to CML-CP or had partial responses.

The results from blast crisis CML suggest partial dependence of the *BCR-ABL* clone and the likely acquisition of further molecular abnormalities that are not targeted by imatinib (Druker*, et al* 2001). O'Brien and colleagues later went on to report a comparative study of imatinib mesylate versus combination therapy with IFNα and cytarabine in newly diagnosed CP-CML patients, and found that 97% of these patients achieved a complete haematological remission by 18 months after imatinib mesylate was employed. In addition, a complete cytogenetic response was achieved in greater than 76% of CML patients, and more than 90% had no evidence of disease progression after 18 months (O'Brien*, et al* 2003). However, to properly assess the

overall survival increases after treatment with imatinib, long-term studies are required.

As previously mentioned, although patients with accelerated or blast phase CML show similar morphological and clinical changes with imatinib therapy, it is not effective in advanced disease (Druker, *et al* 2001, Druker, *et al* 1996). Furthermore, a drawback of imatinib is the emergence of acquired resistance, such as point mutations in the kinase domain or amplification of the *BCR-ABL* gene (Barthe, *et al* 2001, Branford, *et al* 2002, Gorre, *et al* 2001, Hochhaus, *et al* 2001, Mahon, *et al* 2000, von, *et al* 2002). The development of second generation ATP-competitive ABL kinase inhibitors are attempting to overcome some of the problems associated with imatinib resistance and advanced phase disease. These compounds include dasatinib (Sprycel®), nilotinib (Tasigna®), bosutinib (SKI-606) and INNO-406 (NS-187), all of which display increased potency against BCR-ABL. Nilotinib is a novel aminopyrimidine inhibitor (Weisberg, *et al* 2005) and has been developed as a selective potent BCR-ABL inhibitor. It has shown a good outcome in patients with CP-CML following imatinib resistance and intolerance (Kantarjian, *et al* 2007), in fact it has been shown in pre-clinical studies to be 30-fold more potent than imatinib. Dasatinib is a potent dual BCR-ABL and Src kinase inhibitor that is structurally unrelated to imatinib (Shah, *et al* 2004). Its greatest benefits have been observed in patients with CP-CML with 90% achieving a complete haematological response (Steinberg 2007). Much like imatinib, advanced phase-, blast crisis-CML or Ph positive ALL, all showed lower responses. A clinical trial comparing high dose imatinib (800mg/day) and dasatinib showed both produced complete haematological responses however, more achieved a major haematological response with dasatinib (32% versus 7%). Dasatinib has also been found to be effective against KIT and PDGFR, in addition to overcoming resistance exhibited by 18/19 imatinib resistant mutations reported (Shah, *et al* 2006). Neither of the second generation targeted compounds mentioned here are able to overcome the resistance displayed by the T135I mutant but nevertheless they offer the potential for combined therapy and additional hope in cases where imatinib resistance or intolerance has been observed.

Other therapies are also continuing to be developed, such as those that target downstream signalling molecules using farnesyl transferase inhibitors (Peters, *et al* 2001). It is unclear exactly how these compounds work however they are thought to inhibit the farnesylation of a wide range of target proteins, including *RAS*, ultimately resulting in cell growth arrest. In preclinical models these inhibitors showed great potency against tumour cells, however in clinical trials their activity was far less than expected. It is possible that inhibition of farnesylation is sufficient for deactivation of H-Ras, yet insufficient for K-Ras which may be activated by geranylgeranylation thereby suppressing the effect of the farnesyl transferase inhibitors (Appels, *et al* 2005). Another area attracting the development of many haematological based inhibitors is apoptosis-centered therapies. Many agents have been developed that generally act by stimulating pro-apoptotic proteins and which are at various stages of preclinical and clinical development,. In particular, compounds have been devised that act against Akt, over activity of which can enable leukaemic cells to proliferate and survive. Akt is activated downstream of several oncogenes including BCR-ABL. An Akt inhibitor has been described that is relatively selective but unfortunately in preclinical trials proved weak (Mitsiades, *et al* 2002). Heat shock 90 (HSP90) inhibitors such as 17-AAG have also been shown to have an effect on Akt by acting to destabilise and cause overall reductions in Akt levels (Solit, *et al* 2003).

BCR-ABL was the first human fusion gene to be identified in human malignancy. The leukaemogenic mechanism in CML now serves as a paradigm for other myeloid leukaemias, as subsequent to the discovery of *BCR-ABL*, many other oncogenic fusion genes have been reported, including several that involve tyrosine kinases.

1.5 The tyrosine kinases

Tyrosine kinases are enzymes that catalyse the transfer of the γ-phosphate of ATP to protein tyrosine substrates. The phosphorylation of tyrosine and to a lesser extent

threonine and serine is the major mechanism for reversible regulation of protein activity and function. Phosphorylation by protein kinases is central to cellular signal transduction and is responsible for a number of cell functions, including cell-cycle progression, differentiation, apoptosis, cell metabolism and DNA-replication. There are thought to be 518 putative protein kinase genes within the human genome (Manning, *et al* 2002) of which 90 are tyrosine kinases. These tyrosine kinases are classified as either non-receptor or receptor types (Blume-Jensen and Hunter 2001). The receptor tyrosine kinases (RTKs) contain a transmembrane domain and are involved in extracellular ligand binding. The non-receptor tyrosine kinases (NRTKs) are intracellular and do not span the plasma membrane. There are 58 RTKs and 32 NRTKs in the human genome.

1.5.1 Non-receptor tyrosine kinases (NRTKs)

The NRTKs include SRC, JAK and ABL which are intrinsic to the cytosol and are maintained in an inactive state. Despite performing a similar role in signal transduction, apoptosis and differentiation as the receptor tyrosine kinases, activation is usually dependent on interaction with surface receptors that lack intrinsic catalytic activity. For example, the JAK family of NRTKs is involved in signal transduction by the cytokines IL-3 and G-CSF (Figure 1.2). The NRTKs can also be activated by intracellular signalling by enabling inhibitor dissociation which is how they are usually maintained in an inactive form. Interestingly, while most NRTKs are cytoplasmic, ABL is localised to both the cytosol and nucleus (Taagepera, *et al* 1998).

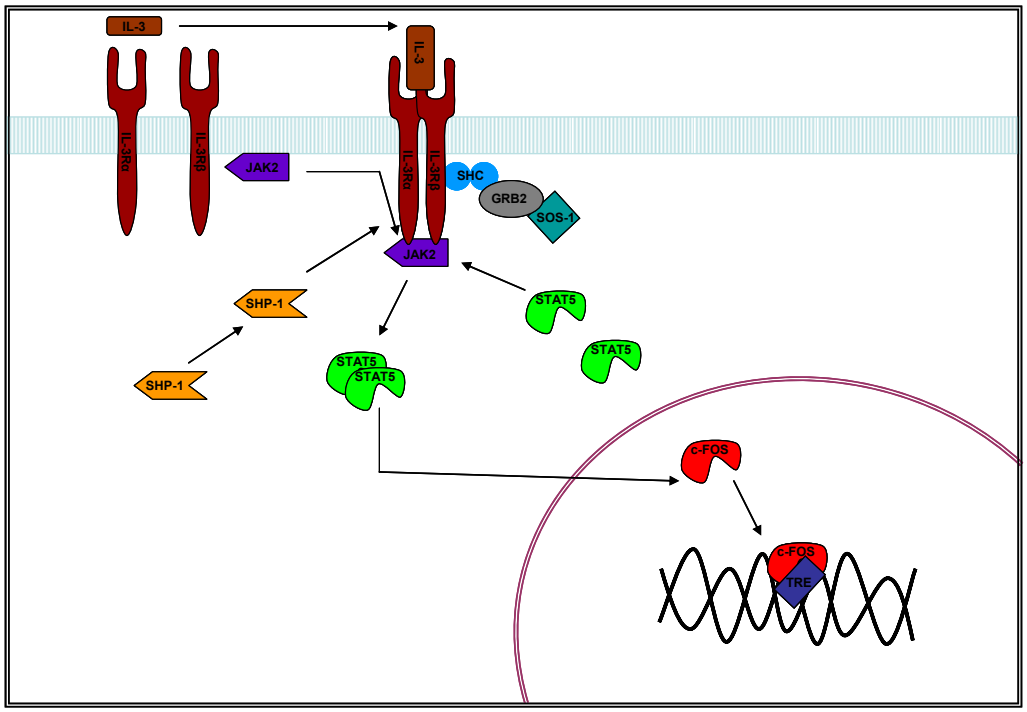


Figure 1. 2: IL-3 signalling pathway detailing JAK2 non-receptor tyrosine kinase signalling. JAK2 binds to the active cytokine receptor IL-3. JAK2 in turn phosphorylates tyrosine residues on the receptor allowing the recruitment of other signalling proteins, such as SHC and STAT5, which are then phosphorylated by JAK2. Activated (phosphorylated) STAT dimers can then accumulate in the cell nucleus and activate transcription.

1.5.2 Receptor tyrosine kinases (RTKs)

The RTKs are transmembrane enzymes that are involved in signal transduction at the cell surface and which have been sub-classified into 20 classes based on their kinase domain sequence (Robinson, *et al* 2000). The class III RTKs includes KIT, FLT3, PDGFRα, PDGFRβ and FMS. They are characterised by five immunoglobulin-like domains in the extracellular ligand-binding region, a single transmembrane domain and two intracellular tyrosine kinase domains (TK1 and TK2) separated by a kinase insert domain (KI) (Schlessinger 2000). They are all thought to play a crucial role in normal haematopoiesis and are thus normally tightly regulated due to the impact they

have on cellular signal transduction. In their inactive states the receptors are thought to exist as monomers embedded in the plasma membrane. Ligand binding to the receptor induces dimerisation, juxtaposing the two catalytic domains causing a conformational change resulting in partial activation of its enzymatic activity (Jiang and Hunter 1999). Full activity follows only after autophosphorylation of key tyrosine residues in the activation loop. These phosphorylated tyrosines then act as docking sites for adaptor proteins that further recruit other downstream cytoplasmic signalling molecules. Subsequently a cascade of signalling events through the JAK/STAT, PLCγ, PI3K and Ras-Raf-MEK/ERK occurs leading to changes in expression of genes involved in the control of cell proliferation, differentiation and cell survival (Kharas and Fruman 2005) (Figure 1.3). Activated receptors are rapidly endocytosed and either degraded or recycled to the plasma membrane.

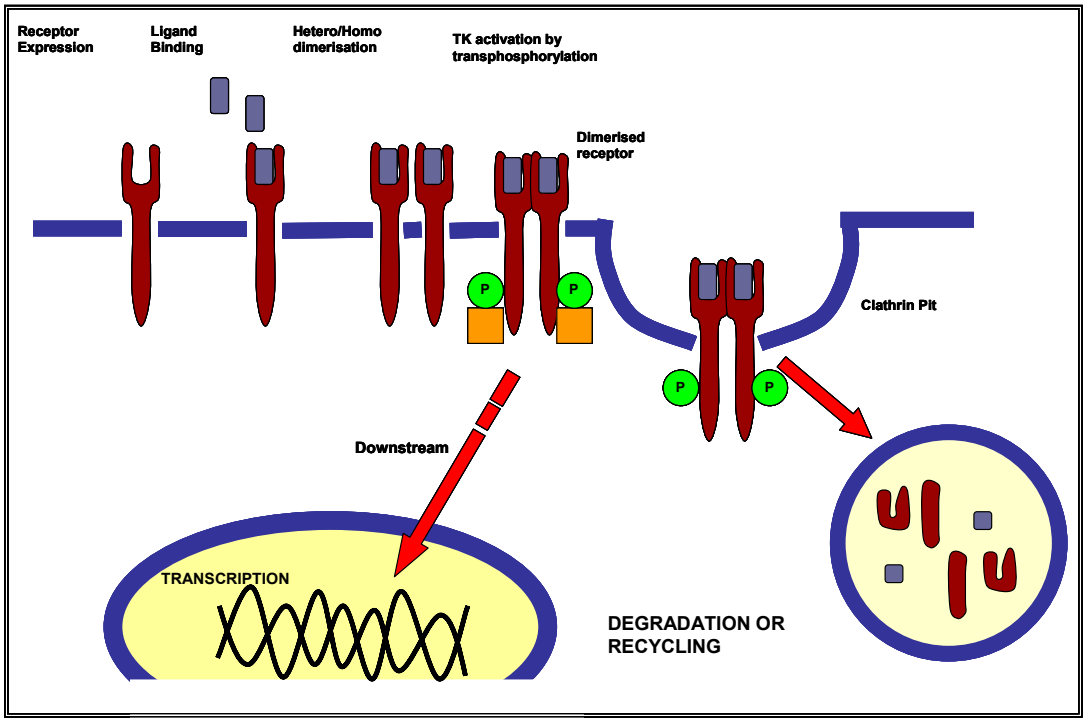


Figure 1. 3: A schematic representation of RTK activation and signalling.

1.6 Molecular pathogenesis and deregulated tyrosine kinases

Following the identification of BCR-ABL in CML, a number of other tyrosine kinases have been implicated in the pathogenesis of leukaemia. In fact, in the past two decades several mechanisms that deregulate tyrosine kinases, such as receptor amplification, chromosomal translocations and point mutations have been identified, with reports of more that 30 deregulated tyrosine kinases associated with cancer (Blume-Jensen and Hunter 2001). In MPD patients in particular, translocation cloning has lead to the identification of numerous fusion genes involving tyrosine kinase genes, for example the fibroblast growth factor receptor 1 (*FGFR1*) on chromosome 8p11. The resulting myeloproliferative disorder is loosely referred to as the 8p11 myeloproliferative syndrome (EMS), also known as stem cell leukaemia-lymphoma (SCLL) (Macdonald, *et al* 2002a).

To date, nine fusion genes associated with translocations involving *FGFR1* have been identified in MPD patients. The translocations reported include chromosomes 13q, 9q, 6q, 22q, 17q, 7q, 12p, 19q and 12q resulting in the respective fusion of *ZMYM2 (ZNF198/RAMP/FIM*), *CEP110*, *FGFR1OP (FOP)*, *BCR*, *MYO18A, Trim24 (TIF1), FGFR1OP2, HERV-K* and *CPSF6* to *FGFR1* (Belloni, *et al* 2005, Demiroglu, *et al* 2001, Grand, *et al* 2004a, Guasch, *et al* 2003a, Guasch, *et al* 2000, Hidalgo-Curtis, *et al* 2008, Popovici, *et al* 1999, Walz, *et al* 2005, Xiao, *et al* 1998). As a consequence, constitutively active fusion proteins are produced that are functionally and structurally similar to BCR-ABL. Other translocations involving tyrosine kinases are reported in MPD patients, including the platelet-derived growth factor receptors alpha (*PDGFR*A) and beta (*PDGRFB)*, class III receptor tyrosine kinases that map to 4q12 and 5q33, respectively (discussed later, section 1.7). Other fusions have also been described in MPDs that involve *JAK2, ABL, FLT3, SYK, ARG* and *TRKC (NTRK3)* (Grand, *et al* 2007, Griesinger, *et al* 2005, Peeters, *et al* 1997, Reiter, *et al* 2005, Vu, *et al* 2006) (Cazzaniga, *et al* 1999, Eguchi, *et al* 1999, Golub, *et al* 1997, Hidalgo-Curtis, *et al* 2008, Kuno, *et al* 2001). The resulting oncoproteins also have

constitutive enzymatic activity and deregulate cell growth in a similar fashion to that generated by BCR-ABL (Cross and Reiter 2002).

In addition to translocation cloning, loss of heterozygosity (LOH) has recently re-emerged as an extremely useful tool for identifying regions harbouring potential candidate oncogenes or tumour suppressor genes in leukaemia. This is exemplified by the recent discovery of the V617F *JAK2* mutation in patients with PV, MF and ET. A study of PV patients in 2002 using microsatellite screening of the whole genome, showed significant loss of heterozygosity (LOH) of chromosome 9p in all PV patients tested (Kralovics, *et al* 2002). LOH was observed without change of copy number, indicating that it occurred by acquired uniparental disomy (UPD) as a consequence of mitotic recombination rather than heterozygous deletion. Within the minimal region of UPD they identified a number of possible candidate genes, including *JAK2*. UPD for other regions has also been associated with gain of function mutations, for example UPD 13q and *FLT3* internal tandem duplications (Fitzgibbon, *et al* 2005, Griffiths, *et al* 2005). Although these results did not actually lead to the discovery of the *JAK2* or *FLT3* mutations, they suggest that screening for LOH/UPD may be a useful approach to identify candidate oncogenes.

At the current time it is important to identify patients with molecular lesions involving imatinib responsive genes. Although designed for use against the *ABL* tyrosine kinase, imatinib has also been shown to inhibit the tyrosine phosphorylation of *ARG*, *KIT*, *FMS*, *PDGFRA* and *PDGFRB*. On identification of a lesion involving one of these genes in a MPD patient, imatinib therapy can be successfully initiated. Translocations involving imatinib responsive genes are commonly described in a minority of MPD patients, notably involving rearrangements of the *PDGFR* genes.

1.7 PDGFRA and PDGFRB rearrangements in MPDs

PDGFRα and PDGFRβ are 2 highly related RTKs with 85% and 75% homology between the two intracellular kinase domains (Matsui, *et al* 1989). PDGF ligand consists of a family of four dimeric isoforms linked by disulphide bonds, PDGF-AA, PDGF-BB, PDGF-CC and PDGF-DD. The N-terminal region of each ligand must be proteolytically removed to enable binding to the receptor (Bergsten, *et al* 2001). PDGFRβ binds the ligands PDGF-BB and PDGF-DD agonistically, whereas PDGFRα binds all the isoforms with the exception of PDGF-DD. While PDGFRs can form homo-and heterodimers between the α and β receptors, it has also been shown to have the ability to dimerise with the epidermal growth factor receptor (EGFR) and the hybrid can also be stimulated by PDGF (Saito, *et al* 2001). Both PDGFR receptors have been linked to haematological malignancies and in particular the MPDs.

1.7.1 Fusions involving PDGFRA

Over-expression and amplification of *PDGFRA* has been described in only a minority of human malignancies (Bozzi, *et al* 2007, Ebert, *et al* 1995, Hermanson, *et al* 1996, Holtkamp, *et al* 2007, MacDonald, *et al* 2001, Puputti, *et al* 2006, Zhang, *et al* 2005), but specific abnormalities have been associated with patients with a CML-like MPD. A rare variant translocation, t(4;22)(q12;q11), resulting in a constitutively active BCR-PDGFRα fusion protein was first detected in a patient with atypical CML (Baxter, *et al* 2002). Following this discovery, Cools et. al., (2003) discovered an interstitial cytogenetically invisible deletion of 4q12 resulting in the fusion of *FIP1L1* to *PDGFRA.* The resulting oncoprotein was found to transform Ba/F3 cells to growth factor independence, similar to BCR-ABL. Although originally found in a patient described clinically with HES, patients with *FIP1L1-PDGFRA* are now classified as CEL since the presence of the fusion indicates clonality (Cools, *et al* 2003a). More recently, the link between *PDGFRA* fusion genes and CML-like MPDs has been

heightened by further reports of new rare *PDGFRA* fusions (Curtis, *et al* 2007, Score, *et al* 2006, Walz, *et al* 2006).

1.7.2 Chronic eosinophilic leukaemia (CEL)

CEL is the clonal proliferation of eosinophilic precursors that results in persistently increased numbers of eosinophils in the blood, bone marrow and peripheral tissues. Organ damage occurs because of leukaemic infiltration or the release of cytokines, enzymes or other proteins by the eosinophils. Unusually, eosinophilic leukaemia shows a marked male predominance, and in particular, HES has been found to have a 9:1 male to female ratio. Due to the difficulty in differentiating CEL and HES, the prevalence of these disorders is unknown, although it is clear that both are rare. However if there is no evidence of clonality of the eosinophils, a diagnosis of HES is preferred, whereas the finding of a clonal myeloid abnormality supports the diagnosis of CEL (Bain 1996, Weller and Bubley 1994). The *FIP1L1-PDGFRA* cytogenetically cryptic fusion gene was first identified in patients with HES/CEL. Patients with the fusion protein show an excellent clinical response to imatinib mesylate (Cools, *et al* 2003a).

A staurosporine derivative, PKC412, has also been developed to inhibit the PDGFRs and other kinases such as KIT and FGFR3 (Chen, *et al* 2005, Gotlib 2005, Growney, *et al* 2005, Seo, *et al* 1999), and although it has not been used in patients with PDGFR-rearrangements yet, it has been shown to overcome imatinib-resistance in a murine model of FIP1L1-PDGFRα induced MPD (Cools, *et al* 2003b). It has also now entered Phase II trials for FLT3 inhibitor therapy in AML with encouraging results however the emergence of primary and secondary resistance does occur in the majority of patients (Stone, *et al* 2005). Other inhibitors such as Sorafenib are also active against FIP1L1-PDGFRA and its most common imatinib-resistant mutant (Lierman, *et al* 2006).

A number of other clonal cytogenetic abnormalities have also been reported in rare cases of CEL, CMML with eosinophilia and aCML primarily involving *PDGFRB*,

FGFR1 and *JAK2*. Acquired reciprocal chromosomal translocations involving the chromosome band 5q31-33 (*PDGFRB*), in particular, are associated with a significant minority of MPD patients.

1.7.3 Fusions involving PDGFRB

Aberrant activation of *PDGFRB* resulting in cell transformation was first substantiated by Waterfield et. al., (1983) (Waterfield*, et al* 1983) who provided evidence that SIS, an oncogenic form of the PDGFR ligand, transformed cells through activation of its receptor. Nevertheless, demonstration of the direct association with *PDGFRB* in leukaemogenesis took another decade, when a rare but recurring cytogenetic abnormality was reported in patients with atypical CML. The abnormality was shown to be a t(5;12)(q33;p13) resulting in an *ETV6-PDGFRB* fusion gene (Golub*, et al* 1994). In an analogous manner to the BCR-ABL oncoprotein, pathogenicity is dependent on both the ETV6 dimerisation motif and the tyrosine kinase activity of PDGFRβ. Reports of additional chromosomal translocations involving *PDGFRB* in patients with MPDs, strengthens the association between myeloid leukaemias and deregulated tyrosine kinases. To date, 21 partner genes have been reported (Table 1.1).

TRANSLOCATION	PARTNER GENE	REFERENCE
t(5;12)(q33;p13)	*ETV6*	(Golub, *et al* 1994, Wlodarska, *et al* 1995)
t(5;10)(q33;q21)	*H4/CCDC6*	(Kulkarni, *et al* 2000, Schwaller, *et al* 2001)
t(5;7)(q33;q11)	*HIP1*	(Ross, *et al* 1998)
t(5;14)(q33;q32)	*CEV14(TRIP11/GMAP210)*	(Abe, *et al* 1997)
t(5;17)(q33;p13)	*Rabaptin-5/RABEP1*	(Magnusson, *et al* 2001)
t(1;5)(q23;q33)	*PDE4DIP (myomegalin)*	(Wilkinson, *et al* 2003)
t(5;14)(q33;q24)	*NIN*	(Vizmanos, *et al* 2004)
t(5;17)(q33;p11.2)	*HCMOGT-1/SPECC1*	(Morerio, *et al* 2004)
t(5;15)(q33;q15)	*TP53BP1*	(Grand, *et al* 2004b)
t(5;14)(q33;q32)	*KIAA1509*	(Levine, *et al* 2005b)
t(5;12)(q33;p13.3)	*ERC1*	(Gorello, *et al* 2008)
t(1;5)(q21;q33)	*TPM3*	(Rosati, *et al* 2006)
t(3;5)(q22;q33)	*WDR48*	SUBMITTED
t(5;12)(q33;q13.13)	*Bin2*	SUBMITTED
t(3;5)(q22;q33)	*GOLGA4*	SUBMITTED
t(5;12)(q31-33;q24)	*GIT2*	(Walz, *et al* 2007)
t(4;5;5)(q23;q31;q33) t(4;5)(q21;q33)	*PRKG2*	(Gallagher, *et al* 2008, Walz, *et al* 2007)
Der(1)t(1;5)(p34;q33),der(5)(1;5)(p34;q15), der(11)ins(11;5)(p12;q15q33)[23];46XY[1]	*GPIAP1*	(Walz, *et al* 2007)
t(5;16)(q33;p13)	*NDEI*	(La Starza, *et al* 2007)
t(5;17)(q33-34;q11.2)	*MYO18A*	(Walz, *et al* 2009)
t(2;5)(p21;q33)	*SPTBN1*	(Gallagher, *et al* 2008)

Table 1. 1: 5q33 translocations in leukaemia involving *PDGFRB*

1.7.4 Chronic myelomonocytic leukaemia (CMML)

CMML is a clonal disorder of a bone marrow stem cell, in which monocytosis is the hallmark. It is, however, a heterogeneous disorder; in some cases, the white blood cell (WBC) count is elevated due to neutrophilia and monocytosis with minimal dysplasia, but significant splenomegaly, and the disease is more reminiscent of a CMPD than MDS. In other patients, although monocytosis is present, neutrophil counts are reduced or normal and there is prominent dysplasia and no splenomegaly, so that the disease resembles MDS rather than CMPD. The median age at diagnosis of CMML is 65-75 years, with a male to female ratio reported to be between 1.5-3:1 (Germing*, et al* 2000, Onida*, et al* 2002). The most frequent recurring chromosomal abnormalities include trisomy 8, monosomy 7/deletion (7q) and structural abnormalities of 12p and i(17q); approximately 20-40% of CMML patients have detectable clonal abnormalities (Fenaux*, et al* 1996, Toyama*, et al* 1993).

A subset of CMML cases, generally those associated with marked eosinophilia, have also been associated with *PDGFRB* fusion genes (Golub*, et al* 1994, Kulkarni*, et al* 2000, Magnusson*, et al* 2001, Sawyers and Denny 1994, Schwaller*, et al* 2001). Mutations of *RAS* are also detected, in approximately a third of CMML cases (Willman 1998). Mutations of *FLT3* may also be present in this group, a recent study by Shih et. al, (2002) found 3 out of 7 CMML patients carried the internal tandem duplication (Shih*, et al* 2002).

Apperley et. al., (2002) has also provided evidence that patients with rearrangements of the *PDGFRB* gene respond well to treatment with imatinib mesylate (Apperley*, et al* 2002). These findings make it increasingly important to gain an accurate molecular diagnosis of *PDGFRA* and *PDGFRB* gene rearrangements, as it is likely they will play a role in the disease management of such patients.

1.7.5 JAK2 and the myeloproliferative disorders

Aberrant tyrosine kinase signalling has been implicated as the pathogenic mechanism of so many of the MPDs, it has not come as a surprise to find the presence of an acquired point mutation in the *Janus Kinase 2* (*JAK2*) gene predominantly in patients with PV, IMF and ET. The *JAK2* mutation itself has been the single most important discovery in the pathogenesis of MPDs in the last decade. The guanine to thymine substitution in exon 12 of *JAK2* changes a highly conserved valine at position 617 to phenylalanine (V617F). This alteration is within the pseudokinase (JH2) domain and acts by disrupting the inhibitory regulation of *JAK2*, leaving it constitutively activated (Baxter*, et al* 2005, James*, et al* 2005, Jones*, et al* 2005, Kralovics*, et al* 2005a, Levine*, et al* 2005a). The activated kinase can then bind to a receptor and recruit STATs independent of haematopoietic growth factors. Five groups have independently confirmed the identification of the V617F mutation. The numbers of PV, IMF and ET patients found to have the mutation is remarkable, and will now act as an invaluable diagnostic tool to distinguish between the MPD and other reactive conditions such as secondary thrombocytosis and erythrocytosis (Table 1.2). Interestingly, this mutation does not occur in non-haematological malignancies and is uncommon in myeloid disorders other than the classic *BCR-ABL* negative MPDs (Scott*, et al* 2005).

	Baxter et.al., Lancet (2005)		Levine et. al., Can. Cell (2005)		Kralovics R et. al., NEJM (2005)		Jones et. al., Blood (2005)		James C et. al., Nature (2005)	
Disease	**No. of cases**	**No. of V617F positive**	**No. of cases**	**No. of V617F positive**	**No. of cases**	**No. of V617F positive**	**No. of cases**	**No. of V617F positive**	**No. of cases**	**No. of V617F positive**
Polycythaemia vera	73	71(97%)	164	121(74%)	128	83(65%)	72	58(81%)	45	40(89%)
Essential thrombocythaemia	51	29(57%)	115	37(32%)	93	21(23%)	59	24(41%)	7	3 (43%)
Chronic idiopathic myelofibrosis	16	8(50%)	46	16(35%)	23	13(57%)	35	15(43%)	21	9(43%)

Table 1. 2: Percentages of patients found to have the V617F mutation categorised by disease entity.

However, a number of questions regarding *JAK2* V617F remain including: how an identical mutation can cause three separate clinical diseases, and why some patients are heterozygous for the mutation and others are homozygous. If those with homozygous mutations are the result of two separate steps, why is it that the first is enough to cause disease features alone?

Since the discovery of the *JAK2* V617F mutation a considerable amount of investigation has been performed trying to answer these, among other questions, with no definitive conclusions. It is odd that only a single point mutation occurs in this gene at this position. If this codon is prone to mutation why is only one amino acid substitution observed? What makes this single mutation activate *JAK2* in the majority of patients with these disorders given the fact that other activating mutations such as T875N have been found in acute megakaryoblastic leukemia (Mercher*, et al* 2006). Work has recently been published looking at this anomaly: Dusa et. al., 2008 have used random site-directed mutagenesis to mutate codon 617 to each of the 20 amino acid possibilities and tested their ability to induce constitutive signalling *in vitro* in Ba/F3 cells expressing the erythropoietin receptor. Only 4 mutations, V617W, V617M, V617I and V617L were able to induce cytokine independence and constitutive downstream signalling. Additionally, V617W was the only mutant that produced constitutive activity comparable to that of V617F, and could generate an MPD in mice that displayed predominant erythrocytosis and megakaryocytic proliferation (Dusa*, et al* 2008). Nevertheless, to achieve this amino acid substitution would require a three base pair change (GTC>TTG) and this is unlikely to occur naturally in patients.

Since this initial finding of V617F, other mutations have been identified in *JAK2* exon 12 in PV patients negative for the *JAK2* V617F mutation. In contrast, these mutations are varied in terms of their molecular basis, with substitutions, complex deletions and insertions identified (Butcher*, et al* 2008, Scott*, et al* 2007b). These mutations account for one third of all bona fide V617F negative PV patients although

this figure will probably turn out to be greatly reduced with screening of increased numbers (Scott, *et al* 2007a).

In vitro and *in vivo* data suggests *JAK2* V617F is sufficient for the development of PV, ET and IMF, however this does not account for how three phenotypically related, but clinically distinct MPDs can arise from an identical *JAK2* mutation. Additional genetic and epigenetic events have been hypothesised to contribute to this phenotypic pleiotropy. It is possible there are still unidentified mutations that contribute to the pathogenesis of these disorders or there may be unknown initiating events which precede the acquisition of the *JAK2* V617F mutation. Pardanani et. al., 2007 screened for genetic differences in allele frequencies of SNPs in *JAK2*, *EPOR*, *GSCF* and *MPL* and found a number of SNP and haplotype associations linked with PV and not ET or IMF in the *JAK2* gene. This group concluded that genetic variation in MPDs may contribute to phenotypic diversity in the presence of specific mutations (Pardanani, *et al* 2007). In addition, a study by Bellanne-Chantelot et. al., 2006 looked at familial MPDs. Somatic *JAK2* V617F mutations were identified in some, but not all MPD patients. In fact some families showed both *JAK2* V617F-positive and -negative affected members, and a small number of relatives without a diagnosis of PV, ET or MF had endogenous erythroid colony formation in the absence of the *JAK2* V617F allele, suggesting perhaps that unidentified mutations may be contributing to the pathogenesis of PV, ET and MF, regardless of *JAK2* mutational status (Bellanne-Chantelot, *et al* 2006). In addition to phenotypic pleiotropy and inherited SNPs, epidemiological data and family studies have also suggested that inherited factors may also predispose to these specific MPDs (Landgren, *et al* 2008, Rumi, *et al* 2008). This work has been expanded and a strong association between *JAK2* V617F disease and specific constitutional JAK2 haplotypes has recently been identified, providing the first evidence in MPDs of inherited genetic factors influencing the risk of acquiring a specific somatic mutation (Jones, *et al* 2009, Kilpivaara, *et al* 2009, Olcaydu, *et al* 2009).

It has also been suggested that the *JAK2* V617F mutation may arise in different progenitor cell populations thus accounting for the difference in clinical phenotype

between the three separate disorders found to harbour this mutation. Jaimeson et. al., 2006 recently used flow cytometry to isolate different progenitor cell populations from PV patients and showed the mutation was observed in all cell fractions providing evidence that PV arises in haematopoietic stem cells (HSCs) and involves myeloid, erythroid and megakaryocytic lineages (Jamieson, *et al* 2006) and thus, like CML, V617F positive MPDs are believed to arise from an early haemopoietic progenitor or stem cell.

A further interesting outcome has been the discovery that homozygous clones occur in most patients with PV but are never found in patients with ET. In addition to this observation is that ET *JAK2* V617F positive patients have phenotypic similarities of PV for instance, increased haemoglobin, increased numbers of neutrophils, bone marrow erythropoiesis and granulopoiesis and a propensity to venous thromboses. ET patients lacking the *JAK2* V617F mutation also differ in that they have more pronounced and isolated thrombocytosis, less active erythropoiesis and granulopoiesis. From these in-depth studies it has been suggested that V617F positive thrombocythaemia and polycythaemia could be viewed as a continuum of disease and not as two separate entities, with depleted iron stores, genetic modifiers and erythropoietin homeostasis tipping the disease in favour of ET and sex (mainly biased toward men), genetic modifiers and homozygosity for V617F pushing the disease balance toward PV (Campbell, *et al* 2005).

Whilst these questions still remain unresolved, work has at least progressed toward possible therapeutic strategies that will hopefully become clinically effective. Targeted potent inhibitors of JAK2 are currently in development for the treatment of MPD (Geron, *et al* 2008, Hexner, *et al* 2008, Pardanani 2008). Most of the compounds that are undergoing preclinical and clinical testing inhibit both wild-type JAK2 and JAK2 V617F. JAK2 signalling is important in many cellular functions hence inhibition of wild-type may result in unwanted haematopoietic side effects. Additionally JAK2 is one of a closely related family and wild-type JAK3 inhibition has to be considered due to its association between loss of function mutations and

severe combined immunodeficiency (Macchi, *et al* 1995, Russell, *et al* 1995). Nevertheless many compounds are now in clinical trials and early results indicate objective improvements in many patients but, disappointingly, only minimal reductions in the size of the neoplastic clone.

1.8 Summary

The *BCR-ABL* paradigm in CML has proved a useful model for the study of other MPDs and aberrant tyrosine kinase signalling has been identified as a common factor driving myeloproliferation. Chromosomal translocations have been instrumental in the identification of numerous deregulated kinases, however it is apparent from the discovery of the V617F *JAK2* mutation that LOH/UPD analysis is also a valuable tool for identifying chromosomal regions containing novel oncogenes.

The efficacy of targeted therapy of deregulated tyrosine kinases makes it critically important to identify patients harbouring these molecular lesions. This is especially true for mutations involving *PDGFRA* or *PDGFRB* as these are imatinib sensitive.

1.9 Statement of aims

The general aim of this study is to further our knowledge about the molecular pathogenesis of the MPDs and to pursue the hypothesis that these diseases are caused primarily by deregulation of tyrosine kinases. My specific aims are:

i) To characterise acquired chromosomal translocations seen in association with MPDs

ii) To develop screening procedures capable of detecting overexpression of specific tyrosine kinase genes and to investigate the molecular basis for any observed overexpression

iii) To investigate MPDs for regions of acquired uniparental disomy as an indicator towards possible oncogenes or tumour suppressor genes.

2 METHODS

2.1 Sample processing

2.1.1 Preparation of blood and bone marrow samples

Patient samples were either (i) referred from Consultant Haematologists throughout the UK for routine diagnostic analysis (usually the detection of specific acquired genetic abnormalities or determination of *in vitro* sensitivity to imatinib), or (ii) provided by collaborators from abroad. For research purposes, UK samples were anonymised following diagnostic analysis and reporting of results. Samples from abroad were provided in coded form. Use of these samples for research was approved by the Salisbury and South Wiltshire LREC (LREC reference 05Q2008/6; mechanisms and consequences of tyrosine kinase activation in chronic myeloproliferative disorders and related conditions; chief investigator Professor NCP Cross).

Patient material usually ranged from 5-50ml of peripheral blood, or 1-10ml bone marrow aspirate. In some cases fixed cell suspensions for cytogenetics were also provided. Mononuclear cells (MNCs) and granulocytes were separated by centrifugation over Lymphoprep^{TM} (Axis Shield, Oslo, Norway). The cell counts of the peripheral blood and bone marrow samples were determined using a Neubauer type haemocytometer (BDH, Poole, UK). $2x10^7$ cells were diluted in Hanks' balanced salt solution (HBSS) to a maximum volume of 25ml and this suspension was then layered over an equal volume of lymphoprep in a 50ml propylene tube (Sarstedt, UK). Centrifugation at 1,800rpm for 30 minutes without a brake created an interface comprised of MNCs. The MNCs were transferred to a new tube using a sterile pasteur pipette (BDH, Poole, UK). The recovered cells were then washed twice in RPMI 1640 supplemented with 10% foetal calf serum (FCS), L-glutamine and penicillin-streptomycin before counting. To obtain the granulocytes, the remainder of the lymphoprep mixture from the original 50ml tube was discarded,

leaving the red cell pellet. To the pellet, ice-cold red cell lysis buffer (RCLB) was added to lyse the red cells. The tubes were then placed on ice for 10 minutes (intermittent mixing), and then spun for 5 minutes at 1600rpm. Before counting, the white cell pellet obtained was washed twice in supplemented RPMI (as before). For RNA extraction $1x10^7$ cells were lysed in either 1ml guanidinium thiocyanate (GTC; the composition of all solutions are given in section 2.5) freshly mixed with β-mercaptoethanol at a ratio of 7µl of β-mercaptoethanol (Sigma) per 1ml GTC lysate or $1.5x10^7$ cells were lysed in 1.5ml TRIZOL® reagent (Invitrogen, UK) and stored at -70°C. For DNA extraction, cell pellets of $1\text{-}2.5x10^7$ cells were frozen at -70°C. Excess MNCs were cryopreserved in aliquots of between 2.5 and $6x10^7$ cells. Each cell aliquot was spun down and resuspended in 500µl of supplemented RPMI (10% FCS, L-glutamine and penicillin-streptomycin). An equal volume of freezing mix (composition detailed in section 2.5) is added dropwise to the resuspended cells. Finally, 1ml of this mixture was again added dropwise to a cryopreservation tube and frozen overnight at -70°C before being transferred to the liquid nitrogen stores.

2.1.2 Enriching cultures for T-cells

To investigate if a mutation has been acquired, DNA prepared from a granulocytic fraction can be compared to DNA from a T-cell fraction. T-cells are lymphoid in origin and are not thought to be involved in myeloproliferative disease. Phytohaemagglutinin (PHA) is a lectin extracted from white kidney beans and is used to stimulate mitotic division of lymphocytes maintained in cell culture.

Peripheral blood was first separated by centrifugation over Lymphoprep™ (Axis Shield, Oslo, Norway) into mononuclear cells (MNCs) and granulocytes as detailed in section 2.1.1. The cell count of the mononuclear cells was determined using a Neubauer type haemocytometer (BDH, Poole, UK). Two cultures of $5x10^6$ cells were then resuspended in 5ml of RPMI 1640 supplemented with 10% foetal calf serum (FCS), L-glutamine and penicillin-streptomycin in a Nunclon® polystyrene flat-sided cell culture tube (Scientific Laboratory Supplies Ltd, Nottingham). To one culture

500µl of myeloid cytokine mix was added (GM-mix, for components see 2.5) and to the other 50µl of PHA. After 3 days incubation at 37 ºC the cell suspension was spun at 1500rpm for 5 minutes. The cell pellet was then resuspended in 2ml of supplemented RPMI 1640 (as above) and the cells counted (as previously). Two aliquots of 4 x10^4 cells were used to produce two post-PHA statspins. The remaining cells were spun at 1500rpm for 5 minutes and DNA made from the cell pellet according to the protocol in section 2.1.5.2. The statspins were taken to our Pathology department where both the pre- and post-PHA slides were stained using May-Grunwald Giemsa stain in an automated machine. (This used to involve fixing the cells in methanol for 15 minutes, staining in May-Grunwald for 5 minutes, then staining in Giemsa for 10 minutes. The slide was then rinsed first in buffer pH 6.8, followed by a 50/50 buffer/acetone rinse, dehydration in acetone twice and cleared using three xylene washes). This stain enables different types of cell (erythrocytes, lymphocytes, etc.) and components within the cell, to be distinguishable from one another by their colour and shape when viewed under a microscope. Lymphocytes look like circular cells with large round blue nuclei that take up ~60% of the cytoplasm with no vessicles. Pre-and post-PHA statspins were used to assess the level of T cell enrichment. DNA made from post-PHA incubation was used as a control for the investigation of acquired mutations in the myeloid cell fraction.

2.1.3 RNA extraction

RNA was extracted from GTC samples using the Qiagen RNeasy kit (Qiagen, West Sussex). 350µl of GTC was mixed with an equal volume of 70% ethanol, mixed well, and applied to the Qiagen column then left to stand for 30 minutes. The column was spun for 10 seconds at 10000rpm in a Genofuge 16M centrifuge. The column was then washed with the buffers provided in the kit: 650µl of RW1 for 10 seconds at 10000rpm, and then twice with 500µl RPE, the first time for 10 seconds at 10000rpm and the second wash at 14000rpm for 2 minutes. Columns were then left in a class II tissue culture cabinet with the lids open for 10 minutes to allow evaporation of any

remaining ethanol. The bound RNA was then eluted from the column with 25µl distilled water. RNA was usually directly reverse transcribed into cDNA.

RNA was extracted from samples in TRIZOL® using chloroform followed by the Qiagen RNeasy kit. Briefly, to 350µl TRIZOL® sample 350µl of chloroform (Sigma Aldrich) was added and vortexed for 15 seconds. The sample was then spun at 14,000 rpm for 5 minutes at room temperature. Centrifugation separates the solution into an aqueous phase and an organic phase, the RNA remains in the aqueous phase. The aqueous phase was transferred to a clean eppendorf and washed again with 350µl chloroform as before. The final aqueous phase was carefully transferred to a clean eppendorf containing 350 µl of 70% ethanol, mixed well and then applied to an RNeasy column. The protocol continues as specified above once on the column.

2.1.4 cDNA synthesis

After elution from the column the RNA was placed on ice and then heated to 65°C for 5 minutes. The cDNA synthesis was performed at 37°C for 2 hours utilising 1-5µg RNA in 25µl (GTC samples) or 50µl (TRIZOL® samples) water. The RNA was combined with 21µl (GTC samples) or 42µl (TRIZOL® samples) of cDNA mix, which resulted in final concentrations of 50mM Tris pH8.3, 75mM KCl, 3mM $MgCl_2$, 1mM DTT, 1mM dATP, dCTP, dTTP, dGTP, as well as 100µg/µl of random $pd(N)_6$ hexamers (Amersham Pharmacia, Amersham, UK), 150 units of Moloney Murine Leukaemia Virus (MMLV) reverse transcriptase (Invitrogen, UK) and 15 units/ml of RNasin (Promega, UK). Synthesis was terminated by heating to 65°C for 10 minutes. cDNA was stored at -20°C.

2.1.5 DNA Extraction

2.1.5.1 DNA Extraction from fixed cell suspension

Fixed cell suspensions were centrifuged at 4000 rpm for 10 minutes in a Thermo IEC, Micromax centrifuge. The cell pellet was then resuspended in 1ml of sterile PBS before spinning the pellet as before and then the washing was repeated three times. After the third centrifugation step the pellet was resuspended in 200µl of sterile PBS (independent of the number of cells) and the Qiagen DNeasy Blood & Tissue kit protocol: 'Purification of total DNA from animal blood or cells' was followed starting from step 1c. To the resuspended pellet in 200µl of PBS, 20µl proteinase K and 200µl of Buffer AL (without ethanol added) was added. This was mixed thoroughly by vortexing before incubation at 56ºC for 10 minutes. 200µl ethanol (96-100%) was then added and the sample mixed thoroughly by vortexing, this mixture was then added to the DNeasy Mini spin column and centrifuged at 8000rpm for 1 minute. The collection tube containing the flow through was then discarded and the spin column containing the bound DNA placed in a new collection tube. 500µl of Buffer AW1 was added to the spin column and centrifuged at 8000 rpm for 1 minute. Again the collection tube was discarded and the spin column placed in a new collection tube, before the addition of 500µl Buffer AW2 and centrifugation at 14,000 rpm for 20 minutes to dry the DNeasy membrane. The collection tube was discarded again and the spin column placed in a sterile 1.5ml microcentrifuge tube. 50µl Buffer AE was added directly to the DNeasy membrane and incubated at room temperature for 1 minute before centrifugation for 1 minute at 8000rpm. To increase the DNA yield the elution step above was repeated.

2.1.5.2 Extraction from cell pellets

Cell pellets from peripheral blood and bone marrow samples were stored at -70ºC. Depending on the cell count the volume below was either doubled (>$2x10^7$ cells) or halved (<$5x10^6$ cells). A $1x10^7$ pellet was thawed quickly by hand before the addition of 500µl resuspension buffer (RSB – see section 2.5 for components), 10µl 10% SDS and 3µl proteinase K (Qiagen, West Sussex). The cell pellet was then incubated at 37ºC overnight. On the second day, the pellet was first checked for proper digestion, if not digested properly a further 2µl of proteinase K was added and left for as long as

necessary. When properly digested, 150µl of 6M NaCl was added and the sample shaken vigorously by hand for 20 seconds. The digested cell pellet was then spun at room temperature in a Thermo IEC, Micromax centrifuge at 14,000rpm for 20 minutes. The supernatant was then removed to a clean sterile bijoux tube and 1.5 times the volume of the supernatant was added of 100% ethanol. The bijoux was inverted to obtain a 'hairball' of DNA, and hooked out using a sterile needle into a sterile eppendorf containing 70% ethanol. The size of the DNA 'hairball' was assessed and an appropriate volume (50-1000µl) of sterile Tris-EDTA buffer (Promega, UK) was aliquoted into a sterile 2ml screw cap tube, to this the washed 'hairball' of DNA was added and left to resuspend properly in the TE buffer before use.

2.1.5.3 DNA Extraction from Guanidinium Thiocyanate

This protocol uses the Qiagen DNA Mini Kit 'Blood and Body Fluid spin protocol' (Qiagen, West Sussex). To an equal volume of GTC, 200µl of Buffer AL was added and mixed by pulse-vortexing for 15 seconds. The sample was then incubated at 56ºC for 10 minutes and briefly centrifuged in a Thermo IEC, Micromax centrifuge to remove the drops from the inside lid. The mixture was then carefully applied to the QIAamp spin column (in a 2ml collection tube) without wetting the rim and the lid closed before centrifugation at 8000rpm for 1 minute. The collection tube containing the filtrate was discarded and the spin column placed in a new collection tube. 500µl of Buffer AW1 was added to the spin column and spun at 8000rpm for 1 minute, the filtrate discarded and 500µl of Buffer AW2 added to the column and spun at 14,000rpm for 3 minutes. To remove excess ethanol the filtrate was discarded and the column spun for a further minute at 14,000rpm. The spin column was then placed in a sterile microcentrifuge tube and 200µl Buffer AE added, incubated for 1 minute at room temperature and then spun at 8000rpm for 1 minute.

2.1.5.4 DNA Extraction from Mouthbrush samples

DNA obtained from mouthbrush samples was used as a control for the investigation of acquired mutations. In general, two mouthbrush specimens were received for each patient. To one mouthbrush tube 200µl of resuspension buffer (see section 2.5) was added. The lid of the other mouthbrush is snapped off and the mouthbrush added to the tube containing the first mouthbrush and 200µl of RSB. The tube was then vortexed for 30 seconds before using a sterile Pasteur pipette to remove the 200µl to a sterile 2ml tube. To this, 10µl of 10% SDS and 3µl of Proteinase K (Qiagen, West Sussex) were added, before incubation at 37°C overnight. The following day 150µl of 6M NaCl was added to the sample, vortexed to mix and then spun at 4°C at 14,000rpm for 20 minutes. The supernatant was then collected into a new sterile bijoux and 1.5 times the volume of 100% ethanol (Sigma Aldrich, UK) added. A hairball of DNA is rarely visible and so the sample was split between two sterile eppendorfs and frozen at -20°C for >2 hours. The sample was then spun again at 4°C at 14,000rpm for 20 minutes and each pellet resuspended in 25µl of 1 x Tris-EDTA buffer. These were left to resuspend for 1 hour at 37°C before combining in a labelled sterile 2ml screw cap tube.

2.1.5.5 DNA extraction from haematological stained slides

DNA extracted from haematological stained slides was used as a control for the investigation of acquired mutations and to investigate when the patient acquired the mutation. The dried material on the slide was first moistened with a drop of sterile PBS (Sigma Aldrich, UK). To a sterile eppendorf 200µl of sterile PBS was added. The cytological material was then scraped into this tube from the slide using a sterile scalpel blade. The resulting ‘sludge’ was mixed by pipetting up and down before adding 20µl of proteinase K (Qiagen, West Sussex) and 200µl of buffer AL and further mixing by pulse-vortexing for 15 seconds. The sample was then left overnight at 56°C to digest before continuing the following day with the Qiagen Blood and

Body Fluid Spin Protocol from step 5. As there is only a small amount of DNA to be obtained, a carrier DNA was always used. Qiagen recommend poly dA:dT carrier as this works best with the Qiagen columns, 10µg of polydA:dT (Sigma, Poole) was added after the overnight digestion and then continued as stated above.

2.1.5.6 GenomiPhi V2 DNA Amplification Kit

The Illustra™ GenomiPhi™ V2 DNA Amplification Kit is part of the Phi29 DNA polymerase family of products (GE Healthcare, UK) for mini-scale whole genome DNA amplification. Our samples are often precious and whole genome amplification offers the possibility to greatly expand the amount of usable material. A typical DNA yield of ~47 µg DNA can be achieved in less than two hours from an input of 1-10 ng genomic DNA with an average product length of >10 kb. The template must be at least 1ng/µl and requires an initial denaturation. On ice in thin-walled PCR tubes, 1µl (typically 10 ng) of template genomic DNA was combined with 9µl of sample buffer. The sample was then denatured at 95ºC for 3 minutes before cooling to 4ºC. Whilst the sample was denaturing a master mix combining 9µl of reaction buffer with 1µl of enzyme per sample was made up on ice. This was then added to the cooled denatured sample on ice prior to incubation at 30ºC for 90 minutes. The enzyme was then heat inactivated at 65ºC for 10 minutes before cooling to 4ºC. To test the quality of the genomic DNA, a control gene PCR was set up (See section 2.2.3).

2.2 Amplification and analysis

2.2.1 Reverse Transcriptase PCR (RT-PCR)

RT-PCRs were undertaken in a volume of 10-25µl. All amplifications were performed for between 30-32 cycles in a MJ Research Inc. thermocycler (single platform or Tetrad) in thin-walled PCR tubes. PCR mixes contained 0.5µM primers, 0.2µM each of dCTP, dATP, dGTP and dTTP, 1x Buffer (10mM Tris-Cl pH8.3, 50mM KCl, 0.1% gelatin), 1.75mM $MgCl_2$ and 0.05U Taq DNA polymerase

(Invitrogen). Typically multiplex PCRs were performed with 0.2μM each primer and $MgCl_2$ at a concentration of 2.5mM.

2.2.2 Microsatellite PCR

DNA samples were amplified with a series of fluorescently labelled primer pairs flanking highly polymorphic microsatellite markers on specific chromosomes of interest. All amplifications were performed in a MJ Research Inc. thermocycler (Tetrad) in thin walled 0.2ml 96-well plate format. The PCRs were undertaken in 10μl volumes with 25-50ng of genomic DNA. All primers were designed to anneal at 55°C and standard cycling parameters were used with 32 amplification cycles. The products were analysed on an ABI 3100 genetic analyser using the program Genotyper 2.0. The peak heights were compared to controls and scored as heterozygous for each marker if two peaks of similar intensity were clearly visible, or if the ratio of the two peaks was similar to normal controls. If only one peak was visible a patient was scored as homozygous. To allow for the presence of background normal cells, acquired homozygosity was also scored if one peak was one third or smaller than the expected size compared to normal controls.

2.2.3 Control gene PCR

The control gene PCR is a multiplex PCR capable of amplifying 5 different sized products (100bp, 200bp, 300bp, 400bp, 500bp & 600bp). This PCR is used to check the quality of genomic DNA, especially after GenomiPhi amplification. PCRs were undertaken in a volume of 50μl. All amplifications were performed for 35 cycles in a MJ Research Inc. thermocycler (Single platform or Tetrad) in thin-walled PCR tubes. The primers were combined to make two mixes: Mix A (25μl of 100bp.F, 100bp.R, 200bp.F, 200bp.R, 300bp.F, 300bp.R, 400bp.F & 400bp.R) and Mix B (50μl 600bp.F & 600bp.R) (van Dongen*, et al* 2003). PCR mixes contained 0.5μM Mix A primers, 1μM Mix B primers, 0.5μM of dNTPs, 1x Buffer, 2mM $MgCl_2$ and 1U Ampli*Taq*

Gold DNA polymerase (Applied Biosystems, UK). The PCR was cycled as follows: 94°C for 7 minutes hotstart; 94°C for 1 minute; 60°C for 1 minute; 72°C for 1 minute for 35 cycles and an extension of 72°C for 7 minutes. The products were run on a 2% agarose gel to allow separation of the bands.

2.2.4 Powerplex® 16 System

The Powerplex® 16 kit (Promega, UK) allows the co-amplification and three colour detection of sixteen loci; fifteen short tandem repeat loci plus Amelogenin for sex determination. The system encompasses the following loci: Penta E & D, D18S51, D21S11, THO1, D3S1358, FGA, TPOX, D8S1179, vWA, CSF1PO, D16S539, D7S820, D13S317, D5S818 and Amelogenin as described in detail at http://docs.appliedbiosystems.com/search.taf. All sixteen loci are co-amplified in a single tube and analysed in a single injection on an Applied Biosystems 3100. This kit was designed to allow a high discriminatory power for cases of identity and paternity. I used this kit as a method of confirming that DNA samples at different time points were indeed from the same individual.

The Powerplex® 16 PCR was performed using the following mix per sample; 1.25µl Gold ST*R 10x Buffer, 1.25µl 10x Primer pair mix, 0.4µl Amplitaq Gold® DNA polymerase (Applied Biosystems, UK), 8.3µl sterile water. To 11.2µl of master mix 1.2µl of DNA at 0.9ng/µl was added. PCR was performed according to the following amplification protocol: 95°C for 11 minutes, 96°C for 1 minute, 94°C for 30 seconds, ramp 68 seconds to 60°C (hold for 30 seconds, ramp 50 seconds to 70°C (hold for 45 seconds), cycle the last two steps 10 times, 90°C for 30 seconds, ramp 60 seconds to 60°C (hold for 30 seconds), ramp 50 seconds to 70 °C (hold for 45 seconds), and cycle the last two steps 20 times, final extension at 60°C for 30 minutes and a 4°C soak.

After PCR the products were prepared for the ABI 3100. An additional lane was incorporated into the plate to include the Powerplex® allelic ladder which contains

the most common alleles for each locus. A master mix was made up of 9µl highly deionised formamide and 0.5µl Promega ILS600 size standard. The master mix was aliquoted into a 96 well plate (specific to the ABI 3100 machine) and 0.5 µl of each sample was then added. The samples were run under specific ABI 3100 settings for 45 minutes. Genotypes were then assigned by comparing the sizes obtained for the unknown samples with the sizes obtained for the alleles in the allelic ladder. Different samples can then be compared against each other based on allele specificity at each locus.

2.2.5 Long-Range PCR

Amplification of products larger than 2.5kb was performed using long-range PCR. The High Fidelity PCR Master kit (Roche Diagnostics, Lewes, UK) was used to amplify genomic DNA up to 10kb in length. All primers were designed with a melting temperature of 66°C with approximately 50% GC content. Typically 100ng of DNA was amplified in the PCR master kit with primers at a final concentration of 0.5µM. PCR was performed according to the following amplification protocol, allowing increased elongation times for the length of product: 94°C for 20 seconds, 66°C for 4 minutes, 68°C for 12 minutes, cycled 10 times, then 94°C for 20 seconds, 66°C for 40 seconds and 68°C for 14 minutes, cycled 22 times, final extension at 72°C for 10 minutes.

2.2.6 Bubble-PCR

Bubble PCR (Zhang, *et al* 1995) is a one sided PCR technique capable of amplifying an unknown sequence that lies adjacent to a known sequence. The bubble oligonucleotide is composed of two strands that are complementary at both ends but non-complementary in the middle. For all primer sequences see appendix I, section I.

Sample DNA was first digested using two separate restriction enzymes (*HaeIII* and *RsaI*) designed to blunt end cut around the known sequence and regularly throughout the rest of the genome. The bubble oligonucleotide was then blunt end ligated to the digested genomic DNA. The ligated DNA was then denatured at 72°C for 5 minutes before removal of unligated bubble oligonucleotides using the QiaQuick PCR purification kit (Qiagen, West Sussex). A sample specific primer, in this case PDAI12-R3, was annealed to 1µl of the ligated DNA and extended using Amplitaq Gold (Applied Biosystems, UK) to the end of the bubble oligonucleotide (if there was no bubble sequence the rest of the genome to which bubbles had been ligated would also be amplified with the bubble primer). A single step Hotstart PCR using the same PDAI12-R3 primer and a forward NVAMP-1 primer, complementary to the bubble oligonucleotide, was then performed to amplify this strand. This results in a copy of the known, unknown (if present) and bubble sequence. The copied, non-complementary part of the bubble sequence is only present if the specific DNA was amplified. This strand was then subsequently amplified using Hotstart PCR with a bubble specific primer. A nested PCR was then performed to increase the specificity of the reaction using NVAMP-2 and PDAI12-R4. Rearranged bands were excised and gel purified with QiaQuick Gel Purification kit (Qiagen) and either sequenced directly or cloned and sequenced.

2.2.7 Gel Electrophoresis

The protocol was altered accordingly dependent on the size of the gel former and the percentage gel required. For a 1% agarose gel in a gel tank with a capacity of 100ml, 1g agarose (Bioline, UK) was added to 100ml 1x Tris-Borate EDTA (TBE) buffer (Sigma Aldrich, UK). This was then heated using a microwave to melt the agarose, before cooling slightly and adding 0.2mg/ml ethidium bromide to visualise the DNA. The melted gel was then poured into a gel former with combs and allowed to set for at least 20 minutes. Ensuring the gel was set the combs and formers were then removed and the gel and electrodes covered by 1 x TBE buffer. To 5µl of sample or PCR product 1-2µl of Orange-G sucrose loading buffer was mixed before loading the

whole amount into a well of the gel. One lane was always loaded with the 1kb plus DNA ladder (Invitrogen, UK) to enable sizing of products. The samples were run for 15 minutes or longer depending on the expected product size and percentage gel before visualising the bands under UV light.

2.2.8 Denaturing high-performance liquid chromatography (dHPLC)

Amplification of each exon of each gene to be analysed was performed from approximately 50-100ng of patient genomic DNA. A high fidelity DNA Taq polymerase was used to ensure that PCR did not introduce any sequence errors. PCRs were undertaken in a volume of 75µl. All amplifications were performed for 35 cycles in a MJ Research Inc. thermocycler (Tetrad) in thin-walled 12-strip PCR tubes. PCR mixes contained 0.5µM primers, 0.2µM each of dCTP, dATP, dGTP and dTTP, 1x GeneAmp PCR Buffer II (100mM Tris-HCl, pH8.3, 500mM KCl), 1.5mM $MgCl_2$ and 0.05U AmpliTaq Gold DNA polymerase (Applied Biosystems, Warrington, UK). The primers were designed at least 50 base pairs away from each end of the exon, with a maximum size of 500 base pairs for each amplicon. In some cases exons very close together were co-amplified with the intervening intron present. The primer sequences and the PCR conditions are in the appendix I, sections II & III.

Following amplification, the PCR product from each patient for each exon was mixed equally with amplified product from the same exon from human placental DNA (Sigma-Aldrich, Poole, UK). The mixed products were then denatured at 95°C for 5 minutes and then slowly reannealed to each other by reducing the temperature by 1.5°C every 22 seconds to room temperature allowing stable formation of heteroduplexes. Conditions for dHPLC analysis were calculated using the manufacturer's software WAVEmaker v4.1. The computer predicts the temperature(s) over which the DNA fragment will be between 50 and 100% reannealed. It is estimated that more than 95% of heteroduplexes are detected. For longer products or those that have a high GC content, there may be several temperatures required to profile the whole fragment. The temperatures required for

heteroduplex analysis of each amplicon are given in the appendix, section III. Before each dHPLC run a plasmid DNA control supplied by the manufacturer (WAVE mutation standard; Transgenomics, Crewe, UK), with a known melting profile was tested to ensure consistency of results between runs. For each exon normal placental DNA PCR products were run as a normal control, again to ensure consistency of results for each exon between experiments. All results were analysed using the WAVEmaker v4.1 software. Any heteroduplexes observed, indicated by shifts in the traces or split peaks, were subsequently re-amplified and sequenced to look for the presence of mutations.

2.2.9 Multiplex Ligation-dependent Probe Amplification (MLPA)

MLPA is a method to establish the copy number of up to 45 nucleic acid sequences in one single reaction (Schouten, *et al* 2002). It uses genomic DNA to detect amongst other things copy number. Deletions and amplifications of a gene (or part of) will usually not be detected by sequence analysis of PCR amplified gene fragments as a normal copy is still present. Amplification products are separated by sequence type electrophoresis. As only one pair of PCR primers is used, MLPA reactions result in a very reproducible gel pattern with fragments ranging from 130 to 490 bp. Comparison of this gel pattern to that obtained with a control sample indicates which sequences show an aberrant copy number.

1µl (50ng) of DNA was added to 4µl of Tris-EDTA buffer in a 0.2ml PCR tube and overlaid with oil (this prevents evaporation overnight) before running on a MJ thermocycler tetrad 'Denature' programme (98°C 5 minutes; 25°C Hold). Whilst the DNA is denaturing the following mix was made up. For each sample, to 1.5µl relevant probe mix, 1.5µl MLPA buffer was added and mixed gently. 3µl of this mix was added to the denatured sample and the probes bound overnight at 60°C (95°C 1 minute; 60°C Hold). Following the overnight binding of the probes, on ice 3µl of ligase buffer A and B, 25µl water and 1 unit ligase were combined for each sample and mixed. The 60°C overnight run was stopped and the samples held at 54°C for at

least 2 minutes before adding 32µl of the above mix to each sample and running the 'Ligate' programme (54°C 15 minutes; 98°C 5 minutes; 4°C Hold). Whilst the probes were ligating, the PCR mix was made up. For each sample 2µl SALSA buffer (MRC Holland, Amsterdam), 1µl SALSA primers, 1µl Enzyme buffer, 15.75µl water and 0.25µl (2.5 units) SALSA polymerase were mixed carefully and then aliquoted into new 0.2ml PCR tubes on ice. To this PCR mix 5µl of the ligation mix was added and run on a 'Cycles' programme on a tetrad (95°C 30 seconds; 60°C 30 seconds; 72°C 1 minute; cycle 33 times; 72°C 20 minutes; 4°C hold). After PCR the samples were analysed on an ABI 3100 by mixing 1µl PCR product to 9µl of formamide and 0.1µl of ROX500 size standard (Applied Biosystems, UK).

Analysis of the samples was performed using Applied Biosystems Genescan™ software using a GS500 size standard. Following this Applied Biosystems Genotyper™ software was used. Based on the number of probes in the MLPA kit the software can be manipulated. The number of peaks was specified in addition to selecting 'highest' peaks, ensuring only the correct number of highest peaks, within the size range, are selected. All peaks were labelled with the peak height before exporting the data into a table. MLPA analysis sheets for each MLPA probe kit have been set up in Microsoft excel. The data from the genotyper table was imported into the excel sheet and samples were compared to controls. The peak height of each probe is compared against every other probe and against the control. This gives a data sheet of numbers where 1 is normal, a horizontal row of 0.5 indicates a deletion and horizontal row of 1.5 a duplication.

2.2.10 Exo-SAP and sequencing

Exonuclease I and shrimp alkaline phosphatase (SAP), two hydrolytic enzymes, are combined at a ratio of 1:4 and used for the preparation of PCR products for fluorescent-based sequencing. Exonuclease I (New England Biolabs, UK) degrades any residual single stranded primers and extraneous single-stranded DNA, whilst SAP (Roche Diagnostics Ltd, UK) hydrolyses the remaining dNTPs from the PCR

reaction. On ice, 1µl of exonuclease I/SAP mix was added to 2µl of PCR product and incubated at 37°C for 15 minutes, followed by 65°C for 15 minutes to denature the enzymes. The product from this reaction was diluted 10 fold and then used directly for sequencing.

2.2.10.1 Sequencing of PCR products

Exonuclease I/SAP purified PCR products were sequenced using a final concentration of 1µM of a single primer (either forward or reverse) and the BigDye terminator kit v3.1 or v1.1 (Applied Biosystems, Foster City CA, USA). Dependent on numbers of samples to be sequenced, reactions were either cleaned up using ABgene® Dye Terminator Removal Kit (Abgene Limited, UK) and dried using a MJ Research Inc. thermocycler (Tetrad), or by using the Montage SEQ96 Cleanup Kit (Millipore, Watford, UK). The first is a gel-based separation matrix and the latter a membrane-based size-exclusion filtration system and requires final resuspension in 15µl of washing buffer. DNA from both methods was resuspended in 10µl deionised formamide and run on an ABI 3100 or ABI3130xl sequencer.

2.2.10.2 Sequencing following cloning

Plasmid DNA was sequenced using 1µl of mini-prepped DNA (approximately 100-500ng) for each PCR (See section 2.2.11.3). Sequencing was performed using the forward and reverse M13 vector primers for each plasmid DNA (two reactions). The single stranded inserts were amplified with BigDye terminator kit v3.1 or v1.1 (Applied Biosystems, Foster City CA, USA), cleaned up using the Montage SEQ96 Cleanup Kit (Millipore, Watford, UK) and resuspended in 15µl wash buffer. To the resuspended DNA 10µl deionised formamide was added and the reactions run on an ABI 3100 or ABI3130xl sequencer.

2.2.11 Cloning PCR products

PCR products were cloned by ligation into a plasmid vector using the TA or TOP10 cloning kits according to the manufacturer's instructions (Invitrogen, Paisley, UK).

2.2.11.1 Ligations

On ice, 1-2µl PCR product (depending on the intensity of the product observed on an agarose gel), 1µl salt solution (1.2M NaCl, 0.06M $MgCl_2$), 1µl pCR®4 vector (10ng/µl), and sterile water to a final volume of 5µl were combined. The reaction was then mixed gently, and incubated at room temperature for 5 minutes. This reaction relies on *Taq* polymerase nontemplate-dependent terminal transferase activity which adds a single deoxyadenosine to the 3'end of PCR products. The vector supplied in the kit is linearised and has complementary single, overhanging 3' deoxythymidine residues allowing PCR inserts to ligate efficiently. The ligation itself relies on the energy conserved by formation of a covalent bond between the 3'phosphate of the cleaved strand and the tyrosyl residue of the enzyme topoisomerase I which binds to and cleaves the phosphodiester backbone in certain regions of DNA.

2.2.11.2 Transformations

Chemically competent TOP10 cells provided with the Invitrogen kit were slowly thawed on ice. 2µl of the ligation reaction was added and stirred gently with a pipette tip to mix. The cells were then incubated on ice for 30 minutes. The remaining ligations were stored at -70°C. After 20 - 30 minutes, the cells were heat-shocked at 42°C for 30 seconds and immediately placed on ice for at least 2 minutes. 250µl of SOC medium (2% Tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 20mM glucose) was then added to each vial, and the cells were then shaken horizontally at 37°C for 1 hour. LB/agar plates with the appropriate antibiotic added were spread with 25µl of 50mg/ml X-Gal (Promega, Madison, USA)

and allowed to dry. Varying amounts of each transformation mix were spread on individual plates. Once the liquid was absorbed the plates were inverted and placed in a 37°C incubator overnight. After incubation the plates were placed at 4°C to allow for colour development. White colonies were picked for rapid screening of inserts by PCR using the following miniprep protocol.

2.2.11.3 Qiagen™ Miniprep kit

Picked colonies were grown up overnight in single 10ml culture tubes with 2ml L-Broth (LB) and the appropriate antibiotic. The following day cells from 1.5ml culture medium were extracted using the Qiagen™Miniprep kit (Qiagen Ltd, West Sussex) and the remaining 500µl is stored in LB/Glycerol (3:1) solution at -70°C. The 1.5ml of culture medium was spun down in an eppendorf at 10,000rpm for 10 minutes in a Thermo IEC, microcentrifuge. The supernatant was then carefully removed leaving only the pellet which was resuspended in buffer P1, ensuring no clumps remained before adding the P2 lysis buffer. The tubes were then inverted gently 4-6 times to ensure complete lysis before addition of the neutralisation buffer P3. On addition of buffer P3 each tube was immediately inverted 4-6 times; the solution became cloudy and white as the denatured proteins, cellular debris and chromosomal DNA precipitate out of solution. The tubes were then spun at 14,000rpm for 10 minutes, in a Thermo IEC microcentrifuge, and the supernatant carefully transferred to an appropriately labelled QIAprep spin column. The column was then spun at 14,000rpm for 1 minute allowing the DNA to bind to the column, and then washed twice for 1 minute 14,000rpm using firstly 500µl wash PB, followed by 750µl wash PE. The column was then spun for an additional 5 minutes at 14,000rpm to remove all residual PE buffer. To collect the DNA, 25µl of distilled water was added to the centre of the column, now in a new sterile labelled eppendorf, and spun for a further 1 minute at 14,000rpm.

2.2.11.4 Digests and PCR

5µl of each mini-prepped plasmid DNA was digested with 10 units Eco*RI*, (Eco*RI* sites flank the cloning site in pCR2.1 or pCR 4.0 vector), 2µl 5x Eco*RI* buffer and the appropriate amount of sterile water up to 10µl. Digests were performed at 37°C for 2-4 hours. The digested clones were run on a 1% agarose gel to determine the insert size and cloning efficiency. The plasmid DNA was then sequenced according to the protocol in section 2.2.10.2.

2.2.12 5' Rapid amplification of cDNA ends (5'RACE)

Fusion genes were identified using the 5'RACE GeneRacer™Kit (Invitrogen, Paisley, UK). In brief, after RNA extraction using the Qiagen RNeasy kit (Qiagen Ltd, West Sussex, UK), approximately 5µg of total RNA was dephosphorylated, decapped and ligated to the GeneRacer™ RNA oligonucleotide (Figure 2.1). The ligated RNA was reverse transcribed using Superscript II™ reverse transcriptase (Invitrogen, Paisley, UK) and random primers. The initial 5'RACE PCR was performed using the 5'GeneRacer™ primer and a reverse primer from the relevant exon of the tyrosine kinase gene in question, using the High Fidelity PCR Master kit (Roche Diagnostics, Lewes, UK). The second step, nested PCR was performed using the 5'GeneRacer™ nested primer and a nested primer designed in the tyrosine kinase gene. After nested PCR, if a product was visible by gel electrophoresis, the product was cloned using the TOPO TA cloning kit for Sequencing (Invitrogen, Paisley, UK), and sequenced.

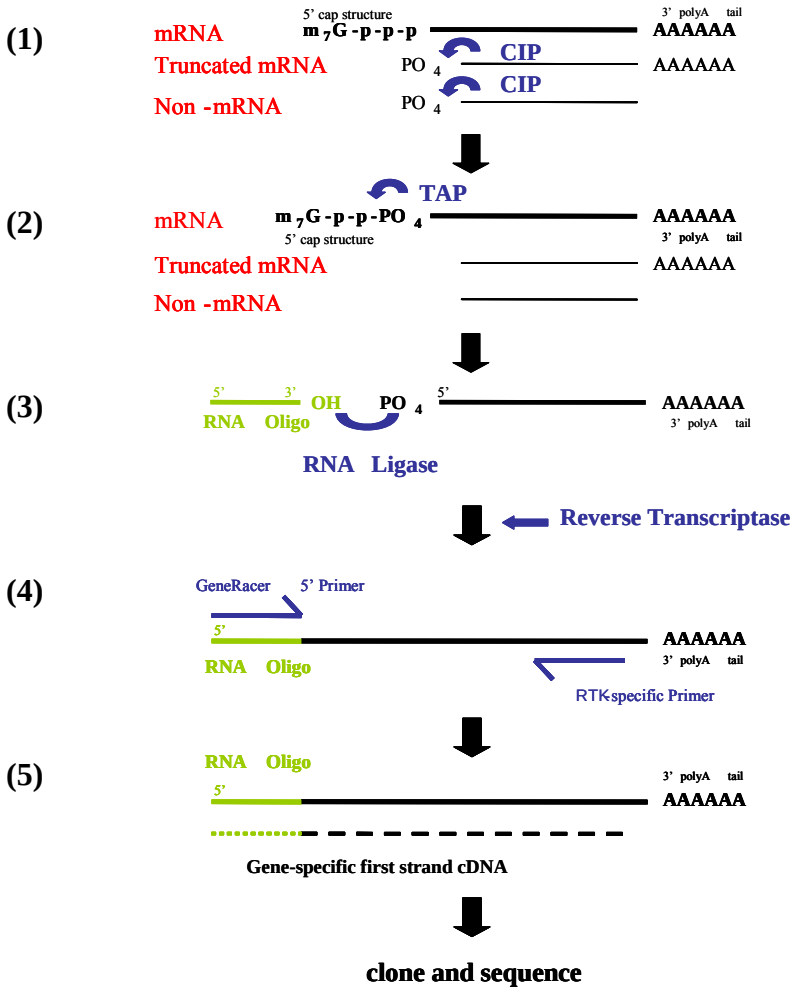


Figure 2. 1: A schematic representation of the process of 5' Rapid Amplification of cDNA Ends (5'RACE). (1) Calf intestinal phosphatase (CIP) is first used for dephosphorylation of truncated or non-mRNA. Full-length mRNA is protected from dephosphorylation due to its 7-methylguanosine cap structure. The truncated and non-mRNAs are thus eliminated; (2) Tobacco alkaline phosphatase (TAP) is then used to remove the cap structure of the full-length mRNA; (3) RNA ligase is then employed to ligate an RNA oligonucleotide of known sequence to the 5' end of the full-length mRNA. Random primers are then used to reverse transcribe the ligated full-length RNA; (4) 5'RACE PCR is then performed using an oligonucleotide complementary to the ligated sequence combined with a gene-specific primer; (5) The product obtained can then be cloned and sequenced.

2.2.13 Pyrosequencing

Pyrosequencing™ is a technique used for accurate and quantitative analysis of DNA sequences that is based on the detection of released pyrophosphate (PPi) during DNA synthesis. A cascade of enzymatic reactions occurs resulting in the production of visible light that is proportional to the number of nucleotides incorporated. The enzymatic cascade begins with a nucleic acid polymerisation reaction as a result of nucleotide incorporation by polymerase, inorganic PPi is released. The released PPi is subsequently converted to ATP by ATP sulphurylase. ATP is the energy required by luciferase to oxidize luciferin and generate light. The amount of light is then detected by a photodiode, photomultiplier tube, or a charge-coupled device camera (CCD) camera (Figure 2.2).

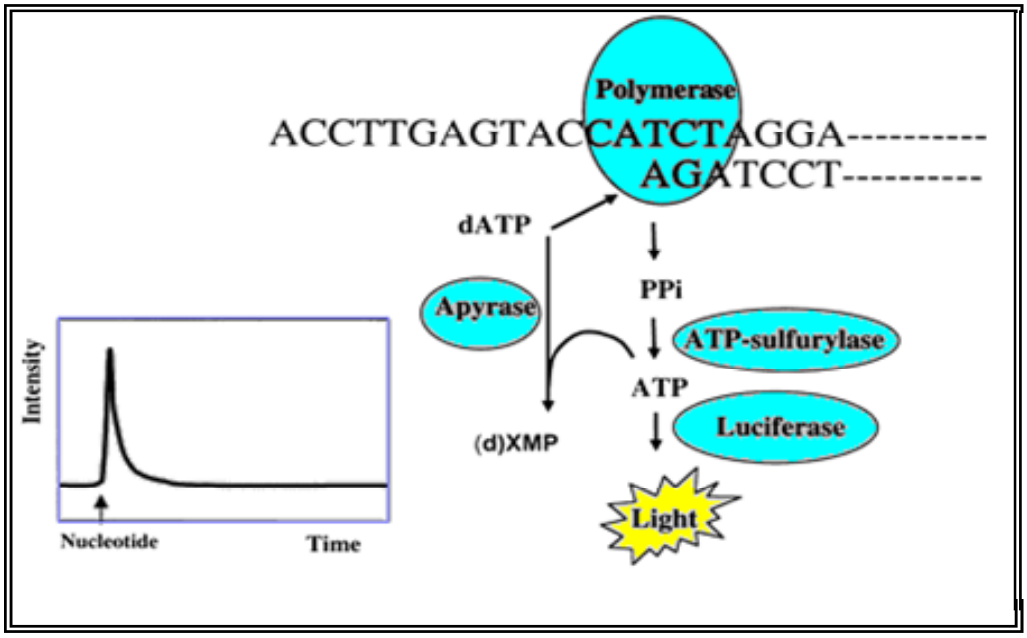


Figure 2. 2: Schematic representation of the progress of the enzyme reaction in pyrosequencing (Ronaghi 2001). A primed DNA template and four enzymes are placed in a well of a microtiter plate. The four different nucleotides are added stepwise to the reaction and incorporation is followed using the enzyme ATP sulphurylase and luciferase. The nucleotides are continuously degraded by nucleotide-degrading enzyme apyrase,

allowing addition of the subsequent nucleotide. dXMP indicates one of the four nucleotides.

Pyrosequencing was performed as follows: the amplicon in question was first generated in a 50 µL reaction volume containing 10 pmol each of the forward and reverse PCR primers, 0.2 mM deoxynucleoside triphosphates (Amersham), 2.5 mM $MgCl_2$, 1x Buffer II (Applied Biosystems), 1 U of AmpliTaq Gold (Applied Biosystems), and 1 µL (~20 ng) of DNA. PCR was performed in a Tetrad with the following conditions; 94 °C for 7 min, followed by 45 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s; 1 cycle at 72 °C for 7 min; and a final hold at 15 °C. All samples were analysed in duplicate. Pyrosequencing reactions were performed according to the manufacturer's instructions, using the Gold PSQ 96 SNP Reagent Kit (Biotage AB). The sequence was read and peaks heights quantified by the AQ Software.

2.3 Fluorescence in situ hybridisation (FISH)

2.3.1 Maxiprep DNA isolation from BAC clones

BACs were initially ordered from the Sanger Institute, Hinxton. The clones arrived as LB/Agar stabs supplemented with the appropriate antibiotic. BACs were also obtained from the 37K clone set which was subsequently acquired in house. BACs were colony purified before use by streaking out using a sterile loop onto a LB/Agar plate containing the relevant antibiotic. After overnight incubation, a single colony was picked and used to inoculate 5mls L-Broth, again containing the appropriate antibiotic. The cultures were incubated overnight in a shaking incubator (225rpm) at 37°C.

The starter culture was then used to inoculate 30ml of L-Broth medium with antibiotic in a 50ml fluted flask (Triple Red, Buckinghamshire, UK) and incubated as before. To harvest the bacteria, the overnight culture was transferred to a 50ml falcon

tube and spun at 3,000rpm in a RC5C Sorvall centrifuge. After centrifugation the supernatant was completely removed and the cell pellet was resuspended in 1.8ml of Solution I (see section 2.5 for all solution details). The resuspended pellet was then transferred to a plastic centrifuge tube. The cell suspension was then lysed by adding 1.8ml of freshly prepared Solution II and left to stand for 5 minutes at room temperature. Bacterial proteins and DNA were then precipitated out by adding 1.8ml of Solution III and leaving the tubes on ice for 5 minutes. This was then centrifuged at 10,000rpm for 10 minutes at 4°C in the RC5C Sorvall centrifuge (SS-34 rotor). Leaving the cell debris, the supernatant was then carefully removed by pipette, and aliquoted into a number of sterile eppendorf tubes containing 700µl chloroform (Sigma-Aldrich, Haverhill). After shaking to mix the chloroform-vector DNA mixture, they were spun at 13,000rpm for 5 minutes in a Thermo IEC, Micromax centrifuge. The supernatant was then removed and transferred to pre-prepared eppendorfs containing 800µl of ice-cold isopropanol (Sigma-Aldrich, Haverhill), inverted to mix and then placed at -20°C overnight to maximise DNA precipitation.

The following day, the tubes were removed from the freezer and centrifuged at 14,000rpm for 15 minutes in a Thermo IEC, Micromax centrifuge. The supernatant was then poured off and 500µl of 70% ethanol (BDH, Dorset) was added, the pellet was washed by inversion before centrifuging again at 14,000rpm for 5 minutes. The supernatant was then carefully removed by pipette, and the pellet allowed to air dry until it become translucent. Each pellet was then resuspended in 40µl of sterile water.

2.3.2 Labelling

1µg BAC DNA was labelled by nick-translation in the presence of 10U DNA polymerase I (Invitrogen, Paisley, UK) and 0.05U DNase 1 (Invitrogen, Paisley, UK). The labelling reaction was performed at 16°C for 1.75 hours using a standard labelling mix (see section 2.5). The reaction was stopped by denaturing the enzymes at 68°C for 10 minutes. 10µl of the product was then run on a 1% agarose gel to check the efficiency of the labelling reaction.

2.3.3 Slide making

Fixed cell material (approximately 5 x 10^6 cells) was spun down and washed with 3:1 methanol:acetic acid fixative and finally resuspended in 1-2mls of fixative. In a humidity-controlled room, one drop of suspension was dropped onto a microscope slide from a short height and the slide was allowed to dry. The slides were checked for spreading of nuclei, under phase contrast light microscopy. The slides were left at room temperature overnight, or exposed to UV for 10 seconds to age them before use.

2.3.4 Probe preparation

Labelled DNA was co-precipitated using an equal volume of 1.05mg/ml COT-1 DNA (EurX, Poland), one tenth the volume of 3M sodium acetate pH5.2 and twice the total volume of 100% ethanol and then frozen for 1 hour at -20°C. The labelled DNA was then spun down at 14,000rpm for 20 minutes, washed several times in 70% ethanol, and resuspended in 12µl hybridisation buffer (see section 2.5).

2.3.5 Slide preparation

Slides were incubated for 2 hours at 37°C in 2 x SSC pH 7.0 (for all solution details see section 2.5), then treated with 100ng/ml pepsin in 0.01M HCl for 10 minutes. Following the pepsin digestion the slides were washed three times with phosphate buffered saline (PBS), before being fixed in a 1% paraformaldehyde PBS for 10 minutes. The PBS wash was repeated (as described above), and then the slides were dehydrated through an ethanol series of 2 minutes in 70% ethanol, 2 minutes 90% ethanol and 2 minutes 100% ethanol. The slides were then dried. The previously prepared probe was then aliquoted onto a cover slip, placed on the dried slide and sealed with vulcanising rubber solution (Weldtite, Barton-on-Humber). When the

rubber solution had dried the slides were denatured on a 75°C hotplate for 5 minutes, and hybridised in a moist chamber for 48 hours at 37°C.

2.3.6 Post hybridisation and visualisation

Following hybridisation the cover slips were removed and the slides washed briefly in 2 x SSC. Washes were then performed at 42°C: 3 x 5 minutes washes in 50% formamide/2 x SSC pH7.0, and 3 x 5 minutes washes in 2 x SSC. Post-hybridisation biotin and digoxygenin signals were detected using the following combined antibody layers: 1) 1/100 sheep anti-digoxigenin (1mg/ml, Roche Diagnostics, Germany), 2) 1/400 texas red avidin (2mg/ml, Vector Laboratories Inc, California, USA) and 1/200 rabbit anti-sheep FITC (1mg/ml, Vector Laboratories Inc, California, USA), 3) 1/100 goat anti-avidin biotinylated (1mg/ml, Vector Laboratories Inc, California, USA), and swine anti-rabbit FITC (1.1mg/ml, Dako, Denmark) and 4) 1/400 texas red avidin (Lichter and Cremer 1992). All dilutions were made with SSCTB (see section 2.5 for solution composition). The slides were incubated for 30 minutes to 1 hour at room temperature in a dark moist chamber, and then washed 3 x 5 minutes in SSCT (4 x SSC, 0.1% Triton-X, pH7.0) between each incubation. After application of all antibody layers the slides were rinsed with PBS and allowed to dry. The slides were then mounted, under a coverslip, with Vectorshield (Vector Laboratories, California, USA) containing 1μg/ml 4,6-diamidino-2-phenylindole (DAPI). Slides were examined using an Axioskop microscope (Zeiss, Germany). The images were captured using a CCD camera and image analysis software (MacProbe v4.3, Applied Imaging International, UK).

2.4 Tissue culture

2.4.1 Imatinib sensitivity assay

2.4.1.1 Thawing of primary haematopoietic cells

Peripheral blood mononuclear cells (MNCs) were separated using lymphoprep (Axis-Shield, Oslo, Norway) and grown either directly or following cryopreservation in methylcellulose supplemented with the recombinant cytokines; SCF (20ng/ml), GM-CSF (1ng/ml) and IL-3 (5ng/ml)+ (Stem Cell Technologies Ltd., Vancouver, British Columbia, Canada). Frozen vials typically contained between $2x10^7$ and $6x10^7$ cryopreserved cells. The cryopreserved MNCs were thawed quickly by incubating at 37°C, before transferring the cells to a 50ml tube. Dropwise, 20ml of pre-warmed thawing solution (section 2.5) was then added. The tube was topped up to 50ml with RPMI, supplemented with 10% foetal calf serum (FCS), L-glutamine and penicillin-streptomycin (as before), and then spun at 1,500rpm for 5 minutes. After centrifugation, the supernatant was discarded and 200,000U of DNase I was added directly to the cell pellet. The tube was gently shaken to disrupt the pellet and to prevent the formation of clumps. To the pellet, 10ml supplemented RPMI was then added and the cells centrifuged, as before. The DNaseI and RPMI wash was then repeated before finally discarding the supernatant and resuspending the cells in 10ml RPMI. To the resuspended cells, 200,000U DNaseI was added and the cells were left for one hour at room temperature. The cells were finally spun and resuspended in RPMI for counting and viability. The cells were counted using an equal volume of trypan blue (Sigma), to differentiate between live and dead cells.

2.4.1.2 Colony assay

A cell count of between 3×10^5 and 2×10^6 can be used to set up the *in vitro* assay that measures the effect of imatinib on colony growth (Crescenzi, *et al* 2007). After counting, the cells were spun down and resuspended in 1ml of RPMI supplemented

with 10% foetal calf serum (FCS), L-glutamine and penicillin-streptomycin (as before). The assay was set up using zero, 1μM or 5μM imatinib mesylate (Novartis, Basel, Switzerland), in triplicate. To each 3ml aliquot of methylcellulose (one for each concentration of imatinib) an equal volume of cell suspension was added to the top. The amount of imatinib required for each concentration was then calculated from a 1mM stock solution and also added to the top of the methylcellulose aliquot. After the addition of the cells and the imatinib, the methylcellulose was mixed thoroughly using a 1ml syringe.

From each aliquot of methylcellulose containing equal cell numbers, 3 x 1ml aliquots were dispensed into 3 separate 40 x 12mm tissue culture dishes (Fisher Scientific, Leicestershire). All the 40 x 12mm tissue culture dishes were then separated into 2 large tissue culture trays (144 x 21mm), ensuring at least one of each concentration were in separate trays in case of infection. A small 40 x 12mm dish was placed in the middle of the large tray with sterile water in it to prevent the methylcellulose drying out. The dishes were then left in a 37ºC incubator to grow for 7 days before the number of colonies was counted. The assay was counted again on day 14. An index of response was then calculated, as the mean response at 1 and 5μM at days 7 and 14 compared to the untreated control. This allows quantitative comparison of the response of different patients. For comparison, normal controls had been previously tested by Dr Andy Chase and the average response index was found to be 0.4. A number of CML patients, acting as positive controls were also tested by Dr Chase, these showed a much greater inhibition with a response index of 0.1. On the basis of these controls an index of below 0.2 was used to define a positive response as previously described (Crescenzi, *et al* 2007).

2.4.2 Midi prep & nucleotide removal kit

2.4.2.1 Midi-prep

The PureYield™ Plasmid Midiprep System (Promega, UK) is used for the isolation of large quantities of high-quality plasmid DNA that can be specifically used in

eukaryotic transfection and in vitro expression experiments. It contains an endotoxin removal wash that removes substantial amounts of protein, RNA and endotoxin contaminants from purified plasmid DNA.

A subculture was grown up overnight in 5ml L-Broth plus the relevant antibiotic before spiking a larger 25 or 45ml volume of L-broth plus antibiotic and leaving this to grow again overnight. In brief, preparation and lysis of the bacterial cell cultures was performed by pelleting the cells at 10,000 x g for 10 minutes in a Sorvall centrifuge. The supernatant was then discarded and the cells resuspended in 3ml Cell Resuspension Solution. To this 3ml of Cell Lysis Solution is added and mixed by inversion 3-5 times before leaving at room temperature for 3 minutes to incubate. Following this 5ml Neutralisation Solution was added to the lysed cells, mixed by inversion 3-5 times before allowing the lysate to settle in an upright position for 2-3 minutes. The DNA was then purified. The lysate was poured into a PureYield™ clearing column in a clean 50ml falcon (Sarstedt, UK) incubated for 2 minutes to allow the cellular debris to rise to the surface before centrifuging at 3,000 x g for 5 minutes (repeating if the not all the lysate has passed through the column). The filtered lysate was then poured into a PureYield™ binding column in a clean 50ml falcon and centrifuged for 3 minutes at 3,000 x g. The column-bound DNA was first washed with 5ml endotoxin removal wash solution and centrifuged as above, the supernatant was discarded and the column washed with 20ml column wash solution for 5 minutes at 3,000 x g. To ensure complete removal of ethanol the supernatant was discarded and the column spun for an additional 10 minutes at 3,000 x g. Elution of the DNA was performed by placing the binding column in a new falcon and adding 300µl nuclease-free water, centrifuging at 1,500 x g for 5 minutes and then transferring the filtrate to a 1.5ml tube.

2.4.2.2 QIAquick nucleotide removal kit

To ensure the midi-prepped DNA is of high enough quality for transfection protocols it is further cleaned up using a QIAquick nucleotide removal kit (Qiagen, West Sussex) according to the manufacturer's instructions for use with a microcentrifuge.

2.4.2.3 Transfection by electroporation

Transfection is a method of introducing nucleic acids into eukaryotic cells by non-viral means. There are various transfection protocols published, all of which are based on neutralising or obviating the issue of introducing negatively charged molecules such as DNA into cells with negatively charged membranes. Chemicals such as calcium phosphate or cationic lipid-based reagents are used to coat the DNA resulting in neutralising or creating an overall positive charge improving the ability of the DNA to cross the membrane. Physical methods, such as electroporation, are based on simply punching a hole in the membrane allowing the DNA to be introduced directly to the cytoplasm. Electroporation was first used for gene transfer studies into mouse cells (Wong and Neumann 1982). It works by creating a transient pore that allows the DNA to be introduced into the cell by disturbing the cell membrane using an electrical pulse (Shigekawa and Dower 1988). Electroporation often requires more cells than chemical methods due to substantial cell death, thus initial optimisation is critical to ensure good transfection efficiency against cell viability.

The murine haemopoietic cell line Ba/F3 was used to transfect in the newly identified fusions involving the tyrosine kinase gene PDGFRβ. The Ba/F3 mouse pro B cells are dependent on IL-3 for growth and have an extremely low rate of spontaneous reversion to factor independence (Palacios and Steinmetz 1985). In addition, Ba/F3 cells express no detectable endogenous PDGFRβ (Drummond-Barbosa, *et al* 1995, Petti, *et al* 1997).

2.4.2.4 Electroporation

The day prior to electroporation, the Ba/F3 cells were split. The following day the cells were washed twice with sterile PBS and then counted using a Neubauer type haemocytometer (BDH, Poole, UK). The cells were then resuspended in pre-warmed RPMI at a concentration of 1 x 10^7 cell per ml. 400µl of these cells (4 x 10^6) was then put into a 0.4cm Gene-pulser Cuvette (BioRad, Hertfordshire) and 16µg of the DNA to be integrated with the cells added. A mock transfection of the cells with no DNA and a vector only transfection (pcDNA 3.1) were carried out in the same way. The Gene Pulser was set to 140kV (Programme 10, mammalian, HL60) and the cuvette inserted into the electroporating pod (only one possible orientation). The cells are pulsed once by pressing the red button on the front of the Gene-Pulser machine before placing the cells in the cuvette on ice for 10 minutes. After this time the cells were transferred into a sterile 25cm^2 flask (Fisher Scientific, Leicestershire) containing 10mls of RPMI 1640 medium with 10% FCS and 10% WEHI (supernatant for IL-3) and left for 2-3 days to allow the cells to recover.

2.4.2.5 Antibiotic selection

After the recovery period the cells were plated out at two different concentrations (2 x 10^3 and 5 x 10^3 cells) in 96 well plates (VWR International, Leicestershire) with selective media, in this case Geneticin® (Invitrogen, Paisley) at a concentration of 1mg/ml. The cells were counted as described previously and plated out using a multichannel pipette with 100µl per well (80 wells total). The two outer rows of each 96 well plate were filled with 100µl sterile water to prevent evaporation. The plates were left at least two weeks to allow selection. Wells with Geneticin® resistant clones were observed under phase microscopy as healthy clumps of cells. The clone positions were noted and then picked into 24 well plates (Fisher Scientific, Leicestershire) with 2 ml of RPMI 1640 medium containing 10% FCS and 10% WEHI supplemented with 1mg/ml Geneticin®. 96 well plates with control cells were expected to be completely dead, whereas 96 well plates with pcDNA 3.1 vector

control (containing Geneticin® resistance) were expected to show a significant proportion of resistant clones.

2.4.2.6 IL-3 independent growth

To confirm if the gene of interest confers IL-3 independent growth, the clones were replated over time into media with 1mg/ml Geneticin® without IL-3. The cells were subjected to a withdrawal of IL-3 over a two week period, beginning with a 50% reduction in IL-3, followed by 25%, 10%, 5%, 2% and finally no IL-3. Clones can only survive without IL-3 if the gene of interest has conferred this ability upon the cells.

2.5 Solutions

Freezing mix	4ml RPMI supplemented with 10% foetal calf serum, L-glutamine and penicillin-streptomycin, 4ml DMSO and 12ml foetal calf serum.
Guanidinium Thiocyanate (GTC)	4M Guanidinium thiocyanate, 5mM EDTA, 25mM citrate pH7.0, 0.5% sarcosyl. 7µl β-mercaptoethanol per 1ml of GTC was added fresh prior to use.
GM-mix	Granulocyte Macrophage-Colony Stimulating factor (GM-CSF) 1ng/ml final concentration, Stem Cell Factor (SCF) 20ng/ml final concentration, Granulocyte- Colony Stimulating factor (G-CSF) 100ng/ml final concentration & IL-3 5ng/ml final concentration (All from First

	Link Ltd, Birmingham, UK) in a volume of 50ml dH_2O.
Hybridisation buffer (HB)	50% formamide, 2 x SSC, 20% dextran sulphate, at pH 7.0.
L-Broth	25g L-Broth per litre dH_2O. Autoclave before use.
LB/Agar	3g Agar per 200ml L-Broth. Autoclave before use.
6M NaCl	350.64g NaCl dissolved in 1L dH_2O.
1% Parafomaldehyde PBS	PBS was prepared according to manufacturers instructions. One tablet was dissolved in 100ml of dH_2O, plus 1g Paraformaldehyde. Concentrated sodium hydroxide is used to aid dissolving.
Phosphate Buffered Saline (PBS)	PBS was prepared according to manufacturers instructions (Oxoid, Basingstoke, UK). One tablet was dissolved in 100ml of dH_2O.
1 x PBS/Tween	PBS was prepared as above with the addition of 4ml 25% Tween.
Proteinase K	50mg/ml in dH_2O.
Red Cell Lysis Buffer (RCLB)	155mM $KHCO_3$, 0.1mM EDTA, pH7.4 at 4°C.

Resuspension Buffer (RSB)	0.075M NaCl, 0.024M EDTA, pH8.0.
10% SDS	10g SDS dissolved in 100ml dH_2O.
Solution I	5mM glucose, 25mM Tris-Cl pH8.0, 10mM EDTA pH8.0
Solution II	0.2M NaOH, 1% SDS
Solution III	5M potassium acetate, 3M glacial acetic acid
20 x SSC	175.3g NaCl and 88.2g sodium citrate were dissolved in 800mls water. The pH was adjusted to 7.0 using concentrated HCl and the volume brought up to 1 litre.
SSCT	4 x SSC, 0.1% Triton-X, pH 7.0.
SSCTB	2 x SSC, 0.1% Triton-X, 0.03g/ml BSA.
Standard labelling mix	0.5M Tris-Cl pH7.8, 0.05M $MgCl_2$, 0.1M β-mercaptoethanol, 20µM dATP, 20µM dCTP, 20µM dGTP, 20µM dTTP, 0.5µg BSA and 25µg of dUTP conjugated with biotin or digoxygenin.
Thawing Solution	20mls HBSS (Invitrogen, Paisley, UK), 100,000U DNase I (Invitrogen, Paisley, UK) and 200U sodium heparin (CP Pharmaceuticals, Wrexham, UK).

2.6 List of probes used for FISH

RP11-100O5	centromeric of *PDGFRB*, 5q33
RP11-2303L4	telomeric of *PDGFRB*, 5q33
RP11-467K9	centromeric of *GOLGA4*, 3p22
RP11-158E7	telomeric of *GOLGA4*, 3p22
RP11-434C1	telomeric of *ETV6*, 12p13
RP11-525I3	centromeric of *ETV6*, 12p13.

3 IDENTIFICATION OF NOVEL FUSION GENES ASSOCIATED WITH VISIBLE CHROMOSOME ABNORMALITIES

3.1 Introduction

As described in the introduction, activation of tyrosine kinases is a recurrent theme in the pathogenesis of myeloproliferative disorders (MPDs). However, the molecular basis of disease in many individuals remains obscure. In this chapter I describe the identification of new tyrosine kinase fusion genes that arise as a consequence of chromosomal translocations.

3.2 A novel PDGFRB fusion gene

3.2.1 Patient details

A patient (from here on referred to as Patient 1) was referred to us with a diagnosis of atypical CML, and pronounced peripheral blood eosinophilia (27%). The bone marrow aspirate showed markedly increased cellularity with granulocytic hyperplasia (90%, with left shift, but no excess of blasts) with marked increase in eosinophils and their precursors. Erythropoiesis was reduced but normoblastic. He was also described as having persistent breathlessness. Cytogenetics was performed after the initial referral, and a t(3;5)(p21;q33) was identified. The clinical phenotype of the patient combined with the karyotype suggested that the *PDGFRB* gene at 5q33 might be rearranged. Fluorescence *in situ* hybridisation (FISH) was first employed to determine if this was indeed the case.

3.2.2 Two-colour fluorescence in situ hybridisation (FISH)

FISH is commonly used to detect translocation breakpoints (Baxter, *et al* 2003). The *PDGFRB* probes were labelled with differently coloured fluorochromes, and signals appear as two red/green (yellow) adjacent signals in the nucleus when hybridised to normal interphase cells (Figure 3.1). In cells with a translocation and breakpoint in *PDGFRB* the signals hybridise to the normal *PDGFRB* allele creating a yellow fusion signal, with two separate signals indicating the 'split' in the translocated *PDGFRB* allele. FISH with the BAC clones RP11-100O5 and CTD-2303L4 confirmed the rearrangement of the *PDGFRB* gene. The BAC clone RP11-100O5 lies centromeric to *PDGFRB*, and was labelled with Biotin-16-2'-deoxy-uridine-5'-triphosphate (Biotin-16-dUTP) (Roche Mannheim, Germany). CTD-2303L4 is telomeric to *PDGFRB* and was labelled with alkali stable digoxigenin-dUTP. Interphase FISH showed a separation of the red and green signals in some cells (Figure 3. 1). Some separation of signals may occur when the chromosomes are uncondensed at interphase, therefore I counted 100 cells and found 56% of interphase cells to have a separation in signals, compared to <10% of cells with split signals seen with these probes in normal controls (Baxter, *et al* 2003). These FISH results confirmed rearrangement of the *PDGFRB* gene.

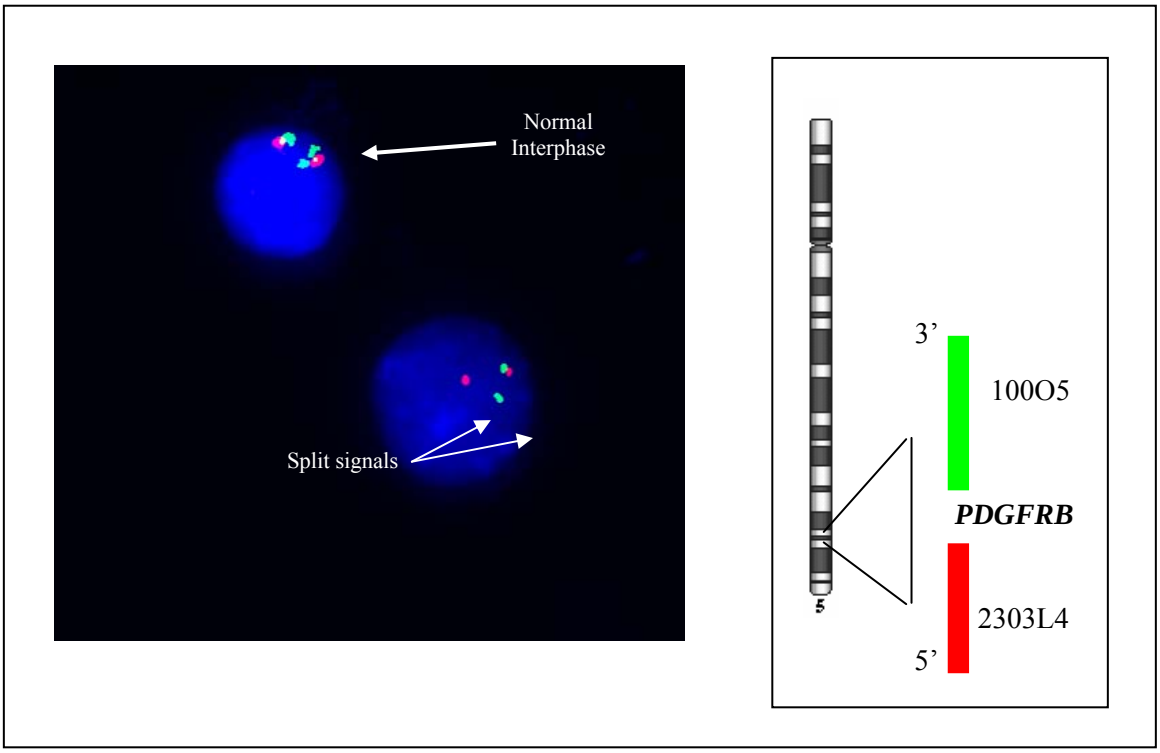


Figure 3. 1: FISH analysis of patient sample P1157. *PDGFRB* BACs RP11-100O5 (green) and CTD-2303L4 (red) hybridised to two interphase cells: (A) Normal cell with two pairs of signals fused, (B) A cell carrying a *PDGFRB* rearrangement, one fused pair and one separated pair of signals is observed. The schematic to the right indicates the position of the probes relative to *PDGFRB* and the labelling for each clone.

3.2.3 5'Rapid Amplification of cDNA Ends (5'RACE)

Having established the disruption of *PDGFRB* by FISH I set out to determine if the t(3;5) resulted in a novel fusion gene. The majority of tyrosine kinase fusions involve the joining of a partner gene to the kinase in a 5' to 3' orientation. To detect the 'unknown,' 5' partner gene sequence I performed 5'Rapid Amplification of cDNA Ends (5'RACE) using the Invitrogen GeneRacer™Kit. The basic principle of this kit is to ligate a known sequence onto the 5' end of the unknown mRNA, enabling subsequent PCR to be performed using an oligonucleotide that is complementary to the ligated sequence.

Three independent RNA samples for the t(3;5) patient were available, designated P1157, E38 and E112. All three samples were analysed in duplicate by 5'RACE PCR. I performed the 5'RACE PCRs with the High Fidelity PCR Master kit (Roche Diagnostics, Lewes, UK) according to manufacturer instructions, using the 5'GeneRacer™ primer and a reverse primer in *PDGRFB* exon 15.1R (for all primer sequences see appendix I, section IV). The amplification was performed under long-range PCR conditions (according to that specified in section 2.2.5) to enable fragments up to 10 kb to be amplified. No products were observed on a 1 % agarose gel after the first round PCR. I then performed a nested PCR using the same PCR conditions. For the nested PCR, I employed the 5'GeneRacer™ nested primer and a nested *PDGFRB* primer, exon 14.1R. As shown in Figure 3.2, several different size bands were obtained after nested PCR, two of which were very strong on the gel (indicated by the red arrows) and were sized at approximately 1 kb and 2.5 kb respectively. A third reaction yielded two medium intensity bands at approximately 0.5 kb plus a slightly larger weaker band (indicated by the white arrow).

Despite the fact that the amplified products were not consistent in size between samples, I cloned all three products using the TOPO TA cloning kit. Although visible colonies were observed on all recombinant plates, none of the picked colonies grew overnight, implying contamination, possibly via the SOC medium. Due to the strong sharp band achieved from nested 5'RACE PCR I went on to directly sequence the product from P1157 (Figure 3.2 - indicated by a star) in duplicate using the 5'GeneRacer™ nested primer (forward) and the *PDGFRB* exon 14.1R primer (reverse).

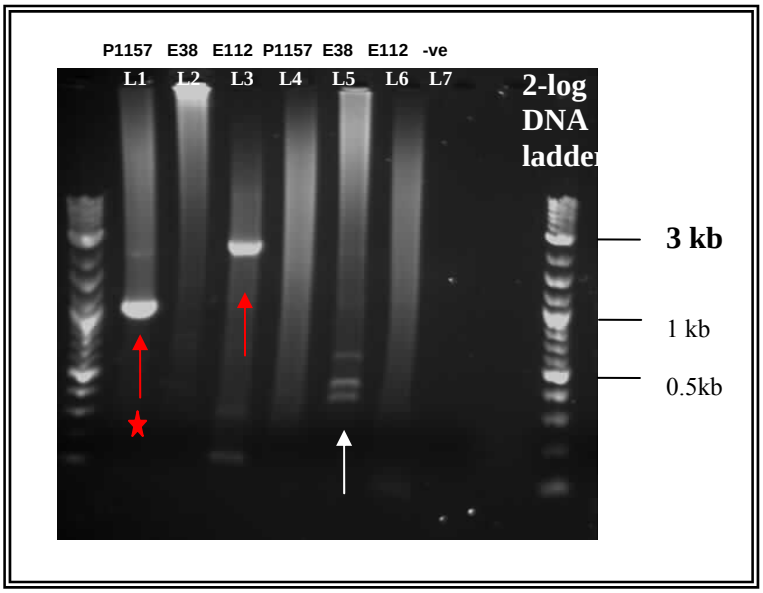


Figure 3. 2: 1% Agarose gel showing nested 5'RACE PCR for all three samples, in duplicate, of patient 1. Three different size bands were observed one in each of the samples; P1157 (L1), E112 (L3) and E38 (L5). The negative control shows no product (L7).

The sequences were first analysed by BLAT analysis (http://genome.ucsc.edu) against the whole human genome. The sequence was 630 bp in length and found to have 140 bp homology to chromosome 3, and 305 bp homology to chromosome 5. The chromosome 5 sequence was as expected, *PDGFRB* and the chromosome 3 sequence was 98.6% homologous to the gene *GOLGA4. GOLGA4* is a member of the Golgin subfamily A and is located at chromosome 3p22.3 and is close to the cytogenetic breakpoint in the t(3;5). Using the accession numbers retrieved from the UCSC genome browser I then went on to BLAST the sequence (www.ncbi.nlm.nih.gov/) first against *PDGFRB* (NM_002609), to check the fragment was homologous to the region expected from the primers used and to identify the breakpoint. The sequence was found to diverge at nucleotide 1936 of *PDGFRB* (Figure 3. 3).

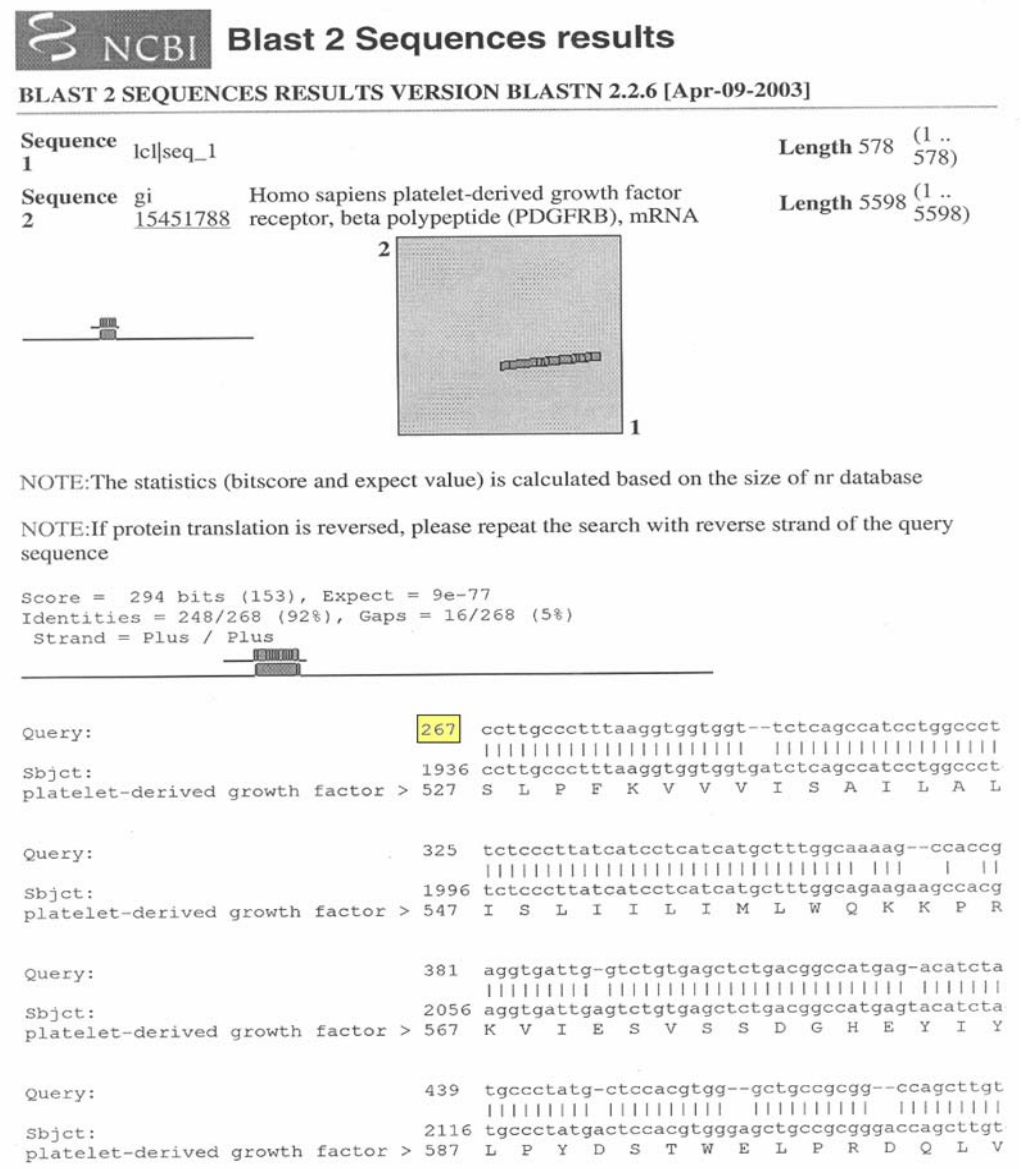


Figure 3. 3: BLAST results – comparison of *PDGFRB* cDNA (NM_002609) against sequenced product from nested 5'RACE PCR (P1157). Note that position 267 and subsequent sequence matches *PDGFRB*.

BLAST analysis was then performed on the same sequence against *GOLGA4* (NM_002078) and this also matched, ending at nucleotide 1513 (Figure 3.4). There were a number of bases not assigned by the sequencing programme due to dye spots in the original sequencing, I manually filled these in after BLAST analysis and determined the breakpoint to be at nucleotide 1520 in *GOLGA4*.

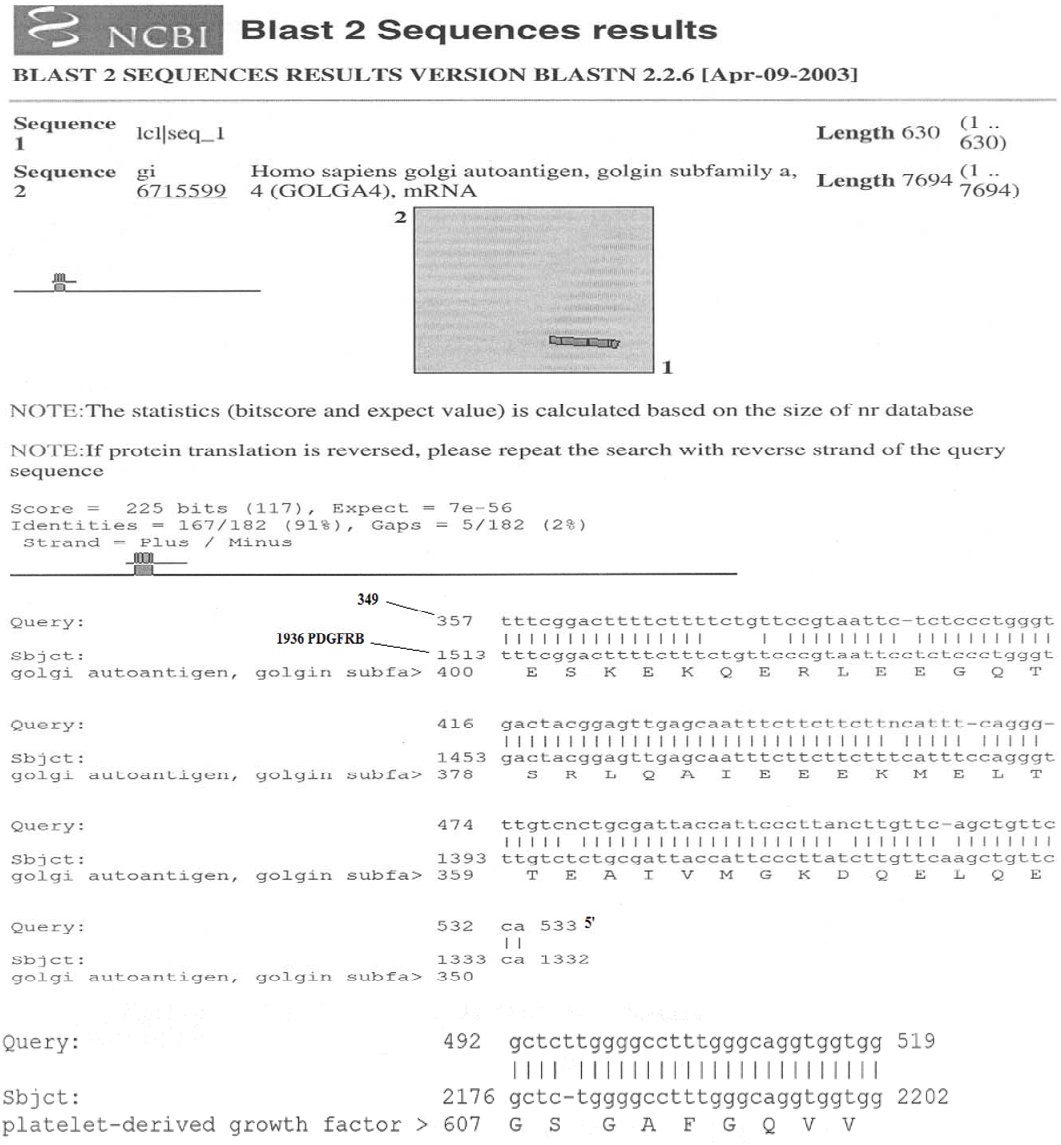


Figure 3. 4: BLAST results – comparison of *GOLGA4* (NM_002078) against sequenced product from nested 5'RACE PCR (P1157). Note sequence ends at nucleotide 1513.

The individual full-length mRNA sequences for *PDGFRB* and *GOLGA4* were then copied into DNA Strider™ (a software program used to analyse DNA sequences). The open reading frame for each sequence was determined and the sequence translated, using both the DNA Strider™ software compared with the reported protein

sequence on the NCBI website. The results from the NCBI alignment between the RACE product sequence and the mRNA sequence were then used to establish the location of the breakpoint in relation to the exons of each gene. A fusion of exon 10 *GOLGA4* to exon 11 *PDGFRB* was revealed. Subsequent DNA Strider™ analysis of the fused mRNA sequence showed the fusion is in-frame (Figure 3.5A). To confirm the mRNA fusion I used reverse transcriptase PCR (RT-PCR), before going on to clone the genomic breakpoint. This confirmation is important as 5'RACE-PCR can give rise to a variety of artifacts.

3.2.4 Confirmation of the GOLGA4-PDGFRB fusion and cloning of the genomic breakpoints.

To confirm the *GOLGA4-PDGFRB* fusion, I employed RT-PCR using the primers *GOLGA4* exon 8.1F and *PDGFRB* exon 12.1R with the High Fidelity PCR master kit according to manufacturer's conditions (see appendix I, section VI for all primer sequences). The PCR was performed on a generic 66°C annealing temperature PCR protocol. The in-frame fusion was detectable by single step RT-PCR, as seen in Lane 5, Figure 3. 5B. Using the same High Fidelity PCR kit and amplification protocol I used RT-PCR with *PDGFRB* exon 9.1F and *GOLGA4* exon 10.1R to try and amplify the reciprocal translocation. *PDGFRB-GOLGA4* reciprocal products were not detectable by single step, Lane 6, Figure 3. 5B or nested PCR.

Having confirmed the in-frame mRNA, I went on to amplify the genomic breakpoint to further substantiate the presence of this fusion. To do this I retrieved the sequences for the introns where I presumed the breakpoint to be in each gene (intron 10 of *PDGFRB* and intron 10 of *GOLGA4*, the intron directly before or after the exon in which the mRNA breakpoint was observed), focusing on these before considering any complexities, such as alternative splicing. The *PDGFRB* intron is about 3.2 kb and I designed eight reverse primers approximately every 250 bp. In some cases the primers were separated by up to 800 bp due to areas of repetitive sequence (for all primer sequences see appendix I, section VII). *GOLGA4* intron 10 is approximately

13kb and for this intron I designed a primer approximately every 1kb, again avoiding areas of repeat sequence (for all primer sequences see appendix I, section VIII). After checking the primers on the NCBI nucleotide BLAST site to ensure they only annealed to the gene in question, I set up a number of amplification reactions for the patient.

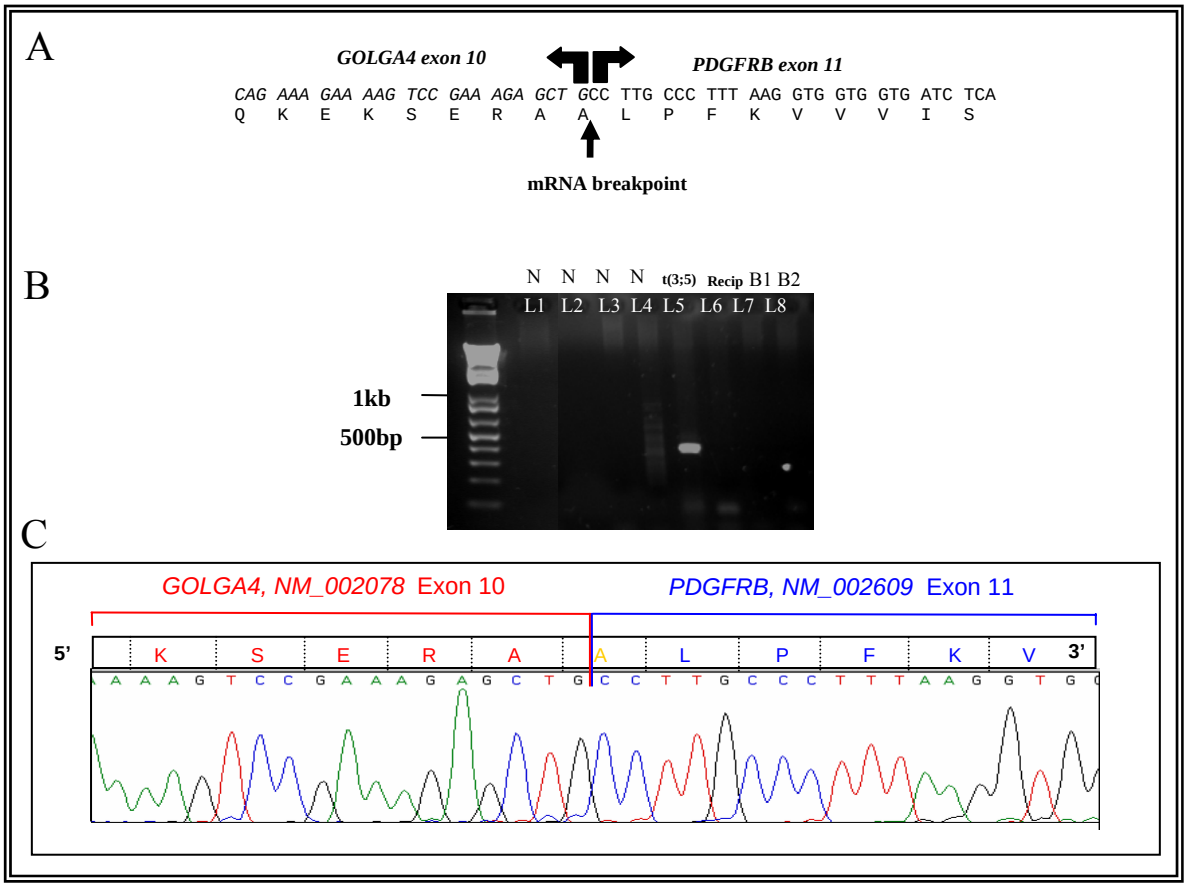


Figure 3. 5: Characterisation of the *GOLGA4-PDGFRB* mRNA fusion. (A) Schematic showing sequence and amino acid translation of the in-frame fusion, (B) Amplification of the *GOLGA4-PDGFRB* fusion by single step RT-PCR: Lanes 1-4 represent normal controls, L5 represent the patient (sample P1157), L6 represents the reciprocal breakpoint PCR (sample P1157) and L7-8 represent no DNA controls, (C) Electropherogram showing the sequencing of the single step confirmatory PCR (L5) and the mRNA junction sequence of the in-frame fusion.

All amplifications were carried out on a long range PCR program (for details of PCR program see section 2.2.5). I began by setting up fifteen reactions according to Table

3. 1, the first seven using a common *GOLGA4* forward primer and the remaining eight using a common *PDGFRB* reverse primer. No product was amplified from any of these PCRs.

Reaction No.	***Common Primer***	***Variable primer***
1	cGOLGA4 exon 8.1F	PDB.I11.7R
2	cGOLGA4 exon 8.1F	PDB.I11.6R
3	cGOLGA4 exon 8.1F	PDB.I11.5R
4	cGOLGA4 exon 8.1F	PDB.I11.4R
5	cGOLGA4 exon 8.1F	PDB.I11.3R
6	cGOLGA4 exon 8.1F	PDB.I11.2R
7	cGOLGA4 exon 8.1F	PDB.I11.1R
8	PDB.I11.1R	gGOLGA4.I9.1F
9	PDB.I11.1R	gGOLGA4.I9.2F
10	PDB.I11.1R	gGOLGA4.I9.3F
11	PDB.I11.1R	gGOLGA4.I9.4F
12	PDB.I11.1R	gGOLGA4.I9.5F
13	PDB.I11.1R	gGOLGA4.I9.6F
14	PDB.I11.1R	gGOLGA4.I9.7F
15	PDB.I11.1R	gGOLGA4.I9.8F

Table 3. 1: First set of primers used to amplify the genomic breakpoint in the t(3;5) patient. The successful primer pair is highlighted.

As a result, I designed another *GOLGA4* primer in exon 10 (cGOLGA4 exon 10.1F) closer to the mRNA breakpoint, also enabling any breakpoint between the end of the exon and the first intronic primer to be discovered (appendix I, section IX). I carried out the same amplifications, as shown in Table 3.1, but substituted the *GOLGA4* exon 8 primer for the new exon 10 primer. On this occasion a band sized at >2.5 kb was observed as result of amplification with the primer pair gGOLGA4.I9.7F, and PDB.I11.1R (Figure 3.6, Lane 15).

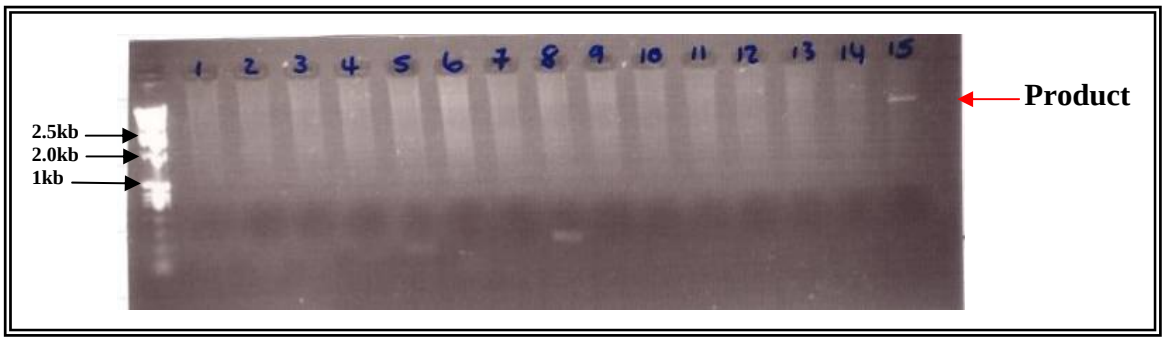


Figure 3. 6: 1% Agarose gel of the amplification products of the *GOLGA4-PDGFRB* genomic fusion. Lanes 1-7 represent the common cGOLGA4 exon 10.1F primer combined with each of the *PDGFRB* intronic primers (7-1R respectively), Lanes 8-15 represent the common PDGFRB primer I11.1R combined with cGOLGA4 exon 10.1F through to gGOLGA4.I9.7F in ascending order. Note the product in lane 15.

I cloned the amplified band using the TOPO TA cloning kit (Invitrogen) and then sequenced it. In the forward reaction, the traces showed sequence reading from gGOLGA4 I9.7F into *GOLGA4* intronic sequence up to nucleotide 9797 (according to my DNA Strider map of intron 10 *GOLGA4*). The reverse reaction showed only *PDGFRB* sequence from the mRNA breakpoint at base 1936 reading into *PDGFRB* for 33 base pairs (bp). Following this result I used a primer internal to gGOLGA4 I9.7F, gGOLGA4.I9.9F (at nucleotide 9684), combined with each of the *PDGFRB* intronic primers (1aR-7R), performing the PCR from the cloned product (lane 15, Figure 3.6). I used a generic 66°C annealing temperature PCR program with a high fidelity Hotstart *Taq* DNA polymerase (Qiagen). Figure 3.7 represents the resulting genomic fusion products from this amplification: all but one lane produced a band (L7) indicating that the breakpoint was before *PDGFRB* intronic primer PDB.I11.7R.

Using the plasmid DNA from the previously cloned product, I then directly sequenced across the genomic breakpoint using PDBI11.6R as this had produced the smallest product size of all the primer pairs. A fusion of *GOLGA4* intron 10 to *PDGFRB* intron 10 was discovered (Figure 3.8A).

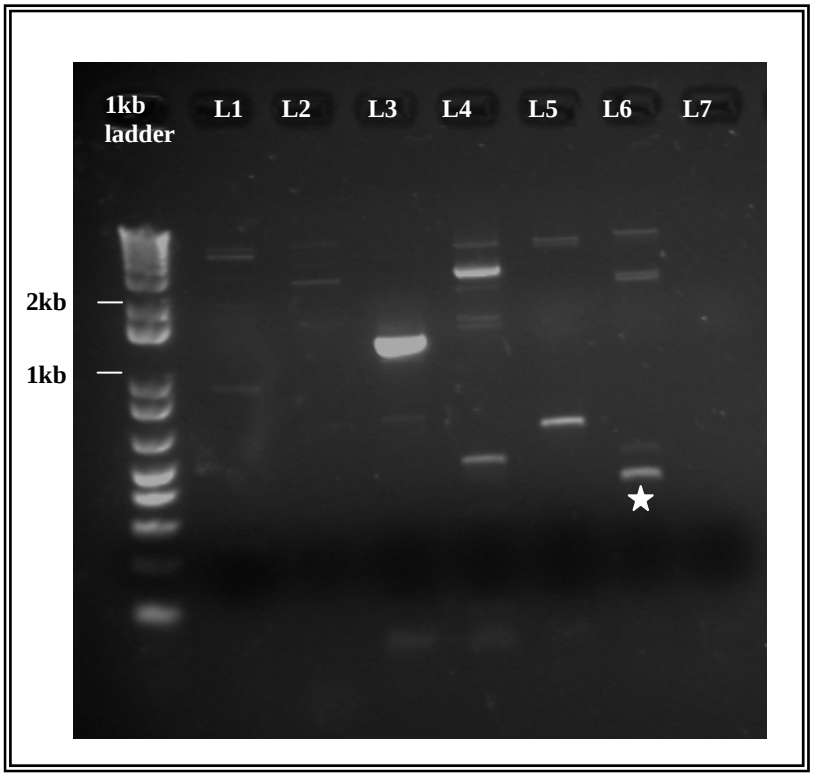


Figure 3. 7: 1% Agarose gel of the amplification products of the *GOLGA4-PDGFRB* genomic fusion amplified from plasmid DNA diluted 1:1000. Lanes 1-7 represent the common cGOLGA4 exon 10.1F primer combined with each of the *PDGFRB* intronic primers (1aR-7R respectively).

To confirm the genomic breakpoint I designed primers closer to the breakpoint in *GOLGA4* and *PDGFRB* and re-amplified the fusion using patient DNA in a single step PCR reaction, using the High Fidelity PCR Master Mix. The forward genomic primer used was in *GOLGA4* intron 10 (7F) in combination with a new *PDGFRB* reverse primer in intron 11 (Brkpt.R) (see appendix I, section X for primer sequences). This confirmed the presence of the fusion of intron 10 *GOLGA4* to *PDGFRB* intron 10 (Figure 3.8A). Using the same High Fidelity PCR kit and amplification protocol I used PDB.I11.recip.F and PDB.I11.recip.2F combined with a common intronic *GOLGA4* primer, gGOLGA4.I9.recip.R to amplify the reciprocal translocation (for primer sequences see appendix I, section XI). *PDGFRB-GOLGA4* reciprocal products were not detectable by single step or nested PCR (Lanes 3 and 4,

Figure 3.8B). This may have been because of a deletion, commonly seen at the site of the chromosomal breakpoints.

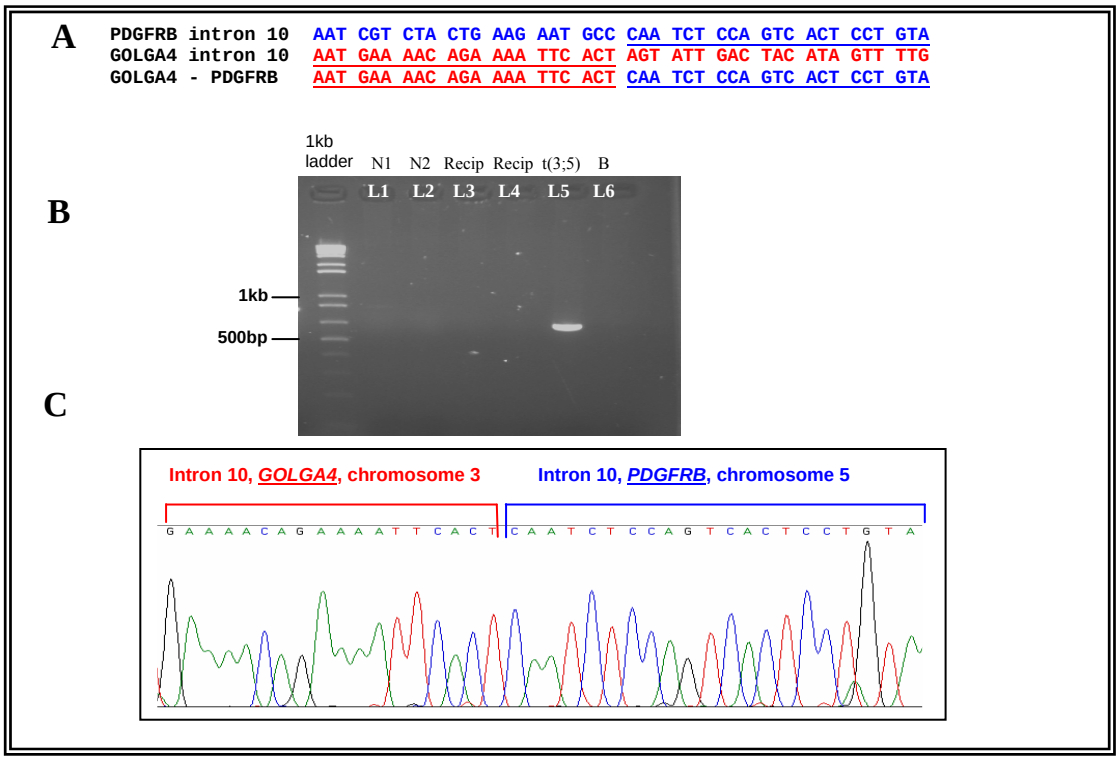


Figure 3. 8: *GOLGA4-PDGFRB* genomic fusion in patient 1 with t(3;5). (A) Schematic showing the individual sequences and the *GOLGA4-PDGFRB* fusion, (B) Amplification of *GOLGA4-PDGFRB* by single step PCR; Lanes 1-2 represent normal controls, lanes 3-4 the reciprocal *PDGFRB-GOLGA4* genomic fusion, lane 5 the *GOLGA4-PDGFRB* genomic fusion and a no DNA control in lane 6 (C) Electropherogram showing the junction sequence in the intron of each gene.

3.2.5 Consequence of the t(3;5)

The experiments performed indicate that the t(3;5) fuses *GOLGA4* exon 10 to *PDGFRB* exon 11. The predicted length of the fusion mRNA is 3662 bp with an open reading frame of 2976 bp that is predicted to generate a chimaeric protein of 992 amino acids, of which 412 amino acids are derived from GOLGA4 and 580 amino acids are derived from PDGFRβ (Figure 3.9). The predicted molecular weight of the protein is 251 kDa.

Normal GOLGA4 is a 2230 amino acid protein. At its c-terminus is a GRIP domain that targets GOLGA4 to the trans-golgi network (TGN) membranes in a G-protein dependent manner (see Figure 3.9 – indicated in yellow). GRIP domain proteins in general contain extensive coiled-coil regions (represented by the blue regions – Figure 3.9), that enable them to self associate and form rod-like homodimers (Luke, *et al* 2003). The chimaeric protein formed from the *GOLGA4-PDGFRB* fusion is expected to be structurally very similar to other tyrosine kinase fusions reported. It retains the entire catalytic domain of the tyrosine kinase fused to a portion of GOLGA4 containing substantial regions of coiled-coil domains.

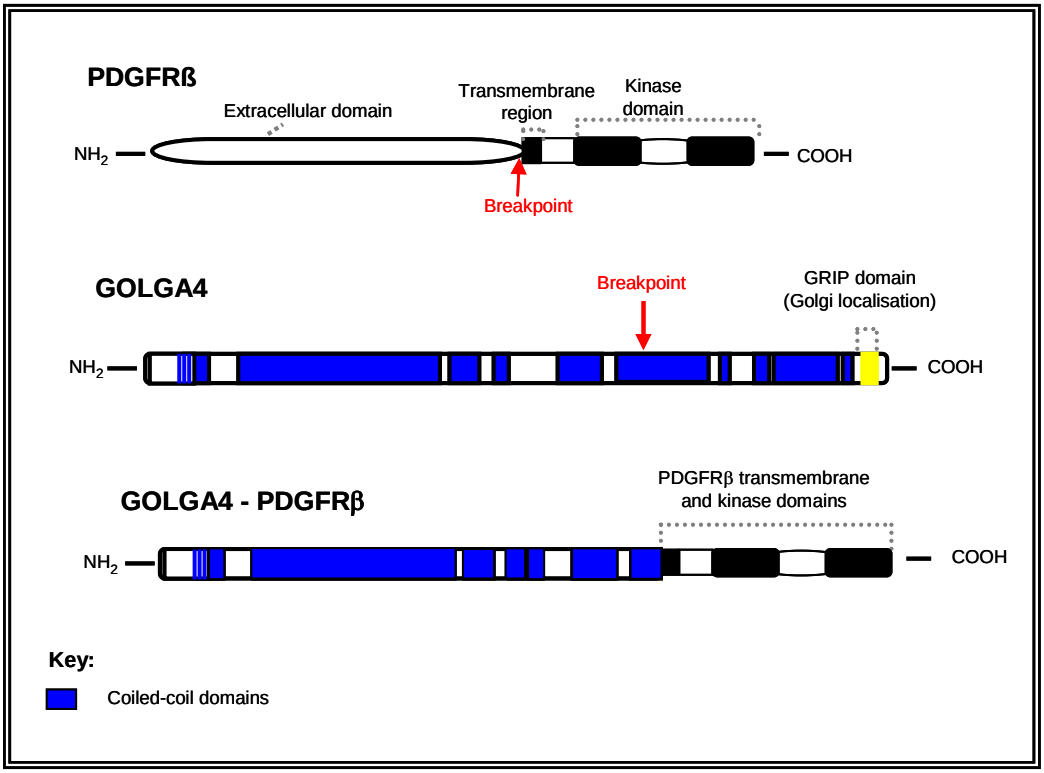


Figure 3. 9: Schematic representation of the normal PDGFRβ, GOLGA4 and GOLGA4-PDGFRβ. The fusion protein retains the entire catalytic tyrosine kinase domain of PDGRFβ and has a substantial proportion of the amino terminal GOLGA4 partner protein containing coiled-coil domains. These oligomerisation motifs in GOLGA4 are likely to be essential for the putative constitutive tyrosine kinase activity of the fusion and the subsequent downstream signalling through proliferation and survival pathways.

3.2.6 Identification of a second MPD patient with a GOLGA4-PDGFRB fusion

Following the identification of the *GOLGA4-PDGFRB* fusion, a second patient with an eosinophilic MPD and a t(3;5)(p21-25;q31-35) was referred to us for analysis (patient reference number 2).

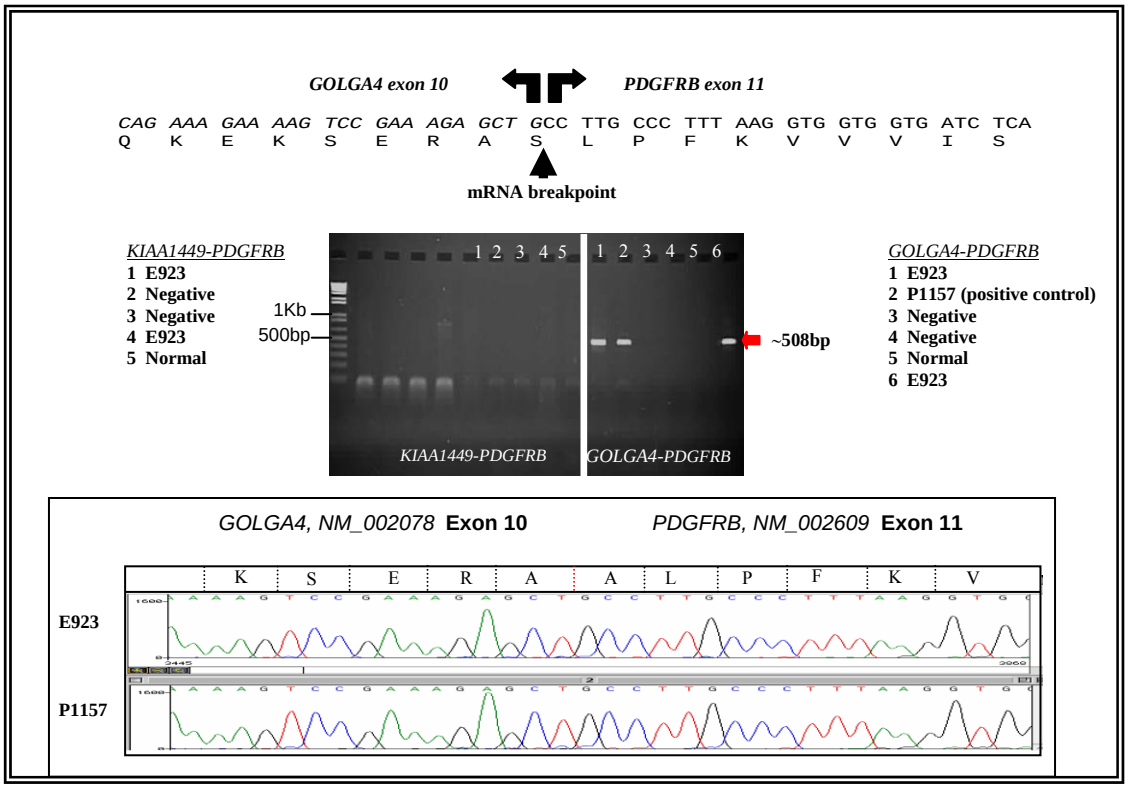


Figure 3. 10: Characterisation of the *GOLGA4-PDGFRB* mRNA fusion in Patient 2. (A) Schematic showing sequence and amino acid translation of the in-frame fusion, (B) The gel picture on the left represents the amplification of the *KIAA1449-PDGFRB* fusion; lanes 1 & 4 are patient 2, lane 5 a normal control and finally the no DNA controls are shown in lanes 2 and 3. Amplification of the *GOLGA4-PDGFRB* fusion by single step PCR is represented by the picture on the right, lanes 1 and 5 represent patient 2, lane 2 patient 1 used as a positive control, lane 3 and 4 are no DNA controls and lane 5 a normal control. A band of approximately 500bp is visible for the *GOLGA4-PDGFRB* fusion only and is the same size as that for the positive control; (C) Electropherogram comparing the mRNA junction sequence of the in-frame fusion in both patients with a *GOLGA4-PDGFRB* fusion gene. The upper trace corresponds to patient 2 (E923) and patient 1 (P1157) is shown in the lower trace.

In this case, I performed RT-PCR to analyse the patient sample for specific 3p partner genes fused to *PDGFRB*. I tested for the *GOLGA4-PDGRFB* fusion, using the primers cGOLGA4 exon 7.1F and *PDGFRB* exon 12.1R (for primer sequences see

appendix I, section XII). A second variant *PDGFRB* fusion gene had also been cloned in our group, involving a gene called at 3p22 called *WDR48* (also known as *KIAA1449*). This was also tested for using RT-PCR using the primers KIAA1449.2F combined with PDGFRB.exon15.2R (for primer sequences see appendix I, section XIII). Patient was found to be positive for *GOLGA4-PDGFRB* and negative for *WDR48-PDGFRB.* On sequencing of the product, I determined the cDNA breakpoint was the same as that for the first patient (Figure 3.10A, B and C). Again, *PDGFRB-GOLGA4* reciprocal products were not detectable by single step amplification.

3.2.7 Two-colour fluorescence in situ hybridisation (FISH)

The BAC clones RP11-100O5 and CTD-2303L4 were again used to confirm the rearrangement involved the *PDGFRB* gene. 100 interphase cells were counted and of these 52% showed a split in signals, in agreement with the PCR result (Figure 3.11A). The 52% separation in signals observed by the *PDGFRB* FISH is well above the false positive rate of 3% as determined by Baxter et. al., (2003) confirming the rearrangement of *PDGFRB* (Baxter*, et al* 2003). To further confirm the PCR result, two-colour FISH was also performed using a BAC clone that lies centromeric to the *PDGFRB* breakpoint, RP11-100O5 and a chromosome 3p BAC RP11-158E7 that lies telomeric to the *GOLGA4* breakpoint. Interphase FISH verified the presence of the *GOLGA4-PDGFRB* fusion, as the signals appeared to be fused in 36% of cells (Figure 3.11B), also indicating the fusion initially found by RT-PCR to be real.

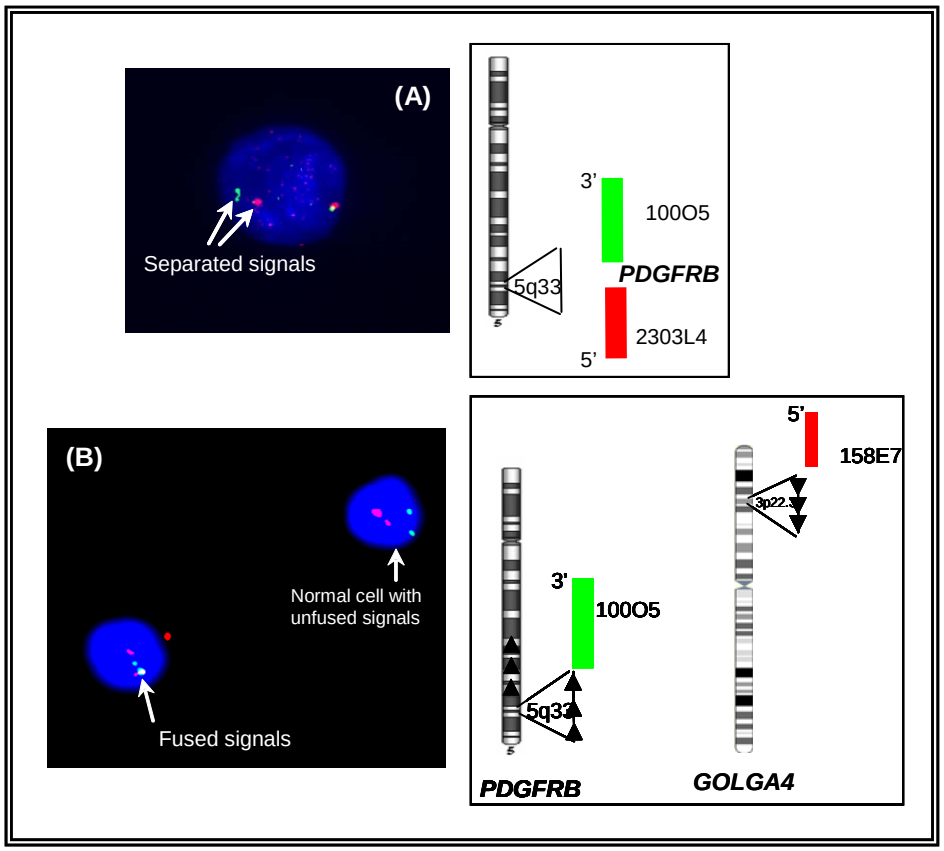


Figure 3. 11: Fluorescent *in situ* hybridisation for *PDGFRB* and *GOLGA4*. (A) *PDGFRB* BAC RP11 100O5 (green) and *GOLGA4* RP11-158E7 (red) BAC hybridised to an interphase from patient E923, note the fused signal observed in one of the cells; (B) *PDGFRB* BACs RP11-100O5 (green) and CTD-2303L4 (red) hybridised to an interphase from patient E923, the cell shows one pair of signals separated.

3.2.8 Imatinib sensitivity assay

Patients with rearrangements of *PDGFRB* usually show a good clinical response to imatinib (Apperley, *et al* 2002). To specifically test if the novel *GOLGA4-PDGFRB* fusion was sensitive to imatinib, I analysed colony-forming unit, granulocyte-macrophage growth (CFU-GM) production for cryopreserved cells from patient 2, in the presence of 0μmol/L, 1μmol/L or 5μmol/L imatinib. A normal control individual and a *BCR-ABL* positive CML patient were tested for comparison. Peripheral blood from the normal control showed a reduction in colony numbers of approximately

50% with 1μmol/L imatinib at 7 and 14 days (Figure 3.12A) and a further 10% reduction was observed at 5μmol/L imatinib. This degree of non-specific inhibition of CFU-GM numbers has been found previously with a large number of other normal controls, as previously tested by Dr Andy Chase.

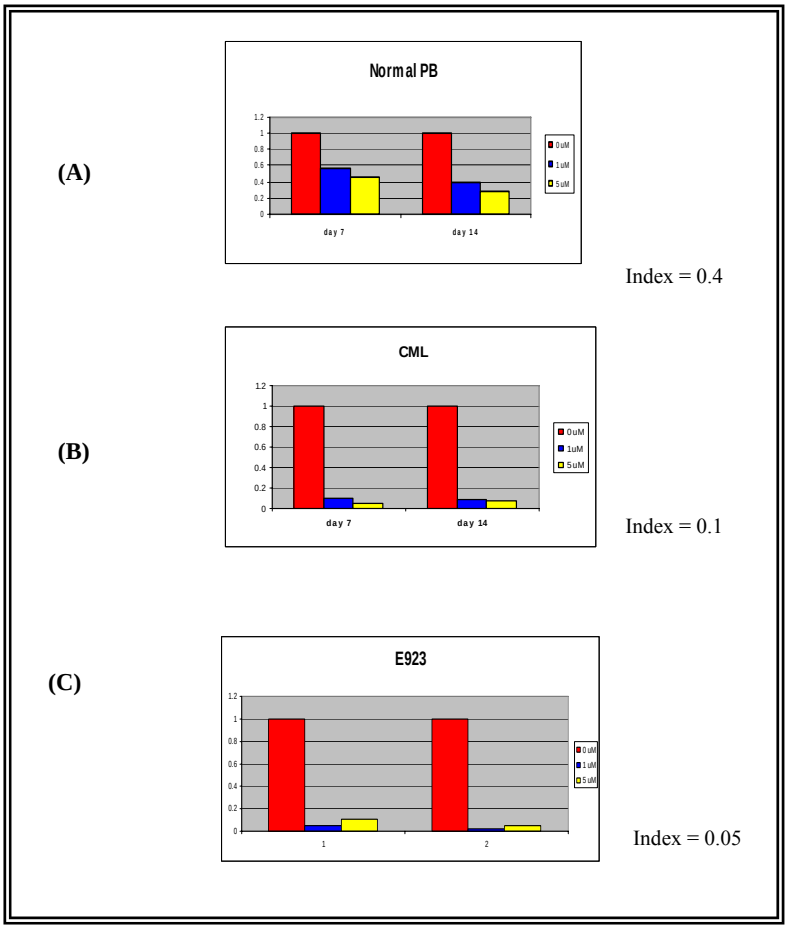


Figure 3. 12: CFU-GM growth of blood mononuclear cells from (A) normal control, (B) a CML patient and (C) patient 2.

The positive control CML patient, showed a much greater reduction in the number of CFU-GM after exposure to imatinib. Approximately 90% growth inhibition was observed in the presence of 1μmol/L and 5μmol/L imatinib at both 7 and 14 days (Figure 3.12B). This replicated similar data obtained in multiple imatinib sensitive CML samples. Marked inhibition was also observed for CFU-GM grown from the

t(3;5) patient, with approximately 95% reduction in colony numbers at day 7 and 14 in the presence of 1μmol/L and 5μmol/L imatinib (Figure 3.12C). The index of response (see methods, section 2.4.1.2) for the t(3;5) patient was 0.05, markedly lower than the 0.2 control which has been used to distinguish imatinib responders from imatinib non-responders. These data suggest that the t(3;5)-associated disease is indeed sensitive to imatinib.

3.2.9 Discussion

I have identified the gene found fused in a recurrent t(3;5) by using a combination of FISH and translocation cloning. I have discovered a novel partner gene, *GOLGA4*, fused to the tyrosine kinase gene *PDGFRB*. The breakpoint was analogous to all other reported fusions involving *PDGFRB,* with the exception of the t(5;14) (Vizmanos*, et al* 2004). The translocation breakpoint on chromosome 3 lies in intron 10 of *GOLGA4.* GOLGA4 contains extensive coiled-coils domains that are thought to act as oligomerisation motifs essential for GOLGA4-PDGFRβ constitutive activation.

GOLGA4 (Golgin-245) was first identified by two groups that were probing cDNA libraries with sera from patients with Sjrogens syndrome (an autoimmune disease) (Fritzler*, et al* 1995, Kooy*, et al* 1992). Golgin-245 is a member of the Golgin family of coiled-coil proteins, associated with the Golgi apparatus. Golgins play a role in tethering events in membrane fusion and as structural supports for Golgi cisternae. Golgin-245 belongs to a subfamily known as the GRIP domain proteins. Van Valkenburgh et. al., (2001) showed that the GRIP domain binds the ARF-like GTPases Arl1 and Arl3 with significantly higher affinity than Rab 6. Recently Rao et. al., (2003), provided evidence for the function of the Arl proteins. Through yeast GFP-GRIP fusion proteins they were able to show a GTPase cycle involving Arl3p that regulates the golgi localisation of Arl1p which in turn binds the GRIP domain and recruits the golgin to the golgi (Rao*, et al* 2003). Although, it has been suggested that Golgin-245 might function as a vesicle coat, it is perhaps more likely that it is

involved in the organisation of membrane subcompartments of the trans-Golgi network (TGN) (Brown, *et al* 2001, Luke, *et al* 2003). Nevertheless, the function of Golgin-245 remains to be fully elucidated.

To date eighteen fusion partners for *PDGFRB* have been reported: *ETV6, TRIP11, HIP1, CCDC6, RABEP1, PDE4DIP, SPECC1, NIN, KIAA1509, TP53BP1, TPM3, NDE1, ERC1, PRKG2, GPIAP1*, *GIT2, SPTBN1* and *MYO18A* (Abe, *et al* 1997, Gallagher, *et al* 2008, Gorello, *et al* 2008, Grand, *et al* 2004b, Kulkarni, *et al* 2000, La Starza, *et al* 2007, Levine, *et al* 2005, Magnusson, *et al* 2001, Morerio, *et al* 2004, Rosati, *et al* 2006, Ross, *et al* 1998, Vizmanos, *et al* 2004, Walz, *et al* 2009, Walz, *et al* 2007, Wilkinson, *et al* 2003, Wlodarska, *et al* 1995), with a further three fusion partners known within our laboratory, including *GOLGA4* (Hidalgo-Curtis et. al., submitted 2009). With the exception of the t(5;14)(q33;q24), all breakpoints within *PDGFRB* occur in intron 10 and this is also the case with the novel *GOLGA4-PDGFRB* fusion detailed here. The partner proteins in these tyrosine kinase fusions typically contribute dimerisation motifs leading to the constitutive activation of the receptor and inappropriate activation of growth signalling and anti-apoptotic pathways (Cross and Reiter 2002). In an analogous manner to the BCR-ABL oncoprotein, constitutively active tyrosine kinase fusion proteins deregulate haematopoiesis.

Interestingly GOLGA4 shows a functional overlap with the previously identified PDGFRβ fusion partners BIN2, HIP1, RABEP1 and SPTBN1. BAR domain proteins like BIN2 are involved in membrane curvature and receptor endocytosis (Ge and Prendergast 2000). Other fusion partners also demonstrate links to intracellular sorting of receptors, a well known example is BCR. BCR is a component of the endosomal sorting complex required for transport (ESCRT) (Olabisi OO 2006). MPDs are well known to be characterised by hyper-responsiveness to growth factors. Suppression of BCR in HeLa cells by small interfering RNA impairs epidermal growth factor receptor turnover, and it is conceivable that as components of fusions,

partner proteins such as BIN2, WDR48 and GOLGA4 may contribute to defective receptor internalisation and sorting.

Patients with rearrangements of the *PDGFRB* gene respond well to treatment with imatinib mesylate (Apperley*, et al* 2002). The identification of *PDGFRB* gene rearrangements is therefore critical for disease management. The identification of PDGFRB fusions in these two cases has enabled them to be treated has and permitted their response to be monitored by RT-PCR and FISH, techniques which can help to provide early clinical indications of resistance to imatinib therapy.

3.3 IDENTIFICATION OF TWO NOVEL PDGFRA FUSION GENES

The *PDGFRA* gene is located at 4q12 and to date has been described fused to six different partners: *BCR*, *FIP1L1*, *KIF5B, CDK5RAP2, STRN* and *ETV6* (Baxter*, et al* 2002, Cools*, et al* 2003, Curtis*, et al* 2007, Score*, et al* 2006, Walz*, et al* 2006). As with *PDGFRB* fusions, identification of patients that harbour *PDGFRA* fusions is particularly important as they respond well to targeted therapy with imatinib (Cools*, et al* 2003, Reiter*, et al* 2007). *PDGFRA* fusions are most commonly associated with chronic eosinophilic leukaemia (CEL), with the most frequent abnormality being seen in approximately 12% of cases with a provisional diagnosis of idiopathic hypereosinophilic syndrome (Cools*, et al* 2003, Griffin*, et al* 2003), is the *FIP1L1-PDGFRA* fusion gene. Here I describe two patients with imatinib responsive CEL who presented with acquired chromosomal rearrangements of 4q.

3.3.1 Patient details

Case 1 was diagnosed with CEL and has been described in further detail by Musto *et al*, 2004. In brief, the patient was a 64 year old male who presented with an eosinophil count of $12x10^9$/L and a bone marrow indicative of an MPD. Cytogenetic analysis revealed 60% of bone marrow metaphases harboured a t(2;4)(p24;q12).

Daily treatment with 100mg imatinib was commenced and following treatment haematologic and cytogenetic remission was achieved. This dose was continued for 9 months and then progressively reduced to 100mg on alternate days followed by 100mg twice a week. Following 24 months of imatinib, the patient refused further therapy despite not having any significant side effects. After 14 months after discontinuation of treatment elevated eosinophil counts were again detected.

Case 2 was another male, 51 years of age, who presented with late onset asthma and Type 2 diabetes. Again the patient exhibited an elevated eosinophil count of 12.9 x 10^9/L and a hypercellular bone marrow suggestive of an MPD. Cytogenetics revealed 100% of bone marrow metaphases carried a t(4;12)(q2?3;p1?2). For over 7 years the patient was managed with hydroxycarbamide but with recurrent problems due to anaemia, thrombocytopenia and very poor control of the same translocation: 46,XY,t(4;12)(q2?3;p1?2)[14]/46,XY[6]. Following the molecular diagnosis as detailed by the work done below, hydroxycarbamide was stopped and imatinib started at 100mg daily. Within four weeks, the full blood counts completely normalised, resulting in significant resolution of symptoms (apart from the asthma) and greatly improved quality of life. Following 9 months of treatment, cytogenetic analysis was again performed and indicated complete cytogenetic remission (30 metaphases).

Upon referral it was known that both cases were negative for *BCR-ABL* and *FIP1L1-PDGFRA* by RT-PCR or FISH analysis using standard procedures. The presence of a 4q translocation in conjunction with persistent eosinophilia suggested that *PDGFRA* might be involved.

3.3.2 Bubble PCR

To establish if the 4q12 translocations resulted in a novel fusion genes involving *PDGFRA*, I employed bubble PCR. Bubble PCR (Zhang, *et al* 1995) is a one sided PCR technique capable of amplifying an unknown sequence that lies adjacent to a known sequence (see section 2.2.6 for more details and appendix I, section I for

primer sequences). Like 5'RACE, this relies on the fact that fusion of tyrosine kinases involves the joining of a partner gene to the kinase in a 5' to 3' orientation. This technique also relies on the tight clustering of genomic breakpoints observed in *PDGFRA* fusions, with almost all involving a truncated exon 12 (Reiter*, et al* 2007). The technique is also dependent on specific restriction enzyme sites distal to the fusion point but equally within an amplifiable distance. The 4 bp cutters Hae*III* and Rsa*I* are used to target this region in the *PDGFRA* gene and in combination with the primer PDAI12-R4 target genomic regions of 978 bp and 615 bp, respectively (Figure 3.13).

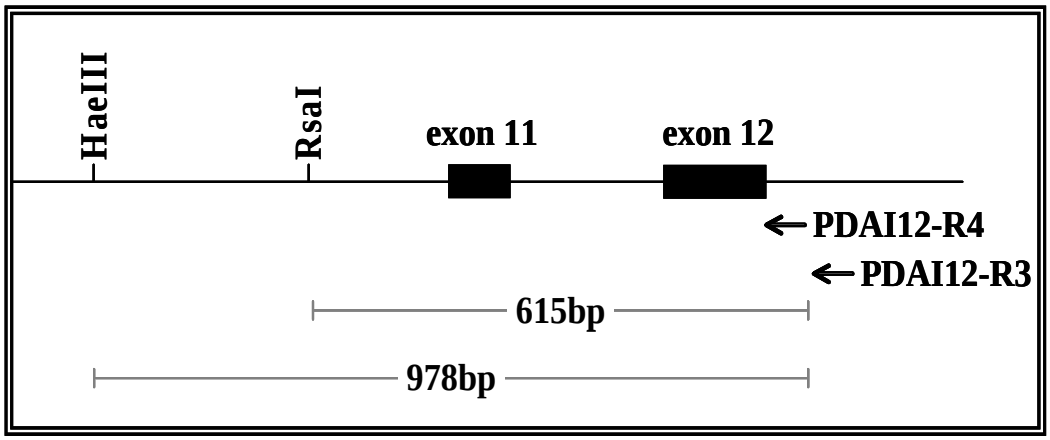


Figure 3. 13: Schematic of bubble PCR strategy showing PCR primers and restriction sites in the region where all breakpoints have been reported to fall. Arrows indicate primer locations.

On average these enzymes cut approximately every 256 bp, thus products from any rearranged *PDGFRA* allele are generally expected to be smaller than the wild-type allele. An advantage of using this approach, over that of the 5'RACE kit, is that it relies on genomic DNA not RNA. This was fortunate as the only material available for case 1 was a stained haematological slide. I employed bubble PCR and detected rearranged bands were detected for both cases 1 and 2 (Figure 3.14).

The rearranged *PDGFRA* alleles were subsequently subcloned into pCR 4-TOPO vector and sequenced using BigDye terminator v1.1 (Applied Biosystems, UK). Sequencing of the rearranged allele for case 1 identified a novel fusion between *STRN* (NM_003162) and *PDGFRA*, whereas a novel fusion between *ETS*-variant gene 6 (ETV6, NM_001987) and *PDGFRA* was found for case 2.

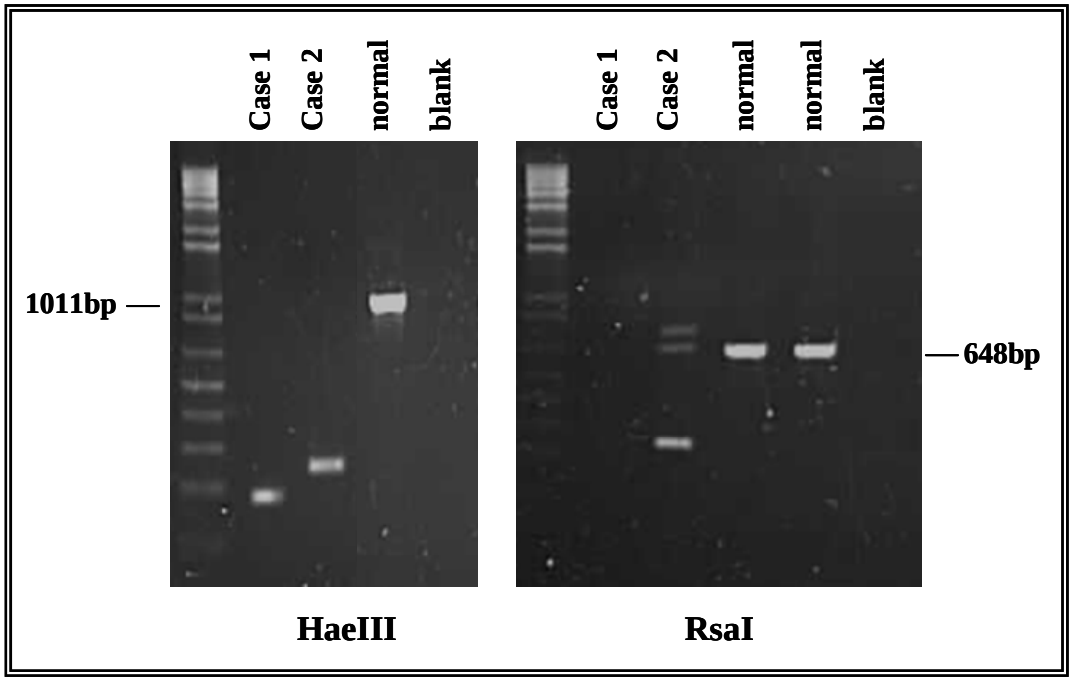


Figure 3. 14: Agarose gel picture of nested bubble PCR after digestion with Hae*III* and Rsa*I*. The normal controls show the expected germline bands of 1011 bp and 648 bp respectively (the distance between primer PDAI12-R4 and the relevant restriction site plus 33 bp contributed by the bubble oligo). Smaller bands derived from the rearranged allele are visible for cases 1 and 2 after Hae*III* digestion and for case 2 after Rsa*I* digestion. The normal allele is often out competed due to size in the amplification reaction.

3.3.3 Confirming the STRN-PDGFRA fusion

The rearranged allele resulted in a fusion of *STRN* intron 6 and a truncated *PDGFRA* exon 12. Hence, primers were designed to confirm the presence of the fusion (see appendix I, section XVI). Hot-start amplifications were performed for between 30-35

cycles in a MJ Research Inc. Tetrad thermocycler with a 66°C annealing temperature. The genomic *STRN-PDGFRA* fusion was amplified using STRN-Fusion-Ex6-1F and PDGFRA-Fusion-Ex12-2R. A 285 bp product was amplified from the patient DNA but not from control DNA. A reciprocal *PDGFRA-STRN* fusion product was not detected with primers PDGFRA-recip-Ex11-1F and STRN-Recip-Int6-1R with either single step or nested PCR (appendix I, section XVI for primer sequences) (Figure 3.15).

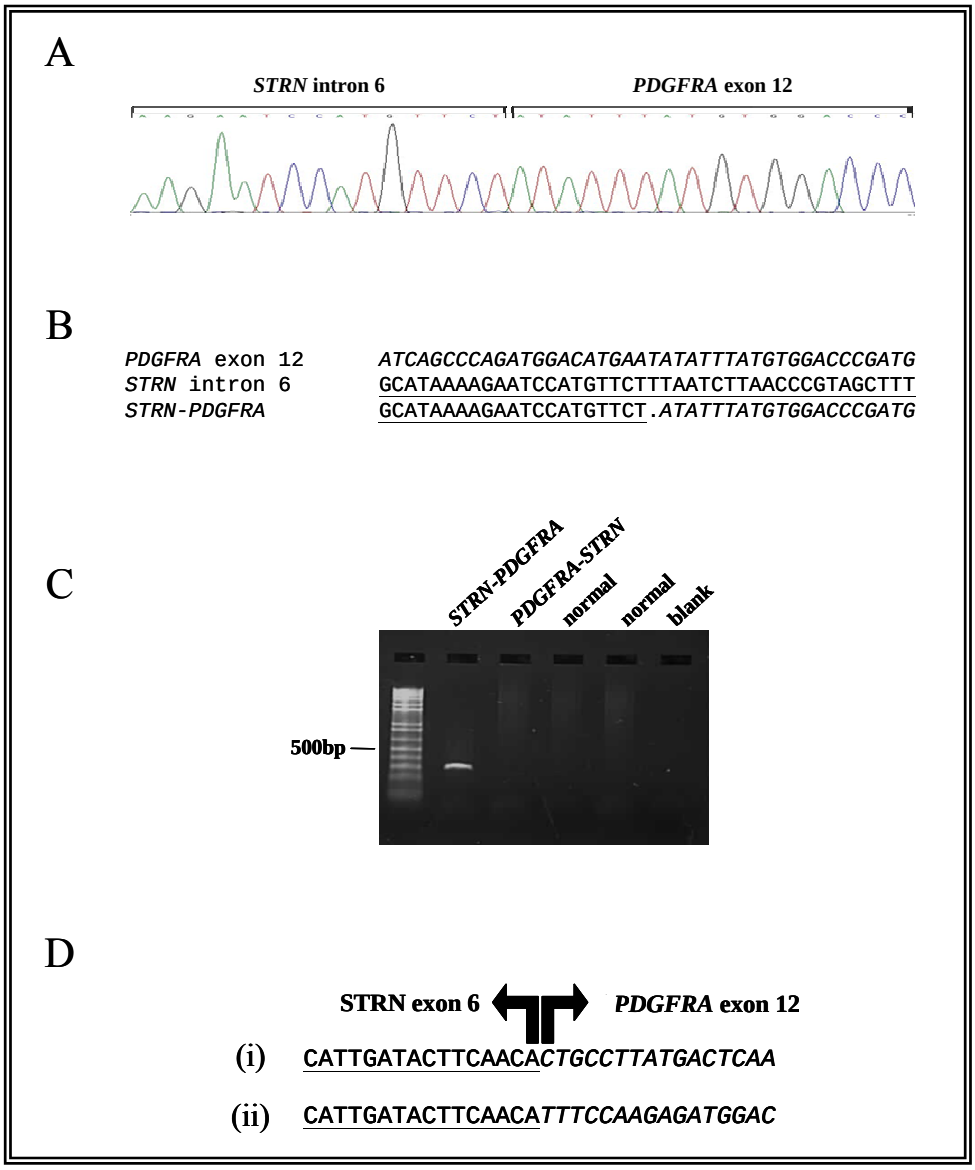


Figure 3. 15: Characterisation of the *STRN-PDGFRA* fusion gene. (A) Electropherogram of the genomic DNA junction amplified by bubble PCR; (B) Sequence of *PDGFRA* exon

12 and *STRN* intron 6 (underlined) flanking the breakpoints, aligned with *STRN-PDGFRA* genomic fusion sequence; (C) Amplification of the *STRN-PDGFRA* but not the reciprocal genomic DNA fusion from case 1 versus normal controls. (D) Predicted in frame *STRN-PDGFRA* mRNA fusions.

Due to the lack of adequate pre-treatment patient material, identification of the mRNA fusion was not possible. Nonetheless, analogous *PDGFRA* breakpoints for *FIP1L1-PDGFRA* usually result in the use of cryptic splice sites within *PDGFRA* exon 12 downstream of the breakpoint. Two potential AG splice acceptor sites were identified and would result in in-frame mRNA fusions with *STRN* exon 6. It is likely that either or both of these might represent the real mRNA fusion in case 1 (Figure 3.15D).

3.3.4 Consequence of the STRN-PDGFRA fusion

The t(2;4) is likely to fuse *STRN* exon 6 to part of *PDGFRA* exon 12. The predicted length of the fusion mRNA is dependent on the *PDGFRA* splice site used and I refer to the most likely mRNA fusions as (i) and (ii) (see Figure 3.15D). The predicted length of the mRNA fusion (i) is 5309 base pairs with an open reading frame of 2328 bp. The predicted length of the mRNA fusion (ii) is 5289 bp with an open reading frame of 2304 bp. The t(2;4)(p24;q12) is thus predicted to generate a chimaeric protein of (i) 776 amino acids or (ii) 768 amino acids, of which 265 amino acids are derived from Striatin (Figure 3.16). The predicted molecular weight of the protein is (i) 187.5 kDA or (ii) 189.4 kDa. Normal Striatin is a 780 amino acid protein. At its C-terminus is a region of WD40 repeats and its N-terminus a coiled-coil domain. This coiled-coil region is thought to be crucial for homo- and hetero-oligomerisation of the protein. Striatin is a scaffolding protein involved in calcium dependent signalling (Bartoli, *et al* 1999) and is also thought to have a putative role in trafficking (Baillat, *et al* 2001). It is highly expressed in the brain, in particular the neurons, and has been shown to directly bind caveolin (Gaillard, *et al* 2001). The chimaeric protein formed from the *STRN-PDGFRA* fusion is structurally very similar to other reported tyrosine kinase fusions. The entire catalytic domain of the tyrosine kinase is retained and

found fused to a portion of Striatin containing coiled-coil domains which probably mediate dimerisation of the fusion protein.

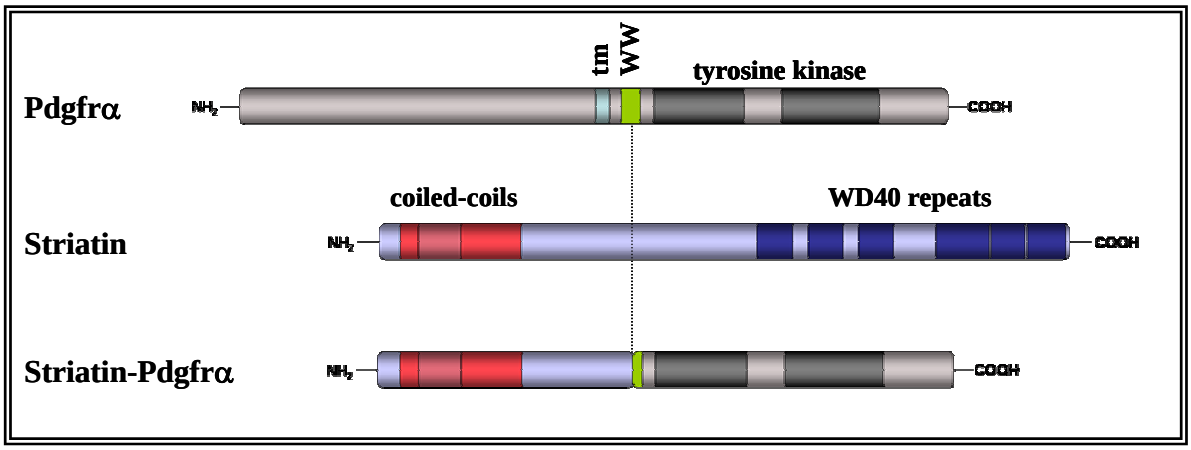


Figure 3. 16: Structure of Pdgfrα, Striatin and the predicted fusion protein. Relevant domains are shown: transmembrane (tm), WW-like (WW) and tyrosine kinase domains for Pdgfrα; coiled-coil region and WD40 repeats for Striatin.The partially deleted autoinhibitory WW domain of Pdgfrα is shown in Striatin-Pdgfrα .

3.3.5 Confirming the ETV6-PDGFRA fusion

As mentioned previously, the sequencing of the rearranged bubble PCR band from case 2 identified a second novel fusion between *ETV6*, intron 6 and *PDGFRA* intron 11 (Figure 3.17A and B). To confirm the fusion a 210 bp product (Figure 3.17C) was specifically amplified from case 2 peripheral blood genomic DNA using breakpoint-specific primers ETV6.Fusion.Int6.4F and PDGFRA.fusion.Ex12.1R on a standard 66°C Hotstart PCR amplification, a band was not observed in control DNA (for all primer sequences see appendix I, section XVII). A reciprocal *PDGFRA-ETV6* fusion product was not detected with PDGFRA-recip-Ex11-1F and ETV6-Recip-Ex7-1R at either single or nested PCR. Expression of *ETV6-PDGFRA* was confirmed by single-step from case 2, using primers ETV6.Fusion.Ex6.1F and PDGFRA.Fusion.Ex12.2R, yielding a 276 bp product that was not observed in normal control cDNA (Figure 3.27C). Sequencing of this product revealed an in-frame whole exon fusion between

ETV6 exon 6 and *PDGFRA* exon 12 (Figure 3.17D). To further confirm the presence of the *ETV6-PDGFRA* fusion, two-colour interphase FISH analysis was performed.

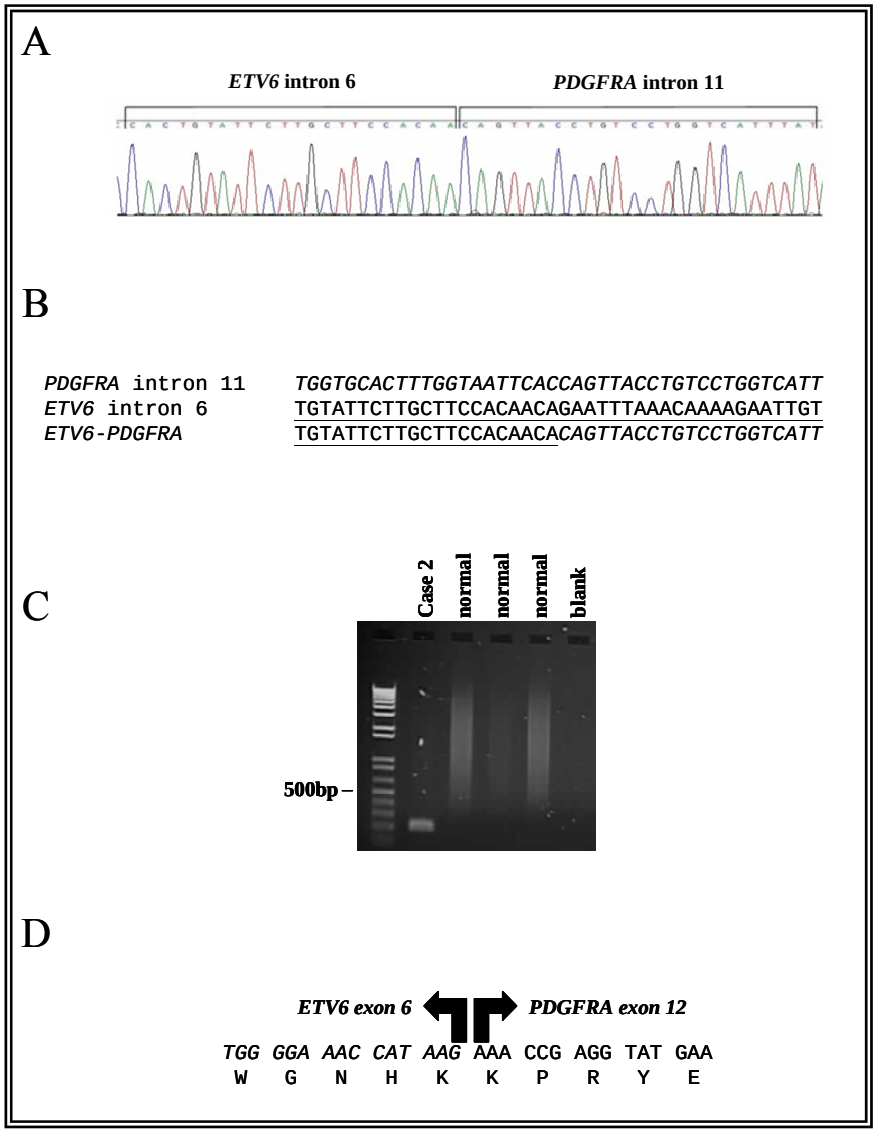


Figure 3. 17: Characterisation of the *ETV6-PDGFRA* fusion. (A) Electropherogram of the genomic DNA junction amplified by bubble PCR; B) Sequence of *PDGFRA* intron 12 (italics) and *ETV6* intron 6 (underlined) flanking the breakpoints, aligned with the *ETV6-PDGFRA* genomic fusion sequence; (C) Amplification of the *ETV6-PDGFRA* fusion from genomic DNA but not the reciprocal genomic DNA fusion from case 2, versus normal controls. (D) Predicted in-frame *ETV6-PDGFRA* mRNA fusion as determined by sequence analysis of case 2 RT-PCR product.

3.3.6 Two-colour fluorescent in situ hybridisation

FISH was only performed on Case 2 due to availability of material. FISH was executed using the bacterial artificial chromosome (BAC) clones RP11-24O10 (3' of *PDGFRA*) and RP11-434C1 (5' of *ETV6*) selected from sequences flanking *PDGFRA* and *ETV6* (www.ensembl.org) or commercially available probe sets. BAC DNA was grown, extracted and labelled by nick translation with digoxygenin or biotin-16-dUTP (Roche, East Sussex) and hybridised using standard procedures. On normal cells the combined BACs gave four separate signals corresponding to the normal chromosomes 4 and 12. However, in 56% of cells for case 2, overlapping probe signals consistent with co-localisation of *ETV6* and *PDGFRA* were observed (Figure 3.18).

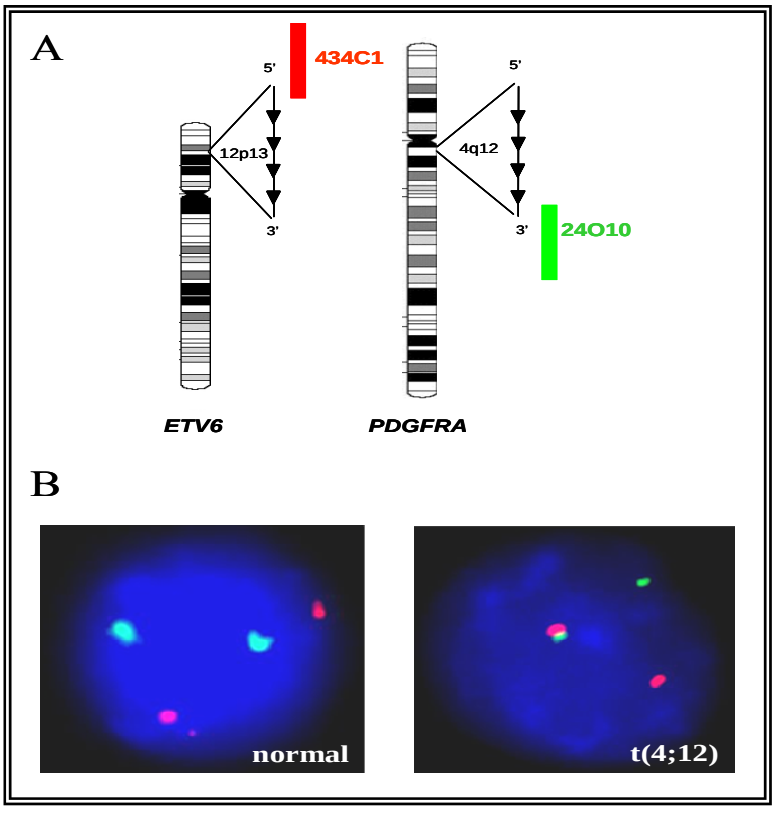


Figure 3. 18: (A) Relative positions of FISH probes with the orientation of each gene indicated by the arrows. (B) A normal cell shows four distinct signals whereas 56% of cells from case 2 showed a fused red/green signal indicative of the t(4;12).

3.3.7 Consequence of the ETV6-PDGFRA fusion

The t(4;12) fuses *ETV6* exon 6 to *PDGFRA* exon 12 resulting in a novel fusion gene. The predicted length of the fusion mRNA is 6015 bp with an open reading frame of 2769 bp. *ETV6-PDGFRA* fusion is predicted to encode a chimaeric protein of 923 amino acids, of which 384 amino acids are derived from ETV6 and 539 amino acids derived from PDGFRα (Figure 3.19). The predicted molecular weight of the protein is 229.3 kDa. Normal ETV6 is a 452 amino acid protein with two functional domains: a C-terminus pointed domain (PNT) that plays a role in protein-protein interactions including self association and interactions with other proteins. At its C-terminus is the ETS domain that is crucial for DNA-binding. The entire catalytic domain of PDGFRα is retained and found fused to almost the entire ETV6 protein containing the whole pointed domain and a part of the ETS domain.

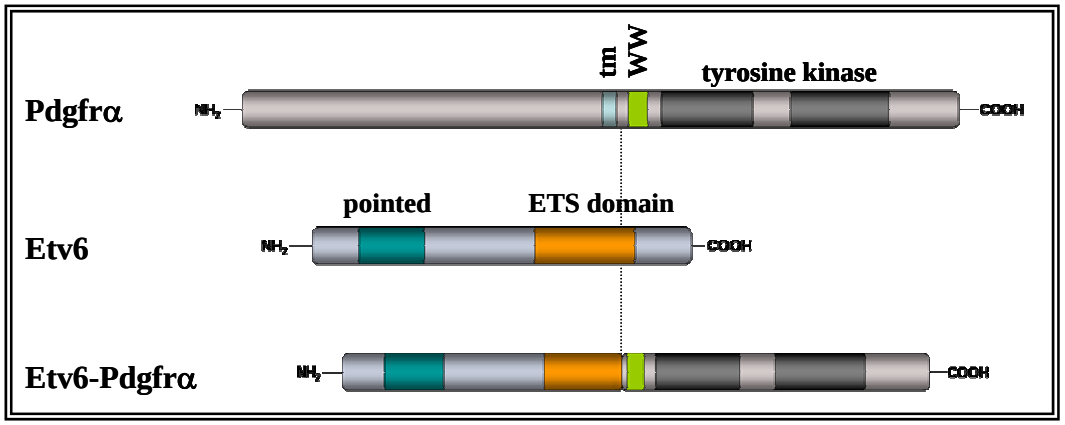


Figure 3. 19: Structure of PDGFRα, ETV6 and the predicted fusion protein. The pointed and ETS domains are shown for ETV6. The dotted vertical line indicates the breakpoint region in each protein.

3.3.8 Detection of MRD and response to Imatinib in Cases 1 & 2

Characterisation of the fusion in each case enabled a molecular assessment of minimal residual disease to be made. Minimal residual disease (MRD) was analysed for cases 1 and 2 by two step nested PCR to maximize sensitivity of detection. Primers used for the detection of the *STRN-PDGFRA* fusion were: STRN-Ex6-1F and PDGFRA-Ex12-2R and for nested PCR, STRN-Ex6-2F and PDGFRA-STRN-Ex12-1R. Primers used for the *ETV6-PDGFRA* fusion were: ETV6-Ex6-1F and PDGFRA-Ex12-2R and for nested PCR ETV6.Ex6.2F and the same PDGFRA.STRN.Ex12.1R (see appendix I, section XVIII for all MRD primers). Peripheral blood for cases 1 and 2 was analysed approximately 28 and 9 months after starting imatinib, respectively. MRD was not detected using cDNA or genomic DNA for case 1, indicating the achievement of molecular remission. For case 2, analysis was only performed using genomic DNA. Of four replicate reactions, two were positive indicating low level disease, consistent with a major molecular response.

3.3.9 Discussion: PDGFRA fusions

Over the last decade, our understanding of the molecular pathophysiology of IHES and CEL has advanced dramatically. *PDGFRA* fusions are the most commonly associated with CEL, with the most frequent being *FIP1LI-PDGFRA* which is seen in up to 12% of cases provisionally diagnosed as IHES (Cools, *et al* 2003, Griffin, *et al* 2003). Using bubble-PCR analysis I have characterised the genomic breakpoints in two CEL cases with chromosome 4q rearrangements and identified *STRN* at 2p22 and *ETV6* at 12p13 as novel *PDGFRA* fusion partners. This increases the number of known *PDGFRA* partner genes to six, with *PDGFRA* being the tenth tyrosine kinase reported to fuse to *ETV6*. It has not been possible to estimate the frequency of these abnormalities since these fusions were referred specifically for investigation of 4q rearrangements and were not part of a series that underwent systematic cytogenetic investigation. However it is thought they probably only account for < 1% of CEL cases.

STRN encodes striatin, a 780 amino acid protein. It is a member of the calmodulin-binding WD repeat family of proteins. It contains four protein-protein interaction domains, a caveolin binding domain (55-63 amino acids), a coiled-coil domain (70-116 amino-acids), a calcium dependent calmodulin-binding domain (149-166 amino acids) and a series of WD repeat motifs (419-780 amino acids) (22). Striatin is thought to be a scaffolding protein that targets proteins such as the oestrogen receptor to the cell membrane allowing the assembly of other proteins required for rapid activation (Lu*, et al* 2004). Interestingly, in case 1 the fusion identified between intron 6 *STRN* to a truncated exon 12 of *PDGFRA*, results in the loss of a large region of WD repeat motifs within striatin. However, the N-terminal coiled-coil domain remains and it is this domain that could act as a necessary oligomerisation motif, constitutively activating the PDGFRα-derived tyrosine kinase moiety. Alternatively, truncation of the WW-like domain may be the critical activating event (Figure 3.26) (Stover*, et al* 2006).

ETV6 is a member of the *ets* family of transcription factors and is frequently rearranged in both myeloid and lymphoid malignancies (Bohlander 2005). It was first cloned by Golub et al (1994) as a result of an *ETV6-PDGFRB* fusion gene in a patient with chronic myelomonocytic leukaemia (CMML) (Golub*, et al* 1994). More then 40 distinct translocations involving ETV6 have been identified cytogenetically with greater then 20 partners described so far (Odero*, et al* 2001, Rowley 1999). *ETV6* fusion partners can be categorised loosely into fusions with 1) protein tyrosine kinases; 2) transcription factors and 3) unproductive fusions, where ectopic expression of nearby genes not normally expressed within the haematopoietic system are thought to be expressed (Bohlander 2005, Cools*, et al* 2002). There is also evidence that *ETV6* functions as a tumour suppressor gene in *RUNX-ETV6* positive childhood ALL, where in 70% of cases the wild-type or normal *ETV6* allele is deleted. In cases where a deletion is not observed a point mutation can be found that abrogates *ETV6* expression (Cave*, et al* 1997, Griesinger*, et al* 2002, Montpetit*, et al* 2002, Patel*, et al* 2003).

BCR was the first described PDGFRα fusion partner. *BCR-PDGFRA* was amplified in two patients with an atypical chronic myeloid leukaemia (aCML) and a t(4;22) (Baxter, *et al* 2002). Two more cases with varied haematological malignancies have since been discovered and whilst the fusion breakpoints are diverse in BCR (exons 1, 7, 12 and 17) all show fusion to *PDGFRA* exon 12, ensuring the entire catalytic domain of PDGFRα is retained (Safley, *et al* 2004, Trempat, *et al* 2003). Analysis of *FIP1L1-PDGFRA* breakpoints show *FIP1L1* breakpoints are widely clustered in exons 7–10 and consistently result in truncation of *PDGFRA* between the two conserved tryptophan residues (the WW domain) of exon 12 (Cools, *et al* 2003, Griffin, *et al* 2003, Klion, *et al* 2003, La Starza, *et al* 2005, Pardanani, *et al* 2003, Roche-Lestienne, *et al* 2005, Vandenberghe, *et al* 2004). Consistent with all other reported *PDGFRA* fusions, I found that *STRN* disrupts the region encoding the negative regulatory WW-like domain of PDGFRα. In contrast, for *ETV6-PDGFRA* the WW-like domain is intact (Figure 3.29). This is the first case of a *PDGFRA* fusion where disruption of *PDGFRA* exon 12 within the WW domain has not occurred. The *ETV6-PDGFRA* fusion fuses *ETV6* exon 6 to the entire length of *PDGFRA* exon 12. In a recent paper by Stover and colleagues the significance of the breakpoint restrictions in *PDGFRA* was demonstrated by the production of a truncated exon 12 *PDGFRA* construct. Despite the absence of any partner gene, truncation of the PDGFRα WW-like domain was sufficient to activate the kinase and transform BaF/3 cells to growth factor independence. This paper may in part explain why the *ETV6-PDGFRA* fusion is still leukaemogenic, despite the fact that the breakpoint in *PDGFRA* does not disrupt the WW domain. An artificial *ETV6-PDGFRA* fusion that retained the region encoding the WW-like domain transformed Ba/F3 cells to growth factor independence, indicating that autoinhibition by the WW-like domain can be overcome by enforced homodimerisation. However the growth of the transformed cells was slower than those transformed by similar fusions in which the WW-like domain was truncated (Stover, *et al* 2006). It remains unclear why the WW-like domain is truncated with other fusion partners that provide an oligomerisation domain such as BCR, and also why truncation of this domain is not usually seen for fusions involving PDGFRβ.

Both patients are now being treated with imatinib. Case 1 has achieved complete chromosomal and molecular remission whilst on treatment, and Case 2 has been treated with low dose imatinib (100mg/day) which has induced a rapid haematological improvement however, follow up is currently too short to assess cytogenetic or molecular remission. In summary, I have identified two novel *PDGFRA* fusion genes in male patients presenting with an imatinib-responsive eosinophilic myeloproliferative disorder (MPD). Like *PDGFRB*, *PDGFRA* also fuses to diverse partner genes in CEL.

3.4 A NOVEL FGFR1 FUSION GENE

A common theme among the myeloproliferative disorders is the diverse tyrosine kinase fusion genes. In addition to those involving *PDGFRA* and *PDGFRB*, recurrent translocations involving chromosome band 8p11 and the tyrosine kinase *FGFR1* are also described in myeloid malignancies. The 8p11 myeloproliferative syndrome (EMS), also known as stem cell leukaemia-lymphoma (SCLL) is a chronic MPD that commonly presents with eosinophilia, T-cell proliferation and acquired reciprocal chromosomal translocations with recurrent breakpoints at chromosome band 8p11 that disrupt the receptor tyrosine kinase *FGFR1*. To date 8 fusion genes have been identified in patients with EMS. The translocations are t(8;13)(p11;q12), t(8;9)(p11;q33), t(6;8)(q27;p11), t(8;22)(p11;q22), t(8;19)(p12;q13.3), t(8;17)(p11;q23), t(7;8)(q34;p11) and t(8;12)(p11;p11) which fuse *ZMYM2 (ZNF198/RAMP/FIM*), *CEP110*, *FGFR1OP (FOP)*, *BCR*, *HERV-K, MYO18A, TRIM24 (TIF1)* and *FGFR1OP2* to *FGFR1*, respectively (Belloni, *et al* 2005, Demiroglu, *et al* 2001, Grand, *et al* 2004a, Guasch, *et al* 2003, Guasch, *et al* 2000, Popovici, *et al* 1999, Walz, *et al* 2005, Xiao, *et al* 1998). Imatinib however is inactive against *FGFR1* and cannot benefit patients with EMS. Other compounds that target FGFR1 have been developed including SU5402, PD173074, and TKI258. They function by blocking the ATP binding site in a manner analogous to imatinib, but none are currently in clinical use (Chase, *et al* 2007, Demiroglu, *et al* 2001,

Mohammadi, *et al* 1997). These compounds have been shown to specifically inhibit the growth of ZNF198-FGFR1 and BCR-FGFR1 transformed cells, resulting in increased levels of apoptosis (Chase, *et al* 2007). In this study, I have investigated a patient that presented with EMS and an acquired chromosomal rearrangement of 8p11, resulting in the fusion of a new partner gene.

3.4.1 Patient details

The patient referred was a 75-year-old female who presented with cervical, left axillary and inguinal lymphadenopathy. A computerised tomography (CT) scan showed right pleural effusion, ascites (abnormal build up of fluid around the abdomen), splenomegaly, and para-aortic nodes. Axillary node biopsy showed effacement of normal architecture by a diffuse population of mononuclear cells. The majority expressed strong surface CD3, but there was a scattered population of cells expressing neutrophil elastase amongst other myeloid markers. The presenting blood count showed haemoglobin of 140 g/L, WBC $16x10^9$ L^{-1}, with neutrophilia ($12x10^9$ L^{-1}) and eosinophilia ($0.8x10^9$ L^{-1}) plus platelets of $375x10^9$ L^{-1}. Hypercellular bone marrow was observed with an increase in eosinophils and eosinophil precursors but no increase in lymphoid cells or blasts. Cytogenetic analysis revealed 46,XX,t(8;12)(p11;q15)[19]/46, XX. There was rapid clinical deterioration accompanied by pyrexia (fever), lethargy, weight loss, and a widespread rash (biopsy was inconclusive on histology). Intensive chemotherapy was not offered and although there was a transient clinical response to steroids, the patient died 10 weeks later.

3.4.2 Fluorescence in situ hybridisation

FGFR1 FISH had been performed and previously reported as rearranged in this patient (Sohal, *et al* 2001). The breakpoints for all cases involving the *FGFR1* gene thus far reported are tightly clustered within or close to exon 9 of this gene, further investigation was made easier. I explored the possibility that *FGFR1* was rearranged by 5'RACE PCR.

3.4.3 5'Rapid Amplification of cDNA Ends (5'RACE) PCR

The partner gene was identified using the Gene-Racer™ Kit (Invitrogen). Briefly, this involves ligating a known sequence onto the 5' end of the unknown mRNA, enabling subsequent PCR to be performed using an oligonucleotide that is complementary to the ligated sequence. A single RNA sample (E0003) was available for the t(8;12) patient and this was analysed using 5'RACE PCR primers and nested *FGFR1* primers as previously described (Grand, *et al* 2004a). I performed the 5'RACE PCRs with the High Fidelity PCR Master kit (Roche Diagnostics) enabling amplification of up to 3 kb with an annealing temperature of 66°C, according to manufacturer instructions. Products were cloned with the TOPO TA Cloning Kit for Sequencing (Invitrogen) and sequenced. As shown on Figure 3.30, sequencing of the rearranged 5'RACE band revealed a fusion between *CPSF6* intron 8 (accession number NM_007007) and *FGFR1* at nucleotide 1272 from ATG (accession number M34185). The *CPSF6* breakpoint within intron 8 results in the inclusion of a portion of the *CPSF6* intron in the fusion mRNA, this fragment of intronic *CPSF6* maintains the reading frame with *FGFR1* and encodes 11 amino acids.

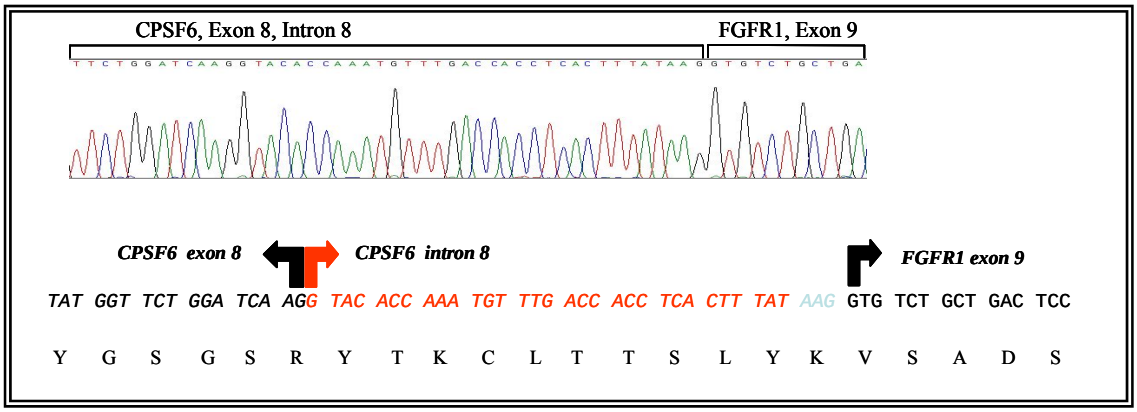


Figure 3. 20: Electropherogram (top) and sequence (bottom) of the *CPSF6-FGFR1* mRNA fusion determined by cloning and sequencing a 5'RACE PCR product. Region highlighted in red indicates incorporated *CPSF6* intronic sequence.

3.4.4 Confirming the CPSF6-FGFR1 fusion

To confirm the *CPSF6-FGFR1* fusion, a 224 bp product was specifically amplified from patient cDNA but not from control cDNA. The genomic breakpoint product of 543 bp was also co-amplified in the same reaction and sequenced (Figure 3.31A). A reciprocal *FGFR1-CPSF6* fusion product was not detected with either single step or nested PCR. The *CPSF6-FGFR1* fusion mRNA was confirmed by sequencing. The genomic breakpoint was confirmed by sequencing the larger amplified 543bp band, this is the same as the mRNA fusion due to inclusion of *CPSF6* intronic sequence (detailed below Figure 3.31B).

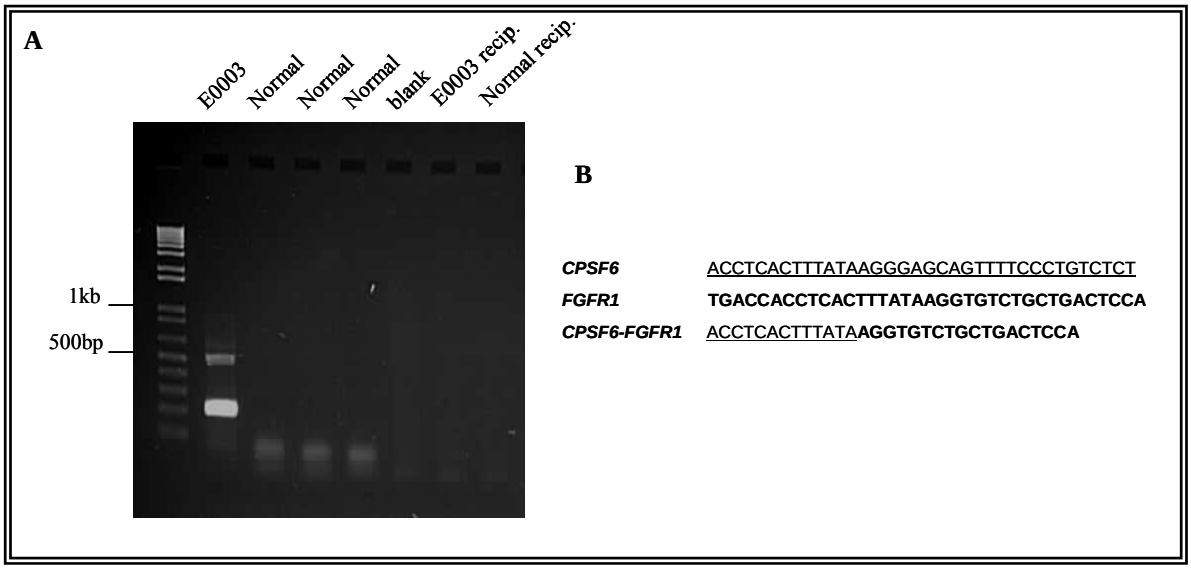


Figure 3. 21: (A) Amplification of the *CPSF6-FGFR1* fusion specifically from cDNA in the t(8;12) patient (E0003) but not the normal controls. The reciprocal fusion product was not detected; (B) Sequences surrounding the genomic breakpoints in *CPSF6* and *FGFR1* (bold).

3.4.5 Consequence of the t(8;12)

The t(8;12) fuses *CPSF6* intron 8 to *FGFR1* exon 10 and is the ninth reported *FGFR1* fusion gene. The *CPSF6-FGFR1* mRNA fusion is predicted to encode 5171 bp with an open reading frame (ORF) of 2688 bp including necessary intronic *CPSF6* sequence to

maintain the ORF. The chimaeric protein is predicted to encode an 896 amino acid protein of which 501 amino acids are derived from CPSF6 and 395 amino acids from FGFR1. The predicted molecular weight of the protein is 223.5 kDa. The chimaeric protein retains the N-terminal region and the RNA recognition motif (RRM) from CPSF6 fused to the entire tyrosine kinase domain of FGFR1 (Figure 3.32).

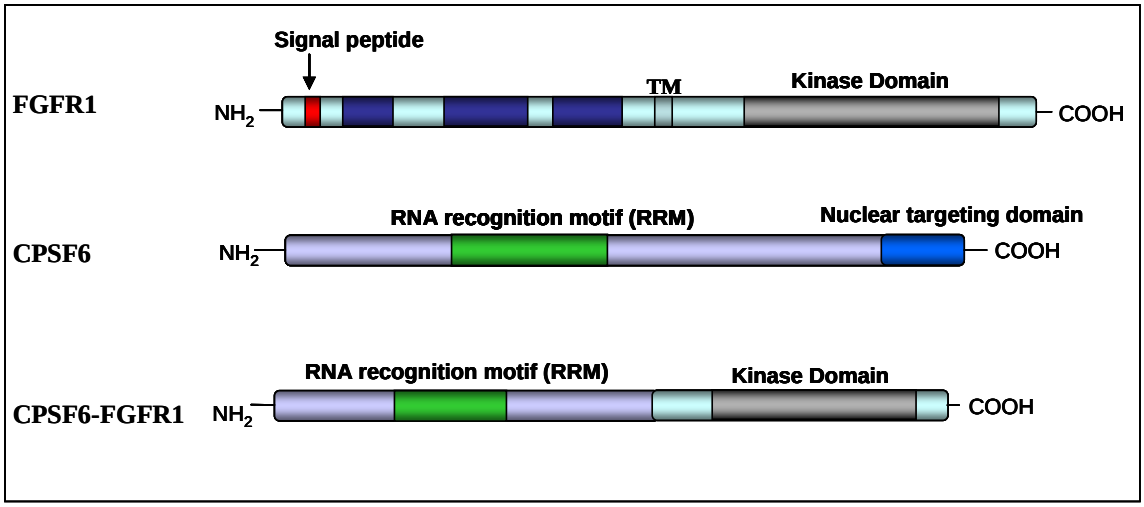


Figure 3. 22: Domain structure of FGFR1, CPSF6, and the predicted fusion protein. Relevant protein domains are detailed: RNA recognition motif (RRM), nuclear targeting domain for CPSF6 and the transmembrane domain, and tyrosine kinase domain for FGFR1.

3.4.6 Discussion

CPSF6 is the ninth partner gene found fused to *FGFR1*. All *FGFR1* fusions to date involve a breakpoint in exon 9 which disrupts the FRS2 binding domain (Hoch and Soriano 2006). No recognisable oligomerisation motifs are identifiable in CPSF6, however it has been demonstrated previously that RNA recognition motifs (RRM), such as the one retained in CPSF6-FGFR1, may mediate homodimerisation (Handa, *et al* 2006, Simpson, *et al* 2004) and thus may activate the fusion protein.

CPSF6 is a cleavage and polyadenylation specificity factor. It plays a major role in pre-mRNA processing and has a domain organisation (N-terminal RNA binding

motif, a proline-rich middle and a C-terminal alternating charge domain that is arginine rich) similar to the superfamily of arginine/serine (R/S) -rich splicing factors which are also involved in pre-mRNA splicing. These R/S splicesomal proteins are thought to bind pre-mRNA and tether other proteins to the RNA via these R/S-domains and commit the pre-mRNA to splicing. It has been suggested that CPSF6 could interact via its alternating charge domain with one or more R/S proteins or R/S related splicing factors (Caceres, *et al* 1997). In conjunction with this CPSF6 has also been found in purified splicesomes (Dettwiler, *et al* 2004, Rappsilber, *et al* 2002, Zhou, *et al* 2002). During pre-mRNA processing, CPSF6 binds the highly conserved polyadenylation site (AAUAAA) that is present in almost every mRNA precursor, as well as interacting with the poly-A binding protein II (PABII) that binds the growing poly-A tail. Together CPSF6 and PABII activate poly (A) polymerase (PAP), by directing the enzyme to the RNA allowing polyadenylation of the upstream cleavage product (Ruegsegger, *et al* 1998).

Interestingly, SFPQ and NonO are also found purified in splicosomes along with CPSF6 (Dettwiler, *et al* 2004, Rappsilber, *et al* 2002, Zhou, *et al* 2002). SFPQ has recently been described as a partner protein found fused to the Abl tyrosine kinase in a patient with B-cell progenitor acute lymphoblastic leukaemia (ALL) (Hidalgo-Curtis, *et al* 2008). It is required early in splicesome formation and is a pre-mRNA splicing factor that binds the poly-pyrimidine tract of mammalian introns. The localisation of both CPSF6 and SFPQ suggests a possible interrelated function in transcription and processing of pre-mRNAs of these proteins. Of further interest is another FGFR1 fusion partner ZMYM2 (ZNF198) (Xiao, *et al* 1998), that immunoprecipitates with SFPQ (Kasyapa, *et al* 2005). ZMYM2 is thought to act as a possible scaffold protein bringing RNA-binding proteins together to facilitate pre-mRNA splicing (Kasyapa, *et al* 2005). In addition, FIP1L1, which fuses to PDGFRα in chronic eosinophilic leukemia (Cools, *et al* 2003), also plays a role in polyadenylation and pre-mRNA splicing (Kaufmann, *et al* 2004, Preker, *et al* 1995). All these interconnected partner proteins found fused to tyrosine kinases in human leukaemia are involved in pre-mRNA processing, the precise reasons for the recurrent

fusion to FGFR1 of this sub-group of interacting partner proteins involved with transcription elongation is still unclear (Hidalgo-Curtis, *et al* 2008).

3.4 Discussion

In this chapter I describe the discovery of four new partner proteins found fused to three different receptor tyrosine kinases in patients diagnosed with a MPD. All four novel fusions represent rare chromosomal abnormalities with only one of these fusions being recurrent (*GOLGA4-PDGFRB*).

The identification of rare fusions involving tyrosine kinase genes is clearly important as they are likely to be responsive to targeted therapy with imatinib or other compounds. The functional role of the partner protein is often overlooked, as well as providing dimerisation motifs accumulating evidence suggests that these partner proteins may play additional roles modifying intracellular signalling and disease phenotype (Million, *et al* 2004).

Groups of partner proteins with similar functions and interactions appear to fuse to tyrosine kinases more frequently than expected. Examples include proteins involved in pre-mRNA processing (Hidalgo-Curtis, *et al* 2008), endocytosis (this thesis) and a number that are targeted to the centrosome (Delaval, *et al* 2005b). Partner proteins provide both oligomerisation and targeting motifs for the fusion protein. The *FIG-ROS* fusion found in glioblastoma and the *FGFR1OP-FGFR1* fusion found in patients with 8p11 myeloproliferative syndrome (EMS) are directed by the partner protein to specific sub-cellular compartments, including the centrosome and the Golgi (Charest, *et al* 2003, Delaval, *et al* 2005b). In the case of FIG-ROS, this is thought to be essential for the transforming ability of the fusion.

Delaval et. al., (2005) describe the targeting of *FGFR1OP-FGFR1*, an oncogenic tyrosine kinase, to the centrosome thus activating signalling pathways at this organelle allowing cells to sustain continuous cell cycles. This group have also

identified a second *FGFR1* partner gene localised to the centrosome, and suggest that the fusion partners may not only provide dimerisation domains but also target oncogenic kinases to a specific cellular locations necessary for transformation. They hypothesise that these proteins may exert their oncogenic activity through dysregulation of cell processes associated with the centrosome. Thus further research into the subcellular localisation of oncoproteins indicates that emerging links between partner protein function and localisation may not all be coincidental. Numerous studies have been performed that look specifically at the subcellular localisation of constitutively activated kinases, Wetzler et. al., (1993) investigated the BCR-ABL protein in leukaemic cells and found the fusion resides exclusively in the cytoplasm, unlike the wild type ABL protein that is only found in the nucleus (Wetzler*, et al* 1993). In addition, Ollendorf et. al., (1999) whom further characterised the FIM-FGFR1 fusion product as a result of the t(8;13) translocation, studied the subcellular localisation of this fusion protein, and in comparison to wild type FIM, they found that it was also entirely cytoplasmic (Ollendorff*, et al* 1999). It is thought that localisation of these fusion proteins to the cytosol is enough to abrogate their entry into the endocytic pathway, and hence prevent lysosomal degradation.

Functionally, the largest related group of partner proteins is those linked to intracellular trafficking and endocytosis. There is increasing evidence of a relationship emerging between endocytosis and leukaemia. Removal of specific receptors from the cell surface by endocytosis and their degradation in lysosomes is an established method for the attenuation of receptor signalling. Genes encoding proteins associated with the endocytic pathway are repeatedly being identified as targets of chromosomal rearrangements in human haematological malignancies. The identification of *GOLGA4-PDGFRB*, adds further weight to this emerging relationship. It is possible that subversion of endocytic control, possibly caused by such oncogenic fusion proteins, could play an important role in hyperproliferative conditions.

Further evidence of the role of endocytic pathways in haematological malignancies is becoming increasingly apparent. At least three reported translocations involving *PDGFRB*, identified in patients with aCML/CMML (Table 3.2), make in-frame fusions with components of the endocytic machinery. These include *Rabaptin-5*, *Hip1* and *PDE4DIP*. Rab5 in its active GTP-bound form recruits a number of effectors to the endosome membranes. Rabaptin-5 is proposed to function as molecular linker between Rab5 and Rab4 to coordinate endocytic and recycling traffic (Magnusson*, et al* 2001, Vitale*, et al* 1998). Zhu et. al., (2004) has also recently shown through biochemical and functional analysis that the Rab5-Rabaptin-5 complex abrogates endosome fusion if disrupted (Zhu*, et al* 2004). The *Rabaptin-5-PDGFRB* fusion results in a loss of the Rab5 binding site and thus might obstruct formation of the Rab5-Rabaptin-5 complex, consequently halting endosome fusion.

A second relevant fusion is *Hip1-PDGFRB* (Ross*, et al* 1998). Huntington-interacting protein 1 (Hip1) was the first component of the endocytic machinery to be directly implicated in tumour formation. It interacts with AP-2, phosphoinositides and clathrin and is thought to play a role in clathrin-mediated endocytosis. Over-expression of Hip1 causes cell transformation and increased cell proliferation and the transformed cells also appear to have less clathrin at the plasma membrane. Hyun et. al., (2004) also found that over-expression of *Hip1* inhibited the degradation of ligand-stimulated growth factor receptors. Transient transfection of 293T cells with *Hip1* stabilised pools of *PDGFRB* following ligand-induced endocytosis, thus it is proposed that *Hip1* may stabilise growth factor receptor levels via altered intracellular trafficking. Hip1 also contains an ENTH (epsin N-terminal homology) domain, that like other ENTH domains binds predominantly to phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P_2), and is important in regulating clathrin-mediated endocytosis. Hyun et. al., (2004) suggest that the ENTH domain may be the factor involved in regulating growth factor endocytosis and thus could be critical in the protein's ability to promote cell growth or survival (Hyun and Ross 2004). This has been further supported by work done using a mutant construct lacking the ENTH domain, which when transfected into cells induced apoptosis. The lipid binding

Protein	Proposed function in receptor downregulation	Cancer type	Chromosomal translocation	References
Eps15	Ubiquitin-binding protein found in PM clathrin-coated pits and in clathrin coats in sorting endosomes. Possible role in receptor recruitment.	Myeloid leukaemia	ALL/HRX gene; t(1;11)(p32;q23)	(Floyd and De 1998)
Endophilin-2	Putative lysophosphatidic acid acyl transferase. Binds dynamin-2 and synaptojanin. Involved in formation of clathrin-coated vesicles. Member of the membrane-bending/GTPase BAR domain	Myeloid leukaemia	ALL/HRX gene; t(11;19)(q23;p13)	(Floyd and De 1998)
CALM	Non-neuronal homologue of AP-180. Binds clathrin and phosphoinositides. Strong activator of clathrin polymerisation.	Myeloid leukaemia	AF-10; t(10;11)(p13;q14)	(Floyd and De 1998)
Rabaptin-5	Effector of Rab5. Complexed to the Rab5 GDP/GTP exchange factor Rabex5. Required for endosomal fusion.	CMML	PDGFRβ; t(5;17)(q33;p13)	(Magnusson, *et al* 2002)
Hip1	Homologue of End4/Sla2, which interacts with actin-binding proteins and is required for endocytosis in *Saccharomyces cerevisae*. Role in clathrin-mediated endocytosis. Interacts with AP-2, clathrin and phosphoinositides	CMML	PDGFRβ; t(5;7)(q33;q11.2)	(Ross, *et al* 1998)
PDE4DIP/Myomegalin	Functions as an anchor to localise components of the cAMP-dependent pathway to the Golgi/centrosome. The cAMP-dependent pathway is thought to be essential for intracellular signalling.	Eosinophilic CMPD	PDGFRβ; t(1;5)(q23;q33)	(Wilkinson, *et al* 2003)
BIN2	Member of the BAR domain proteins. BAR domain drives membrane curvature, implicating a role in vesicular trafficking. BAR also shares significant homology to GTPase binding domain of Arfaptin 2, an effector of Rho-and Arf-like GTPases. Arfaptin 2 is known to be involved in intracellular transport, especially endocytosis.	Eosinophilic CMPD	PDGFRβ; t(5;12)(q22.1;q33)	Hidalgo-Curtis et. al., SUBMITTED
WDR48	WD-repeat endosomal protein.	Eosinophilic CMPD	PDGFRβ; t(3;5)(p22.2;q33)	Hidalgo-Curtis et. al., SUBMITTED
Golgin-245	Golgin family of coiled-coil proteins, associated with the Golgi apparatus. Plays a role in tethering events in membrane fusion and as structural supports for Golgi cisternae (subcompartments of the TGN). May also function as a vesicle coat.	aCML	PDGFRβ; t(3;5)(p22;q33)	Hidalgo-Curtis et. al., SUBMITTED
KIF5B	Important for microtubule-based endosome-motility, by regulating early endosome trafficking. If inhibited reduced endosome motility and vesicle fission is observed.	Eosinophilic CMPD	PDGFRα complex karyotype	(Score, *et al* 2006)
FIG	Encodes a protein that peripherally associates with the golgi apparatus, and likely plays a role in golgi function.	Glioblastoma	del 6(q21)	(Charest, *et al* 2003)

Table 3.2: Components of endocytic machinery found as partners of oncogenic fusion proteins.

characteristics of the ENTH domains were also analysed and unusually found to preferentially bind 3-phosphoinositides not (PtdIns(4,5)P_2). Thus, it is suggested that *Hip1* acts downstream of receptor uptake to inhibit trafficking of receptor tyrosine kinases to the lysosome for subsequent degradation. It is speculated that this occurs by binding inositol lipid determinants of the early endosome by their ENTH domains, acting to either stabilise the early endosome, favouring continued signalling, or by stimulating growth factor recycling to the surface. This work certainly provides evidence of a strong link between a partner gene of *PDGFRB* and the targeting of the endocytic pathway.

Other examples of oncoproteins involving members of the endocytic pathway include *AF-1p* (Bernard*, et al* 1994) and *EEN* (So*, et al* 1997) both of which are partners found fused to the C-terminal region of the *MLL* gene. In addition, endocytic proteins fused to RTKs in other malignancies have been reported, in particular papillary thyroid carcinomas. Rabes et. al., (2000) observed a fusion between the *RET* receptor tyrosine kinase and *GOLGA5* a golgi integral membrane protein (Bascom*, et al* 1999, Rabes*, et al* 2000). Further evidence to support the hypothesis that subversion of endocytic pathways may be important in leukaemogenesis will be described in Chapter 5.

Deregulated, constitutively activated tyrosine kinases are proving to be the main driving force behind the MPDs, as substantiated in this chapter with the identification of four new tyrosine kinase fusion genes: *GOLGA4-PDGFRB, STRN-PDGFRA, ETV6-PDGFRA* and *CPSF6-FGFR1.* Moreover, there is now accumulating evidence to suggest an additional role for the partner proteins in modifying intracellular signalling and disease phenotype. Functionally similar groups of partner proteins linked to the centrosome (Delaval*, et al* 2005a), transcription elongation (Hidalgo-Curtis*, et al* 2008) and intracellular trafficking (Hidalgo-Curtis et. al., submitted 2009) have been identified that are thought to contribute not only the dimerisation motif but may impair endocytosis and even direct the oncoprotein to specific intracellular compartments, all of which may be critical to their transforming ability.

4 IDENTIFICATION OF DEREGULATED EXPRESSION OF TYROSINE KINASE GENES

4.1 Introduction

There is increasing evidence to suggest a role for deregulated tyrosine kinases in the pathogenesis of MPDs. It is particularly important to recognise patients with *PDGFRA* or *PDGFRB* fusions as these cases are likely to be responsive to imatinib (Apperley, *et al* 2002, Cools, *et al* 2003a, Gotlib 2005, Jones and Cross 2004, Klion, *et al* 2004, Pardanani, *et al* 2003). Identification of patients who are positive for *PDGFR* fusions is therefore critical to appropriate clinical management. However these genes have been shown to fuse to diverse partners in MPDs and it seems likely that additional partner genes are yet to be discovered. Furthermore many of these cases are very rare and the breakpoint diversity of many fusions is unknown. Consequently it is difficult to screen for all *PDGFR* fusions.

Our research group has studied patients for *in vitro* sensitivity to imatinib, as mentioned above. A number of patients have been found that appear to be responsive to imatinib but are negative for the known common fusions (e.g. *FIP1L1-PDGFRA* and *ETV6-PDGFRB*). This leaves a number of patients where the molecular lesion has not been characterised, but we suspect involves an imatinib responsive gene. Cytogenetic analysis may suggest that either *PDGFRA* or *PDGFRB* might be involved, but some fusions are cytogenetically cryptic (e.g. *FIP1L1-PDGFRA*) and several translocation targets have been identified on chromosome 5q that are distinct from *PDGFRB,* for example *FLT4* fusions found in a minority of MDS patients (Armstrong, *et al* 1993). Thus cytogenetics alone cannot be used to determine unambiguously whether a patient might be sensitive to imatinib or not. FISH may be used to look for rearrangements of specific genes, but these results may be misleading due to cryptic or complex rearrangements (Kulkarni, *et al* 2000). Consequently there

is a need for new methods to detect *PDGFRA* and *PDGFRB* fusion genes. Here I describe the design and application of two generic RT-PCR assays to detect over-expression of *PDGFR* genes, and an mRNA-based Multiplex Ligation-dependent Probe Amplification (MLPA) test. The main aim was to search for new cryptic gene fusions that might be responsive to imatinib.

4.1.1 PDGFRB rearrangements

Disruption of *PDGFRB* was first described by Golub et. al., (1994), as a consequence of a t(5;12)(q33;p13), fusing the 5'end of *ETV6* (otherwise known as *TEL*) to the 3'end, encoding the entire tyrosine kinase domain, of *PDGFRB* (Golub*, et al* 1994). Subsequently, twenty other translocations involving *PDGFRB* have been reported (See Table 1.1, Section 1.7.3) including three identified by our laboratory unpublished at present. The breakpoint for twenty out of the twenty one reported oncogenic fusions is within intron 10 of *PDGFRB*. To date, *NIN-PDGFRB* is the only reported fusion gene that exhibits a different genomic breakpoint (Vizmanos*, et al* 2004). The partner genes fuse in frame to the entire *PDGFRB* tyrosine kinase domain. Each partner protein contains an oligomerisation motif that dimerises the receptor tyrosine kinase, leading to aberrant phosphorylation of downstream proteins promoting growth and resistance to apoptosis.

Under normal conditions *PDGFRB* mRNA expression is tightly regulated, but when fused to the partner gene expression of the tyrosine kinase portion is under the control of the partner gene promoter and so may be over-expressed compared to the normal, full length *PDGFRB* mRNA. I developed a PCR-based assay designed to detect the expression of the tyrosine kinase domain of *PDGFRB* compared to a control band spanning the exons upstream and downstream of the common reported breakpoints.

4.1.2 PDGFRA rearrangements

Disruption of *PDGFRA* was first described by Baxter et. al., (2002), as a consequence of a t(4;22)(q12;q11), fusing the 5'end of *BCR* to the 3'end encoding the entire tyrosine kinase domain of *PDGFRA*. This is a rare variant translocation identified initially in only two patients with a CML-like MPD (Baxter*, et al* 2002). Five further rearrangements of *PDGFRA* have been identified and the breakpoints within *PDGFRA* have been found to be very tightly clustered in all cases, resulting in a truncated *PDGFRA* exon 12 (Baxter*, et al* 2002, Cools*, et al* 2003a, Curtis*, et al* 2007, Score*, et al* 2006, Walz*, et al* 2007). The most common *PDGFRA* rearrangement is the fusion of *FIP1L1* to *PDGFRA* in HES patients described by Cools et. al., (2003). The fusion is generated by an interstitial deletion on chromosome 4q12. This fusion is cytogenetically cryptic but can be detected by FISH or RT-PCR (Cools*, et al* 2003a, Pardanani*, et al* 2003). Detection of this fusion is also complicated by the complex alternative splicing and variability of genomic breakpoints in *FIP1LI*, in addition to the low expression of the fusion gene in many cases and the significant variability in the proportion of cells involved in the malignant clone (Chung*, et al* 2006, Cools*, et al* 2003a). I designed a second assay in the same manner as that for *PDGFRB*, but for *PDGFRA* in an attempt to detect novel cryptic rearrangements. It was also hoped that this assay might address the limitation of current procedures in detecting *FIP1L1-PDGFRA*. I designed a second simple multiplex RT-PCR assay to determine the relative expression of the *PDGFRA* kinase domain which is retained in all fusions. This was compared to expression of the intact receptor, which is usually not detectable or weakly detectable by standard RT-PCR from peripheral blood or bone marrow leukocytes. Samples with a possible *PDGFRA* fusion gene were expected to over-express the kinase domain relative to the normal receptor mRNA.

4.2 Methods

4.2.1 Design of the PDGFRB RT-PCR assay

Three primers were designed, two that flank the region where the common breakpoints are reported (PDB Ex9/10.F and PDB Ex15.1R), and a third after the exon in which all the breakpoints are reported (PDBEx13.1F) (Figure 4.1). Primers were designed to anneal at 66°C and the PCR was undertaken in a volume of 10µl, due to limited cDNA available for each patient. The number of patients eligible for screening prompted me to perform all the amplifications in plate format or in 0.2ml strips tubes in a MJ Research Inc. thermocycler (Tetrad). After optimisation, the PCR protocol was finalised to: 0.4µM *PDGFRB* exon 9/10.F, 0.1µM *PDGFRB* exon 13.1F, 0.5µM *PDGFRB* exon 15.1R, 0.2mM dNTP, 1x buffer (Hotstar buffer containing 15mM $MgCl_2$), 2.5mM $MgCl_2$ (total), 0.4U Hotstart *Taq* DNA polymerase (Qiagen, West Sussex, UK). The PCR cycling parameters use a 15 minute 95°C hot-start to activate the enzyme and a 30 second, 66°C annealing temperature for stringency. The PCR was cycled 30 times.

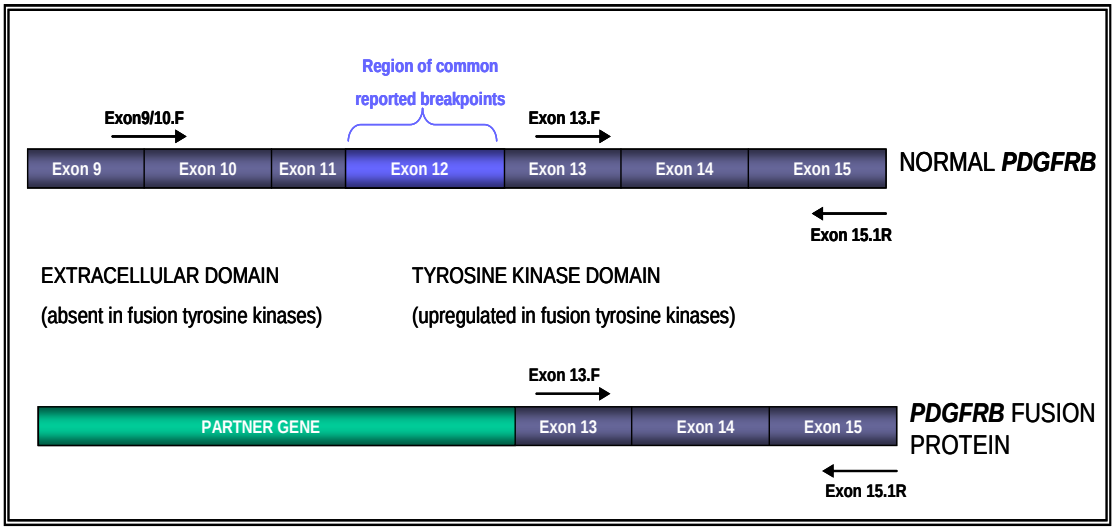


Figure 4. 1: A diagram representing the position of each of the primers in the multiplex, in relation to the common reported breakpoints in *PDGRFB* and to the tyrosine kinase and extracellular domains of *PDGFRB*.

A patient with normal *PDGFRB* expression is expected to generate two bands of equal intensity (assuming both PCRs worked with the same efficiency), and a patient with a rearrangement that results in over-expression of the region encoding the tyrosine kinase would exhibit a strong smaller band derived from both the normal receptor mRNA and the fusion mRNA, and a faint larger band derived solely from the normal receptor mRNA (Figure 4.2A).

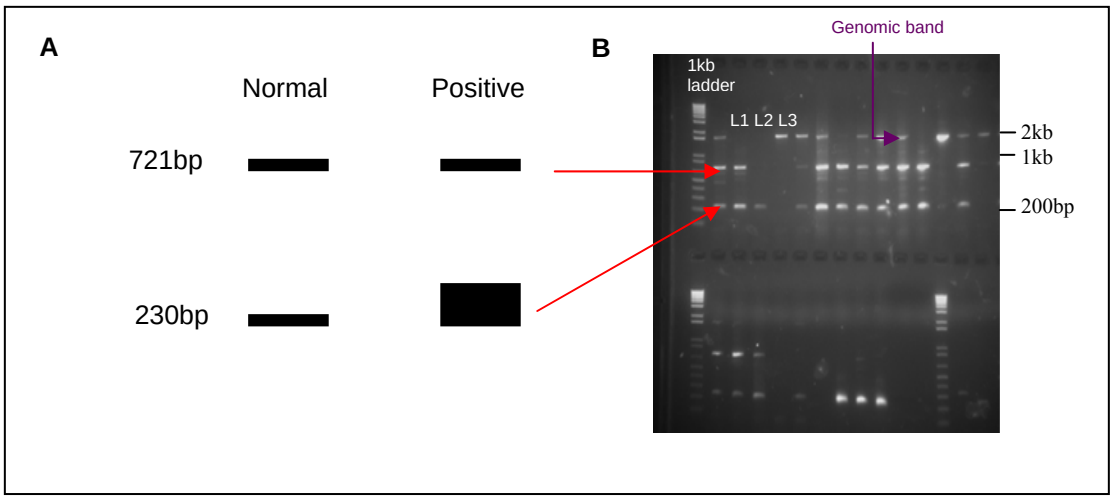


Figure 4. 2: *PDGFRB* RT-PCR assay, a schematic of the expected banding pattern. (A) The banding pattern expected from initial design of the *PDGFRB* multiplex, (B) Gel picture of observed bands from a number of patients and controls screened using the *PDGFRB* multiplex.

Figure 4.2B represents an example of a gel of the *PDGFRB* RT-PCR assay. Lane 1 (L1) represents a normal individual with an additional band visible at 2kb. I cloned and sequenced this aberrant band and found it contained intronic sequence, and was therefore a genomic DNA contaminant. Residual genomic DNA is often found in the cDNA. I altered the PCR conditions by reducing the amount of the 5' primer in the extracellular domain but was unsuccessful at eliminating this band. I also redesigned

the primer, finally resorting to the Ex9/10F primer. This primer was designed to sit across the exon boundaries of exon 9 and 10 to prevent amplification of any genomic DNA. Although this primer helped to reduce the intensity of this contaminating band it was not successful at eradicating it completely. With further work I showed the assay was still informative and that this band did not interfere with interpretation of the results, and can therefore be disregarded. Lane 2 (L2) represents my normal control, for which I cloned each of the fragments and combined them equally to create an artificial mixture. In this lane as expected, the 2kb genomic band is absent. Finally, lane 3 (L3) is my positive control, created again through cloning, from a patient with a novel *PDGFRB* rearrangement. The positive control shows only the bottom band, representing the tyrosine kinase portion of the gene. Some patients showed this same pattern, with the absence of the larger control band, others showed the differential expression expected (Lanes 1 and 4) (Figure 4.3).

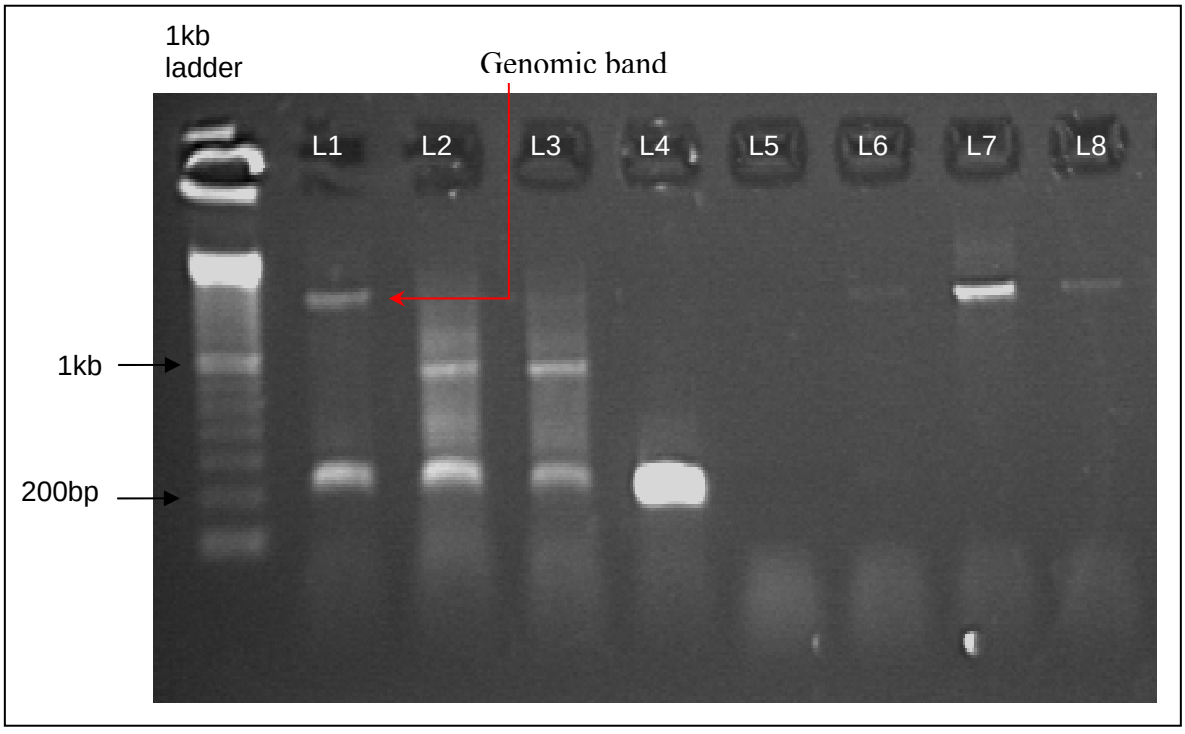


Figure 4. 3: A representative agarose gel picture of patients screened by the *PDGFRB* multiplex. Lane 1 (L1) represents a positive patient where the tyrosine kinase band is present only; Lanes 2 and 3 (L2 and L3) are negative patients where both bands are clearly visible; Lane 4 (L4) is the positive control created from cloning of a patient with a

known *PDGFRB* fusion. Lane 5 (L5) shows the no DNA control, and Lanes 6, 7 and 8 (L6, L7, and L8) represent patients where the genomic band has out-competed the presence of the other bands and were classed as fails.

4.2.2 Design of the PDGFRA RT-PCR assay

Three primers were again designed, two that flank the region where the common breakpoints have been reported (PDA Ex10.1F and PDA Ex15.1R). The majority of breakpoints in *PDGFRA* have been reported in exon 12, therefore the third primer, designed to prime after the region of common breakpoints was located in exon 13 (PDA Ex13.1F) (Figure 4.4A). As before, all primers were designed to anneal at 66°C, the PCR volume was 10µl and the amplifications were in plate or strip-tube format, for 30 cycles, in a MJ Research Inc. thermocycler (Tetrad). After optimisation the PCR protocol was finalised to: 0.05µM *PDGFRA* exon 10.1F, 0.2µM *PDGFRA* exon 13.2F, 0.4µM *PDGFRA* exon 15.1R, 0.2µM dNTP, 1x buffer (Hotstar buffer containing 15mM $MgCl_2$), 2.5mM $MgCl_2$ (total), 0.4U Hotstart *Taq* DNA polymerase (Qiagen, West Sussex, UK). The PCR cycling parameter again uses a 15 minute 95°C hot-start to activate the enzyme and a 30 second, 66°C annealing temperature for stringency.

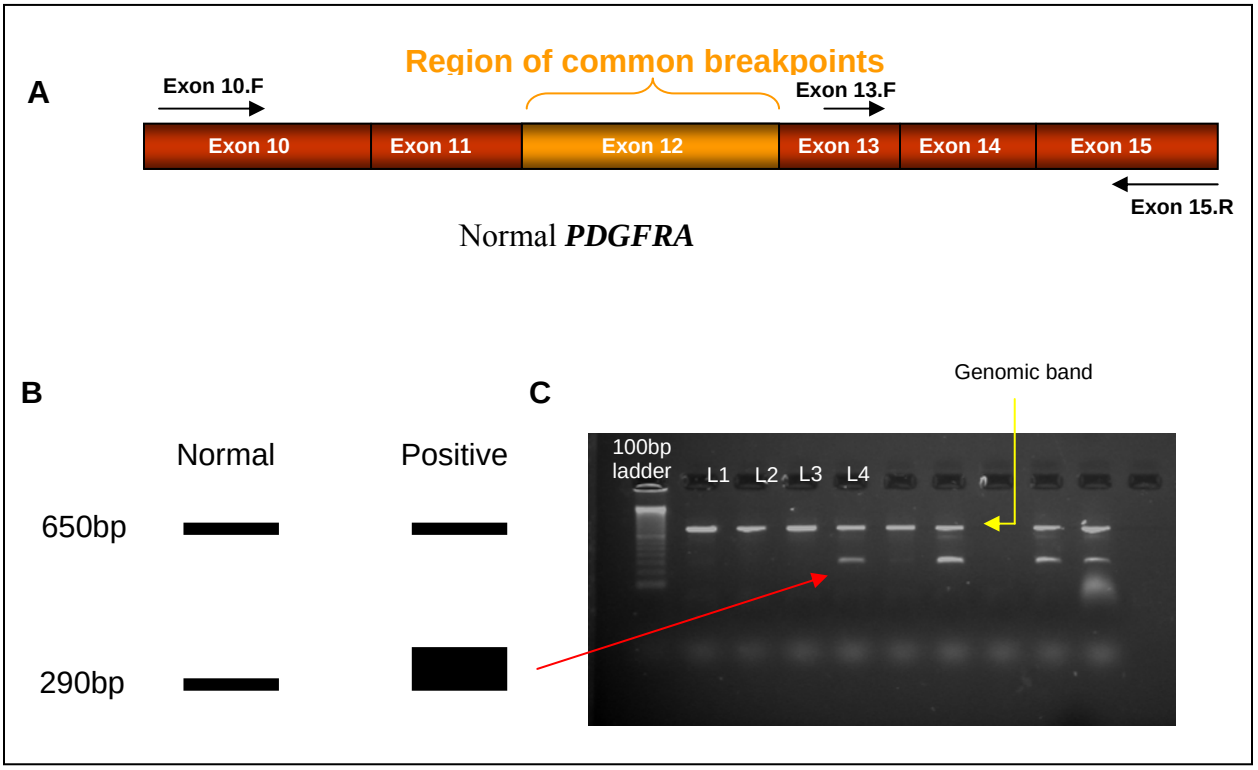


Figure 4. 4: *PDGFRA* RT-PCR assay. (A) A schematic representing the position of each primer in the multiplex in relation to the common reported breakpoints in *PDGRFA*, (B) The banding pattern expected from initial design of the *PDGFRA* multiplex, (C) Gel picture of observed bands from a number of patients screened using the *PDGFRA* multiplex.

A patient with normal *PDGFRA* expression is expected to yield two bands of equal intensity and a patient with a rearrangement that results in over-expression of the tyrosine kinase would exhibit a strong smaller band and a faint larger band as before (Figure 4.4B). Lanes 1-3 show normal individuals, however only one band was present, and initially presumed to be the larger control band. On analysis of this band, it was clear the product was not the size expected from the primer pair. Upon cloning and sequencing I found that this band was again due to genomic DNA contamination. Alternate primers and differing PCR conditions failed to yield a correct sized control band with no intron, or the expression of both mRNA-derived bands in a normal individual. This is almost certainly because *PDGFRA* is not normally expressed or

expressed at very low levels in peripheral blood leukocytes. I therefore ordered a cell line, MG-63 that has been shown to express normal *PDGFRA*. MG-63 cDNA produced, as expected, two bands of 650 and 290 bp (Figure 4.5). Although the lower band is stronger than the upper band, both are clearly visible. Eol-1 cells express *FIP1L1-PDGFRA* and only the smaller 290 bp band derived from the region retained in the fusion is visible (Figure 4.5). Most normal individuals didn't show either of the *PDGFRA* mRNA-derived bands, but instead the 1049 bp genomic DNA fragment encompassing exons 13–15 was amplified (Figure 4.5). Although an unexpected product, this band served as a useful internal control for successful amplification. On screening positive controls I also discovered that although differential expression was not observed, the smaller band (290bp) was visible only in patients with confirmed *PDGFRA* rearrangements (Figure 4.4C, Lane 4). Thus, the RT-PCR assay allowed identification of over-expression of the tyrosine kinase domain of *PDGFRA*.

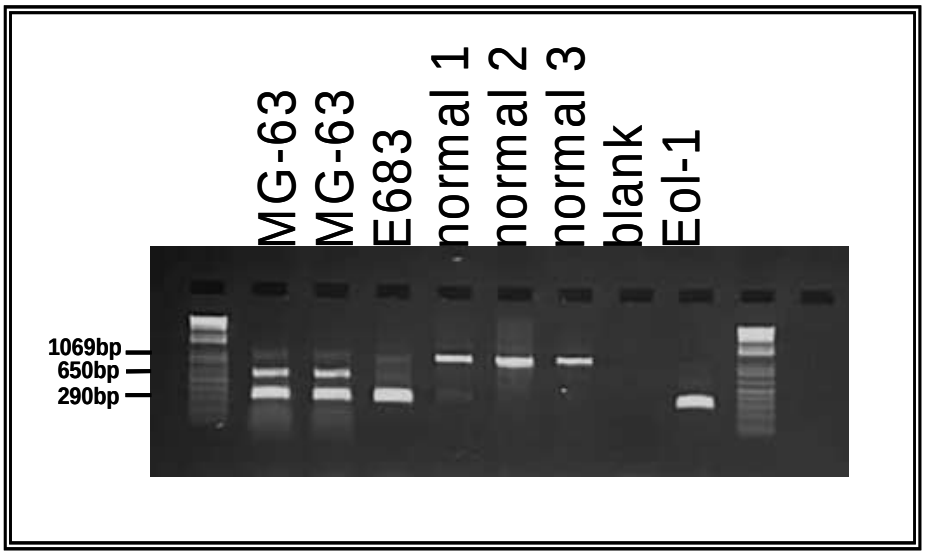


Figure 4. 5: MG-63 cells express normal *PDGFRA* and both 650 bp and 290 bp bands are visible. Eol-1 cells express *FIP1L1-PDGFRA* but not normal *PDGFRA* and only the smaller 290 bp band is visible. Normal individuals 1–3 express *PDGFRA* very weakly and the predominant amplification product (1049 bp) is exons 13–15 from genomic DNA. Patient E683 shows the 290 bp band but not the 650 bp band, suggesting the presence of a *PDGFRA* fusion gene (discussed in more detail later).

4.3 Results

4.3.1 Identification of patients with possible PDGFRB rearrangements

Having optimised the *PDGFR* RT-PCR assays I went on to screen a cohort of 199 MPD patients for over-expression of *PDGFRB*. Of these 18 were identified that were regarded as possibly positive for a *PDGFRB* rearrangement on the basis of increased intensity of the smaller band. Of these 18, 9 were strongly positive with markedly increased relative intensity (Table 4.1) and are considered in detail below.

Identifier	**Diagnosis/Referral Reason**	**Previously identified fusion gene**
P936	EMS	*ZNF198-FGFR1*
P981	MPD? 8p11	*FGFR1OP-FGFR1*
P997	MPD t(5;10)	*CCDC6-PDGFRB*
P1218	MPD t(5;12)	*ETV6-PDGFRB*
E285	MDS transformed t(5;12)	*ETV6-PDGFRB*
E287	HES	No
E354	Eosinophilia	No
E508	?HES/SM	No
E521	MPD	No

Table 4. 1: Patients observed to have *PDGFRB* over-expression using the *PDGFRB* multiplex RT-PCR assay.

The samples were screened blind and the multiplex was able to identify all three patients with previously characterised *PDGFRB* fusion genes (P997, P1218 and E285), of which two had been found to harbour *ETV6-PDGFRB* and the third *CCDC6-PDGFRB* Unexpectedly two patients with over-expression of *PDGFRB* had previously been found to have rearrangements of *FGFR1* (P936 and P981). Finally,

four patients (E287, E354, E508 and E521) had been previously tested negative by RT-PCR for *ETV6-PDGFRB* and were investigated in more detail.

My initial analysis focused on patient E521 as this case had the most stored material. I performed FISH and 5'RACE and discovered the patient had a typical *ETV6-PDGFRB* rearrangement, despite the previous RT-PCR negative result. Only after I had identified the *ETV6-PDGFRB* fusion in this patient did we receive notification of the patient's karyotype. The fusion is a consequence of a complex rearrangement involving chromosomes 2, 5 and 12 (Figure 4.6). It is not obvious from the karyotype that this fusion might exist but the involvement of 5q33 is suggestive of a *PDGFRB* rearrangement. After this discovery the original diagnostic RT-PCR was repeated and the patient sample was found to be positive. Why this fusion was missed initially is unclear however, he had been treated and responded to imatinib. Of the remaining three patients, E508 and E354 were studied extensively performing *PDGFRB* split apart FISH and 5'RACE to detect any possible *PDGFRB* rearrangements. No evidence for disruption of this gene was found. Very little material was available for study of the fourth case (E287).

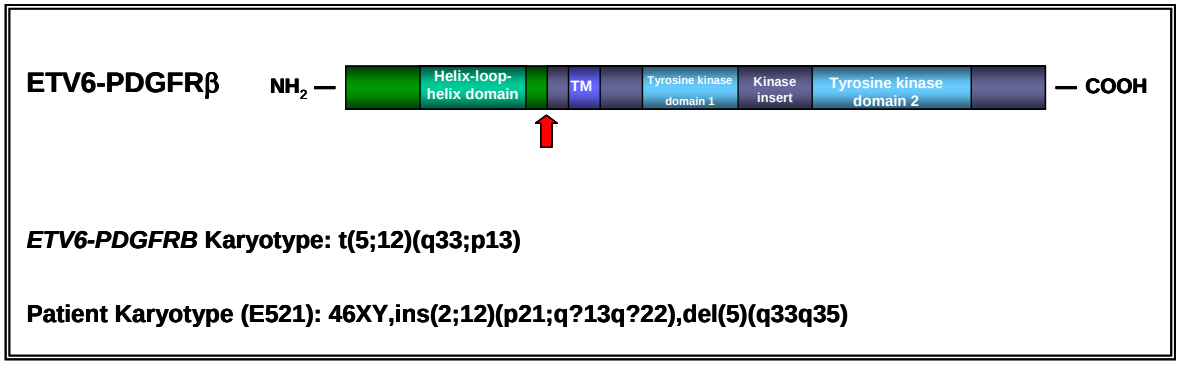


Figure 4. 6: Structure of the ETV6-PDGFRβ fusion with the usual karyotype associated with this fusion and the complex karyotype observed in the bone marrow of patient E521 by conventional cytogenetics

Gain-of-function point mutations have also been discovered in certain receptor tyrosine kinases that act to constitutively activate the kinase, such as the well

characterised V617F change in *JAK2* associated with PV and other MPDs. Additionally, point mutations and small deletions in the activation loop and juxtamembrane domains of *PDGFRA* and *KIT* have been commonly reported in patients with gastrointestinal stromal tumours (Heinrich*, et al* 2003, Hirota*, et al* 1998) and more recently Chiara et. al, (2004) reported the identification of a mutation in the activation loop of the murine *PDGFRB* gene. The mutation conferred transforming ability to the embryonic fibroblasts (Chiara*, et al* 2004). Based on this, I screened two patients with available material over-expressing *PDGFRB* using denaturing High Performance Liquid Chromatography (dHPLC) and sequencing. I began by screening the key regions of *PDGFRB*, those encoding the juxtamembrane domain and the activation loop (exons 12-18). For both patients each exon was individually amplified and the PCR product mixed in a 1:1 ratio with normal placental DNA PCR product for each individual exon. The samples were then heated to 95°C and slowly allowed to reanneal on a reducing temperature gradient allowing the formation of heteroduplexes. When run through a chromatography column under partially denaturing conditions changes in sequence usually result in extra peaks or shifted peaks compared to normal controls. No activating mutations were identified in either patient in the juxtamembrane or activation loop of *PDGFRB.* I screened all the remaining exons of *PDGFRB* to exclude the possibility of a mutation in these regions (exons 1-11 and exons 19-23). Both patients lacked any sequence variants.

A colleague (Kathy Waghorn) had also designed a DNA Multiplex ligation-specific probe amplification (MLPA) assay for selected tyrosine kinase genes. MLPA is a probe based technique that allows multiple targets to be amplified with only a single primer pair (for more detail see Figure 4.8). Amplification of the probe is dependent upon the number of target sites present in the sample, permitting detection of genomic deletions and amplifications. Each probe is fluorescently labelled and has a unique length allowing multiple amplicons to be analysed in one reaction and separated by size by capillary electrophoresis. Comparison of the peak pattern obtained for a given sample against reference samples allows the relative quantity of each amplicon to be determined. The ratio obtained identifies the amount of target

sequence present in the sample DNA. The MLPA assay designed by Kathy included the *PDGFRB* gene, and was used to look for possible duplications and deletions of specific tyrosine kinase genes. Following the *PDGRFB* over-expression RT-PCR result, E508 (only patient with available DNA) was analysed using MLPA. No deletions or duplications were observed (Figure 4.7).

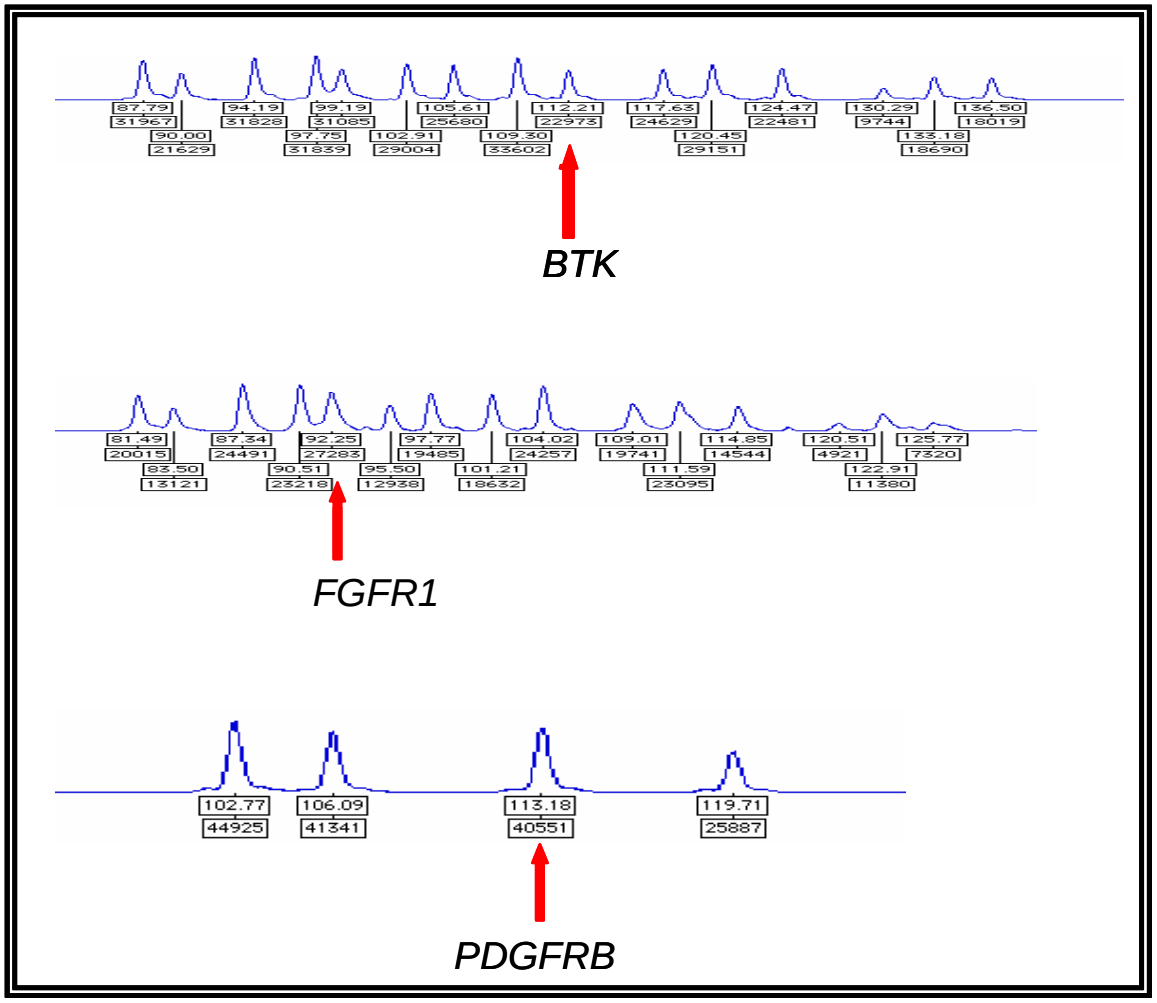


Figure 4. 7: MLPA traces of patient E508. MLPA traces showing (A) Normal control displaying full set of candidate tyrosine kinase genes in order: *PDGFRA, MET, PDGFRB, RET, FGFR1, SYK, FLT3, ROS, BTK, RRP4, MSTA, TRKB, EPHB4, MPL* and *FGFR3* (Note: *BTK* is an X-linked gene used as control), (B) patient with *FGFR1* duplicated proving its discriminatory power, (C) E508, set of four genes: *EPHB4,* MPL, PDGFRB and MET, used when looking for rearrangements of a specific gene.

4.3.2 Identification of patients with possible PDGFRA rearrangements

A total of 270 cases with IHES or persistent unexplained eosinophilia were referred to the Wessex Regional Genetics Laboratory, Salisbury for routine molecular analysis between May 2002 and November 2005. Of these, 229 patients were screened using the *PDGFRA* multiplex. The patients were simultaneously screened for the more common *FIP1L1-PDGFRA* fusion using single and nested RT-PCR. All 12 cases in this study group who tested *FIP1L1-PDGFRA* positive by RT-PCR were also positive by the *PDGFRA* assay, that is, the 290bp band was clearly visible and the 650bp band was either not visible or was barely visible (Table 4.2), confirming the multiplex worked to identify rearrangements of *PDGFR*A. There were no cases where a *FIP1L1-PDGFRA* fusion determined by standard RT-PCR was not identified by my *PDGFRA* expression assay. Interestingly one patient (E293) turned out to be a sample from a *FIP1L1-PDGFRA* positive case that had been treated with imatinib and was therefore referred for minimal residual disease analysis**.** Both single and nested RT-PCR of *FIP1L1-PDGFRA* proved negative, however my assay observed over-expression of *PDGFRA*. This is unusual, as it is expected that the nested RT-PCR will be more sensitive than the *PDGFRA* assay. However, analysis of *FIP1L1-PDGFRA* breakpoints by a colleague has highlighted breakpoints in regions downstream of the common *FIP1L1* breakpoint. Due to the position of the nested RT-PCR primers within *PDGFR*A, nested RT-PCR using standard primer sets would not be able to identify this case efficiently and thus my *PDGFRA* assay proved useful in this case for treatment analysis. The diagnosis of the *FIP1L1-PDGFRA* fusion by RT-PCR specific for the fusion itself is often very difficult and part of the reason may be the fact that some cases have aberrant breakpoints. Some patients are found to be positive only by nested PCR, even at diagnosis.

Identifier	*FIP1L1-PDGFRA* RT-PCR result
243 [1]	Single step positive
293 [1]	Negative
458	Nested positive
555	Single step positive
591 [2]	Nested positive
606	Nested positive
614	Single step positive
636	Single step positive
646 [2]	Nested positive
759	Single step positive
794	Single step positive
905	Single step positive
939	Single step positive
1025	Single step positive

[1] Same patient
[2] Same patient

Table 4. 2: Validation of the *PDGFRA* multiplex. Patients that are *FIP1L1-PDGFRA* positive were also positive with the *PDGFRA* RT-PCR multiplex.

A further 14 patients were identified in the screen of the 217 *FIP1L1-PDGFRA*-negative samples to have over-expression of *PDGFRA* (~6%) (Table 4.3). These patients were not characterised molecularly and therefore the over-expression of *PDGFRA* observed by the multiplex offered a possible explanation to the pathogenetic cause of the patients' leukaemia.

Case number	Referral reason
94	HES
200	Idiopathic HES
271	?MPD ?HES
289	Eosinophilia
344	HES
349	HES
421	? HES
459	HES
539	Eosinophilia
639	? CML/MPD
680	? CML
683/684	Eosinophilic leukaemia
748	MPD/MDS
843	?MPD

Table 4. 3: *FIP1L1-PDGFRA* negative cases that tested positive for the *PDGFRA* RT-PCR assay.

A colleague was carrying out work to characterise the exact genomic breakpoints of patients with *FIP1L1-PDGFRA* using bubble PCR (see section 2.2.6 for more details). Also included in this analysis was patient E683/E684, who had a complex rearrangement; 46XY, del(3)(p21), add(4)(q12),-10,13q?,+der(?)(?→cen→ ?::4q12→ 4q28.3::10q11.2→ 10qter) and who also tested positive by the *PDGFRA* multiplex in the blood sample (E683) but, interestingly, not the bone marrow (E684). My colleague went on to discover a previously undescribed fusion involving a novel partner gene to *PDGFRA*. The partner gene is a member of the Kinesin family, *KIF5B* and is located at chromosome 10p11.22. The patient was treated with imatinib and six months later had achieved complete cytogenetic and molecular negativity

(Score, *et al* 2006). At the time of publication this was the third *PDGFRA* fusion partner and highlighted the utility of this the over-expression assay to detect rare *PDGFRA* fusions.

Sufficient material was available for further analysis of six of the fourteen cases found to over-express the *PDGFRA* kinase domain. Because of the success of bubble-PCR in identifying *PDGFRA* rearrangements, I used this technique and found a rearrangement in one case with both restriction enzymes. Unexpectedly, sequencing of these rearranged bands revealed *FIP1L1* sequence. Reanalysis of patient cDNA confirmed the presence of *FIP1L1-PDGFRA* mRNA that had been missed on initial analysis. The fusion was detectable by nested but not single-step RT-PCR and this case highlights the problems in reliably detecting *FIP1L1-PDGFRA* (Score, *et al* 2009). The reason for *PDGFRA* over-expression in the remaining cases remains obscure.

4.3.3 Over-expression assays: discussion

Detection of *PDGFRA* and *PDGFRB* rearrangements is of increasing importance but both standard RT-PCR and FISH assays may miss positive cases. Through the blind screening of patients using the newly designed *PDGFRA* and *PDGFRB* multiplexes, I have been able to prove that both assays work well to detect overexpression of these kinases. Both multiplex PCRs identified all patients with previously characterised *PDGFR* rearrangements.

To date, only one patient (E521) found positive by the *PDGFRB* multiplex has actually proved to have a 5q33 rearrangement. The patient was clinically diagnosed as having a CMPD, that was later found to be as a consequence of a previously undetected t(5;12)(q33;p13). None of the remaining patients with over-expression of *PDGFRB* have yielded identification of an oncogene or chromosomal rearrangement and the basis for over-expression remains ambiguous. Surprisingly two positive

patients were known to have rearrangements of the *FGFR1* gene. The reason for this is unclear but it is known that active receptor tyrosine kinases activate complex intracellular signalling networks that result in a variety of cell responses including cell maintenance, mitogenesis, migration and differentiation. Thus it is possible that the *FGFR1* fusions might activate *PDGFRB*, however it is unclear why this should lead to specific over-expression of the kinase domain. Potentially this could be an artifact, for example if *PDGFRB* in this context was alternatively spliced such that the normal full length product was not detected. Alternatively, it is known that *PDGFRA* can be expressed from an internal promoter. (Mosselman*, et al* 1994, Mosselman*, et al* 1996) and it is possible that *PDGFRB* might also have an internal promoter, transcripts from which might be scored in my assay as a potential fusion.

Of the 7 patients found to over express *PDGFRA,* one patient who responded to imatinib was investigated in detail and found by bubble-PCR to have a previously undescribed fusion involving a novel partner gene, *KIF5B* at chromosome 10p11.22. Little is known about *KIF5B* however it is a member of the Kinesin superfamily. The kinesins are microtubule-based motor proteins that play important roles in organelle transport and cell division. They have an N-terminal motor domain which converts ATP into a motile force along the microtubules (plus-end direction only), a central alpha-helical rod domain enabling the protein to dimerise and a globular C-terminal domain, that allows interaction with possible organelle receptors. Xia et. al., (1998) have demonstrated in mice that *KIF5B* is ubiquitously expressed, unlike many of the other family members that are expressed only in neuronal tissues (Xia*, et al* 1998).

Confirmation of both the genomic breakpoint and the cDNA breakpoint were later performed, and the fusion mRNA demonstrated by single step PCR amplification. The complex rearrangement involving chromosomes 2, 4 and 10 is predicted to generate a chimaeric fusion protein that is structurally similar to other tyrosine kinase fusions. Therefore, it is likely that the *KIF5B-PDGFRA* fusion deregulates haematopoiesis in a manner analogous to *BCR-ABL* in CML patients. A schematic

representation of the normal and chimaeric fusion proteins is presented in (Figure 4.8).

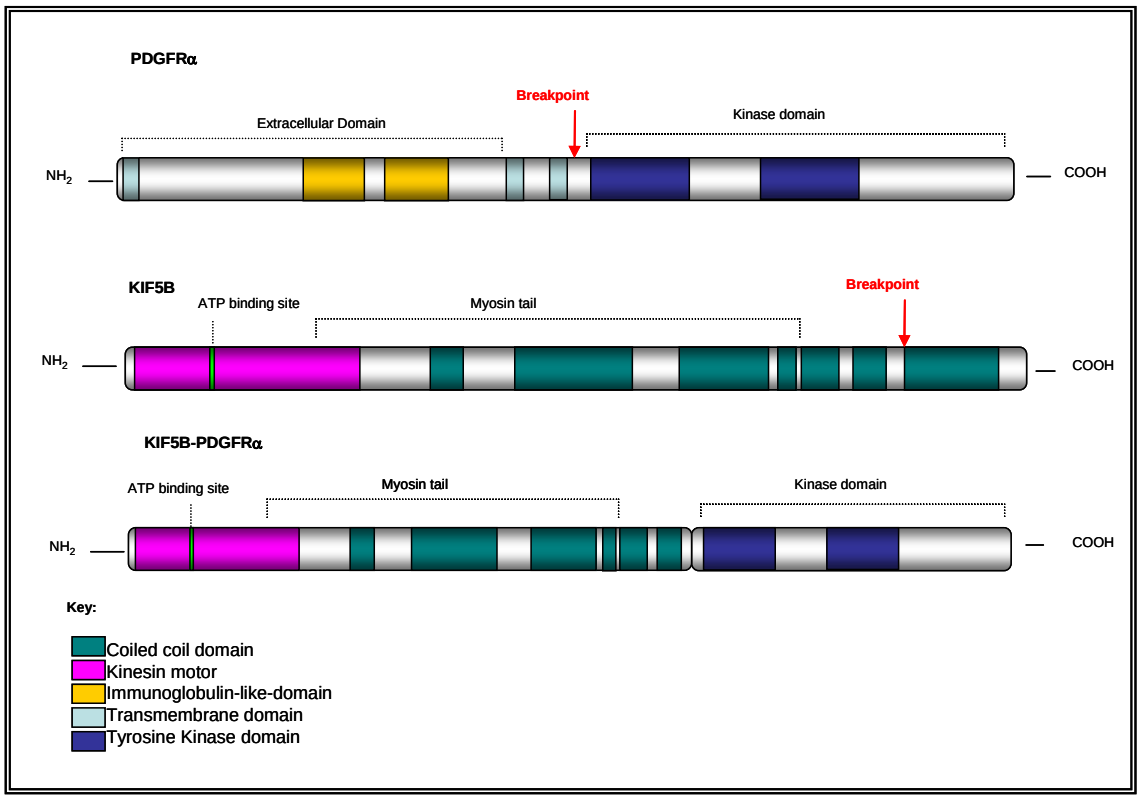


Figure 4. 8: Schematic representation of the normal PDGFRα, KIF5B and the KIF5B-PDGFRα fusion.

The *PDGFRA* RT-PCR assay has not only pointed the way to a new fusion but also proved to be extremely useful for detecting patients with *FIP1L1-PDGFRA*. This more common fusion is predominantly tested for diagnostically by RT-PCR using single and nested primers, however studies by a colleague have highlighted the problem of breakpoints further downstream of the common *FIP1L1* breakpoint (Score, *et al* 2009). Clearly, further research is needed to design an optimal protocol for detection of *FIP1L1-PDGFRA* in patients at diagnosis or after treatment, and my data indicated that simple expression analysis may be very helpful in this regard. In

addition it seems likely that further screening of MPD patients for over-expression of *PDGFRA* and *PDGFRB* will identify new uncommon gene fusions.

Also frequently described in myeloid malignancies are recurrent translocations involving chromosome band 8p11. The 8p11 myeloproliferative syndrome (EMS), also known as stem cell leukaemia-lymphoma (SCLL), is associated with disruption of *FGFR1*. EMS is a CMPD that commonly presents with eosinophilia and associated T-cell proliferation. To date, nine fusion genes have been identified associated with translocations t(8;13)(p11;q12), t(8;9)(p11;q33), t(6;8)(q27;p11), t(8;22)(p11;q22), t(8;17)(p11;q23), t(7;8)(q34;p11), t(8;12)(p11;p11), t(8;19)(p12;q13.3) and t(8;12)(p11;q15) which fuse *ZYMYM2 (ZNF198/RAMP*/*FIM*), *CEP110*, *FGFR1OP (FOP)*, *BCR*, *MYO18A, TRIM24 (TIF1), FGFR1OP2*, *HERV-K* and *CPSF6* to *FGFR1*, respectively (Belloni*, et al* 2005, Demiroglu*, et al* 2001, Grand*, et al* 2004a, Guasch*, et al* 2000, Guasch*, et al* 2003b, Hidalgo-Curtis*, et al* 2008, Popovici*, et al* 1999, Walz*, et al* 2005, Xiao*, et al* 1998).

Imatinib is inactive against *FGFR1* and cannot benefit patients with EMS. Nevertheless, other compounds have been developed such as SU5402, PD173074 and TKI258 that function by blocking the ATP binding site analogous to imatinib, but also possess anti-FGFR activity (Chase*, et al* 2007, Demiroglu*, et al* 2001, Mohammadi*, et al* 1997). These compounds have been shown to specifically inhibit the growth of ZNF198-FGFR1 and BCR-FGFR1 transformed cells, and are therefore more likely to be important in the treatment of EMS patients, although currently no compound is currently available for clinical use.

To encompass the *FGFR1* gene I first thought about designing another RT-PCR assay analogous to those for the PDGF receptors to assess over-expression in uncharacterised MPD patients. However, this would mean testing every patient three times for each of the assays, so instead I designed an MLPA mRNA detection/quantification assay. MLPA has a further advantage in that it has the potential to be expanded further to include other targets.

4.4 *Multiplex Ligation-dependent Probe Amplification (MLPA)*

MLPA is a probe-based technique used routinely to detect genomic deletions and amplifications (Schouten, *et al* 2002) (Figure 4.9) however it can also be used for mRNA profiling. It is a technique capable of establishing the relative copy number of up to 45 nucleic acid sequences in one single reaction by yielding amplification products of unique size per target. Products are separated by automated fluorescent-based sequence electrophoresis and the copy number of each target is determined from the relative intensities of the products compared to those obtained from controls. As only one pair of PCR primers is used, MLPA results in a very reproducible pattern. Furthermore, MLPA is a fast technique that is very simple to perform and needs only 20ng of DNA. This seemed like an ideal procedure to screen for three tyrosine kinase genes in one assay. In addition more genes can easily be added and it is easily adapted to screen large numbers of patients.

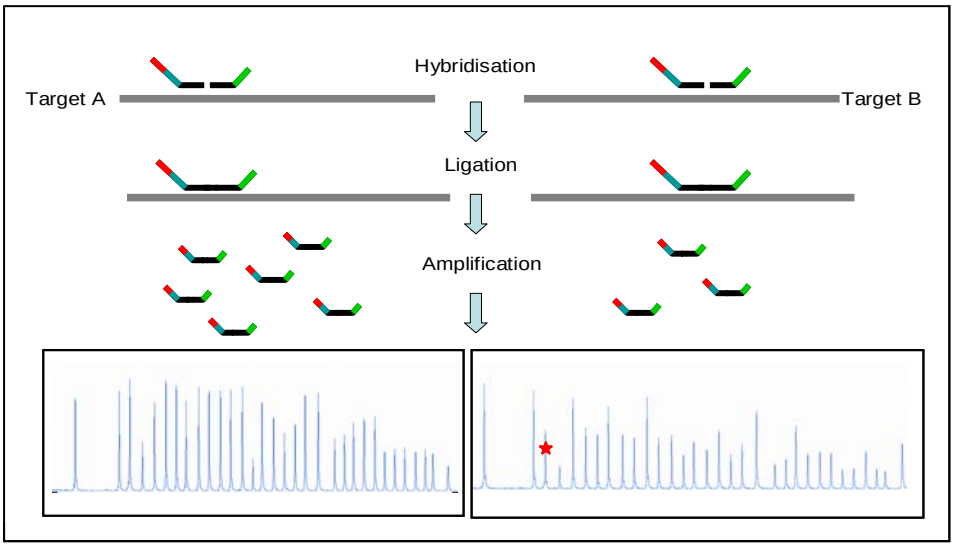


Figure 4. 9: A schematic representation of Multiplex Ligation-dependent Probe Amplification (MLPA). Target A represents a normal MLPA trace, the probe is represented in black, the 'stuffer' sequence in grey (alters the size), and the primers are shown in red

and green. The 5' primer (red) is fluorescently labelled. The probes are first hybridised to the target sequence, these are then ligated together and finally amplified. The amplified products are analysed on an ABI 3100 genetic analyser, a representative trace is shown at the bottom. Target B represents a control where some of the peaks are deleted, the star shows an example of a reduction in peak height compared to the normal.

4.4.1 Methods – Probe design

MLPA depends on the use of progressively longer oligonucleotide probes in order to generate locus-specific amplicons of increasing size that can be resolved by electrophoresis. Schouten et. al., 2002 originally designed these long oligonucleotides using a series of proprietary M13-based vectors, however it is now possible to chemically synthesise long oligonucleotides (Schouten, *et al* 2002). I used a number of rules designated for the design of synthetic MLPA probes published by MRC Holland (www.mrc-holland.com/). An important consideration when designing synthetic probes is that the second probe of each pair must be 5' phosphorylated and significant variation in the quality of long oligonucleotides and the percentage phosphorylation between batches has to be expected.

For each gene I had to design four probes, two that abutted each other within the region encoding the tyrosine kinase domain and two that also abutted each other but were located in the region encoding the extracellular domain. The extracellular peak was then compared to the intracellular peak to determine if there was any over-expression of the kinase domain (retained in gene fusions) compared to the extracellular domain (only present in full length normal transcripts), compared to that seen in controls. As my MLPA assay was based on cDNA I needed to consider specifically how I might design the probes so they would detect only cDNA and not genomic DNA. It was recommended that these probes had an exon boundary close to the ligation site (within 7 nucleotides), thus genomic DNA sequences will not be detected as an intron separates the hybridisation of these two probes. A number of other basic rules have been proposed for optimal probe design: 1) The length of the

two immediately adjacent hybridising sequences must be greater than 21 nucleotides with a predicted melting temperature of more than 70 °C; 2) the first nucleotide following the PCR primer-sequence should preferably not be an adenine as this results in low peak size of the probe amplification product. Specifically the peak size increases in the order A<T<G<C with the difference between A and C greater than two-fold; 3) no more than three G or C residues adjacent to the ligation site; 4) probe target sequences must not overlap as probes will compete for binding and 5) the probes should be 45-60% GC content. An example of one of my MLPA cDNA oligonucleotides is detailed below (Figure 4.10) showing each of the components.

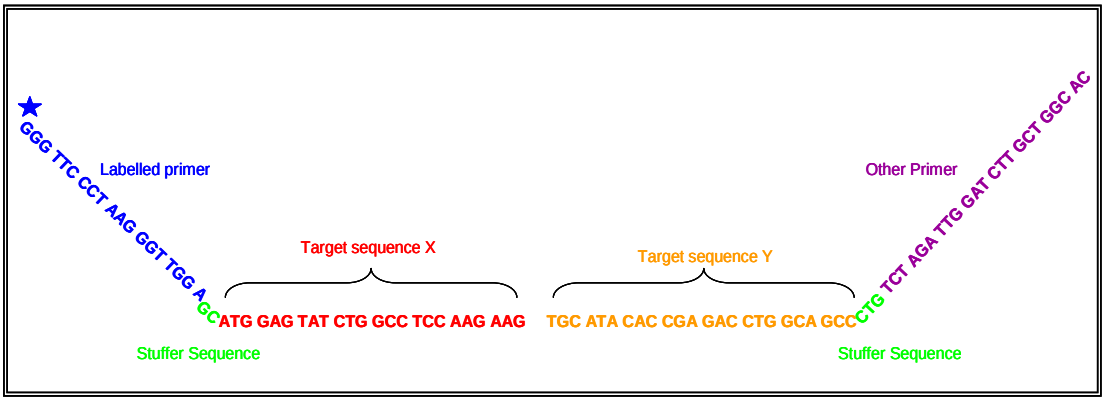


Figure 4. 10: Schematic representing the *FGFR1* intracellular domain probe. The target sequence spans exons 12 and 13. The total size of the MLPA oligonucleotide is 95bp, this includes the complement primer sequences (labelled primer sequence in blue, unlabelled primer sequence in purple), the stuffer sequence (green) and the adjacent target probe sequences (red and orange).

The primer sequences for all MLPA cDNA oligonucleotides are exactly the same allowing amplification of all probes in a single reaction. In addition, only one primer (5'primer) is labelled and this is always with 6-FAM which enables the peaks to be distinguished using fluorescent based capillary electrophoresis. The primer sequences are the same as MRC Holland enabling us to be able to use the kit components from this company. The stuffer sequence differs in size enabling each of the probes to be differentiated by size on electrophoresis. The stuffer sequence must fit with the rules

detailed above and not hybridise to the adjacent target sequence. All other probes are detailed in Appendix I, section XIX.

4.4.2 Methods – Optimisation of the assay

It was anticipated that when samples with known fusion genes were tested the tyrosine kinase portion would appear amplified compared to the extracellular domain. My assay contained 8 probes, two each for the tyrosine kinase genes: *FGFR1, PDGFRB, PDGFRA* and *ABL*. *ABL* was included partly because it is routinely used as a control for real time quantitative PCR assessments of minimal residual disease but also because it is a target of leukaemia-associated translocations. These were added to MRC-Holland MLPA kit components and then used to test sample cDNA according to the manufacturer's protocol (see section 2.2.9). The probes were prepared before use. The cDNA MLPA oligonucleotides arrive lyophilised and were initially dissolved in sterile dH_2O and diluted to achieve a final concentration of 100pmol/µl. To prevent most of the PCR primers being consumed, low amounts of the probe oligonucleotides must be used and the suggested working concentration of each primer per reaction is between 1-4 fmol. To determine the amount of each probe needed the following calculations were made so a stock solution containing all 16 probes could be prepared for use in the MLPA reaction:

1) $\frac{1.33 \text{ fmol}}{\text{Stock concentration (100pmol/µl)}}$ X 200µl = 2.66 µl (volume of primer required)

2) 2.66µl of each probe (16 X 2.66µl = 42.56µl) and make up to 200µl with 157.44µl sterile dH_2O

3) A 1:50 dilution of this probe stock is then required (20µl probe stock + 980µl sterile dH_2O)

4) Followed by a 1:10 dilution of step 3 (10µl from 1:50 dilution + 90µl sterile dH_2O) to make a final 1:500 dilution. This stock solution is used for the MLPA reaction.

To optimise the assay, available cell line cDNA with known fusions was first used along with a number of normal cDNA controls to enable me to identify any probes that gave false positive or negative results. The cell lines used were EOL-1, K562, SD1, Jurkat, HL60, HEL, and 5637. EOL-1 is an AML cell line which carries *FIP1L1-PDGFRA*; K562 is a CML blast crisis cell line carrying the *BCR-ABL* b3a2 fusion gene; SD1 is a pre B-cell ALL which also carries a *BCR-ABL* fusion (e1a2); Jurkat is a human T-cell leukaemia; HL60 is an AML cell line with amplified *MYC*; HEL is an erythroleukaemia cell line containing the *JAK2* V617F mutation, and finally 5637 is a urinary bladder carcinoma cell line. From these controls it was anticipated that only EOL-1, SD1 and K562 would show any differential expression due to known fusion genes. The remainder should act as normal controls however problems were encountered due to the low expression of some of the targets in normal controls (Figure 4.11).

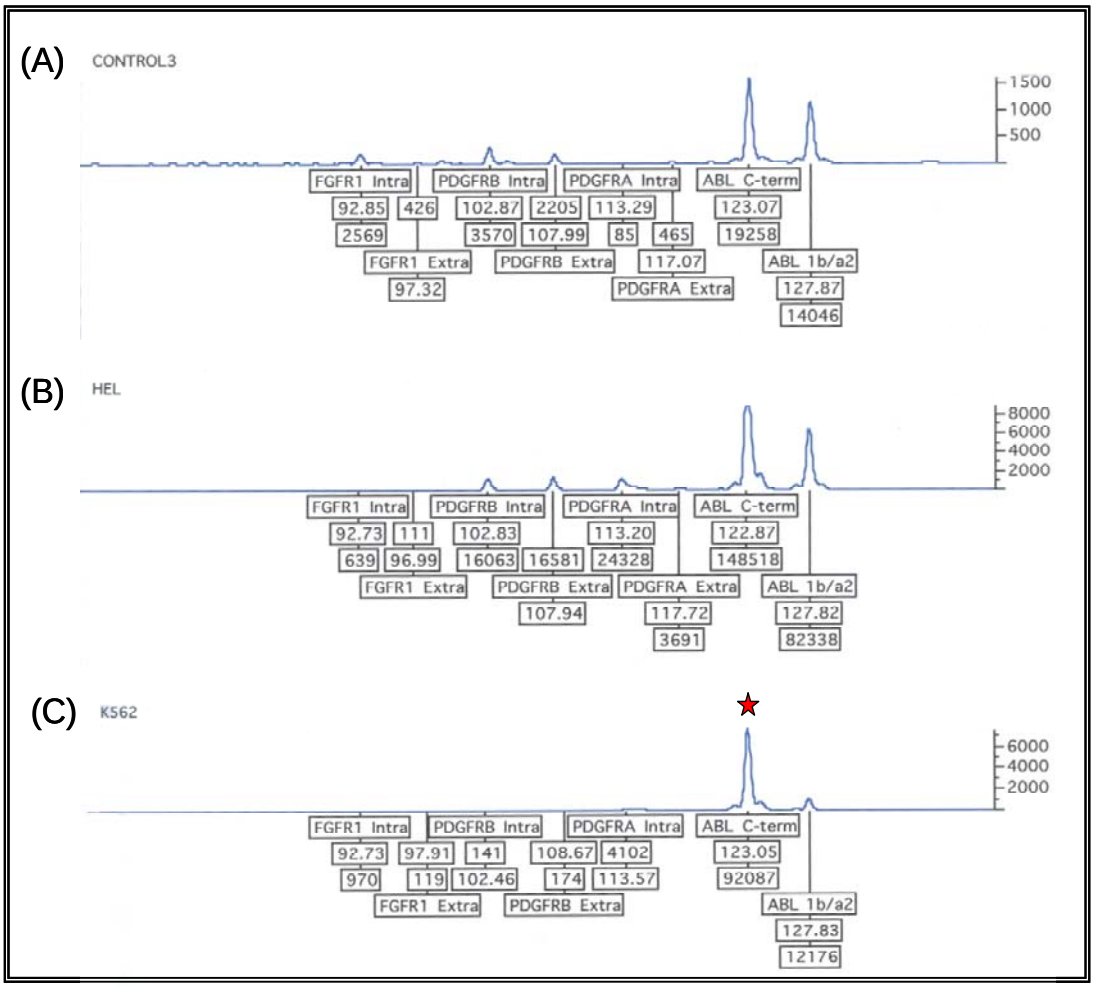


Figure 4. 11: MLPA run 1, electrophoretic traces. (A) A normal control cDNA note the low expression of all probes for most RTKs other than *ABL;* (B) the HEL cell line also predicted to act as a normal control, note the similarity with A, and (C) K562 CML blast crisis cell line carrying a *BCR-ABL* fusion gene, as expected the C-terminal region of *ABL* is clearly amplified (star) compared to *ABL* 1b/a2 extracellular domain.

Because of its relatively high expression in normal controls, the *ABL* probes were used as a reference control except for cases where a known ABL fusion was present, in which case the *FGFR1* probes were used as a control. The table below shows an example of the Jurkat cell line (normal relative to RTK fusions) (Table 4.4). The expected ratio for all extracellular/intracellular peaks is 1 but this was only seen for *FGFR1*, and a substantial difference was seen for *PDGFRA*.

Gene	Intracellular		Extracellular			RATIO
	Peak area	Normalised	Peak area	Normalised		Intra/Extra
FGFR1	84234	0.6415678	33229	0.56473487		1.136051
PDGFRB	317	0.0024144	209	0.00355201		0.679737
PDGFRA	20012	0.1524213	3059	0.05198844		2.93183
ABL control	131294		58840			

Table 4. 4: Table showing the MLPA results for the cell line Jurkat. Left to right: The RTK probe; the intracellular peak area determined by capillary electrophoresis; the normalised value when divided by the corresponding control *ABL* peak (shown at bottom); the extracellular peak area the normalised extracellular value when divided by the corresponding *ABL* control peak, and finally the ratio produced from the normalised values comparing the intra- and extracellular domains. Based on the cut offs used for DNA analysis, a relative value of 1 would indicate a normal 1:1 intra:extracellular expression, a value of 0.5 would indicate significant under expression of the intracellular domain and a value of 1.5 or above would indicate over-expression of the kinase (intracellular domain) compared to the extracellular domain.

I decided to repeat the MLPA assay using the same samples to determine if the assay was reproducible or whether the initial results were biased for any technical reasons. Below is a table showing four of cell lines that were tested twice (Table 4.5). Only the urinary bladder cancer cell line (5637) appeared to produce reproducible results. Unexpectedly, EOL-1 with a known *FIP1L1-PDGFRA* fusion shows over-expression of *PDGFRA* on the first run as expected, but virtually no expression on the repeat assay. It was unclear why the results were so variable, however DNA quality and quantity on commercially produced MRC-Holland kits is known to be of utmost importance and it seems likely that cDNA quality and quantity is even more difficult to control for.

Gene	EOL1	EOL1	K562	K562	Jurkat	Jurkat	5637	5637
FGFR1	0.808339	2.321555	8.151261	1.650277	2.534954	3.1048	2.935601	3.947915
PDGFRB	174323	2355	0.810345	3.462349	1.516746	1	1.074525	1.190534
PDGFRA	392	0.001842	4102	0.001351	6.542007	28615	4.304769	4.050285
ABL control	2.49143	3.098622	7.562993	13.85285	2.231373	1.90531	1.503822	1.519456

Table 4. 5: A table showing the comparison between the first MLPA run (black) and the repeat MLPA run (red). Only four cell lines have been shown here for comparison (other cell line data not shown). Only the 5637 bladder cancer cell line shows reproducible results, the remainder are highly variable.

Further experiments were conducted to try and improve the reproducibility using a series of freshly prepared cDNAs from normal controls and Eol1 cells. The results were more promising with the EOL-1 electrophoretic trace showing a clear over-amplification of the intracellular kinase domain of *PDGFRA* consistent with the presence of the *FIP1L1-PDGFRA* fusion gene (Figure 4.12A). Even the normal control traces looked cleaner and although expression of both portions of every RTK gene was not always observed, the *ABL* control was always present (Figure 4.12B). Despite cleaner traces, the intracellular:extracellular ratio was rarely calculated as 1 and in fact this still varied considerably, especially where one of the probes was lacking completely. Moreover, in most cases the *ABL* control looked differentially expressed implying the *ABL* 1b/a2 probe may not be as sensitive as its counterpart.

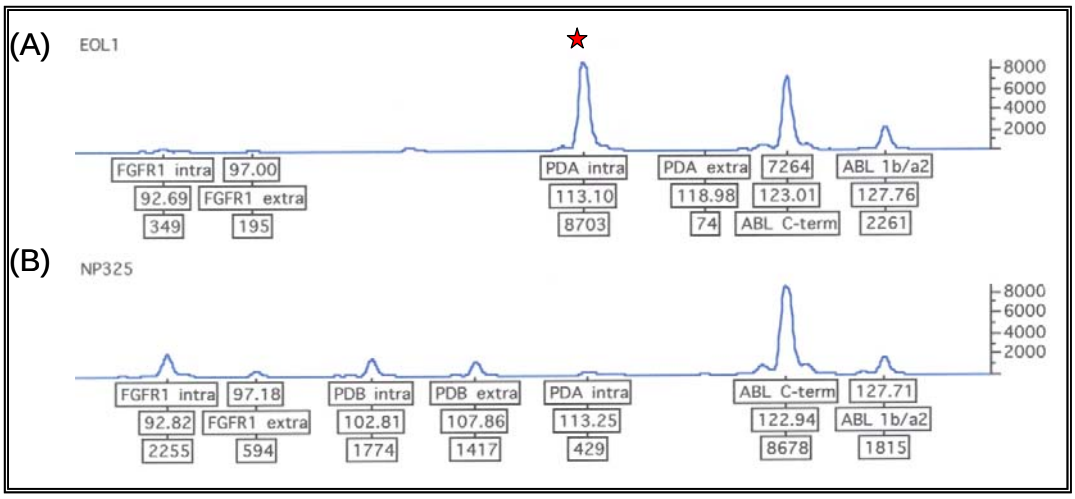


Figure 4. 12: MLPA run 3, electrophoretic traces. (A) EOL-1 cell line showing a clear amplification of the *PDGFRA* intracellular domain, the extracellular domain is barely visible; compare this to (B) a normal control where expression of every probe is visible yet very tightly regulated. Note however the differential expression of the *ABL* probes.

Fresh cDNA samples were then made from patient samples with available material and known fusion genes involving *PDGFRA, PDGFRB, ABL* and *FGFR1* (n=15). All the samples with known fusions showed an amplification of the intracellular domain of the relevant RTK. The results were very encouraging, although CML cases showed a high degree of variation. Figure 4.13 shows an example MLPA trace showing over-amplification for each of the tyrosine kinases in the assay.

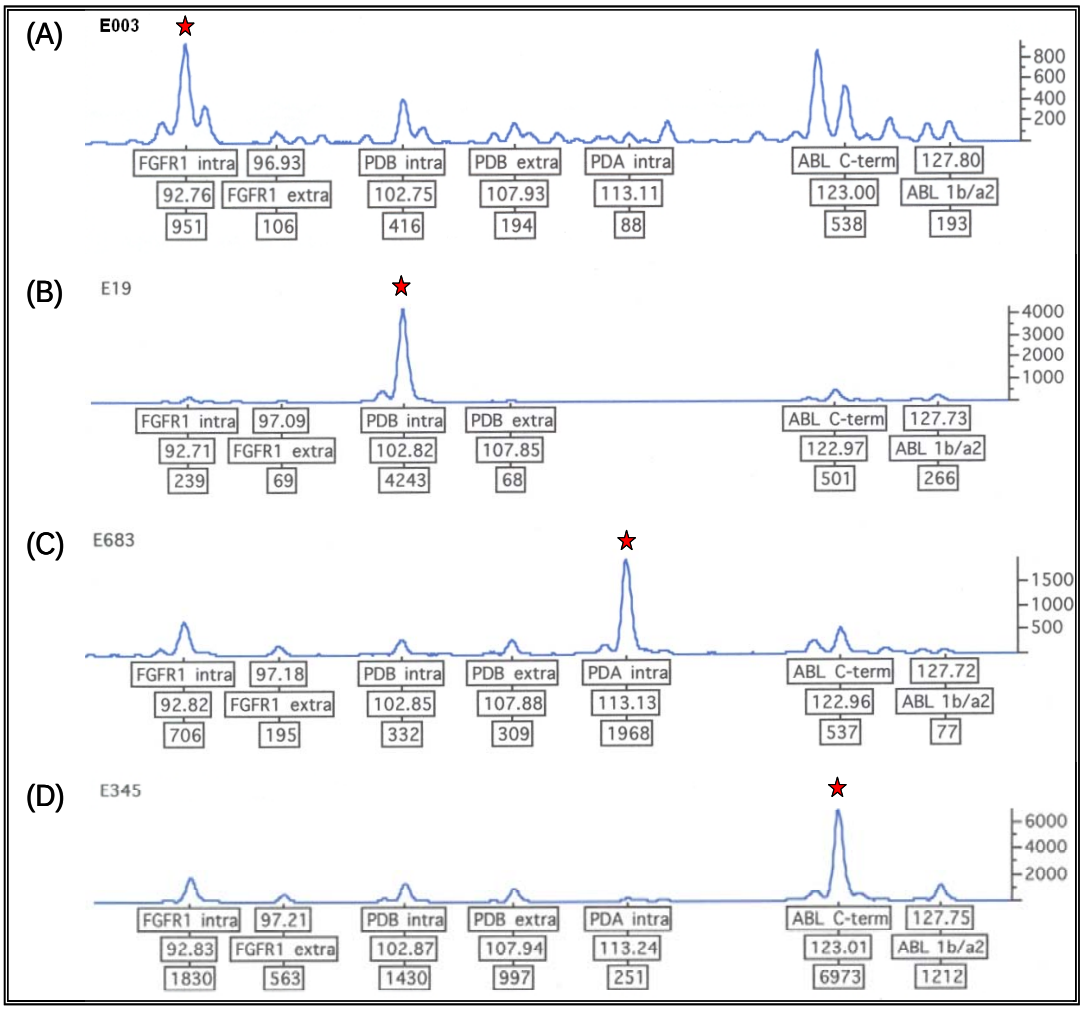


Figure 4. 13: MLPA run 4, electrophoretic traces. (A) A patient with a known *FGFR1* fusion gene, note the over-amplification of FGFR1 intracellular probe compared to the extracellular probe (ratio of ~9); (B) a patient with a *ETV6-PDGFRB* fusion gene only the intracellular *PDGFRB* probe is highly over-expressed; (C) a patient with a *PDGFRA* fusion gene, the extracellular *PDGFRA* probe is not even visible, and finally (D) a CML patient with a *BCR-ABL* fusion showing clear over-expression of the *ABL* C-terminal probe.

It is important to note here that although over-amplification was observed for the intracellular domain of the correct RTK, in all cases the ratios of the other probes were rarely 1.0 and the CML patients often showed *ABL* and *PDGFRA* intracellular domain over-expression. As a result of this work I decided to concentrate only on

PDGFRA, PDGFRB and *FGFR1*, and use the *ABL* probe only as a control for normalisation.

It is clear from these results that sample quality and probe efficiency influence the results as well as any genuine difference in expression of the target sequences. In addition, distortions in the data and substantial variations between replicates were seen when peak sizes were zero or very low. Based on the results I therefore created some rules based on +/- 5 standard deviations of the values seen for all normal controls compared to cases known to harbour specific fusions (Table 4.6).

Gene	Extracellular =0, or very low	Extra & Intracellular present
FGFR1	Fail if extracellular not >50	Positive is >10.87
PDGFRB	Positive if intracellular is >200, ideally >500	Positive is >5.10
PDGFRA	Positive if intracellular is >200	Positive is > 4.39

Table 4. 6: Table shows rules for analysing unknown patients to be screened for possible over-expression of the kinase domain (intracellular) compared to a control band (extracellular) based on normal and positive control data. The data has been split according to detection of both probes and only one probe for each RTK.

4.4.3 Validation of cDNA MLPA assay

Having created these rules I had a colleague blind my samples so I could repeat the MLPA assay and determine its predictive value. These were the same 14 patient samples with known fusions plus four normal controls and the EOL-1 cell line. By applying the rules I was able to identify all four patients with known *FGFR1* rearrangements, three of four *PDGFRB* rearranged patients and four of five *PDGFRA* rearranged patients (including the EOL-1 cell line). The two patients missed with a *PDGFRA* or *PDGFRB* fusion were scored as normal indicating a false negative result of approximately 15%. Nevertheless, all the normal controls were scored as normal therefore no false positives were identified. Analysis of a further blinded series

correctly identified 9 of 9 cases with *PDGFRB* rearrangements, and a false positive result was seen in one out of sixteen normal controls (6%).

4.4.4 Results

A number of patients (n=16) were then screened that had been previously found by our laboratory to be sensitive to imatinib *in vitro* or were of particular interest for other reasons such as the presence of a particular chromosome translocation. All of these patients had an unknown molecular pathogenesis however the sensitivity to imatinib implies they harbour a fusion gene or mutation affecting a tyrosine kinase gene that has a response to imatinib such as *PDGFRA* or *PDGFRB.* Other cases included 5 that were chosen due to possible *PDGFRB* involvement, including one patient that had been tested previously by Western blot and shown to have massive over representation of the PDGFRβ protein. Another patient had been found positive by my *PDGFRA* RT-PCR multiplex and the remaining were imatinib responders *in vitro*.

Of the sixteen patients I found only 4 to have an unusual expression of the intracellular domains of one of the tyrosine kinases. Three of these patients were *in vitro* imatinib responders, E1009, E281 and E57 and all showed over-expression of *PDGFRA.* The final patient (E1334) showed over-expression of *PDGFRB*. The sample that was screened due to over-expression of *PDGFRA* by the multiplex assay proved to be normal by MLPA analysis, and also by bubble-PCR. E1334, the patient with over-expression of *PDGFRB* was analysed further using 5'RACE PCR but no *PDGFRB* fusion was identified. Insufficient material was available for further investigation of the other cases.

More patients were screened using the MLPA assay (n=39), of whom 16 were further candidate imatinib responders, 18 were found previously positive by *PDGFRA* RT-PCR multiplex and 5 new patients with unusual referrals.

I found 2 of the 16 imatinib responders consistently failed, 8 were shown to be normal, 5 had over-expression of *PDGFRA* and one was borderline *FGFR1* positive. All the *PDGFRA* positive patients were followed up by bubble-PCR and found to be normal. The borderline positive *FGFR1* case was later repeated and found to be normal. The 18 patients with *PDGFRA* over-expression first identified by RT-PCR multiplex were analysed and only 10 of these were confirmed by MLPA assay to also over-express *PDGFRA*. Of the remaining 8 patients, 5 appeared normal and three consistently failed. These patients were also followed up by bubble-PCR to determine if any unusual fusions might be identified but none were discovered. Finally, of the 5 new referrals, 2 were *PDGFRA* positive, 2 were borderline *FGFR1* positive and one was normal. The *PDGFRA* over-expressing patients were analysed by bubble-PCR but no rearrangements were identified. The patients with *FGFR1* over-expression were analysed using 5'RACE PCR as they were both referred with query 8p11 myeloproliferative syndrome (EMS). Again no rearrangements of the *FGFR1* gene were discovered.

In summary, although the MLPA expression assay was validated with a large number of positive controls and normal controls it is far from a perfect screening test and yielded a large number of apparently false positives. Nevertheless, subsequent analysis demonstrated that the test could be useful.

4.4.4.1 A cryptic ETV6-PDGFRB fusion gene

The outcome of the t(5;12)(q31-33;p13) is the *ETV6-PDGFRB* fusion and, although rare, is the most frequent of the known *PDGFRB* fusion genes (Steer and Cross 2002). In patients with a t(5;12)(q31-33;p13), confirmation of the presence of *ETV6-PDGFRB* is particularly important since many cases with this translocation have a distinct molecular aetiology, with upregulation of *IL-3* expression and no involvement of *PDGFRB* (Cools*, et al* 2002). Only cases with a *PDGFRB* fusion are expected to be responsive to imatinib.

The fusion of *ETV6* exon 4 spliced to *PDGFRB* exon 11 is found in the great majority of positive cases, and is typically detected by RT-PCR using primers directed to *ETV6* exons 3 or 4 and *PDGFRB* exon 11 (David*, et al* 2007). In our laboratory, of 10 positive cases detected with these primer combinations, only one had a variant fusion in which *ETV6* exon 4 was fused to *PDGFRB* exon 9. A patient with idiopathic myelofibrosis and an *ETV6* exon 7-*PDGFRB* exon 10 fusion was also recently reported and it was suggested that this novel transcript may have accounted for the somewhat unusual clinical phenotype seen in this case (Tokita*, et al* 2007). Importantly, both these variants are detectable using the standard primer sets above. Here I present an unusual *ETV6-PDGFRB* case that was missed by standard screening.

4.4.4.2 Patient details

The patient was a 19 year old male (E1026) who presented to the outpatient department with a fever for two months. On examination he was found to have generalised lymphadenopathy, hepatosplenomegaly and bony swellings over the left clavicle and right tibia, however there were no bleeding manifestations. A total leukocyte count exceeding 100 x 10^9/L with a left shift was observed on peripheral blood smear. Bone marrow aspiration showed an increase in the basophils and eosinophils, suggestive of an MPD. Fine needle aspiration showed infiltration of the lymph node by blasts that were negative for myeloperoxidase and periodic acid Schiff on cytochemistry. Flow cytometry of the peripheral blood and lymph node aspirate was positive for myeloid markers (CD11c, 13, 15 and 33) along with CD45 and myeloperoxidase. *BCR-ABL* was not detected by RT-PCR and cytogenetics showed a t(5;12)(q31-33;p13), however *ETV6-PDGFRB* was also not detected by RT-PCR. A diagnosis of atypical CML in accelerated phase was made.

Despite being negative for *ETV6-PDGFRB* the patient was started on imatinib mesylate (600mg in divided doses) and exhibited a complete haematological and cytogenetic response to treatment (including complete regression of the lymph nodes, bony swellings and hepatosplenomegaly) at 3 months post treatment. One year after diagnosis the patient continues in remission on 600mg/day imatinib.

4.4.4.3 cDNA MLPA analysis of patient E1026

In view of the excellent response to imatinib, we performed further analysis to determine if *PDGFRB* was involved using the cDNA MLPA assay (material was not available for molecular cytogenetic analysis). This assay detected relative over-expression of the *PDGFRB* kinase domain in this patient (E1026) but not in a number of normal controls, suggesting the presence of a *PDGFRB* fusion (Figure 4.14).

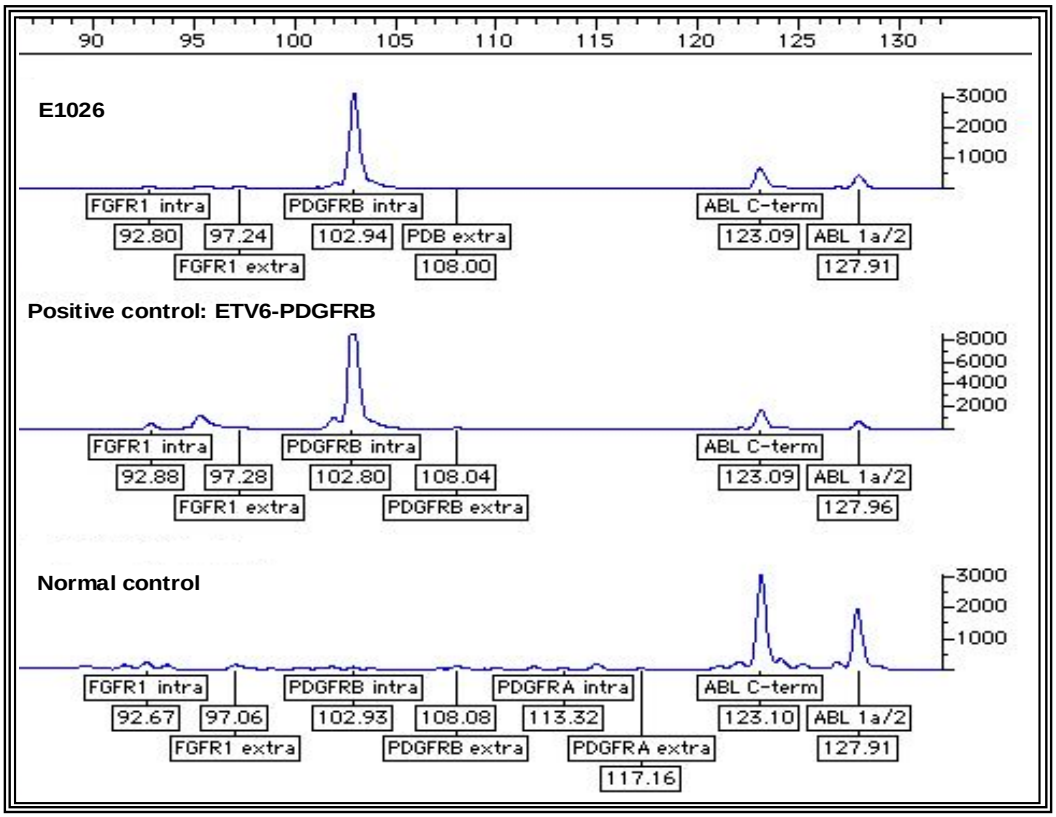


Figure 4. 14: The cDNA MLPA assay. Probes are in order *FGFR1* tyrosine kinase domain, *FGFR1* 3'extracellular domain, *PDGFRB* tyrosine kinase domain, *PDGFRB* 3'extracellular domain, *PDGFRA* tyrosine kinase domain, *PDGFRA* 3'extracellular domain. Under normal circumstances expression of *FGFR1, PDGFRB* and *PDGFRA* is low level in peripheral blood leukocytes. Over-expression of the *PDGFRB* intracellular domain is apparent in both the previously known *ETV6-PDGFRB* case (middle) and also E1026 (top), compared to the normal control (bottom).

4.4.4.4 5'RACE PCR to identify the potential PDGFRB fusion gene in patient E1026

I employed 5'rapid amplification of cDNA ends PCR (Invitrogen, Paisley) to characterise the fusion, employing reverse primers derived from sequences located downstream of known *PDGFRB* breakpoint locations (Appendix I, section IV). The amplification was performed under long-range PCR conditions (section 2.2.5) to enable fragments up to 10kb to be amplified. After nested PCR a smear was observed with 4 or 5 possible bands visible, ranging from 250bp – 650bp. The negative control was completely blank indicating one of these bands may have been derived from a *PDGFRB* fusion (Figure 4.15).

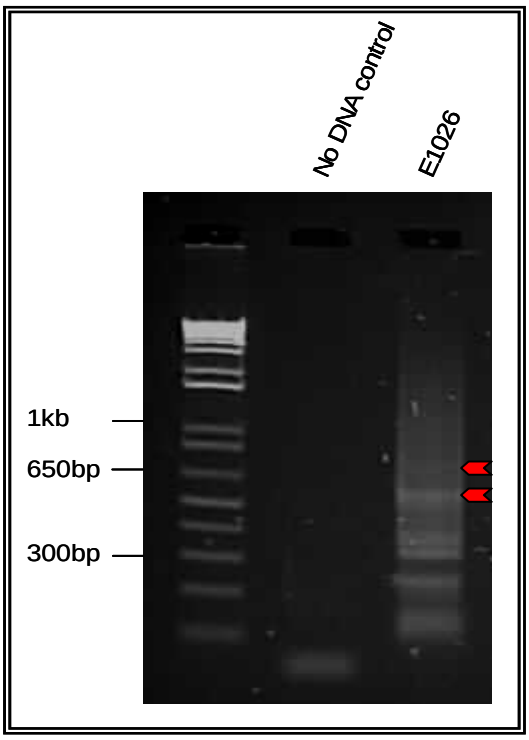


Figure 4. 15: Agarose gel of nested 5'RACE PCR for patient E1026 versus a control. Lane 1 indicates a no DNA control, only a primer dimer smear is visible below 100bp; Lane 2 shows the nested PCR result after E1026 5'RACE PCR a smear is visible containing several different sized bands ranging from 250-650bp. Red arrows indicate the bands that were cloned and sequenced.

The resulting PCR products were cloned and the digests revealed only two different sized bands one at approximately 500bp and the other at approximately 650bp (indicated by red arrows, Figure 4.15). Sequencing of the smaller product turned out to be *PDGFRB* and vector sequence only, however sequencing of the larger 650bp product revealed an in-frame mRNA fusion between full length *ETV6* exon 7, 34bp derived from *ETV6* intron 7 and a truncated *PDGFRB* exon 12 (Figure 4.16). This fusion involves novel breakpoints, never previously reported.

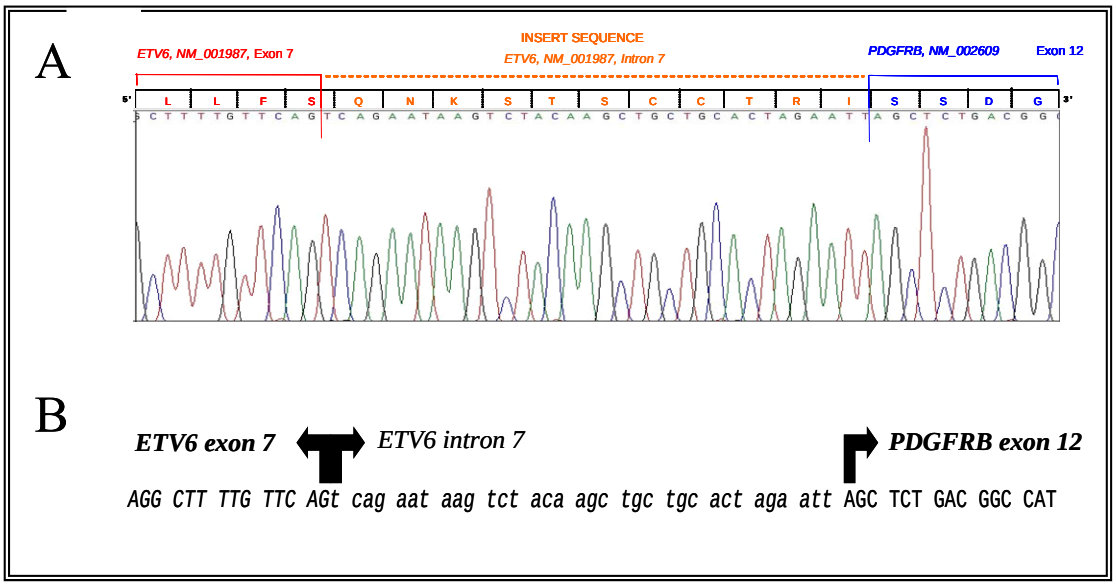


Figure 4. 16: (A) Electropherogram of the cDNA junction amplified by 5'RACE PCR; (B) Sequence of the novel *ETV6-PDGFRB* fusion showing the insertion of *ETV6* intronic sequence ensuring the fusion reads in-frame.

4.4.4.5 Confirmation of the novel ETV6-PDGFRB fusion

The presence of the fusion was confirmed by RT-PCR using the primers *ETV6* exon 6.2F and *PDGFRB* exon 14.1R with the High Fidelity PCR master kit according to manufacturer's conditions (see appendix I, section XX for primer sequences). The PCR was performed on a generic 66°C annealing temperature PCR protocol. The in-frame fusion was detectable by single step RT-PCR, as seen in Lane 1, Figure 4.17. RT-PCR of the fusion revealed two bands the upper most of which was of the expected size, 415bp. On cloning and sequencing the lower band turned out to be an *ETV6-PDGFRB* out-of frame fusion between part of *ETV6* exon 7 and *PDGFRB* exon 13. Using the same High Fidelity PCR kit and amplification protocol, I used RT-PCR to try and amplify the reciprocal translocation. *PDGFRB-ETV6* reciprocal products were not detectable by single or nested PCR, Lane 2, Figure 4.17.

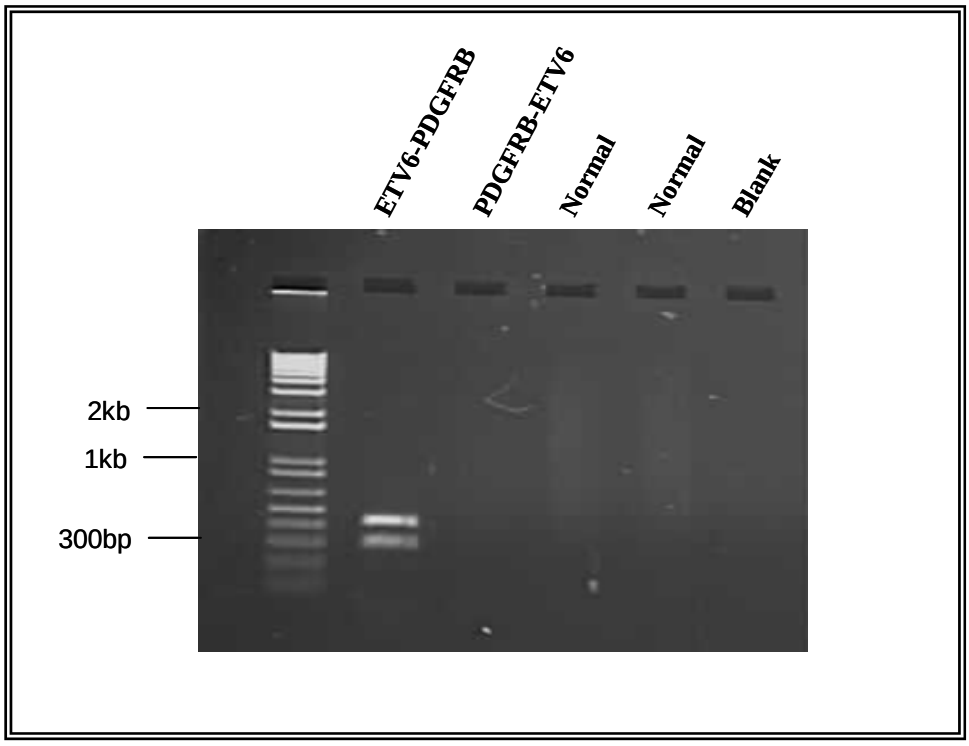


Figure 4. 17: Amplification of the *ETV6-PDGFRB* fusion by single step RT-PCR: Lane 1 represents patient E1026 fusion RT-PCR; Lane 2 represents the reciprocal breakpoint PCR in patient E1026, Lanes 3 & 4 show normal controls, and Lane 5 a no DNA control. Note the reciprocal breakpoint was not observed at single-step or nested PCR (data not shown). The fusion RT-PCR shows two bands of 300 and 415 base pairs, these represent an out-of-frame fusion and the confirmed in-frame fusion**.**

I also went on to determine the genomic breakpoints of this fusion for further confirmation using the primers ETV6.E1026.Int7.1F and ETV6.E1026.Int7.2F with PDGFRB exon 12R (appendix I, section XX). Both PCRs resulted in a product that was subsequently cloned and sequenced. I was able to show the same fusion breakpoints between exon 12 of *PDGFRB* and intron 7 of *ETV6* in both PCR products (Figure 4.18).

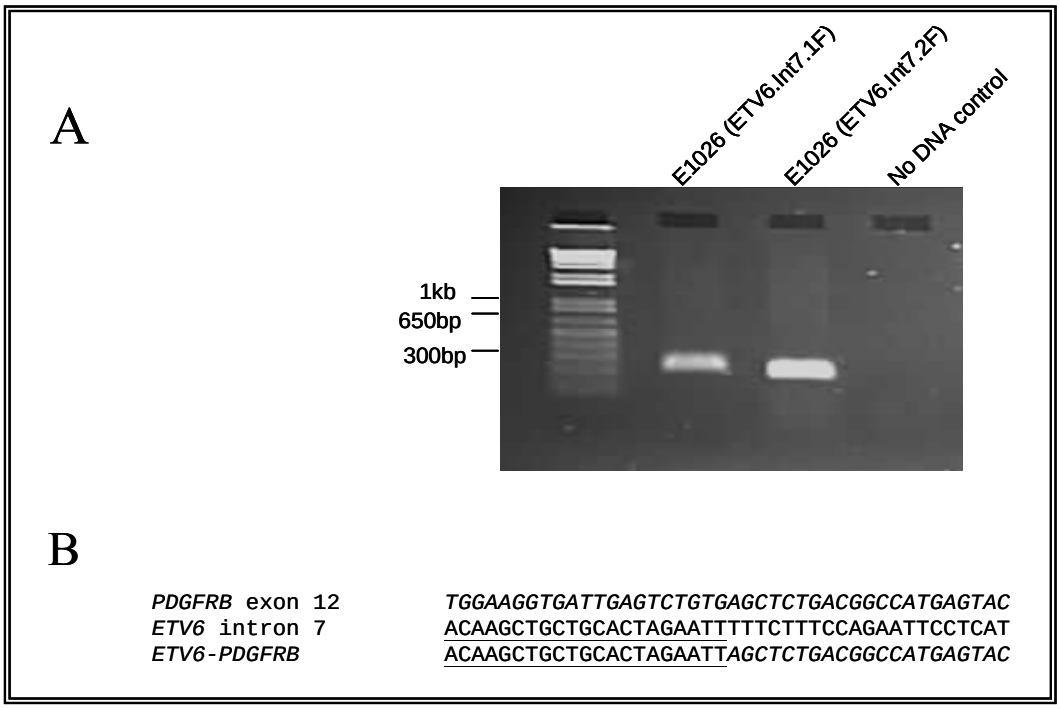


Figure 4. 18: (A) Amplification of the genomic *ETV6-PDGFRB* novel fusion. Lane 1 is amplification of the genomic breakpoint using primer ETV6.Int7.1F, Lane 2 is the genomic breakpoint amplified using a second primer ETV6.Int7.2F, both at single step combined with primer PDGFRB exon12.1R; (B) Schematic showing the individual sequences for *PDGFRB* (italics), *ETV6* and the *ETV6-PDGFRB* fusion sequence.

4.4.4.6 Further ETV6-PDGFRB analysis

The *PDGFRB* sequence retained in the novel fusion is downstream of the PCR primer we had initially used for RT-PCR, explaining why *ETV6-PDGFRB* fusion was not detected. To ascertain if any other similar cases had been missed, 10 cases with a t(5;12) that tested negative for *ETV6-PDGFRB* using the standard primers were re-analysed using ETV6.Ex2F and PDGFRB.Ex14R. All those tested proved negative, indicating that this variant is not common. Nevertheless, as a result of the finding of this novel *ETV6-PDGFRB* fusion the above primers are now routinely used to screen t(5;12) cases referred to our laboratory.

4.4.4.7 ETV6-PDGFRB - discussion

In summary, patient E1026 tested negative by RT-PCR for the *ETV6-PDGFRB* fusion, but following the finding that his disease responded to imatinib, MLPA and 5'RACE analysis showed the patient to have a novel *ETV6-PDGFRB* variant with breakpoints outside those commonly reported.

The unusual *ETV6-PDGFRB* fusion is reminiscent of those seen for *FIP1L1-PDGFRA*, for which breaks almost always fall within *PDGFRA* exon 12 and *FIP1L1* intron-derived sequence is frequently incorporated into the mature mRNA (Cools*, et al* 2003a). As a consequence the region encoding the PDGFRα auto-inhibitory WW-like domain is disrupted, resulting in partner protein independent activation of the kinase moiety (Stover*, et al* 2006). The disruption of PDGFRβ WW-like domain in our case is unexpected in view of the fact that *ETV6* encodes a dimerisation domain that is required for transformation by 'normal' *ETV6-PDGFRB* (Carroll*, et al* 1996).

Interestingly, this case presented with advanced phase disease, something not normally associated with this fusion. The fusion itself retained *ETV6* exons 1-7, in common with that described by Tokita et. al., (2007) in a case that also presented with advanced features, specifically those of chronic idiopathic myelofibrosis. In contrast to the most common fusion reported that involves *ETV6* exons 1-4, these variants are predicted to encode a chimaeric protein that retains the *ETV6* internal and ETS DNA-binding domain (Figure 4.19). It is possible that retention of this domain leads to a more aggressive phenotype however it would be premature to draw any clear conclusions from just two cases (Tokita*, et al* 2007).

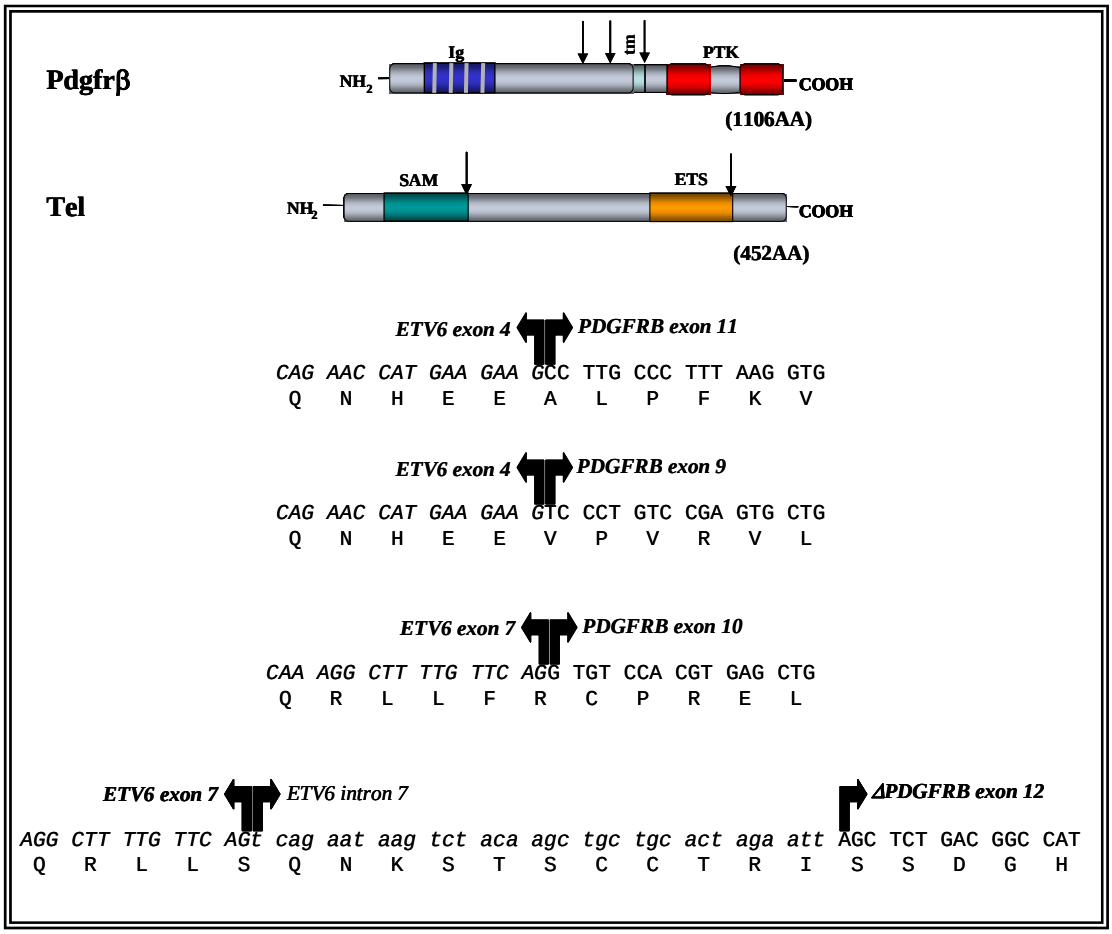


Figure 4. 19: Schematic of PDGFRβ and the Tel (ETV6) proteins showing relevant domains and arrows indicating reported breakpoints. Below, are the sequences of *ETV6-PDGFRB* fusions described to date, with the most common variant at the top and the fusion in case E1026 at the bottom.

It is also noteworthy however that our case responded well to imatinib despite being treated in advanced phase. Good responses to imatinib have been described for *FIP1L1-PDGFRA* associated chronic eosinophilic leukemia in transformation (Metzgeroth, *et al* 2007) and this may be a general feature of diseases associated with *PDGFR* fusions.

It is clear from this case that the breakpoint diversity for *ETV6-PDGFRB* fusions is greater than originally reported and hence diagnostic screens for this fusion should

employ primers capable of detecting all variants in order to provide an accurate molecular diagnosis allowing for appropriate clinical management.

4.4.5 MLPA Discussion

MLPA is a technique that utilises probes to detect deletions and amplifications of a gene. I designed a cDNA MLPA assay to detect over-expression of the kinase domain of specific tyrosine kinase genes relative to a control probe spanning a number of exons upstream of the common reported breakpoints. The basis of the assay is that under normal conditions the mRNA expression of each tyrosine kinase is tightly regulated, but when fused to a partner gene the expression level of the tyrosine kinase portion of the gene may be increased as a consequence of gene fusion. By screening patients of unknown aetiology I hoped to be able to discover cases with cryptic or novel fusion genes. I designed the assay to incorporate both *PDGFR* RT-PCR assays into one reaction whilst also accommodating for the *FGFR1* and *ABL* tyrosine kinases. MLPA is ideal for screening multiple tyrosine kinase genes in one assay, and hence is economical on volume of material required. In addition, it is designed to screen high numbers of patients quickly and reproducibly.

The assay was validated using numerous normal control cDNA samples sourced from a non-leukaemic background along with a number of cell lines and patient samples with previously known fusion genes involving one of the tyrosine kinase genes within the assay. Most positive controls were identified, with over-amplification of the intracellular domain of the correct RTK in these cases however, the intracellular:extracellular ratios of the other unaffected probes were rarely 1.0, presumably because of differences in probe efficiencies. The false negative result was 15%, where known positive controls were scored as normal. Analysis of more positive controls was more successful with nine out of nine patients identified correctly. The normal controls were all scored as normal with only one exception that may have been related to poor cDNA quality.

A major drawback of the cDNA MLPA assay was the low expression of the tyrosine kinases under normal circumstances which made reproducibility difficult. The assay was also very dependent on the quality of the cDNA and it appeared that some probes were more sensitive than others. This was particularly apparent when I tested known CML patients for validation of the assay. *ABL* over-amplification often appeared concomitantly with over-expression of the *PDGFRA* intracellular domain making these patients difficult to distinguish. Despite redesigning new *ABL* probes to try and combat this problem differential expression of the *ABL* probes themselves was always observed. As a result only the *PDGFRA, PDFGRB* and *FGFR1* probes were used to try and identify possible fusion genes, and the *ABL* probe was used as a control for normalisation. Despite the drawbacks the results were collated and rules of analysis created based on plus or minus five times the standard deviation for all normal controls compared to positive controls.

Having validated the assay it was used to screen MPD patients of unknown molecular pathogenesis from three basic categories: imatinib responders as previously determined by a colleague using *in vitro* liquid and colony assays (n=26), patients found to over-express *PDGFRA* or *PDGFRB* by RT-PCR assays (n=24) and finally patients with cytogenetic or other indicators of possible gene fusions (n=6). Apparent over-expression was seen for several cases but only one (2% of all cases) proved to be positive on further investigation

4.5 Discussion

Persistent aberrant signalling of PDGF and its receptors is thought to be essential for the survival of the many cancer cell types. These findings have influenced the pharmaceutical and research communities to design agents that selectively block PDGFR signalling (Levitzki 2004). Imatinib mesylate has received the most attention and has potent inhibitory activity against both PDGFRα and PDGFRβ. Patients with MPDs and rearrangements of *PDGFRA* and *PDGFRB*, in particular, have been

shown to respond very well to treatment with this compound (Apperley, *et al* 2002, David, *et al* 2007). However, a shortcoming of imatinib has been the emergence of acquired resistance in the form of point mutations in the kinase domain in a subset of cases. Furthermore some patients show primary or secondary resistance for unknown reason and others are intolerant of therapy. The development of second generation ATP-competitive tyrosine kinase inhibitors (TKI) has begun to try and overcome these problems. These include nilotinib, a novel aminopyrimidine inhibitor (Weisberg, *et al* 2005). In pre-clinical studies nilotinib has been shown to be 30-fold more potent than imatinib in treating chronic-phase CML, and Stover et. al., (2005) have also showed it is effective at inhibiting proliferation *in vitro* of *ETV6-PDGFRB, FIP1L1-PDGFRA* and imatinib resistant *ETV6-PDGFRB*/T681I, but not *FIP1L1-PDGFRA*/T674I mutants (Stover, *et al* 2005). A second TKI, dasatinib is a potent dual Src/Abl kinase inhibitor that is 67-fold more potent than imatinib (Shah, *et al* 2004). Its greatest benefits have been observed in patients with CP-CML with 90% achieving a complete haematological response (Steinberg 2007). Dasatinib is also effective against KIT and PDGFRs, in addition to overcoming resistance exhibited by 18/19 imatinib resistant mutations reported (Shah, *et al* 2006). Sunitinib has also been studied for its effectiveness against PDGFRs although most of this work has centered on PDGFR-associated solid tumours such as colon cancer, non-small cell lung carcinoma and melanoma. However, preclinical trials have shown this TKI has both potent and robust anti-angiogenic and anti-tumour activity (Chow and Eckhardt 2007). Finally, other inhibitors have also been developed that have shown activity against PDGFR, such as sorafenib. This is a small molecule B-Raf inhibitor currently used for the treatment of renal cell carcinoma. It has been shown *in vitro* to be a potent inhibitor of *ETV6-PDGFRB* by inducing cell cycle block and apoptosis (Lierman, *et al* 2007).

To date, 21 *PDGFRB* and 6 *PDGFRA* fusion genes have been reported. It seems likely that additional *PDGFR* rearrangements are yet to be discovered, particularly in view of occasional reports of cases that are responsive to imatinib despite being negative for known fusions and, in most cases, cytogenetic indicators of variant

fusions (i.e. rearrangements involving 4q12 or 5q31-33). To search for new fusions I designed RT-PCR and MLPA assays to detect over-expression of the region encoding the kinase domain of selected tyrosine kinases. This region is retained in gene fusions and may be overexpressed relative to the normal, full length product. Both assays were validated using cases with known rearrangements and normal controls and although there is no doubt that the assays could be considerably improved, e.g. by testing and incorporating other primer combinations, both were applied to screen cases of unknown molecular aetiology.

I screened 229 MPD patients negative for *FIPL1-PDGFRA* by *PDGFRA* RT-PCR and found a total of 14 that over-expressed the kinase domain of *PDGFRA.* Of these cases two yielded identification of unknown fusion genes. The first was a previously unreported novel *KIF5B-PDGFRA* fusion as a consequence of a complex cytogenetic rearrangement involving chromosomes 2, 4 and 10. The second case proved to have a *FIP1L1-PDGFRA* fusion that had not previously been detected by standard RT-PCR. The remaining cases for whom material was available (n=12) were investigated further by bubble-PCR as *PDGFRA* rearrangements are particularly suited to this technique since the breakpoints in all cases reported to date are clustered within exon 12 of this gene. No *PDGFRA* rearrangements were identified after screening these patients in this way, suggesting a false positive rate for the expression assay of ~6%.

I also screened 199 MPD patients looking for over-expression of the *PDGFRB* kinase domain; of these I found 9 to be strongly positive. Surprisingly, two of these were later discovered to have known *FGFR1* rearrangements suggesting interplay between the signalling cascades affecting both these tyrosine kinase genes. Of the remaining 7 patients only one was found to harbour an unknown *PDGFRB* rearrangement. This in fact turned out to be the most common of the *PDGFRB* fusions, *ETV6-PDGFRB.* It is unclear why the patient was initially screened and found negative by standard RT-PCR but clearly the usefulness of screening all MPD patients using this more generalised RT-PCR has proved its worth. Other patients with available material were

subsequently screened by 5'RACE PCR but no further rearrangements were discovered. This gives a false positive rate of 4%.

The MLPA assay proved to be less reliable: of 56 MPD patients screened (some of whom were also screened by the RT-PCR assays above) 25 were found to over-express either *PDGFRA, PDGFRB* or *FGFR1*. *PDGFRA* over-expression was the most frequent (n=20), 2 over-expressed *PDGFRB* and three were borderline for over-expression of *FGFR1*. 5'RACE PCR and bubble-PCR of all patients where material was available was performed and with one exception no rearrangements were detected. The exception was a patient referred with a t(5;12) who turned out to have a variant *ETV6-PDGFRB* fusion with breakpoints that could not be detected with standard primers.

The most important conclusion to be drawn from these results is that cryptic fusions of *PDGFRA*, *PDGFRB* and *FGFR1* are very uncommon in MPDs, even in cases that responded to imatinib but tested negative for known fusions. This is because although apparent over-expression of specific kinase domains were seen by the RT-PCR and/or MPLA assays, further analysis failed to reveal gene fusions in most cases. Although RACE PCR might miss some fusions, particularly if the partner gene is very large, we have previously found that bubble PCR is highly effective for isolating genomic *PDGFRA* fusions since the breakpoints within this gene are tightly clustered. Most apparent over-expressors involved *PDGFRA* and thus it seems likely that most were false positives. A second conclusion is that although both assays proved to be useful in identifying new fusions in the research setting, their false positive rates are unacceptable for routine diagnostic use because of the amount of effort required to follow these cases up. It seems likely that the false positive rates could be reduced by the use of further primer sets in the MPLA assay and possibly the more rigorous definitions of cutoffs to define positive and negative results, It is also possible that more quantitative techniques such as real time PCR may be useful to measure the relative expression of the regions encoding the extracellular and

intracellular domains of tyrosine kinases. Other systems such as exon expression arrays might also be used, but would be expensive for routine use.

Overall it is estimated that tyrosine kinase fusions account for only about 2-5% of atypical MPD cases (Reiter, *et al* 2007). Other abnormalities of tyrosine kinase or downstream signalling components have also been described, e.g V617F *JAK2* (10% of cases), *NRAS* mutations (12.5% of cases) and activating *FLT3* or *KIT* mutations (5% of cases) (Jones, *et al* 2005). Collectively these abnormalities account for little over a quarter of all atypical MPD cases (Figure 4.20).

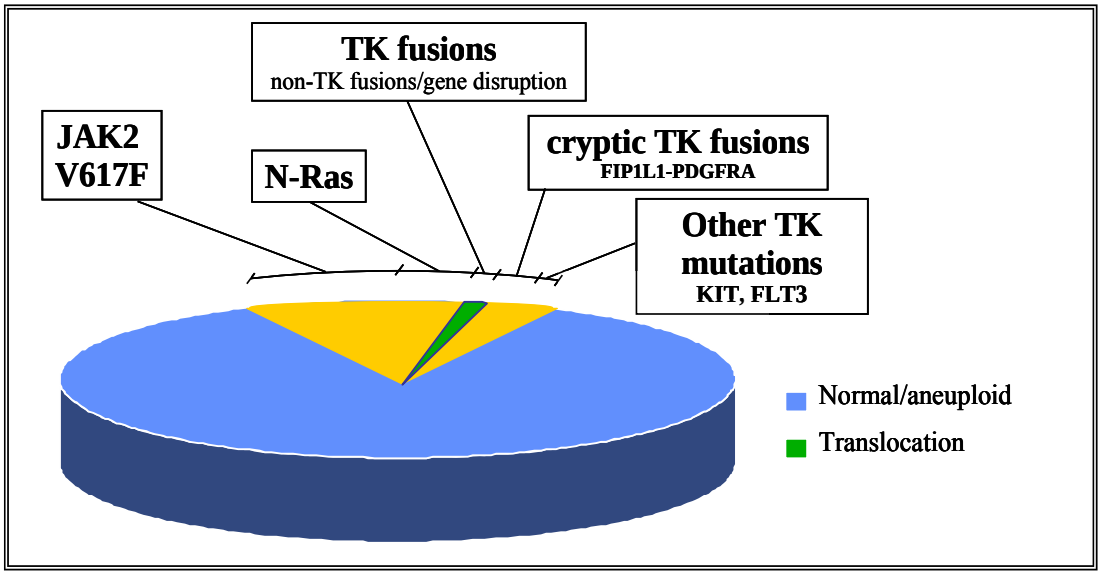


Figure 4. 20: Pie chart showing the molecular pathogenesis of atypical MPD patients (i.e. atypical CML, CMML, CEL, MPD-U, and MDS/MPD-U). The blue region represents atypical MPD patients that have a normal or aneuploid karyotype and are of unknown aetiology. The section in yellow indicates collectively the number of patients defined molecularly with tyrosine kinase mutations or mutations of downstream signaling components. The section in green shows the small number of patients found specifically to have fusion genes involving tyrosine kinases.

If cryptic tyrosine fusions are uncommon, then other abnormalities must be responsible for the majority of atypical MPDs for whom the molecular pathogenesis

is currently unknown. There are several possibilities: first, they may harbour gene fusions of other tyrosine kinases that I have not investigated. Second, they may harbour activating mutations of tyrosine kinases that would not have been detected by my over-expression assays. Third, they may harbour activating mutations in other signalling components that are either upstream or downstream of tyrosine kinases. Fourth, they may harbour mutations in completely different pathways. In the next Chapter I describe a novel approach aimed at identifying mutations in atypical MPDs.

5 SCREENING FOR NEW ONCOGENES MARKED BY ACQUIRED UNIPARENTAL DISOMY

5.1 Introduction

5.1.1 Loss of heterozygosity (LOH) & acquired uniparental disomy (aUPD)

Loss of heterozygosity (LOH) refers to a widespread phenomenon in cancer whereby normal inherited heterozygous markers in a genomic or chromosomal region are reduced to homozygosity in the malignant clone. LOH frequently marks the presence of a recessive tumour suppressor gene, the classic example being retinoblastoma in which an inherited inactivating mutation in the Rb gene along with flanking heterozygous markers are all reduced to homozygosity in retinal tumours (Knudson 1971). LOH may originate from diverse mutational events such as deletion of the wild-type allele to yield a hemizygous state, mitotic recombination, localised gene conversion or non-dysjunction with either loss or reduplication of the mutant chromosome (Figure 5.1). There are various techniques discussed in detail below that can identify LOH and help to distinguish between the underlying mutational mechanism. In the last few years the discovery of large stretches of somatically acquired LOH associated with no net change in copy number, termed as acquired isodisomy or acquired uniparental disomy (aUPD) has emerged as a common theme, initially in polycythaemia vera but subsequently in acute leukaemia and other malignancies (Irving, *et al* 2005, Kralovics, *et al* 2002, Raghavan, *et al* 2005).

Classically, uniparental disomy (UPD) refers to the finding of an individual inheriting both homologues of a chromosome pair from only one of their parents as a result of an error in meiosis. It is associated with a range of inherited developmental disorders such as Beckwith-Wiedemann syndrome (UPD of chromosome 11p15) and Prader-

Willi syndrome (25% have maternal UPD of chromosome 15). My work is concerned with identifying regions of aUPD that are thought to be caused by mitotic recombination and selection for one of the derivative products. Recent evidence, reviewed below, has suggested that aUPD is usually associated with specific oncogenic mutations. Definition of novel, minimal regions of aUPD may therefore be an important new route to identify novel oncogenes.

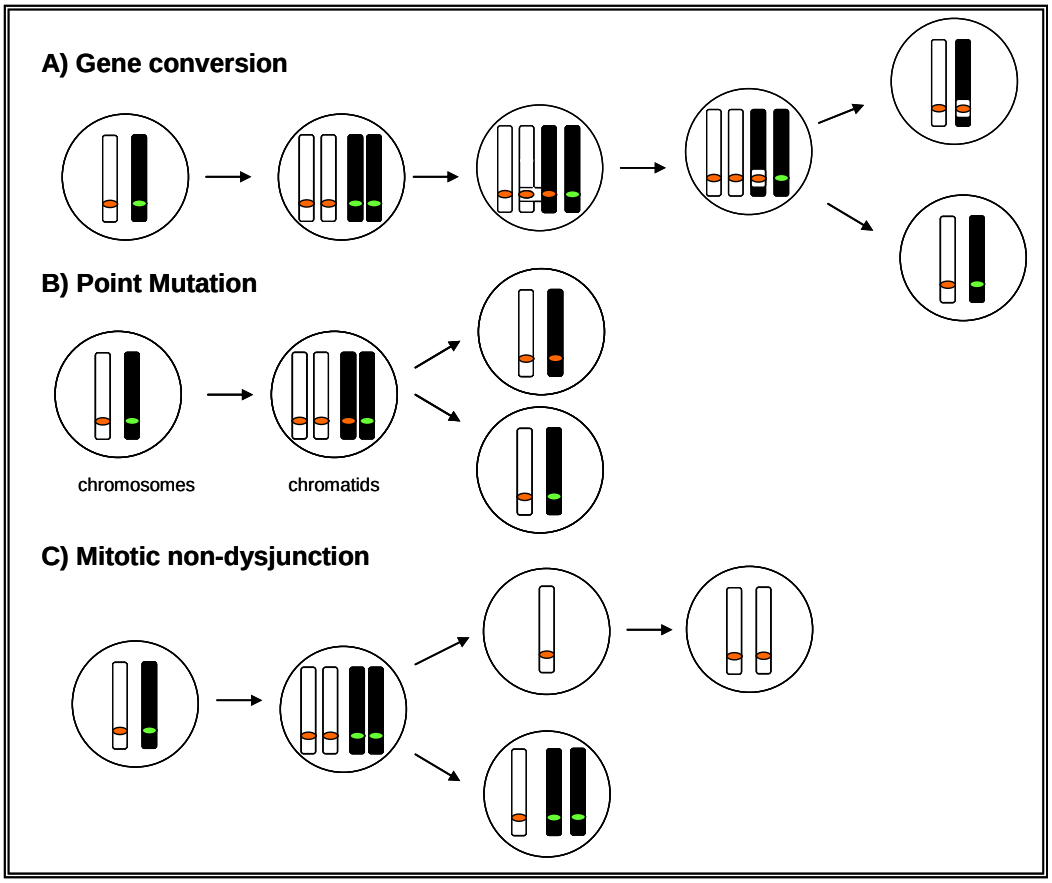


Figure 5. 1: Schematic illustrating the principal classical mutational events that lead to LOH: A) Interallelic gene conversion, a non-reciprocal exchange occurring between two alleles that reside on homologous chromosomes resulting in loss of the normal allele as shown in the last cell at the top; B) A point mutation may occur through a reading error during semi-conservative replication resulting in an additional mutant chromatid, represented by the end cell at the top, finally C) Mitotic non-dysjunction. This example shows non-dysjunction followed by re-duplication of the mutant chromosome.

5.1.2 aUPD and MPDs

The first description of aUPD in MPDs was reported by Kralovics and colleagues in 2002 (Kralovics, *et al* 2002). Previously, studies of rare familial polycythaemia vera (PV) had suggested a possible pathogenic role for loss of heterozygosity (LOH) in this disease. As a consequence they performed whole genome LOH analysis using a series of microsatellite markers on a small number of PV subjects and identified LOH at three new genomic loci, by far the most frequent of which was at 9p. They then showed that there was no change in copy number at 9p, suggesting that mitotic recombination was the mechanism of LOH i.e. aUPD rather than allele loss. The prevalence of 9p aUPD was observed to be approximately 33% of all PV cases, and is therefore the most frequent chromosomal lesion seen in these patients. This group mapped the minimal aUPD region and identified the possible candidate genes within this region, which included the *JAK2* tyrosine kinase. Although this group went on to screen candidate genes in this region including a partial screen of *JAK2*, no mutations were observed and the significance of their observations was unclear.

5.1.3 aUPD & the discovery of the JAK2 V617F acquired mutation

It took a further 3 years before the observations of 9p aUPD in PV came to fruition. In 2005 four groups reported the discovery of an acquired G to T mutation at nucleotide 1849 in exon 12 of the *JAK2* gene that results in the substitution of a valine by a phenylalanine (V617F) (Baxter, *et al* 2005, James, *et al* 2005, Kralovics, *et al* 2005a, Levine, *et al* 2005a). The four groups took three different approaches to find this mutation:

(i) James et al., took a functional approach. They observed that the formation of erythropoietin (Epo)-independent colony formation (EEC), one of the biological hallmarks of PV, was impaired by both JAK2 inhibitors and also short interfering RNA (siRNA) against JAK2. They went on to screen the *JAK2* gene and identified

the mutation. They then showed that introduction of V617F into mouse bone marrow cells by retroviral-mediated gene transfer gave rise to a PV-like phenotype which was not seen in controls, thus clearly implicating V617F as causative in the disease process. This seminal discovery again highlighted the importance of aberrant tyrosine kinase signalling in the pathogenesis of MPDs.

(ii) Based largely on the recurrent findings of abnormal tyrosine kinases in atypical MPDs, both Baxter et. al., 2005 and Levine et. al., 2005 performed mutational screens of genes encoding tyrosine kinases.

(iii) Kralovics et. al., 2005 followed up on their earlier work that identified 9p UPD in PV. Analysis of further patients led to a smaller minimally affected region that still included *JAK2*. Despite having been apparently excluded previously, *JAK2* remained an obvious candidate because of its known importance in erythropoiesis. Consequently, this gene was re-screened and the V617F mutation identified in a small region that was apparently not covered in their previous analysis (Figure 5.2).

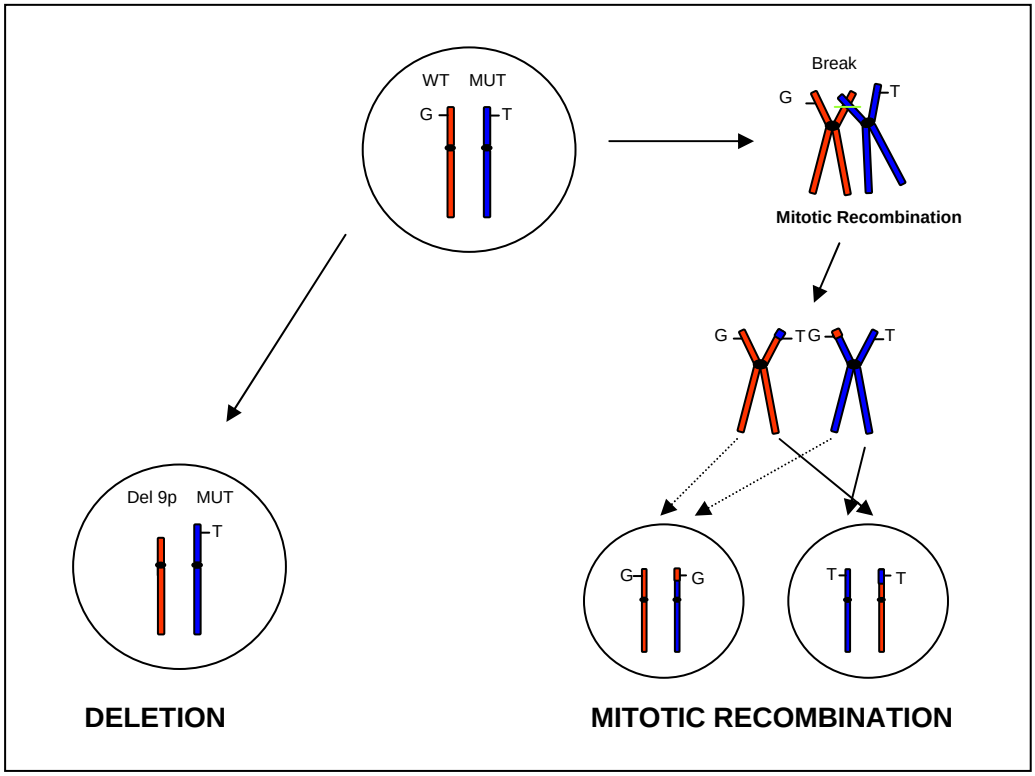


Figure 5. 2: Mechanisms and consequences of 9p LOH in PV. The chromosome 9 with the wild-type *JAK2* sequence (G) is shown in red, and the chromosome 9 with the G→T substitution (T) is shown in blue. Circles signify the nuclei of cells. Loss of the telomeric part of the wild-type chromosome 9p by deletion is shown on the left, as a potential mechanism for LOH. Alternatively, on the right mitotic recombination is depicted, as a second mechanism of 9p LOH. The events that may occur during mitosis are represented, as are the resulting cell progeny.

At the time our group was also screening tyrosine kinase genes for mutations, mainly in atypical MPDs. Once the finding of V617F *JAK2* emerged, we looked at a series of cases for the presence of this mutation. In total, we investigated 480 MPDs and found that the percentage of PV, ET and IMF patients positive for the mutation correlated well with the other studies. Interestingly the mutation was also seen in a subset of atypical MPDs: 30/152 (20%) atypical or unclassified MPD, 2/134 (2%) idiopathic hypereosinophilic syndrome were V617F positive but the mutation was not seen in any patients with systemic mastocytosis (SM), AML, secondary erythrocytosis or in normal individuals (Jones*, et al* 2005).

5.1.4 The association between JAK2 V617F and 9p aUPD

Work performed by Amy Jones using both direct sequencing and pyrosequencing, observed a wide range in the proportion of *JAK2* alleles that carried the V617F mutation. In many patients the mutation appeared to be homozygous, i.e. the residual wild type allele was completely or almost completely lost. Homozygosity was only seen in PV and MF and strikingly, the frequency was similar to that of aUPD as described previously (Kralovics*, et al* 2002). To investigate the relationship between homozygosity for V617F and aUPD 9p in more detail, I used microsatellite PCR analysis to look for chromosome 9p LOH in a series of patients. Microsatellites are repetitive sequence elements that are arranged in tandem and dispersed throughout almost the entire genome. The length of the core repeats are highly variable, and form allelic variants that can be analysed using PCR. I used 8 primer pairs flanking the

dinucleotide (CA) highly polymorphic microsatellite markers on chromosome 9p (Table 5.1).

Chromosome band	Marker	Heterozygosity (http://www.gdb.org/)
	D9S1779	0.64
p24.3	D9S1858	0.58
p24.2	D9S288	0.86
p24.2	D9S1813	0.84
p24.1	***JAK2***	
p24.1	D9S286	0.88
p23	D9S254	0.75
p22.2	D9S157	0.85
p21	D9S171	0.80

Table 5. 1: Microsatellite marker name, location on chromosome 9p in relation to the *JAK2* gene, and the relative heterozygosity of each microsatellite marker.

I analysed 30 MPD cases that did not have the V617F mutation and 27 V617F homozygous patients. The products were analysed on an ABI3100 using the Genotyper 2.0 program. If only one peak was visible a patient was scored as homozygous. To allow for the presence of background normal cells, homozygosity was also scored if one peak was one third or smaller than the expected size compared to controls. A significant tract of homozygosity was scored if four or more consecutive markers abutted or encompassing *JAK2* were homozygous, since the probability of this occurring by chance was well under 5% as determined by known the rates of heterozygosity (http://www.gdb.org/). The results of the analysis are

summarised in Table 5.2. Of the 27 V617F homozygotes, 26 were found to have chromosome 9 UPD (96%). Of the 30 *JAK2* negative PV, IMF and ET patients, only four were found to have significant homozygosity (13%) ($P<0.0001$, Fisher's exact test). This indicates that LOH is significantly associated with homozygosity for the V617F mutation.

To determine if the observed homozygosity was due to aUPD or LOH, MLPA was performed to look for copy number by another member of our group. One copy of *JAK2* would indicate LOH due to either a whole chromosome deletion or specific chromosomal region, whereas two copies would indicate aUPD. All seven cases analysed by MLPA that were homozygous for at least four 9p microsatellites and were homozygous for the V617F mutation were shown to have two copies of *JAK2*, consistent with chromosome 9 aUPD (Kralovics, *et al* 2002). Whilst my analysis was ongoing Kralovics et. al., (2005) published similar data showing a strong correlation between V617F homozygosity and 9p aUPD (Kralovics, *et al* 2005a).

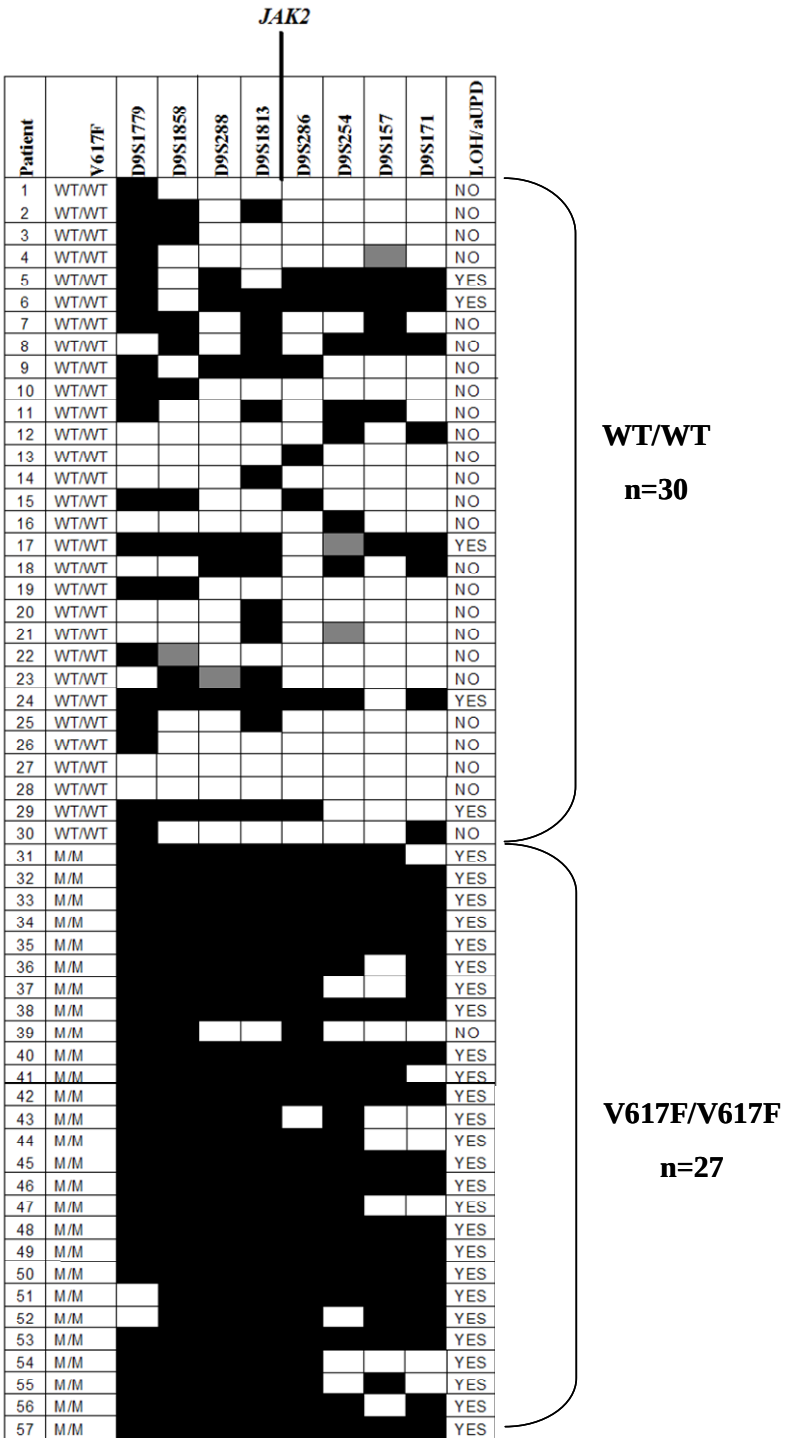


Table 5. 2: 9p LOH in V617F homozygous patients compared to V617F negative cases. The V617F genotype of the patient is indicated on the right as either homozygous normal (wt/wt; patients 1-30) or homozygous mutant (M/M; patients 31-57). The microsatellite markers are shown in order from the most telomeric (D9S1779) to the

most centromeric (D9S171) and are scored as homozygous (black), heterozygous (white) or not done (grey).

5.1.5 Acquired UPD is not limited to MPDs

More recently it has become apparent that aUPD is widespread in haematological malignancies, for example AML (Raghavan, *et al* 2005), ALL (Irving, *et al* 2005) chronic lymphocytic leukaemia (Pfeifer, *et al* 2007) and multiple myeloma (Walker and Morgan 2006) plus also solid tumours such as basal cell carcinomas (Teh, *et al* 2005), prostate cancer (Torring, *et al* 2007) and colorectal cancer (Gaasenbeek, *et al* 2006).

In contrast to the initial studies in PV which used microsatellite analysis, these more recent studies used genome-wide microarray single nucleotide polymorphism (SNP) analysis. For example, Raghavan et. al., showed that approximately 20% of normal karytotype AMLs exhibited large regions of homozygosity that was not associated with visible chromosomal abnormalities. They were also able to show that a previously identified homozygous *CEBPA* mutation coincided with a region of large scale aUPD of chromosome 19 (Raghavan, *et al* 2005). Further studies in AML determined that aUPD for chromosome 13q was associated with activating *FLT3* mutations (Griffiths, *et al* 2005), aUPD 11p with inactivating *WT1* mutations and aUPD 21q with *RUNX1* mutations (Fitzgibbon, *et al* 2005). Mutations in other malignancies have also been linked to homozygosity for specific mutations, e.g. the glucocorticoid receptor gene in pediatric ALL (Irving, *et al* 2005), *PTPN11* gene in ALL associated with Noonan syndrome (Karow, *et al* 2007), *CDKN2A* and *TP53* in follicular lymphoma (Fitzgibbon, *et al* 2007) and *NF1* in JMML (Flotho, *et al* 2007).

These studies show a strong association between regions of aUPD and specific acquired mutations in oncogenes or tumour suppressor genes. In reports where paired samples have been studied, it appears the mutation precedes the aUPD and hence the aUPD is a 'second hit' leading to homozygosity of an oncogenic mutation associated

with disease progression (Fitzgibbon, *et al* 2007, Fitzgibbon, *et al* 2005). Global UPD screens may therefore be a new approach to help identify new oncogenes and tumour suppressor genes in malignancy.

5.1.6 Aims of my analysis

The primary aim of my analysis was to determine if aUPD was present in atypical MPDs and, if so, to follow up any findings to try and identify new pathogenetic lesions in these diseases. As described above, MPDs are characterised by deregulated tyrosine kinases and therefore our principal hypothesis was that novel activating tyrosine kinase abnormalities were likely to be the root cause of many or all atypical MPDs. Acquired UPD has been associated with two activated tyrosine kinases – *JAK2* and *FLT3* – and therefore identification of new regions of aUPD may therefore be a particularly good method to provide pointers towards new kinase mutations. In addition, since this was an entirely new and potentially very powerful method of analysis, I also took the opportunity to analyse a range of other cases of haematological disorders.

5.2 SNP CHIPS TO IDENTIFY REGIONS OF UPD

5.2.1 SNP chips

SNPs are believed to be the most common sequence variation in the human genome, occurring every approximately 1,200bp (Sachidanandam, *et al* 2001) and thus there are likely to be more than 2 million polymorphic sites in the human genome. The identification of SNPs that are roughly evenly spaced has enabled the realisation of whole genome SNP screens. The advent of high density single SNP chips has revolutionised the field of genome-wide allelic screening as they can accurately and reproducibly detect copy number and loss of heterozygosity simultaneously. A few reports have compared LOH differences observed by microsatellite analysis versus

those found by SNP array and demonstrated excellent concordance however, as expected, additional smaller regions of LOH were detected by SNP arrays due to their higher resolution (Dumur, *et al* 2003, Lindblad-Toh, *et al* 2000, Wang, *et al* 2004b).

5.2.2 Plan of analysis and methods

When I started this project there were two platforms available: Affymetrix 10K chips, i.e. chips that harboured 10,000 SNPs across the genome, and Affymetrix 100K chips, which actually consisted of 2 x 50K chips that each contained 50,000 genome-spanning SNPs. The studies that had been published at the time had used 10K arrays and this density was easily sufficient to identify regions of UPD (Irving, *et al* 2005, Raghavan, *et al* 2005). However I wanted to explore the possibility of looking at more detail at copy number changes and in particular small deletions or amplifications that may mark the presence of cryptic fusion genes. Because of the relatively high cost of the technique I elected to use just one 50K chip per sample in order to maximise the number of cases that could be analysed.

The generic method of sample preparation for SNP chip use was devised by Kennedy et. al., (2003) (Kennedy, *et al* 2003). The whole-genome sampling assay (WGSA) is based on a simple restriction enzyme digestion, followed by linker-ligation of a common known adaptor sequence to every fragment generated by digestion. Using a single primer complementary to this adaptor, multiple loci can be simultaneously amplified. Hence, the genomic DNA is converted by PCR into amplified fragments of reduced complexity that can be easily hybridised to the SNP arrays (Figure 5.3). This method was aided by the completion of the human genome project enabling prediction of restriction fragments that should amplify consistently and SNPs that reside within these fragments. Oligonucleotides corresponding to these SNPs are then synthesised onto high-density microarrays. The *Xba*I 50K mapping array has 58,960 SNPs spaced across the whole-genome with 99.1% of the genome within 500Kb of a SNP, 91.6% within 100Kb and 40% within 10kb (Matsuzaki, *et al* 2004). For each SNP there are 40 oligonucleotides (20 perfect matches and 20 mismatches) these are

scattered across the chip in order to minimise any problems arising from patchy hybridisation.

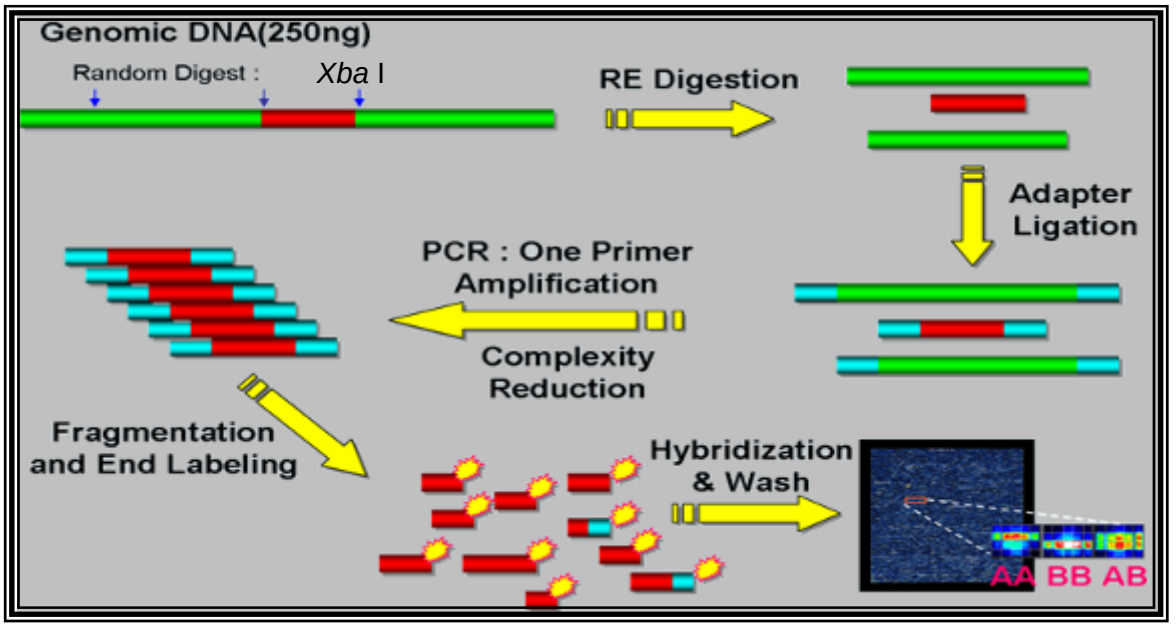


Figure 5. 3: Schematic representing in brief how the Affymetrix SNP chips work adapted from the website of ImaGenes GmbH, Berlin, Germany. Genomic DNA is first digested using a restriction enzyme that cut regularly throughout the genome, such as *Xba*I or *Hind*III. An adapter sequence (represented in blue) is then ligated on to both ends of every fragment. A primer is designed to the adapter sequence that is capable of amplifying all the digested and ligated fragments. These fragments are then end labeled and hybridised to the chip. The results are typed as homozygous (AA or BB) or heterozygous (AB), and a measure of the intensity of the hybridisation signal relates to the copy number of the SNP.

We do not have the local facility to perform Affymetrix analysis ourselves and therefore the laboratory elements were performed as a service by the RZPD in Berlin (now ImaGenes GmbH). The RZPD performed the labelling and hybridisation to the *Xba*I 50K Affymetrix GeneChip SNP array using our samples. After hybridisation, the array was washed and stained using an Affymetrix Fluidics station. The staining protocol is designed to amplify the signal from the annealed probe. Scanning was

then carried out using an Agilent GeneArray Scanner and the raw data was provided to us as a cabinet (.CAB) file of approximately 50Mb for each sample.

The .CAB file is first read by the Affymetrix GeneChip® Operating System (GCOS) in conjunction with the library files for the 50K mapping array, which contain the details of all SNP and probe locations (http://www.affymetrix.com/index.affx). Further analysis to genotype each call was then carried out using Affymetrix GeneChip® Genotyping Analysis Software (GTYPE 4.1). This is specifically designed to give highly accurate, automated SNP allele calls for the mapping arrays, including the 50K SNP array. The programme includes a new genotyping algorithm, designated BRLMM that produces a quality score for each genotyping call (Rabbee and Speed 2006). Copy-number estimations were then determined using the Affymetrix GeneChip® Chromosome Copy Number Analysis Tool (CNAT 4.0) which is integrated into GTYPE 4.1. The copy number algorithm jointly uses genotype information and perfect match (PM) probe intensity and discrimination ratios between paired PM and mismatched (MM) probe intensity values to identify and estimate genetic copy number changes. Values from an experimental sample are compared with SNP-specific distributions derived from a reference set containing over 100 normal individuals (provided by Affymetrix) to gain statistical power. Once the data was fully analysed using Affymetrix GeneChip® software, it was exported to Excel spreadsheets developed in Newcastle by Professor Andy Hall which displays LOH and copy number changes in ideogram format for individual samples or across multiple samples for the ready detection of common areas of deletion.

5.2.3 Patient samples

DNA from specific patient groups was selected, quantified, tested by agarose gel electrophoresis to check for degradation and 250ng sent to the RZPD. The samples I sent for analysis (in four batches) are shown in Table 5.3. For the cases of CML in blast crisis (CML-BC) cases, samples were selected that had at least 75% blasts. No specific criteria were employed for other patient subgroups, although for

aCML/aMPD I focused preferentially on cases with a high leucocyte count, i.e > $20x10^9$/L.

Leukaemia subgroup	**Nos. sent for analysis**
aCML/aMPD	30
CML: myeloid blast crisis	10
CML: lymphoid blast crisis	10
JAK2 V617F negative IMF	18
JAK2 V617F positive IMF	10
JAK2 V617F negative PV	10

Table 5. 1: Number of samples sent to the RZPD for 50K SNP chip analysis from each leukaemic subgroup.

Although my main aim was to investigate the pathogenesis of aMPD/aCML patients, I also took the opportunity to investigate other related patient groups for whom the pathogenesis was completely or largely unknown.

5.2.4 SNP array results

5.2.4.1 Overview of results

Data was returned from all samples along with a series of quality metrics, the most important of which was the SNP call rate. Overall a median of 98.2% SNPs gave readable calls with a range from 91.5% - 99.6% for individual samples. Only 6 samples were below 95%. Analysis of the data led to the identification of copy number and LOH anomalies and the results are summarised in detail in Table 5.4

below. Using the analysis software it was easy to observe the different causes of LOH when the affected regions were relatively large, i.e. deletions, amplifications or copy number neutral loss (aUPD). In Table 5.4 these have been classified as copy number (CN) changes and acquired uniparental disomy (aUPD).

The main aim of my study was to identify large regions of aUPD as an indicator towards novel oncogenes or tumour suppressor genes. However since we did not have constitutional DNA from most of our cases we could not formally prove that regions of homozygosity were acquired and not inherited. Instead we inferred that they are highly likely to be acquired from the patterns of homozygosity seen in normal individuals. A recent study used SNP arrays to genotype 276 DNA samples derived from lymphocytes of elderly neurologically normal subjects. They found extended regions of homozygosity (contiguous tracts >5Mb) in 9.5% patients but structural genomic variations in 66.9%. Of the patients exhibiting extended homozygosity, 42.3% showed multiple tracts ranging from 5-39Mb that were attributed to parental consanguinity. Regions of homozygosity were rarely greater than 15-20Mb in size (Simon-Sanchez, *et al* 2007). To avoid regions that may be inherited rather than acquired, we therefore focused primarily on larger regions of copy neutral LOH, which we arbitrarily defined as greater than 20Mb running to a telomere (Figure 5.4), although regions >10Mb running to a telomere were also considered. Smaller, interstitial regions of LOH were also identified within most subgroups however, as previously mentioned, the lack of constitutional matched DNA for any of the samples made it impossible to know if these were acquired or inherited. I therefore discounted these regions from the initial analysis.

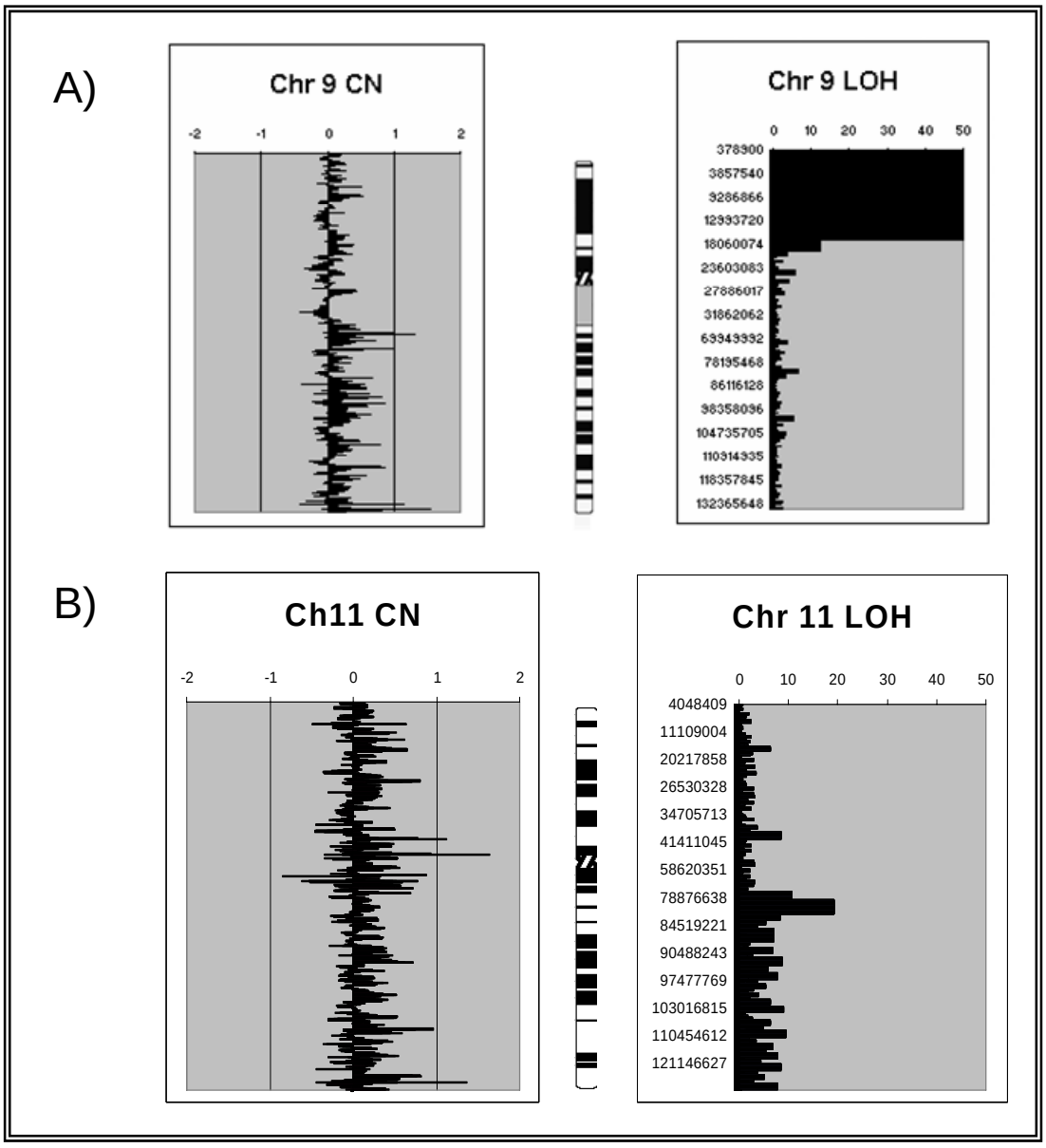


Figure 5. 1: Ideograms showing UPD. A) The graph on the left shows the copy number trace and indicates no large abnormalities. The graph on the right shows zygosity and shows a large block of homozygous SNP calls at 9p. Since there is no copy number change it is likely that the homozygosity arose by aUPD; B) Apparently patchy 11q aUPD observed in some patients that may to be due to poor DNA quality or a high proportion of background normal cells affecting the SNP calls.

Leukaemic subgroup	**Identifier**	**aUPD**	**CN imbalances**
aCML/aMPD	E110, E129, E639	7q	
	P951	1p (~15Mb)	
	E160	17q	
	E484, E632, E1191 (granulocyte pellet)	11q	
	E632	8p (~15Mb)	
	P1304	20q	
	E180	whole chr 13	
	E155		
	E1445	21q	
	P1213		Del 20q, 7q, Del 17q
	E674		Del Chr 7
	P981, E139		Del 8p
Myeloid blast crisis	mBC3	15p (~10Mb)	Del 17p
	mBC6		Del 2q, Del 11p
	mBC10		Del 7q
Lymphoid blast crisis	LyBC1		Gain Chr1, Del 16q
	LyBC5		Del 9p, Del 11q and Del 20q.
	LyBC8		Del 7p, Del 9p
	LyBC10		Del 16q, Gain 1q

	LyBC11		Del 7p, Del 9p
***JAK2* V617F negative IMF**	E996	1pUPD	
	E887	9pUPD	
	04/1345	11qUPD	
	04/4885		Gain 21q, Gain 22
	01/4124		Gain 1q, Del 7q
	01/1755		Gain Chr8
***JAK2* V617F positive (homozygous) IMF**	E1936	9p	
	E1530	9p	
	E1302	9p	
	E1190	9p	
	E904	9p	Del 20q
	0604544	9p	
	0512361	Low level 9p	
	00/1180	9p and 5q	
	04/3974	9p	

Table 5. 2: Summary of SNP array results. aUPD was scored as regions of LOH greater than 20Mb running to a telomere with no change in copy number. Regions of extended copy neutral homozygosity that were between 10Mb and 20Mb but which extended to a telomere are also shown. Copy number (CN) anomalies (deletions and amplifications) are also shown.

Cytogenetic information was only available for some cases and in general the correspondence between visible cytogenetic abnormalities and SNP array results, particularly deletions, was very good. However since the main focus of my analysis was to identify regions of aUPD and other groups have already shown a good correlation between array results and karyotype, I did not perform a detailed comparison of the two techniques. Candidate regions of aUPD for each patient subgroup are considered in detail in the following sections.

5.2.4.2 JAK2 V617F negative Polycythaemia Vera

About 5% of PV patients are negative for the V617F *JAK2* mutation. By screening a small subset of V617F negative patients I hoped to identify another region of recurrent aUPD or imbalance that might indicate a novel chromosomal area harbouring a new candidate gene. However I observed no regions of aUPD or copy number changes in the 10 cases I analysed.

Subsequently it has emerged that many V617F negative PV cases have mutations in exon 12 of the *JAK2* gene (Butcher, *et al* 2008, Pietra, *et al* 2008, Scott, *et al* 2007a, Scott, *et al* 2007b, Williams, *et al* 2007) and further analysis by our group indicated that two of the cases analysed were exon 12 mutation positive. However in contrast to V617F, these exon 12 mutations are not generally reduced to homozygosity by aUPD.

5.2.4.3 JAK2 V617F negative myelofibrosis

Of the 18 cases analysed, two were apparently positive for regions on 1p and 9p, respectively. The 9p abnormality was a surprise as this is the location of *JAK2*, and negativity for V617F was confirmed. One possibility was that the homozygous tract was marking the presence of a novel *JAK2* mutation, however since the patient was local a mouthbrush sample was obtained in order to see if the small region of homozygosity was really acquired. Rather than performing expensive array analysis

on the mouthbrush sample, microsatellite PCR markers were used to assess the region of 9p LOH in both the mouthbrush DNA sample, the PB DNA sample as well as three normal controls (Table 5.5).

Chr location	Marker	E887 (PB)	E2357 (MOUTHBRUSH)	NORMAL1	NORMAL2	NORMAL3
TEL	D9S1779	HOM	HOM	HET	HOM	HOM
9p24.3	D9S913	HOM	HOM	HOM	HET	HET
9p24.3	D9S1858	HOM	HOM	HET	HET	HET
9p24.3	D9S143	HOM	HOM	HET	HOM	HOM
9p24.2	D9S1871	HET	HET	HET	HET	HET
9p24.2	D9S228	HOM	HOM	HOM	HOM	HET
9p24.2	D9S288	HET	HET	HET	HET	HET
9p24.2	D9S178	FAIL	FAIL	FAIL	FAIL	FAIL
9p24.2	D9S1813	HET	HET	HET	HET	HET
9p24.1	JAK2					

Table 5. 3: Results of microsatellite analysis on the MF case with 9p homozygosity.

This analysis revealed first that the region of homozygosity did not include *JAK2* (something that was also obvious from the SNP array results) and second, that the region is also present in the mouthbrush DNA and is thus inherited not acquired.

The remaining patient, E996 had 1p UPD (Figure 5.5A). At about this time, a report by Pikman et. al., (2006) detailed the discovery of a mutation in the thrombopoietin receptor (MPL). The somatic activating mutation in the transmembrane domain of MPL (W515L or W515K) was found in 9% of the *JAK2* V617F-negative MF patients. Subsequently other variants were found that affected the same residue. The reason this was of interest to me is that this gene is located on chromosome 1p. I investigated the region of 1p UPD in this patient to confirm if this region harboured this gene and indeed this was the case.

We had an assay set up for investigating this mutation by pyrosequencing so I used this to initially test for W515L or W515K mutations (Figure 5.5B). The pyrosequencing trace shows my patient (E996) has a significantly higher percentage of the ‘T’ allele compared to the normal control indicating the G>T mutation

(W515L). This was further confirmed by sequencing (Figure 5.5C), which clearly indicates that the normal G has been completely replaced by a T. This therefore provides another example of aUPD being associated with a particular oncogenic change.

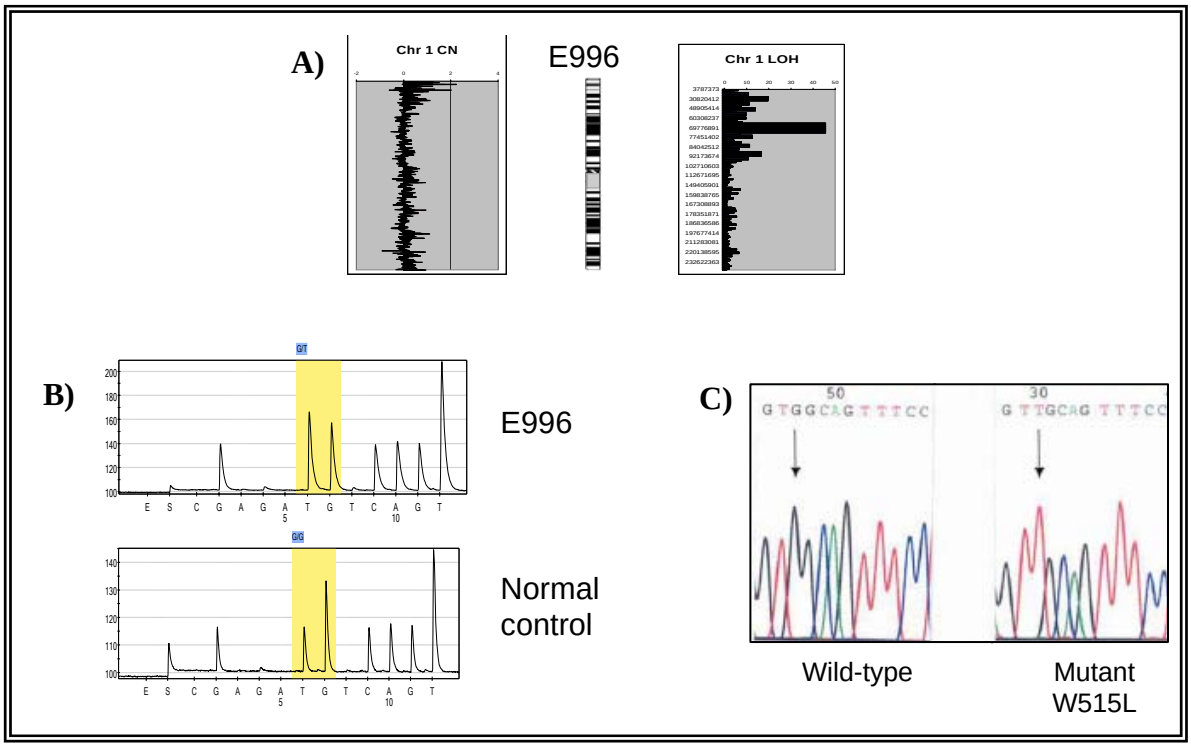


Figure 5. 5: A) the two graphs generated by the customised excel tool supplied by Professor Hall detailing the SNP chip results for chromosome one in this patient (E996). The graph on the left shows the copy number, whilst the graph on the right details regions of loss of heterozygosity; B) shows the pyrograms for E996 (above) versus a normal control (below). E996 shows a greater proportion of the T allele compared to the normal control indicating the G>T mutation; C) the mutation was further confirmed by sequencing. The trace on the left shows the wild-type sequence and the trace on the right E996 showing the homozygous G>T mutation.

5.2.4.4 JAK2 V617F positive myelofibrosis

It is unclear why the V617F mutation is associated with distinct clinical phenotypes, (PV, ET and MF) but it has been suggested that these are multistep diseases with other genetic abnormalities that remain to be identified. In particular, PV and ET may evolve to MF and the appearance of new cytogenetic abnormalities during disease evolution suggests that the transition may be driven by new mutations. I therefore examined 10 V617F positive MF cases with high levels of the mutant clone (>50%) to see if any other abnormalities were apparent.

Of the 10 *JAK2* cases, 9 exhibited 9p LOH, all of which showed no change in copy number. The single case that did not show 9p LOH was only 53% V617F positive and hence was probably heterozygous. In addition to 9p, one case also had clear 5q aUPD (Figure 5.6). There are no known haematologically relevant oncogenes at 5q and currently the significance of this finding remains unclear.

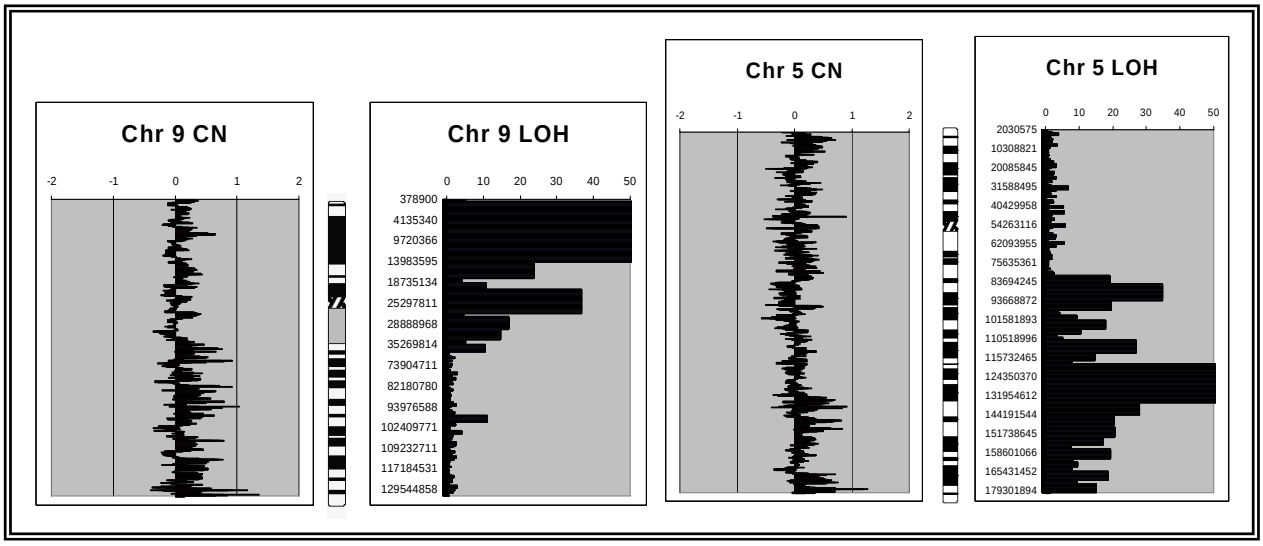


Figure 5.6: Figure showing the SNP chip results for Patient 00/1180. The patient shows 9p aUPD and 5q aUPD determined by the large region of LOH on the graphs on the right, compared to the left-hand graphs depicting copy number changes of which there are none.

In addition, a second patient had an additional copy number imbalance, specifically deletion of 20q. Deletion of 20q is a known recurrent finding in a range of MPDs.

5.2.4.5 CML in transformation

Although *BCR-ABL* is probably the sole abnormality that drives chronic phase CML, it is widely believed that a series of additional genetic events precipitate transformation to acute leukaemia, or blast crisis. I analysed 20 cases (10 in lymphoid transformation and 10 in myeloid transformation) to see if these putative changes are associated with aUPD. In fact, none of the lymphoid cases showed copy neutral LOH and just one of the myeloid cases showed one possible region at the tip of 15p that was approximately 10 Mb in size. Unfortunately there was no chronic phase or constitutional DNA available to confirm that this was acquired. Several copy number changes were identified that correlated with the results of chromosome analysis.

5.2.4.6 aCML/aMPDs

Strikingly, this subgroup showed a high frequency of aUPD compared to other patient groups. Of the 30 cases analysed, 12 (40%) were affected by aUPD (including two cases with regions estimated to be 15 Mb in size). In one case (E1191) it was possible to show that the abnormality was indeed acquired by array analysis of DNA extracted from mononuclear cells (MNC; includes B-cells and T-cells that are not usually involved in the disease) and DNA extracted from granulocytes pellet (which would be expected to be predominantly clonal). As shown on Figure 5.7, 11q LOH is clearly visible in granulocyte DNA but is not visible or barely detectable in MNC DNA.

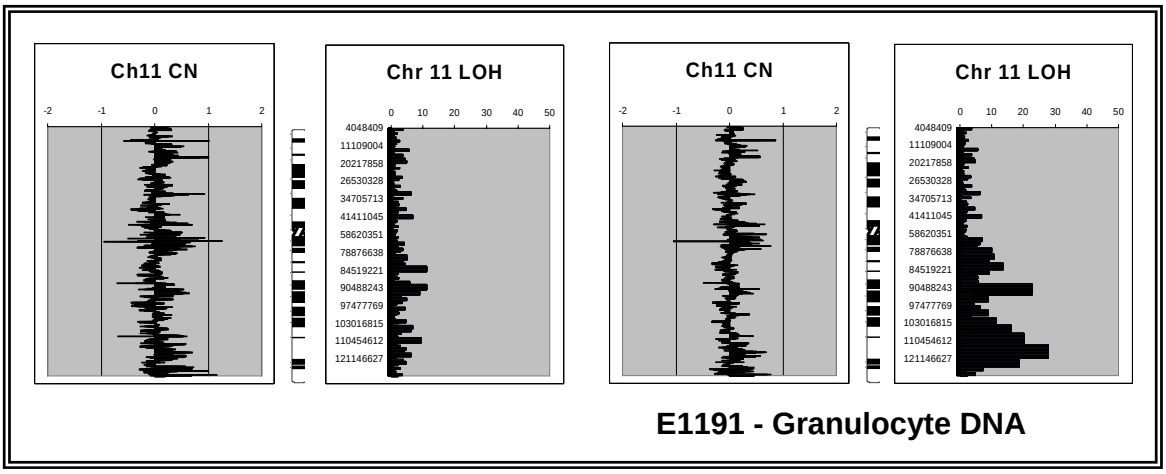


Figure 5.7: Comparison of MNC and granulocyte DNA for patient E1191. No copy number changes are observed in either DNA sample (first and third panels) however clear 11q LOH is observed in the DNA extracted from granulocytes (fourth panel) that is absent in MNC DNA (second panel), indicating an acquired abnormality.

In total, eight different chromosomal regions were affected by aUPD suggesting substantial genetic heterogeneity in these diseases. Recurrent regions were identified at 7q and 11q (3 cases of each, i.e. 10% of cases) and single cases with aUPD were seen at 1p, 8p, whole chromosome 13, 17q, 20q and 21q.

One patient, E632, showed both whole arm 8p aUPD and 11q aUPD. This patient also showed many regions of interstitial LOH, far more than any other patient (Figure 5.8). Limited clinical history suggests this patient may be consanguineous, suggesting that at least some of these multiple homozygous regions may be autozygous, i.e. inherited rather than acquired.

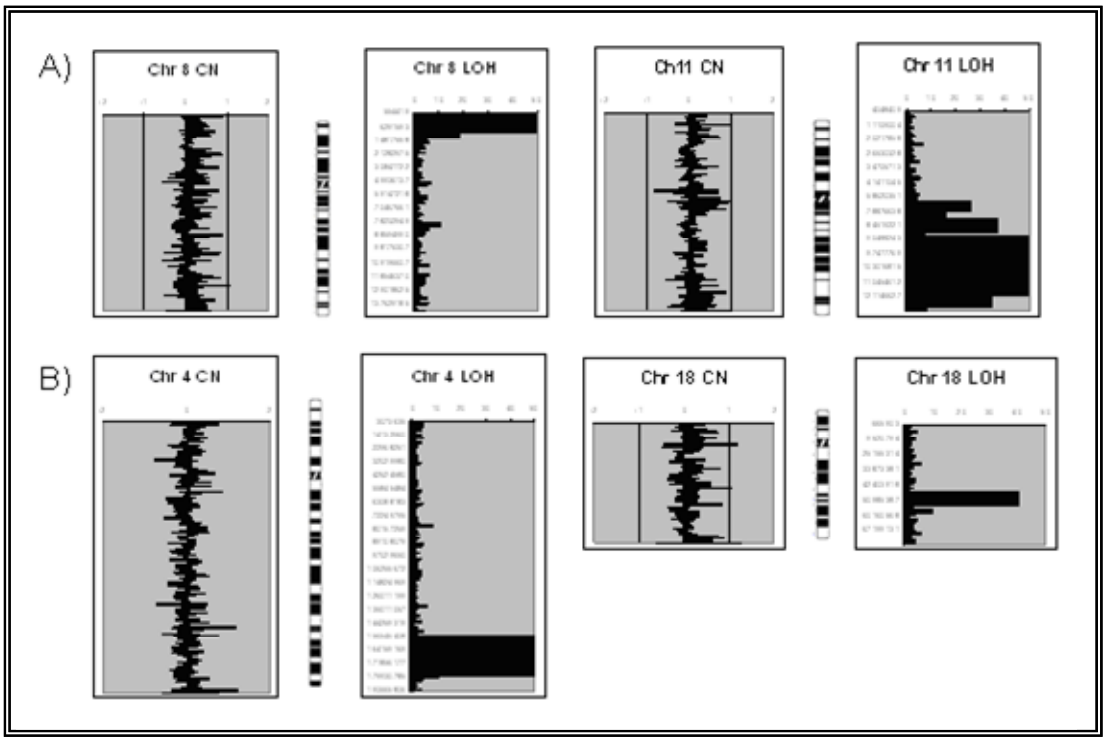


Figure 5. 8: Ideograms depicting E632 SNP array results. A) Top panels show tracts of homozygosity at 11q and 8p, both of which extend to a telomere; B) bottom panels show an example of two of the ten interstitial regions of homozygosity.

One patient (E180) showed whole chromosome 13 aUPD (Figure 5.9a). The receptor tyrosine kinase *FLT3* is located on this chromosome and is the most frequently mutated gene in acute myelogenous leukemia (AML), usually associated with an internal tandem duplication (ITD) of the region encoding the juxtamembrane domain. Furthermore it has already been reported that aUPD for chromosome 13 is associated with *FLT3* mutations in AML (Griffiths, *et al* 2005). The ITD is within exon 14 and/or 15 of the *FLT3* gene, so primers were designed initially to amplify exon 14. As shown on (Figure 5.9b), patient E180 is indeed homozygous for a *FLT3* ITD.

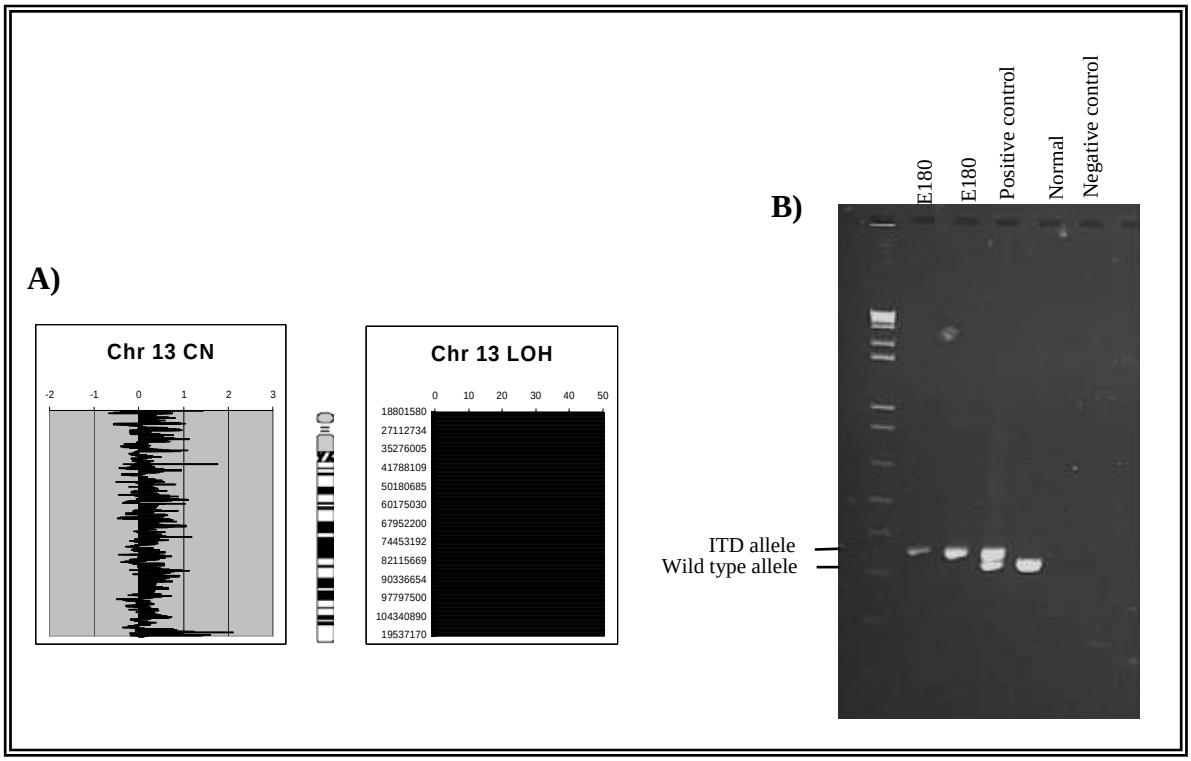


Figure 5. 9: A) Analysis of aMPD patient E180. The graph on the left shows there has been no copy number change, whilst the graph on the right indicates complete chromosome 13 homozygosity; B) PCR assay for FLT3 ITD. Lanes 1 & 2 represent patient E180, lane 3 is a positive control (an AML patient with a known FLT3 ITD and lane 4 is a normal control. The bottom band is the normal FLT3, the top band is FLT3 with an ITD.

Other regions of aUPD could not be simply linked to known mutations. My analysis focused on the recurrent regions at 7q and 11q and these are discussed in detail below. The abnormalities at 1p, 8p, 17q, 20q and 21q were not investigated. The region at 17q may target *NF1*, a gene that has long been associated with juvenile myelomonocytic leukaemia (JMML) (Side, *et al* 1998). More recently, Flotho et. al., (2007) screened JMML patients and found regions (>55Kb) of copy number neutral 17q LOH leading to loss of the normal *NF1* allele that was associated with an increased incidence of JMML in NF1 patients. However, *NF1* is a large gene with 58 exons and therefore very expensive to screen for mutations (Flotho, *et al* 2007).

In addition to large regions of copy number neutral LOH, copy number imbalances were also observed in aMPD cases. One patient exhibited three deletions of the long arms of chromosomes 20, 17 and 7. As discussed previously deletions of chromosome 20q are a common abnormality associated with myeloid disorders as are deletions of 7q (Luna-Fineman, *et al* 1995). The 17q deletion may again be associated with an *NF1* mutation. Another patient was also found to have a deletion of entire chromosome 7, and two patients had a deletion of 8p. Deletions of 8p have been described previously in association with MDS and the evolution of CML (Kerndrup, *et al* 1987, Pedersen and Kerndrup 1986, Wada, *et al* 1994).

5.3 RECURRENT aUPD OF CHROMOSOME 7q

5.3.1 7q abnormalities in myeloid disorders

LOH for chromosome 7, commonly in the form of monosomy 7 or deletion of the long arm of chromosome 7 is a very frequent occurrence in a variety of human malignancies including gastric cancer (Weng, *et al* 2006), breast cancer (Zenklusen, *et al* 1994), prostate cancer (Latil, *et al* 1995), colon cancer and primary cell carcinoma of the head and neck (Zenklusen, *et al* 1995). It has been hypothesised that a tumour suppressor gene is located at 7q that is ubiquitously expressed thereby accounting for the numerous cancers associated with abnormalities of chromosome 7.

More relevant to my work is the consistent reports of chromosome 7 LOH in myeloid abnormalities (Dohner, *et al* 1998, Fischer, *et al* 1997, Kere, *et al* 1989, Kere, *et al* 1987, Le Beau, *et al* 1996, Lewis, *et al* 1996, Liang, *et al* 1998, Tosi, *et al* 1999). Monosomy 7 and del(7q) are frequently found associated with MDS, AML and juvenile CML (Aktas, *et al* 1999, Hasle, *et al* 1999, Mantadakis, *et al* 1999, Tosi, *et al* 1996), as well as therapy related MDS (t-MDS) and therapy related AML (t-AML) occurring after treatment with alkylating agents or exposure to certain carcinogens (Rowley 1976). Although much work has been performed in defining critical regions of LOH based on deletions and translocations, no group has yet identified a definitive

tumor suppressor gene. Furthermore, at least three distinct minimal regions of LOH have been identified suggesting that multiple loci may be important. Some have also argued that the complexity of 7q rearrangements might indicate that this region is prone to breakage and instability as a consequence of different genetic factors that play a general role in oncogenesis rather than the presence of a specific tumour suppressor gene involved in myeloid disorders (Basirico, *et al* 2003, Gonzalez, *et al* 2004).

5.3.2 Recurrent 7q aUPD identified by SNP chip analysis

As a result of screening 30 aCML/aMPD patients using Affymetrix 50K SNP arrays I identified three cases with varying sized regions of aUPD of chromosome 7q (Figure 5.10).

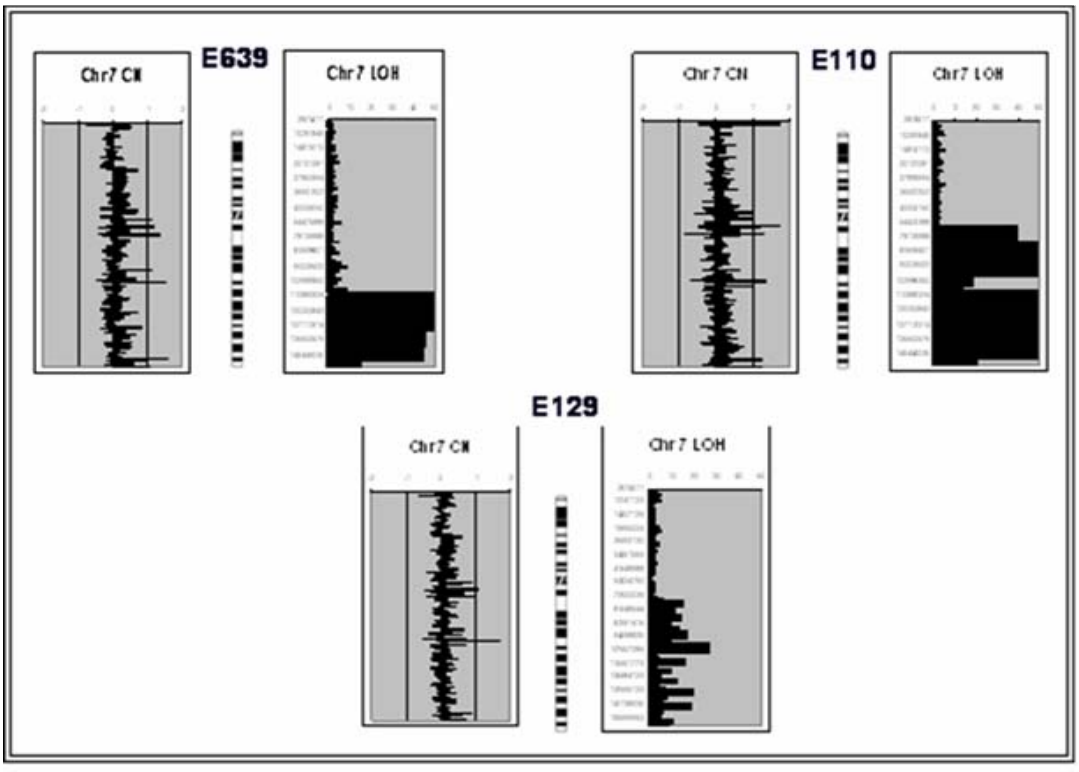
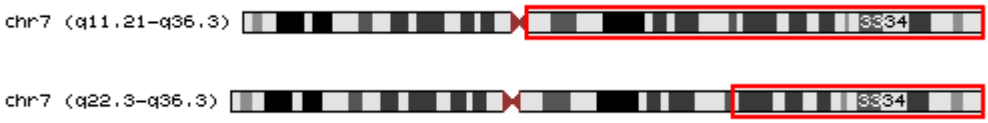


Figure 5.10: Copy number and LOH graphs produced for chromosome 7 for three aCML/aMPD patients: E639, E129 and E110. In all three cases the copy number graph shows no

copy number changes. All three LOH charts show varying sized regions of LOH affecting the q arm of chromosome 7.

No copy number alterations or aUPD was observed elsewhere in the genome of these patients. Cytogenetics was only known for E639 and this was normal. Patients E110 and E129 had regions of aUPD spanning approximately 98Mb, almost the entire length of the long arm of chromosome 7, specifically 7q11.21-tel and 7q11.22-tel, respectively (Figure 5.11). The third patient has a smaller region of partial 7q aUPD spanning 53Mb from cytogenetic bands 7q22.3-tel. This region represents the minimal overlapping region of aUPD observed in all three patients (Figure 5.11).



Patient Identifier	Region of LOH/Chromosomal region	Size of region of LOH
E110	63,695,332-158,554,645 (q11.21-qter)	94,859,313 (95Mb)
E129	71,178,702-158,554,645 (q11.22-qter)	87,375,943 (87Mb)
E639	105,954,360-158,554,645 (q22.3-qter)	52,600,285 (53Mb)

Figure 5. 11: Schematic representation of chromosome 7 and the regions of aUPD (highlighted in red) observed in the three aCML/aMPD patients E110, E129 and E639.

5.3.3 Screening for mutations in 7q aUPD patients

The minimal region 53Mb is still very large and encompasses approximately 400 genes so prioritising candidate genes for screening was of primary concern. Due to the growing evidence linking aberrant tyrosine kinase signalling to the MPDs, I began by screening genes encoding tyrosine kinases and hoped to find mutations that were homozygous as a consequence of the aUPD.

5.3.3.1 Candidate tyrosine kinases located within the region of 7qUPD

There are 3 tyrosine kinase genes in the minimal region, *MET* (7q31.2), *EPHA1* (7q34-35) and *EPHB6* (7q34). The *MET* proto-oncogene, encoding the tyrosine kinase receptor for the hepatocyte growth factor (HGF), controls cell growth, invasiveness, and protection from apoptosis. Deregulated MET signalling has long been associated with tumour progression and metastasis in a variety of tumours (Jeffers, *et al* 1996), as well as being deregulated in a few cases of human leukaemia and lymphoma (Jucker, *et al* 1994). Activating *MET* mutations have also been identified in hereditary and sporadic forms of papillary renal carcinoma (Schmidt, *et al* 1997). Recently, missense mutations have also been identified in this gene in a variety of human cancers (Lee, *et al* 2000, Ma, *et al* 2003). This seemed like an ideal candidate to screen for possible mutations in these three patients with 7q UPD.

The Ephrin receptors and their ligands, the ephrins, regulate numerous developmental and cellular processes, particularly in the nervous system. They are divided into two groups based on their sequence similarities and structures, the ephrin-A (EFNA) class, which are anchored to the membrane by a glycosylphosphatidylinositol linkage, and the ephrin-B (EFNB) class, which are transmembrane proteins. They make up the largest subgroup of receptor tyrosine kinases. Although relatively little is known about these receptors over-expression and loss of expression of some family members have been related to specific cancers (Noren, *et al* 2006, Song, *et al* 2007).

5.3.3.2 Expression analysis

It would be expected that any relevant mutated gene would be expressed in haematopoietic progenitor cells and probably in myeloid cells. As a first step towards prioritising these three genes, the expression of each was checked by RT-PCR using cDNA made from peripheral blood (PB) and bone marrow (BM) from 5 cases with other haematological disorders. cDNA-specific primers were designed to amplify the kinase domain (See Appendix I, section XXI) for all three receptor tyrosine kinases. Standard PCR conditions were used (annealing temperature of 66°C and 30 cycles). Clear products for *EPHA1* and *EPHB6* were seen in all samples but no amplification was seen for *MET*. I therefore focused initially on analysis of the two Ephrin receptors.

Genomic DNA primers were designed, again to anneal at 66°C, to cover exons encoding the transmembrane and kinase domain as these are known hotspots for oncogenic mutations (for all primer sequences see Appendix I, section XXII). PCR products were directly sequenced and compared to the RefSeq entries for each gene. No mutations were identified so further primers were designed to cover all exons of each gene. As summarised on Table 5.6, no abnormalities were detected.

Exon number	**E110**	**E129**	**E639**
EPHA1.5'UTR.F	N	N	N
EPHA1.Int1.1F	N	N	N
EPHA1.Int2.1F	N	N	N
EPHA1.Int3.1F	T>C @479 HOM V160A	T>C @479 HOM V160A	T>C @479 HOM V160A
EPHA1.Int4.1F	N	N	N
EPHA1.Int5.1F	N	N	N
EPHA1.Int6.1F	N	N	N
EPHA1.Int7.1F	N	N	N
EPHA1.Int8.1F	N	N	N
EPHA1.Int9.1F	N	N	N
EPHA1.Int10.1F	N	N	N
EPHA1.Int11.1F	N	N	N
EPHA1.Int12.1F	N	N	N
EPHA1.Int13.1F	N	N	N
EPHA1.Int14.1F	N	N	N
EPHA1.Int15.1F	N	N	N
EPHA1.Int16.1F	N	N	N

EPHA1.Int17.1F	N	N	N
EPHB6.Int 5'UTR.1F	N	N	N
EPHB6.Int 1.1F	N	N	N
EPHB6.Int 2.1F	N	N	N
EPHB6.Int3.1F	N	N	N
EPHB6.Ex4.1F	N	N	N
EPHB6.Int4.1F	N	N	N
EPHB6.Int5.1F	N	N	N
EPHB6.Int6.1F	N	N	N
EPHB6.Int7.1F	N	N	N
EPHB6.Int8.1F	N	N	N
EPHB6.Int9.1F	N	N	N
EPHB6.Int10.1F	N	N	N
EPHB6.Int11.1F	N	N	N
EPHB6.Int12.1F	N	N	N
EPHB6.In13.1F	N	N	N
EPHB6.Int14.1F	N	N	N
EPHB6.Int15.1F	N	N	N
EPHB6.Int16.1F	N	N	N

Table 5. 6: Mutation screening of EPHA1 and EPHB6. The variant found in all three patients in exon 4 of EPHA1 is a known SNP reported in various databases. Exon 1 of EPHA1 proved difficult to amplify, so it was cloned and ten colonies sequenced for each patient.

Although the *MET* tyrosine kinase was not obviously expressed in the PB or BM of controls, the possibility that its expression might be limited to a small population of myeloid precursor cells plus the strong association of activated tyrosine kinases with MPDs persuaded me that I should screen this gene as well. Genomic DNA primers were designed for the entire coding region of *MET* (Appendix I, section XXII). An identical intronic sequence change was identified upstream of exon 6 but was ruled out as causative due to the location, 60bp away from the exon start site. Other variants were identified in exons 20 and 21 but corresponded to known reported SNPs. The results are summarised in Table 5.7.

Exon no.	Exon	E110	E129	E639
Int1.F	2	N	N	N
Ex2F	2	N	N	N
Int2.F	3	N	N	N
Int3.F	4	N	N	N
Int4.F	5	N	N	N
Int5.F	6	82799C>T (Intronic)	82799C>T (Intronic)	82799C>T (Intronic)
Int6.F	7	N	N	N
Int7.F	8	N	N	N
Int8.F	9	N	N	N
Int9.F	10	N	N	N
Int10.F	11	N	N	N
Int12.R	12	N	N	N
Int12.F	13	N	N	N
Int13.F	14	9 clones = N	7 clones = N	9 clones = N
Int14.F	15	N	N	N
Int15.F	16	N	N	N
Int16.F	17	N	N	N
Int17.F	18	N	N	N
Int18.F	19	N	N	N
Int19.F	20	123310 C>T: 1304 D>D	123310 C>T: 1304 D>D	123310 C>T: 1304 D>D
Int20.F	21	HOM G>A: 1357 A>A & HOM G>A: 1382 P>P	HET G>A: 1357 A>A & HET G>A: 1382 P>P	HOM G>A: 1357 A>A & HET G>A: 1382 P>P

Table 5. 7: Mutational analysis of the MET gene in the three aMPD patients with 7q aUPD. All exons were directly sequenced apart from exon 14 which was subcloned and then sequenced.

Although it is conceivable that either *EPHA1*, *EPHB6* or *MET* could be activated by another mechanism which cannot be detected by the strategy I used, e.g. a deep intronic mutation giving rise to a splicing abnormality or an internal deletion/duplication that encompasses at least one of the primer binding sites, it is much more likely that these genes are normal. Since these are the only tyrosine kinases on 7q, I therefore focused on the prioritisation of other candidate genes.

5.3.3.3 Other candidate genes

To prioritise further genes I decided to look at those located within the minimal region of aUPD that impact on tyrosine kinase signalling pathways and for which there was a known link with leukaemia or cancer. There were no obvious candidates specifically implicated in leukaemia but there were two genes encoding signalling components more generally associated with cancer: *BRAF* and *PIK3CG*.

The *BRAF* gene is located at 7q34 and seemed an obvious candidate as it encodes a member of the RAS-RAF-MEK-ERK-MAP kinase signalling pathway which is responsible for mediating the transduction of mitogenic signals from the cell membrane to the nucleus. It is a cytoplasmic member of the serine-threonine protein kinase family that is regulated by binding to RAS. Davies et. al., (2002) were the first to report mutations in this gene by high throughput sequence analysis. They screened genomic DNA from a number of different cancer cell lines and identified *BRAF* somatic missense mutations in a wide variety of cancers, notably malignant melanoma (66% of cases). All mutations were within the kinase domain, with a single missense mutation (V599E) accounting for over 80% of the mutations observed. Of 53 leukaemia/lymphoma cell lines analysed no mutations were identified. However, it has been suggested that these mutations are commonly found in cancers that are also known to have a high frequency of *RAS* mutations (Davies, *et al* 2002). *RAS* mutations are seen in a substantial proportion of JMML and CMML patients (Bowen, *et al* 2005, Emanuel 2004, Emanuel 2008). Additionally, Christiansen et. al., (2005) screened 29 AML patients and identified 3 (21%) with *BRAF* mutations so it was entirely possibly that mutations in this gene may be found in patients with MPDs. Since the original report, *BRAF* mutations have been identified in a wide variety of human cancers including papillary thyroid carcinomas (Espinosa, *et al* 2007), non-small cell lung carcinomas (Brose, *et al* 2002), and colorectal carcinomas (Oikonomou and Pintzas 2006) with all mutations being located in either exons 11 or 15.

Based on these reports I decided to screen my aMPD patients with 7q aUPD for these two *BRAF* exons. Primers were taken from Brose et. al., (2002) and performed as suggested, using standard PCR conditions with a 58°C annealing temperature and 30 cycles, however no sequence variants were found. These negative results fit with a recent report by de Vries et. al., (2007) who screened 65 JMML patients for mutations of the *BRAF* gene but found no mutations (de Vries, *et al* 2007).

The second candidate I considered was *PIK3CG* (PI3-kinase p110 subunit gamma), located at 7q22.3. *PIK3CG* is highly related to *PIK3CA* at 3q26, which is commonly mutated in colorectal cancer (Oikonomou and Pintzas 2006) and other cancer types including bladder tumours, breast & ovarian cancer, and brain tumours (Bachman, *et al* 2004, Broderick, *et al* 2004, Campbell, *et al* 2004, Lopez-Knowles, *et al* 2006). Most of the mutations are clustered in two hotspots within the gene, as seen in Figure 5.12. The mutations are thought to confer resistance to apoptosis and facilitate tumour invasion of solid tumours.

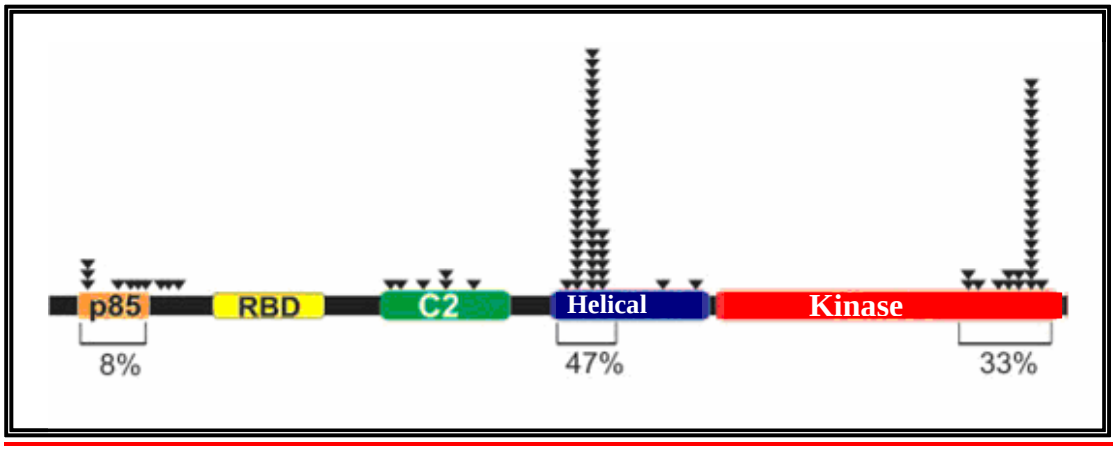


Figure 5. 12: Adapted from Samuels et. al., (2004) (Samuels and Velculescu 2004). Clustering of reported mutations in PIK3CA identified in various human cancers. The arrow heads indicate the location of reported missense mutations. The percentage of mutations identified in each tumour group from this paper is indicated.

Although no mutations have been described in *PIK3CG*, I decided to analyse the regions that were homologous to the hotspots in *PIK3CA*. Following alignment of the

two genes, I designed primers to amplify exons 2 and 11 of *PIK3CG*, which are homologous to exons 9 and 20 of *PIK3CA*. All primers were designed to anneal at 66°C using standard PCR conditions with 30 cycles (Appendix I, section XXIII) but again no mutations were found.

Having eliminated the likely hotspots in *BRAF* and *PIK3CG*, I considered full mutation screens and also created a shortlist of possible new candidate genes for analysis (Table 5.8). However at this point, the analysis of genes within the second recurrent region of aUPD at 11q was generating interesting results and this is where I focused my attention.

Accession No.	**Gene Name**	**Chr. Location**	**Description**
NM_005515	*HLXB9*	q36.3	Homeobox protein HB9. Fused to ETV6 in infant AML. Activates IL6 via PI3K in Hodgkins Lymphoma. Also fused to MYB. Mutations found in Currarino syndrome.
NM_002847; NM_130843; NM_130842	*PTPRN2*	q36.3	Receptor-type tyrosine-protein phosphatase N2 precursor. It is a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signalling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation.
NM_194071	*CREB3L2*	q33-q34	cAMP responsive element binding protein 3-like. The chimeric FUS/CREB3l2 gene is specific for low-grade fibromyxoid sarcoma. It belongs to a family of transcription factors. Fused to FUS in liposarcomas
NM_178562	*TSPAN33*	q32.1	Tetraspanin 33, alternatively known as Penumbra whose mouse homologue has growth suppressive activity
NM_170606; NM_021230	*MLL3*	q36.1	Myeloid/lymphoid or mixed-lineage leukemia 3. It is involved in transcriptional co-activation. May be involved in leukaemogenesis and developmental disorder.
NM_198267; NM_019071	*ING3*	q31.31	Inhibitor of growth protein 3 (p47ING3 protein). The protein is similar to ING1, a tumor suppressor protein that can interact with TP53, inhibit cell growth, and induce apoptosis. This gene can activate p53 trans-activated promoters, including promoters of p21/waf1 and bax. Over-expression of this gene has been shown to inhibit cell growth and induce apoptosis. Allelic loss and reduced

			expression of this gene were detected in head and neck cancers.
NM_004935	*CDK5*	q36.1	Cell division protein kinase 5 (Tau protein kinase II catalytic subunit). Probably involved in the control of the cell cycle.

Table 5. 8: Table showing the gene name, accession number and chromosomal location of some potential candidate genes on 7q. A brief list of known facts relevant to why these genes were short listed is also included.

5.4 Recurrent aUPD of chromosome 11q

5.4.1 11q abnormalities in myeloid disorders

Like chromosome 7, chromosome 11 has also long been reported as a region commonly affected by imbalances in cancer. Both gains and losses of chromosome 11 have been described. Gains of 11q are less common than losses but are often seen in neuroblastoma cell lines (Carr, *et al* 2007). LOH of this region is frequent and associated with cervical, ovarian, breast, colon and skin cancers (Bethwaite, *et al* 1995, Carter, *et al* 1994, Foulkes, *et al* 1993, Hampton, *et al* 1994, Keldysh, *et al* 1993, Tomlinson, *et al* 1993). By far the most frequent region of imbalance reported is centered round the cytogenetic band 11q23, in solid tumours as well as haematological malignancies such as T-cell prolymphocytic leukemia (T-PLL), B-cell chronic lymphocytic leukemia (B-CLL) and mantle cell lymphoma (MCL). Amplifications of specific regions of 11q have also been identified in cases of AML and MDS. One group identified 11q genomic gain in more than 40% of AML cases with high level amplification again centered on 11q23, whilst another group reported the rare but recurrent occurrence of 11q amplifications of 11q23 and 11q13 in AML and MDS cases (Rucker, *et al* 2006, Zatkova, *et al* 2006). 11q LOH associated with leukaemia has also been described (Siu, *et al* 2000, Takeuchi, *et al* 2003).

5.4.2 Recurrent 11q aUPD identified by SNP chip analysis

From screening 30 aMPD patients with Affymetrix 50K SNP arrays I identified three aCML/aMPD patients with varying sized regions of aUPD at chromosome 11q (Figure 5.13).

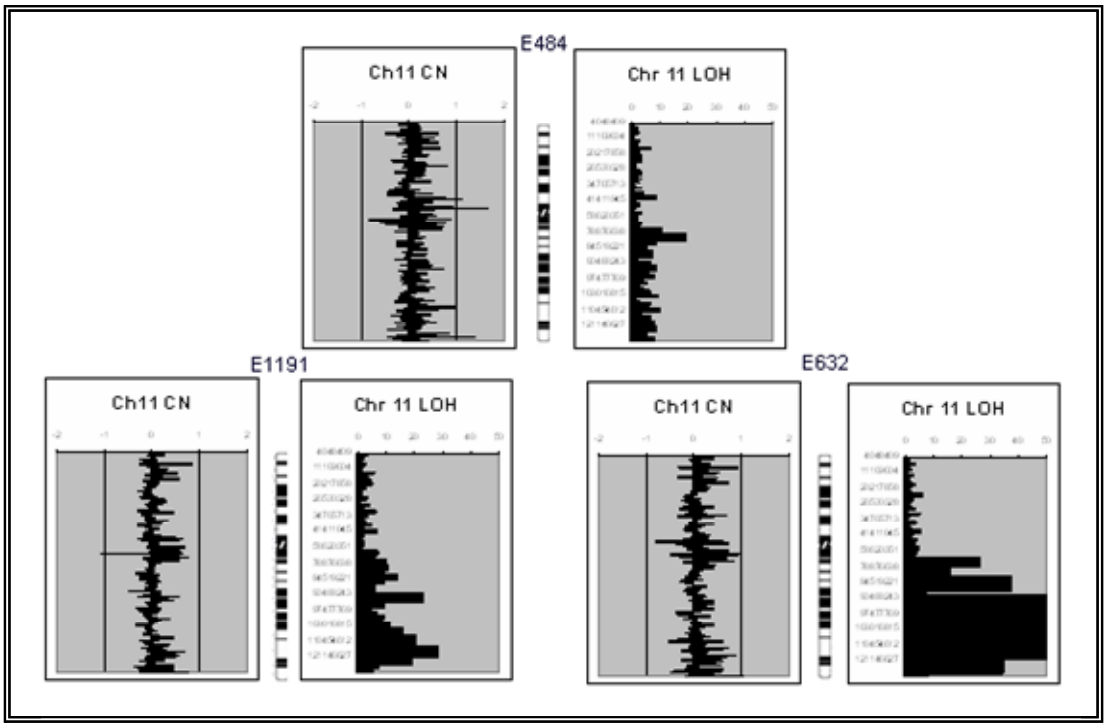
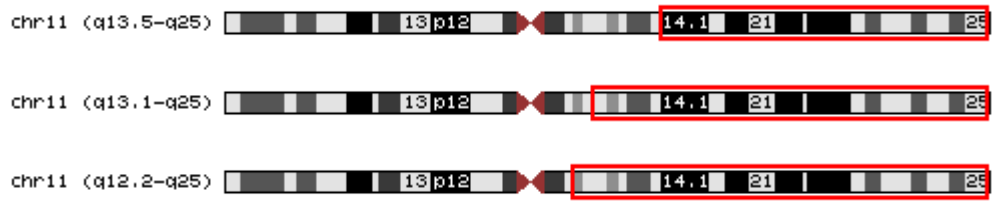


Figure 5. 13: 50K SNP array analysis of chromosome 11 in three aCML/aMPD patients: E484, E632 & E1191. The graphs on the left represent the copy number distribution observed across the chromosome whilst the graphs on the right hand side represent the loss of heterozygosity observed. All three have varying degrees of LOH with no corresponding copy number alterations, indicating aUPD.

Previous cytogenetics was only available for one of the patients, E484, and this was found to be normal. All three patients had partial 11q aUPD of varying size as

illustrated on Figure 5.14. E632 and E1191 had the largest regions spanning approximately 69Mb, almost the entire long arm of chromosome 11. Cytogenetically, these correspond to 11q13.1-qter and 11q12.2-qter in E632 and E1191, respectively (Figure 5.14). E484 had a slightly smaller region of aUPD spanning 57Mb and covering cytogenetic bands 11q13.5-qter. This therefore represents the minimal region and includes about 350 known genes.



Patient Identifier	Region of LOH/Chromosomal region	Size of region of LOH
E484	**76,606,755-134,082,843 (q13.5-qter)**	**57476088 (57Mb)**
E632	**64,698,446 -134,082,843 (q13.1-qter)**	**69384397 (69Mb)**
E1191 (Granulocyte)	**61,241,557-134,082,843 (q12.2-qter)**	**67413720 (67Mb)**

Figure 5. 14: Schematic representation of chromosome 11 and the regions of aUPD observed in the three aCML/aMPD patients E484, E632 & E1191.

5.4.3 Screening 11q aUPD patients

5.4.3.1 Candidate genes located within the region of 11q aUPD

Interestingly no tyrosine kinase genes are located within the minimal region of 11q aUPD, or indeed anywhere on 11q. Instead, I looked for genes associated with leukaemia or cancer in general within the minimal region. As expected there were several possible candidates (Table 5.9). However, only a few really stood out due to

previous reports of defects associated with specific leukaemias or very specific pathway links to receptor tyrosine kinases (those highlighted in yellow).

Accession No.	**Gene Name**	**Chr. Location**	**Description**
NM_002576	*PAK1*	q14.1	p21-activated kinase 1, a serine/threonine kinase.
NM_138292	*ATM*	q22.3	Belongs to the PI3/PI4-kinase family. It is an important cell cycle checkpoint kinase. Defects in *ATM* contribute to T-cell acute lymphoblastic leukemia (TALL) and T-prolymphocytic leukemia (TPLL). Defects in *ATM* also contribute to B-cell non-Hodgkin lymphomas (BNHL), including mantle cell lymphoma (MCL) and B-cell chronic lymphocytic leukemia (BCLL).
NM_003904	*ZNF259*	q23.2	May be a signalling molecule that communicates mitogenic signals from the cytoplasm to the nucleus. Binds to the EGF and PDGF receptors.
NM_005933	*MLL*	q23.3	Myeloid/lymphoid or mixed-lineage leukemia. Possibly acts as a transcriptional regulatory factor. Numerous chromosomal aberrations involving MLL are a cause of acute leukaemias. To name a few, translocation t(1;11)(q21;q23) with MLLT11/AF1Q; translocation t(3;11)(p21;q23) with NCKIPSD/AF3p21; translocation t(11;19)(q23;p13.3) with MLLT1/ENL; translocation t(11;19)(q23;p23) with GAS7; translocation t(X;11)(q13;q23) with MLLT7/AFX1.
NM_005188	*CBL*	q23.3	Participates in signal transduction in haematopoietic cells. Adapter protein that functions as a negative regulator of many signalling pathways that starts from receptors at the cell surface. Acts as an E3 ubiquitin-protein ligase, which accepts ubiquitin from specific E2 ubiquitin-conjugating enzymes, and then transfers it to substrates promoting their degradation by the proteasome. Recognises activated receptor tyrosine kinases, including PDGFA, EGF and CSF1, and terminates signalling. Can be converted to an oncogenic protein by deletions or mutations that disturb its ability to down-regulate RTKs
NM_032873.3	*STS-1*	q24.1	Encodes a protein that contains an ubiquitin associated domain at the N-terminus. The encoded protein was found to inhibit endocytosis of epidermal growth factor receptor (EGFR) and platelet-derived growth factor receptor.
NM_005238	*ETS1*	q24.3	v-ets erythroblastosis virus E26 oncogene. ETS is responsible for erythroblast and fibroblast transformation. The juxtaposition of the interferon and c-ETS-1 proto-oncogene may be involved in the pathogenesis of human monocytic leukemia.

NM_032811	*TBRG1*	q24.2	Transforming growth factor-beta (TGF-beta) is perhaps the most potent endogenous negative regulator of haematopoiesis.
NM_001165	*BIRC3*	q22.2	Inhibits apoptosis by binding to tumor necrosis factor receptor-associated factors TRAF1 and TRAF2, probably by interfering with activation of ICE-like proteases.
NM_033306	*CASP4*	q22.3	This gene encodes a protein that is a member of the cysteine-aspartic acid protease (caspase) family. Sequential activation of caspases plays a central role in the execution-phase of cell apoptosis.
NM_004347	*CASP5*	q22.3	As above. Over-expression of the active form of this enzyme induces apoptosis in fibroblasts.
NM_001223	*CASP1*	q22.3	As above.
NM_006235	*POU2AF1*	q23.1	A chromosomal aberration involving POU2AF1/OBF1 may be a cause of a form of B-cell leukemia. Translocation t(3;11)(q27;q23) with BCL6.
NM_002017	*FLI-1*	q24.3	Friend leukemia virus integration 1. A chromosomal aberration involving FLI1 is a cause of Ewing sarcoma. Translocation t(11;22)(q24;q12) with EWS.

Table 5.9: Initial list of candidate genes on 11q.

Of the three highlighted genes (yellow), I decided to begin with the mixed-lineage leukemia (*MLL*) gene due to its strong links with acute leukaemia, and its location at 11q23.

5.4.3.2 Screening the Mixed lineage leukaemia (MLL) gene

The *MLL* gene, otherwise known as *ALL-1*, *HTRX1* or *HRX* amongst others, encodes a DNA-binding protein that methylates histone H3 lysine 4 as well as positively regulating expression of numerous *HOX* genes, which are in turn essential for haematopoietic development (Krivtsov and Armstrong 2007). Rearrangements of *MLL* were first recognised as a consequence of chromosomal translocation in AML (Ziemin-van der Poel, *et al* 1991). In fact there have now been reports of translocations involving *MLL* and greater than 50 different fusion partners (Meyer, *et al* 2006). In addition, partial tandem duplications (PTD) within *MLL* gene have also been reported in AML with a normal karyotype or trisomy 11 (Caligiuri, *et al* 1994,

Schichman, *et al* 1994). Numerous in frame *MLL*-PTDs have been identified and named according to the fused exons, the most frequent being e9/e3 (under previous nomenclature this was e2/e6), followed by e11/e3 (e2/e8) and e10/e3 (e2/e7). The *MLL*-PTD is found in approximately 5-20% of adult AML with a normal karyotype and significantly higher proportion (25-54%) in adult AML patients with trisomy 11. Although the precise role of the *MLL*-PTD in leukaemogenesis remains a matter of debate (Marcucci, *et al* 1998, Schnittger, *et al* 1998, Weisser, *et al* 2005) I decided to see if the 11q aUPD patients were positive for the *MLL*-PTD.

The single step primers, 3.1c and E3AS, were ordered according to the Schichman et. al., (1994) paper, and nested primers, MLLint and 6.1, were later ordered according to the Schnittger et. al., (1998) paper. RT-PCR was performed as described in Schnittger et. al., (1998) using cDNA from each of the three 11q aUPD patients plus a positive control, the eosinophilic cell line EOL-1 that is known to harbour a *MLL*-PTD (Schichman, *et al* 1994, Schnittger, *et al* 1998). Applied Biosystems AmpliTaq Gold and buffer were used as an alternative to that stated in the paper. The results of the RT-PCR are shown below in Figure 5.15. As can be seen, the EOL-1 cell line clearly shows an *MLL*-PTD by standard single step PCR, whilst none of the aMPD patients with 11q UPD show any sign of a *MLL*-PTD by either single step or nested RT-PCR (gel picture not shown).

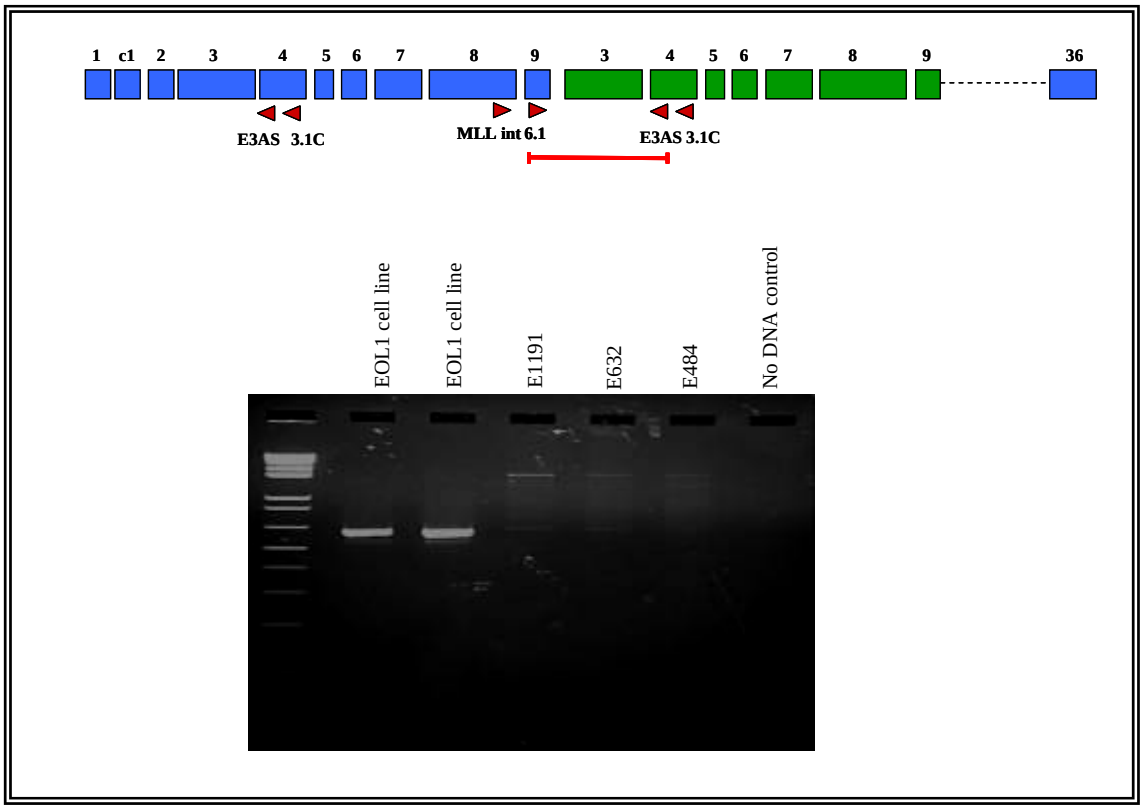


Figure 5. 15: Exon structure of the 5' region of the MLL gene illustrating the common e9/e3 PTD (only the new nomenclature is detailed). The boxes represent individual exons, those in blue are the WT and those in green indicate the PTD. The red arrows indicate the primers as taken from Schnittger et. al.,(1998). The gel picture below shows an Invitrogen 1kb plus ladder (lane 1) and the product observed from all three patients with 11q aUPD (lanes 3-5) versus the cell line EOL-1 (lanes 1-2). The EOL-1 cell line has been shown to harbour a MLL-PTD so was used as a positive control (Saito H 1985).

MLL is a very large gene and point mutations have not been reported so I went on to consider other candidates. The *ATM* gene seemed a strong possibility with mutations reported in this gene responsible for genomic instability and an increased risk of cancer. *ATM* gene mutations are seen in subtypes of lymphoid leukaemia and lymphoma (Boultwood 2001) but have not been reported in myeloid disease. The third gene that stood out was *CBL*. This gene encodes a member of the Casitas B-

lineage lymphoma gene family of ubiquitin ligases which function as negative regulators of several receptor protein tyrosine kinase signalling pathways. CBL achieves this by ubiquitinating activated receptor tyrosine kinases, thus promoting their sorting for endosomal degradation or recycling to the plasma membrane. Whilst deliberating which of these candidate genes to screen next, two interesting papers were published reporting the finding of *CBL* mutations in AML. Sargin et al., found a missense mutations in 1 of 150 cases analysed and demonstrated that the mutation had transforming activity (Sargin, *et al* 2007). Cagliuri et. al., (2007) found abnormalities in 4 of 12 cases, none of which were functionally analysed (Caligiuri, *et al* 2007, Langdon, *et al* 1989). The reason for the marked discrepancy in the proportion of mutation positive cases is unclear, however the technical details for the second paper were very scant and in particular it is unclear whether the sequence changes were verified or not. Nevertheless, given these findings and the known interaction of CBL with tyrosine kinases, I therefore decided to screen this gene next.

5.4.3.3 The CBL gene – introduction

The *CBL* proto-oncogene (originally called c-CBL) was first discovered as the cellular homologue of *v-CBL*, a viral transforming gene from the Cas NS-1 murine retrovirus that causes pre-B cell lymphomas and myelogenous leukaemias in mice (Langdon, *et al* 1989). It is a large, ubiquitously expressed, principally cytoplasmic protein. There are three members of the CBL family: CBL, CBLB and CBLC, all of which play an important role in regulating signalling by many receptor tyrosine kinases. Structurally, the CBL family members are modular proteins that contain a conserved N-terminal tyrosine kinase binding domain (TKB) that has been shown to bind to over 40 proteins involved in cellular regulation. For example it has been shown to bind phosphorylated cytoplasmic kinases such as ZAP70, and several receptor tyrosine kinases such as EGFR and PDGFR. They also have a RING finger domain that has been shown to interact with and recruit the E3 ubiquitin-ligase enzymes (E3) as well as other negative regulators of signalling. Finally, they all contain a proline-rich region that interacts with SH3-domain containing signalling

proteins such as GRB2, SRC and CIN85. As a consequence of this multifunctionality, CBL has been shown to play both a positive and negative role in signalling.

The negative regulatory role of CBL is summarised in the diagram below (Figure 5.16). In brief, activated receptor tyrosine kinases recruit CBL, and CBL mediates the ubiquitination of the receptor on lysine residues. Ubiquitination is the signal that targets the receptors for internalisation via clathrin coated pits, followed by endocytic sorting resulting in either recycling or degradation of the receptor (Shtiegman and Yarden 2003).

In addition, CBL proteins may also play a positive role in mediating positive RTK signalling events to downstream effectors. Schmidt et. al., (2005) have shown that phosphorylation of CBL results in it binding to positive signalling regulators such as SHP-2, Gab2 and PI3-kinase via its C-terminal multi-adaptor domain (Schmidt and Dikic 2005). In other words, CBL can act as a scaffold to the recruitment of downstream effectors of tyrosine kinases.

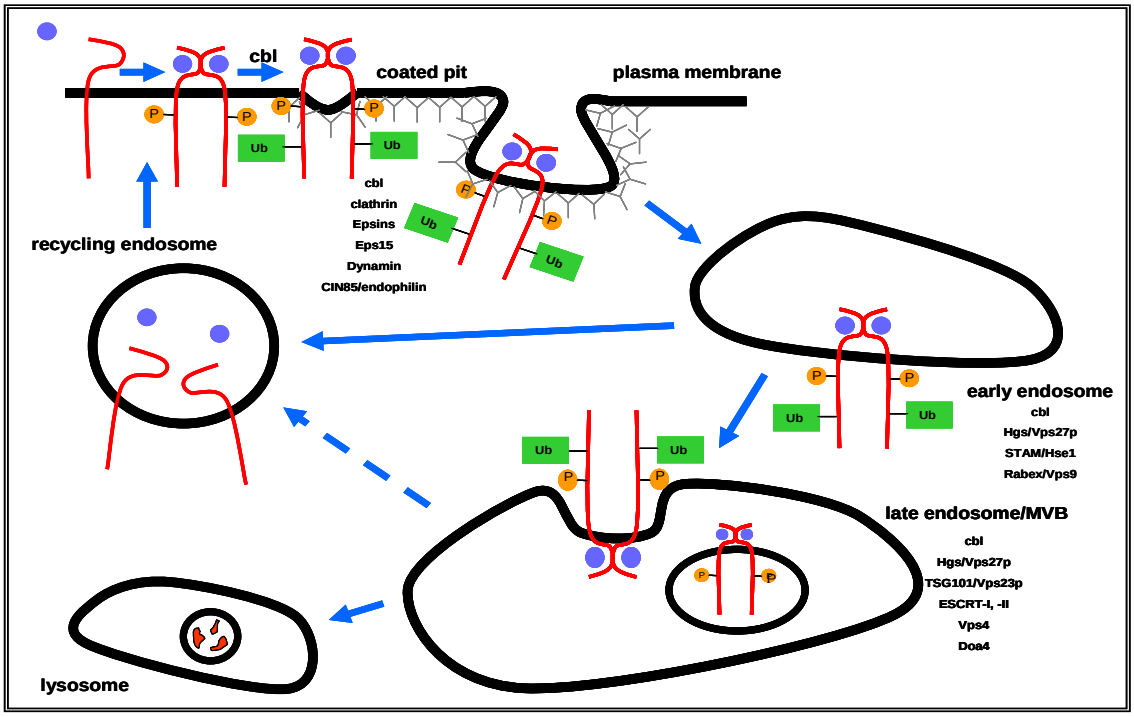


Figure 5. 16: (Adapted from Marmor & Yarden, Oncogene, 2004) (Marmor and Yarden 2004) Receptor tyrosine kinase (RTK) negative regulation and the role of CBL. RTK stimulation with ligand induces dimerisation and hence activation. This leads to autophosphorylation and recruitment of CBL, which itself is then phosphorylated and ubiquitinates the RTK. The ubiquitylated receptors are then sorted into clathrin-coated pits by a multiprotein complex that includes coat adaptors such as Epsin and Eps15. Fission of clathrin-coated vesicles is mediated by a GTPase, dynamin. In addition, CBL-mediated recruitment of CIN85 and endophilin is thought to promote negative membrane curvature and invagination. Shedding of clathrin, a decrease in the internal pH and the accumulation of hydrolytic enzymes, signals progression through the endocytic pathway. The RTK trafficking from early to late endosomes is dependent on its continued association with CBL and its sustained ubiquitination. The endocytic vesicle finally uncoats (facilitated by synaptojanin) and fuses with an early endosome, by a mechanism catalysed by the small GTPase Rab5 and its effectors. From here, the RTK is sorted into intraluminal vessicles of the endosome, in a process mediated by ESCRT-I, ESCRT-II and ESCRT-III. The RTK is probably dephosphorylated and deubiquitinated during this sorting step. The multivesicular body (MVB)/endosomal carrier vesicle is formed in an Annexin-II-dependent fashion, which ultimately fuses with a late endosome or a lysosome. Fusion of the

MVB with the lysosome results in degradation of the contents of the internal vesicles. Recycling of receptors back to the plasma membrane can occur throughout the endocytic pathway, albeit with decreasing efficiency.

5.4.3.4 The CBL gene – screening 11q aUPD patients

Mutations of *CBL* in AML patients were restricted to the RING domain, encoded by exons 8 and 9. Primers were therefore designed to anneal at 66°C within the introns surrounding these exons to allow the whole exons to be screened using genomic DNA (for primer sequences see Appendix I, section XXIV). PCRs were undertaken in a volume of 25µl using standard conditions and yielded exon 8 and exon 9 PCR products of 424bp and 369bp, respectively (Figure 5.17).

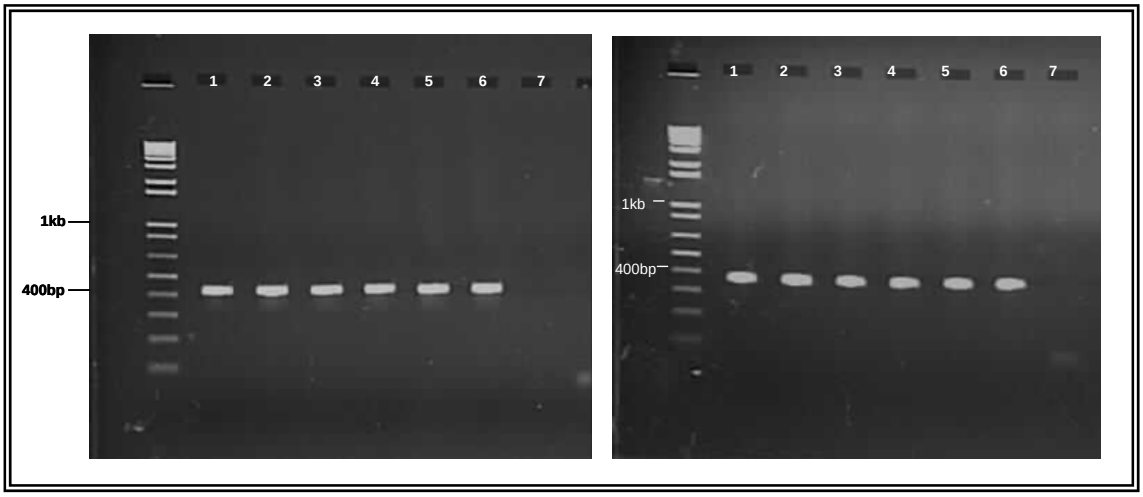


Figure 5. 17: Agarose gels showing amplification of CBL exon 8 (left) and 9 (right) for all three aMPD patients with 11qUPD. In both cases, the 1Kb plus ladder (Invitrogen) is shown in the first lane to accurately size the products. For both gels: Lane 1, E1191; Lane 2, E632; Lane 3, E484, Lanes 4-6 normal controls, and Lane 7 no DNA control. The correct sized product for CBL exon 8 (~424bp) and exon 9 (~360bp) is seen for all patients and controls.

Direct sequencing identified changes in all three patients. In two patients, E484 and E1191, a T>C change at nucleotide 1277 in exon 8 (Figure 5.18), was seen that is

predicted to change leucine 380 to a proline (L380P). The mutation for E1191 appeared to be predominant with only a small residual wild type T peak visible on the sequence traces. E484 showed a more substantial T peak, which is consistent with the relatively low LOH peaks detected by array analysis as shown in Figure 5.13 above.

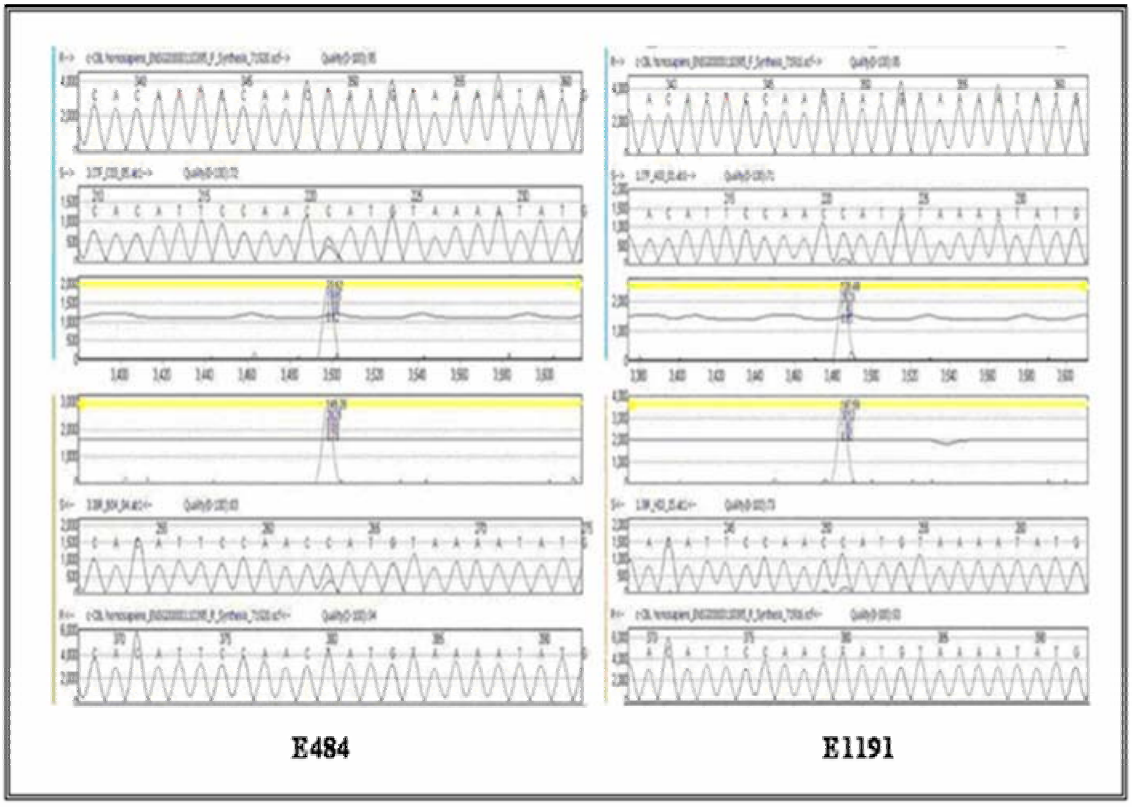


Figure 5. 18: Electropherograms showing part of CBL exon 8 amplified from aCML/aMPD patients E484 (left) and E1191 (right). Each panel show 6 traces: those at the top and bottom (traces 1 and 6) represent the reference CBL sequence (NM_005188). Traces 2 and 5 represent patient sequence in forward and reverse orientations, respectively. The traces in the centre (traces 3 and 4) show a comparison between the reference and the patient sequence highlighting any sequence changes with a peak. Both patients show a T>C mutation at nucleotide 1277 which is predicted to result in the substitution of the amino acid leucine by proline.

The third case (E632) was also found to have a sequence change, this time in exon 9. The mutation is a C>G missense change at nucleotide 1387 that is predicted to result

in the substitution of the proline 417 by alanine (P417A; Figure 5.19). The wild type C residue was barely detectable, consistent with the high level of LOH seen by array analysis. Both sequence changes were different to those described in AML patients.

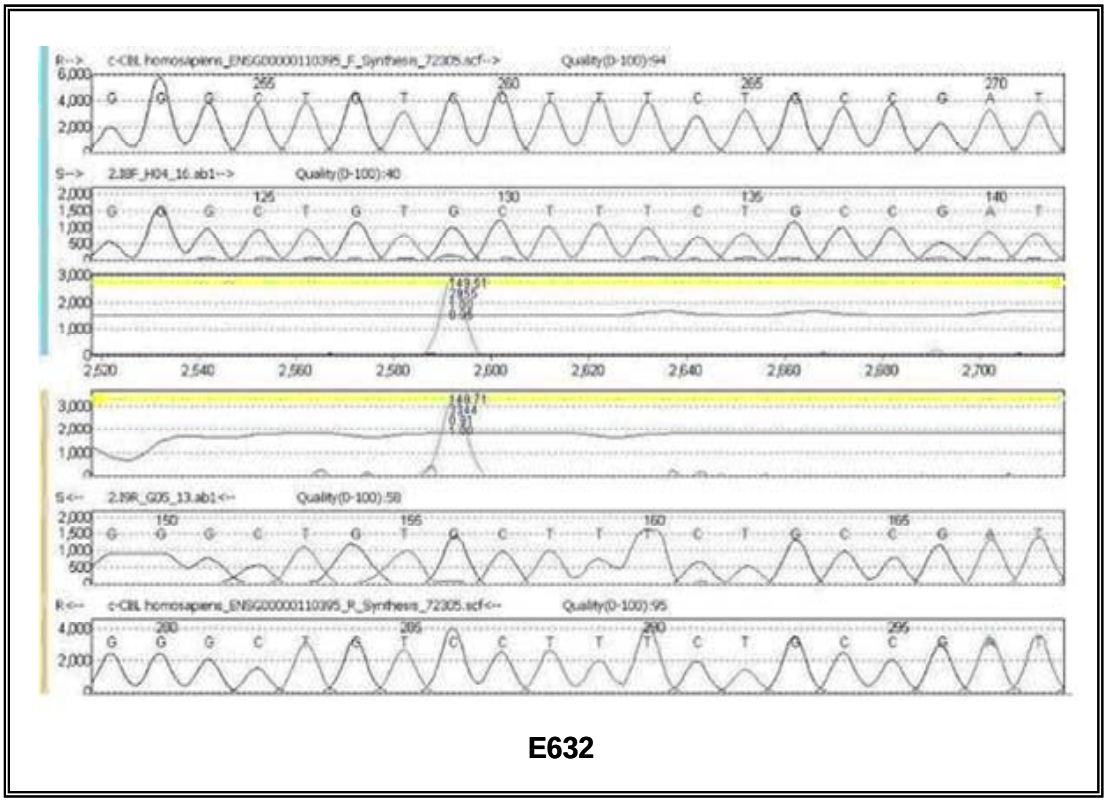


Figure 5. 19: Sequence analysis of patient E632 showing a portion of CBL exon 9. The order of the traces is the same as Figure 5.18, with traces 3 and 4 showing the C>G mutation at nucleotide 1387 that results in a P417A amino acid substitution.

5.4.3.5 Screening other MPD patients for CBL mutations

Following the detection of the *CBL* mutation in all three 11q aUPD patients I wanted to next determine if mutations of this gene might be more widespread in both typical and atypical MPDs. I therefore screened 547 patients as detailed in Table 5.10 by sequence analysis of exons 8 and 9.

I identified mutations in a further 23 patients, making 26 in total. Surprisingly, one patient was found to have a mutation in both exons 8 and exon 9 making a total of 15 cases with exon 8 mutations and 12 cases with exon 9 mutations (Table 5.10). Comparison with MPD subtype indicates that *CBL* mutations are seen in cases with CMML (12.8%), aCML/aMPD (7.9%) and IMF (5.7%).

Disease classification	***No. screened***	***Total no. of patients with mutations***	***CBL ex8 mutations***	***CBL ex9 mutations***	***% with CBL mutation***
PV	74	0	0	0	0%
IMF	53	3	1	2	5.7%
ET	24	0	0	0	0%
SM	60	0	0	0	0%
CMML/MDS	78	10	8	2	12.8%
aCML/aMPD	152	12	6	7	7.9%
HES/Eosinophilia	96	1	0	1	1%
CNL	9	0	0	0	0%
AML	1	0	0	0	0%
TOTAL	**547**	**26 patients (27 mutations)**	**15**	**12**	**4.9%**

Table 5. 10: Summary of CBL mutations according to disease subtype.

The mutations are summarised in Table 5.11. Most were missense changes, although I identified some candidate splice site mutations (n=4) and a heterozygous deletion (n=1). The majority of these sequence changes were novel however I did detect the R420Q mutation originally found in an AML patient (Sargin, *et al* 2007) in 4 of my 26 patients. This mutation was present in two IMF, one aCML and one unclassified aMPD cases and thus does not appear to be associated with a specific disease subtype. R420 is highly conserved in all CBL proteins and is the predicted contact

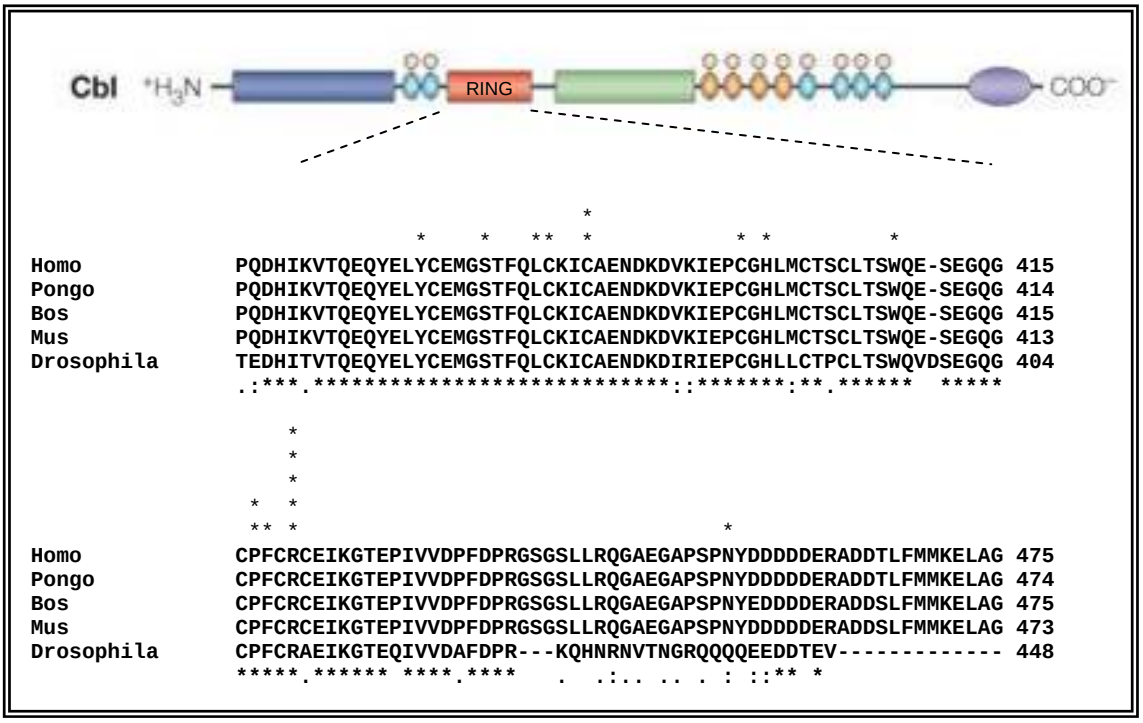
point of interaction between CBL and E2 ubiquitin-conjugating enzymes (Zheng, *et al* 2000) (Figure 5.20). This mutation was shown by Sargin et. al., (2007) to prevent FLT3 receptor internalisation and ubiquitination and is the first human *CBL* mutation causatively linked to the development of leukaemia. I also identified one patient with an amino acid alteration of the same codon, 420 that results in an arginine to leucine (R420L) substitution.

Patient Identifier	**Diagnosis**	**Exon 8 amplicon**	**Exon 9 amplicon**
E484	aCML	71930T>T/C: 380L>L/P	N
E632	aCML	N	72252C>G:417P>A
E1191	aCML	71930T>T/C: 380L>L/P	N
E001	aCML	71918C>C/T:376S>S/F	N
E088	CMML	71933G>G/A:381C>C/Y	N
E106	aMPD	71902T>T/C:371Y>Y/H	462R>R/X
E110	CMML	HET DEL GGTAC (DEL Exon 8)	N
E435	aMPD	N	72252C>G:417P>A
NP277	aMPD	N	72262G>G/T:420R>R/L
04/1345	IMF	N	72262G>G/A:420R>R/Q
E879	IMF	71930T>T/C:380L>L/P	N
E2167	CMML	71977T>G:396C>G	N
E492	CEL	N	A>AG -2
04/4357	IMF	N	72262G>G/A R420R/Q
E140	CMML	72022C>CT (+4)	N
E2015	CMML	71886G>GC -1 (DEL Exon 8)	N
P42	aCML	72023 G>GA (+5)	N
E674.2	aCML	N	72262 G>GA R420R/Q
P607	aMPD	N	72262 G>GA R420R/Q
0607621	CMML	N	72253C>C/T 417P>P/L
1383	CMML	N	72363A>AG: 454N>N/D
1457	CMML	71941T>TC: 384C>C/R	N
1617	CMML	71983C>CT: 398H>H/Y	N
826	CMML	71942G>GA 384C>C/Y	N
03/2231	aCML	72015G>C: W408C	N
06/5197	aCML	N	72257C>CA: F418F/L

Table 5. 11: Summary of CBL mutations identified in this study.

Another oncogenic *CBL* mutation (C381A) in the same region of the protein has been reported in an animal model (Andoniou, *et al* 1994, Thien, *et al* 2001). Although I did not detect this mutation in any of my patients, I did discover a mutation that occurred in the same codon, C381Y. This residue is also highly conserved among the CBL

proteins and hence is likely to have a causative affect equivalent to constitutive receptor tyrosine kinase signalling (Figure 5.20).




```
                                        *
                          *    *   **  *          * *        *
Homo            PQDHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTSWQE-SEGQG 415
Pongo           PQDHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTSWQE-SEGQG 414
Bos             PQDHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTSWQE-SEGQG 415
Mus             PQDHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTSWQE-SEGQG 413
Drosophila      TEDHITVTQEQYELYCEMGSTFQLCKICAENDKDIRIEPCGHLLCTPCLTSWQVDSEGQG 404
                .:***.****************************::*******:**.******  *****

                    *
                    *
                    *
                 *  *
                 ** *                                *
Homo            CPFCRCEIKGTEPIVVDPFDPRGSGSLLRQGAEGAPSPNYDDDDDERADDTLFMMKELAG 475
Pongo           CPFCRCEIKGTEPIVVDPFDPRGSGSLLRQGAEGAPSPNYDDDDDERADDTLFMMKELAG 474
Bos             CPFCRCEIKGTEPIVVDPFDPRGSGSLLRQGAEGAPSPNYEDDDDERADDSLFMMKELAG 475
Mus             CPFCRCEIKGTEPIVVDPFDPRGSGSLLRQGAEGAPSPNYDDDDDERADDSLFMMKELAG 473
Drosophila      CPFCRAEIKGTEQIVVDAFDPR---KQHNRNVTNGRQQQQEEDDTEV------------- 448
                *****.****** ****.****   .  .:.. .. . : ::** *
```

Figure 5. 20: Schematic showing the CBL protein domain structure at the top and a 5 species alignment of the RING domain below (Homo sapiens, Pongo Pygmaeus, Bos Taurus, Mus musculus & Drosophila melanogaster). The stars in black at the bottom of each alignment indicate the homology between species. The asterisks at the top of each alignment indicate the site of mutations found in this study.

All other missense mutations affect novel residues in the same region. In addition I identified 5 other sequence changes. The 5bp heterozygous deletion of GGTAC in patient E110 involved the last base pair of exon 8 and was similar to the deletion identified previously in the AML MOLM-13 cell line (Caligiuri, *et al* 2007) (Figure 5.21).

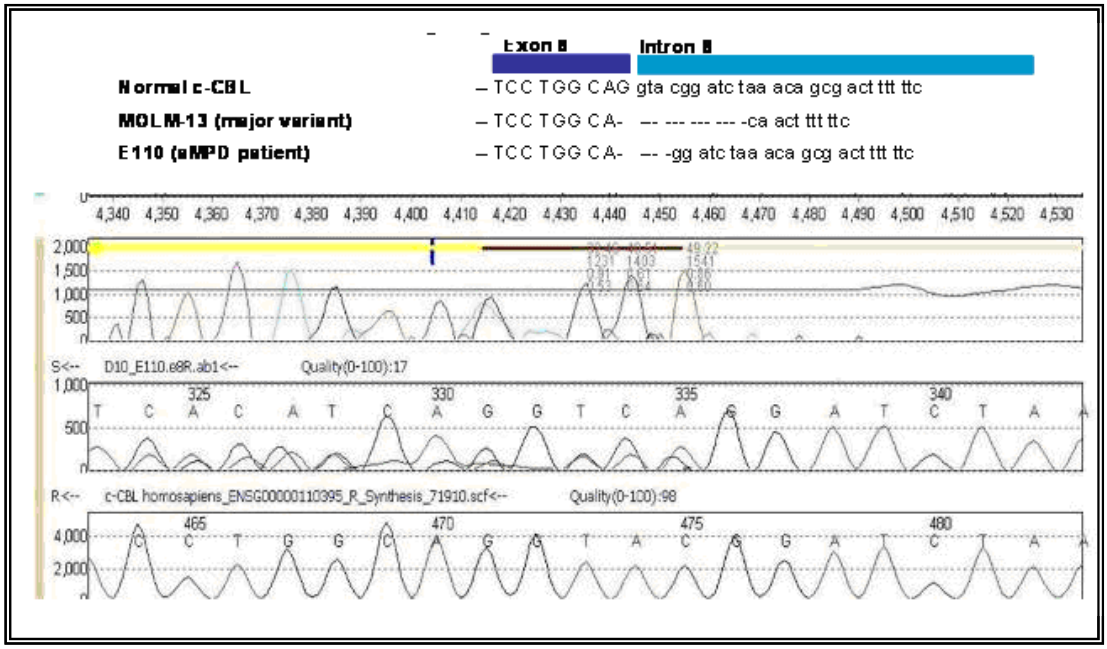


Figure 5. 21: The 5bp deletion in patient E110. Top panel: Comparison of the 14bp deletion in the MOLM-13 cell line with the deletion in patient E110. Bottom panel: the sequencing electropherogram showing the 5bp deletion.

In addition four sequence changes in patients P42, E140, E492 and E2015 were identified in the non-coding regions of exons 8 and 9 (Figure 5.22). It is possible that these mutations represent new splice donor or acceptor sites thus altering the protein structure.

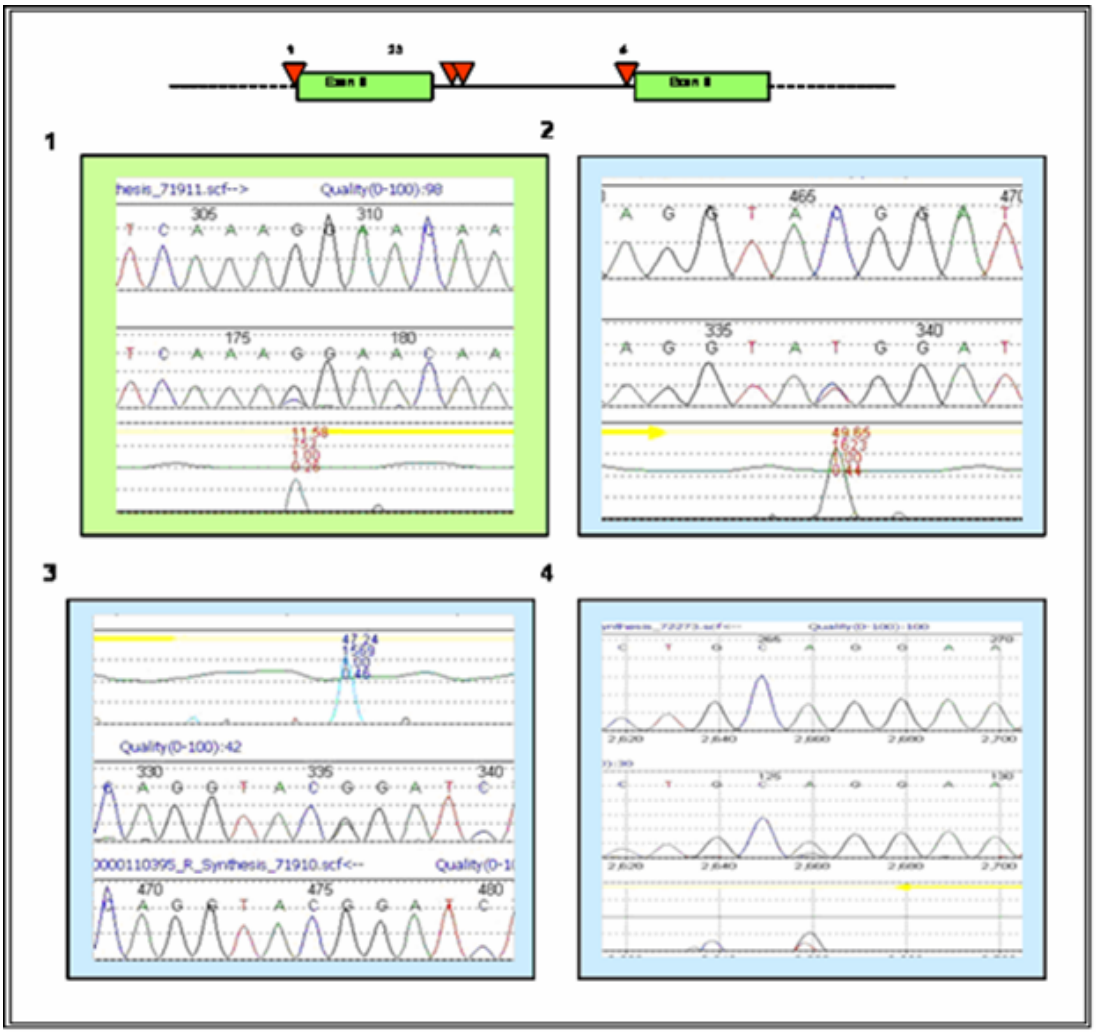


Figure 5. 22: Schematic showing the locations and sequencing electropherograms of all four intronic CBL mutations. The trace highlighted in green (1) was discovered in intron 7 a G>C mutation at -1 of exon 8 in patient E2015. The traces highlighted in blue represent the three non-coding mutations identified in intron 8; (2) is the +4C>T substitution in patient E140, (3) shows the +5G>C mutation found in patient P42, and (4) depicts patient E492 and the -2A>G mutation.

The mutation scan that yielded these sequence changes focused exclusively on exons 8 and 9. Upon identification of a mutation the patient DNA was subjected to repeat PCR for the exon in question and resequenced to confirm the mutation was in fact present. Given the strong association of the mutations with an aCML/CMML/MF phenotype, all other *CBL* exons were amplified and screened for mutations in a total

of 70 of aCML/CMML/MF patients that were negative for exon 8 or 9 mutations. Although a number of known SNPs were identified, no other mutations were found.

5.4.3.6 11q LOH studies

I aimed to determine if the *CBL* mutations were usually associated with 11q aUPD. Apart from the three initial cases, the other mutation positive cases had not been analysed by SNP arrays. Since sequence analysis is poorly quantitative and arrays are expensive, I used microsatellite PCR analysis to look for chromosome 11q LOH in all 26 CBL mutation positive samples. In total I used 9 fluorescently labelled primer pairs targeting highly polymorphic microsatellite markers spanning chromosome 11q (Table 5.12).

Chromosome band	**Marker**	**Heterozygosity (http://www.gdb.org/)**
11q13.1	D11S1883	0.74
11q14.1	D11S937	0.88
11q21	D11S4182	0.73
11q23.2	D11S1885	0.71
11q23.3	D11S925	0.85
11q	D11S4107	0.78
11q24.1	D11S982	0.77
11q24.2	D11S934	0.85
11q25	CA rpt	unknown

Table 5. 12: Microsatellite marker name, location on chromosome 11q and heterozygosity where known (www.gdb.org/).

The products were analysed on an ABI3100 using the Genotyper 2.0 program. As previously described for *JAK2*, if only one peak was visible a patient was scored as homozygous. To allow for the presence of background normal cells, homozygosity was also scored if one peak was one third or smaller than the expected size compared to controls (Jones, *et al* 2005). Of the 26 *CBL* mutation positive patients tested, 9 (including all three known 11q LOH patients (E484, E632 & E1191) displayed patterns that were suggestive of LOH (Table 5.13).

Patient I.D	D11S 1883	D11S 937	D11S 4182	D11S 1885	D11S 925	D11S 4107	D11S 982	D11S 934	11q25 CA rpt	Tel
E484										
E632										
E1191										
E001										
E088										
E106										
E110										
E435										
NP277										
04/1345										
E879										
E2167										
E492										
04/4357										
E140										
E2015										
P42										
E674.2										
P607										
0607621										
1383										
1457										
1617										
826										
03/2231										
06/5197										

Table 5. 13: 11q LOH in CBL mutation positive patients. The microsatellite markers are shown in order from the most centromeric (D11S1883) to the most telomeric (11q25 CA repeat) and are scored as homozygous (black) or heterozygous (white). The marker highlighted in yellow is the first that is distal to CBL.

Of these nine, six showed complete homozygosity at all microsatellite loci tested (the three 11q LOH patients detected by arrays plus E001, E2167 and 1457). The others (E879, 1617 and 03/2231) showed varying degrees of homozygosity however all ran from a centromeric to telomeric marker, consistent with mitotic recombination. A further three patients (04/1345, P42 & E140) also had regions of homozygosity over a region of 4 or 5 microsatellite markers however they did not extend to a telomere. Unfortunately constitutional DNA was not available from any case to determine if these regions were acquired or not, however the mean heterozygosity of the markers employed was 0.79 and thus the chance of four sequential homozygous calls randomly is $(0.21)^4 = 0.0019$, assuming that all are independent of each other.

Patient E106 was heterozygous for all markers. This is the only patient for whom I identified two *CBL* mutations, one in exon 8 the other in exon 9. The exon 8 mutation is missense and predicted to result in a unique Y371H amino acid substitution. The exon 9 mutation is also unique to this patient, R462X, and is also the only nonsense mutation identified. I carried out a PCR to determine if these mutations were on the same or different chromosomes as this may tell us more about how they are acting. I designed two intronic primers that would anneal at 66°C, and amplified both exons 8 and 9 of the *CBL* gene (see Appendix I, section XXVII) in a single amplicon of 786bp. This was then sub-cloned and several colonies sequenced. I was able to show that the mutations were never in the same clone indicating that the mutations were on different chromosomes. One possibility is that the truncating mutation serves to reduce the Y371H mutation to functional homozygosity.

5.4.3.7 Copy number analysis

I next went on to determine whether the *CBL* locus was present in two copies, which would implicate aUPD as the cause of the homozygosity, or one copy which would implicate a deletion. Multiplex ligation-dependent probe amplification (MLPA) is a robust technique used to look at copy number differences and a number of different kits have been developed by MRC Holland (www.mlpa.com) for genes of diagnostic interest. In particular, the Marfan probeset 1 (P065) kit contains 4 control probes at

11q including one at 11q23.3 the exact cytogenetic location of *CBL*. I therefore used this kit to examine *CBL* copy number and the results are summarised below in Table 5.14.

Patient I.D	MLPA results
E484	N
E632	N
E1191	N
E001	N
E088	N
E106	N
E110	no material
E435	no material
NP277	N
04/1345	N
E879	N
E2167	N
E492	N
04/4357	no material
E140	no material
E2015	N
P42	N
E674.2	N
P607	N
0607621	no material
1383	no material
1457	N
1617	N
826	no material
03/2231	N
06/5197	N

Table 5. 14: Patient identifier versus MLPA copy number result. N= normal (no copy number alteration) indicating the patient has two copies of chromosome 11q.

Of the 26 patients with *CBL* mutations, only 19 were analysed by MLPA due to the paucity of suitable material for some cases (MLPA is very sensitive to DNA quality). The three original 11q LOH patients identified by SNP array analysis were included as controls and all three were confirmed by MLPA to have two 11q chromosomes. Of the remaining 16 patients, all apparently had two intact chromosomes 11q including the six cases for whom I found LOH by microsatellite analysis. This

implies the LOH is likely to be due to mitotic recombination following an initial *CBL* mutation.

5.4.3.8 Are the CBL mutations acquired?

It is important to determine if the *CBL* mutations were acquired or constitutional. An acquired mutation implies the genetic change occurred in a somatic cell and provided it with a selective advantage and thus giving rise to a clone. Since this required obtaining fresh samples, it was only possible to perform the analysis on three cases.

A fresh blood sample was received from E1191 who was positive for the L380P mutation. I set up a phytohaemagglutinin (PHA) culture to enrich for T-cells and compared DNA prepared from a granulocytic fraction to DNA from a T-cell fraction. Previously studies in our group have shown that growth in PHA typically results in >80% T-cells after 5 days. As shown in Figure 5.23, the mutation is present in granulocytes but absent from T-cells.

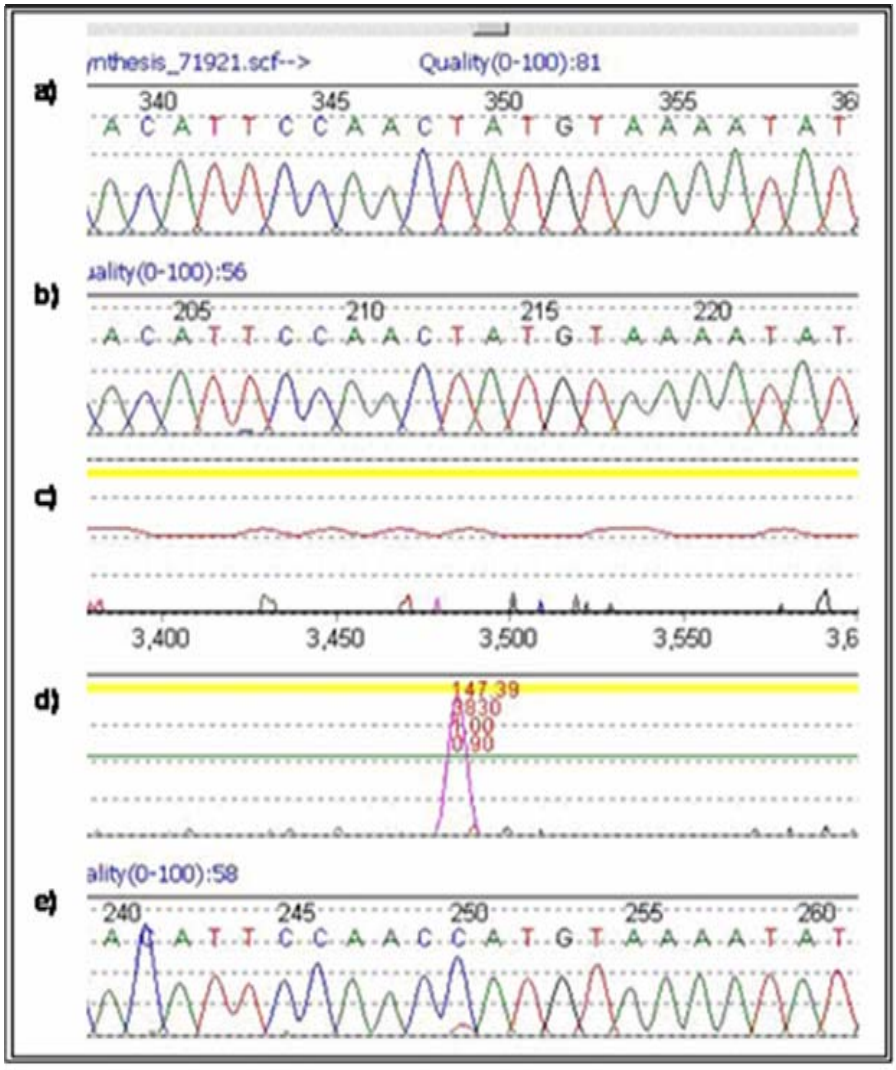


Figure 5. 23: Electropherogram for patient E1191 showing a) CBL exon 8 reference sequence; b) T-cell enriched DNA; c) comparison of a and b showing no differences in sequence; d) comparison between reference sequence and granulocyte DNA with the peak indicating a sequence difference, and e) granulocytic DNA sequence showing the T>C mutation.

Patient E879 also had the L380P mutation and in this case we were able to obtain a mouthbrush sample. Sequencing showed again that the mutation was only present in the PB DNA sample and not in the epithelial (mouthbrush) fraction (Figure 5.24).

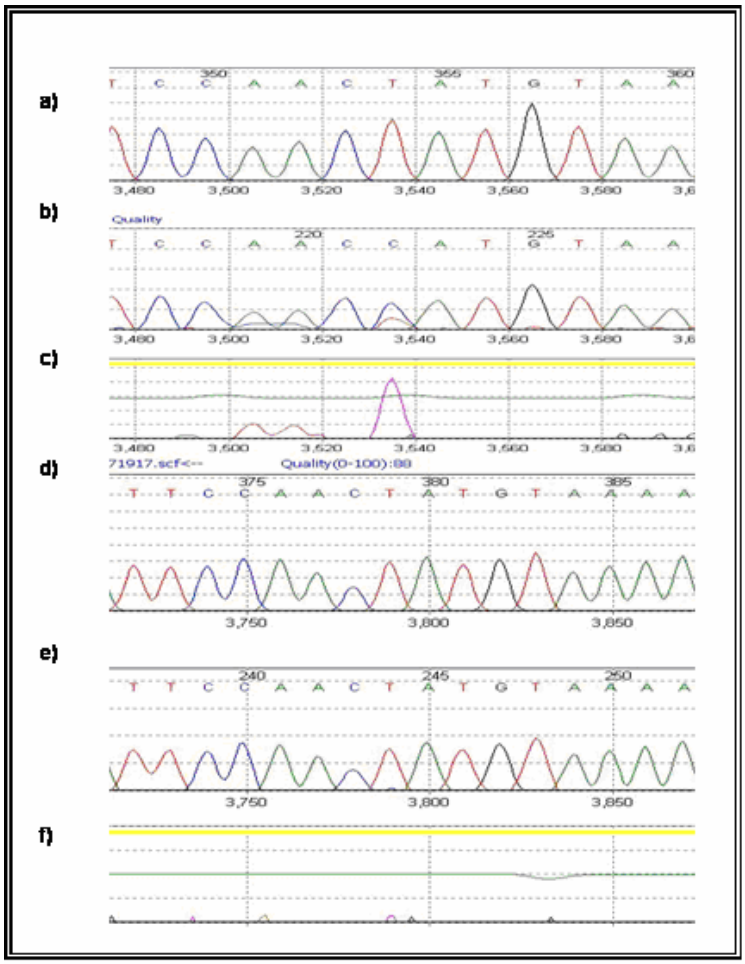


Figure 5. 24: Electropherogram depicting a) CBL reference sequence, part of exon 8; b) E879 PB DNA sample sequence trace showing T>C missense mutation; c) comparison between the reference sequence and E879 sequence, the peak indicates the position of the mutation; d) CBL reference sequence; e) mouthbrush DNA sequence, and f) comparison between the reference and the mouthbrush sequence showing no differences.

Finally, patient 04/1345 with a R420Q mutation was analysed. The sample I had originally analysed from this case was taken in 2004 when MF was diagnosed, however he had originally presented with ET in 1989. I was able to obtain and extract DNA from bone marrow slides taken in 1989, 1992, 1995 and 1997 and found no evidence of the *CBL* mutation. Unfortunately no slides were available between 1997 and 2004, however these findings suggest that the mutation may have appeared on transformation from ET to MF.

In conclusion, I have been able to determine that the *CBL* mutation was acquired in three of three cases analysed. It is likely that the other 23 patients also acquired *CBL* mutations however this remains a matter of conjecture.

5.4.3.9 Analysis of non-coding CBL mutations

As mentioned above, intronic sequence changes were identified in four cases and a 5bp deletion was found in a further patient. To determine if these variants affected splicing, I performed RT-PCR on the three cases for which material was available: E140 and E2015 with intronic changes and E110 with the heterozygous 5bp deletion.

Primers were designed over two exon boundaries to ensure they would only amplify the coding sequence and not any residual genomic DNA if present in the cDNA sample. The primers were designed to anneal at 66°C and cover exons 6/7 and exons 10/11, thus ensuring the region comprising exons 7-10 would be amplified (for primer sequences see Appendix I, section XXV). A normal control was also run alongside and the resulting amplifications are shown in Figure 5.25.

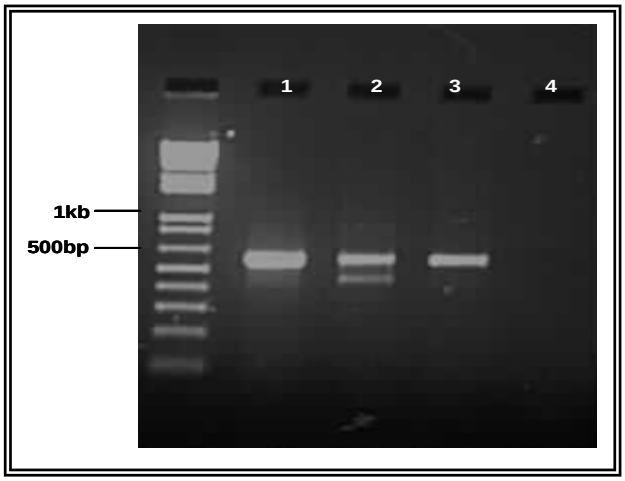


Figure 5. 25: Agarose gel picture. Lane 1 represents E140; Lane 2, E2015; Lane 3, normal control, and Lane 4, no DNA control run alongside the 1kb Plus ladder (Invitrogen) for sizing. The normal control (Lane 3) is the expected size (~585bp) and when compared to E140 it is clear this patient also only has this normal band suggesting no splicing differences are created with the presence of this mutation. The other patient, E2015 exhibits two bands a normal 585bp band and a smaller band that upon cloning and sequencing revealed a deletion of exon 8.

Patient E140 showed a single band that corresponds to the size of the normal control. Upon sequencing I confirmed exons 7-10 were fully intact thus it appears the variant at 72022C>C/T(+4) does not affect the structure of the protein. Although I did not have material to prove this, it seems likely that the 72023G>G/A(+5) variant seen in patient P42 is also likely to be of no consequence. To test this possibility, I put both these altered sequences into a splice site predictor programme (www.fruitfly.org/seq_tools/splice.html) which predicts if an unknown sequence change might create a better or different splice acceptor or donor site compared to the reference sequence (Reese MG 1997). Neither of the mutations showed improved splice donor or acceptor sites supporting the idea that they are non-functional and are presumably rare polymorphisms or pathogenetically unimportant passenger mutations.

Patients E110 (not shown) and E2015 both exhibited double bands by RT-PCR. Cloning and sequencing showed that the larger band corresponded to the normal product and the smaller band resulted from the complete deletion of exon 8 in both cases. It therefore appears that both patients harbour heterozygous mutations with one normal allele and one allele with a deletion of exon 8. The mutation in MOLM-13 also causes skipping of exon 8 and siRNA targeting the exon 7-9 mutant transcript in this cell line was found to be growth inhibitory strongly suggesting that it is indeed of functional importance (Caligiuri, *et al* 2007).

5.4.3.10 The relationship between CBL and other mutations in MPDs

It is possible that the *CBL* mutants provide a growth advantage by interfering with the normal negative regulation of tyrosine kinase signalling, however it is unclear if they might act as sole abnormalities or if they co-operate with other acquired genetic events. All the cases I studied were negative for known tyrosine kinase fusion genes but they had not been tested for other oncogenic mutations seen in MPDs such as the *FLT3* ITD, *JAK2* V617F and *NRAS* mutations. These mutations have been reported in approximately 5%, 10-15% and 10% of aMPD cases, respectively (Jones, *et al* 2005).

The *FLT3* ITD analysis was performed as described previously. The *JAK2* V617F mutation analysis was carried out by a colleague using a specific ARMS test (Jones, *et al* 2005). For *NRAS* I designed primers to exons 2 and 3 which includes the two mutation hotspots at codons 12, 13 and codon 61 (for primer sequences see Appendix I, section XXVI). All *CBL* mutation positive patients were analysed and the results are shown in Table 5.15.

Patient ID	*JAK2* status	*NRAS* status	*FLT3* ITD status
E484	Normal	N	Normal
E632	Normal	N	Normal
E1191	Normal	N	Normal
E001	Normal	N	Normal
E088	Normal	N	Normal

E106	Normal	N	Normal
E110	Normal	N	Normal
E435	Normal	N	Normal
NP277	Normal	N	Normal
04/1345	Normal	N	Normal
E879	Het Positive	N	Normal
E2167	Normal	N	Normal
E492	Normal	N	Normal
04/4357	Normal	N	Normal
E140	Normal	N	Normal
E2015	Normal	N	Normal
P42	Normal	N	Normal
E674.2	Normal	N	Normal
P607	Normal	N	Normal
0607621	Normal	N	Normal
1383	Normal	N	Normal
1457	Normal	N	Normal
1617	Normal	N	Normal
826	Normal	N	Normal
03/2231	Normal	N	Normal
06/5197	Normal	N	Normal

Table 5. 15: Results of JAK2 V617F, NRAS & FLT3 ITD mutation analysis in CBL mutated cases. Het = heterozygous mutation.

Only one patient was found to harbour another mutation: E879 was heterozygous positive for *JAK2* V617F. Since this finding was unusual, i.e. only seen in a single case, it raised the question whether both mutations arose in the same clone or not. To address this issue a fresh blood sample (taken approximately 2.5 years after the original sample) was obtained and set up in methylcellulose cultures in order to grow haematopoietic colonies for clonal analysis. However in contrast to typical V617F positive cases, the colony growth was very poor and, furthermore, analysis of the blood leucocytes showed the *CBL* mutation was present as expected but V617F *JAK2* was no longer detectable. This raised the question of a technical error at some point in the analysis. To address this, I used the Promega Powerplex® 16 kit to determine if the two samples, plus also other samples we had from this case, were indeed from the same individual. This kit amplifies 16 highly polymorphic loci across the genome and was designed to allow a high discriminatory power for analysis of identity and paternity. I tested four different samples (E40, E879, E2763 and E2772) received at

various time points from this individual All four samples concurred at all 16 loci, proving they were in fact all from the same individual. These samples and others that were received were analysed for V617F *JAK2* and *CBL* as shown on Table 5.16.

Sample ID	**Date**	**Sample type**	***JAK2* mutation**	***CBL* mutation**
E1064	1998	BM Slide	YES (5% V617F)	NO
E40	2000	PB	YES	NO
E2922	2003	BM slide	not tested	NO
E879	2005	PB	YES (7% V617F)	YES
E2763	2008	BM	NO	YES
E2772	2008	PB	NO	YES
E2915	2008	mouthbrush	NO	NO

Table 5. 16: Time course of disease in patient E879 showing the disappearance of JAK2 V617F and the appearance of the CBL mutation over time.

The early samples received were all *JAK2* V617F positive (albeit at relatively low levels) and *CBL* L380P negative. Between 2005 and 2008 this appears to have changed with the *CBL* L380P clone becoming more dominant and the *JAK2* V617F clone being lost. The proportion of *CBL* mutation positive alleles as called by the mutation surveyor sequencing programme was 86% in 2005 and had increased to 100% by 2008. This transition coincided with an increase in leukocyte numbers that has proved difficult to control by standard therapy. This situation is reminiscent of the evolution of V617F positive PV to acute leukaemia in which it has been reported that roughly half of cases have leukaemia blasts that are V617F negative (Campbell, *et al* 2006, Theocharides, *et al* 2007). One possible explanation for this unexpected finding is that an unknown mutation results in a clinically covert clonal expansion prior to the acquisition of V617F. Another explanation is that there may be a hereditary predisposition to acquisition of independent mutations. In our case, the fact that V617F disappeared must presumably mean that the *CBL* mutation was acquired in a V617F negative clone and that it had a relative growth advantage.

One other case had two mutations: E2167 was originally diagnosed as systemic mastocytosis (SM) before evolving to CMML and had previously tested positive for

the *KIT* D816V mutation before I found a *CBL* mutation in the same sample. Unfortunately it was not possible to obtain any further samples from this case.

5.5 DISCUSSION

5.5.1 SNP arrays to identify new pathogenetic abnormalities

LOH is a genetic mechanism that creates a homozygous or hemizygous somatic cell by loss of the wild-type allele. This can be as a result of various mutational events such as a deletion, mitotic recombination, localised gene conversion, point mutation, or non-dysjunction. My aim has been to identify new oncogenes or tumour suppressor genes that are marked by these regions of LOH, and specifically regions of aUPD, or copy number neutral LOH formed as a consequence of mitotic recombination.

Regions of aUPD have been described in many human malignancies. Furthermore, these regions are now being commonly associated with oncogenic mutations, the first example of which was the discovery of 9p aUPD and the *JAK2* V617F mutation (Kralovics, *et al* 2002, Kralovics, *et al* 2005b). Since these initial reports, other oncogenic mutations have been associated with aUPD involving distinct chromosomal regions, such as *CEBPA* and chromosome 19 (Raghavan, *et al* 2005), *FLT3* mutations and 13q (Griffiths, *et al* 2005), *NF1* and 17q (Flotho, *et al* 2007), *WT1* mutations and 11p and finally *RUNX1* and 21q (Fitzgibbon, *et al* 2005). Moreover, reports where paired samples have been studied often show that the mutation precedes the aUPD and hence the aUPD represents a ‘second hit’ associated with disease progression (Fitzgibbon, *et al* 2007, Fitzgibbon, *et al* 2005). These studies all highlight the potential for discovering new pathogenetic abnormalities by the finding of regions of aUPD. Obviously this approach has limitations, e.g. it will not detect cryptic fusion genes or point mutations that do not have a selective advantage when homozygous. However I was particularly encouraged by the

observations that activating mutations of two tyrosine kinases are subject to aUPD and that these are the gene class that is most commonly mutated in MPDs.

There are various techniques that can aid identification of these regions of LOH however SNP arrays have the advantage of being able to detect copy number in addition to zygosity. In my study, I employed 50K SNP arrays and focused almost exclusively on regions of aUPD however it was clear that large copy number changes were also detected that correlated with cytogenetic data. Although I did not investigate this aspect in any detail, others have found that copy number information generated by 50K arrays is relatively poor but is greatly improved by using arrays of a higher density, in particular the Affymetrix 500K arrays (Mullighan, *et al* 2007) and more recently SNP 6.0 arrays which include 1.8 million markers. However for the purpose of identifying aUPD the lower density arrays are sufficient.

In total 88 samples were sent to the RZPD German Resource Centre for Affymetrix 50K SNP array analysis, including 30 aMPDs (predominantly aCML), 28 IMF, 10 PV, 10 myeloid & 10 lymphoid CML blast crisis patients. Results were successfully obtained for all samples. The principal weakness of my study was the absence of constitutional DNA, e.g. from mouthbrush swabs or remission DNA and so it was not possible to determine if tracts of homozygosity were indeed acquired or whether they were inherited. Consequently I did not attempt to analyse the smaller regions of homozygosity in any detail, although I did determine that there were no obvious recurrent abnormalities that were restricted to particular patient subgroups and also no recurrent regions that might target tyrosine kinases.

I was fortunate that my principal study group – atypical MPDs – were particularly informative with 40% of cases showing large regions of copy number neutral homozygosity. Because such large regions are not seen in normal individuals it is reasonable to infer that they are most likely acquired, although I was only able to prove this in three cases. I found eight different chromosomal regions affected by aUPD of which two were recurrent (7q and 11q; 3 cases each) and not seen in other

patient subgroups. These regions are discussed in detail below. Acquired UPD was also seen at 1p, 8p, whole chromosome 13, 17q, 20q and 21q. The 1p abnormality was linked to an oncogenic *MPL* 515 mutation and the patient with whole chromosome 13 aUPD was homozygous for a *FLT3* ITD. It is likely that these two oncogenic changes are the targets of the aUPD, i.e. homozygosity for these mutations confers a growth advantage over the heterozygotes and so following random mitotic recombination the homozygous clone has a selective advantage. However I cannot exclude the possibility that there are other relevant abnormalities in the affected regions that are being selected for.

In the other patient subgroups, aUPD was uncommon. I was particularly hopeful that SNP analysis might be helpful in understanding disease progression in CML. CML-BC may be myeloid or lymphoid in phenotype and the molecular basis for transformation is poorly defined (Calabretta and Perrotti 2004). It is likely, but not certain, that disease progression involves several genetic steps and that different combinations of a limited number of genes are involved in different patients. The genes involved in myeloid transformation may be distinct from those in lymphoid transformation, but cytogenetic and molecular evidence (Chase, *et al* 2001, Johansson, *et al* 2002) suggest that there is an appreciable overlap in the molecular basis of myeloid blast crisis and AML and similarly in the molecular basis of lymphoid blast crisis and ALL. aUPD has been reported in both AML and ALL (Fitzgibbon, *et al* 2005, Irving, *et al* 2005, Raghavan, *et al* 2005) and therefore it was a disappointment that all BC cases were negative, apart from a possible small region at the tip of 15p in one case with myeloid disease. As mentioned above, the copy number resolution of the 50K arrays was insufficient to detect small deletions such as those recently described to target Ikaros in lymphoid BC and Ph+ALL (Mullighan, *et al* 2007).

Similarly, patients with PV and MF were largely uninformative, although one case with V617F positive MF had extensive aUPD on 5q. The significance of this finding is unclear, although it suggests that there may be a gene on 5q that co-operates with

V617F to induce progression to MF. One possibility is the recently identified 5q- MDS gene RPS14 (Ebert, *et al* 2008), however this is perhaps unlikely given that RPS14 appears to exert its effects through heterozygous deletion rather than homozygous mutations. Instead of analysing this single abnormality I focused on the recurrent regions in aCML/aMPD patients.

5.5.2 7q aUPD

I identified three patients with varying sized regions of aUPD at chromosome 7q with no further copy number alterations or aUPD observed elsewhere in the genome. This is the first report of recurrent 7q aUPD amongst MPDs although deletions of 7q and monosomy 7 are associated with diverse human cancers such as gastric and breast cancer (Weng, *et al* 2006, Zenklusen, *et al* 1994), lymphoid leukaemia (Dascalescu, *et al* 1999, Hernandez, *et al* 1997, Offit, *et al* 1995) and also occasionally in myeloid malignancies such as MDS, AML and juvenile CML (Aktas, *et al* 1999, Hasle, *et al* 1999, Mantadakis, *et al* 1999, Tosi, *et al* 1996).

The region of 7q aUPD in each patient varied, however the minimal overlapping region spanned approximately 53Kb covering cytogenetic bands 7q22.3-qter. This is a colossal amount of sequence and is thought to comprise some 400 different genes. Much work has already been performed on cases in the literature trying to define critical regions of deletion, yet no group has yet identified any definitive tumor suppressor gene(s) that is thought to be located within this region. However I was more concerned with identifying activating mutations, based on the strong linkage between aberrant tyrosine kinase signalling and MPDs. The most obvious candidate genes were the three tyrosine kinases *EPHA1*, *EPHB6* and *MET* but no mutations were identified. Partial screens of *BRAF* and *PI3KCG* were also negative. These genes were selected because they participate in tyrosine kinase signalling pathways and are either known to be mutated in cancer (*BRAF*) or highly related to a gene that is mutated in cancer (*PI3KCG* is related to *PI3KCA*). Since these are such good

candidates it might be worth screening the complete sequence of all exons in case the previously described hotspots are limited to epithelial cell tumours. A similar situation has been described previously: the activating D816V *KIT* mutation is seen in >80% of cases of systemic mastocytosis but in gastrointestinal stromal tumours, D816V is very uncommon and instead activating mutations are seen elsewhere in the same gene. I looked for other genes on 7q but it is easy to find large numbers of candidates, none of which were obviously outstanding.

Subsequent to my work, a number of papers have been published by Gondek and colleagues (Gondek, *et al* 2007a, Gondek, *et al* 2007b, Gondek, *et al* 2008). This group's main focus is on MDS and MDS/MPD using 250K SNP arrays. Copy number neutral LOH was seen in 30% of cases, particularly 1p, 1q, 4q, 7q and 11q. Figure 5.26 shows the regions of 7q aUPD in my cases compared to the five identified by Gondek and colleagues, three of which had CMML, one AML and one refractory anaemia (RA, a subtype of MDS).

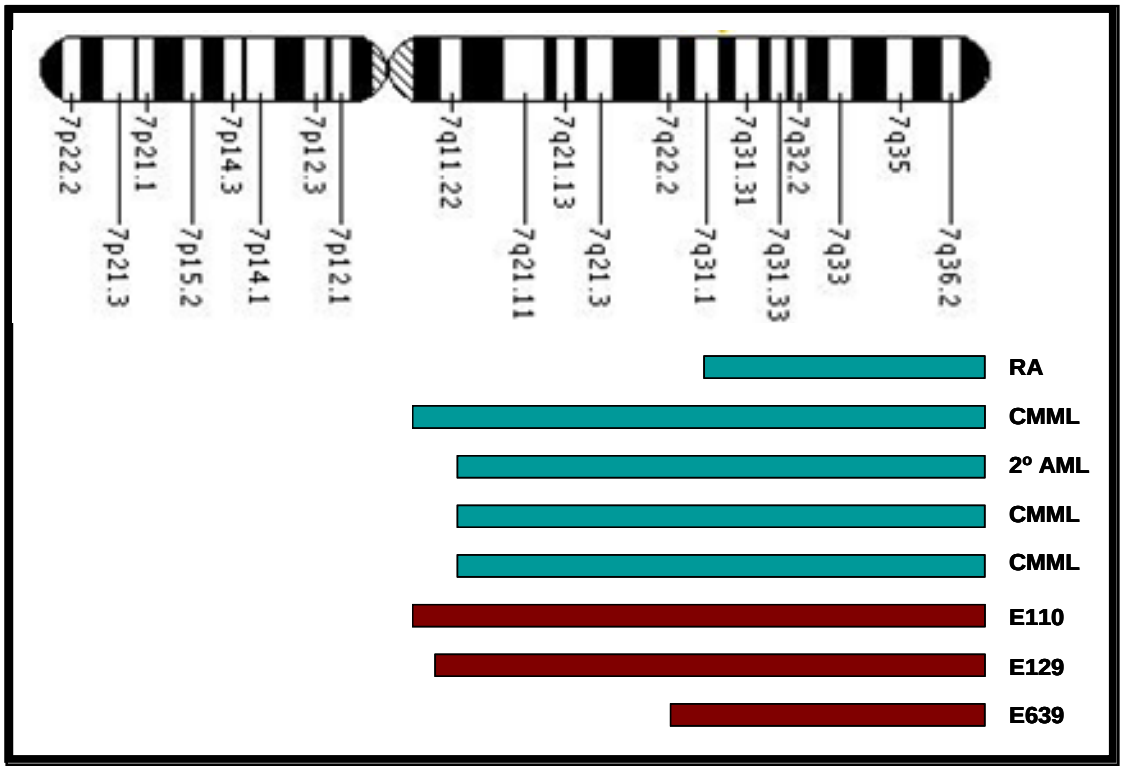


Figure 5. 26: Schematic detailing chromosome 7 and the regions of aUPD identified by Gondek & colleagues (2008), highlighted in green and my aCML/aMPD patients highlighted in red.

The smallest area of overlapping aUPD is now represented by the Gondek patient with RA, but the minimal region is only marginally smaller than that defined by E639. It will be important to analyse more cases to try and narrow the region down and it might be expected that with enough cases the proximal end of the minimal region might converge to a point close to the key gene (Figure 5.27). However this may well require analysis of a very large number of cases.

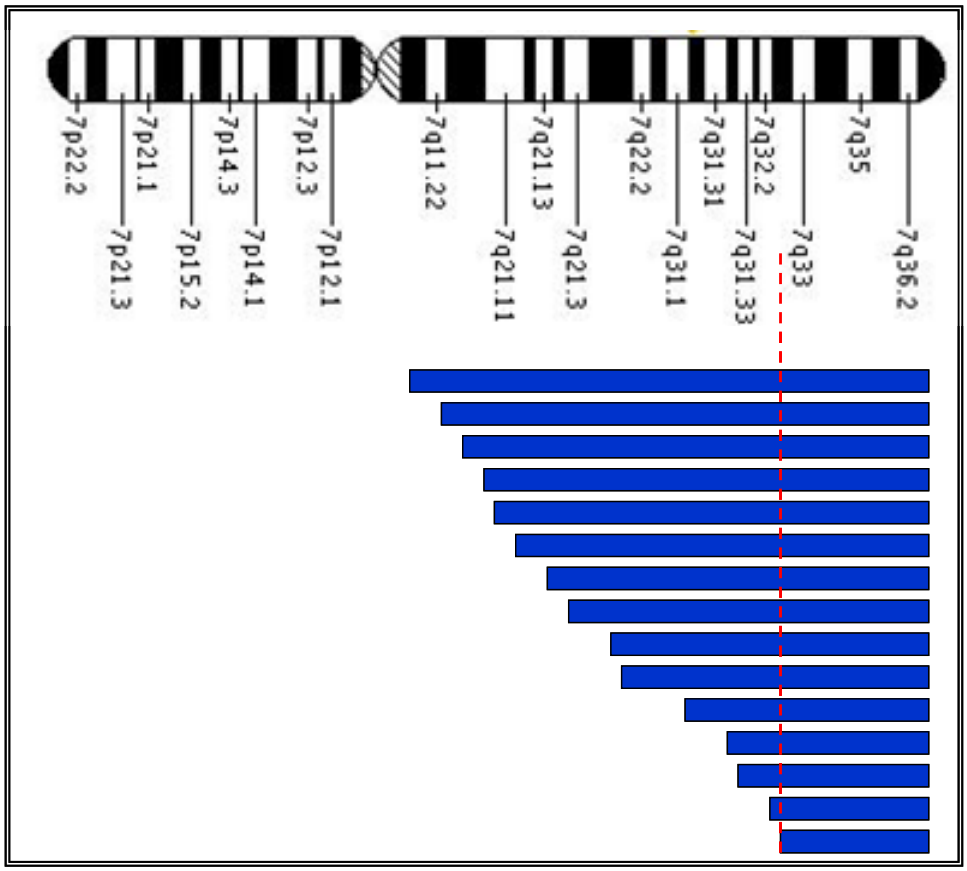


Figure 5. 27: Schematic detailing a theoretical analysis of chromosome 7 and how identification of more regions of aUPD might narrow down the region and converge at the proximal end (indicated by the red dashed line) near the gene of interest.

Whether the underlying abnormality is the same as the elusive 7q candidate gene believed to be at the root of recurrent deletions in AML is unclear, but since aUPD for 7q is uncommon in AML (Fitzgibbon, *et al* 2005, Gondek, *et al* 2008, Raghavan, *et al* 2005) it may well be that the two abnormalities are different. Over the years much work has been performed in myeloid disorders with 7q deletions and translocations trying to narrow down a minimal region in the hope of highlighting a specific mutated oncogenic or tumour suppressor gene. Some people have suggested that the complexity of 7q rearrangements might simply indicate that this is a region prone to breakage and instability, and thus a synergy of different genetic factors may play a general role in oncogenesis rather than a gene harbouring a specific mutation. However it seems more likely that there is a specific target for the cases affected by

aUPD. Interestingly, one of the 7qUPD cases (E110) was found to have a heterozygous deletion of the *CBL* gene. Assuming that this *CBL* abnormality is functionally important, this might indicate that the 7q gene functionally co-operates with *CBL* to produce an aCML phenotype.

Although the analysis of further cases by arrays will help to define a smaller minimal region, identification of the critical target will probably require the sequence analysis of a large number of genes to identify mutations that are homozygous in cases with 7q aUPD and absent or heterozygous in cases without 7q aUPD. Although this could be achieved by the stepwise of candidate genes such as those listed in Table 5.8, very soon it should be possible to systematically sequence all genes in the minimal region using a combination of pull-down hybridisation arrays covering all 7q exons in the critical region and next generation sequencing, which currently has the capacity to sequence approximately 10^8 -10^9 bases per run depending on the platform employed (Chan 2005, Mardis 2008, Morozova and Marra 2008, Schuster 2008, von Bubnoff 2008, Wheeler, *et al* 2008).

5.5.3 11q aUPD

Chromosome 11 has also been reported in the literature as a region commonly affected by imbalances in human malignancy. Reports of LOH are far more frequent than deletions, and have been associated with cervical, ovarian, breast, colon and skin cancers (Bethwaite, *et al* 1995, Carter, *et al* 1994, Foulkes, *et al* 1993, Hampton, *et al* 1994, Keldysh, *et al* 1993, Tomlinson, *et al* 1993). Nevertheless, deletions and translocations of 11q appear to be more common than LOH in haematological malignancies. LOH has rarely been reported to be associated with leukaemia and so the identification of a recurrent region of 11q aUPD in aMPD patients is particularly interesting and suggests an abnormality specifically associated with an MDS/MPD phenotype.

The smallest overlapping region of aUPD was represented by patient E484 and spans 57Mb covering cytogenetic bands 11q13.5-qter. This covers almost three quarters of the long arm of chromosome 11 and is thought to comprise greater than 350 genes. I had anticipated that the identification of regions of aUPD might provide a pointer towards novel tyrosine kinase mutations, however there are no tyrosine kinase genes on 11q and the three at 7q were eliminated by sequence analysis. A number of candidate genes were then short-listed according to association with leukaemia, and screening for the MLL partial tandem duplication (PTD) was carried out as a priority, but this was found to be negative in all three cases with 11q aUPD. Because of the strong association of MPDs and deregulated tyrosine kinase signalling, the next step was to look at genes that are known to impact on this process.

The *CBL* gene is located at 11q23.3 and encodes an ubiquitin ligase which functions as a negative regulator of several receptor protein tyrosine kinase signalling pathways. It acts by inducing ubiquitination of the RTK, thus promoting their sorting for endosomal degradation or recycling, however CBL also acts in a positive way by serving as a scaffold for the recruitment of downstream signalling substrates. CBL was an obvious candidate for mutation screening, particularly as two papers were published that reported the discovery of rare *CBL* mutations in AML patients (Caligiuri, *et al* 2007, Sargin, *et al* 2007). Based on the clustering of mutations in these cases I screened the three 11q aUPD cases and found that two had changes in exon 8 predictive of an L380P missense mutation whilst the third had a mutation of exon 9, P417A.

To determine if these mutations were more widespread among the MPDs I screened 547 patients, including cases with PV, IMF, ET, SM, HES, CMML and aMPD. I identified a further 23 patients (total 26 including 11q aUPD patients) with *CBL* mutations, 15 of that were missense and one nonsense as illustrated on Figure 5.28. One patient had a mutation in both exons 8 and 9 so in total there were 15 exon 8 mutations and 12 exon 9 mutations. The mutations were strongly associated with CMML (12.8%) and aCML/aMPD (7.9%) and in 3 cases I was able to demonstrate that they were acquired. Further analysis is being taken by other members of the

group to try and determine the specific clinical features associated with *CBL* mutations and their impact on prognosis.

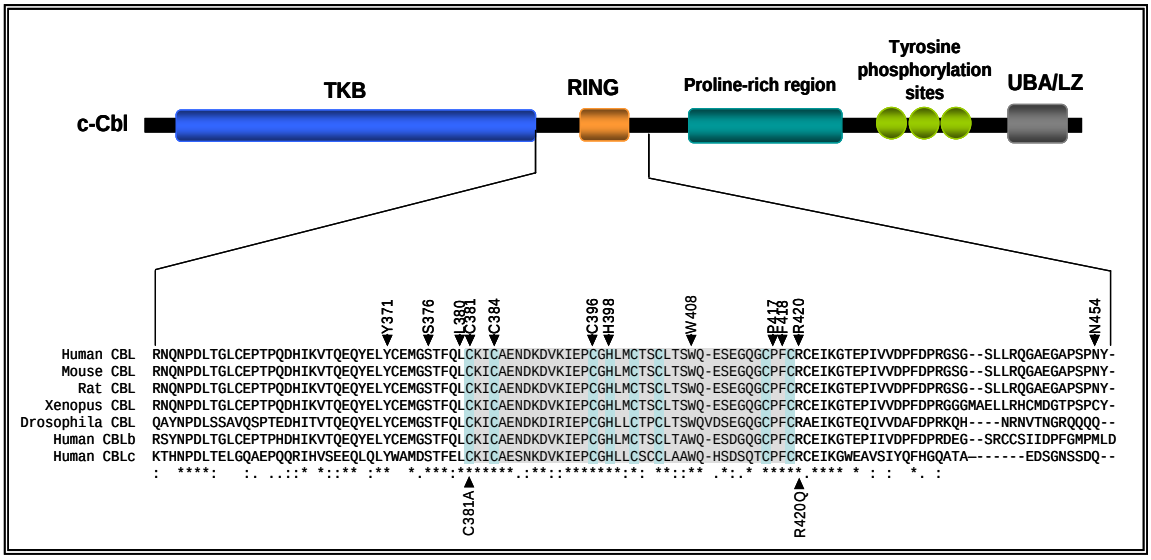


Figure 5. 28: Location of CBL missense mutations. Top panel: schematic showing the protein structure of CBL detailing the tyrosine kinase binding (TKB), RING, proline-rich (P-rich) and ubiquitin-associated/leucine zipper (UBA/LZ) domains. Bottom panel: ClustalW alignment of the entire linker and RING domains plus part of the distal sequence. The ring domain is shaded in grey. Residues affected by missense mutations identified by this study to date are indicated by arrowheads above the clustalW alignment. Those underneath the alignment are previously reported mutations.

I found four cases with R420Q, the same mutation described by Sargin et. al., (2007) but the majority of mutations was novel. Codon 420 is highly conserved in all CBL proteins and is the predicted contact point of interaction between CBL and the E2 ubiquitin-conjugating enzyme (Zheng, *et al* 2000). Functional work by Sargin et al. showed this mutant co-operates with overexpression of FLT3 to induce growth factor independence of 32D cells. R420Q prevents FLT3 receptor internalisation and ubiquitination and is the first description of a *CBL* mutation that is very likely to be a causative event in the development of human leukaemia. There have also been previous reports of animal models with oncogenic mutations in the RING finger

domain or linker sequence of *CBL*, at codon 381 (C381A) (Andoniou, *et al* 1994, Thien, *et al* 2001). I also discovered a mutation of the same codon (C381Y).

It is thought that alteration of signals for sorting tyrosine kinases to the lysosomes may contribute to the progression of neoplasia or even be an initiating event (Bache, *et al* 2004, Peschard, *et al* 2004). For example, several oncogenic variants of EGFR, MET, CSF1R and KIT have been shown to escape lysosomal degradation by a number of methods such as loss of PTB-mediated CBL binding sites (Peschard and Park 2003), or through the formation of oncoproteins caused by chromosomal translocations. The *TPR-MET* (Park, *et al* 1986) fusion gene, found originally in a mutagenised human osteogenic sarcoma cell line, results in deletion of sequence encoding the MET juxtamembrane domain that includes a binding domain site for CBL. As a consequence, the TPR-MET fusion escapes ubiquitination (Peschard, *et al* 2001) suggesting that its inability to be targeted to the lysosomes for degradation may contribute to its pathogenicity. Analysis performed by Sargin & colleagues proved that R420Q prevented ubiquitination and internalisation of not just FLT3 but also PDGFR and EGFR, suggesting that its oncogenicity might be mediated by failure to properly regulate many different receptors.

The critical analysis that is missing from my study is the functional determination of the consequences of the missense mutations plus the abnormalities leading to skipping exon 8. It was my intention to undertake this analysis but there was insufficient time and this is now being undertaken by another member of the group. The work is being conducted in collaboration with Sargin et al. who have provided the key reagents: pMY based plasmids expressing wild type and R420Q CBL (as a positive control) plus parental 32D cells and 32Ds expressing wild type FLT3. 32D cells are murine myeloid precursor cells that normally require IL-3 for growth but 32D/FLT3 can be propagated in the presence of FLT3 ligand.

Preliminary studies have included transfection of the WT CBL and R420Q expression vectors into 32D and 32D/FLT3 and show they are transforming in our

hands, as well as producing stable clones. To do this, a retroviral transfection protocol was used. The pMY vectors were developed from pMX (originally based on MuLV), which utilise pCMVs long terminal repeat (LTR) required for replication, integration, and expression of viral RNA. They also express green fluorescent protein (GFP) and thus transfected cells can be selected by fluorescent-activated cell sorting (FACS). The selected cells are expanded and proliferation assays (MTS assay and manual cell counting after trypan blue staining) performed after starving the cells of FLT3 ligand overnight. Sargin et al. showed that the R420Q mutant expressed in 32D/FLT3 continued to proliferate under these conditions whereas all other combinations failed to proliferate.

Sargin et. al., also demonstrated that R420Q resulted in prolongation of signalling. This will be addressed by starving transfected cells of cytokines and then stimulating them briefly with FLT3 ligand. Samples will be taken over a time course and Western blotting used to look at levels of phospho-ERK compared to total ERK. Other substrates will also be examined; in particular STAT5 which is known to be an important tyrosine kinase substrate in haematopoietic cells (Kato, *et al* 2005, Spiekermann, *et al* 2002).

Sargin et al. also demonstrated that the R420Q was deficient in ubiquitination activity and thus the mutant has lost its ability to negatively regulate tyrosine kinases by inducing receptor turnover but has presumably retained its ability to act as a scaffold for the recruitment of other signalling proteins. Loss of ubiquitin ligase activity has been confirmed for some of the mutants by transiently transfecting them into 293T cells along with FLT3 and HA-tagged ubiquitin. After 48 hours, FLT3 ligand was used to stimulate the receptor which was then immunoprecipitated and blotted with anti-HA. The strength of the signal indicated the degree of ubiquitination which was found to be greatly reduced for most CBL mutants (Grand FH 2009).

Another aspect that needs further investigation is functional consequences of homozygosity of *CBL* mutations. Both functionally important mutations that had

been identified previously (R420Q and C381A) were believed to act in a dominant negative fashion, however the finding of aUPD strongly suggests that there is a further selective advantage by reduction to homozygosity and that therefore the mutations are co-dominant. Following the array analysis, I tested *CBL* mutation positive cases by a combination of microsatellite PCR and MLPA and found a further 6 of the 26 patients exhibited 11q aUPD. Of the 9 patients in total (includes the 3 aMPD patients diagnosed by SNP array) 5 had aCML/aMPD, 3 CMML and 1 IMF (Table 5.17).

Patient Identifier	**CBL mutation**
E484	L380P
E632	P417A
E1191	L380P
E001	S376F
E879	L380P
E2167	C396G
1457	C384R
1617	H398Y
03/2231	W408C

Table 5. 17: Nine patients with 11q aUPD and a CBL mutation.

The majority of mutations associated with 11q aUPD are found in exon 8 of *CBL* gene with the exception of E632 (P417A). It also appears that the L380P mutation is significantly associated with homozygosity as all three patients identified with this

mutation also exhibited 11q aUPD. This is not true of the other cases where the same mutation or codon has been altered, such as codon 417 and 384, for which some cases were homozygous and others were heterozygous. Although it is conceivable that another, unknown locus on 11q is the target of aUPD, the finding of one case with a Y371H mutation and a stop codon on opposite alleles strongly suggests that homozygous *CBL* mutations do indeed confer a direct selective advantage.

There may be an analogy between the *CBL* mutations and JAK2 *V617F*, which is also subject to reduction to homozygosity by aUPD in a similar proportion of cases. Homozygosity for V617F is relatively common in PV and MF but rare in ET however many patients with all three disease types are heterozygous. When large numbers of cases are analysed it appears that homozygosity is associated with a slightly more aggressive disease and it would be interesting to determine if cases with homozygous *CBL* mutants were also more aggressive. Interestingly, most 'heterozygous' PV patients have a proportion of haematopoietic colonies that are homozygous for V617F. This is not seen in ET and thus strengthens the association between V617F zygosity and disease phenotype (Scott, *et al* 2005). It will be interesting to test haematopoietic colonies from *CBL* mutated cases to see if they are homozygous. This might indicate that the observed heterozygosity was only apparent and in fact caused by a mixture of homozygous mutated cells with normal cells that are not part of the malignant clone.

Finally there is the question of whether *CBL* mutations are sufficient to give rise to disease by themselves, or if they require other co-operating genetic events. The requirement for over-expression of a tyrosine kinase for transformation of 32D cells may simply reflect the paucity of receptor tyrosine kinase gene expression in this cell line or, alternatively, may indicate that the mutants are only weakly transforming and require the co-operation of other events to give rise to clinically manifest disease. Consistent with the latter hypothesis, I identified two cases in which *CBL* mutations were acquired following progression of a pre-existing MPD. One of these presented with a relatively low level V617F *JAK2* mutation that disappeared on progression to

CBL L380P positive disease, indicating that the two mutations must have arisen independently in different clones. It remains to be established, however, if *CBL* mutations are usually secondary events or whether the two cases we identified are exceptional.

Further evidence for the requirement for additional co-operating genetic events comes from published mouse studies in which neither complete *CBL* knockout nor heterozygous knock-in of the RING finger inactivating mutant C379A (mouse equivalent of C381) resulted in evidence of an MPD or other malignancy (Rathinam, *et al* 2008, Thien, *et al* 2005). However closer analysis revealed subtle haemopoietic perturbations that clearly relate to the pathogenesis of MPDs. Specifically, haemopoietic stem cells (HSC) from these animals were increased in number, hyper-responsive to thrombopoietin and more potent than wild type HSC in repopulating the haematopoietic system. These effects were associated with increased STAT5 accumulation and phosphorylation. Hyper-responsiveness to growth factors and stem cell involvement are two of the cardinal features of MPDs. Furthermore, the association between STAT5 activation and MPDs is well documented (Cain, *et al* 2007, Kotecha, *et al* 2008). Whilst it is possible that overt haematological disease might have been induced by a homozygous knock-in mutant, these findings suggest that *CBL* mutations may not be solely responsible for a full MPD phenotype. Possibly this could be tested further by transfecting bone marrow cells from *CBL* knockout mice, which are viable and fertile (Murphy, *et al* 1998). Nevertheless, since the CBL mutants appear to be acting by enhancing tyrosine kinase signalling, it may prove possible to treat patients in the future with tyrosine kinase inhibitors.

6 DISCUSSION

The unifying theme of my research is the development and use of different approaches to identify deregulated tyrosine kinases and associated signalling components in MPDs. The paradigm of *BCR-ABL* in CML paved the way for the identification of other deregulated tyrosine kinase genes in the MPDs, although it took several years for the association between MPDs and deregulated signalling to emerge. Initially I focused on the analysis of chromosome translocations, subsequently moving onto screens for over-expression of selected tyrosine kinase genes and finally to genome-wide approaches that took into account the fact that genes other than tyrosine kinase may be responsible for the substantial proportion of MPDs of unknown pathogenesis.

6.1 Novel fusion genes and roles of partner proteins

In this thesis I have described the identification of six novel fusion genes. Four of these were completely novel with partner genes identified by me (*GOLGA4*, *STRN* and *CPSF6*) or by another member of the group (*KIF5B*) that have not been previously described as being involved in leukaemia. All four are very rare and as yet only *GOLGA4-PDGFRB* has been found in more than one patient. The fifth resulted from fusion of a known partner protein, *ETV6,* to *PDGFRA.* The final fusion was a variant *ETV6-PDGFRB* fusion with rare breakpoints never reported before.

Five of the six fusions involved *PDGFR* genes and are therefore diagnostically important because of the known sensitivity of both PDGFRα and PDGFRβ to targeted therapy with the tyrosine kinase inhibitor imatinib. Indeed all five received imatinib, of who four achieved complete cytogenetic remission, and for the fifth case the follow up is still too short to assess remission status. The remaining fusion involved the *FGFR1* gene, known to be involved in a specific MPD referred to as

8p11 myeloproliferative syndrome (EMS) that is refractory to imatinib but may be sensitive to inhibitors that are active against FGFR1 (Chase, *et al* 2007) although none are currently available for clinical use.

Fusions of *PDGFRA*, *PDGFRB* and *FGFR1* are all known to be associated with MPDs. Translocations involving *PDGFRB* appear the most diverse with reports of partner genes at more than 20 chromosomal locations (Figure 6.1). Only one, *ETV6-PDGFRB*, has been found relatively frequently with the remainder being very rare and many only described in one or two patients. *FGFR1* in EMS has been reported to fuse to 8 partners in addition to the one I found. The three *PDGFRA* fusions I found take the current total to six, by far the most common of which is *FIP1L1-PDGFRA*.

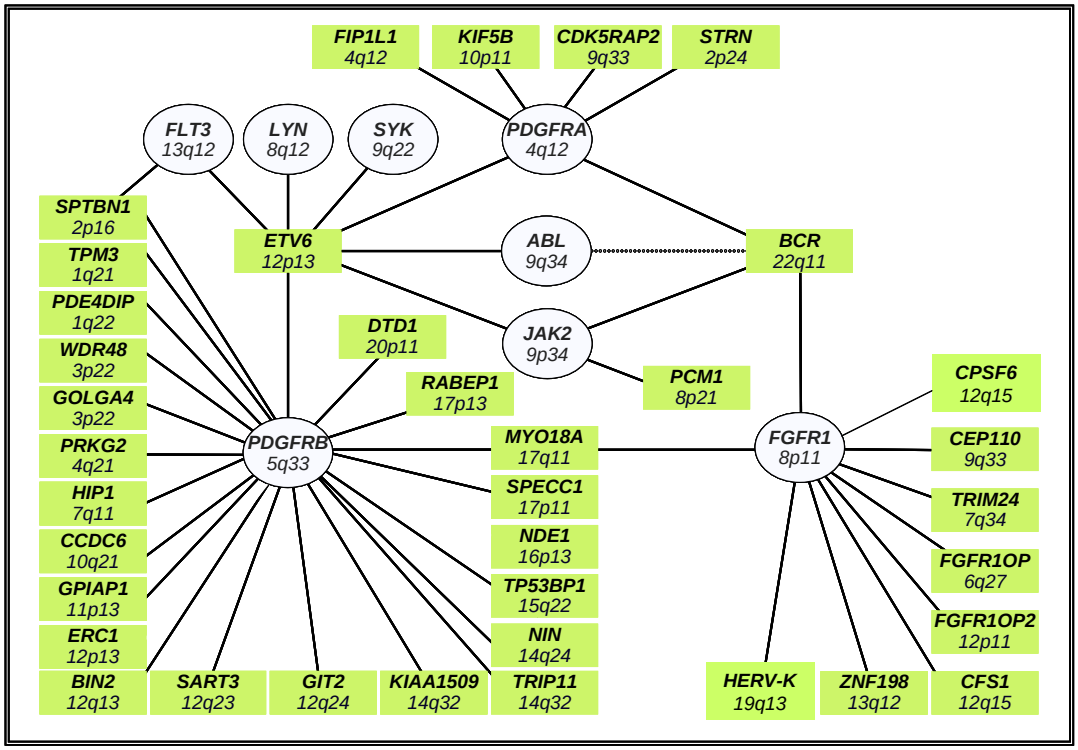


Figure 6. 1: Spider diagram illustrating translocations associated with MPDs and related disorders. The tyrosine kinase genes are indicated in white and the fusion partner genes in green.

I took a number of different approaches to identify the fusion genes:

- Both *GOLGA4-PDGFRB* and *CPSF6-FGFR1* were identified following initial FISH analysis to confirm the involvement of the respective tyrosine kinase genes, both of which were suspected from the cytogenetic results and clinical phenotype. Once this was confirmed I performed 5'RACE PCR from the kinase to identify the partner gene.

- *STRN-PDGFRA* and *ETV6-PDGFRA* were cloned by bubble PCR from genomic DNA, exploiting the fact that the genomic breakpoints for all reported *PDGFRA* fusions are very tightly clustered. This approach was particularly useful for the *STRN-PDGFRA* case since no RNA was available, and in fact the fusion was cloned from DNA extracted from a single bone marrow slide. As above, involvement of *PDGFRA* in both cases was suspected by the finding of a translocation involving 4q12 in patients with a MPD and eosinophilia.

- *KIF5B-PDGFRA* was identified following a screen for over-expression of *PDGFRA* by the generic RT-PCR assay I developed. Of 200 patients screened, just one strongly over-expressed *PDGFRA* (although unknown to us at the time, the patient had a complex rearrangement involving 4q12). The fusion was identified by bubble PCR from genomic DNA.

- The variant *ETV6-PDGFRB* case was referred with a t(5;12) karyotype and a clinical response to imatinib, strongly suggesting involvement of *PDGFRB*, however RT-PCR for *ETV6-PDGFRB* was negative, Over-expression of *PDGFRB* was confirmed by the MLPA RT-PCR assay and therefore I performed 5'-RACE PCR leading to the unexpected finding of the variant fusion.

All the fusions I found conformed broadly to the consensus structure of tyrosine kinase fusions following the paradigm of BCR-ABL (Figure 6.2). The N-terminal region of the partner gene was fused in-frame to the C-terminal region of the kinase, including the entire catalytic domain.

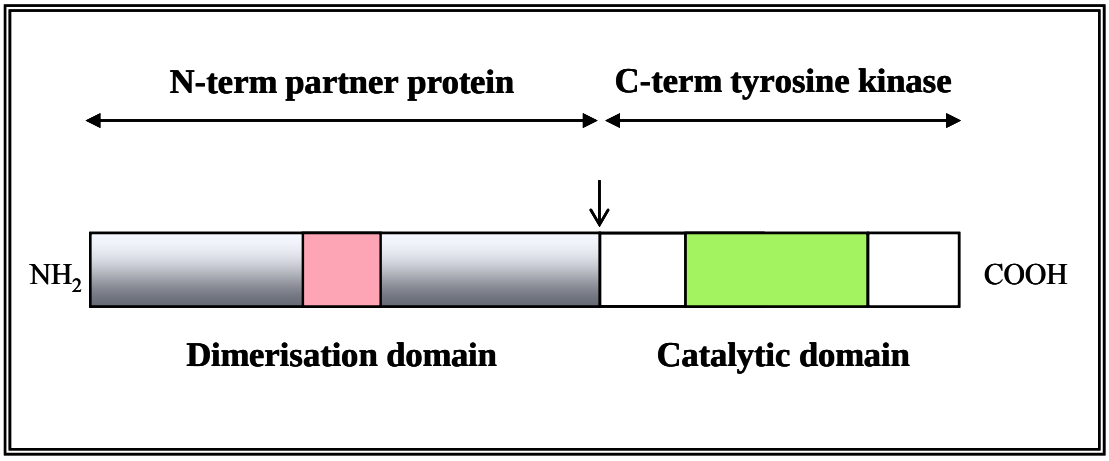


Figure 6. 2: Schematic showing the universal structure of tyrosine kinase fusions based on the BCR-ABL paradigm. The N-terminal dimerisation domain of the partner is fused to the entire catalytic C-terminal domain of the tyrosine kinase.

The partner genes are widely expressed, usually in all normal tissues but certainly in haematopoietic cells and thus provide a promoter to drive transcription of the fusion. All partner proteins analysed to date (except FIP1L1) contain one or more dimerisation domains that are also required for the transforming activity of the fusion proteins. Homotypic interactions between these domains lead to dimerisation or oligomerisation of the fusion protein, thus mimicking the normal process of ligand-mediated dimerisation that is essential for the activation of normal tyrosine kinases. Although I was unable to test my fusions functionally, all the partner proteins had plausible dimerisation domains that are retained in the fusion genes, e.g.

- GOLGA4: extensive coiled-coil domains (Barr 1999, Kjer-Nielsen, *et al* 1999, Lu, *et al* 2006, Luke, *et al* 2003, Munro and Nichols 1999, Panic, *et al* 2003, Wu, *et al* 2004).

- STRN: an N-terminus coiled-coil region is thought to be crucial for homo- and hetero-oligomerisation of the protein and hence of the fusion protein (Gaillard, *et al* 2006).
- CPSF6: no classic domains but the RNA recognition motif (RRM) may mediate homodimerisation (Handa, *et al* 2006, Simpson, *et al* 2004).
- KIF5B: a central alpha-helical rod domain that enables the protein to dimerise (Huang, *et al* 1994).
- ETV6: the pointed (helix-loop-helix) domain has been established functionally as a self association motif that is essential for the transforming activity of ETV6-PDGFRβ (Carroll, *et al* 1996, Golub, *et al* 1994).

The normal functions of the partner proteins are diverse, however with the characterisation of increasing numbers of tyrosine kinase fusions, some intriguing patterns have emerged. Indeed it is possible that the partner proteins are not just contributing oligomerisation motifs and ubiquitous promoter expression but may also be playing a more direct role in cell dysregulation, as discussed below.

Deregulated endocytosis. Several translocations involving *PDGFRA* or *PDGFRB* generate in-frame fusions with components of the endocytic machinery e.g. *Rabaptin-5*, *Hip1* and *PDE4DIP*. My work adds three further endocytosis-related partners, *GOLGA4, KIF5B* and *STRN* and, in addition, a further two *PDGFRB* partner genes submitted for publication (*Bin2* and *WDR48*) identified in our laboratory also have endocytic links. Endocytosis is recognised as a fundamental mechanism of signal attenuation by committing receptors to degradation or recycling back to the plasma membrane. Potentially, interference with endocytosis may affect normal receptor down-regulation, resulting in hyper-proliferation (Mosesson, *et al* 2008). Furthermore, if signalling can occur at the endosomes, as has been reported by Wang Y et al., (2004) for Rabaptin5-PDGFRβ, then this may suggest a mechanism for prolonged signalling (Wang, *et al* 2004a).

GOLGA4 (tGolgin1; Golgin-245). tGolgin-1 is a member of the Golgin family of coiled-coil proteins associated with specific sub-domains of the *trans*-Golgi network (TGN) and is thought to play a role in Rab6-regulated tethering events in membrane fusion and as a structural support for the Golgi cisternae (Barr 1999, Kjer-Nielsen*, et al* 1999, Munro and Nichols 1999).

The TGN itself is a series of interconnected tubules and vesicles located at the *trans* face of the golgi stack and its purpose is to process and sort glycoproteins and glycolipids at the periphery of the endosomal pathway (Figure 6.3A). The Golgins are further characterised by a highly conserved C-terminal GRIP domain that localises the proteins to the TGN. Over-expression of the isolated GRIP domain has been shown to result in alteration of the pericentriolar distribution of the TGN integral membrane and coat components whilst concomitantly inhibiting vesicular transport from the TGN to the plasma membrane (PM). This suggests that these proteins help maintain the integrity of the TGN by regulating resident protein localisation (Yoshino*, et al* 2003). tGolgin-1 is suggested therefore to be a tubulovesicular carrier that emerges from the TGN and hence is likely to be a regulator of transport from the TGN to the recycling endosome and PM. Recently Lieu et. al., (2008) have further defined the role of tGolgin-1 and shown that not only is it essential for intracellular trafficking but it might additionally act as a vesicle coat, however its complete function still remains to be fully elucidated (Lieu*, et al* 2008).

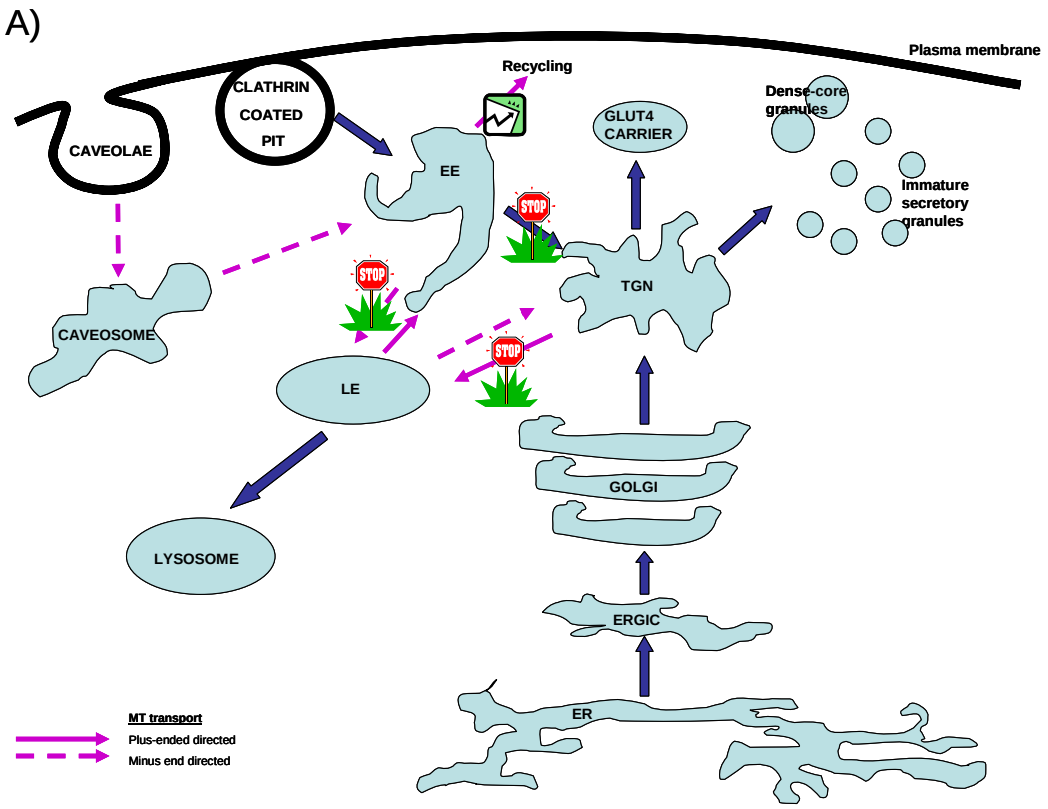


Figure 6. 3A: Schematic showing anterograde and endocytotic/retrograde transport through cellular organelles. 'Stop signs' indicate which pathways might be affected by the GOLGA4-PDGFRβ fusion, leading to possible aberrant endocytosis and increased cellular transformation. tGolgin-1 is a tubulovesicular carrier likely to be a regulator of transport from the TGN to the recycling endosome and PM. This is a retrograde pathway that allows signalling proteins to escape the degradative environment of the late endosomes and lysosomes and to reach internal sites where they can interact with their intracellular targets or be exported back to the PM via the recycling endosomes permiting increased signaling (represented by arrow in green box). EE, early endosome; ERGIC, ER-Golgi intermediate compartment; Glut4, a glucose transporter molecule; LE, late endosome; RE, recycling endosome, and TGN, *trans*-Golgi network.

It can only be speculated upon how the fusion of tGolgin-1 and PDGFRβ might abrogate endocytosis and enhance the oncogenic potential of the fusion. One such hypothesis is that the partner protein plays a role in localising the fusion to a particular subcellular compartment, in this case the TGN. However for this specific

fusion the c-terminal GRIP domain (essential for localisation to the Golgi) is lost and it is unclear where in the cell the fusion would be. It has also been shown that in general tyrosine kinase fusions are found to be cytosolic due to deletion of the N-terminal signalling peptide from the kinase (Lamorte and Park 2001). Hence they avoid entry to the endocytic pathway anyway, thereby presumably preventing degradation and prolonging signalling.

A second hypothesis concerns the method of uptake of the fusion/constitutively activated receptor. It has been suggested that differential clathrin and caveolae mediated endocytic routes of the same cargo may regulate signalling capacity and endocytic fate (Di Guglielmo*, et al* 2003, Sigismund*, et al* 2005). Clathrin mediated endocytosis is a form of membrane invagination and vesicle scission of receptors clustering in clathrin coated pits. The released vesicles discard the coat before fusing with tubulovesicular compartments such as the early endosomes where the cargo is sorted for either degradation in the lysosomes or recycling back to the PM. The caveolae pathway also involves membrane invagination and vesicle budding, however the vesicles are caveolin-positive and are directed first to larger caveolin-enriched organelles termed caveosomes before being further targeted to the early endosomes, thereby allowing convergence with the clathrin mediated pathway. Both pathways are a means of signal attenuation however it is possible that one or the other might promote recycling over degradation further compounded by the role of tGolgin-1 as regulator of transport from the TGN to the recycling endosome and plasma membrane.

Finally and perhaps most interestingly, is the role of tGolgin-1 in retrograde transport from the endosomes to the TGN. This pathway is regulated by specific SNAREs and Rab proteins signifying it is a pre-existing cellular pathway not just a subversive pathway. This retrograde pathway may allow signalling proteins to escape the degradative environment of the late endosomes and lysosomes to reach internal sites where they can interact with their intracellular targets or be exported back to the PM via the recycling endosomes (Figure 6.3A) (Mallard*, et al* 2002). Therefore, efficient

retrograde transport from the early endosomes to the TGN is important for the recycling of endogenous proteins (Lieu, *et al* 2007). A role for Golgin-1 in this pathway has been proposed by Yoshino et. al., (2005), who showed that recruitment of microtubule motors (such as the kinesins) to the Golgi requires tGolgin1-dependent retrograde transport from the endosomes. This was observed after experiments over-expressing the GRIP domain were performed that resulted in specific disruption of TGN morphology, protein localisation and function. Concomitantly, vesicular transport from the TGN to the PM was inhibited (Yoshino, *et al* 2005). This suggests the GRIP domain proteins, such as tGolgin-1 function in TGN maintenance probably by regulating recycling from the endosomes. Thus the TGN requires a functional GRIP domain, shown not only by over-expression of GRIP domains competing for binding, but also by inactivity observed with N-terminal tGolgin-1 fragments that lack the GRIP domain, such as might be observed with the GOLGA4-PDGFRβ fusion. This would imply the fusion may not be recycled back to the PM but instead retained at the TGN or in the early endosomes where conceivably signalling can still occur, aiding oncogenicity.

Yoshino et. al., 2005 have also shown that cells depleted of tGolgin-1 mis-sort internalised glycolipids to the late endosomes/lysosomes and accumulate aberrant endosomal structures, highlighting a possible secondary outcome of disrupting membrane recycling. These results are consistent with the function of golgin-97 (Lu, *et al* 2004a) and the dominant-negative GRIP-domain over-production in which endosome to TGN recycling is inhibited (Yoshino, *et al* 2003). The consequence of such mis-sorting to the late endosomes would be eventual depletion from the Golgi of recycling glycosphingolipids, so proteins and receptors cannot be modified and shipped back to the PM. This could mean prolonged signalling and reduced degradation and could also perturb multiple other pathways if required proteins are not modified or transported where required, adding to cellular transformation.

KIF5B. KIF5B is an N-terminal (plus-end motor) kinesin that is a member of the superfamily of Kinesin-1 molecular motor proteins that are responsible for

microtubule-dependent transport of cargo in eukaryotic cells (Goldstein and Yang 2000). These motor proteins are powered by the hydrolysis of ATP enabling them to move cargo such as vesicles, proteins and lipids throughout the cell. The plus-end and minus-end microtubule specific motors are responsible for opposing movement along the filaments within the cell (Hirokawa and Takemura 2005). *KIF5B* in particular, is important for microtubule-based endosome-motility, by regulating early endosome trafficking and has been further implicated specifically in lysosomal and mitochondrial transport (Nakata and Hirokawa 1995, Tanaka*, et al* 1998).

It is important to highlight the link between the Golgi and the motor protein Kinesin to comprehend how these proteins found fused to a PDGFR tyrosine kinase are related not only to endocytosis but also how these proteins interact at the cellular level. As discussed previously, proteins sequestered at the TGN are selectively sorted and packaged into distinct carrier vesicles destined for either the PM or endosomal compartments. The Golgi is a dynamic structure and to maintain this level of activity it is sustained by a complex cytoskeletal matrix composed of microtubules and actin-spectrin networks as well as by temporal localised proteins such as the Golgins. The interaction between these filaments and the Golgi membrane is further mediated by diverse families of 'motor proteins' such as the kinesins. Interestingly, these two endocytic proteins interact at the Golgi and together help maintain the overall integrity of the Golgi structure whilst also supporting protein transport. Overwhelming evidence now suggests that *KIF5B* is specifically involved in transporting vesicles from the TGN to the PM or endosomes further localising both *KIF5B* and tGolgin1 to the same subcellular domain: the TGN. Additionally, if *KIF5B* expression is reduced the Golgi apparatus has been shown to switch from an extended ribbon to a much more compact structure (Feiguin*, et al* 1994) with ensuing inefficient protein modification and sorting, and decreased trafficking. An additional layer of complexity within the trafficking of vesicles from the TGN is the requirement of different motor proteins depending on the cargo being transported and the specific cell type, for example epithelial or fibroblasts. Due to the constitutive activity of the kinase caused by the *KIF5B-PDGFRA* fusion it is likely that *KIF5B* expression is enhanced. This may cause specific proteins to be modified and

transported far more efficiently than others, resulting in an unbalanced accumulation of specific proteins and a reduction in others creating generalised cell defects perhaps adding to the transforming potential of this fusion.

Another hypothesis for this fusion's oncogenicity is related to *KIF5B*'s role in plus-end motility of vesicles. It has been shown to work in association with the minus-end motility kinesin *KIFC1*, and if either Kinesin is inhibited reduced vesicle fission occurs suggesting that the opposing forces from the activity of both motors are required for fission to occur (Nath*, et al* 2007). This could imply that vesicles containing activated receptors are unable to be sorted for degradation to the lysosomes due to inefficient vesicle fission. The active receptors may then be able to continue signalling increasing the transformation potential and disrupting receptor attenuation via lysosomal degradation (Figure 6.3B).

KIF5B is a tetramer composed of two heavy chains and two light chains that together form a rod-like molecule consisting of two globular head domains, a stalk domain and a fan-like tail domain. The head domain is thought to bind to the microtubules, and the fan-like ends are thought to be associated with vesicles and membranous organelles requiring transport (Hirokawa*, et al* 1989). The N-terminal motor/head domain is present within the fusion with *PDGFRA* so it still retains the ability to move along the microtubules, however with the c-terminal fan-like region absent it may not be able to interact with the vesicles, organelles and proteins requiring intracellular transport. Transport is very important for the proper targeting of lipids, proteins and organelles to distinct subcellular domains without this function it is plausible that more generalised defects within the cell may occur. This also implies the fusion itself would remain cytosolic, possibly bound to the microtubule network, avoiding endocytosis resulting in increased signalling.

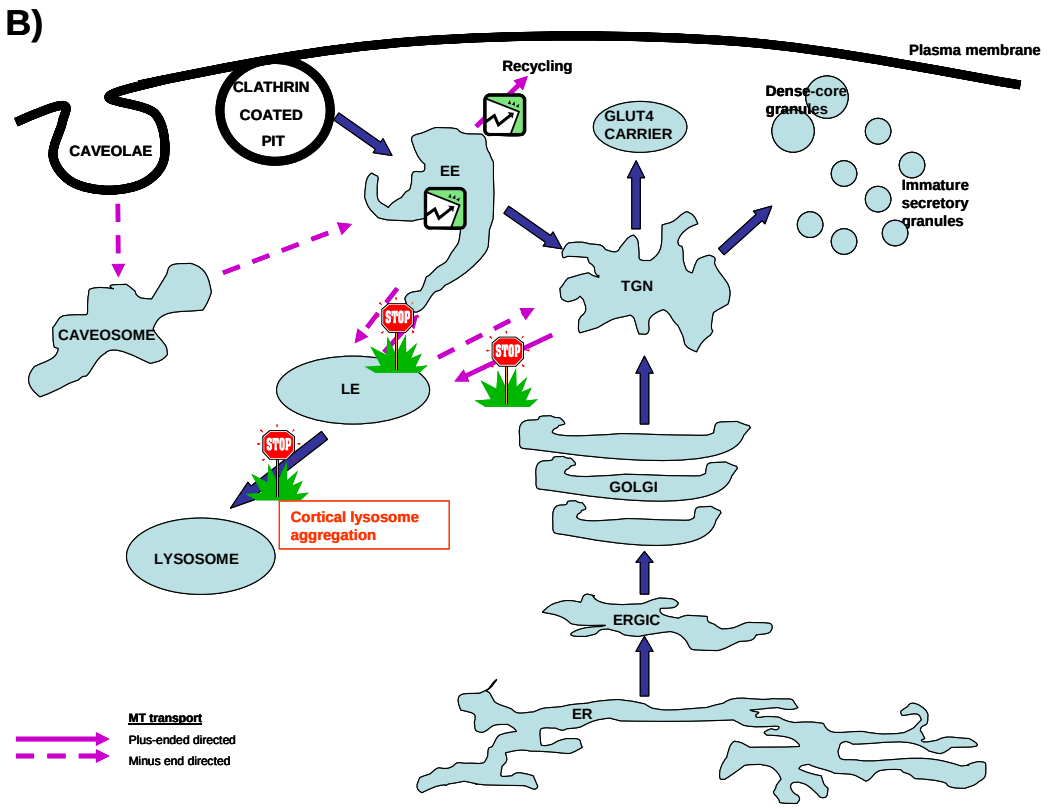


Figure 6.3 B: Schematic showing anterograde and endocytotic/retrograde transport through cellular organelles. 'Stop signs' indicate which pathways might be affected by the KIF5B-PDGFRα fusion, leading to possible aberrant endocytosis and increased cellular transformation. The plus-end and minus-end microtubule specific motors are responsible for opposing movement along the filaments within the cell, represented by purple arrows. KIF5B is specifically involved in TGN-to-PM or endosome trafficking, enhanced expression of this protein caused by the fusion may cause specific proteins to be modified and transported far more efficiently than others, resulting in an unbalanced accumulation of specific proteins and a reduction in others creating generalised cell defects. Vesicles containing the activated fusion may also be unable to be sorted for degradation to the lysosomes due to inefficient vesicle fission. The active receptors may then be able to continue signalling within the EE or be recycled (represented by arrow in green box) increasing the transformation potential and disrupting receptor attenuation via lysosomal degradation. It is also speculated that overexpression of *KIF5B* may result in continuous transportation to the required subcellular domains and subsequent enhanced signalling and cell turnover and reduced lysosomal degradation. EE, early endosome; ERGIC, ER-

Golgi intermediate compartment; Glut4, a glucose transporter molecule; LE, late endosome; RE, recycling endosome, and TGN, *trans*-Golgi network.

Additionally, Cardoso et. al., (2009) have recently reported up-regulation of *KIF5B* mRNA in several types of cancer including bladder cancer, advanced gastric cancer, squamous cell carcinoma and BRCA1-associated breast cancer. (Cardoso, *et al* 2009). They were also able to show that when *KIF5B* was depleted *in vitro* HeLa cells acquired an elongated cellular phenotype concomitant with significant growth inhibition and cell death. It could be speculated that over-expression of *KIF5B* may have the opposite effect with effective continuous transportation to the required subcellular domains and subsequent enhanced signalling and cell turnover. Further work from this group implicates that *KIF5B* may function in transporting lysosomes to the plasma membrane destined for exocytosis and moreover may play a role in positioning lysosomes at distinct sites in the cytoplasm, particularly perinuclear. Consequently, depletion of *KIF5B* would allow other motor proteins to transport the lysosomes to the cortical areas of the cell resulting in lysosomal aggregation.

STRN. Striatin is a multimodular, WD-repeat, and calmodulin-binding protein. It is found in the cytosol in addition to being associated with membranes, and is composed of several different, colinearly aligned protein-protein association modules. These include a caveolin-1 binding motif, a coiled-coil structure, a Ca^{2+}-calmodulin-binding motif and a large WD-repeat domain (Bartoli, *et al* 1998, Castets, *et al* 1996, Gaillard, *et al* 2001). Striatin is a scaffolding protein involved in calcium dependent signalling (Bartoli, *et al* 1999) with a putative role in trafficking (Baillat, *et al* 2001), whilst also being shown to directly bind caveolin-1 (Gaillard, *et al* 2001).

To further define the role of Striatin, Baillat et. al., 2001 used a yeast two-hybrid strategy to search for partners and identified phocein, an intracellular protein. Its distribution mirrors that of striatin, due to its association and a lack of transmembrane binding motifs. Interestingly, phocein is also highly localised to the Golgi complex indicating a link between this organelle and striatin (Baillat, *et al* 2001). It is hypothesised that phocein may be a component of a novel vesicular coat allowing

transport to the Golgi complex and further adding to striatins' role in membrane trafficking.

Additionally, it is speculated that caveolin-1 binding the caveolin binding site of striatin may directly regulate its function by concentrating it within caveolin-enriched domains, concordant with its putative role in membrane trafficking (Baillat*, et al* 2001). The caveolae themselves are thought to be subcellular domains that regulate among other things, intracellular Ca^{2+}- concentration and Ca^{2+}–dependent signal transduction (Isshiki and Anderson 1999). This links both the Ca^{2+}-calmodulin binding domain of striatin and the caveolin-1 binding domain. It is also well known that caveolins interact with signalling proteins from the plasma membrane, such as tyrosine kinases, Ras and G-proteins (Chang*, et al* 1994, Lisanti*, et al* 1994, Sargiacomo*, et al* 1993), and is consistent with a role for caveolae in sequestering signalling molecules to attenuate signalling after receptor stimulation, similar to clathrin coated pits. This further addresses the role of striatin as a possible signalling transduction component due to its links and enrichment at the caveolae. Furthermore, it has been shown that transformation of NIH3T3 cells by various oncogenes leads to a reduction in cellular levels of caveolin, suggesting functional alterations in caveolae may play a critical role in oncogenic transformation, perhaps by disrupting contact inhibition and negatively regulating the activation state of the p42/44 MAP kinase cascade (Engelman*, et al* 1997, Galbiati*, et al* 1998, Koleske*, et al* 1995). Additionally the reduction of caveolae in transformed cells could further prevent the down-regulation of growth-promoting signals, thus contributing to uncontrolled proliferation. It is possible that the STRN-PDGFRα fusion, by interacting with caveolin via the caveolin-1 binding site of striatin, is either maintained at the cell surface constitutively active, or is sequestered from the cell surface and internalised via the caveolae where it may be transported through the endocytic pathway and recycled due to its interaction with caveolin back to the cell surface thereby allowing continuous signalling and promoting further to cell transformation (Figure 6.3C).

In conclusion, it is possible this protein with its calmodulin-binding site a WD-repeat, scaffolding domain and a caveolin-1 binding motif enhances cell transformation, adding to the growing list of proteins now known to have a dual capacity in endocytosis and signalling (Di Fiore and De Camilli 2001).

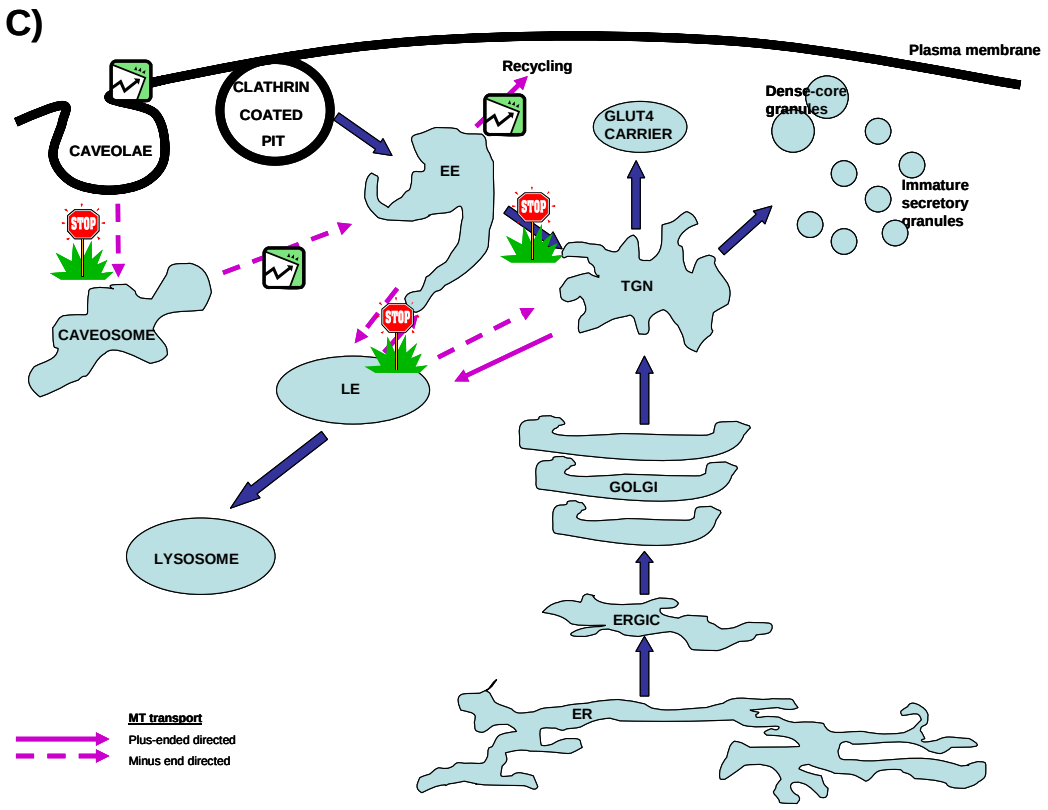


Figure 6.3 C: Schematic showing anterograde and endocytotic/retrograde transport through cellular organelles. 'Stop signs' indicate which pathways might be affected by the STRN-PDGFRα fusion and lead to possible aberrant endocytosis and increased cellular transformation. Reduction of caveolae in transformed cells is thought to prevent the down-regulation of growth-promoting signals, thus contributing to uncontrolled proliferation. It is speculated that the striatin caveolin-1 binding site within STRN-PDGFRα enables it to be either retained at the cell surface constitutively active (represented by arrow in green box), due to reduced levels of caveolae or is sequestered from the cell surface and internalised via the caveolae and recycled allowing continuous signaling whilst promoting cell transformation. EE, early endosome; ERGIC, ER-Golgi intermediate compartment; Glut4, a glucose transporter molecule; LE, late endosome; RE, recycling endosome, and TGN, *trans*-Golgi network.

Finally, the finding of mutations in *CBL*, a protein that is normally involved in attenuating receptor tyrosine kinase signalling, further strengthens the hypothesis that subversion of endocytic control is pathogenetically important. These mutations have been shown to prevent internalisation and ubiquitination of the RTKs FLT3, PDGFRα/β and EGFR, leading to prolonged signalling (Sargin*, et al* 2007).

Pre-mRNA processing. The characterisation of a second novel fusion gene: *CPSF6-FGFR1* in parallel with the discovery in our laboratory of a novel *ABL* fusion: *SFPQ-ABL* led us to the identification of a second functionally related group of partner proteins (Hidalgo-Curtis*, et al* 2008). Both *CPSF6* and *SFPQ* encode proteins that are involved in pre-mRNA processing. SFPQ is an essential factor involved in RNA splicing and was first identified in a complex with polypyrimidine tract-binding protein (PTB)(Patton*, et al* 1993). Interestingly, CPSF6 co-localises in purified spliceosomes with SFPQ and another splicing factor NonO, linking these proteins to both transcription and processing of pre-mRNAs (Dettwiler*, et al* 2004, Rappsilber*, et al* 2002, Zhou*, et al* 2002). ZMYM2 (ZNF198), a partner protein commonly fused to FGFR1, is also thought to be involved in this process, possibly by acting as a scaffolding protein that brings RNA-binding proteins together to facilitate pre-mRNA splicing (Kasyapa*, et al* 2005). Strikingly, ZMYM2 has also been found to immunoprecipitate with SFPQ (Kasyapa*, et al* 2005). Adding further weight to this group of proteins is FIP1L1, which fuses to PDGFRα in CEL (Cools*, et al* 2003a). Not only is this one of the most common fusion genes described in eosinophilic MPD patients, FIP1L1 has also been show to have a role in polyadenylation and pre-mRNA splicing (Kaufmann*, et al* 2004, Preker*, et al* 1995).

How these functionally related proteins might impact on neoplastic growth is not clear. It is possible that the fusions may interfere with normal pre-mRNA processing complexes thereby inhibiting processing of many transcripts. Alternatively they may result in increased processing. Both splicing and polyadenylation are known to be important control points with regard to gene expression and it is conceivable that

alterations could disrupt the fine balance between proliferation, differentiation and apoptosis of these cells by altering mature mRNA levels of key genes. Consistent with this idea, it has been shown that high levels of poly(A)polymerase (PAP), the enzyme that causes elongation of the poly(A) tail, are associated with proliferating cells and that the levels reduce in cells undergoing apoptosis (Thomadaki, *et al* 2008).

Centrosomal proteins. A third group of functionally related partner proteins has also been observed by Delaval et. al., (2005). This group initially observed that FGFR1OP is a centrosomal protein and that the FGFR1OP-FGFR1 fusion is also centrosomal, Furthermore, FGFR1OP1-FGFR1 was shown to be active at the centrosome and activate multiple signalling pathways (Delaval, *et al* 2005a). Several other centrosomal partner proteins have been described, e.g. CEP110 fuses to FGFR1, PCM1 fuses to JAK2, CDK5RAP2 fuses to PDGFRα, NDE1 and NIN fuse to PDGFRβ.

The centrosome consists of a pair of centrioles surrounded by pericentriolar material and plays an integral role in directing the organisation of the cytoplasmic microtubules and the assembly of the mitotic spindles (Rieder, *et al* 2001). Aberrations in the number of centrosomes lead to mitotic defects and cause chromosome mis-segregation and centrosome abnormalities and are therefore frequent in many common cancers (Nigg 2002). Delaval et. al,. hypothesise that the centrosome, which is linked to the microtubules, is close to the nucleus, and is connected to the Golgi apparatus and the proteasome, could serve to integrate multiple signalling pathways controlling cell division, cell migration and cell fate. Abnormal kinase activity at the centrosome may be an efficient way to pervert cell division in malignancy. In support of this idea, it has been reported that an unidentified centrosomal mechanism controls the number of neurons generated by neural precursor cells (Bond, *et al* 2005). It is conceivable that similar mechanisms might operate in haematopoietic development.

These three functional clusters of partner proteins are striking and, as discussed above, may directly impact on cellular processes and promote transformation. However since the clusters impact on such functionally diverse processes, it is perhaps unlikely that any of the postulated interactions are critical for transformation, although they might produce cellular changes that modify the phenotype. In support of this, it has been observed that the clinical features in patients with FGFR1 rearrangements are slightly different depending on the identity of the partner protein (Macdonald, *et al* 2002b). However other interpretations are also possible, for example it might be that endocytic, pre-mRNA processing and centrosomal proteins are simply large families, widely expressed and tend to contain dimerisation domains. It is also possible that altered subcellular localisation may be important, by enabling the kinase to interact with novel substrates. For example, the *FIG-ROS* fusion found in glioblastoma has been shown to be directed by the partner protein to the Golgi (Charest, *et al* 2003), a localisation that is thought to be essential to its transforming ability. However whether this is due to interactions with novel substrates or whether specific localisation increases the effective concentration of the fusion protein above a critical threshold is unclear.

6.2 Routine detection of tyrosine kinase fusions

Detection of tyrosine kinase fusions is important for several reasons. First, the finding of a fusion such as *FIP1L1-PDGFRA* in a cytogenetically normal patient confirms the diagnosis of a clonal neoplastic disorder. Second, specific fusions are indicators for specific targeted therapies, e.g. patients with *PDGFRA* or *PDGFRB* translocations respond very well to imatinib. Third, the involvement of particular tyrosine kinases may be prognostically important, e.g. patients with *FGFR1* translocations have a strong propensity to progress to AML and should be considered as candidates for allogeneic bone marrow transplantation. It is clearly important therefore that diagnostic procedures are available to detect fusion tyrosine kinases accurately. The problem arises from the sheer number of different fusions that have been described,

which makes it impractical and very expensive to screen for each by RT-PCR. Furthermore it is likely that additional fusions with novel partners remain to be discovered. For many cases it would be possible to use FISH with split-apart probes however rare translocations often mask complex rearrangements at the molecular level and consequently split-apart FISH may fail to detect a rearrangement (Kulkarni*, et al* 2000, Reiter*, et al* 1998).

To help address this problem, and also to try and screen for new fusions, I developed two generic RT-PCR assays to detect over-expression of *PDGFR* genes. The assays were designed to look for over-expression of the tyrosine kinase domain relative to an extracellular control band and which might therefore indicate the presence of a *PDGFR* fusion gene. Following validation using samples with known *PDGFR* fusions, I screened approximately 200 patients for both *PDGFRA* and *PDGFRB* over-expression. Of the six patients found to over-express *PDGFRB*, one had a previously undetected variant *ETV6-PDGFRB* fusion that had arisen as a result of a complex translocation. Unexpectedly, 2 had *FGFR1* rearrangements. The reason for this is unclear and requires further investigation, initially to see if this is a consistent finding in cases with *FGFR1* fusions. Of the remaining 3 patients, only 2 had material available for further study and no abnormalities were found. Over-expression of *PDGFRA* was identified in 14 patients. Of the 7 that had sufficient material to investigate further with bubble PCR one had the novel *KIF5B-PDGFRA* fusion, one had a previously undetected *FIP1LI-PDGFRA* fusions and five showed no abnormalities.

Overall then, the generic RT-PCR assays led to the identification of fusions that had not been detected previously and is therefore clearly a useful diagnostic tool. However it does produce quite a large number of false positives that require extensive further investigation. It would be desirable to have a technique that reduced these, and was also more readily extendable to screen for fusions involving additional tyrosine kinases. I therefore designed a cDNA MLPA assay to look at expression of the four kinases most commonly involved in MPDs: *ABL*, *FGFR1*, *PDGFRA* and

PDGFRB. MLPA is a probe-based technique that was originally designed to detect deletions and amplifications of multiple target genomic DNA sequences in a single reaction. The cDNA MLPA assay was designed on the same principle as that of the generic RT-PCR assays i.e. to look at differential expression of the tyrosine kinase portion of the gene (which would be retained from any fusion) compared to a second probe that hybridised to a portion of the extracellular domain or 5' region that would not be retained in the fusion.

The assay was validated using normal control cDNA samples, cell lines and patient cDNA with known tyrosine kinase gene fusions. The majority of positive controls were correctly identified, however the intracellular:extracellular ratio of the other unaffected probes was rarely 1.0 as predicted. Furthermore due to the differing levels of expression of each gene it proved difficult to define a single set of conditions that effectively covered all four genes and, in addition, the assay was very dependent on the quality of the cDNA and probe sensitivity. Nevertheless, this technique did lead to the discovery of a novel variant *ETV6-PDGFRB* fusion in one case who over-expressed the *PDGFRB* kinase domain. Overall, the cDNA MLPA seemed to be more useful as a general screen allowing the analysis of 3 tyrosine kinase genes in one test limiting sample volume which is always precious, however it was time consuming and less reproducible. The generic RT-PCRs worked relatively well for the two targets that are clinically most important, *PDGFRA* and *PDGFRB*, however over time they also proved temperamental with sensitivity to cDNA quality, primer age and the thermocycler that was used. The yield of previously unknown positive cases from either technique was low and perhaps the most effective approach would be to use the generic RT-PCR assays to screen for *PDGFRA* over-expression associated with *FIP1L1-PDGFRA* in cases with eosinophilia, and to screen for *PDGFRA* and *PDGFRB* abnormalities in cases with a complex karyotype. This strategy is unlikely to miss any true positives but will minimise the number of false positives. Other generic strategies to detect gene fusions are also possible, e.g. exon arrays or real time PCR assays to look at expression across target genes, but these would be much more expensive.

6.3 Loss of heterozygosity analysis as an approach to identify genes harbouring activating mutations

The main focus of my work was to identify new abnormalities in atypical MPDs. LOH is a genetic mechanism that creates a homozygous or hemizygous cell by loss of the wild-type allele as a result of various mutational events. These include deletion, mitotic recombination, localised gene conversion, point mutation, or non-dysjunction. My objective was to identify new oncogenes or tumour suppressor genes that are marked by these regions of LOH, and specifically regions of acquired uniparental disomy (aUPD), or copy number neutral LOH, formed as a consequence of mitotic recombination. Uniparental disomy (UPD) refers to a mechanism in which both copies of a chromosome pair have originated from one parent. This can result in diverse clinical conditions as a consequence of homozygosity for recessive mutations or aberrant patterns of imprinting. Somatic UPD has also been shown to arise as a consequence of recombination occurring during mitotic cell divisions. This acquired UPD (aUPD) has been shown be associated with pathogenetic driver mutations in genes within the region of homozygosity such as *JAK2* V617F, *FLT3* ITD, *CEBPA* and *WT1* mutations (Fitzgibbon*, et al* 2005, Griffiths*, et al* 2005, Kralovics*, et al* 2005b, Raghavan*, et al* 2005). I aimed to search for new regions of aUPD as a means to identify regions harbouring genes that are mutated in aMPDs and transformed CML.

Numerous techniques are capable of identifying regions of LOH however SNP arrays have the advantage as they can simultaneously detect copy number and zygosity. For my study, I utilised 50K SNP arrays which are entirely adequate for detecting large regions of aUPD although the resolution of copy number changes turned out to be very limited. A major drawback of my study was that paired sample analysis was not possible due to the lack of constitutional DNA, thus I was unable to determine if tracts of homozygosity were acquired or whether they were inherited. Fortunately though I found that 40% of aMPD cases had large (>20Mb) tracts of copy neutral

homozygosity that are very unlikely to have been inherited. However, aUPD was uncommon in the other patient subgroups (CML-BC, MF and V617F negative PV).

In aMPDs, 8 distinct regions of aUPD were identified indicating considerable genetic heterogeneity in these disorders. However two novel, recurrent regions were seen at 7q and 11q, each in 10% of cases. Initially, because of the strong association between MPDs and tyrosine kinases, I fully re-sequenced the three tyrosine kinase genes on 7q, (*EPHA1*, *EPHB6* and *MET*) but no mutations were found. There are no tyrosine kinase genes on 11q so it was apparent that both abnormalities must be targeting other genes. I then went on to select candidates at both locations for further analysis. Both regions were relatively large and contained several hundred genes, including a large number of candidates that could potentially be involved in malignancy. I focused mainly on those that were known to encode signal transduction components and were either known oncogenes (*BRAF*, *CBL*) or closely related to known oncogenes (*PI3KCG*). I also checked for the *MLL* ITD as this gene is widely involved in leukaemia. This analysis led to the finding of *CBL* mutations in all three cases with 11q aUPD.

The *CBL* gene is located at 11q23.3 and encodes a widely expressed ubiquitin ligase that functions as a negative regulator of several receptor protein tyrosine kinases by inducing ubiquitination of the receptors, thereby promoting their sorting for endosomal degradation or recycling. In addition, CBL plays a positive role in signalling by serving as a scaffold for the recruitment of downstream substrates. The mutations I found predicted missense substitutions in the RING/linker domain, L380P in two cases and P417A in the third. This region is responsible for the ubiquitin ligase activity and it is likely that these mutations, like the R420Q substitution previously described in AML, have transforming activity and are therefore causative driver mutations (Sargin, *et al* 2007).

Having found *CBL* mutations in the three 11q aUPD cases, I next screened 547 patients, including those with PV, IMF, ET, SM, HES, CMML, aCML and other

aMPDs. A further 23 patients with *CBL* mutations were identified, the majority of which were missense changes, plus one nonsense mutation and some intronic changes, two of which were shown to result in skipping of exon 8. Although I found the R420Q in 4 cases, the majority of mutations were novel. The mutations were associated predominantly with a diagnosis of aCML, CMML and MF i.e. highly related and relatively aggressive diseases. In three cases for which suitable material was available, the mutations were shown to be acquired. In two cases, the mutations were acquired during progression of a pre-existing MPD indicating that in these cases at least, *CBL* mutations were secondary events.

There are a number of obvious future avenues of investigation that should be undertaken following the results of my studies, e.g.

- Functional analysis to prove that the *CBL* mutations I found are indeed transforming. This has been achieved for some of the mutants by testing if they transform 32D cells over-expressing *FLT3*, as has been described for R420Q (Grand 2009, Sargin*, et al* 2007).
- Investigation of genetic events that might co-operate with *CBL* mutations. One possibility is over-expression of one or more tyrosine kinases, and potentially this could be investigated by gene expression profiling. Another is indicated by the fact that one of my cases has both a *CBL* mutation and 7q aUPD suggesting that they might co-operate. Once the 7q aUPD mutation is identified this can be formally tested, probably most appropriately in mouse models.
- If *CBL* mutants function by increasing tyrosine kinase signalling then it is likely that the malignant clone might be susceptible to inhibition by tyrosine kinase inhibitors. This could be investigated by testing the effects of inhibitors against patient and control cells.
- The clinical phenotype associated with *CBL* mutations needs to be more closely defined. Ultimately it should be possible to see if any clinical features are correlated with specific mutations and also since aUPD is effectively a

second hit, it is possible that homozygous mutations are associated with more aggressive disease.

- Identification of the gene underlying 7q aUPD. Other members of the group have eliminated other candidate genes and are planning to sequence the coding sequence (exons) of all genes in the region using custom designed pull down arrays and Illumina sequencing.
- Investigation of smaller regions of aUPD that may have been missed by my analysis. Efficient algorithms have been described to pick up low level aUPD that is missed by standard approaches (Yamamoto*, et al* 2001) and this, in combination with the current state-of-the-art SNP 6.0 arrays, may prove to be a powerful approach to identify novel regions harbouring mutated genes.

6.4 Current knowledge of the pathogenesis of atypical MPDs

Prior to my findings it was estimated that at least part of the genetic basis of approximately a third of aMPDs was understood, with roughly 2% having tyrosine kinase fusions, 5% having *FLT3* ITD or *KIT* D816V, 12.5% V617F *JAK2* and 12.5% activating *NRAS* mutations. (Cross 2006, Jones*, et al* 2005). Clearly a lot of detail remains to be filled in, especially the targets of 7q and other regions of aUPD plus how different mutations may co-operate together, but following my work, I estimate that we now know something about two thirds of cases (Figure 6.4) with *CBL* mutations, 7q aUPD and other regions of aUPD each seen in roughly 10% of aMPDs. Further works needs to be performed to determine which combinations of mutations occur in individual patients, how they co-operate and how they might be targeted therapeutically.

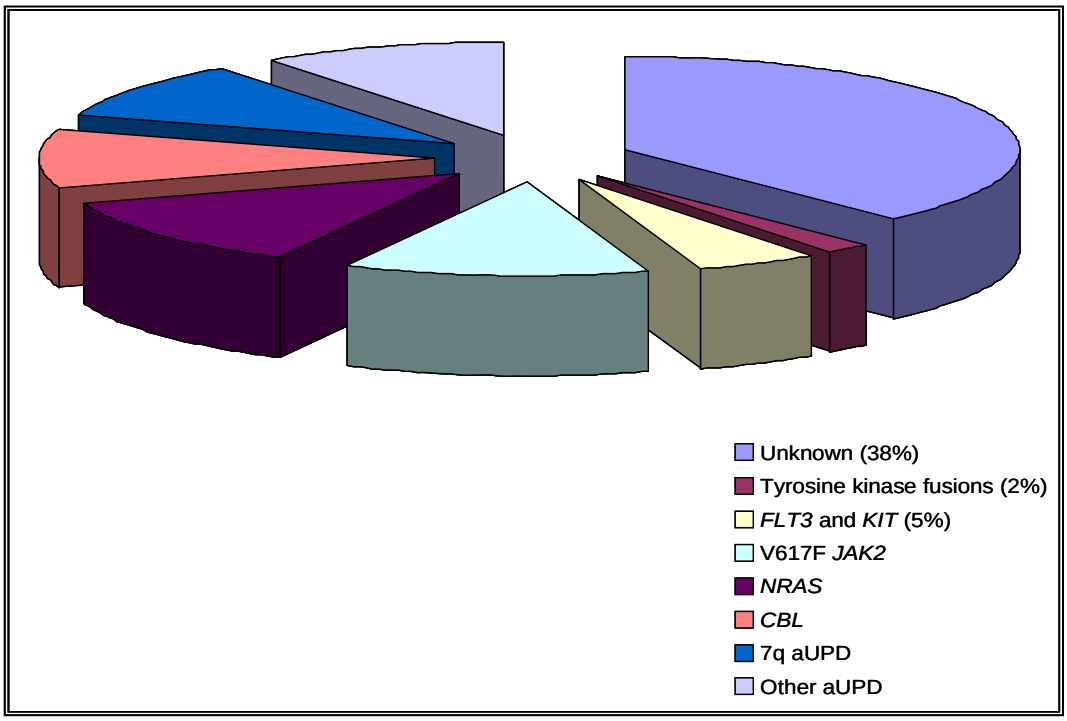


Figure 6. 4: Pie chart illustrating the proportion of aMPD cases related to known genetic variants, such as V617F and CBL mutations aUPD compared to unknown pathogenicity.

List of References

Abe, A., Emi, N., Tanimoto, M., Terasaki, H., Marunouchi, T. & Saito, H. (1997) Fusion of the platelet-derived growth factor receptor beta to a novel gene CEV14 in acute myelogenous leukemia after clonal evolution. *Blood,* **90,** 4271-4277.

Adamson, J.W., Fialkow, P.J., Murphy, S., Prchal, J.F. & Steinmann, L. (1976) Polycythemia vera: stem-cell and probable clonal origin of the disease. *N.Engl.J.Med,* **295,** 913-916.

Adjei, A.A. (2001) Blocking oncogenic Ras signaling for cancer therapy. *J Natl Cancer Inst,* **93,** 1062-1074.

Aktas, D., Yenicesu, I., Hicsonmez, G. & Tuncbilek, E. (1999) Loss of maternal allele in a child with myelodysplastic syndrome and monosomy 7. *Am J Hematol,* **62,** 49-51.

Andoniou, C.E., Thien, C.B. & Langdon, W.Y. (1994) Tumour induction by activated abl involves tyrosine phosphorylation of the product of the cbl oncogene. *EMBO J,* **13,** 4515-4523.

Appels, N.M., Beijnen, J.H. & Schellens, J.H. (2005) Development of farnesyl transferase inhibitors: a review. *Oncologist,* **10,** 565-578.

Apperley, J.F., Gardembas, M., Melo, J.V., Russell-Jones, R., Bain, B.J., Baxter, E.J., Chase, A., Chessells, J.M., Colombat, M., Dearden, C.E., Dimitrijevic, S., Mahon, F.X., Marin, D., Nikolova, Z., Olavarria, E., Silberman, S., Schultheis, B., Cross, N.C. & Goldman, J.M. (2002) Response to imatinib mesylate in patients with chronic myeloproliferative diseases with rearrangements of the platelet-derived growth factor receptor beta. *N.Engl.J.Med,* **347,** 481-487.

Armstrong, E., Kastury, K., Aprelikova, O., Bullrich, F., Nezelof, C., Gogusev, J., Wasmuth, J.J., Alitalo, K., Morris, S. & Huebner, K. (1993) FLT4 receptor tyrosine kinase gene mapping to chromosome band 5q35 in relation to the t(2;5), t(5;6), and t(3;5) translocations. *Genes Chromosomes Cancer,* **7,** 144-151.

Bache, K.G., Slagsvold, T. & Stenmark, H. (2004) Defective downregulation of receptor tyrosine kinases in cancer. *EMBO J,* **23,** 2707-2712.

Bachman, K.E., Argani, P., Samuels, Y., Silliman, N., Ptak, J., Szabo, S., Konishi, H., Karakas, B., Blair, B.G., Lin, C., Peters, B.A., Velculescu, V.E. & Park, B.H. (2004) The PIK3CA gene is mutated with high frequency in human breast cancers. *Cancer Biol Ther,* **3,** 772-775.

Baillat, G., Moqrich, A., Castets, F., Baude, A., Bailly, Y., Benmerah, A. & Monneron, A. (2001) Molecular cloning and characterization of phocein, a protein found from the Golgi complex to dendritic spines. *Mol Biol Cell,* **12,** 663-673.

Bain, B.J. (1996) Eosinophilic leukaemias and the idiopathic hypereosinophilic syndrome. *Br.J.Haematol.,* **95,** 2-9.

Barnes, D.J. & Melo, J.V. (2002) Cytogenetic and molecular genetic aspects of chronic myeloid leukaemia. *Acta Haematol.,* **108,** 180-202.

Barr, F.A. (1999) A novel Rab6-interacting domain defines a family of Golgi-targeted coiled-coil proteins. *Curr Biol,* **9,** 381-384.

Barthe, C., Cony-Makhoul, P., Melo, J.V. & Mahon, J.R. (2001) Roots of clinical resistance to STI-571 cancer therapy. *Science,* **293,** 2163.

Bartoli, M., Monneron, A. & Ladant, D. (1998) Interaction of calmodulin with striatin, a WD-repeat protein present in neuronal dendritic spines. *J Biol Chem,* **273,** 22248-22253.

Bartoli, M., Ternaux, J.P., Forni, C., Portalier, P., Salin, P., Amalric, M. & Monneron, A. (1999) Down-regulation of striatin, a neuronal calmodulin-binding protein, impairs rat locomotor activity. *J Neurobiol,* **40,** 234-243.

Bascom, R.A., Srinivasan, S. & Nussbaum, R.L. (1999) Identification and characterization of golgin-84, a novel Golgi integral membrane protein with a cytoplasmic coiled-coil domain. *J.Biol.Chem.,* **274,** 2953-2962.

Basirico, R., Pirrotta, R., Fabbiano, F., Mirto, S., Cascio, L., Pagano, M., Cammarata, G., Magrin, S. & Santoro, A. (2003) Submicroscopic deletions at 7q region are associated with recurrent chromosome abnormalities in acute leukemia. *Haematologica*, **88,** 429-437.

Baxter, E.J., Hochhaus, A., Bolufer, P., Reiter, A., Fernandez, J.M., Senent, L., Cervera, J., Moscardo, F., Sanz, M.A. & Cross, N.C. (2002) The t(4;22)(q12;q11) in atypical chronic myeloid leukaemia fuses BCR to PDGFRA. *Hum.Mol.Genet.,* **11,** 1391-1397.

Baxter, E.J., Kulkarni, S., Vizmanos, J.L., Jaju, R., Martinelli, G., Testoni, N., Hughes, G., Salamanchuk, Z., Calasanz, M.J., Lahortiga, I., Pocock, C.F., Dang, R., Fidler, C., Wainscoat, J.S., Boultwood, J. & Cross, N.C. (2003) Novel translocations that disrupt the platelet-derived growth factor receptor beta (PDGFRB) gene in BCR-ABL-negative chronic myeloproliferative disorders. *Br.J.Haematol.,* **120,** 251-256.

Baxter, E.J., Scott, L.M., Campbell, P.J., East, C., Fourouclas, N., Swanton, S., Vassiliou, G.S., Bench, A.J., Boyd, E.M., Curtin, N., Scott, M.A., Erber, W.N. & Green, A.R. (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet,* **365,** 1054-1061.

Bellanne-Chantelot, C., Chaumarel, I., Labopin, M., Bellanger, F., Barbu, V., De Toma, C., Delhommeau, F., Casadevall, N., Vainchenker, W., Thomas, G. &

Najman, A. (2006) Genetic and clinical implications of the Val617Phe JAK2 mutation in 72 families with myeloproliferative disorders. *Blood,* **108,** 346-352.

Belloni, E., Trubia, M., Gasparini, P., Micucci, C., Tapinassi, C., Confalonieri, S., Nuciforo, P., Martino, B., Lo-Coco, F., Pelicci, P.G. & Di Fiore, P.P. (2005) 8p11 myeloproliferative syndrome with a novel t(7;8) translocation leading to fusion of the FGFR1 and TIF1 genes. *Genes Chromosomes.Cancer,* **42,** 320-325.

Bennett, J.H. (1845) Case of hypertrophy of the spleen and liver in which death took place from suppuration of the blood. *Edinburgh Medical and Surgical Journal,* **64,** 413-423.

Bergsten, E., Uutela, M., Li, X., Pietras, K., Ostman, A., Heldin, C.H., Alitalo, K. & Eriksson, U. (2001) PDGF-D is a specific, protease-activated ligand for the PDGF beta-receptor. *Nat.Cell Biol.,* **3,** 512-516.

Bernard, O.A., Mauchauffe, M., Mecucci, C., Van den, B.H. & Berger, R. (1994) A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4, AF-9 nor ENL. *Oncogene,* **9,** 1039-1045.

Bethwaite, P.B., Koreth, J., Herrington, C.S. & McGee, J.O. (1995) Loss of heterozygosity occurs at the D11S29 locus on chromosome 11q23 in invasive cervical carcinoma. *Br J Cancer,* **71,** 814-818.

Blume-Jensen, P. & Hunter, T. (2001) Oncogenic kinase signalling. *Nature,* **411,** 355-365.

Bohlander, S.K. (2005) ETV6: a versatile player in leukemogenesis. *Semin Cancer Biol,* **15,** 162-174.

Bond, J., Roberts, E., Springell, K., Lizarraga, S., Scott, S., Higgins, J., Hampshire, D.J., Morrison, E.E., Leal, G.F., Silva, E.O., Costa, S.M., Baralle, D., Raponi, M.,

Karbani, G., Rashid, Y., Jafri, H., Bennett, C., Corry, P., Walsh, C.A. & Woods, C.G. (2005) A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. *Nat.Genet.,* **37,** 353-355.

Boultwood, J. (2001) Ataxia telangiectasia gene mutations in leukaemia and lymphoma. *J Clin Pathol,* **54,** 512-516.

Bowen, D.T., Frew, M.E., Hills, R., Gale, R.E., Wheatley, K., Groves, M.J., Langabeer, S.E., Kottaridis, P.D., Moorman, A.V., Burnett, A.K. & Linch, D.C. (2005) RAS mutation in acute myeloid leukemia is associated with distinct cytogenetic subgroups but does not influence outcome in patients younger than 60 years. *Blood,* **106,** 2113-2119.

Bozzi, F., Tamborini, E., Negri, T., Pastore, E., Ferrari, A., Luksch, R., Casanova, M., Pierotti, M.A., Bellani, F.F. & Pilotti, S. (2007) Evidence for activation of KIT, PDGFRalpha, and PDGFRbeta receptors in the Ewing sarcoma family of tumors. *Cancer,* **109,** 1638-1645.

Branford, S., Rudzki, Z., Walsh, S., Grigg, A., Arthur, C., Taylor, K., Herrmann, R., Lynch, K.P. & Hughes, T.P. (2002) High frequency of point mutations clustered within the adenosine triphosphate-binding region of BCR/ABL in patients with chronic myeloid leukemia or Ph-positive acute lymphoblastic leukemia who develop imatinib (STI571) resistance. *Blood,* **99,** 3472-3475.

Broderick, D.K., Di, C., Parrett, T.J., Samuels, Y.R., Cummins, J.M., McLendon, R.E., Fults, D.W., Velculescu, V.E., Bigner, D.D. & Yan, H. (2004) Mutations of PIK3CA in anaplastic oligodendrogliomas, high-grade astrocytomas, and medulloblastomas. *Cancer Res,* **64,** 5048-5050.

Brose, M.S., Volpe, P., Feldman, M., Kumar, M., Rishi, I., Gerrero, R., Einhorn, E., Herlyn, M., Minna, J., Nicholson, A., Roth, J.A., Albelda, S.M., Davies, H., Cox, C., Brignell, G., Stephens, P., Futreal, P.A., Wooster, R., Stratton, M.R. & Weber, B.L.

(2002) BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res,* **62,** 6997-7000.

Brown, D.L., Heimann, K., Lock, J., Kjer-Nielsen, L., van, V.C., Stow, J.L. & Gleeson, P.A. (2001) The GRIP domain is a specific targeting sequence for a population of trans-Golgi network derived tubulo-vesicular carriers. *Traffic.,* **2,** 336-344.

Butcher, C.M., Hahn, U., To, L.B., Gecz, J., Wilkins, E.J., Scott, H.S., Bardy, P.G. & D'Andrea, R.J. (2008) Two novel JAK2 exon 12 mutations in JAK2V617F-negative polycythaemia vera patients. *Leukemia,* **22,** 870-873.

Caceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L. & Krainer, A.R. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol,* **138,** 225-238.

Cain, J.A., Xiang, Z., O'Neal, J., Kreisel, F., Colson, A., Luo, H., Hennighausen, L. & Tomasson, M.H. (2007) Myeloproliferative disease induced by TEL-PDGFRB displays dynamic range sensitivity to Stat5 gene dosage. *Blood,* **109,** 3906-3914.

Calabretta, B. & Perrotti, D. (2004) The biology of CML blast crisis. *Blood,* **103,** 4010-4022.

Caligiuri, M.A., Briesewitz, R., Yu, J., Wang, L., Wei, M., Arnoczky, K.J., Marburger, T.B., Wen, J., Perrotti, D., Bloomfield, C.D. & Whitman, S.P. (2007) Novel c-CBL and CBL-b ubiquitin ligase mutations in human acute myeloid leukemia. *Blood,* **110,** 1022-1024.

Caligiuri, M.A., Schichman, S.A., Strout, M.P., Mrozek, K., Baer, M.R., Frankel, S.R., Barcos, M., Herzig, G.P., Croce, C.M. & Bloomfield, C.D. (1994) Molecular rearrangement of the ALL-1 gene in acute myeloid leukemia without cytogenetic evidence of 11q23 chromosomal translocations. *Cancer Res,* **54,** 370-373.

Campbell, I.G., Russell, S.E., Choong, D.Y., Montgomery, K.G., Ciavarella, M.L., Hooi, C.S., Cristiano, B.E., Pearson, R.B. & Phillips, W.A. (2004) Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res,* **64,** 7678-7681.

Campbell, P.J., Baxter, E.J., Beer, P.A., Scott, L.M., Bench, A.J., Huntly, B.J., Erber, W.N., Kusec, R., Larsen, T.S., Giraudier, S., Le Bousse-Kerdiles, M.C., Griesshammer, M., Reilly, J.T., Cheung, B.Y., Harrison, C.N. & Green, A.R. (2006) Mutation of JAK2 in the myeloproliferative disorders: timing, clonality studies, cytogenetic associations, and role in leukemic transformation. *Blood,* **108,** 3548-3555.

Campbell, P.J., Scott, L.M., Buck, G., Wheatley, K., East, C.L., Marsden, J.T., Duffy, A., Boyd, E.M., Bench, A.J., Scott, M.A., Vassiliou, G.S., Milligan, D.W., Smith, S.R., Erber, W.N., Bareford, D., Wilkins, B.S., Reilly, J.T., Harrison, C.N. & Green, A.R. (2005) Definition of subtypes of essential thrombocythaemia and relation to polycythaemia vera based on JAK2 V617F mutation status: a prospective study. *Lancet,* **366,** 1945-1953.

Cardoso, C.M., Groth-Pedersen, L., Hoyer-Hansen, M., Kirkegaard, T., Corcelle, E., Andersen, J.S., Jaattela, M. & Nylandsted, J. (2009) Depletion of kinesin 5B affects lysosomal distribution and stability and induces peri-nuclear accumulation of autophagosomes in cancer cells. *PLoS ONE,* **4,** e4424.

Carr, J., Bown, N.P., Case, M.C., Hall, A.G., Lunec, J. & Tweddle, D.A. (2007) High-resolution analysis of allelic imbalance in neuroblastoma cell lines by single nucleotide polymorphism arrays. *Cancer Genet Cytogenet,* **172,** 127-138.

Carroll, M., Tomasson, M.H., Barker, G.F., Golub, T.R. & Gilliland, D.G. (1996) The TEL/platelet-derived growth factor beta receptor (PDGF beta R) fusion in chronic myelomonocytic leukemia is a transforming protein that self-associates and activates PDGF beta R kinase-dependent signaling pathways. *Proc.Natl.Acad.Sci.U.S.A,* **93,** 14845-14850.

Carter, S.L., Negrini, M., Baffa, R., Gillum, D.R., Rosenberg, A.L., Schwartz, G.F. & Croce, C.M. (1994) Loss of heterozygosity at 11q22-q23 in breast cancer. *Cancer Res,* **54,** 6270-6274.

Castets, F., Bartoli, M., Barnier, J.V., Baillat, G., Salin, P., Moqrich, A., Bourgeois, J.P., Denizot, F., Rougon, G., Calothy, G. & Monneron, A. (1996) A novel calmodulin-binding protein, belonging to the WD-repeat family, is localized in dendrites of a subset of CNS neurons. *J Cell Biol,* **134,** 1051-1062.

Cave, H., Cacheux, V., Raynaud, S., Brunie, G., Bakkus, M., Cochaux, P., Preudhomme, C., Lai, J.L., Vilmer, E. & Grandchamp, B. (1997) ETV6 is the target of chromosome 12p deletions in t(12;21) childhood acute lymphocytic leukemia. *Leukemia,* **11,** 1459-1464.

Cazzaniga, G., Tosi, S., Aloisi, A., Giudici, G., Daniotti, M., Pioltelli, P., Kearney, L. & Biondi, A. (1999) The tyrosine kinase abl-related gene ARG is fused to ETV6 in an AML-M4Eo patient with a t(1;12)(q25;p13): molecular cloning of both reciprocal transcripts. *Blood,* **94,** 4370-4373.

Chan, E.Y. (2005) Advances in sequencing technology. *Mutat Res,* **573,** 13-40.

Chang, W.J., Ying, Y.S., Rothberg, K.G., Hooper, N.M., Turner, A.J., Gambliel, H.A., De Gunzburg, J., Mumby, S.M., Gilman, A.G. & Anderson, R.G. (1994) Purification and characterization of smooth muscle cell caveolae. *J Cell Biol,* **126,** 127-138.

Charest, A., Lane, K., McMahon, K., Park, J., Preisinger, E., Conroy, H. & Housman, D. (2003) Fusion of FIG to the receptor tyrosine kinase ROS in a glioblastoma with an interstitial del(6)(q21q21). *Genes Chromosomes.Cancer,* **37,** 58-71.

Chase, A., Grand, F.H. & Cross, N.C. (2007) Activity of TKI258 against primary cells and cell lines with FGFR1 fusion genes associated with the 8p11 myeloproliferative syndrome. *Blood,* **110,** 3729-3734.

Chase, A., Huntly, B.J. & Cross, N.C. (2001) Cytogenetics of chronic myeloid leukaemia. *Best.Pract.Res.Clin.Haematol.,* **14,** 553-571.

Chen, J., Lee, B.H., Williams, I.R., Kutok, J.L., Mitsiades, C.S., Duclos, N., Cohen, S., Adelsperger, J., Okabe, R., Coburn, A., Moore, S., Huntly, B.J., Fabbro, D., Anderson, K.C., Griffin, J.D. & Gilliland, D.G. (2005) FGFR3 as a therapeutic target of the small molecule inhibitor PKC412 in hematopoietic malignancies. *Oncogene,* **24,** 8259-8267.

Chiara, F., Goumans, M.J., Forsberg, H., Ahgren, A., Rasola, A., Aspenstrom, P., Wernstedt, C., Hellberg, C., Heldin, C.H. & Heuchel, R. (2004) A gain of function mutation in the activation loop of platelet-derived growth factor beta-receptor deregulates its kinase activity. *J.Biol.Chem.,* **279,** 42516-42527.

Chow, L.Q. & Eckhardt, S.G. (2007) Sunitinib: from rational design to clinical efficacy. *J Clin Oncol,* **25,** 884-896.

Chung, K.F., Hew, M., Score, J., Jones, A.V., Reiter, A., Cross, N.C. & Bain, B.J. (2006) Cough and hypereosinophilia due to FIP1L1-PDGFRA fusion gene with tyrosine kinase activity. *Eur Respir J,* **27,** 230-232.

Cools, J., DeAngelo, D.J., Gotlib, J., Stover, E.H., Legare, R.D., Cortes, J., Kutok, J., Clark, J., Galinsky, I., Griffin, J.D., Cross, N.C.P., Tefferi, A., Malone, J., Alam, R., Schrier, S.L., Schmid, J., Rose, M., Vandenberghe, P., Verhoef, G., Boogaerts, M., Wlodarska, I., Kantarjian, H., Marynen, P., Coutre, S.E., Stone, R. & Gilliland, D.G. (2003a) A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *New England Journal of Medicine,* **348,** 1201-1214.

Cools, J., Mentens, N., Odero, M.D., Peeters, P., Wlodarska, I., Delforge, M., Hagemeijer, A. & Marynen, P. (2002) Evidence for position effects as a variant ETV6-mediated leukemogenic mechanism in myeloid leukemias with a t(4;12)(q11-q12;p13) or t(5;12)(q31;p13). *Blood*, **99,** 1776-1784.

Cools, J., Stover, E.H., Boulton, C.L., Gotlib, J., Legare, R.D., Amaral, S.M., Curley, D.P., Duclos, N., Rowan, R., Kutok, J.L., Lee, B.H., Williams, I.R., Coutre, S.E., Stone, R.M., DeAngelo, D.J., Marynen, P., Manley, P.W., Meyer, T., Fabbro, D., Neuberg, D., Weisberg, E., Griffin, J.D. & Gilliland, D.G. (2003b) PKC412 overcomes resistance to imatinib in a murine model of FIP1L1-PDGFR alpha-induced myeloproliferative disease. *Cancer Cell*, **3,** 459-469.

Cragie (1845) *Edinburgh Medical and Surgical Journal*, **64,** 400-412.

Crescenzi, B., Chase, A., Starza, R.L., Beacci, D., Rosti, V., Galli, A., Specchia, G., Martelli, M.F., Vandenberghe, P., Cools, J., Jones, A.V., Cross, N.C., Marynen, P. & Mecucci, C. (2007) FIP1L1-PDGFRA in chronic eosinophilic leukemia and BCR-ABL1 in chronic myeloid leukemia affect different leukemic cells. *Leukemia,* **21,** 397-402.

Cross, N.C.P. & Reiter, A. (2002) Tyrosine kinase fusion genes in chronic myeloproliferative diseases. *Leukemia*, **16,** 1207-1212.

Cross, N.R., A (2006) *'BCR-ABL negative chronic myeloid leukaemia' in Hematologic Malignancies: Myeloproliferative disorders.* Springer.

Cucuianu, A. (2001) Cell darwinism, apoptosis, free radicals and haematological malignancies. *Med Hypotheses*, **56,** 52-57.

Curtis, C.E., Grand, F.H., Musto, P., Clark, A., Murphy, J., Perla, G., Minervini, M.M., Stewart, J., Reiter, A. & Cross, N.C. (2007) Two novel imatinib-responsive

PDGFRA fusion genes in chronic eosinophilic leukaemia. *Br J Haematol,* **138,** 77-81.

Daley, G.Q., Van Etten, R.A. & Baltimore, D. (1990) Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science,* **247,** 824-830.

Dascalescu, C.M., Peoc'h, M., Callanan, M., Jacob, M.C., Sotto, M.F., Gressin, R., Sotto, J.J. & Leroux, D. (1999) Deletion 7q in B-cell low-grade lymphoid malignancies: a cytogenetic/fluorescence in situ hybridization and immunopathologic study. *Cancer Genet Cytogenet,* **109,** 21-28.

David, M., Cross, N.C., Burgstaller, S., Chase, A., Curtis, C., Dang, R., Gardembas, M., Goldman, J.M., Grand, F., Hughes, G., Huguet, F., Lavender, L., McArthur, G.A., Mahon, F.X., Massimini, G., Melo, J., Rousselot, P., Russell-Jones, R.J., Seymour, J.F., Smith, G., Stark, A., Waghorn, K., Nikolova, Z. & Apperley, J.F. (2007) Durable responses to imatinib in patients with PDGFRB fusion gene-positive and BCR-ABL-negative chronic myeloproliferative disorders. *Blood,* **109,** 61-64.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B.A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G.J., Bigner, D.D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J.W., Leung, S.Y., Yuen, S.T., Weber, B.L., Seigler, H.F., Darrow, T.L., Paterson, H., Marais, R., Marshall, C.J., Wooster, R., Stratton, M.R. & Futreal, P.A. (2002) Mutations of the BRAF gene in human cancer. *Nature,* **417,** 949-954.

de Klein, A., van Kessel, A.G., Grosveld, G., Bartram, C.R., Hagemeijer, A., Bootsma, D., Spurr, N.K., Heisterkamp, N., Groffen, J. & Stephenson, J.R. (1982) A

cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature,* **300,** 765-767.

de Vries, A.C., Stam, R.W., Kratz, C.P., Zenker, M., Niemeyer, C.M. & van den Heuvel-Eibrink, M.M. (2007) Mutation analysis of the BRAF oncogene in juvenile myelomonocytic leukemia. *Haematologica,* **92,** 1574-1575.

Delaval, B., Lelievre, H. & Birnbaum, D. (2005a) Myeloproliferative disorders: the centrosome connection. *Leukemia,* **19,** 1739-1744.

Delaval, B., Letard, S., Lelievre, H., Chevrier, V., Daviet, L., Dubreuil, P. & Birnbaum, D. (2005b) Oncogenic tyrosine kinase of malignant hemopathy targets the centrosome. *Cancer Res.,* **65,** 7231-7240.

Demiroglu, A., Steer, E.J., Heath, C., Taylor, K., Bentley, M., Allen, S.L., Koduru, P., Brody, J.P., Hawson, G., Rodwell, R., Doody, M.L., Carnicero, F., Reiter, A., Goldman, J.M., Melo, J.V. & Cross, N.C. (2001) The t(8;22) in chronic myeloid leukemia fuses BCR to FGFR1: transforming activity and specific inhibition of FGFR1 fusion proteins. *Blood,* **98,** 3778-3783.

Dettwiler, S., Aringhieri, C., Cardinale, S., Keller, W. & Barabino, S.M. (2004) Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J Biol Chem,* **279,** 35788-35797.

Dexter, T.M. & Spooncer, E. (1987) Growth and differentiation in the hemopoietic system. *Annu Rev Cell Biol,* **3,** 423-441.

Di Fiore, P.P. & De Camilli, P. (2001) Endocytosis and signaling. an inseparable partnership. *Cell,* **106,** 1-4.

Di Guglielmo, G.M., Le, R.C., Goodfellow, A.F. & Wrana, J.L. (2003) Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nat.Cell Biol.,* **5,** 410-421.

Dohner, K., Brown, J., Hehmann, U., Hetzel, C., Stewart, J., Lowther, G., Scholl, C., Frohling, S., Cuneo, A., Tsui, L.C., Lichter, P., Scherer, S.W. & Dohner, H. (1998) Molecular cytogenetic characterization of a critical region in bands 7q35-q36 commonly deleted in malignant myeloid disorders. *Blood,* **92,** 4031-4035.

Druker, B.J., Sawyers, C.L., Kantarjian, H., Resta, D.J., Reese, S.F., Ford, J.M., Capdeville, R. & Talpaz, M. (2001) Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *N.Engl.J.Med,* **344,** 1038-1042.

Druker, B.J., Tamura, S., Buchdunger, E., Ohno, S., Segal, G.M., Fanning, S., Zimmermann, J. & Lydon, N.B. (1996) Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat.Med,* **2,** 561-566.

Drummond-Barbosa, D.A., Vaillancourt, R.R., Kazlauskas, A. & DiMaio, D. (1995) Ligand-independent activation of the platelet-derived growth factor beta receptor: requirements for bovine papillomavirus E5-induced mitogenic signaling. *Mol Cell Biol,* **15,** 2570-2581.

Dumur, C.I., Dechsukhum, C., Ware, J.L., Cofield, S.S., Best, A.M., Wilkinson, D.S., Garrett, C.T. & Ferreira-Gonzalez, A. (2003) Genome-wide detection of LOH in prostate cancer using human SNP microarray technology. *Genomics,* **81,** 260-269.

Dusa, A., Staerk, J., Elliott, J., Pecquet, C., Poirel, H.A., Johnston, J.A. & Constantinescu, S.N. (2008) Substitution of pseudokinase domain residue Val-617 by large non-polar amino acids causes activation of JAK2. *J Biol Chem,* **283,** 12941-12948.

Ebert, B.L., Pretz, J., Bosco, J., Chang, C.Y., Tamayo, P., Galili, N., Raza, A., Root, D.E., Attar, E., Ellis, S.R. & Golub, T.R. (2008) Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. *Nature,* **451,** 335-339.

Ebert, M., Yokoyama, M., Friess, H., Kobrin, M.S., Buchler, M.W. & Korc, M. (1995) Induction of platelet-derived growth factor A and B chains and over-expression of their receptors in human pancreatic cancer. *Int.J.Cancer,* **62,** 529-535.

Eguchi, M., Eguchi-Ishimae, M., Tojo, A., Morishita, K., Suzuki, K., Sato, Y., Kudoh, S., Tanaka, K., Setoyama, M., Nagamura, F., Asano, S. & Kamada, N. (1999) Fusion of ETV6 to neurotrophin-3 receptor TRKC in acute myeloid leukemia with t(12;15)(p13;q25). *Blood,* **93,** 1355-1363.

Emanuel, P.D. (2004) Juvenile myelomonocytic leukemia. *Curr Hematol Rep,* **3,** 203-209.

Emanuel, P.D. (2008) RAS pathway mutations in juvenile myelomonocytic leukemia. *Acta Haematol,* **119,** 207-211.

Engelman, J.A., Wykoff, C.C., Yasuhara, S., Song, K.S., Okamoto, T. & Lisanti, M.P. (1997) Recombinant expression of caveolin-1 in oncogenically transformed cells abrogates anchorage-independent growth. *J Biol Chem,* **272,** 16374-16381.

Espinosa, A.V., Porchia, L. & Ringel, M.D. (2007) Targeting BRAF in thyroid cancer. *Br J Cancer,* **96,** 16-20.

Faderl, S., Talpaz, M., Estrov, Z., O'Brien, S., Kurzrock, R. & Kantarjian, H.M. (1999) The biology of chronic myeloid leukemia. *N.Engl.J.Med,* **341,** 164-172.

Feiguin, F., Ferreira, A., Kosik, K.S. & Caceres, A. (1994) Kinesin-mediated organelle translocation revealed by specific cellular manipulations. *J Cell Biol,* **127,** 1021-1039.

Fenaux, P., Morel, P. & Lai, J.L. (1996) Cytogenetics of myelodysplastic syndromes. *Semin.Hematol.,* **33,** 127-138.

Fialkow, P.J., Faguet, G.B., Jacobson, R.J., Vaidya, K. & Murphy, S. (1981) Evidence that essential thrombocythemia is a clonal disorder with origin in a multipotent stem cell. *Blood,* **58,** 916-919.

Fialkow, P.J., Gartler, S.M. & Yoshida, A. (1967) Clonal origin of chronic myelocytic leukemia in man. *Proc.Natl.Acad.Sci.U.S.A,* **58,** 1468-1471.

Fialkow, P.J., Jacobson, R.J. & Papayannopoulou, T. (1977) Chronic myelocytic leukemia: clonal origin in a stem cell common to the granulocyte, erythrocyte, platelet and monocyte/macrophage. *Am.J.Med,* **63,** 125-130.

Fialkow, P.J., Jacobson, R.J., Singer, J.W., Sacher, R.A., McGuffin, R.W. & Neefe, J.R. (1980) Philadelphia chromosome (Ph1)-negative chronic myelogenous leukemia (CML): a clonal disease with origin in a multipotent stem cell. *Blood,* **56,** 70-73.

Fischer, K., Frohling, S., Scherer, S.W., McAllister Brown, J., Scholl, C., Stilgenbauer, S., Tsui, L.C., Lichter, P. & Dohner, H. (1997) Molecular cytogenetic delineation of deletions and translocations involving chromosome band 7q22 in myeloid leukemias. *Blood,* **89,** 2036-2041.

Fitzgibbon, J., Iqbal, S., Davies, A., O'Shea, D., Carlotti, E., Chaplin, T., Matthews, J., Raghavan, M., Norton, A., Lister, T.A. & Young, B.D. (2007) Genome-wide detection of recurring sites of uniparental disomy in follicular and transformed follicular lymphoma. *Leukemia,* **21,** 1514-1520.

Fitzgibbon, J., Smith, L.L., Raghavan, M., Smith, M.L., Debernardi, S., Skoulakis, S., Lillington, D., Lister, T.A. & Young, B.D. (2005) Association between acquired uniparental disomy and homozygous gene mutation in acute myeloid leukemias. *Cancer Res.,* **65,** 9152-9154.

Flotho, C., Steinemann, D., Mullighan, C.G., Neale, G., Mayer, K., Kratz, C.P., Schlegelberger, B., Downing, J.R. & Niemeyer, C.M. (2007) Genome-wide single-nucleotide polymorphism analysis in juvenile myelomonocytic leukemia identifies uniparental disomy surrounding the NF1 locus in cases associated with neurofibromatosis but not in cases with mutant RAS or PTPN11. *Oncogene,* **26,** 5816-5821.

Floyd, S. & De, C.P. (1998) Endocytosis proteins and cancer: a potential link? *Trends Cell Biol.,* **8,** 299-301.

Foulkes, W.D., Campbell, I.G., Stamp, G.W. & Trowsdale, J. (1993) Loss of heterozygosity and amplification on chromosome 11q in human ovarian cancer. *Br J Cancer,* **67,** 268-273.

Fritzler, M.J., Lung, C.C., Hamel, J.C., Griffith, K.J. & Chan, E.K. (1995) Molecular characterization of Golgin-245, a novel Golgi complex protein containing a granin signature. *J.Biol.Chem.,* **270,** 31262-31268.

Gaasenbeek, M., Howarth, K., Rowan, A.J., Gorman, P.A., Jones, A., Chaplin, T., Liu, Y., Bicknell, D., Davison, E.J., Fiegler, H., Carter, N.P., Roylance, R.R. & Tomlinson, I.P. (2006) Combined array-comparative genomic hybridization and single-nucleotide polymorphism-loss of heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers. *Cancer Res,* **66,** 3471-3479.

Gaillard, S., Bailly, Y., Benoist, M., Rakitina, T., Kessler, J.P., Fronzaroli-Molinieres, L., Dargent, B. & Castets, F. (2006) Targeting of proteins of the striatin family to dendritic spines: role of the coiled-coil domain. *Traffic,* **7,** 74-84.

Gaillard, S., Bartoli, M., Castets, F. & Monneron, A. (2001) Striatin, a calmodulin-dependent scaffolding protein, directly binds caveolin-1. *FEBS Lett,* **508,** 49-52.

Galbiati, F., Volonte, D., Engelman, J.A., Watanabe, G., Burk, R., Pestell, R.G. & Lisanti, M.P. (1998) Targeted downregulation of caveolin-1 is sufficient to drive cell transformation and hyperactivate the p42/44 MAP kinase cascade. *EMBO J,* **17,** 6633-6648.

Gallagher, G., Horsman, D.E., Tsang, P. & Forrest, D.L. (2008) Fusion of PRKG2 and SPTBN1 to the platelet-derived growth factor receptor beta gene (PDGFRB) in imatinib-responsive atypical myeloproliferative disorders. *Cancer Genet Cytogenet,* **181,** 46-51.

Ge, K. & Prendergast, G.C. (2000) Bin2, a functionally nonredundant member of the BAR adaptor gene family. *Genomics,* **67,** 210-220.

Germing, U., Gattermann, N., Strupp, C., Aivado, M. & Aul, C. (2000) Validation of the WHO proposals for a new classification of primary myelodysplastic syndromes: a retrospective analysis of 1600 patients. *Leuk.Res.,* **24,** 983-992.

Geron, I., Abrahamsson, A.E., Barroga, C.F., Kavalerchik, E., Gotlib, J., Hood, J.D., Durocher, J., Mak, C.C., Noronha, G., Soll, R.M., Tefferi, A., Kaushansky, K. & Jamieson, C.H. (2008) Selective inhibition of JAK2-driven erythroid differentiation of polycythemia vera progenitors. *Cancer Cell,* **13,** 321-330.

Goldstein, L.S. & Yang, Z. (2000) Microtubule-based transport systems in neurons: the roles of kinesins and dyneins. *Annu Rev Neurosci,* **23,** 39-71.

Golub, T.R., Barker, G.F., Lovett, M. & Gilliland, D.G. (1994) Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell,* **77,** 307-316.

Golub, T.R., Barker, G.F., Stegmaier, K. & Gilliland, D.G. (1997) The TEL gene contributes to the pathogenesis of myeloid and lymphoid leukemias by diverse molecular genetic mechanisms. *Curr Top Microbiol Immunol,* **220,** 67-79.

Gondek, L.P., Dunbar, A.J., Szpurka, H., McDevitt, M.A. & Maciejewski, J.P. (2007a) SNP array karyotyping allows for the detection of uniparental disomy and cryptic chromosomal abnormalities in MDS/MPD-U and MPD. *PLoS ONE,* **2,** e1225.

Gondek, L.P., Tiu, R., Haddad, A.S., O'Keefe, C.L., Sekeres, M.A., Theil, K.S. & Maciejewski, J.P. (2007b) Single nucleotide polymorphism arrays complement metaphase cytogenetics in detection of new chromosomal lesions in MDS. *Leukemia,* **21,** 2058-2061.

Gondek, L.P., Tiu, R., O'Keefe, C.L., Sekeres, M.A., Theil, K.S. & Maciejewski, J.P. (2008) Chromosomal lesions and uniparental disomy detected by SNP arrays in MDS, MDS/MPD, and MDS-derived AML. *Blood,* **111,** 1534-1542.

Gonzalez, M.B., Gutierrez, N.C., Garcia, J.L., Schoenmakers, E.F., Sole, F., Calasanz, M.J., San Miguel, J.F. & Hernandez, J.M. (2004) Heterogeneity of structural abnormalities in the 7q31.3 approximately q34 region in myeloid malignancies. *Cancer Genet Cytogenet,* **150,** 136-143.

Gorello, P., La Starza, R., Brandimarte, L., Trisolini, S.M., Pierini, V., Crescenzi, B., Limongi, M.Z., Nanni, M., Belloni, E., Tapinassi, C., Gerbino, E., Martelli, M.F., Foa, R., Meloni, G., Pelicci, P.G. & Mecucci, C. (2008) A PDGFRB-positive acute myeloid malignancy with a new t(5;12)(q33;p13.3) involving the ERC1 gene. *Leukemia,* **22,** 216-218.

Gorre, M.E., Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P.N. & Sawyers, C.L. (2001) Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science,* **293,** 876-880.

Gotlib, J. (2005) Molecular classification and pathogenesis of eosinophilic disorders: 2005 update. *Acta Haematol.,* **114,** 7-25.

Grand, E.K., Grand, F.H., Chase, A.J., Ross, F.M., Corcoran, M.M., Oscier, D.G. & Cross, N.C. (2004a) Identification of a novel gene, FGFR1OP2, fused to FGFR1 in 8p11 myeloproliferative syndrome. *Genes Chromosomes.Cancer,* **40,** 78-83.

Grand, F.H., Burgstaller, S., Kuhr, T., Baxter, E.J., Webersinke, G., Thaler, J., Chase, A.J. & Cross, N.C. (2004b) p53-Binding protein 1 is fused to the platelet-derived growth factor receptor beta in a patient with a t(5;15)(q33;q22) and an imatinib-responsive eosinophilic myeloproliferative disorder. *Cancer Res.,* **64,** 7216-7219.

Grand FH, H.-C.C., Ernst T, Zoi K, McGuire C, Kreil S, Jones A, Score J, Metzgeroth G, Oscier D, Hall A, Brandts C, Serve H, Reiter A, Chase AJ, Cross NC (2009) Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms. *Blood,* **113,** 6182-6192.

Grand, F.H., Iqbal, S., Zhang, L., Russell, N.H., Chase, A. & Cross, N.C. (2007) A constitutively active SPTBN1-FLT3 fusion in atypical chronic myeloid leukemia is sensitive to tyrosine kinase inhibitors and immunotherapy. *Exp Hematol,* **35,** 1723-1727.

Grand, F.H.e.a. (2009) *Haematologica*.

Griesinger, F., Hennig, H., Hillmer, F., Podleschny, M., Steffens, R., Pies, A., Wormann, B., Haase, D. & Bohlander, S.K. (2005) A BCR-JAK2 fusion gene as the result of a t(9;22)(p24;q11.2) translocation in a patient with a clinically typical chronic myeloid leukemia. *Genes Chromosomes.Cancer,* **44,** 329-333.

Griesinger, F., Janke, A., Podleschny, M. & Bohlander, S.K. (2002) Identification of an ETV6-ABL2 fusion transcript in combination with an ETV6 point mutation in a T-cell acute lymphoblastic leukaemia cell line. *Br J Haematol,* **119,** 454-458.

Griffin, J.H., Leung, J., Bruner, R.J., Caligiuri, M.A. & Briesewitz, R. (2003) Discovery of a fusion kinase in EOL-1 cells and idiopathic hypereosinophilic syndrome. *Proc.Natl.Acad.Sci.U.S.A,* **100,** 7830-7835.

Griffiths, M., Mason, J., Rindl, M., Akiki, S., McMullan, D., Stinton, V., Powell, H., Curtis, A., Bown, N. & Craddock, C. (2005) Acquired isodisomy for chromosome 13 is common in AML, and associated with FLT3-itd mutations. *Leukemia,* **19,** 2355-2358.

Growney, J.D., Clark, J.J., Adelsperger, J., Stone, R., Fabbro, D., Griffin, J.D. & Gilliland, D.G. (2005) Activation mutations of human c-KIT resistant to imatinib are sensitive to the tyrosine kinase inhibitor PKC412. *Blood***,** (zur Publikation angenommen).

Guasch, G., Delaval, B., Arnoulet, C., Xie, M.J., Xerri, L., Sainty, D., Birnbaum, D. & Pebusque, M.J. (2003a) FOP-FGFR1 tyrosine kinase, the product of a t(6;8) translocation, induces a fatal myeloproliferative disease in mice. *Blood,* **103,** 309-312.

Guasch, G., Mack, G.J., Popovici, C., Dastugue, N., Birnbaum, D., Rattner, J.B. & Pebusque, M.J. (2000) FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33). *Blood,* **95,** 1788-1796.

Guasch, G., Popovici, C., Mugneret, F., Chaffanet, M., Pontarotti, P., Birnbaum, D. & Pebusque, M.J. (2003b) Endogenous retroviral sequence is fused to FGFR1 kinase in the 8p12 stem-cell myeloproliferative disorder with t(8;19)(p12;q13.3). *Blood,* **101,** 286-288.

Gunz, F.W. (1977) The epidemiology and genetics of the chronic leukaemias. *Clin.Haematol.,* **6,** 3-20.

Hampton, G.M., Mannermaa, A., Winqvist, R., Alavaikko, M., Blanco, G., Taskinen, P.J., Kiviniemi, H., Newsham, I., Cavenee, W.K. & Evans, G.A. (1994) Loss of heterozygosity in sporadic human breast carcinoma: a common region between 11q22 and 11q23.3. *Cancer Res,* **54,** 4586-4589.

Handa, N., Kukimoto-Niino, M., Akasaka, R., Kishishita, S., Murayama, K., Terada, T., Inoue, M., Kigawa, T., Kose, S., Imamoto, N., Tanaka, A., Hayashizaki, Y., Shirouzu, M. & Yokoyama, S. (2006) The crystal structure of mouse Nup35 reveals atypical RNP motifs and novel homodimerization of the RRM domain. *J Mol Biol,* **363,** 114-124.

Hasle, H., Arico, M., Basso, G., Biondi, A., Cantu Rajnoldi, A., Creutzig, U., Fenu, S., Fonatsch, C., Haas, O.A., Harbott, J., Kardos, G., Kerndrup, G., Mann, G., Niemeyer, C.M., Ptoszkova, H., Ritter, J., Slater, R., Stary, J., Stollmann-Gibbels, B., Testi, A.M., van Wering, E.R. & Zimmermann, M. (1999) Myelodysplastic syndrome, juvenile myelomonocytic leukemia, and acute myeloid leukemia associated with complete or partial monosomy 7. European Working Group on MDS in Childhood (EWOG-MDS). *Leukemia,* **13,** 376-385.

Heinrich, M.C., Corless, C.L., Duensing, A., McGreevey, L., Chen, C.J., Joseph, N., Singer, S., Griffith, D.J., Haley, A., Town, A., Demetri, G.D., Fletcher, C.D. & Fletcher, J.A. (2003) PDGFRA activating mutations in gastrointestinal stromal tumors. *Science,* **299,** 708-710.

Hermanson, M., Funa, K., Koopmann, J., Maintz, D., Waha, A., Westermark, B., Heldin, C.H., Wiestler, O.D., Louis, D.N., von, D.A. & Nister, M. (1996) Association of loss of heterozygosity on chromosome 17p with high platelet-derived growth factor alpha receptor expression in human malignant gliomas. *Cancer Res.,* **56,** 164-171.

Hernandez, J.M., Mecucci, C., Michaux, L., Criel, A., Stul, M., Meeus, P., Wlodarska, I., Van Orshoven, A., Cassiman, J.J., De Wolf-Peeters, C. & Van den

Berghe, H. (1997) del(7q) in chronic B-cell lymphoid malignancies. *Cancer Genet Cytogenet,* **93,** 147-151.

Hexner, E.O., Serdikoff, C., Jan, M., Swider, C.R., Robinson, C., Yang, S., Angeles, T., Emerson, S.G., Carroll, M., Ruggeri, B. & Dobrzanski, P. (2008) Lestaurtinib (CEP701) is a JAK2 inhibitor that suppresses JAK2/STAT5 signaling and the proliferation of primary erythroid cells from patients with myeloproliferative disorders. *Blood,* **111,** 5663-5671.

Hidalgo-Curtis, C., Chase, A., Drachenberg, M., Roberts, M.W., Finkelstein, J.Z., Mould, S., Oscier, D., Cross, N.C. & Grand, F.H. (2008) The t(1;9)(p34;q34) and t(8;12)(p11;q15) fuse pre-mRNA processing proteins SFPQ (PSF) and CPSF6 to ABL and FGFR1. *Genes Chromosomes Cancer,* **47,** 379-385.

Hirokawa, N., Pfister, K.K., Yorifuji, H., Wagner, M.C., Brady, S.T. & Bloom, G.S. (1989) Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration. *Cell,* **56,** 867-878.

Hirokawa, N. & Takemura, R. (2005) Molecular motors and mechanisms of directional transport in neurons. *Nat Rev Neurosci,* **6,** 201-214.

Hirota, S., Isozaki, K., Moriyama, Y., Hashimoto, K., Nishida, T., Ishiguro, S., Kawano, K., Hanada, M., Kurata, A., Takeda, M., Muhammad, T.G., Matsuzawa, Y., Kanakura, Y., Shinomura, Y. & Kitamura, Y. (1998) Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. *Science,* **279,** 577-580.

Hoang, T. (2004) The origin of hematopoietic cell type diversity. *Oncogene,* **23,** 7188-7198.

Hoch, R.V. & Soriano, P. (2006) Context-specific requirements for Fgfr1 signaling through Frs2 and Frs3 during mouse development. *Development,* **133,** 663-673.

Hochhaus, A., Kreil, S., Corbin, A., La, R.P., Lahaye, T., Berger, U., Cross, N.C., Linkesch, W., Druker, B.J., Hehlmann, R., Passerini, C., Corneo, G. & D'Incalci, M. (2001) Roots of clinical resistance to STI-571 cancer therapy. *Science,* **293,** 2163.

Holtkamp, N., Ziegenhagen, N., Malzer, E., Hartmann, C., Giese, A. & von Deimling, A. (2007) Characterization of the amplicon on chromosomal segment 4q12 in glioblastoma multiforme. *Neuro Oncol,* **9,** 291-297.

Hooberman, A.L., Carrino, J.J., Leibowitz, D., Rowley, J.D., Le Beau, M.M., Arlin, Z.A. & Westbrook, C.A. (1989) Unexpected heterogeneity of BCR-ABL fusion mRNA detected by polymerase chain reaction in Philadelphia chromosome-positive acute lymphoblastic leukemia. *Proc.Natl.Acad.Sci.U.S.A,* **86,** 4259-4263.

Huang, T.G., Suhan, J. & Hackney, D.D. (1994) Drosophila kinesin motor domain extending to amino acid position 392 is dimeric when expressed in Escherichia coli. *J Biol Chem,* **269,** 32708.

Huettner, C.S., Zhang, P., Van Etten, R.A. & Tenen, D.G. (2000) Reversibility of acute B-cell leukaemia induced by BCR-ABL1. *Nat.Genet.,* **24,** 57-60.

Hyun, T.S. & Ross, T.S. (2004) HIP1: trafficking roles and regulation of tumorigenesis. *Trends Mol.Med,* **10,** 194-199.

Irving, J.A., Bloodworth, L., Bown, N.P., Case, M.C., Hogarth, L.A. & Hall, A.G. (2005) Loss of heterozygosity in childhood acute lymphoblastic leukemia detected by genome-wide microarray single nucleotide polymorphism analysis. *Cancer Res.,* **65,** 3053-3058.

Isshiki, M. & Anderson, R.G. (1999) Calcium signal transduction from caveolae. *Cell Calcium,* **26,** 201-208.

Jacobson, R.J., Salo, A. & Fialkow, P.J. (1978) Agnogenic myeloid metaplasia: a clonal proliferation of hematopoietic stem cells with secondary myelofibrosis. *Blood,* **51,** 189-194.

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. & Vainchenker, W. (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature,* **434,** 1144-1148.

Jamieson, C.H., Gotlib, J., Durocher, J.A., Chao, M.P., Mariappan, M.R., Lay, M., Jones, C., Zehnder, J.L., Lilleberg, S.L. & Weissman, I.L. (2006) The JAK2 V617F mutation occurs in hematopoietic stem cells in polycythemia vera and predisposes toward erythroid differentiation. *Proc Natl Acad Sci U S A,* **103,** 6224-6229.

Jeffers, M., Rong, S. & Woude, G.F. (1996) Hepatocyte growth factor/scatter factor-Met signaling in tumorigenicity and invasion/metastasis. *J Mol Med,* **74,** 505-513.

Jiang, G. & Hunter, T. (1999) Receptor signaling: when dimerization is not enough. *Curr.Biol.,* **9,** R568-R571.

Johansson, B., Fioretos, T. & Mitelman, F. (2002) Cytogenetic and molecular genetic evolution of chronic myeloid leukemia. *Acta Haematol,* **107,** 76-94.

Jones, A.V., Chase, A., Silver, R.T., Oscier, D., Zoi, K., Wang, Y.L., Cario, H., Pahl, H.L., Collins, A., Reiter, A., Grand, F. & Cross, N.C. (2009) JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. *Nat Genet,* **41,** 446-449.

Jones, A.V. & Cross, N.C. (2004) Oncogenic derivatives of platelet-derived growth factor receptors. *Cell Mol.Life Sci.,* **61,** 2912-2923.

Jones, A.V., Kreil, S., Zoi, K., Waghorn, K., Curtis, C., Zhang, L., Score, J., Seear, R., Chase, A.J., Grand, F.H., White, H., Zoi, C., Loukopoulos, D., Terpos, E., Vervessou, E.C., Schultheis, B., Emig, M., Ernst, T., Lengfelder, E., Hehlmann, R., Hochhaus, A., Oscier, D., Silver, R.T., Reiter, A. & Cross, N.C. (2005) Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. *Blood*.

Jucker, M., Gunther, A., Gradl, G., Fonatsch, C., Krueger, G., Diehl, V. & Tesch, H. (1994) The Met/hepatocyte growth factor receptor (HGFR) gene is overexpressed in some cases of human leukemia and lymphoma. *Leuk Res*, **18,** 7-16.

Kantarjian, H.M., Giles, F., Gattermann, N., Bhalla, K., Alimena, G., Palandri, F., Ossenkoppele, G.J., Nicolini, F.E., O'Brien, S.G., Litzow, M., Bhatia, R., Cervantes, F., Haque, A., Shou, Y., Resta, D.J., Weitzman, A., Hochhaus, A. & le Coutre, P. (2007) Nilotinib (formerly AMN107), a highly selective BCR-ABL tyrosine kinase inhibitor, is effective in patients with Philadelphia chromosome-positive chronic myelogenous leukemia in chronic phase following imatinib resistance and intolerance. *Blood,* **110,** 3540-3546.

Karow, A., Steinemann, D., Gohring, G., Hasle, H., Greiner, J., Harila-Saari, A., Flotho, C., Zenker, M., Schlegelberger, B., Niemeyer, C.M. & Kratz, C.P. (2007) Clonal duplication of a germline PTPN11 mutation due to acquired uniparental disomy in acute lymphoblastic leukemia blasts from a patient with Noonan syndrome. *Leukemia,* **21,** 1303-1305.

Kasyapa, C.S., Kunapuli, P. & Cowell, J.K. (2005) Mass spectroscopy identifies the splicing-associated proteins, PSF, hnRNP H3, hnRNP A2/B1, and TLS/FUS as interacting partners of the ZNF198 protein associated with rearrangement in myeloproliferative disease. *Exp Cell Res*, **309,** 78-85.

Kato, Y., Iwama, A., Tadokoro, Y., Shimoda, K., Minoguchi, M., Akira, S., Tanaka, M., Miyajima, A., Kitamura, T. & Nakauchi, H. (2005) Selective activation of

STAT5 unveils its role in stem cell self-renewal in normal and leukemic hematopoiesis. *J Exp Med,* **202,** 169-179.

Kaufmann, I., Martin, G., Friedlein, A., Langen, H. & Keller, W. (2004) Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J,* **23,** 616-626.

Keldysh, P.L., Dragani, T.A., Fleischman, E.W., Konstantinova, L.N., Perevoschikov, A.G., Pierotti, M.A., Della Porta, G. & Kopnin, B.P. (1993) 11q deletions in human colorectal carcinomas: cytogenetics and restriction fragment length polymorphism analysis. *Genes Chromosomes Cancer,* **6,** 45-50.

Kennedy, G.C., Matsuzaki, H., Dong, S., Liu, W.M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J., Liu, W., Yang, G., Di, X., Ryder, T., He, Z., Surti, U., Phillips, M.S., Boyce-Jacino, M.T., Fodor, S.P. & Jones, K.W. (2003) Large-scale genotyping of complex DNA. *Nat Biotechnol,* **21,** 1233-1237.

Kere, J., Donis-Keller, H., Ruutu, T. & de la Chapelle, A. (1989) Chromosome 7 long-arm deletions in myeloid disorders: terminal DNA sequences are commonly conserved and breakpoints vary. *Cytogenet Cell Genet,* **50,** 226-229.

Kere, J., Ruutu, T., Lahtinen, R. & de la Chapelle, A. (1987) Molecular characterization of chromosome 7 long arm deletions in myeloid disorders. *Blood,* **70,** 1349-1353.

Kerndrup, G., Pedersen, B. & Bendix-Hansen, K. (1987) Specific minor chromosome deletions in myelodysplastic syndromes: clinical and morphologic correlations. *Cancer Genet Cytogenet,* **26,** 227-234.

Kharas, M.G. & Fruman, D.A. (2005) ABL oncogenes and phosphoinositide 3-kinase: mechanism of activation and downstream effectors. *Cancer Res.,* **65,** 2047-2053.

Kilpivaara, O., Mukherjee, S., Schram, A.M., Wadleigh, M., Mullally, A., Ebert, B.L., Bass, A., Marubayashi, S., Heguy, A., Garcia-Manero, G., Kantarjian, H., Offit, K., Stone, R.M., Gilliland, D.G., Klein, R.J. & Levine, R.L. (2009) A germline JAK2 SNP is associated with predisposition to the development of JAK2(V617F)-positive myeloproliferative neoplasms. *Nat Genet,* **41,** 455-459.

Kjer-Nielsen, L., Teasdale, R.D., van Vliet, C. & Gleeson, P.A. (1999) A novel Golgi-localisation domain shared by a class of coiled-coil peripheral membrane proteins. *Curr Biol,* **9,** 385-388.

Klion, A.D., Noel, P., Akin, C., Law, M.A., Gilliland, D.G., Cools, J., Metcalfe, D.D. & Nutman, T.B. (2003) Elevated serum tryptase levels identify a subset of patients with a myeloproliferative variant of idiopathic hypereosinophilic syndrome associated with tissue fibrosis, poor prognosis, and imatinib responsiveness. *Blood,* **101,** 4660-4666.

Klion, A.D., Robyn, J., Akin, C., Noel, P., Brown, M., Law, M., Metcalfe, D.D., Dunbar, C. & Nutman, T.B. (2004) Molecular remission and reversal of myelofibrosis in response to imatinib mesylate treatment in patients with the myeloproliferative variant of hypereosinophilic syndrome. *Blood,* **103,** 473-478.

Knudson, A.G., Jr. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A,* **68,** 820-823.

Koleske, A.J., Baltimore, D. & Lisanti, M.P. (1995) Reduction of caveolin and caveolae in oncogenically transformed cells. *Proc Natl Acad Sci U S A,* **92,** 1381-1385.

Kooy, J., Toh, B.H., Pettitt, J.M., Erlich, R. & Gleeson, P.A. (1992) Human autoantibodies as reagents to conserved Golgi components. Characterization of a peripheral, 230-kDa compartment-specific Golgi protein. *J.Biol.Chem.,* **267,** 20255-20263.

Kotecha, N., Flores, N.J., Irish, J.M., Simonds, E.F., Sakai, D.S., Archambeault, S., Diaz-Flores, E., Coram, M., Shannon, K.M., Nolan, G.P. & Loh, M.L. (2008) Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. *Cancer Cell,* **14,** 335-343.

Kralovics, R., Guan, Y. & Prchal, J.T. (2002) Acquired uniparental disomy of chromosome 9p is a frequent stem cell defect in polycythemia vera. *Exp.Hematol.,* **30,** 229-236.

Kralovics, R., Passamonti, F., Buser, A.S., Teo, S.S., Tiedt, R., Passweg, J.R., Tichelli, A., Cazzola, M. & Skoda, R.C. (2005a) A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N.Engl.J.Med,* **352,** 1779-1790.

Kralovics, R., Teo, S.S., Buser, A.S., Brutsche, M., Tiedt, R., Tichelli, A., Passamonti, F., Pietra, D., Cazzola, M. & Skoda, R.C. (2005b) Altered gene expression in myeloproliferative disorders correlates with activation of signaling by the V617F mutation of Jak2. *Blood,* **106,** 3374-3376.

Krivtsov, A.V. & Armstrong, S.A. (2007) MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer,* **7,** 823-833.

Kulkarni, S., Heath, C., Parker, S., Chase, A., Iqbal, S., Pocock, C.F., Kaeda, J., Cwynarski, K., Goldman, J.M. & Cross, N.C. (2000) Fusion of H4/D10S170 to the platelet-derived growth factor receptor beta in BCR-ABL-negative myeloproliferative disorders with a t(5;10)(q33;q21). *Cancer Res.,* **60,** 3592-3598.

Kuno, Y., Abe, A., Emi, N., Iida, M., Yokozawa, T., Towatari, M., Tanimoto, M. & Saito, H. (2001) Constitutive kinase activation of the TEL-Syk fusion gene in myelodysplastic syndrome with t(9;12)(q22;p12). *Blood,* **97,** 1050-1055.

La Starza, R., Rosati, R., Roti, G., Gorello, P., Bardi, A., Crescenzi, B., Pierini, V., Calabrese, O., Baens, M., Folens, C., Cools, J., Marynen, P., Martelli, M.F., Mecucci,

C. & Cuneo, A. (2007) A new NDE1/PDGFRB fusion transcript underlying chronic myelomonocytic leukaemia in Noonan Syndrome. *Leukemia,* **21,** 830-833.

La Starza, R., Specchia, G., Cuneo, A., Beacci, D., Nozzoli, C., Luciano, L., Aventin, A., Sambani, C., Testoni, N., Foppoli, M., Invernizzi, R., Marynen, P., Martelli, M.F. & Mecucci, C. (2005) The hypereosinophilic syndrome: fluorescence in situ hybridization detects the del(4)(q12)-FIP1L1/PDGFRA but not genomic rearrangements of other tyrosine kinases. *Haematologica,* **90,** 596-601.

Lamorte, L. & Park, M. (2001) The receptor tyrosine kinases: role in cancer progression. *Surg.Oncol.Clin.N.Am.,* **10,** 271-288, viii.

Landgren, O., Goldin, L.R., Kristinsson, S.Y., Helgadottir, E.A., Samuelsson, J. & Bjorkholm, M. (2008) Increased risks of polycythemia vera, essential thrombocythemia, and myelofibrosis among 24,577 first-degree relatives of 11,039 patients with myeloproliferative neoplasms in Sweden. *Blood,* **112,** 2199-2204.

Langdon, W.Y., Hartley, J.W., Klinken, S.P., Ruscetti, S.K. & Morse, H.C., 3rd (1989) v-cbl, an oncogene from a dual-recombinant murine retrovirus that induces early B-lineage lymphomas. *Proc Natl Acad Sci U S A,* **86,** 1168-1172.

Latil, A., Cussenot, O., Fournier, G., Baron, J.C. & Lidereau, R. (1995) Loss of heterozygosity at 7q31 is a frequent and early event in prostate cancer. *Clin Cancer Res,* **1,** 1385-1389.

Le Beau, M.M., Espinosa, R., 3rd, Davis, E.M., Eisenbart, J.D., Larson, R.A. & Green, E.D. (1996) Cytogenetic and molecular delineation of a region of chromosome 7 commonly deleted in malignant myeloid diseases. *Blood,* **88,** 1930-1935.

Lee, J.H., Han, S.U., Cho, H., Jennings, B., Gerrard, B., Dean, M., Schmidt, L., Zbar, B. & Vande Woude, G.F. (2000) A novel germ line juxtamembrane Met mutation in human gastric cancer. *Oncogene,* **19,** 4947-4953.

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. & Gilliland, D.G. (2005a) Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell,* **7,** 387-397.

Levine, R.L., Wadleigh, M., Sternberg, D.W., Wlodarska, I., Galinsky, I., Stone, R.M., DeAngelo, D.J., Gilliland, D.G. & Cools, J. (2005b) KIAA1509 is a novel PDGFRB fusion partner in imatinib-responsive myeloproliferative disease associated with a t(5;14)(q33;q32). *Leukemia,* **19,** 27-30.

Levitzki, A. (2004) PDGF receptor kinase inhibitors for the treatment of PDGF driven diseases. *Cytokine Growth Factor Rev.,* **15,** 229-235.

Lewis, S., Abrahamson, G., Boultwood, J., Fidler, C., Potter, A. & Wainscoat, J.S. (1996) Molecular characterization of the 7q deletion in myeloid disorders. *Br J Haematol,* **93,** 75-80.

Liang, H., Fairman, J., Claxton, D.F., Nowell, P.C., Green, E.D. & Nagarajan, L. (1998) Molecular anatomy of chromosome 7q deletions in myeloid neoplasms: evidence for multiple critical loci. *Proc Natl Acad Sci U S A,* **95,** 3781-3785.

Lichter, P. & Cremer, T. (1992) Chromosome analysis by non-isotopic in situ hybridisation. p. 157. IRL, Oxford, UK.

Lierman, E., Folens, C., Stover, E.H., Mentens, N., Van Miegroet, H., Scheers, W., Boogaerts, M., Vandenberghe, P., Marynen, P. & Cools, J. (2006) Sorafenib is a potent inhibitor of FIP1L1-PDGFRalpha and the imatinib-resistant FIP1L1-PDGFRalpha T674I mutant. *Blood,* **108,** 1374-1376.

Lierman, E., Lahortiga, I., Van Miegroet, H., Mentens, N., Marynen, P. & Cools, J. (2007) The ability of sorafenib to inhibit oncogenic PDGFRbeta and FLT3 mutants and overcome resistance to other small molecule inhibitors. *Haematologica,* **92,** 27-34.

Lieu, Z.Z., Derby, M.C., Teasdale, R.D., Hart, C., Gunn, P. & Gleeson, P.A. (2007) The golgin GCC88 is required for efficient retrograde transport of cargo from the early endosomes to the trans-Golgi network. *Mol Biol Cell,* **18,** 4979-4991.

Lieu, Z.Z., Lock, J.G., Hammond, L.A., La Gruta, N.L., Stow, J.L. & Gleeson, P.A. (2008) A trans-Golgi network golgin is required for the regulated secretion of TNF in activated macrophages in vivo. *Proc Natl Acad Sci U S A,* **105,** 3351-3356.

Lindblad-Toh, K., Tanenbaum, D.M., Daly, M.J., Winchester, E., Lui, W.O., Villapakkam, A., Stanton, S.E., Larsson, C., Hudson, T.J., Johnson, B.E., Lander, E.S. & Meyerson, M. (2000) Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat Biotechnol,* **18,** 1001-1005.

Lisanti, M.P., Scherer, P.E., Tang, Z. & Sargiacomo, M. (1994) Caveolae, caveolin and caveolin-rich membrane domains: a signalling hypothesis. *Trends Cell Biol,* **4,** 231-235.

Lopez-Knowles, E., Hernandez, S., Malats, N., Kogevinas, M., Lloreta, J., Carrato, A., Tardon, A., Serra, C. & Real, F.X. (2006) PIK3CA mutations are an early genetic alteration associated with FGFR3 mutations in superficial papillary bladder tumors. *Cancer Res,* **66,** 7401-7404.

Lu, L., Tai, G. & Hong, W. (2004a) Autoantigen Golgin-97, an effector of Arl1 GTPase, participates in traffic from the endosome to the trans-golgi network. *Mol Biol Cell,* **15,** 4426-4443.

Lu, L., Tai, G., Wu, M., Song, H. & Hong, W. (2006) Multilayer interactions determine the Golgi localization of GRIP golgins. *Traffic,* **7,** 1399-1407.

Lu, Q., Pallas, D.C., Surks, H.K., Baur, W.E., Mendelsohn, M.E. & Karas, R.H. (2004b) Striatin assembles a membrane signaling complex necessary for rapid, nongenomic activation of endothelial NO synthase by estrogen receptor alpha. *Proc Natl Acad Sci U S A,* **101,** 17126-17131.

Luke, M.R., Kjer-Nielsen, L., Brown, D.L., Stow, J.L. & Gleeson, P.A. (2003) GRIP domain-mediated targeting of two new coiled-coil proteins, GCC88 and GCC185, to subcompartments of the trans-Golgi network. *J.Biol.Chem.,* **278,** 4216-4226.

Luna-Fineman, S., Shannon, K.M. & Lange, B.J. (1995) Childhood monosomy 7: epidemiology, biology, and mechanistic implications. *Blood,* **85,** 1985-1999.

Lyon, M.F. (1988) The William Allan memorial award address: X-chromosome inactivation and the location and expression of X-linked genes. *Am.J.Hum.Genet.,* **42,** 8-16.

Ma, P.C., Kijima, T., Maulik, G., Fox, E.A., Sattler, M., Griffin, J.D., Johnson, B.E. & Salgia, R. (2003) c-MET mutational analysis in small cell lung cancer: novel juxtamembrane domain mutations regulating cytoskeletal functions. *Cancer Res,* **63,** 6272-6281.

Macchi, P., Villa, A., Giliani, S., Sacco, M.G., Frattini, A., Porta, F., Ugazio, A.G., Johnston, J.A., Candotti, F., O'Shea, J.J. & et al. (1995) Mutations of Jak-3 gene in patients with autosomal severe combined immune deficiency (SCID). *Nature,* **377,** 65-68.

Macdonald, D., Reiter, A. & Cross, N.C. (2002a) The 8p11 myeloproliferative syndrome: a distinct clinical entity caused by constitutive activation of FGFR1. *Acta Haematol,* **107,** 101-107.

Macdonald, D., Reiter, A. & Cross, N.C.P. (2002b) The 8p11 myeloproliferative syndrome: a distinct clinical entity caused by constitutive activation of FGFR1. *Acta Haematol.,* **107,** 101-107.

MacDonald, T.J., Brown, K.M., LaFleur, B., Peterson, K., Lawlor, C., Chen, Y., Packer, R.J., Cogen, P. & Stephan, D.A. (2001) Expression profiling of medulloblastoma: PDGFRA and the RAS/MAPK pathway as therapeutic targets for metastatic disease. *Nat.Genet.,* **29,** 143-152.

Magnusson, M.K., Meade, K.E., Brown, K.E., Arthur, D.C., Krueger, L.A., Barrett, A.J. & Dunbar, C.E. (2001) Rabaptin-5 is a novel fusion partner to platelet-derived growth factor beta receptor in chronic myelomonocytic leukemia. *Blood,* **98,** 2518-2525.

Magnusson, M.K., Meade, K.E., Nakamura, R., Barrett, J. & Dunbar, C.E. (2002) Activity of STI571 in chronic myelomonocytic leukemia with a platelet-derived growth factor beta receptor fusion oncogene. *Blood,* **100,** 1088-1091.

Mahon, F.X., Deininger, M.W., Schultheis, B., Chabrol, J., Reiffers, J., Goldman, J.M. & Melo, J.V. (2000) Selection and characterization of BCR-ABL positive cell lines with differential sensitivity to the tyrosine kinase inhibitor STI571: diverse mechanisms of resistance. *Blood,* **96,** 1070-1079.

Mallard, F., Tang, B.L., Galli, T., Tenza, D., Saint-Pol, A., Yue, X., Antony, C., Hong, W., Goud, B. & Johannes, L. (2002) Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform. *J Cell Biol,* **156,** 653-664.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T. & Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science,* **298,** 1912-1934.

Mantadakis, E., Shannon, K.M., Singer, D.A., Finklestein, J., Chan, K.W., Hilden, J.M. & Sandler, E.S. (1999) Transient monosomy 7: a case series in children and review of the literature. *Cancer,* **85,** 2655-2661.

Marcucci, G., Strout, M.P., Bloomfield, C.D. & Caligiuri, M.A. (1998) Detection of unique ALL1 (MLL) fusion transcripts in normal human bone marrow and blood: distinct origin of normal versus leukemic ALL1 fusion transcripts. *Cancer Res,* **58,** 790-793.

Mardis, E.R. (2008) Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet,* **9,** 387-402.

Marmor, M.D. & Yarden, Y. (2004) Role of protein ubiquitylation in regulating endocytosis of receptor tyrosine kinases. *Oncogene,* **23,** 2057-2070.

Matsui, T., Heidaran, M., Miki, T., Popescu, N., La Rochelle, W., Kraus, M., Pierce, J. & Aaronson, S. (1989) Isolation of a novel receptor cDNA establishes the existence of two PDGF receptor genes. *Science,* **243,** 800-804.

Matsuzaki, H., Dong, S., Loi, H., Di, X., Liu, G., Hubbell, E., Law, J., Berntsen, T., Chadha, M., Hui, H., Yang, G., Kennedy, G.C., Webster, T.A., Cawley, S., Walsh, P.S., Jones, K.W., Fodor, S.P. & Mei, R. (2004) Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nat Methods,* **1,** 109-111.

McGahon, A., Bissonnette, R., Schmitt, M., Cotter, K.M., Green, D.R. & Cotter, T.G. (1994) BCR-ABL maintains resistance of chronic myelogenous leukemia cells to apoptotic cell death. *Blood,* **83,** 1179-1187.

McWhirter, J.R. & Wang, J.Y. (1997) Effect of Bcr sequences on the cellular function of the Bcr-Abl oncoprotein. *Oncogene,* **15,** 1625-1634.

Mercher, T., Wernig, G., Moore, S.A., Levine, R.L., Gu, T.L., Frohling, S., Cullen, D., Polakiewicz, R.D., Bernard, O.A., Boggon, T.J., Lee, B.H. & Gilliland, D.G. (2006) JAK2T875N is a novel activating mutation that results in myeloproliferative disease with features of megakaryoblastic leukemia in a murine bone marrow transplantation model. *Blood,* **108,** 2770-2779.

Metzgeroth, G., Walz, C., Score, J., Siebert, R., Schnittger, S., Haferlach, C., Popp, H., Haferlach, T., Erben, P., Mix, J., Muller, M.C., Beneke, H., Muller, L., Del Valle, F., Aulitzky, W.E., Wittkowsky, G., Schmitz, N., Schulte, C., Muller-Hermelink, K., Hodges, E., Whittaker, S.J., Diecker, F., Dohner, H., Schuld, P., Hehlmann, R., Hochhaus, A., Cross, N.C. & Reiter, A. (2007) Recurrent finding of the FIP1L1-PDGFRA fusion gene in eosinophilia-associated acute myeloid leukemia and lymphoblastic T-cell lymphoma. *Leukemia,* **21,** 1183-1188.

Meyer, C., Schneider, B., Jakob, S., Strehl, S., Attarbaschi, A., Schnittger, S., Schoch, C., Jansen, M.W., van Dongen, J.J., den Boer, M.L., Pieters, R., Ennas, M.G., Angelucci, E., Koehl, U., Greil, J., Griesinger, F., Zur Stadt, U., Eckert, C., Szczepanski, T., Niggli, F.K., Schafer, B.W., Kempski, H., Brady, H.J., Zuna, J., Trka, J., Nigro, L.L., Biondi, A., Delabesse, E., Macintyre, E., Stanulla, M., Schrappe, M., Haas, O.A., Burmeister, T., Dingermann, T., Klingebiel, T. & Marschalek, R. (2006) The MLL recombinome of acute leukemias. *Leukemia,* **20,** 777-784.

Million, R.P., Harakawa, N., Roumiantsev, S., Varticovski, L. & Van Etten, R.A. (2004) A direct binding site for Grb2 contributes to transformation and leukemogenesis by the Tel-Abl (ETV6-Abl) tyrosine kinase. *Mol.Cell Biol.,* **24,** 4685-4695.

Mitelman, F. (1993) The cytogenetic scenario of chronic myeloid leukemia. *Leuk.Lymphoma,* **11 Suppl 1,** 11-15.

Mitelman, F., Levan, G., Nilsson, P.G. & Brandt, L. (1976) Non-random karyotypic evolution in chronic myeloid leukemia. *Int.J.Cancer,* **18,** 24-30.

Mitsiades, C.S., Mitsiades, N., Poulaki, V., Schlossman, R., Akiyama, M., Chauhan, D., Hideshima, T., Treon, S.P., Munshi, N.C., Richardson, P.G. & Anderson, K.C. (2002) Activation of NF-kappaB and upregulation of intracellular anti-apoptotic proteins via the IGF-1/Akt signaling in human multiple myeloma cells: therapeutic implications. *Oncogene,* **21,** 5673-5683.

Mohammadi, M., McMahon, G., Sun, L., Tang, C., Hirth, P., Yeh, B.K., Hubbard, S.R. & Schlessinger, J. (1997) Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science,* **276,** 955-960.

Montpetit, A., Boily, G. & Sinnett, D. (2002) A detailed transcriptional map of the chromosome 12p12 tumour suppressor locus. *Eur J Hum Genet,* **10,** 62-71.

Morerio, C., Acquila, M., Rosanda, C., Rapella, A., Dufour, C., Locatelli, F., Maserati, E., Pasquali, F. & Panarello, C. (2004) HCMOGT-1 is a novel fusion partner to PDGFRB in juvenile myelomonocytic leukemia with t(5;17)(q33;p11.2). *Cancer Res.,* **64,** 2649-2651.

Morozova, O. & Marra, M.A. (2008) Applications of next-generation sequencing technologies in functional genomics. *Genomics,* **92,** 255-264.

Mosesson, Y., Mills, G.B. & Yarden, Y. (2008) Derailed endocytosis: an emerging feature of cancer. *Nat Rev Cancer,* **8,** 835-850.

Mosselman, S., Claesson-Welsh, L., Kamphuis, J.S. & van Zoelen, E.J. (1994) Developmentally regulated expression of two novel platelet-derived growth factor alpha-receptor transcripts in human teratocarcinoma cells. *Cancer Res.,* **54,** 220-225.

Mosselman, S., Looijenga, L.H., Gillis, A.J., van Rooijen, M.A., Kraft, H.J., van Zoelen, E.J. & Oosterhuis, J.W. (1996) Aberrant platelet-derived growth factor alpha-receptor transcript as a diagnostic marker for early human germ cell tumors of the adult testis. *Proc.Natl.Acad.Sci.U.S.A,* **93,** 2884-2888.

Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.H., Relling, M.V., Evans, W.E., Shurtleff, S.A. & Downing, J.R. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature,* **446,** 758-764.

Munro, S. & Nichols, B.J. (1999) The GRIP domain - a novel Golgi-targeting domain found in several coiled-coil proteins. *Curr Biol,* **9,** 377-380.

Murphy, M.A., Schnall, R.G., Venter, D.J., Barnett, L., Bertoncello, I., Thien, C.B., Langdon, W.Y. & Bowtell, D.D. (1998) Tissue hyperplasia and enhanced T-cell signalling via ZAP-70 in c-Cbl-deficient mice. *Mol Cell Biol,* **18,** 4872-4882.

Nakata, T. & Hirokawa, N. (1995) Point mutation of adenosine triphosphate-binding motif generated rigor kinesin that selectively blocks anterograde lysosome membrane transport. *J Cell Biol,* **131,** 1039-1053.

Nath, S., Bananis, E., Sarkar, S., Stockert, R.J., Sperry, A.O., Murray, J.W. & Wolkoff, A.W. (2007) Kif5B and Kifc1 interact and are required for motility and fission of early endocytic vesicles in mouse liver. *Mol Biol Cell,* **18,** 1839-1849.

Nigg, E.A. (2002) Centrosome aberrations: cause or consequence of cancer progression? *Nat Rev Cancer,* **2,** 815-825.

Noren, N.K., Foos, G., Hauser, C.A. & Pasquale, E.B. (2006) The EphB4 receptor suppresses breast cancer cell tumorigenicity through an Abl-Crk pathway. *Nat Cell Biol,* **8,** 815-825.

Nowell & Hungerford (1960) *Science,* **132,** 1497.

O'Brien, S.G., Guilhot, F., Larson, R.A., Gathmann, I., Baccarani, M., Cervantes, F., Cornelissen, J.J., Fischer, T., Hochhaus, A., Hughes, T., Lechner, K., Nielsen, J.L., Rousselot, P., Reiffers, J., Saglio, G., Shepherd, J., Simonsson, B., Gratwohl, A., Goldman, J.M., Kantarjian, H., Taylor, K., Verhoef, G., Bolton, A.E., Capdeville, R. & Druker, B.J. (2003) Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N.Engl.J.Med,* **348,** 994-1004.

Odero, M.D., Carlson, K., Calasanz, M.J., Lahortiga, I., Chinwalla, V. & Rowley, J.D. (2001) Identification of new translocations involving ETV6 in hematologic malignancies by fluorescence in situ hybridization and spectral karyotyping. *Genes Chromosomes Cancer,* **31,** 134-142.

Offit, K., Louie, D.C., Parsa, N.Z., Noy, A. & Chaganti, R.S. (1995) Del (7)(q32) is associated with a subset of small lymphocytic lymphoma with plasmacytoid features. *Blood,* **86,** 2365-2370.

Oikonomou, E. & Pintzas, A. (2006) Cancer genetics of sporadic colorectal cancer: BRAF and PI3KCA mutations, their impact on signaling and novel targeted therapies. *Anticancer Res,* **26,** 1077-1084.

Olabisi OO, M.G., Kostenko EV, Liu Z, Ozer HL, Whitehead IP (2006) Bcr interacts with components of the endosomal sorting complex required for transport-I and is required for epidermal growth factor receptor turnover. *Cancer Research,* **66,** 6250-6257.

Olcaydu, D., Harutyunyan, A., Jager, R., Berg, T., Gisslinger, B., Pabinger, I., Gisslinger, H. & Kralovics, R. (2009) A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. *Nat Genet,* **41,** 450-454.

Ollendorff, V., Guasch, G., Isnardon, D., Galindo, R., Birnbaum, D. & Pebusque, M.J. (1999) Characterization of FIM-FGFR1, the fusion product of the myeloproliferative disorder-associated t(8;13) translocation. *J.Biol.Chem.,* **274,** 26922-26930.

Onida, F., Kantarjian, H.M., Smith, T.L., Ball, G., Keating, M.J., Estey, E.H., Glassman, A.B., Albitar, M., Kwari, M.I. & Beran, M. (2002) Prognostic factors and scoring systems in chronic myelomonocytic leukemia: a retrospective analysis of 213 patients. *Blood,* **99,** 840-849.

Palacios, R. & Steinmetz, M. (1985) Il-3-dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. *Cell,* **41,** 727-734.

Panic, B., Perisic, O., Veprintsev, D.B., Williams, R.L. & Munro, S. (2003) Structural basis for Arl1-dependent targeting of homodimeric GRIP domains to the Golgi apparatus. *Mol Cell,* **12,** 863-874.

Pardanani, A. (2008) JAK2 inhibitor therapy in myeloproliferative disorders: rationale, preclinical studies and ongoing clinical trials. *Leukemia*, **22,** 23-30.

Pardanani, A., Ketterling, R.P., Brockman, S.R., Flynn, H.C., Paternoster, S.F., Shearer, B.M., Reeder, T.L., Li, C.Y., Cross, N.C., Cools, J., Gilliland, D.G., Dewald, G.W. & Tefferi, A. (2003) CHIC2 deletion, a surrogate for FIP1L1-PDGFRA fusion, occurs in systemic mastocytosis associated with eosinophilia and predicts response to Imatinib therapy. *Blood,* **102,** 3093-3096.

Pardanani, A., Lasho, T.L., Finke, C., Hanson, C.A. & Tefferi, A. (2007) Prevalence and clinicopathologic correlates of JAK2 exon 12 mutations in JAK2V617F-negative polycythemia vera. *Leukemia,* **21,** 1960-1963.

Park, M., Dean, M., Cooper, C.S., Schmidt, M., O'Brien, S.J., Blair, D.G. & Vande Woude, G.F. (1986) Mechanism of met oncogene activation. *Cell,* **45,** 895-904.

Patel, N., Goff, L.K., Clark, T., Ford, A.M., Foot, N., Lillington, D., Hing, S., Pritchard-Jones, K., Jones, L.K. & Saha, V. (2003) Expression profile of wild-type ETV6 in childhood acute leukaemia. *Br J Haematol,* **122,** 94-98.

Patton, J.G., Porro, E.B., Galceran, J., Tempst, P. & Nadal-Ginard, B. (1993) Cloning and characterization of PSF, a novel pre-mRNA splicing factor. *Genes Dev,* **7,** 393-406.

Pedersen, B. & Kerndrup, G. (1986) Specific minor chromosome deletions consistently occurring in myelodysplastic syndromes. *Cancer Genet Cytogenet,* **23,** 61-75.

Peeters, P., Raynaud, S.D., Cools, J., Wlodarska, I., Grosgeorge, J., Philip, P., Monpoux, F., Van, R.L., Baens, M., Van den, B.H. & Marynen, P. (1997) Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia. *Blood,* **90,** 2535-2540.

Peschard, P., Fournier, T.M., Lamorte, L., Naujokas, M.A., Band, H., Langdon, W.Y. & Park, M. (2001) Mutation of the c-Cbl TKB domain binding site on the Met receptor tyrosine kinase converts it into a transforming protein. *Mol.Cell,* **8,** 995-1004.

Peschard, P., Ishiyama, N., Lin, T., Lipkowitz, S. & Park, M. (2004) A conserved DpYR motif in the juxtamembrane domain of the Met receptor family forms an

atypical c-Cbl/Cbl-b tyrosine kinase binding domain binding site required for suppression of oncogenic activation. *J Biol Chem,* **279,** 29565-29571.

Peschard, P. & Park, M. (2003) Escape from Cbl-mediated downregulation: a recurrent theme for oncogenic deregulation of receptor tyrosine kinases. *Cancer Cell,* **3,** 519-523.

Peters, D.G., Hoover, R.R., Gerlach, M.J., Koh, E.Y., Zhang, H., Choe, K., Kirschmeier, P., Bishop, W.R. & Daley, G.Q. (2001) Activity of the farnesyl protein transferase inhibitor SCH66336 against BCR/ABL-induced murine leukemia and primary cells from patients with chronic myeloid leukemia. *Blood.,* **97,** 1404-1412.

Petti, L.M., Reddy, V., Smith, S.O. & DiMaio, D. (1997) Identification of amino acids in the transmembrane and juxtamembrane domains of the platelet-derived growth factor receptor required for productive interaction with the bovine papillomavirus E5 protein. *J Virol,* **71,** 7318-7327.

Pfeifer, D., Pantic, M., Skatulla, I., Rawluk, J., Kreutz, C., Martens, U.M., Fisch, P., Timmer, J. & Veelken, H. (2007) Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. *Blood,* **109,** 1202-1210.

Pietra, D., Li, S., Brisci, A., Passamonti, F., Rumi, E., Theocharides, A., Ferrari, M., Gisslinger, H., Kralovics, R., Cremonesi, L., Skoda, R. & Cazzola, M. (2008) Somatic mutations of JAK2 exon 12 in patients with JAK2 (V617F)-negative myeloproliferative disorders. *Blood,* **111,** 1686-1689.

Popovici, C., Zhang, B., Gregoire, M.J., Jonveaux, P., Lafage-Pochitaloff, M., Birnbaum, D. & Pebusque, M.J. (1999) The t(6;8)(q27;p11) translocation in a stem cell myeloproliferative disorder fuses a novel gene, FOP, to fibroblast growth factor receptor 1. *Blood,* **93,** 1381-1389.

Preker, P.J., Lingner, J., Minvielle-Sebastia, L. & Keller, W. (1995) The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell,* **81,** 379-389.

Puil, L., Liu, J., Gish, G., Mbamalu, G., Bowtell, D., Pelicci, P.G., Arlinghaus, R. & Pawson, T. (1994) Bcr-Abl oncoproteins bind directly to activators of the Ras signalling pathway. *EMBO J.,* **13,** 764-773.

Puputti, M., Tynninen, O., Sihto, H., Blom, T., Maenpaa, H., Isola, J., Paetau, A., Joensuu, H. & Nupponen, N.N. (2006) Amplification of KIT, PDGFRA, VEGFR2, and EGFR in gliomas. *Mol Cancer Res,* **4,** 927-934.

Quackenbush, R.C., Reuther, G.W., Miller, J.P., Courtney, K.D., Pear, W.S. & Pendergast, A.M. (2000) Analysis of the biologic properties of p230 Bcr-Abl reveals unique and overlapping properties with the oncogenic p185 and p210 Bcr-Abl tyrosine kinases. *Blood,* **95,** 2913-2921.

Rabbee, N. & Speed, T.P. (2006) A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics,* **22,** 7-12.

Rabes, H.M., Demidchik, E.P., Sidorow, J.D., Lengfelder, E., Beimfohr, C., Hoelzel, D. & Klugbauer, S. (2000) Pattern of radiation-induced RET and NTRK1 rearrangements in 191 post-chernobyl papillary thyroid carcinomas: biological, phenotypic, and clinical implications. *Clin.Cancer Res.,* **6,** 1093-1103.

Raghavan, M., Lillington, D.M., Skoulakis, S., Debernardi, S., Chaplin, T., Foot, N.J., Lister, T.A. & Young, B.D. (2005) Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. *Cancer Res.,* **65,** 375-378.

Rao, M., Rothwell, S.W. & Alving, C.R. (2003) Trafficking of liposomal antigens to the trans-Golgi complex in macrophages. *Methods Enzymol,* **373,** 16-33.

Rappsilber, J., Ryder, U., Lamond, A.I. & Mann, M. (2002) Large-scale proteomic analysis of the human spliceosome. *Genome Res,* **12,** 1231-1245.

Raskind, W.H. & Fialkow, P.J. (1987) The use of cell markers in the study of human hematopoietic neoplasia. *Adv.Cancer Res.,* **49,** 127-167.

Raskind, W.H., Steinmann, L. & Najfeld, V. (1998) Clonal development of myeloproliferative disorders: clues to hematopoietic differentiation and multistep pathogenesis of cancer. *Leukemia,* **12,** 108-116.

Rathinam, C., Thien, C.B., Langdon, W.Y., Gu, H. & Flavell, R.A. (2008) The E3 ubiquitin ligase c-Cbl restricts development and functions of hematopoietic stem cells. *Genes Dev,* **22,** 992-997.

Reese MG, E.F., Kulp D, Haussler D (1997) Improved splice site detection in Genie. *J Comput Biol,* **4,** 311-323.

Reiter, A., Sohal, J., Kulkarni, S., Chase, A., Macdonald, D.H., Aguiar, R.C., Goncalves, C., Hernandez, J.M., Jennings, B.A., Goldman, J.M. & Cross, N.C. (1998) Consistent fusion of ZNF198 to the fibroblast growth factor receptor-1 in the t(8;13)(p11;q12) myeloproliferative syndrome. *Blood,* **92,** 1735-1742.

Reiter, A., Walz, C. & Cross, N.C. (2007) Tyrosine kinases as therapeutic targets in BCR-ABL negative chronic myeloproliferative disorders. *Curr Drug Targets,* **8,** 205-216.

Reiter, A., Walz, C., Watmore, A., Schoch, C., Blau, I., Schlegelberger, B., Berger, U., Telford, N., Aruliah, S., Yin, J.A., Vanstraelen, D., Barker, H.F., Taylor, P.C., O'Driscoll, A., Benedetti, F., Rudolph, C., Kolb, H.J., Hochhaus, A., Hehlmann, R., Chase, A. & Cross, N.C. (2005) The t(8;9)(p22;p24) is a recurrent abnormality in chronic and acute leukemia that fuses PCM1 to JAK2. *Cancer Res.,* **65,** 2662-2667.

Rieder, C.L., Faruki, S. & Khodjakov, A. (2001) The centrosome in vertebrates: more than a microtubule-organizing center. *Trends Cell Biol,* **11,** 413-419.

Robinson, D.R., Wu, Y.M. & Lin, S.F. (2000) The protein tyrosine kinase family of the human genome. *Oncogene,* **19,** 5548-5557.

Roche-Lestienne, C., Lepers, S., Soenen-Cornu, V., Kahn, J.E., Lai, J.L., Hachulla, E., Drupt, F., Demarty, A.L., Roumier, A.S., Gardembas, M., Dib, M., Philippe, N., Cambier, N., Barete, S., Libersa, C., Bletry, O., Hatron, P.Y., Quesnel, B., Rose, C., Maloum, K., Blanchet, O., Fenaux, P., Prin, L. & Preudhomme, C. (2005) Molecular characterization of the idiopathic hypereosinophilic syndrome (HES) in 35 French patients with normal conventional cytogenetics. *Leukemia,* **19,** 792-798.

Ronaghi, M. (2001) Pyrosequencing sheds light on DNA sequencing. *Genome Res,* **11,** 3-11.

Rosati, R., La Starza, R., Luciano, L., Gorello, P., Matteucci, C., Pierini, V., Romoli, S., Crescenzi, B., Rotoli, B., Martelli, M.F., Pane, F. & Mecucci, C. (2006) TPM3/PDGFRB fusion transcript and its reciprocal in chronic eosinophilic leukemia. *Leukemia,* **20,** 1623-1624.

Ross, T.S., Bernard, O.A., Berger, R. & Gilliland, D.G. (1998) Fusion of Huntingtin interacting protein 1 to platelet-derived growth factor beta receptor (PDGFbetaR) in chronic myelomonocytic leukemia with t(5;7)(q33;q11.2). *Blood,* **91,** 4419-4426.

Rowley, J.D. (1973) Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature,* **243,** 290-293.

Rowley, J.D. (1976) The relationship of chromosomal abnormalities to neoplasia. *Adv Pathobiol,* **4,** 67-73.

Rowley, J.D. (1999) The role of chromosome translocations in leukemogenesis. *Semin Hematol,* **36,** 59-72.

Rucker, F.G., Bullinger, L., Schwaenen, C., Lipka, D.B., Wessendorf, S., Frohling, S., Bentz, M., Miller, S., Scholl, C., Schlenk, R.F., Radlwimmer, B., Kestler, H.A., Pollack, J.R., Lichter, P., Dohner, K. & Dohner, H. (2006) Disclosure of candidate genes in acute myeloid leukemia with complex karyotypes using microarray-based molecular characterization. *J Clin Oncol,* **24,** 3887-3894.

Ruegsegger, U., Blank, D. & Keller, W. (1998) Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol Cell,* **1,** 243-253.

Rumi, E., Passamonti, F., Picone, C., Della Porta, M.G., Pascutto, C., Cazzola, M. & Lazzarino, M. (2008) Disease anticipation in familial myeloproliferative neoplasms. *Blood,* **112,** 2587-2588; author reply 2588-2589.

Russell, S.M., Tayebi, N., Nakajima, H., Riedy, M.C., Roberts, J.L., Aman, M.J., Migone, T.S., Noguchi, M., Markert, M.L., Buckley, R.H., O'Shea, J.J. & Leonard, W.J. (1995) Mutation of Jak3 in a patient with SCID: essential role of Jak3 in lymphoid development. *Science,* **270,** 797-800.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R.H., McPherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M.C., Linton, L., Lander, E.S. & Altshuler, D. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature,* **409,** 928-933.

Safley, A.M., Sebastian, S., Collins, T.S., Tirado, C.A., Stenzel, T.T., Gong, J.Z. & Goodman, B.K. (2004) Molecular and cytogenetic characterization of a novel translocation t(4;22) involving the breakpoint cluster region and platelet-derived growth factor receptor-alpha genes in a patient with atypical chronic myeloid leukemia. *Genes Chromosomes.Cancer,* **40,** 44-50.

Saito H, B., A, Ginsburg, M, Minato, K, Ceresi, E, Yamada, K, Machover, D, Breard, J, Mathe, G (1985) Establishment and characterization of a new human eosinophilic leukemia cell

line. *Blood,* **66,** 1233-1240.

Saito, Y., Haendeler, J., Hojo, Y., Yamamoto, K. & Berk, B.C. (2001) Receptor heterodimerization: essential mechanism for platelet-derived growth factor-induced epidermal growth factor receptor transactivation. *Mol.Cell Biol.,* **21,** 6387-6394.

Samuels, Y. & Velculescu, V.E. (2004) Oncogenic mutations of PIK3CA in human cancers. *Cell Cycle,* **3,** 1221-1224.

Sargiacomo, M., Sudol, M., Tang, Z. & Lisanti, M.P. (1993) Signal transducing molecules and glycosyl-phosphatidylinositol-linked proteins form a caveolin-rich insoluble complex in MDCK cells. *J Cell Biol,* **122,** 789-807.

Sargin, B., Choudhary, C., Crosetto, N., Schmidt, M.H., Grundler, R., Rensinghoff, M., Thiessen, C., Tickenbrock, L., Schwable, J., Brandts, C., August, B., Koschmieder, S., Bandi, S.R., Duyster, J., Berdel, W.E., Muller-Tidow, C., Dikic, I. & Serve, H. (2007) Flt3-dependent transformation by inactivating c-Cbl mutations in AML. *Blood,* **110,** 1004-1012.

Sawyers, C.L. & Denny, C.T. (1994) Chronic myelomonocytic leukemia: Tel-a-kinase what Ets all about. *Cell,* **77,** 171-173.

Schichman, S.A., Caligiuri, M.A., Gu, Y., Strout, M.P., Canaani, E., Bloomfield, C.D. & Croce, C.M. (1994) ALL-1 partial duplication in acute leukemia. *Proc Natl Acad Sci U S A,* **91,** 6236-6239.

Schlessinger, J. (2000) Cell signaling by receptor tyrosine kinases. *Cell,* **103,** 211-225.

Schmidt, L., Duh, F.M., Chen, F., Kishida, T., Glenn, G., Choyke, P., Scherer, S.W., Zhuang, Z., Lubensky, I., Dean, M., Allikmets, R., Chidambaram, A., Bergerheim, U.R., Feltis, J.T., Casadevall, C., Zamarron, A., Bernues, M., Richard, S., Lips, C.J., Walther, M.M., Tsui, L.C., Geil, L., Orcutt, M.L., Stackhouse, T., Lipan, J., Slife, L., Brauch, H., Decker, J., Niehans, G., Hughson, M.D., Moch, H., Storkel, S., Lerman, M.I., Linehan, W.M. & Zbar, B. (1997) Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. *Nat.Genet.,* **16,** 68-73.

Schmidt, M.H. & Dikic, I. (2005) The Cbl interactome and its functions. *Nat Rev Mol Cell Biol,* **6,** 907-918.

Schnittger, S., Wormann, B., Hiddemann, W. & Griesinger, F. (1998) Partial tandem duplications of the MLL gene are detectable in peripheral blood and bone marrow of nearly all healthy donors. *Blood,* **92,** 1728-1734.

Schouten, J.P., McElgunn, C.J., Waaijer, R., Zwijnenburg, D., Diepvens, F. & Pals, G. (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.,* **30,** e57.

Schuster, S.C. (2008) Next-generation sequencing transforms today's biology. *Nat Methods,* **5,** 16-18.

Schwaller, J., Anastasiadou, E., Cain, D., Kutok, J., Wojiski, S., Williams, I.R., LaStarza, R., Crescenzi, B., Sternberg, D.W., Andreasson, P., Schiavo, R., Siena, S.,

Mecucci, C. & Gilliland, D.G. (2001) H4(D10S170), a gene frequently rearranged in papillary thyroid carcinoma, is fused to the platelet-derived growth factor receptor beta gene in atypical chronic myeloid leukemia with t(5;10)(q33;q22). *Blood,* **97,** 3910-3918.

Score, J., Curtis, C., Waghorn, K., Stalder, M., Jotterand, M., Grand, F.H. & Cross, N.C. (2006) Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia. *Leukemia,* **20,** 827-832.

Score, J., Walz, C., Jovanovic, J.V., Jones, A.V., Waghorn, K., Hidalgo-Curtis, C., Lin, F., Grimwade, D., Grand, F., Reiter, A. & Cross, N.C. (2009) Detection and molecular monitoring of FIP1L1-PDGFRA-positive disease by analysis of patient-specific genomic DNA fusion junctions. *Leukemia,* **23,** 332-339.

Scott, L.M., Beer, P.A., Bench, A.J., Erber, W.N. & Green, A.R. (2007a) Prevalance of JAK2 V617F and exon 12 mutations in polycythaemia vera. *Br J Haematol,* **139,** 511-512.

Scott, L.M., Campbell, P.J., Baxter, E.J., Todd, T., Stephens, P., Edkins, S., Wooster, R., Stratton, M.R., Futreal, P.A. & Green, A.R. (2005) The V617F JAK2 mutation is uncommon in cancers and in myeloid malignancies other than the classic myeloproliferative disorders. *Blood,* **106,** 2920-2921.

Scott, L.M., Tong, W., Levine, R.L., Scott, M.A., Beer, P.A., Stratton, M.R., Futreal, P.A., Erber, W.N., McMullin, M.F., Harrison, C.N., Warren, A.J., Gilliland, D.G., Lodish, H.F. & Green, A.R. (2007b) JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N Engl J Med,* **356,** 459-468.

Seo, M.S., Kwak, N., Ozaki, H., Yamada, H., Okamoto, N., Yamada, E., Fabbro, D., Hofmann, F., Wood, J.M. & Campochiaro, P.A. (1999) Dramatic inhibition of retinal

and choroidal neovascularization by oral administration of a kinase inhibitor. *Am.J.Pathol.*, **154,** 1743-1753.

Shah, N.P., Lee, F.Y., Luo, R., Jiang, Y., Donker, M. & Akin, C. (2006) Dasatinib (BMS-354825) inhibits KITD816V, an imatinib-resistant activating mutation that triggers neoplastic growth in most patients with systemic mastocytosis. *Blood,* **108,** 286-291.

Shah, N.P., Tran, C., Lee, F.Y., Chen, P., Norris, D. & Sawyers, C.L. (2004) Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science,* **305,** 399-401.

Shigekawa, K. & Dower, W.J. (1988) Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells. *Biotechniques,* **6,** 742-751.

Shih, L.Y., Huang, C.F., Wu, J.H., Lin, T.L., Dunn, P., Wang, P.N., Kuo, M.C., Lai, C.L. & Hsu, H.C. (2002) Internal tandem duplication of FLT3 in relapsed acute myeloid leukemia: a comparative analysis of bone marrow samples from 108 adult patients at diagnosis and relapse. *Blood,* **100,** 2387-2392.

Shtiegman, K. & Yarden, Y. (2003) The role of ubiquitylation in signaling by growth factors: implications to cancer. *Semin Cancer Biol,* **13,** 29-40.

Shtivelman, E., Lifshitz, B., Gale, R.P. & Canaani, E. (1985) Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature,* **315,** 550-554.

Side, L.E., Emanuel, P.D., Taylor, B., Franklin, J., Thompson, P., Castleberry, R.P. & Shannon, K.M. (1998) Mutations of the NF1 gene in children with juvenile myelomonocytic leukemia without clinical evidence of neurofibromatosis, type 1. *Blood,* **92,** 267-272.

Sigismund, S., Woelk, T., Puri, C., Maspero, E., Tacchetti, C., Transidico, P., Di Fiore, P.P. & Polo, S. (2005) Clathrin-independent endocytosis of ubiquitinated cargos. *Proc Natl Acad Sci U S A,* **102,** 2760-2765.

Simon-Sanchez, J., Scholz, S., Fung, H.C., Matarin, M., Hernandez, D., Gibbs, J.R., Britton, A., de Vrieze, F.W., Peckham, E., Gwinn-Hardy, K., Crawley, A., Keen, J.C., Nash, J., Borgaonkar, D., Hardy, J. & Singleton, A. (2007) Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Hum Mol Genet,* **16,** 1-14.

Simpson, P.J., Monie, T.P., Szendroi, A., Davydova, N., Tyzack, J.K., Conte, M.R., Read, C.M., Cary, P.D., Svergun, D.I., Konarev, P.V., Curry, S. & Matthews, S. (2004) Structure and RNA interactions of the N-terminal RRM domains of PTB. *Structure,* **12,** 1631-1643.

Sinclair, P.B., Green, A.R., Grace, C. & Nacheva, E.P. (1997) Improved sensitivity of BCR-ABL detection: a triple-probe three-color fluorescence in situ hybridization system. *Blood,* **90,** 1395-1402.

Siu, L.L., Chan, V., Chan, J.K., Wong, K.F., Liang, R. & Kwong, Y.L. (2000) Consistent patterns of allelic loss in natural killer cell lymphoma. *Am J Pathol,* **157,** 1803-1809.

So, C.W., Caldas, C., Liu, M.M., Chen, S.J., Huang, Q.H., Gu, L.J., Sham, M.H., Wiedemann, L.M. & Chan, L.C. (1997) EEN encodes for a member of a new family of proteins containing an Src homology 3 domain and is the third gene located on chromosome 19p13 that fuses to MLL in human leukemia. *Proc.Natl.Acad.Sci.U.S.A,* **94,** 2563-2568.

Sohal, J., Chase, A., Mould, S., Corcoran, M., Oscier, D., Iqbal, S., Parker, S., Welborn, J., Harris, R.I., Martinelli, G., Montefusco, V., Sinclair, P., Wilkins, B.S., van den, B.H., Vanstraelen, D., Goldman, J.M. & Cross, N.C. (2001) Identification of

four new translocations involving FGFR1 in myeloid disorders. *Genes Chromosomes.Cancer,* **32,** 155-163.

Sokal, J.E., Gomez, G.A., Baccarani, M., Tura, S., Clarkson, B.D., Cervantes, F., Rozman, C., Carbonell, F., Anger, B. & Heimpel, H. (1988) Prognostic significance of additional cytogenetic abnormalities at diagnosis of Philadelphia chromosome-positive chronic granulocytic leukemia. *Blood,* **72,** 294-298.

Solit, D.B., Basso, A.D., Olshen, A.B., Scher, H.I. & Rosen, N. (2003) Inhibition of heat shock protein 90 function down-regulates Akt kinase and sensitizes tumors to Taxol. *Cancer Res,* **63,** 2139-2144.

Song, J.H., Kim, C.J., Cho, Y.G., Kwak, H.J., Nam, S.W., Yoo, N.J., Lee, J.Y. & Park, W.S. (2007) Genetic and epigenetic analysis of the EPHB2 gene in gastric cancers. *APMIS,* **115,** 164-168.

Spiekermann, K., Pau, M., Schwab, R., Schmieja, K., Franzrahe, S. & Hiddemann, W. (2002) Constitutive activation of STAT3 and STAT5 is induced by leukemic fusion proteins with protein tyrosine kinase activity and is sufficient for transformation of hematopoietic precursor cells. *Exp Hematol,* **30,** 262-271.

Steer, E.J. & Cross, N.C. (2002) Myeloproliferative disorders with translocations of chromosome 5q31-35: role of the platelet-derived growth factor receptor Beta. *Acta Haematol.,* **107,** 113-122.

Steinberg, M. (2007) Dasatinib: a tyrosine kinase inhibitor for the treatment of chronic myelogenous leukemia and philadelphia chromosome-positive acute lymphoblastic leukemia. *Clin Ther,* **29,** 2289-2308.

Stone, R.M., DeAngelo, D.J., Klimek, V., Galinsky, I., Estey, E., Nimer, S.D., Grandin, W., Lebwohl, D., Wang, Y., Cohen, P., Fox, E.A., Neuberg, D., Clark, J., Gilliland, D.G. & Griffin, J.D. (2005) Patients with acute myeloid leukemia and an

activating mutation in FLT3 respond to a small-molecule FLT3 tyrosine kinase inhibitor, PKC412. *Blood,* **105,** 54-60.

Stover, E.H., Chen, J., Folens, C., Lee, B.H., Mentens, N., Marynen, P., Williams, I.R., Gilliland, D.G. & Cools, J. (2006) Activation of FIP1L1-PDGFRalpha requires disruption of the juxtamembrane domain of PDGFRalpha and is FIP1L1-independent. *Proc Natl Acad Sci U S A,* **103,** 8078-8083.

Stover, E.H., Chen, J., Lee, B.H., Cools, J., McDowell, E., Adelsperger, J., Cullen, D., Coburn, A., Moore, S.A., Okabe, R., Fabbro, D., Manley, P.W., Griffin, J.D. & Gilliland, D.G. (2005) The small molecule tyrosine kinase inhibitor AMN107 inhibits TEL-PDGFR{beta} and FIP1L1-PDGFR{alpha} in vitro and in vivo. *Blood.*

Taagepera, S., McDonald, D., Loeb, J.E., Whitaker, L.L., McElroy, A.K., Wang, J.Y. & Hope, T.J. (1998) Nuclear-cytoplasmic shuttling of C-ABL tyrosine kinase. *Proc.Natl.Acad.Sci.U.S.A,* **95,** 7457-7462.

Takeuchi, S., Tsukasaki, K., Bartram, C.R., Seriu, T., Zimmermann, M., Schrappe, M., Takeuchi, N., Park, S., Taguchi, H. & Koeffler, H.P. (2003) Long-term study of the clinical significance of loss of heterozygosity in childhood acute lymphoblastic leukemia. *Leukemia,* **17,** 149-154.

Tanaka, Y., Kanai, Y., Okada, Y., Nonaka, S., Takeda, S., Harada, A. & Hirokawa, N. (1998) Targeted disruption of mouse conventional kinesin heavy chain, kif5B, results in abnormal perinuclear clustering of mitochondria. *Cell,* **93,** 1147-1158.

Teh, M.T., Blaydon, D., Chaplin, T., Foot, N.J., Skoulakis, S., Raghavan, M., Harwood, C.A., Proby, C.M., Philpott, M.P., Young, B.D. & Kelsell, D.P. (2005) Genomewide single nucleotide polymorphism microarray mapping in basal cell carcinomas unveils uniparental disomy as a key somatic event. *Cancer Res,* **65,** 8597-8603.

Theocharides, A., Boissinot, M., Girodon, F., Garand, R., Teo, S.S., Lippert, E., Talmant, P., Tichelli, A., Hermouet, S. & Skoda, R.C. (2007) Leukemic blasts in transformed JAK2-V617F-positive myeloproliferative disorders are frequently negative for the JAK2-V617F mutation. *Blood,* **110,** 375-379.

Thien, C.B., Blystad, F.D., Zhan, Y., Lew, A.M., Voigt, V., Andoniou, C.E. & Langdon, W.Y. (2005) Loss of c-Cbl RING finger function results in high-intensity TCR signaling and thymic deletion. *EMBO J,* **24,** 3807-3819.

Thien, C.B., Walker, F. & Langdon, W.Y. (2001) RING finger mutations that abolish c-Cbl-directed polyubiquitination and downregulation of the EGF receptor are insufficient for cell transformation. *Mol Cell,* **7,** 355-365.

Thomadaki, H., Tsiapalis, C.M. & Scorilas, A. (2008) The effect of the polyadenylation inhibitor cordycepin on human Molt-4 and Daudi leukaemia and lymphoma cell lines. *Cancer Chemother Pharmacol,* **61,** 703-711.

Tokita, K., Maki, K., Tadokoro, J., Nakamura, Y., Arai, Y., Sasaki, K., Eguchi-Ishimae, M., Eguchi, M. & Mitani, K. (2007) Chronic idiopathic myelofibrosis expressing a novel type of TEL-PDGFRB chimaera responded to imatinib mesylate therapy. *Leukemia,* **21,** 190-192.

Tomlinson, I.P., Gammack, A.J., Stickland, J.E., Mann, G.J., MacKie, R.M., Kefford, R.F. & McGee, J.O. (1993) Loss of heterozygosity in malignant melanoma at loci on chromosome 11 and 17 implicated in the pathogenesis of other cancers. *Genes Chromosomes Cancer,* **7,** 169-172.

Torring, N., Borre, M., Sorensen, K.D., Andersen, C.L., Wiuf, C. & Orntoft, T.F. (2007) Genome-wide analysis of allelic imbalance in prostate cancer using the Affymetrix 50K SNP mapping array. *Br J Cancer,* **96,** 499-506.

Tosi, S., Harbott, J., Haas, O.A., Douglas, A., Hughes, D.M., Ross, F.M., Biondi, A., Scherer, S.W. & Kearney, L. (1996) Classification of deletions and identification of cryptic translocations involving 7q by fluorescence in situ hybridization (FISH). *Leukemia,* **10,** 644-649.

Tosi, S., Scherer, S.W., Giudici, G., Czepulkowski, B., Biondi, A. & Kearney, L. (1999) Delineation of multiple deleted regions in 7q in myeloid disorders. *Genes Chromosomes Cancer,* **25,** 384-392.

Toyama, K., Ohyashiki, K., Yoshida, Y., Abe, T., Asano, S., Hirai, H., Hirashima, K., Hotta, T., Kuramoto, A. & Kuriya, S. (1993) Clinical implications of chromosomal abnormalities in 401 patients with myelodysplastic syndromes: a multicentric study in Japan. *Leukemia,* **7,** 499-508.

Trempat, P., Villalva, C., Laurent, G., Armstrong, F., Delsol, G., Dastugue, N. & Brousset, P. (2003) Chronic myeloproliferative disorders with rearrangement of the platelet-derived growth factor alpha receptor: a new clinical target for STI571/Glivec. *Oncogene,* **22,** 5702-5706.

van Dongen, J.J., Langerak, A.W., Bruggemann, M., Evans, P.A., Hummel, M., Lavender, F.L., Delabesse, E., Davi, F., Schuuring, E., Garcia-Sanz, R., van Krieken, J.H., Droese, J., Gonzalez, D., Bastard, C., White, H.E., Spaargaren, M., Gonzalez, M., Parreira, A., Smith, J.L., Morgan, G.J., Kneba, M. & Macintyre, E.A. (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia,* **17,** 2257-2317.

Vandenberghe, P., Wlodarska, I., Michaux, L., Zachee, P., Boogaerts, M., Vanstraelen, D., Herregods, M.C., Van Hoof, A., Selleslag, D., Roufosse, F., Maerevoet, M., Verhoef, G., Cools, J., Gilliland, D.G., Hagemeijer, A. & Marynen,

P. (2004) Clinical and molecular features of FIP1L1-PDFGRA (+) chronic eosinophilic leukemias. *Leukemia,* **18,** 734-742.

Vardiman, J.W., Harris, N.L. & Brunning, R.D. (2002) The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood,* **100,** 2292-2302.

Vardiman, J.W., Thiele, J., Arber, D.A., Brunning, R.D., Borowitz, M.J., Porwit, A., Harris, N.L., Le Beau, M.M., Hellstrom-Lindberg, E., Tefferi, A. & Bloomfield, C.D. (2009) The 2008 revision of the WHO classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*.

Velpeau, A. (1827) Sur le resorption du pusaet sur l'alteration du sang dans les maladies clinique de persecution nememant. Premier observation. *Revue medical francaise et etrange,* **2,** 216-240.

Virchow, R. (1845) *Neue Notizen aus dem Gebriete der Natur- und Heilkunde,* **36,** 151-156.

Vitale, G., Rybin, V., Christoforidis, S., Thornqvist, P., McCaffrey, M., Stenmark, H. & Zerial, M. (1998) Distinct Rab-binding domains mediate the interaction of Rabaptin-5 with GTP-bound Rab4 and Rab5. *EMBO J.,* **17,** 1941-1951.

Vizmanos, J.L., Novo, F.J., Roman, J.P., Baxter, E.J., Lahortiga, I., Larrayoz, M.J., Odero, M.D., Giraldo, P., Calasanz, M.J. & Cross, N.C. (2004) NIN, a gene encoding a CEP110-like centrosomal protein, is fused to PDGFRB in a patient with a t(5;14)(q33;q24) and an imatinib-responsive myeloproliferative disorder. *Cancer Res.,* **64,** 2673-2676.

von, B.N., Schneller, F., Peschel, C. & Duyster, J. (2002) BCR-ABL gene mutations in relation to clinical resistance of Philadelphia-chromosome-positive leukaemia to STI571: a prospective study. *Lancet,* **359,** 487-491.

von Bubnoff, A. (2008) Next-generation sequencing: the race is on. *Cell,* **132,** 721-723.

Vu, H.A., Xinh, P.T., Masuda, M., Motoji, T., Toyoda, A., Sakaki, Y., Tokunaga, K. & Sato, Y. (2006) FLT3 is fused to ETV6 in a myeloproliferative disorder with hypereosinophilia and a t(12;13)(p13;q12) translocation. *Leukemia,* **20,** 1414-1421.

Wada, C., Shionoya, S., Fujino, Y., Tokuhiro, H., Akahoshi, T., Uchida, T. & Ohtani, H. (1994) Genomic instability of microsatellite repeats and its association with the evolution of chronic myelogenous leukemia. *Blood,* **83,** 3449-3456.

Wainscoat, J.S. & Fey, M.F. (1990) Assessment of clonality in human tumors: a review. *Cancer Res.,* **50,** 1355-1360.

Walker, B.A. & Morgan, G.J. (2006) Use of single nucleotide polymorphism-based mapping arrays to detect copy number changes and loss of heterozygosity in multiple myeloma. *Clin Lymphoma Myeloma,* **7,** 186-191.

Walz, C., Chase, A., Schoch, C., Weisser, A., Schlegel, F., Hochhaus, A., Fuchs, R., Schmitt-Graff, A., Hehlmann, R., Cross, N.C. & Reiter, A. (2005) The t(8;17)(p11;q23) in the 8p11 myeloproliferative syndrome fuses MYO18A to FGFR1. *Leukemia,* **19,** 1005-1009.

Walz, C., Curtis, C., Schnittger, S., Schultheis, B., Metzgeroth, G., Schoch, C., Lengfelder, E., Erben, P., Muller, M.C., Haferlach, T., Hochhaus, A., Hehlmann, R., Cross, N.C. & Reiter, A. (2006) Transient response to imatinib in a chronic eosinophilic leukemia associated with ins(9;4)(q33;q12q25) and a CDK5RAP2-PDGFRA fusion gene. *Genes Chromosomes Cancer,* **45,** 950-956.

Walz, C., Haferlach, C., Hanel, A., Metzgeroth, G., Erben, P., Gosenca, D., Hochhaus, A., Cross, N.C. & Reiter, A. (2009) Identification of a MYO18A-

PDGFRB fusion gene in an eosinophilia-associated atypical myeloproliferative neoplasm with a t(5;17)(q33-34;q11.2). *Genes Chromosomes Cancer,* **48,** 179-183.

Walz, C., Metzgeroth, G., Haferlach, C., Schmitt-Graeff, A., Fabarius, A., Hagen, V., Prummer, O., Rauh, S., Hehlmann, R., Hochhaus, A., Cross, N.C. & Reiter, A. (2007) Characterization of three new imatinib-responsive fusion genes in chronic myeloproliferative disorders generated by disruption of the platelet-derived growth factor receptor beta gene. *Haematologica,* **92,** 163-169.

Wang, Y., Pennock, S.D., Chen, X., Kazlauskas, A. & Wang, Z. (2004a) Platelet-derived growth factor receptor-mediated signal transduction from endosomes. *J.Biol.Chem.,* **279,** 8038-8046.

Wang, Z.C., Lin, M., Wei, L.J., Li, C., Miron, A., Lodeiro, G., Harris, L., Ramaswamy, S., Tanenbaum, D.M., Meyerson, M., Iglehart, J.D. & Richardson, A. (2004b) Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. *Cancer Res,* **64,** 64-71.

Warmuth, M., Danhauser-Riedl, S. & Hallek, M. (1999) Molecular pathogenesis of chronic myeloid leukemia: implications for new therapeutic strategies. *Ann Hematol,* **78,** 49-64.

Waterfield, M.D., Scrace, G.T., Whittle, N., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C.H., Huang, J.S. & Deuel, T.F. (1983) Platelet-derived growth factor is structurally related to the putative transforming protein p28sis of simian sarcoma virus. *Nature,* **304,** 35-39.

Weisberg, E., Manley, P.W., Breitenstein, W., Bruggen, J., Cowan-Jacob, S.W., Ray, A., Huntly, B., Fabbro, D., Fendrich, G., Hall-Meyers, E., Kung, A.L., Mestan, J., Daley, G.Q., Callahan, L., Catley, L., Cavazza, C., Mohammed, A., Neuberg, D., Wright, R.D., Gilliland, D.G. & Griffin, J.D. (2005) Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. *Cancer Cell,* **7,** 129-141.

Weisser, M., Kern, W., Schoch, C., Hiddemann, W., Haferlach, T. & Schnittger, S. (2005) Risk assessment by monitoring expression levels of partial tandem duplications in the MLL gene in acute myeloid leukemia during therapy. *Haematologica,* **90,** 881-889.

Weller, P.F. & Bubley, G.J. (1994) The idiopathic hypereosinophilic syndrome. *Blood,* **83,** 2759-2779.

Weng, D.S., Li, J.T., Mai, S.J., Pan, Z.Z., Feng, B.J., Feng, Q.S., Huang, L.X., Wang, Q.J., Li, Y.Q., Yu, X.J., Chen, S.P., He, J. & Xia, J.C. (2006) Identification of a new target region on the long arm of chromosome 7 in gastric carcinoma by loss of heterozygosity. *World J Gastroenterol,* **12,** 2437-2440.

Werner, M., Ewig, M., Nasarek, A., Wilkens, L., von, W.R., Tchinda, J. & Nolte, M. (1997) Value of fluorescence in situ hybridization for detecting the bcr/abl gene fusion in interphase cells of routine bone marrow specimens. *Diagn.Mol.Pathol.,* **6,** 282-287.

Westbrook, C.A. & Keinanen, M.J. (1992) Myeloid malignancies and chromosome 5 deletions. *Baillieres Clin.Haematol.,* **5,** 931-942.

Wetzler, M., Talpaz, M., Van Etten, R.A., Hirsh-Ginsberg, C., Beran, M. & Kurzrock, R. (1993) Subcellular localization of Bcr, Abl, and Bcr-Abl proteins in normal and leukemic cells and correlation of expression with myeloid differentiation. *J.Clin.Invest,* **92,** 1925-1939.

Wheeler, D.A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y.J., Makhijani, V., Roth, G.T., Gomes, X., Tartaro, K., Niazi, F., Turcotte, C.L., Irzyk, G.P., Lupski, J.R., Chinault, C., Song, X.Z., Liu, Y., Yuan, Y., Nazareth, L., Qin, X., Muzny, D.M., Margulies, M., Weinstock, G.M., Gibbs, R.A. & Rothberg, J.M. (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature,* **452,** 872-876.

Wilkinson, K., Velloso, E.R., Lopes, L.F., Lee, C., Aster, J.C., Shipp, M.A. & Aguiar, R.C. (2003) Cloning of the t(1;5)(q23;q33) in a myeloproliferative disorder associated with eosinophilia: involvement of PDGFRB and response to imatinib. *Blood,* **102,** 4187-4190.

Williams, D.M., Kim, A.H., Rogers, O., Spivak, J.L. & Moliterno, A.R. (2007) Phenotypic variations and new mutations in JAK2 V617F-negative polycythemia vera, erythrocytosis, and idiopathic myelofibrosis. *Exp Hematol,* **35,** 1641-1646.

Williams, G.T., Smith, C.A., Spooncer, E., Dexter, T.M. & Taylor, D.R. (1990) Haemopoietic colony stimulating factors promote cell survival by suppressing apoptosis. *Nature,* **343,** 76-79.

Willman, C.L. (1998) Molecular genetic features of myelodysplastic syndromes (MDS). *Leukemia,* **12 Suppl 1,** S2-S6.

Wlodarska, I., Mecucci, C., Marynen, P., Guo, C., Franckx, D., La, S.R., Aventin, A., Bosly, A., Martelli, M.F. & Cassiman, J.J. (1995) TEL gene is involved in myelodysplastic syndromes with either the typical t(5;12)(q33;p13) translocation or its variant t(10;12)(q24;p13). *Blood,* **85,** 2848-2852.

Wong, T.K. & Neumann, E. (1982) Electric field mediated gene transfer. *Biochem Biophys Res Commun,* **107,** 584-587.

Wu, M., Lu, L., Hong, W. & Song, H. (2004) Structural basis for recruitment of GRIP domain golgin-245 by small GTPase Arl1. *Nat Struct Mol Biol,* **11,** 86-94.

Xia, C., Rahman, A., Yang, Z. & Goldstein, L.S. (1998) Chromosomal localization reveals three kinesin heavy chain genes in mouse. *Genomics,* **52,** 209-213.

Xiao, S., Nalabolu, S.R., Aster, J.C., Ma, J., Abruzzo, L., Jaffe, E.S., Stone, R., Weissman, S.M., Hudson, T.J. & Fletcher, J.A. (1998) FGFR1 is fused with a novel

zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. *Nat.Genet.*, **18,** 84-87.

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. & Naoe, T. (2001) Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood,* **97,** 2434-2439.

Yoshino, A., Bieler, B.M., Harper, D.C., Cowan, D.A., Sutterwala, S., Gay, D.M., Cole, N.B., McCaffery, J.M. & Marks, M.S. (2003) A role for GRIP domain proteins and/or their ligands in structure and function of the trans Golgi network. *J Cell Sci,* **116,** 4441-4454.

Yoshino, A., Setty, S.R., Poynton, C., Whiteman, E.L., Saint-Pol, A., Burd, C.G., Johannes, L., Holzbaur, E.L., Koval, M., McCaffery, J.M. & Marks, M.S. (2005) tGolgin-1 (p230, golgin-245) modulates Shiga-toxin transport to the Golgi and Golgi motility towards the microtubule-organizing centre. *J Cell Sci,* **118,** 2279-2293.

Young, J.L., Jr., Percy, C.L., Asire, A.J., Berg, J.W., Cusano, M.M., Gloeckler, L.A., Horm, J.W., Lourie, W.I., Jr., Pollack, E.S. & Shambaugh, E.M. (1981) Cancer incidence and mortality in the United States, 1973-77. *Natl.Cancer Inst.Monogr,* 1-187.

Zatkova, A., Schoch, C., Speleman, F., Poppe, B., Mannhalter, C., Fonatsch, C. & Wimmer, K. (2006) GAB2 is a novel target of 11q amplification in AML/MDS. *Genes Chromosomes Cancer,* **45,** 798-807.

Zenklusen, J.C., Thompson, J.C., Troncoso, P., Kagan, J. & Conti, C.J. (1994) Loss of heterozygosity in human primary prostate carcinomas: a possible tumor suppressor gene at 7q31.1. *Cancer Res,* **54,** 6370-6373.

Zenklusen, J.C., Weitzel, J.N., Ball, H.G. & Conti, C.J. (1995) Allelic loss at 7q31.1 in human primary ovarian carcinomas suggests the existence of a tumor suppressor gene. *Oncogene,* **11,** 359-363.

Zhang, J.G., Goldman, J.M. & Cross, N.C. (1995) Characterization of genomic BCR-ABL breakpoints in chronic myeloid leukaemia by PCR. *Br.J.Haematol.,* **90,** 138-146.

Zhang, T., Sun, H.C., Xu, Y., Zhang, K.Z., Wang, L., Qin, L.X., Wu, W.Z., Liu, Y.K., Ye, S.L. & Tang, Z.Y. (2005) Overexpression of platelet-derived growth factor receptor alpha in endothelial cells of hepatocellular carcinoma associated with high metastatic potential. *Clin Cancer Res*, **11,** 8557-8563.

Zheng, N., Wang, P., Jeffrey, P.D. & Pavletich, N.P. (2000) Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell,* **102,** 533-539.

Zhou, Z., Licklider, L.J., Gygi, S.P. & Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature,* **419,** 182-185.

Zhu, G., Zhai, P., Liu, J., Terzyan, S., Li, G. & Zhang, X.C. (2004) Structural basis of Rab5-Rabaptin5 interaction in endocytosis. *Nat.Struct.Mol.Biol.,* **11,** 975-983.

Ziemin-van der Poel, S., McCabe, N.R., Gill, H.J., Espinosa, R., 3rd, Patel, Y., Harden, A., Rubinelli, P., Smith, S.D., LeBeau, M.M., Rowley, J.D. & et al. (1991) Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc Natl Acad Sci U S A,* **88,** 10735-10739.

APPENDIX

I) Bubble-PCR primers

BUB-T:
5’ AAGGATCCTAGTCTAGCTGTCTGTCGAAGGTAAGGAACGGACGAGCACTGAG

BUB-B:
5’ CTCAGTGCTCGTAGTAATCGTTCGCACGAGAATCGCAAGATCTAGGATCCTT

NVAMP1:
5’ TGCTCGTAGTAATCGTTCGCAC

NVAMP2:
5’ GTTCGCACGAGAATCGCAAGAT

PDAI12R3:
5’ AGGTTACCCCATGGAACTTACCA

PDA112R4:
5’ AAGTTGTGTGCAAGGGAAAAGGG

II) PDGFRB Exons 2 to 23

PDBI1F3	5’ AGA GCC AAA GCA AAG GCC AG 3’
PDBI2R3	5’ AGA GCC CAC TGG AAA TCC TG 3’
PDBI2F	5’ CCC CTG TTT CCT GAT GTC TG 3’
PDBI3R	5’ATG CCT CCA GTT GAC AAG AG 3’
PDBI3F	5’ CTT GCC ACA CAG CAA TTC AG 3’
PDBI4R	5’ CTG AGC ATC AGG CCA GAA AG 3’
PDBI4F	5’ GAC CTA AGC CAA TCT CTC TC 3’
PDBI6R	5’ GAA TTG GGG ATT GGG CTG AG 3’
PDBI6F	5’ TGG CCT CCT TTG GGA TTC AG 3’
PDBI7R2	5’ CCT AGG TTT GTG GCT GAA AG 3’
PDBI7F	5’AGT CCT TCC GAC TCT GAC AG 3’
PDBI8R	5’ATC CAT CTC CTG AGT TCC AG 3’
PDBI8F	5’TAA CTG TCC TGA CCC TCC CG 3’

PDBI9R	5’TCC TAG CCA GCT GGG GAC AC 3’
PDBI9F	5’ TCA GTT TCC CTG TCT GCA AG 3’
PDBI10R	5’ GTA GGG ATT GGG ATC GTC AG 3’
PDBI10F	5’ GAT GCC AAA GAT GGG GTG AG 3’
PDBI11R2	5’ ATG TGC CAG TCT TCA CCC AC 3’
PDBI11F	5’ CCT GAG ACT CCC TCT GAT AA
PDBI12R	5’ ACA TGA GGC CTC TCA GGA CT
PDBI12F	5’ CTG GTA GGC CAG GAG CTA AT
PDBI13R	5’ TGG AGG GCT CCA AGG ACT A
PDBI13F	5’ GGG GCA GAA GAG TCA GAA TA
PDBI14R	5’ CCT TGG TGG TGG GCA CTT T
PDBI14F	5’ CCT TGA AGG GAC GCC TGA
PDBI15R	5’ AAA GAT TCA GTC CCT GGC CT
PDBI15F	5’ CTG CAG GCT CTC CTG TGA T
PDBI16R	5’ AGC CTG TTT GGA TGT GGG GT 3’
PDBI16F	5’ TAG CAG GTG ACC CTC TGC TT 3’
PDBI18R	5’ AGG GGC CAG GGA AGG TA 3’
PDBI18F	5’ CAC TCA GCA CCT GTC CTG A 3’
PDBI19R	5’ GGC TGG AGG AGG AAG CAA 3’
PDBI19F	5’ CCC CTG AGC CTG TGT ATC T 3’
PDBI20R	5’ TCC CAA ATG CAT GAG ACT CCA 3’
PDBI20F	5’ AAG GGA ATA TCC AGA GAC AGG A 3’
PDBI21R	5’ TAA ATG CCA GCC CAT CAC GCA 3’
PDBI21F	5’ CGG TTA GAA GAT CCC TGA AG 3’
PDBI22R2	5’ TGT GCA CAA TTT CCT TGG CC 3’
PDBI22F	5’ ATG AGC CGG AGT GTG TGA AG 3’
PDBE23R	5’ GGG ACA GCT GAT AAG GGC AG 3’

III) PCR and dHPLC conditions

GENE	EXON	FORWARD PRIMER	REVERSE PRIMER	ANNEALING TEMPERATURE	DHPLC TEMPERATURE	TIME-SHIFT (SEC)
PDGFRB	2	PDBI1F3	PDBI2R3	62	63, 64	0, -0.5
	3	PDBI2F	PDBI3R	60	61, 62, 63	-0.5, 0, 0.5
	4	PDB13F	PDBI4R	60	59, 60, 61	-0.5, 0, 0.5
	5+6	PDBI4F	PDBI6R	60	60, 61, 63	-0.5, 0, 1
	7	PDBI6F	PDBI7R2	62	64, 65	0.5, 1
	8	PDBI7F	PDBI8R	60	62, 63, 64	-0.5, 0, 0.5
	9	PDBI8F	PDBI9R	64	63, 64, 65	0, 0, 0.5
	10	PDBI9F	PDBI10R	60	64, 65, 66	0, 0.5, 1
	11	PDBI10F	PDBI11R	62	61, 62, 63	-0.5, 0, 0.5
	12	PDBI11F	PDBI12R	62	62, 63, 64	0, 0.5, 1
	13	PDBI12F	PDBI13R	62	60, 62, 63	-0.5, 0, 0.5
	14	PDBI13F	PDBI14R	64	63, 64	-0.5, 0
	15	PDBI14F	PDBI15R	60	62, 63, 64, 66	-0.5, 0, 0.5, 1
	16	PDBI15F	PDBI16R	60	60, 61, 62, 63	-0.5, 0, 0.5, 1
	17+18	PDBI16F	PDBI18R	62	60, 61, 62	-1, -0.5, 0
	19	PDBI18F	PDBI19R	60	60, 62, 63	-1, 0, 0.5
	20	PDBI19F	PDBI20R	62	60, 63, 64	-1.5, 0, 0.5
	21	PDBI20F	PDBI21R	62	60, 61, 62	-1, -0.5, 0
	22	PDBI21F	PDBI22R	58	61, 62, 63	-0.5, 0, 0.5
	23	PDBI22F	PDBI23R	62	62, 63, 64	0, 0.5, 1

IV) PDGFRB

First step
PDGFRB exon 15.1R 5’ agg tag tcc acc agg tct ccg ta 3’
Nested
PDGFRB exon 14.1R 5’ aag ggc ttg ctt ctc act gct gcg 3’

V) Sequencing

M13F 5’ gta aaa cga cgg cca g 3’
M13R 5’ cag gaa aca gct atg ac 3’

VI) Fusion

GOLGA4 exon 8.1F	5’ act tat cac tca gtt gcg tga tg 3’
PDGFRB exon 12.1R	5’ acc ttc cat cgg atc tcg taa cgt 3’

Reciprocal

GOLGA4 exon 10.1R	5’ ttc ctg ttg aag act gat gcg tt 3’
PDGFRB exon 9.1F	5’ aca tca tct ggt ctg cct gca g 3’

VII) PDGFRB

PDB.I11.1R	5’ gag agc agg cca tga gca aac 3’
PDB.I11.1aR	5’ gag acc ccc agc ctg atg aat 3’
PDB.I11.2R	5’ cat gag gtc cta tgg aac tca tt 3’
PDB.I11.3R	5’ tcc cgc tca ggc ttt agt cac 3’
PDB.I11.4R	5’ ggc agc tct aga cag gca cg 3’
PDB.I11.5R	5’ aag gtg tgc tcc aat tat atg ctt 3’
PDB.I11.6R	5’ cag gca tag tgc aca ggt gta 3’
PDB.I11.7R	5’ tgc cag aat gtg ttc cta tgc at 3’

VIII) GOLGA4

(Note: all primers are designed in intron 10, just labelled wrong)

gGOLGA4.I9.1F	5’ tct taa cta tag agg ctc cta ata a 3’
gGOLGA4.I9.2F	5’ gaa gac tga ctt gaa gca gtt tg 3’
gGOLGA4.I9.3F	5’ cag aac tgt tca agt agg aag tg 3’
gGOLGA4.I9.4F	5’ ctt tgg ctc act tat tct gga tgt 3’
gGOLGA4.I9.5F	5’ tct tat cat ggg tac tta gaa taa g 3’
gGOLGA4.I9.6F	5’ cca tgg atg tga gtg aac ctc a 3’
gGOLGA4.I9.7F	5’ caa tgg agc tcg atc cat ttc ag 3’
gGOLGA4.I9.8F	5’ tca atg gta gag gga agg ctg g 3’
gGOLGA4.I9.9F	5’ gca gaa aat tat atc taa aca gga ac 3’
gGOLGA4.I9.10F	5’ gtg cca agt tac aca cct gtt g 3’
gGOLGA4.I9.11F	5’ tta tac agt aca gaa ata ggg aca t 3’

IX) GOLGA4 RT-PCR

cGOLGA4.exon 10.1F	5’ cag atg act acc cag gga gag 3’

X) GOLGA4 genomic fusion

gGOLGA4.I9.7F	5’ caa tgg agc tcg atc cat ttc ag 3’
PDB.I11.Brkpt.R	5’ aac aga tgt ctg gga att agg ca 3’

XI) Reciprocal PDGFRB

PDB.I11.recip.F	5’ gtg agt ccc cag cct cag tg 3’
PDB.I11.recip.2F	5’ atg cat agg aac aca ttc tgg ca 3’

Reciprocal GOLGA4

gGOLGA4.I9.recip.R	5’ gat tcc tat tct aac ttt act g 3’

XII) RT-PCR Fusion

cGOLGA4 exon 7.1F	5’ tcc agc aaa gag tga agc gtc a 3’
PDGFRB exon 12.1R	5’ acc ttc cat cgg atc tcg taa cgt 3’

XIII) KIAA1449

KIAA1449.2F	5’ cat acc gag tcc atg atg aag 3’
PDGFRB.exon15.2R	5’ tgc tgc agg aag gtg tgt ttg ttg 3’

XIV) Bin2-PDGFRB construct primers

Full-length fusion amplification

Bin2-PDB.Fus.1F	5’ agt tgg cag gat ggc aga ggg 3’
Bin2-PDB.Fus.2F	5’ gtt ggc agg atg gca gag g 3’
Bin2-PDB.Fus.2R	5’ acg gcc cct gca gtt ttc tt 3’

Overlapping primers for sequencing

M13F & M13R	(as previous – section V)
Bin.e5.1F	5’ ttg gga aga cta cga gga 3’
Bin.e9.1F	5’ cta cga ggt gat gag caa 3’
PDB.e15.2F	5’ caa caa aca cac ctt cct gca gca 3’
PDB.e18.1F	5’ att aca tct cca aag gca gca cct 3’
PDB.e15.2R	5’ tgc tgc agg aag gtg tgt ttg ttg 3’

pcDNA3.1(+) vector primers

T7 F	5’ taa tac gac tca cta tag gg 3’
BGH R	5’ tag aag gca cag tcg agg 3’

XV) GOLGA4-PDGFRB construct primers

Fragment 1 construct amplification (GOLGA4 only)

NheI.1F	5’ tac gct agc cca tgt tca aga aac tg 3’
cGOLGA4.e10.3R	5’ cct ctc cct ggg tag tca tct g 3’

Overlapping primers for sequencing fragment 1

M13F & M13R	(as previous – section V)
BC/GOL.3F	5’ tgg aac gaa gct taa gta gct ac 3’
BC/GOL.2F	5’ gca cag aag ctc cag ctc cg 3’
cGOLGA4.e6.2F	5’ tgc atc ttt aga gga gaa aga tca 3’

Fragment 2 construct amplification (GOLGA4 only)

cGOLGA4.e7.1F	5’ tcc agc aaa gag tga agc gtc a 3’
PDGFRB.e15.1R	(as previous – section IV)
cGOLGA4.e8.1F	5’ tta tca ctc agt tgc gtg atg ca 3’

XVI) STRN-PDGFRA fusion primers

Fusion

STRN.Fusion.Ex6.1F	5’ tct gag tta aca gat tct gcc tc 3’
PDGFRA.Fusion.Ex12.2R	5’ caa gca cta gtc cat ctc ttg g 3’

Reciprocal

PDGFRA.recip.Ex11.1F	5’acg gtg gct gct gca gtc ct 3’
STRN.Recip.Int6.1R	5’act ata ttc atc tgc ccc ata gaa 3’

XVII) ETV6-PDGFRA fusion primers

Genomic DNA Fusion

ETV6.Fusion.Int6.4F	5’ ctt cta aga agg cag tgg gta at 3’
PDGFRA.fusion.Ex12.1R	5’ cca gcg aat ttc ata cct cgg 3’

mRNA Fusion

ETV6.Fusion.Ex6.1F:	5’ act gta gac tgc ttt ggg att ac 3’
PDGFRA.Fusion.Ex12.2R	5’ caa gca cta gtc cat ctc ttg g 3’

Reciprocal

PDGFRA.recip.Ex11.1F	5’ acg gtg gct gct gca gtc ct 3’
ETV6.Recip.Ex7.1R	5’ ttg tag tag tgg cgc agg gct 3’

XVIII) STRN-PDGFRA & ETV6-PDGFRA MRD primers

STRN Fusion – SINGLE STEP

STRN.Fusion.Ex6.1F	5’ tct gag tta aca gat tct gcc tc 3’
PDGFRA.Fusion.Ex12.2R	5’ caa gca cta gtc cat ctc ttg g 3’

STRN Fusion – NEST STEP

STRN.Ex.6.2F	5’ tct gag tta aca gat tct gcc tc 3’
PDGFRA.STRN.Ex12.1R	5’ cac tag tcc atc tct tgg aaa ct 3’

ETV6 Fusion – SINGLE STEP

ETV6.Fusion.Ex6.1F:	5’ act gta gac tgc ttt ggg att ac 3’
PDGFRA.Fusion.Ex12.2R	5’ caa gca cta gtc cat ctc ttg g 3’

ETV6 Fusion – NEST STEP

ETV6.Ex6.2F	5’att ccg gat agt gga tcc caa c 3’
PDGFRA.STRN.Ex12.1R	5’ cac tag tcc atc tct tgg aaa ct 3’

XIX) MLPA cDNA oligonucleotides

Labelled primer
Target sequence X
Target sequence Y
Stuffer Sequence
Other Primer

1. FGFR1 intracellular (exons 12/13) (95bp)

A: GGG TTC CCT AAG GGT TGG AGC ATG GAG TAT CTG GCC TCC AAG AAG

B: TGC ATA CAC CGA GAC CTG GCA GCC CTG TCT AGA TTG GAT CTT GCT GGC AC

2. FGFR1 extracellular (exons 3/4) (100bp)

A: GGG TTC CCT AAG GGT TGG AGA GCC ATA ACA CCA AAC CAA ACC GTA TGC

B: CCG TAG CTC CAT ATT GGA CAT CCC GAA TCT CTA GAT TGG ATC TTG CTG GCA C

3. PDGFRB intracellular (exons 20/21) (105bp)

A: GGG TTC CCT AAG GGT TGG AGT ACT ACG CCT GCC CAT GCC TCC GAC GAG AT

B: CTA TGA GAT CAT GCA GAA GTG CTG GCA CTA TGT CTA GAT TGG ATC TTG CTG GCA C

4. PDGFRB extracellular (exons 4/5) (110bp)

A: GGG TTC CCT AAG GGT TGG ACT GAC GTA CGC CTA CTA TGT CTA CAG ACT CCA GG

B: TGT CAT CCA TCA ACG TCT CTG TGA CCT AGC TGA GTC TAG ATT GGA TCT TGC TGG CAC

5. PDGFRA intracellular (exons 20/21) (115bp)

A: GGG TTC CCT AAG GGT TGG AGT ATG ACT GTA CCG CCT GAC CAC GCT ACC AGT GAA GT

B: CTA CGA GAT CAT GGT GAA ATG CTG CTG ACA TAG CAG TCT AGA TTG GAT CTT GCT GGC AC

6. PDGFRA extracellular (exons 2/3) (120bp)

A: GGG TTC CCT AAG GGT TGG ACA CTC TAG GTA CGT GTG GTC TTA GGC TGT CTT CTC ACA G

B: GGC TGA GCC TAA TCC TCT GCC AGC GTC AGA GAC CGA TAC TCT AGA TTG GAT CTT GCT GGC AC

7. ABL C-terminal domain (exons 8/9) (125bp)

A: GGG TTC CCT AAG GGT TGG AGC TGA ATC AGA CCT ACC AGG TCT ATG AAC TCA TGC GAG CAT

B: GTT GGC AGT GGA ATC CCT CTG ACC CAG TCA ACA GGC TCG ATC TCT AGA TTG GAT CTT GCT GGC AC

8. ABL 1b start (exons 1b/a2) (130bp)

A: GGG TTC CCT AAG GGT TGG AGC TAC CGC TAC AAG TTC TAG GGT CCA CAC TGC AAT GTT TTT GTG

B: GAA CAT GAA GCC CTT CAG CGG CCA CTA GAT CCA TTA TCG TAT ACT CTA GAT TGG ATC TTG CTG GCA C

XX) Cryptic ETV6-PDGFRB fusion primers

Fusion

ETV6.Ex6.2F	5'ATT CCG GAT AGT GGA TCC CAA C 3'
PDGFRB.Ex14R	5'- CCC CAA CAG GTT GAC CAC GTT CAG -3'

Genomic Fusion

ETV6.E1026.Int7.1F	5' AGG ATA GGA TAG GAT AGA GTT GAA -3'
ETV6.E1026.Int7.2F	5' AAT GCC AGA CTG CAT GAC ATG -3'
PDGFRB.Exon 12R	5'- CCA CGT AGA TGT ACT CAT GGC -3'

XXI) Primer sequences designed to amplify the kinase domain of the RTKs EPHA1, EPHB6 & Met

EPHA1.Ex12F	5' tgg aag gcg tcg tca caa agc 3'
EPHA1.Ex16R	5' cca tcc tca atg ctc ttc ata ac 3'
EPHB6.Ex11F	5' gat cct gct tat atc aag att gag 3'

EPHB6.Ex14R	5’ tat gct ttc cat gtg caa tga cc 3’
Met.Ex15F	5’ gac aag tgc agt atc ctc tga c 3’
Met.Ex19R	5’ cca tcc act tca ctg gca gct 3’

XXII) EPHA1 genomic DNA primers to screen the whole gene

Primer name	**Primer sequence**
EPHA1.5'UTRF	gacgcctccggttcttgaatttga
EPHA1.i1R	aacatggcccctaggaatg
EPHA1.i1F	tctttggtctttaaccccaagcg
EPHA1.i2R	ctctgttcctccacccacctctc
EPHA1.i2F	gggatgatctggtccctgactct
EPHA1.i3R	ccttctctggcacccaataagga
EPHA1.i3F	agaagcggactttgccattgct
EPHA1.i4R	gtactcctgcaatcctcccgagt
EPHA1.i4F	tggtggtggttctgcctgtaaaa
EPHA1.i5R	ccaaccctgacaaagtctcagca
EPHA1.i5F	gaggacatcgagaattgggtggt
EPHA1.i6R	tatggcttcctgcccatcagttt
EPHA1.i6F	gatgggcaggaagccatagaaaa
EPHA1.i7R	ggcatggacactgaagaccatct
EPHA1.i7F	ggctcagcactgcccacttaat
EPHA1.i8R	aggtgtatgtgtgttggggcaga
EPHA1.i8F	gatgggcaggaagccatagaaaa
EPHA1.i9R	ggcatggacactgaagaccatct
EPHA1.Int8.1F	cctacggactgtgagtcccaaca
EPHA1.Int9.1R	tactaggggatgggagggtcgtt
EPHA1.Int9.1F	tctggggaagattcagccagagt
EPHA1.Int10.1R	gatcccttcccagtggcatctc
EPHA1.Int10.1F	cctgatgttggtactccccaacc
EPHA1.Int11.1R	ctgctggtggctctgtgctg
EPHA1.Int11.1F	gggccacagcctgttttctttc
EPHA1.Int12.1R	agctgagggagaccactcatcgt
EPHA1.Int12.1F	gaaagaggggtgaaaccactgga
EPHA1.Int13.1R	cagctagtcctctgccctcctca
EPHA1.Int13.1F	ccatcccctgactcacccact
EPHA1.Int14.1R	catgttacggggcaatgtgaatg
EPHA1.Int14.1F	ccctgacccatgtgccttcttag

EPHA1.Int15.1R	gggaggaccttgggaggaaataa
EPHA1.Int15.1F	tttctcatgggcctccaggttag
EPHA1.Int16.1R	ctgcctctgggttccctagtcct
EPHA1.Int16.1F	gccaaccagatcagcccatgta
EPHA1.Int17.1R	cccagacagctccagctccttac
EPHA1.Ex17.1F	caccatggagtgtgtgctgga
EPHA1.Ex 18.1R	agggttctcctgaaagtgggc

Primer name	**Primer sequence**
EPHB6.Int5'UTR.F	aggatgcagcagcttcttaacca
EPHB6.Int.1R	gcctactccctgtgtgaccctgt
EPHB6.Int.1F	cctgcctcttctggtcttgtgtg
EPHB6.Int.2R	acttccccctttcactgggtctc
EPHB6.Int.2F	gccactgtgtgcccctcttattt
EPHB6.Int3R	aatcctcacccagggtttccttc
EPHB6.Int.3F	ctaaagcttctcctggcccttcc
EPHB6.Int4R	caattgtgtccaccttggtccag
EPHB6.Ex4F	ccgggagaccttcaccctttact
EPHB6.Int.5R	aagtgcaagacgagtgggctct
EPHB6.Int5F	cgaggatacctgcctgctgtgt
EPHB6.Int6R	acttctggcccctctccaagtct
EPHB6.Int6F	ggctccttccttcctcaatcac
EPHB6.Int.7R	ccaggctggagatcaggagtagg
EPHB6.Int.7F	atcctgctcccagggactcctat
EPHB6.Int.8R	tgaggcctatgagctcctaccag
EPHB6.Int 7.1F	cagggagtgagtggctgttaccc
EPHB6.Int 8.1R	actacccacttgggccttcacct
EPHB6.Int 8.1F	gcacatgaacaagcacctgtgag
EPHB6.Int 9.1R	acttggcaggacttgaggtttgg
EPHB6.Int 9.1F	gtcatccccagtcaccgttttgt
EPHB6.Int 10.1R	gcccacttacctccagaaggttg
EPHB6.Int10.1F	ctgcctttcacaacacgctcataa
EPHB6.Int11.1R	cctgctgtgtttggatgggact
EPHB6.Int11.1F	agggccaaagcagggatcag
EPHB6.Int12.1R	cctcgtctcctggaagctttcc
EPHB6.Int12.1F	ccgtctccgatcactgacctct
EPHB6.Int13.1R	gtgtccccaaggctgtgctct
EPHB6.Int13.1F	gtaaaacccttggcatatctgagcac
EPHB6.Int14.1R	ggtgggaaatcagcagtgtctagg
EPHB6.Int14.1F	agcccaccctcttacccaacaa

EPHB6.Int15.1R	agcagggtcaaggtctcct
EPHB6.Int15.1F	agaggagaccttgaccctgcttg
EPHB6.Int16.1R	cagcttctccgagatccccaata
EPHB6.Int16.1F	tattggggatctcggagaagctg
EPHB6.Ex17.1R	atgtcccttctcggaccagtgtc

Primername	**Primersequence**
MET.5UTR.F	tctccggctgtgctaacttcagac
MET.Int1.1R	gtctcttcagtttccgaagccccta
MET.Int1.1F	tccagttgggaagctttatttctga
MET.Int2.2R	tctgaaatgtcacgaacatgcaaa
MET.Int2.1F	gccattatcctccaggctctgaaa
MET.Int3.1R	ggcttcattgtttggctttcagtc
MET.Int3.1F	gaggggaactgttgggtcttcagt
MET.Int4.1R	aagggccaagtgacactggttgta
MET.Int4.1F	ccgttatgacaggatttgcacaca
MET.Int5.1R	catgtatgccagctgttagagattcc
MET.Int5.1F	tgtcccatctttggatctcctgaa
MET.Int6.1R	tgtgacaggccaagacccttttta
MET,Int6.1F	tgcttgctattcaaagcagtcagc
MET.Int7.1R	tggaggggagataaaaacaaaacaaa
MET.Int7.1F	tttgttttgtttttatctcccctcca
MET.Int8.1R	ttctccacacacacacacaaaaca
MET.Int8.1F	gcttccactcaggaaattcccact
MET.Int9.1R	ttggaaagaagccttcaaaagcaa
MET.Int9.2F	tttgtcctgcctgctttacacaca
MET.Int10.2R	tggttagaggcaaagatgcagagc
MET.Int10.1F	tgctatggatgttgccaagctgta
MET.Int11.1R	tttgtgtcccacatttgctacgtt
MET.Int11.1F	acaacatggcctgtgtttgcagta
MET.Int12.1R	aggaatgcaggctgagttgatgag
MET.Int.12.2F	tcctgaaggcagttatgccatttg
MET.Int13.2R	cttcgggcacttacaagcctatcc
MET.Int13.2F	attgtcgtcgattcttgtgtgctg
MET.Int14.2R	ttcaatgtttgggcttcaacaggt
MET.Int.14.1F	gctcttcctgtttcagtccccatt
MET.Int15.1R	ctgctctgtcagttgctttcacca
MET.Int15.1F	cgcagtgctaaccaagttctttcttt

MET.Int16.1R	tgatttttccacaaggggaaagtg
MET.Int.16.1F	ctgcagtcaaaccctcaggacaag
MET.Int17.1R	gagcaggcctattttgaagggatg
MET.Int17.1F	caggcttgagccattaagaccaaa
MET.Int18.1R	tggattgtggcacagagattctgata
MET.Int.18.1F	aggccagatgaaatacttccttcaga
MET.Int19.1R	tgaagaaaactggaattggtggtg
MET.Int19.1F	tggtcacatctctcacctcatctgtc
MET.Int20.1R	cctttgaaggcatttctgta
MET.Int20.1F	gcctgccttcaaagggtctcttac
MET.3UTR.1R	ccaatatgacaggagtgttgtcacg

XXIII) PIK3CG primer sequences

PIK3CG.Codons.542.5.1F	5’ taa gca tca gcc cac ccc tga 3’
PIK3CG.Codons 542.5.1R	5’ ctt gct gtc ccc att tca ctg a 3’
PIK3CG.Codon 1007.2F	5’ ggc tta tct agc cct tcg tca tca 3’
PIK3CG.Codon.1007.2R	5’ tgc aca gtc gat cct ttg tgg tct g 3’
PIK3CG.Codon.1047.1F	5’ cct act gat cat cct gtt ctc c 3’
PIK3CG.Codon.1047.1R	5’ gca ttt gaa atc caa gtt cga tg 3’

XXIV) c-CBL RING finger domain primers

CBL.Int7.1F	5’ tgt ggt ttc act tta aac cct gga 3’
CBL.Int8.1R	5’ gc cag gcc acc cct tgt atc 3’
CBL.Int8.1F	5’ ggc ctg gct ttt ggg tta gg 3’
CBL.Int9.1R	5’ ca caa tgg att ttg cca gtc tcc 3’

XXV) c-CBL cDNA primers

CBL.Ex6/7.F	5’ ggc tc agg gaa ggc ttc ta 3’
CBL.Ex10/11.R	5’ gcc aga agc agc ctt aga agc a 3’

XXVI) NRAS primers

2	f: 5’gat gtg gct cgc caa tta ac 3’ r: 5’tgc ata act gaa tgt ata ccc aaa 3’
3	f: 5’ttg cat tcc ctg tgg ttt tt 3’ r: 5’cct cat ttc ccc ata aag att c 3’
4	f: 5’cct ccc aaa gtg ctg aga tt 3’ r: 5’gca aac tct tgc aca aat gc 3’
5	f: 5’tgg aat ctt atg tcc aca taa aga 3’ r: 5’tca gaa agg gtg tca tat gga 3’

These primers cover the entire coding sequence of the NRAS gene, however only primers to exons 2 & 3 were used to look at the mutation hotspots at codons 12, 13 and 61.

XXVII) c-CBL intronic primers to amplify exons 8 & 9

CBL.i7F 5’ tgt ggt ttc act tta aac cct gga 3’

CBL.i9R 5’ca caa tgg att ttg cca gtc tcc 3’

List of Publications

Two novel imatinib-responsive PDGFRA fusion genes in chronic eosinophilic leukaemia.

Curtis CE, Grand FH, Musto P, Clark A, Murphy J, Perla G, Minervini MM, Stewart J, Reiter A, Cross NC.

Br J Haematol. 2007 Jul;138(1):77-81.

A novel ETV6-PDGFRB fusion transcript missed by standard screening in a patient with an imatinib responsive chronic myeloproliferative disease.

Curtis CE, Grand FH, Waghorn K, Sahoo TP, George J, Cross NC.

Leukemia. 2007 Aug;21(8):1839-41. Epub 2007 May 17.

The t(1;9)(p34;q34) and t(8;12)(p11;q15) fuse pre-mRNA processing proteins SFPQ (PSF) and CPSF6 to ABL and FGFR1.

Hidalgo-Curtis C, Chase A, Drachenberg M, Roberts MW, Finkelstein JZ, Mould S, Oscier D, Cross NC, Grand FH.

Genes Chromosomes Cancer. 2008 May;47(5):379-85.

Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms.

Grand FH, Hidalgo-Curtis CE, Ernst T, Zoi K, Zoi C, McGuire C, Kreil S, Jones A, Score J, Metzgeroth G, Oscier D, Hall A, Brandts C, Serve H, Reiter A, Chase AJ, Cross NC.

Blood. 2009 Apr 22. [Epub ahead of print]

Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia.

Score J, Curtis C, Waghorn K, Stalder M, Jotterand M, Grand FH, Cross NC.

Leukemia. 2006 May;20(5):827-32.

Durable responses to imatinib in patients with PDGFRB fusion gene-positive and BCR-ABL-negative chronic myeloproliferative disorders.

David M, Cross NC, Burgstaller S, Chase A, Curtis C, Dang R, Gardembas M, Goldman JM, Grand F, Hughes G, Huguet F, Lavender L, McArthur GA, Mahon FX, Massimini G, Melo J, Rousselot P, Russell-Jones RJ, Seymour JF, Smith G, Stark A, Waghorn K, Nikolova Z, Apperley JF.

Blood. 2007 Jan 1;109(1):61-4. Epub 2006 Sep 7.

Transient response to imatinib in a chronic eosinophilic leukemia associated with ins(9;4)(q33;q12q25) and a CDK5RAP2-PDGFRA fusion gene.

Walz C, Curtis C, Schnittger S, Schultheis B, Metzgeroth G, Schoch C, Lengfelder E, Erben P, Müller MC, Haferlach T, Hochhaus A, Hehlmann R, Cross NC, Reiter A.

Genes Chromosomes Cancer. 2006 Oct;45(10):950-6.

Detection and molecular monitoring of FIP1L1-PDGFRA-positive disease by analysis of patient-specific genomic DNA fusion junctions.

Score J, Walz C, Jovanovic JV, Jones AV, Waghorn K, Hidalgo-Curtis C, Lin F, Grimwade D, Grand F, Reiter A, Cross NC.

Leukemia. 2009 Feb;23(2):332-9. Epub 2008 Nov 6.

Imatinib sensitivity as a consequence of a CSF1R-Y571D mutation and CSF1/CSF1R signaling abnormalities in the cell line GDM1.

Chase A, Schultheis B, Kreil S, Baxter J, Hidalgo-Curtis C, Jones A, Zhang L, Grand FH, Melo JV, Cross NC.

Leukemia. 2009 Feb;23(2):358-64. Epub 2008 Oct 30.

Detection and molecular monitoring of FIP1L1-PDGFRA-positive disease by analysis of patient-specific genomic DNA fusion junctions.

Score J, Walz C, Jovanovic JV, Jones AV, Waghorn K, Hidalgo-Curtis C, Lin F, Grimwade D, Grand F, Reiter A, Cross NC.

Leukemia. 2009 Feb;23(2):332-9. Epub 2008 Nov 6.

Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders.

Jones AV, Kreil S, Zoi K, Waghorn K, Curtis C, Zhang L, Score J, Seear R, Chase AJ, Grand FH, White H, Zoi C, Loukopoulos D, Terpos E, Vervessou EC, Schultheis B, Emig M, Ernst T, Lengfelder E, Hehlmann R, Hochhaus A, Oscier D, Silver RT, Reiter A, Cross NC.

Blood. 2005 Sep 15;106(6):2162-8. Epub 2005 May 26.

Minimal molecular response in polycythemia vera patients treated with imatinib or interferon alpha.

Jones AV, Silver RT, Waghorn K, Curtis C, Kreil S, Zoi K, Hochhaus A, Oscier D, Metzgeroth G, Lengfelder E, Reiter A, Chase AJ, Cross NC.

Blood. 2006 Apr 15;107(8):3339-41. Epub 2005 Dec 13.

bjh short report

Two novel imatinib-responsive *PDGFRA* fusion genes in chronic eosinophilic leukaemia

Claire E. Curtis,[1] Francis H. Grand,[1] Pellegrino Musto,[2] Andrew Clark,[3] John Murphy,[3] Gianni Perla,[4] Maria M. Minervini,[4] Janet Stewart,[5] Andreas Reiter[6] and Nicholas C. P. Cross[1]

[1]Wessex Regional Genetics Laboratory, University of Southampton, Salisbury, UK, [2]Haematology and Stem Cell Transplantation Unit, CROB – Centro di Riferimento Oncologico della Basilicata, Rionero in Vulture, Italy, [3]Monklands Hospital, Airdrie, UK, [4]Unit of Haematology, IRCCS 'Casa Sollievo della Sofferenza', S. Giovanni Rotondo, Italy, [5]Institute of Medical Genetics, Yorkhill Hospital, Glasgow, UK, and [6]III. Medizinische Universitätsklinik, Medizinische Fakultät Mannheim der Universität Heidelberg, Heidelberg, Germany

Received 16 February 2007; accepted for publication 26 March 2007
Correspondence: Prof. N.C.P. Cross, Wessex Regional Genetics Laboratory, University of Southampton, Salisbury District Hospital, Salisbury SP2 8BJ, UK.
E-mail: ncpc@soton.ac.uk

Summary

We identified two patients with a t(2;4)(p24;q12) and a t(4;12)(q2?3;p1?2), respectively, in association with *BCR-ABL* and *FIP1L1-PDGFRA* negative chronic eosinophilic leukaemia. Molecular analysis revealed a novel *STRN-PDGFRA* fusion for the t(2;4) and *ETV6-PDGFRA* for the t(4;12). The fusions were confirmed by specific amplification of the genomic breakpoints, reverse transcription polymerase chain reaction and fluorescence *in situ* hybridisation. Both patients were treated with imatinib and, following a rapid haematological response, achieved cytogenetic remission and a major molecular response. In conclusion, *PDGFRA* fuses to diverse partner genes in myeloid disorders. Identification of these fusions is important as they are particularly sensitive to imatinib.

Keywords: *PDGFRA*, *ETV6*, *STRN*, chronic eosinophilic leukaemia, imatinib.

FIP1L1-PDGFRA is currently the most common abnormality in chronic eosinophilic leukaemia (CEL), being seen in approximately 12% of cases with a provisional diagnosis of idiopathic hypereosinophilic syndrome (Cools *et al*, 2003; Griffin *et al*, 2003; Score *et al*, 2006). Other abnormalities associated with CEL include a variety of translocations and fusion genes involving *PDGFRA* at 4q12, *PDGFRB* at 5q31-33, *FGFR1* at 8p11 or *JAK2* at 9p24 (Reiter *et al*, 2007). Identification of patients that harbour *PDGFRA* or *PDGFRB* fusions is particularly important as they usually respond very well to targeted therapy with imatinib (Cools *et al*, 2003; Reiter *et al*, 2007). To date, four fusion partners have been identified for *PDGFRA*: *BCR*, *FIP1L1*, *KIF5B* and *CDK5RAP2* (Baxter *et al*, 2002; Cools *et al*, 2003; Score *et al*, 2006; Walz *et al*, 2006). Here, we report two patients with imatinib responsive CEL who presented with acquired chromosomal rearrangements of 4q.

Patients and methods

Patients

Case 1 had CEL and has been described in detail elsewhere (Musto *et al*, 2004). Briefly, a 64-year-old male presented with an eosinophil count of 12×10^9/l and a bone marrow indicative of a myeloproliferative disorder (MPD). Cytogenetics revealed a t(2;4)(p24;q12) in 60% of bone marrow metaphases. Haematological and cytogenetic remission was achieved following treatment with 100 mg imatinib daily. This dose was continued for 9 months and then progressively reduced to 100 mg on alternate days followed by 100 mg twice a week. Following 24 months of imatinib, he refused further therapy despite not having any significant side effects. Elevated eosinophil counts were again detected 14 months after discontinuation of treatment.

  doi:10.1111/j.1365-2141.2007.06628.x

Case 2 was a 51-year-old male who presented with late onset asthma and type 2 diabetes. He had an eosinophil count of $12{\cdot}9 \times 10^9$/l and a hypercellular bone marrow suggestive of an MPD. Cytogenetics revealed a t(4;12)(q2?3;p1?2) in 100% of bone marrow metaphases. The patient was managed with hydroxycarbamide for over 7 years but with recurrent problems due to anaemia, thrombocytopenia and very poor control of the eosinophil count. Cytogenetic analysis at 6 years after diagnosis showed 46,XY,t(4;12)(q2?3;p1?2)[14]/46,XY[6]. Following the molecular diagnosis detailed below, hydroxycarbamide was stopped and imatinib started at 100 mg daily. The full blood counts completely normalised within 4 weeks, resulting in significant resolution of symptoms (apart from the asthma) and greatly improved quality of life. Cytogenetic analysis was performed 9 months later and indicated complete cytogenetic remission (30 metaphases).

Polymerase chain reaction (PCR) and reverse transcription polymerase chain reaction (RT-PCR)

Conditions for amplifying *PDGFRA* rearrangements by bubble PCR have been described in detail elsewhere (Score *et al*, 2006). Primers used to detect *STRN-PDGFRA* DNA were STRN.Fusion.Ex6.1F (5′-TCTGAGTTAACAGATTCTGCCTC3-′) and PDGFRA.Fusion.Ex12.2R (5′-CAAGCACTAGTCCATCTCTTGG3-′) Primers for *ETV6-PDGFRA* mRNA from random hexamer reverse transcribed cDNA were ETV6.Fusion.Ex6.1F: 5′-ACTGTAGACTGCTTTGGGATTAC3-′ and PDGFRA.Fusion.Ex12.2R. Primer sequences for analysis of reciprocal breakpoints and minimal residual disease (MRD) are available on request.

Fluorescence in situ *hybridisation (FISH)*

Fluorescence *in situ* hybridisation (FISH) was performed using bacterial artificial chromosome (BAC) clones selected from sequences flanking *PDGFRA* and *ETV6* (http://www.ensembl.org) or commercially available probe sets. BAC DNA was grown, extracted and labelled by nick translation with digoxygenin or biotin-16-2′deoxy-uridine-5′-triphosphate (Roche, East Sussex, UK) and hybridised using standard procedures. Case 2 was analysed using BACs RP11-24O10 (3′ of *PDGFRA*); RP11-434C1 (5′ of *ETV6*).

Results

Bubble PCR to identify PDGFRA rearrangements

Both cases were negative for *BCR-ABL* and *FIP1L1-PDGFRA* by RT-PCR or FISH analysis using standard procedures. The presence of a 4q translocation in conjunction with persistent eosinophilia suggested that *PDGFRA* might be involved. As the genomic breakpoints for all cases involving this gene thus far reported are tightly clustered within or close to *PDGFRA* exon 12 (Reiter *et al*, 2007) we investigated the possibility that *PDGFRA* was rearranged, by bubble PCR. As shown in Fig 1A, rearranged bands were detected for both cases 1 and 2.

Identification and confirmation of the STRN-PDGFRA fusion

Sequencing of the rearranged allele for case 1 identified a novel fusion between *STRN* (*STRN*; NM_003162) intron 6 and a truncated *PDGFRA* exon 12. To confirm the fusion, a 285 bp product was amplified from patient genomic DNA but not from control DNA. A reciprocal *PDGFRA-STRN* fusion product was not detected with various primer combinations using either single step or nested PCR. Identification of the mRNA fusion was not possible due to lack of adequate patient material, however analogous *PDGFRA* breakpoints for *FIP1L1-PDGFRA* usually result in the use of cryptic splice sites within *PDGFRA* exon 12 downstream of the breakpoint. Two potential AG splice acceptor sites would produce in-frame mRNA fusions with *STRN* exon 6 and it is likely that either or both of these might represent the real mRNA fusion in case 1 (Fig 1B).

Identification and confirmation of the ETV6-PDGFRA fusion

Sequencing of the rearranged bubble PCR band from case 2 identified a second novel fusion between *ets*-variant gene 6 (*ETV6*; NM_001987) intron 6 and *PDGFRA* intron 11. To confirm the fusion a 210 bp product was specifically amplified from case 2 peripheral blood genomic DNA but not from control DNA using breakpoint-specific primers. A reciprocal *PDGFRA-ETV6* fusion product was not detected with various primer combinations with either single or nested PCR. Expression of *ETV6-PDGFRA* was confirmed by single-step PCR from case 2, yielding a 276 bp product that was not observed in normal control cDNA. Sequencing of the product revealed an in-frame whole exon fusion between *ETV6* exon 6 and *PDGFRA* exon 12 (Fig 2A). To further confirm the presence of the *ETV6-PDGFRA* fusion, two-colour interphase FISH analysis was performed. On normal cells BACs RP11-24O10 and RP11-434C1 gave four separate signals corresponding to the normal chromosomes 4 and 12. However, in 56% of cells for case 2, overlapping probe signals consistent with co-localisation of *ETV6* and *PDGFRA* were observed (Fig 2B).

Response to imatinib

For cases 1 and 2, peripheral blood was analysed approximately 28 and 9 months after starting imatinib respectively. MRD was not detected using cDNA or genomic DNA for case 1, indicating the achievement of molecular remission. For case 2, analysis was only performed using genomic DNA. Of four replicate reactions, two were positive indicating low-level disease consistent with a major molecular response.

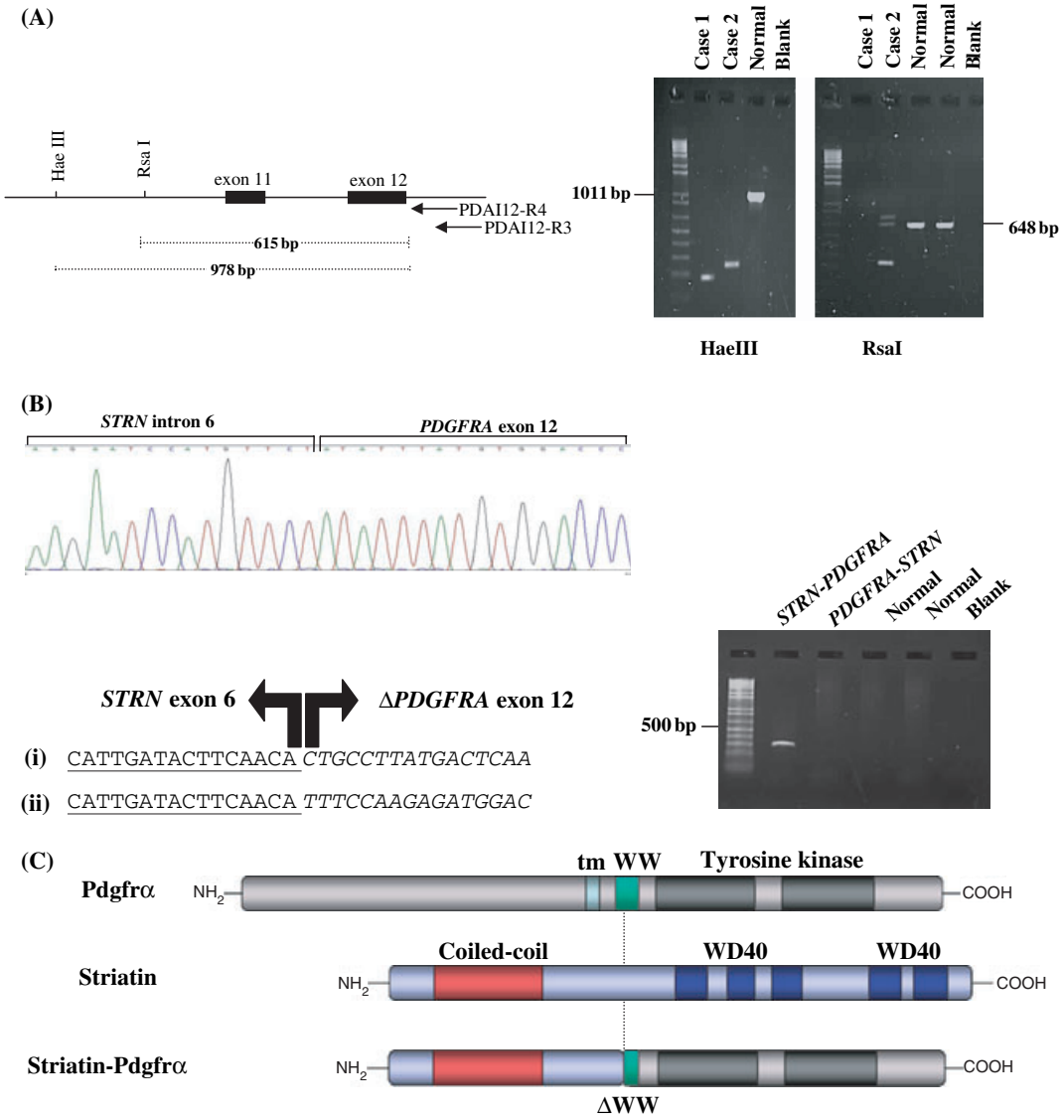


Fig 1. Characterisation of *STRN-PDGFRA*. (A) Left panel: Bubble PCR strategy showing PCR primers and restriction sites in the region where all breakpoints have been reported to fall; right panel: Nested bubble PCR gel after digestion with HaeIII and RsaI. The normal controls show the expected germline bands of 1011 and 648 bp respectively (the distance between primer PDAI12-R4 and the relevant restriction site plus 33 bp contributed by the bubble oligo). Smaller bands are visible for cases 1 and 2 after HaeIII digestion and for case 2 after RsaI digestion derived from the rearranged allele. The normal allele is not always visible due to out competition in the amplification reaction. (B) Electropherogram of the genomic DNA junction amplified by bubble PCR, the two predicted in frame *STRN-PDGFRA* mRNA fusions and gel showing amplification of *STRN-PDGFRA* but not the reciprocal genomic DNA fusion specifically from case 1. (C) Structure of Pdgfrα, Striatin and the predicted fusion protein. Relevant domains are shown: transmembrane (tm), WW-like (WW) and tyrosine kinase domains for Pdgfrα; coiled-coiled region and WD40 repeats for striatin; partially deleted WW domain for striatin-Pdgfrα (*Δ*WW).

Discussion

This study has characterised the genomic breakpoints in two CEL cases with chromosome 4q rearrangements and identified *STRN* at 2p22 and *ETV6* at 12p13 as novel *PDGFRA* fusion partners. This brings the number of known *PDGFRA* partner genes to six, with *PDGFRA* being the tenth tyrosine kinase reported to fuse to *ETV6*. As these fusions were referred specifically for investigation of 4q rearrangements and were not part of a series that underwent systematic cytogenetic investigation, it is not possible to accurately determine the frequency of these abnormalities. However, we estimate that they probably account for <1% of CEL cases.

STRN encodes striatin, a 780 amino acid member of the calmodulin-binding WD repeat family of proteins. It has four protein–protein interaction domains, a caveolin binding domain (55–63aa), a coiled-coil domain (70–116aa), a calcium dependent calmodulin-binding domain (149–166aa) and a series of WD repeat motifs (419–780aa) (Bartoli *et al*, 1998). Striatin is thought to be a scaffolding protein that targets

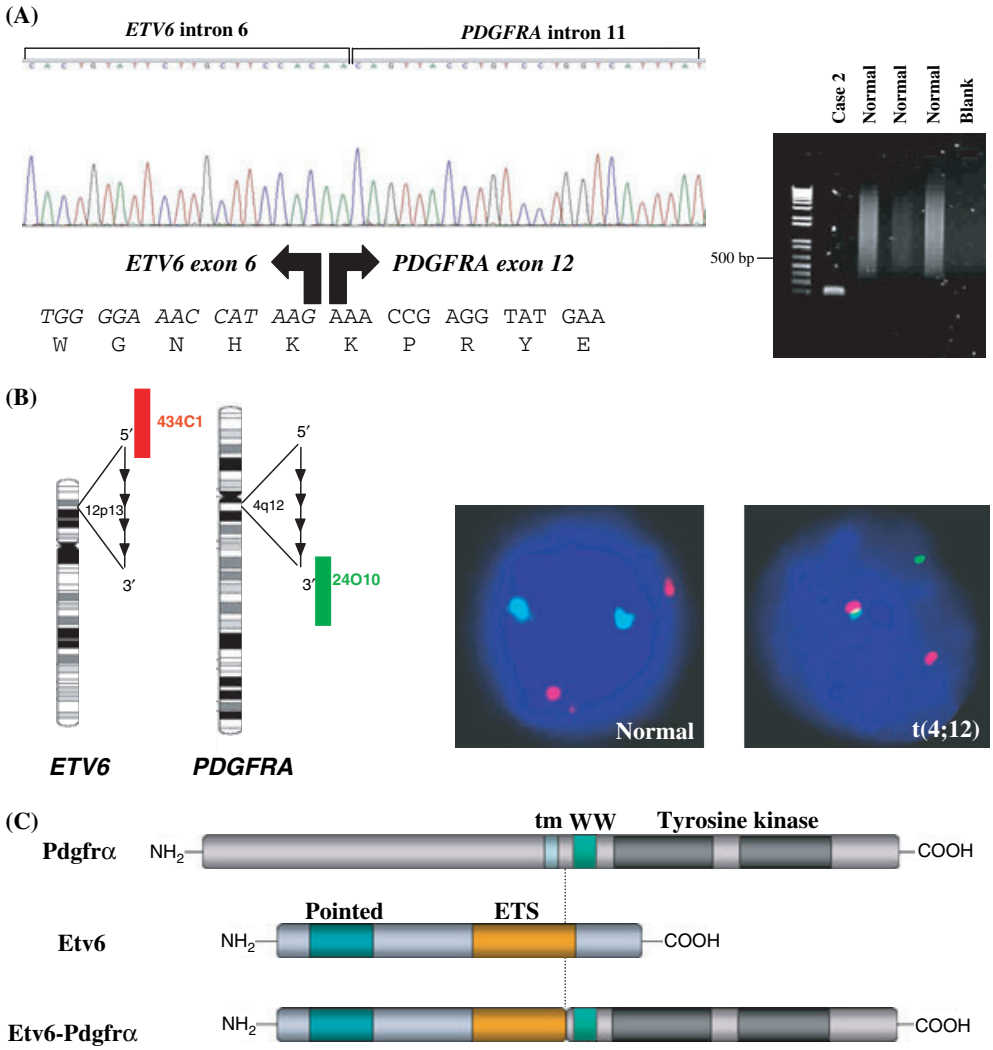


Fig 2. Characterisation of the *ETV6-PDGFRA* fusion. (A) Electropherogram of the genomic DNA junction amplified by bubble PCR; mRNA fusion between *ETV6* exon 6 and *PDGFRA* exon 12 as determined by sequence analysis of case 2 RT-PCR product and gel showing amplification of *ETV6-PDGFRA* specifically from genomic DNA of case 2. (B) Relative positions of FISH probes with the orientation of each gene indicated by the arrows. A normal cell shows four distinct signals whereas 56% of cells from case 2 showed a fused red/green signal indicative of the t(4;12). (C) Structure of Pdgfrα, Etv6 and the predicted fusion protein. The pointed and ETS domains are shown for Etv6; other domains are as shown in Fig 1.

proteins, such as the oestrogen receptor to the cell membrane, facilitating the assembly of other proteins required for rapid receptor activation (Lu *et al*, 2004). In case 1, we identified a DNA fusion of *STRN* intron 6 to a truncated *PDGFRA* exon 12 and thus the large region of WD repeat motifs within striatin are not retained in the fusion. However, the N-terminal coiled-coil domain remains that may act as an oligomerisation motif, constitutively activating the Pdgfrα-derived tyrosine kinase moiety (Fig 1C). Alternatively, truncation of the WW-like domain may be the critical activating event (Stover *et al*, 2006).

Etv6 (*ets* translocation variant 6) is a member of the *ets* family of transcription factors and is frequently rearranged in both myeloid and lymphoid malignancies (Bohlander, 2005). *ETV6* fuses to a second gene at 4q12, *CHIC2/BTL* in undifferentiated AML (Cools *et al*, 1999). Since *PGFRA* and *CHIC2/BTL* are only 200 kb apart, *ETV6-PDGFRA* and *ETV6-CHIC2* would be expected to have a cytogenetically indistinguishable t(4;12)(q12;p13), however our case was reported to have a t(4;12)(q2?3;p1?2) which suggests that the rearrangement may have been more complex.

In common with other *PDGFRA* fusions, we found that *STRN* disrupts the region encoding negative regulatory WW-like domain of Pdgfrα. In contrast, for *ETV6-PDGFRA* the WW-like domain is intact (Fig 2C). Recently, Stover *et al* (2006) showed that truncation of the Pdgfrα WW-like domain is sufficient to activate the kinase in the absence of an oligomerisation domain provided by the partner protein. An artificial *ETV6-PDGFRA* fusion that retained the region encoding the WW-like domain transformed Ba/F3 cells to growth factor independence, indicating that autoinhibition by the WW-like domain can be overcome by enforced

homodimerisation. However, the growth of the transformed cells was slower than those transformed by similar fusions in which the WW-like domain was truncated (Stover *et al*, 2006). It remains unclear why the WW-like domain is truncated with other fusion partners that provide an oligomerisation domain such as Bcr, and also why truncation of this domain is not usually seen for fusions involving Pdgfrβ.

Acknowledgements

This work was supported by the Leukaemia Research Fund.

References

Bartoli, M., Monneron, A. & Ladant, D. (1998) Interaction of calmodulin with striatin, a WD-repeat protein present in neuronal dendritic spines. *Journal of Biological Chemistry*, **273,** 22248–22253.

Baxter, E.J., Hochhaus, A., Bolufer, P., Reiter, A., Fernandez, J.M., Senent, L., Cervera, J., Moscardo, F., Sanz, M.A. & Cross, N.C.P. (2002) The t(4;22)(q12;q11) in atypical chronic myeloid leukaemia fuses BCR to PDGFRA. *Human Molecular Genetics*, **11,** 1391–1397.

Bohlander, S.K. (2005) ETV6: a versatile player in leukemogenesis. *Seminars in Cancer Biology*, **15,** 162–174.

Cools, J., Bilhou-Nabera, C., Wlodarska, I., Cabrol, C., Talmant, P., Bernard, P., Hagemeijer, A. & Marynen, P. (1999) Fusion of a novel gene, BTL, to ETV6 in acute myeloid leukemias with a t(4;12)(q11-q12;p13). *Blood*, **94,** 1820–1824.

Cools, J., De Angelo, D.J., Gotlib, J., Stover, E.H., Legare, R.D., Cortes, J., Kutok, J., Clark, J., Galinsky, I., Griffin, J.D., Cross, N.C., Tefferi, A., Malone, J., Alam, R., Schrier, S.L., Schmid, J., Rose, M., Vandenberghe, P., Verhoef, G., Boogaerts, M., Wlodarska, I., Kantarjian, H., Marynen, P., Coutre, S.E., Stone, R. & Gilliland, D.G. (2003) A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *New England Journal of Medicine*, **348,** 1201–1214.

Griffin, J.H., Leung, J., Bruner, R.J., Caligiuri, M.A. & Briesewitz, R. (2003) Discovery of a fusion kinase in EOL-1 cells and idiopathic hypereosinophilic syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **100,** 7830–7835.

Lu, Q., Pallas, D.C., Surks, H.K., Baur, W.E., Mendelsohn, M.E. & Karas, R.H. (2004) Striatin assembles a membrane signaling complex necessary for rapid, nongenomic activation of endothelial NO synthase by estrogen receptor alpha. *Proceedings of the National Academy of Sciences of the United States of America*, **101,** 17126–17131.

Musto, P., Falcone, A., Sanpaolo, G., Bodenizza, C., Perla, G., Minervini, M.M., Cascavilla, N., Dell'Olio, M., La, S.A., Mantuano, S., Melillo, L., Nobile, M., Scalzulli, P.R., Bisceglia, M. & Carella, A.M. (2004) Heterogeneity of response to imatinib-mesylate (glivec) in patients with hypereosinophilic syndrome: implications for dosing and pathogenesis. *Leukemia & Lymphoma*, **45,** 1219–1222.

Reiter, A., Walz, C. & Cross, N.C.P. (2007) Tyrosine kinases as therapeutic targets in BCR-ABL negative chronic myeloproliferative disorders. *Current Drug Targets*, **8,** 205–216.

Score, J., Curtis, C., Waghorn, K., Stalder, M., Jotterand, M., Grand, F.H. & Cross, N.C.P. (2006) Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia. *Leukemia*, **20,** 827–832.

Stover, E.H., Chen, J., Folens, C., Lee, B.H., Mentens, N., Marynen, P., Williams, I.R., Gilliland, D.G. & Cools, J. (2006) Activation of FIP1L1-PDGFRalpha requires disruption of the juxtamembrane domain of PDGFRalpha and is FIP1L1-independent. *Proceedings of the National Academy of Science of the United States of America*, **103,** 8078–8083.

Walz, C., Curtis, C., Schnittger, S., Schultheis, B., Metzgeroth, G., Schoch, C., Lengfelder, E., Erben, P., Muller, M.C., Haferlach, T., Hochhaus, A., Hehlmann, R., Cross, N.C.P. & Reiter, A. (2006) Transient response to imatinib in a chronic eosinophilic leukemia associated with ins(9;4)(q33;q12q25) and a CDK5RAP2-PDGFRA fusion gene. *Genes, Chromosomes & Cancer*, **45,** 950–956.

References

1 Melief CJ, Toes RE, Medema JP, van der Burg SH, Ossendorp F, Offringa R. Strategies for immunotherapy of cancer. *Adv Immunol* 2000; **75**: 235–282.
2 Woan K, Reddy V. Potential therapeutic role of natural killer cells in cancer. *Expert Opin Biol Ther* 2007; **7**: 17–29.
3 Koehl U, Esser R, Zimmermann S, Tonn T, Kotchetkov R, Bartling T *et al. Ex vivo* expansion of highly purified NK cells for immunotherapy after haploidentical stem cell transplantation in children. *Klinische Padiatrie* 2005; **217**: 345–350.
4 Kobari L, Pflumio F, Giarratana M, Li X, Titeux M, Izac B *et al. In vitro* and *in vivo* evidence for the long-term multilineage (myeloid, B, NK, and T) reconstitution capacity of ex vivo expanded human CD34(+) cord blood cells. *Exp Hematol* 2000; **28**: 1470–1480.
5 Amsellem S, Pflumio F, Bardinet D, Izac B, Charneau P, Romeo PH *et al. Ex vivo* expansion of human hematopoietic stem cells by direct delivery of the HOXB4 homeoprotein. *Nat Med* 2003; **9**: 1423–1427.
6 Haddad R, Guardiola P, Izac B, Thibault C, Radich J, Delezoide AL *et al.* Molecular characterization of early human T/NK and B-lymphoid progenitor cells in umbilical cord blood. *Blood* 2004; **104**: 3918–3926.
7 Carayol G, Robin C, Bourhis JH, Bennaceur-Griscelli A, Chouaib S, Coulombel L *et al.* NK cells differentiated from bone marrow, cord blood and peripheral blood stem cells exhibit similar phenotype and functions. *Eur J Immunol* 1998; **28**: 1991–2002.
8 Dalle JH, Menezes J, Wagner E, Blagdon M, Champagne J, Champagne MA *et al.* Characterization of cord blood natural killer cells: implications for transplantation and neonatal infections. *Pediatr Res* 2005; **57** (5 Part 1): 649–655.

A novel *ETV6-PDGFRB* fusion transcript missed by standard screening in a patient with an imatinib responsive chronic myeloproliferative disease

Leukemia (2007) **21,** 1839–1841; doi:10.1038/sj.leu.2404728; published online 17 May 2007

ETV6-PDGFRB results from the t(5;12)(q31–33;p13) and, although rare, is the most frequent of at least 15 known *PDGFRB* fusion genes. *ETV6-PDGFRB* is usually seen in patients with chronic myelomonocytic leukaemia, atypical chronic myeloid leukaemia (aCML) or chronic eosinophilic leukaemia (CEL), with most patients exhibiting peripheral blood or bone marrow eosinophilia. Although *ETV6-PDGFRB* disease may be controlled with conventional therapy (hydroxyurea, busulphan, interferon-α), many die from transformed disease or other causes. More recently, however, imatinib has emerged as the standard therapy for *PDGFR*-rearranged patients, with most experiencing sustained haematologic and cytogenetic remission, and many achieving molecular remission.[1] Molecular confirmation of a *PDGFR* rearrangement is therefore important for appropriate clinical management and monitoring response to therapy. For the t(5;12)(q31–33;p13), confirmation of the presence of *ETV6-PDGFRB* is particularly important since many cases with this translocation have a distinct molecular aetiology, with upregulation of *IL-3* expression and no involvement of *PDGFRB*.[2]

In the great majority of positive cases, *ETV6* exon 4 splices to *PDGFRB* exon 11 and is typically detected by reverse transcription PCR (RT-PCR) using primers directed to *ETV6* exon 3 or 4 and *PDGFRB* exon 11. Out of 10 positive cases we have detected with these primer combinations, only one had a variant fusion in which *ETV6* exon 4 was fused to *PDGFRB* exon 9 (unpublished data). Recently, a patient with idiopathic myelofibrosis and an *ETV6* exon 7–*PDGFRB* exon 10 fusion was reported and it was suggested that this novel transcript may have accounted for the somewhat unusual clinical phenotype seen in this case.[3] Importantly, both these variants are detectable using the standard primer sets above. Here we report a novel *ETV6-PDGFRB* fusion with breakpoints outside this region.

A 19–year-old male (E1026) presented to the outpatient department with a 2-month history of fever. On examination he was found to have generalized lymphadenopathy, hepatosplenomegaly and bony swellings over the left clavicle and right tibia. There were no bleeding manifestations. A peripheral blood smear showed a total leucocyte count exceeding 100×10^9/l with a left shift. The bone marrow aspirate was suggestive of a myeloproliferative disorder with an increase in the number of basophils and eosinophils. Fine needle aspiration of the lymph node showed infiltration by blasts which cytochemistry revealed to be negative for myeloperoxidase and periodic acid schiff. Flow cytometry of the peripheral blood and lymph node aspirate was positive for myeloid markers (CD 11c, 13, 15 and 33) along with CD45 and myeloperoxidase. *BCR-ABL* was not detected by RT-PCR and cytogenetics showed a t(5;12)(q31–33;p13). A diagnosis of atypical CML in blast crisis was made. Despite the fact that *ETV6-PDGFRB* was not detected by RT-PCR, he was started on imatinib mesylate (300 mg twice a day) after taking informed consent. The patient had a complete haematological and cytogenetic response (including complete regression of the lymph nodes, bony swellings and hepatosplenomegaly) at 3 months post-treatment. He continues in remission on 600 mg/day imatinib, 1 year after diagnosis.

In view of the excellent response to imatinib, we performed further analysis to determine if *PDGFRB* was involved (material was not available for molecular cytogenetic analysis). We have previously reported a simple multiplex RT-PCR assay to identify *PDGFRA* fusions by looking for overexpression of sequence encoding the kinase domain (which is retained in any fusion) relative to that encoding the extracellular domain (which is expressed only by the normal, unrearranged gene).[4] Attempts to establish a similar assay for *PDGFRB* were only partially successful and therefore we developed an analogous strategy using multiplex ligation-dependent probe amplification (MLPA). We designed probes targeting the region encoding the kinase domain and upstream sequences of *PDGFRA*, *PDGFRB* and *FGFR1*. MLPA was performed on patient and control cDNA essentially as described.[5] The cDNA MLPA assay detected relative overexpression of the *PDGFRB* kinase domain in 9 of 9 patients known to harbour *PDGFRB* fusion genes but not in 16 normal controls. Overexpression of the kinase domain was also seen for E1026, suggesting the presence of a *PDGFRB* fusion (Figure 1a).

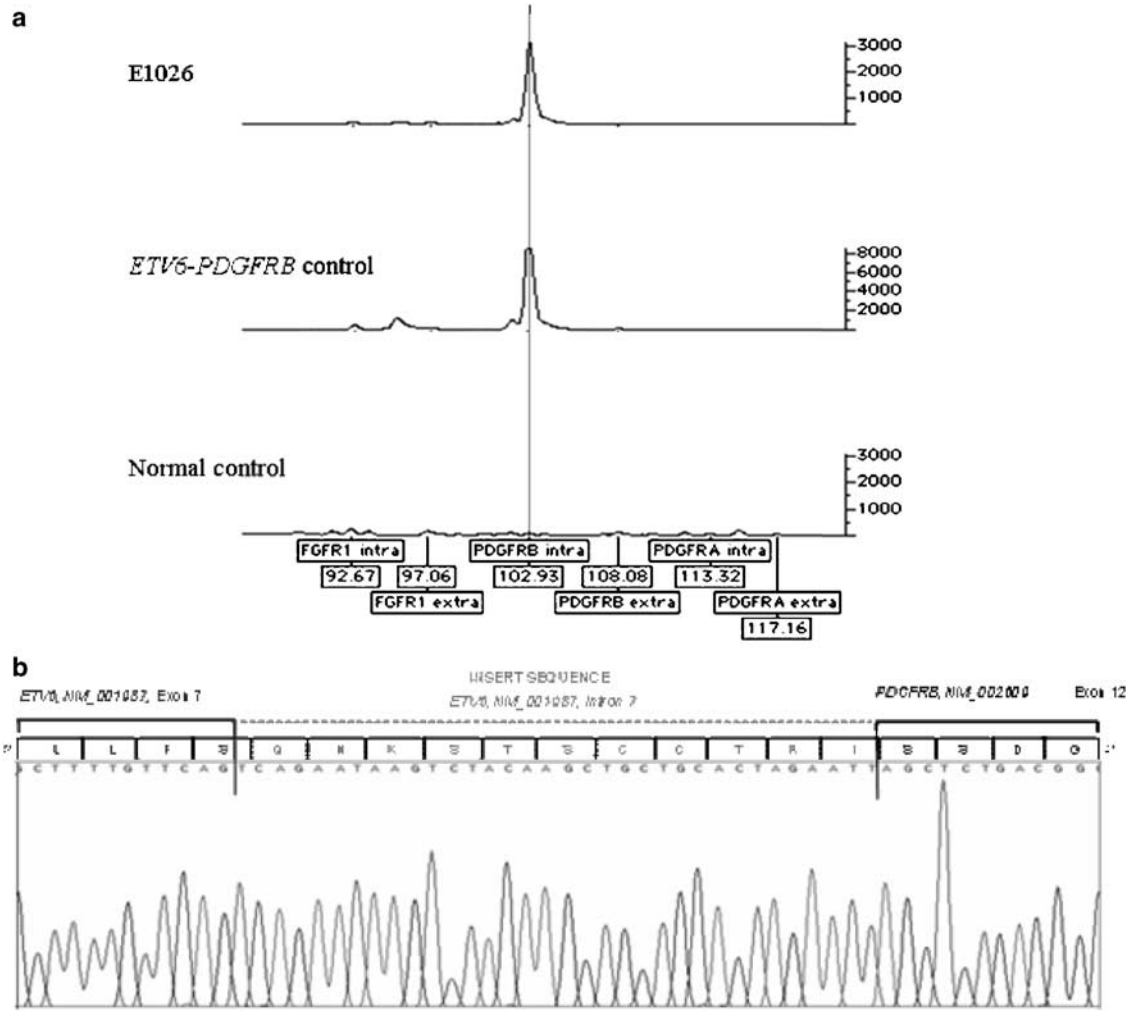


Figure 1 (**a**) The cDNA MLPA assay. Probes are in the following order: *FGFR1* tyrosine kinase domain, *FGFR1* 3′ extracellular domain, *PDGFRB* tyrosine kinase domain, *PDGFRB* 3′ extracellular domain, *PDGFRA* tyrosine kinase domain, *PDGFRA* 3′ extracellular domain. Primer sequences and exact conditions are available on request. Under normal circumstances, expression of *FGFR1, PDGFRB* and *PDGFRA* is low in level in peripheral blood leucocytes. Overexpression of the *PDGFRB* intracellular domain is apparent in both the previously known *ETV6-PDGFRB* case and also E1026. (**b**) Electropherogram of the cDNA junction amplified by 5′rapid amplification of cDNA ends PCR. MLPA, multiplex ligation-dependent probe amplification.

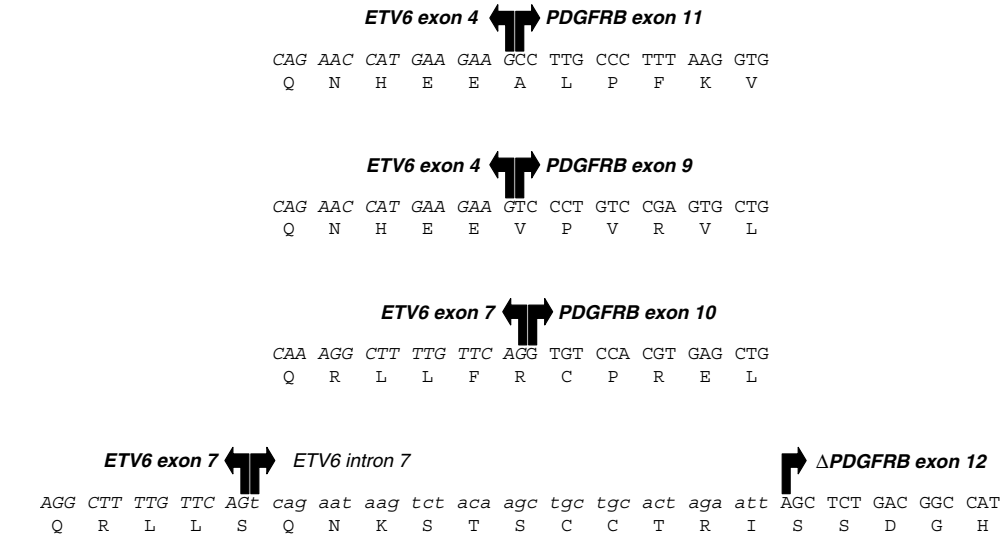


Figure 2 Sequence of *ETV6-PDGFRB* fusions described to date, with the most common variant at the top and the fusion in case E1026 at the bottom.

We used 5′ rapid amplification of cDNA ends PCR (Invitrogen, Paisley, UK) to characterize the fusion, employing reverse primers derived from sequences located downstream of known *PDGFRB* breakpoint locations. The resulting PCR products were cloned and sequencing revealed an in-frame mRNA fusion between full-length *ETV6* exon 7, 34 bp derived from *ETV6* intron 7 and a truncated *PDGFRB* exon 12 (Figures 1b and 2). The presence of the fusion was confirmed by RT-PCR and by amplifying and sequencing fusion point from genomic DNA (not shown).

The *PDGFRB* sequence retained in the novel fusion is downstream of the PCR primer we had initially used, explaining why *ETV6-PDGFRB* was not detected. To determine if we might have missed other similar cases, we analyzed 10 cases with a t(5;12) who tested negative for *ETV6-PDGFRB* using primers ETV6.Ex2F (5′-TCAGGATGGAGGAAGACTCG-3′) and PDGFRB.Ex14R (5′-CCCCAACAGGTTGACCACGTTCAG-3′). All were negative, indicating that this variant is not common, but nevertheless we now routinely use these primers to screen t(5;12) cases.

This unusual fusion in our case is reminiscent of that seen for *FIP1L1-PDGFRA*, for which breaks almost always fall within *PDGFRA* exon 12 and *FIP1L1* intron-derived sequence is frequently incorporated into the mature mRNA.[6] As a consequence, the PDGFRα autoinhibitory WW-like domain is disrupted resulting in partner protein-independent activation of the kinase moiety.[7] The disruption of PDGFRβ WW-like domain in our case is unexpected in view of the fact that *ETV6* encodes a dimerization domain that is required for transformation by 'normal' *ETV6-PDGFRB*.

It is notable that our case presented with advanced phase disease. The fusion retained *ETV6* exons 1–7, in common with that described by Tokita *et al*.[3] in a case that also presented with advanced features. In contrast to the most common fusion involving *ETV6* exons 1–4, these variants are predicted to encode a chimaeric protein that retains the ETV6 internal and ETS DNA-binding domains. It is possible that retention of these domains leads to a more aggressive phenotype; however, it would be premature to draw any clear conclusions from just two cases. It is noteworthy, however, that our case responded well to imatinib despite being treated in advanced phase. Good responses to imatinib have been described for *FIP1L1-PDGFRA*-associated CEL in transformation[8] and this may be a general feature of diseases associated with PDGFR fusions.

In summary, it appears that breakpoint diversity for the *ETV6-PDGFRB* fusion is more common than originally reported. Screens for this fusion should employ primers capable of detecting all variants in order to provide an accurate molecular diagnosis and appropriate clinical management.

Acknowledgements

This work was supported by Leukaemia Research (UK).

CE Curtis[1], FH Grand[1], K Waghorn[1], TP Sahoo[2], J George[2] and NCP Cross[1]
[1]*Wessex Regional Genetics Laboratory, Salisbury District Hospital, Salisbury, UK and*
[2]*Department of Medical Oncology, Kidwai Memorial Institute of Oncology, Bangalore, India*
E-mail: ncpc@soton.ac.uk

References

1 Steer EJ, Cross NCP. Myeloproliferative disorders with translocations of chromosome 5q31–35: role of the platelet-derived growth factor receptor Beta. *Acta Haematol* 2002; **107**: 113–122.
2 Cools J, Mentens N, Odero MD, Peeters P, Wlodarska I, Delforge M *et al*. Evidence for position effects as a variant ETV6-mediated leukemogenic mechanism in myeloid leukemias with a t(4;12)(q11–q12;p13) or t(5;12)(q31;p13). *Blood* 2002; **99**: 1776–1784.
3 Tokita K, Maki K, Tadokoro J, Nakamura Y, Arai Y, Sasaki K *et al*. Chronic idiopathic myelofibrosis expressing a novel type of TEL-PDGFRB chimaera responded to imatinib mesylate therapy. *Leukemia* 2007; **21**: 190–192.
4 Score J, Curtis C, Waghorn K, Stalder M, Jotterand M, Grand FH *et al*. Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia. *Leukemia* 2006; **20**: 827–832.
5 Kreil S, Pfirrmann M, Hareflach C, Waghirn K, Chase A, Hehlamnn R *et al*. Heterogeneous prognostic impact of derivative chromosome 9 deletions in chronic myelogenous leukemia. *Blood* 2007, E-pub April 24.
6 Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J *et al*. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med* 2003; **348**: 1201–1214.
7 Stover EH, Chen J, Folens C, Lee BH, Mentens N, Marynen P *et al*. Activation of FIP1L1-PDGFRalpha requires disruption of the juxtamembrane domain of PDGFRalpha and is FIP1L1-independent. *Proc Natl Acad Sci USA* 2006; **103**: 8078–8083.
8 Metzgeroth G, Walz C, Score J, Siebert R, Schnittger S, Haferlach C *et al*. Recurrent finding of the FIP1L1-PDGFRA fusion gene in eosinophilia-associated acute myeloid leukemia and lymphoblastic T-cell lymphoma. *Leukemia* 2007, E-pub March 22.

Treatment of severe autoimmune hemolytic anemia in B-cell chronic lymphocytic leukemia with alemtuzumab

Leukemia (2007) **21,** 1841–1842; doi:10.1038/sj.leu.2404713; published online 26 April 2007

We read with interest the article by Karlsson *et al*.[1] in *Leukemia* (2007), which reports five patients treated successfully with alemtuzumab (anti-CD52) for refractory autoimmune hemolytic anemia (AIHA) complicating B-cell chronic lymphocytic leukemia (B-CLL). AIHA occurs in about 5% of patients with B-CLL and requires treatment including corticosteroids, cytotoxic drugs, splenectomy or immunosuppressive drugs.[2] Recently, rituximab (anti-CD20) has been used with efficacy in AHAI complicating B-CLL.[3] But relapses are frequent and some patients become refractory.

We report the case of a 69-year-old woman who presented B-CLL complicated by AIHA. B-CLL was diagnosed in 1994 and was treated with different chemotherapy regimens (chlorambucile, cyclophosphamide and then fludarabine) between 1994 and 2002. In October 2003 chlorambucile was introduced for progressing B-CLL; the direct antiglobulin test (DAT) was

The t(1;9)(p34;q34) and t(8;12)(p11;q15) Fuse Pre-mRNA Processing Proteins *SFPQ (PSF)* and *CPSF6* to *ABL* and *FGFR1*

Claire Hidalgo-Curtis,[1] Andrew Chase,[1] Milton Drachenberg,[2] Mark W. Roberts,[2] Jerry Z. Finkelstein,[2] Sarah Mould,[3] David Oscier,[3] Nicholas C. P. Cross,[1] and Francis H. Grand[1]*

[1]Wessex Regional Genetics Laboratory, Salisbury District Hospital and Human Genetics Division, University of Southampton, Southampton, United Kingdom
[2]Long Beach Memorial Medical Center, California, CA 90801
[3]Department of Hematology, Royal Bournemouth Hospital, Bournemouth, UK

We have investigated two patients with acquired chromosomal rearrangements, a male presenting with a t(1;9)(p34;q34) and B cell progenitor acute lymphoid leukemia and a female presenting with a t(8;12)(p11;q15) and the 8p11 myeloproliferative syndrome. We determined that the t(1;9) fused *ABL* to *SFPQ* (also known as *PSF*), a gene mapping to 1p34 that encodes a polypyrimidine tract-binding protein-associated splicing factor. The t(8;12) fused *CPSF6*, a cleavage and polyadenylation specificity factor, to *FGFR1*. The fusions were confirmed by amplification of the genomic breakpoints and RT-PCR. The predicted oncogenic products of these fusions, SFPQ-ABL and CPSF6-FGFR1, are in-frame and encode the N-terminal domain of the partner protein and the entire tyrosine kinase domain and C-terminal sequences of ABL and FGFR1. SFPQ interacts with two FGFR1 fusion partners, ZNF198 and CPSF6, that are functionally related to the recurrent PDGFRα partner FIP1L1. Our findings thus identify a group of proteins that are important for pre-mRNA processing as fusion partners for tyrosine kinases in hematological malignancies.

INTRODUCTION

Diverse tyrosine kinase fusion genes have been identified in hematological malignancies, particularly myeloproliferative disorders (MPD), acute lymphoblastic leukemia (ALL), and certain subtypes of lymphoma. The 8p11 myeloproliferative syndrome (EMS) is a chronic MPD that commonly presents with eosinophilia, T-cell proliferation and acquired reciprocal chromosomal translocations with recurrent breakpoints at chromosome band 8p11 that disrupt the receptor tyrosine kinase *FGFR1*. To date, eight fusion genes have been identified by the molecular investigation of chromosomal rearrangements in patients with EMS; these are the t(8;13)(p11;q12), t(6;8)(q27;p11), t(8;9)(p11;q33), t(8;22)(p11;q22), t(8;19)(p12;q13.3), ins(12;8)(p11;p11p22), t(7;8)(q34;p11), t(8;17)(p11;q23), and which fuse *ZMYM2*, *FGFR1OP*, *CEP110*, *BCR*, *HERV-K*, *FGFR1OP2*, *TRIM24*, *MYO18A*, to *FGFR1*, respectively (Xiao et al., 1998; Popovici et al., 1999; Guasch et al., 2000, 2003; Demiroglu et al., 2001; Grand et al., 2004; Belloni et al., 2005; Walz et al., 2005). The *FGFR1* fusion genes encode constitutively active proteins that are functionally and structurally similar to BCR-ABL. The *ABL* gene has also been described as being fused to multiple partner genes: *BCR* (Groffen et al., 1984), *ETV6* (also known as *TEL*) (Golub, 1997), *NUP214* (Graux et al., 2004), and *EML1* (De Keersmaecker et al., 2005). Although uncommon, these fusions are important to recognise because of their potential to be targeted effectively by small molecular inhibitors (Cross and Reiter, 2006).

For most tyrosine kinase fusions, the partner proteins provide dimerisation motifs that are essential for transforming activity; however, accumulating evidence suggests that partner proteins may play additional roles in modifying intracellular signalling and disease phenotype. Fusions such as FIG-ROS (Charest et al., 2003) and FGFR1OP-FGFR1 (Delaval et al., 2005) are directed by the partner protein to specific intracellular compartments, which may be critical to their transforming ability. BCR fuses to multiple kinases including PDGFRA (Baxter et al., 2002), FGFR1 (Demi-

Supported by: The Leukaemia Research Fund.

*Correspondence to: F. H. Grand, Wessex Regional Genetics Laboratory, Salisbury District Hospital, Salisbury SP2 8BJ, UK. E-mail: f.h.grand@soton.ac.uk

Received 5 November 2007; Accepted 17 December 2007
DOI 10.1002/gcc.20541
Published online 18 January 2008 in Wiley InterScience (www.interscience.wiley.com).

roglu et al., 2001), and ABL (Groffen et al., 1984). BCR-FGFR1 binds GRB2 via BCR Tyr177 and induces a CML-like leukaemia in mice, whereas BCR-FGFR1/Y177F (which lacks the GRB2 binding site) causes EMS-like disease. These results implicate different signalling pathways originating from both the kinase and the fusion partner in the differential pathogenesis of CML and EMS (Roumiantsev et al., 2004).

In this study, we have investigated an acquired t(1;9)(p34;q34), and identified *SFPQ* as a new fusion partner to *ABL*. The second patient presented with EMS and an acquired chromosomal rearrangement of 8p11, resulting in the fusion of a new partner gene, *CPSF6* to *FGFR1*. Both partner genes share a common role associated with pre-mRNA processing and our findings thus identify a group of proteins that are important for pre-mRNA processing as fusion partners for multiple tyrosine kinases in hematological malignancies.

MATERIALS AND METHODS

Patient Details

Case 1: A 22-year-old male patient presented with CD34 positive ALL of pre-B immunophenotype. His full blood count at presentation was as follows: hemoglobin 7.8 g/L; WBC 39.7×10^9 L^{-1} (89% blasts); platelets 25×10^9 L^{-1}. Bone marrow analysis revealed 94% blasts and cytogenetics showed a t(1;9)(p34;q34)[7]/47, idem, +21[6]/46, XY [7]. The patient was treated according to the Children's Cancer Group Protocol, 1961, and subsequently received augmented BFM therapy with doxorubicin and double delayed intensification. He achieved complete remission but suffered extramedullary testicular relapse at 4.5 years. Following orchidectomy and intensive chemotherapy, he remains in complete remission more than 6 years after diagnosis.

Case 2: A 75-year-old female presented with cervical, left axillary, and inguinal lymphadenopathy. A CT scan showed right pleural effusion, ascites, splenomegaly, and para-aortic nodes. Axillary node biopsy showed effacement of normal architecture by a diffuse population of mononuclear cells. The majority expressed strong surface CD3, but there was a scattered population of cells expressing myeloid markers such as neutrophil elastase. The presenting blood count showed a hemoglobin of 140 g/L, WBC 16×10^9 L^{-1}, with neutrophilia (12×10^9 L^{-1}) and eosinophilia (0.8×10^9 L^{-1}) plus platelets of 375×10^9 L^{-1}. The marrow was hypercellular with an increase in eosinophils and eosinophil precursors but no increase in lymphoid cells or blasts. Cytogenetics analysis revealed 46,XX,t(8;12)(p11;q15)[19]/46, XX [1]. There was rapid clinical deterioration accompanied by pyrexia, lethargy, weight loss, and a widespread rash (biopsy was inconclusive on histology). Intensive chemotherapy was not offered and although there was a transient clinical response to steroids, she died 10 weeks later.

5′ Rapid Amplification of cDNA Ends (5′ RACE) PCR

Partner genes were identified using the GeneRacer™ Kit (Invitrogen, Paisley, UK). Briefly, ~5 μg of total RNA extracted using the Qiagen RNeasy kit (Qiagen, Boundary Court, UK) was dephosphorylated, decapped, and ligated to the GeneRacer™ RNA oligo according to the manufacturer's instructions. The ligated RNA was reverse transcribed using Superscript II™ reverse transcriptase (Invitrogen, Paisley, UK) and random primers (100 ng). Single step 5′RACE PCR was performed using the 5′ GeneRacer™ primer from the kit in combination with a reverse primer from ABL exon 4 (1R: 5′-CCA CCG TTG AAT GAT GAA CC-3′) for Case 1. For Case 2, 5′RACE was performed using nested *FGFR1* primers as previously described (Grand et al., 2004). The PCR cycles were designed to amplify fragments up to 3 kb, with an annealing temperature of 66°C using the High Fidelity PCR Master kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. Products were cloned with the TOPO TA Cloning Kit for Sequencing (Invitrogen, Paisley, UK) and sequenced.

RT-PCR Methods

Case 1: The presence of *SFPQ-ABL* mRNA was confirmed on random hexamer reverse transcribed cDNA using primers to *SFPQ* (Exon 5 1F: 5′-GAT GCC TAT CAT GAA CAT CAG GC-3′) with the reverse *ABL* exon 4 1R (see above). Because of limited patient material the genomic breakpoint was initially amplified by long PCR using a forward genomic primer from intron 9 of *SFPQ* (2F: 5′-GTG TAG GCT GCT GTG TTT GCA T-3′) and reverse primer *ABL* exon 4 1R. This product was cloned (Invitrogen, Paisley, UK) and sequenced using a reverse *ABL* primer from intron 4 (5′-TCT TTT GAC AGC CAC ACT GAC AG).

Case 2: The presence of *CPSF6-FGFR1* fusion mRNA was confirmed on random hexamer reverse transcribed cDNA using primers to *CPSF6* (Exon 8

1F: 5′-GGA GTG CTA TTG AGA CAC TGG TA-3′) with the reverse *FGFR1* RACE 1R (5′-AGA AGA ACC CCA GAG TTC ATG GA-3′). These primers were also used to amplify the genomic breakpoint.

Fluorescence In Situ Hybridization Analysis

To our knowledge, no other cases have been reported with a t(1;9)(p34;q34) or t(8;12)(p11;q15). For Case 1, FISH was performed using the Vysis ES BCR-ABL probes as described (Chase et al., 1997). Previous analysis suggested disruption of FGFR1 in Case 2 (Sohal et al., 2001).

RESULTS

Characterization of the t(1;9)

FISH using the Vysis ES probes that immediately flank the *ABL* gene hybridized to the normal copy of chromosome 9, also the der(9) and the der(1), indicating that the translocation targeted *ABL*; however, no *BCR-ABL* fusion was seen by FISH or RT-PCR. All *ABL* fusions reported to date result in the partner gene fusing to *ABL* exons a2 or a3. To identify the t(1;9) partner, we, therefore, performed 5′-RACE PCR from both these exons. These initial attempts failed despite the fact that normal 5′ *ABL* sequence was readily obtained (data not shown). 5′RACE primers were subsequently designed in *ABL* exon 4 (accession number M14752). Sequencing of the products revealed several clones in which exon 9 *SFPQ* (accession number X70944) was fused in frame to *ABL* exon 4 (a4) (Fig. 1a).

Confirmation of the SFPQ-ABL Fusion

The presence of the *SFPQ-ABL* fusion was confirmed initially by RT-PCR. The reciprocal *ABL-SFPQ* product was also detectable by single step PCR. Cloning and sequencing revealed that the reciprocal product fuses *ABL* exon 3 (a3) to an alternative RNA isoform (see GenBank accession number 4685098) of *SFPQ* (Fig. 1a). As shown in Figure 1b, the *SFPQ-ABL* fusion was specifically amplified from patient cells by single step PCR, but was not detectable in normal controls. To further confirm the presence of the fusion the genomic breakpoint was amplified by long PCR, cloned, and sequenced. The genomic sequence near the breakpoint on the derivative chromosome 1 was 3′ of exon 10 of *SFPQ* (GenBank accession number NM_005066.1) and was fused to sequence derived from intron 3 of *ABL* (Fig. 1c). The chimeric fusion mRNA is formed through splicing to exon 9 (nucleotide 2072) of the *SFPQ* gene (NM_005066.1). The *SFPQ-ABL* fusion mRNA is predicted to encode a 1609 amino acid protein of 174 KDa that retains the N-terminal proline rich region and the two RNA recognition motifs (RRM) from SFPQ fused to part of the SH2 domain of ABL (Fig. 1d).

Characterization of the t(8;12)

FGFR1 had previously been reported as rearranged by FISH in this patient (Sohal et al., 2001). Because the breakpoints for all cases involving the *FGFR1* gene thus far reported are tightly clustered within or close to exon 9, we investigated the possibility that *FGFR1* was rearranged by 5′-RACE PCR. As shown on Figure 2a, sequencing of the rearranged 5′RACE band revealed a fusion between *CPSF6* intron 8 (accession number NM_007007) and *FGFR1* at nucleotide 1272 from ATG (accession number M34185).

Confirmation of the *CPSF6-FGFR1* Fusion

To confirm the *CPSF6-FGFR1* fusion, a 224 bp product was specifically amplified from patient cDNA but not from control DNA. The genomic breakpoint product of 543 bp was coamplified in the same reaction and sequenced (Figs. 2b and 2c) The *CPSF6* breakpoint within intron 8 (Fig. 2c) results in the incorporation of part of the *CPSF6* intron in the fusion mRNA; the *CPSF6* intronic fragment maintains the reading frame with *FGFR1* and encodes 11 amino acids.

A reciprocal *FGFR1-CPSF6* fusion product was not detected with either single step or nested PCR (Fig. 2b). The *CPSF6-FGFR1* fusion mRNA is predicted to encode a 895 amino acid protein of 97 KDa, that retains the N-terminal region and the RNA recognition motif (RRM) from CPSF6 fused to the entire tyrosine kinase domain of FGFR1 (Fig. 2d).

DISCUSSION

The *SFPQ-ABL* fusion is unique in that it is the first description of an *ABL* fusion to involve the fourth exon of *ABL* (exon a4). This is a structurally unusual ABL fusion protein as amino acids W127-K183 (accession number M14752) of the SH2 domain are not included whereas all previously described ABL fusion proteins retain the whole SH2 domain. Mouse models with BCR-ABL constructs in which the SH2 domain was inactivated have shown the fusion protein retained the ability to cause leukemia in mice. Under conditions in

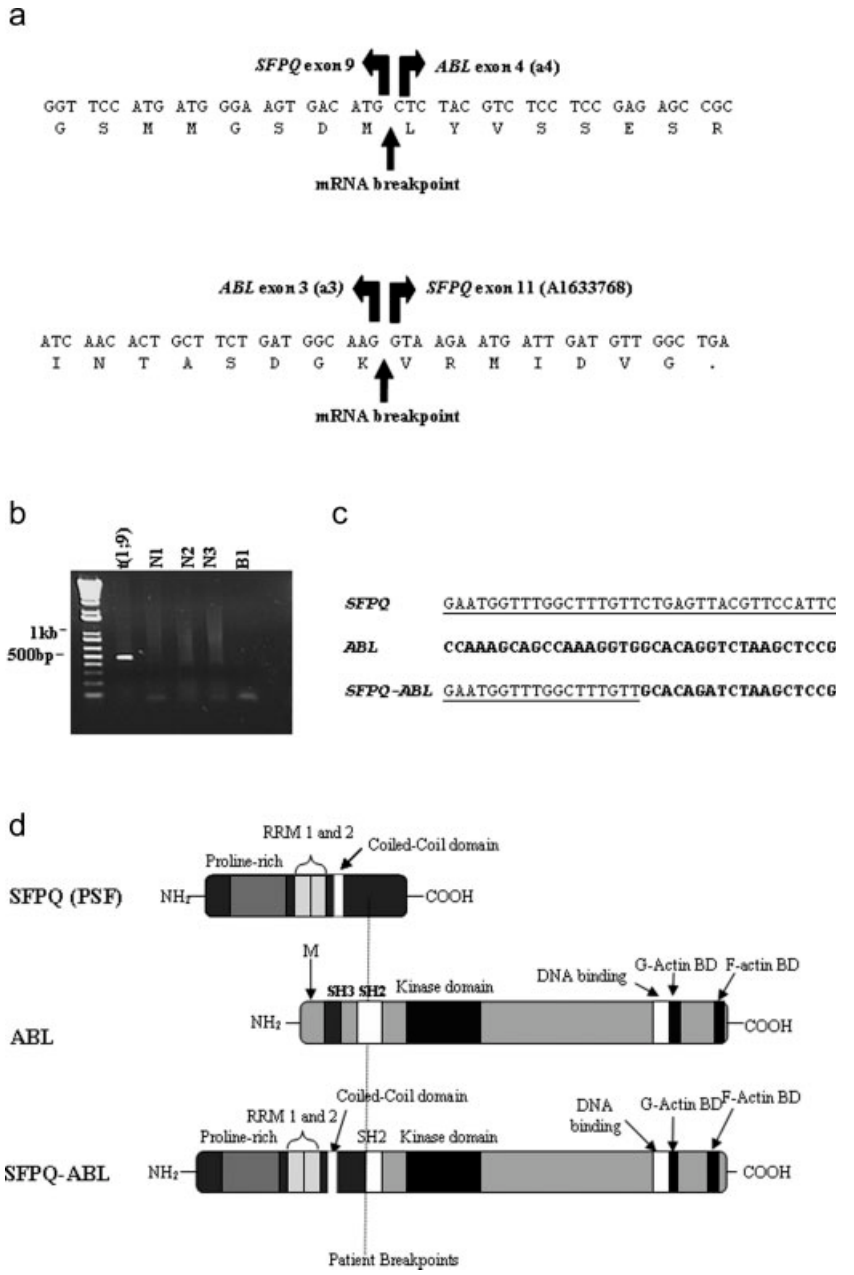


Figure 1. A: Sequence of *SFPQ-ABL* and reciprocal *ABL-SFPQ* mRNA fusions; (B) amplification of *SFPQ-ABL* by single step RT-PCR from the t(1;9) patient but not normal controls (N1-N3). B1 is a blank; (C) sequences surrounding the genomic breakpoints in *SFPQ* (underlined) and *ABL*; (D) schematic representation of SFPQ, ABL, and the SFPQ-ABL fusion protein. The breakpoint within ABL for the great majority (M) of BCR-ABL cases is upstream of the SH3 and SH2 domains and is indicated by an arrow.

which P210 and P190 forms of BCR-ABL cause fatal CML within 4 weeks, P210 SH2 mutants produced CML-like disease in some mice after a significant delay, with others developing a B-lymphoid leukemia (Roumiantsev et al., 2001).

SFPQ is a multifunctional protein able to bind single and double stranded DNA as well as RNA (Shav-Tal et al., 2002). SFPQ was first identified in a complex with polypyrimidine tract-binding protein (PTB), and is an essential factor for RNA splicing. SFPQ has previously been identified in papillary renal cell carcinoma as the fusion partner of the transcription factor TFE3. As with the SFPQ-ABL fusion, SFPQ-TFE3 includes exons 1–9 of SFPQ (NM_005066.1). SFPQ also binds to NPM-ALK (Crockett et al., 2004) and is

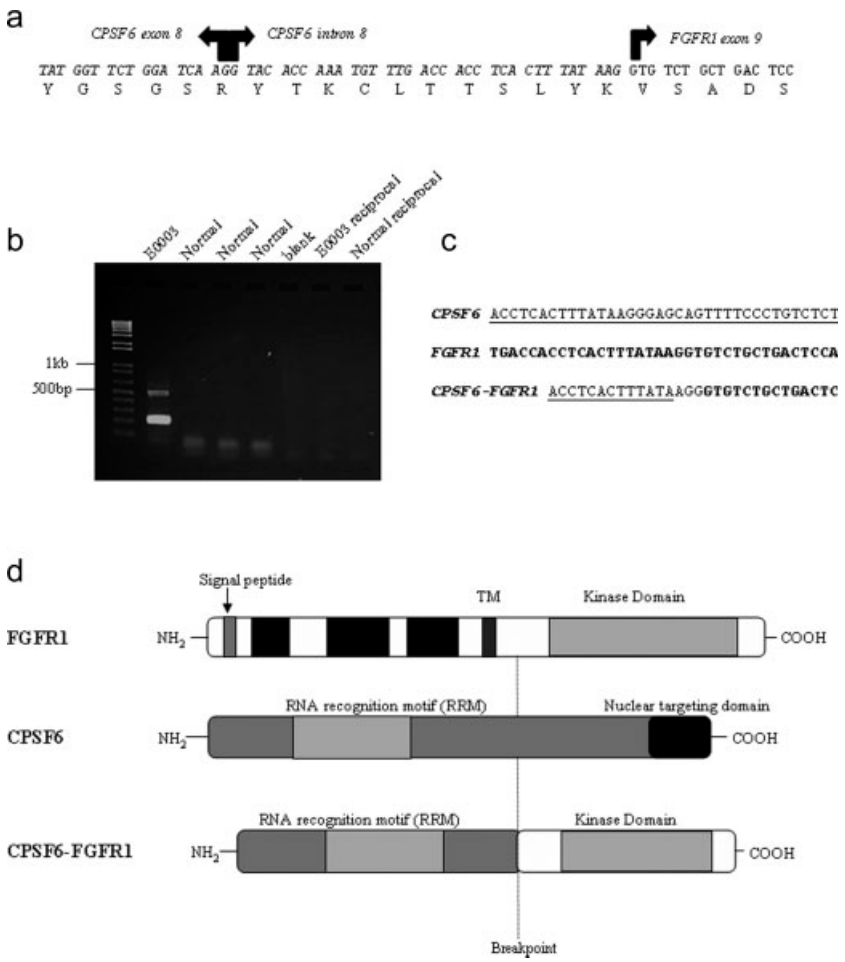


Figure 2. A: Sequence of the *CPSF6-FGFR1* mRNA fusion; (B) amplification of the *CPSF6-FGFR1* fusion specifically from cDNA in the t(8;12) patient (E0003) but not the normal controls. The reciprocal fusion product was not detected (C) Sequences surrounding the genomic breakpoints in *CPSF6* (underlined) and *FGFR1*; (D) domain structure of FGFR1, CPSF6, and the predicted fusion protein. Relevant protein domains are shown: RNA recognition motif (RRM), transmembrane domain, and tyrosine kinase domain for FGFR1.

phosphorylated at Y293F by the NPM-ALK fusion. It has been postulated that SFPQ function is perturbed in NPM-ALK transformed cells (Crockett et al., 2004; Galietta et al., 2007).

As described above, constitutive activation of fusion tyrosine kinases is mediated by an oligomerization domain in the partner protein. The coiled-coil prediction program ncoils (http://www.sanger.ac.uk/cgi-bin/Pfam/swisspfamget.pl?name=P23246) (Lupas et al., 1994) indicated the presence of two coiled-coil domains between SFPQ amino acids 498–529 and amino acids 532–595. It is possible that one or more of these domains is responsible for oligomerization of SFPQ-ABL.

CPSF6 is the ninth partner gene found fused to FGFR1. All these fusions involve FGFR1 exon 9, and therefore disrupt the FRS2 binding domain (Hoch and Soriano, 2006). No recognisable oligomerisation motifs are identifiable in CPSF6, although it has been demonstrated previously that RNA recognition motifs (RRM), such as the one retained in CPSF6-FGFR1, may mediate homodimerization (Simpson et al., 2004; Handa et al., 2006) and it is possible that this domain causes the oligomerization of CPSF6-FGFR1.

CPSF6 has also been found in purified spliceosomes with SFPQ and NonO, linking these proteins to both transcription and processing of pre-mRNAs (Rappsilber et al., 2002; Zhou et al., 2002; Dettwiler et al., 2004). During pre-mRNA processing, CPSF6 binds the highly conserved polyadenylation site AAUAAA that is present in almost every mRNA precursor, as well as interacting with the poly-A binding protein II (PABII) that binds the growing poly-A tail. Together CPSF6 and PABII activate poly (A) polymerase (PAP), by

directing the enzyme to the RNA allowing polyadenylation of the upstream cleavage product (Ruegsegger et al., 1998).

The U1 snRNP has been shown to function in polyadenylation and has been found in complex with SFPQ (Lutz et al., 1998; Liang and Lutz, 2006). Another FGFR1 fusion partner, ZNF198 (Xiao et al., 1998), has been found to immunoprecipitate with SFPQ (Kasyapa et al., 2005). ZNF198 was proposed as a possible scaffold protein bringing RNA-binding proteins together to facilitate pre-mRNA splicing (Kasyapa et al., 2005). In addition, FIP1L1, which fuses to PDGFRα in chronic eosinophilic leukemia (Cools et al., 2003), also plays a role in polyadenylation and pre-mRNA splicing (Preker et al., 1995; Kaufmann et al., 2004). Our findings, thus, identify proteins involved in pre-mRNA processing as a functionally related group that fuse to tyrosine kinase genes in hematological malignancies.

REFERENCES

Belloni E, Trubia M, Gasparini P, Micucci C, Tapinassi C, Confalonieri S, Nuciforo P, Martino B, Lo-Coco F, Pelicci PG, Di Fiore PP. 2005. 8p11 myeloproliferative syndrome with a novel t(7;8) translocation leading to fusion of the FGFR1 and TIF1 genes. Genes Chromosomes Cancer 42:320–325.

Baxter EJ, Hochhaus A, Bolufer P, Reiter A, Fernandez JM, Senent L, Cervera J, Moscardo F, Sanz MA, Cross NC. 2002. The t(4;22)(q12;q11) in atypical chronic myeloid leukaemia fuses BCR to PDGFRA. Hum Mol Genet 11:1391–1397.

Charest A, Kheifets V, Park J, Lane K, McMahon K, Nutt CL, Housman D. Oncogenic targeting of an activated tyrosine kinase to the Golgi apparatus in a glioblastoma. 2003. Proc Natl Acad Sci USA 100:916–921.

Chase A, Grand F, Zhang JG, Blackett N, Goldman J, Gordon M. 1997. Factors influencing the false positive and negative rates of BCR-ABL fluorescence in situ hybridization. Genes Chromosomes Cancer 18:246–253.

Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J, Kutok J, Clark J, Galinsky I, Griffin JD, Cross NC, Tefferi A, Malone J, Alam R, Schrier SL, Schmid J, Rose M, Vandenberghe P, Verhoef G, Boogaerts M, Wlodarska I, Kantarjian H, Marynen P, Coutre SE, Stone R, Gilliland DG. 2003. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. N Engl J Med 348:1201–1214.

Crockett DK, Lin Z, Elenitoba-Johnson KS, Lim MS. 2004. Identification of NPM-ALK interacting proteins by tandem mass spectrometry. Oncogene 23:2617–2629.

Cross NCP, Reiter A. 2006. BCR-ABL negative chronic myeloid leukaemia. In: Melo JV, Goldman JM, editors. Hematologic Malignancies: Myeloproliferative disorders. New York: Springer, pp. 219–234.

Dettwiler S, Aringhier C, Cardinale S, Keller W, Barabino SM. 2004. Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein–protein interactions, and subcellular localization. J Biol Chem 279:35788–35797.

De Keersmaecker K, Graux C, Odero MD, Mentens N, Somers R, Maertens J, Wlodarska I, Vandenberghe P, Hagemeijer A, Marynen P, Cools J. 2005. Fusion of EML1 to ABL1 in T-cell acute lymphoblastic leukemia with cryptic t(9;14)(q34;q32). Blood 105:4849–4852.

Delaval B, Letard S, Lelievre H, Chevrier V, Daviet L, Dubreuil P, Birnbaum D. 2005. Oncogenic tyrosine kinase of malignant hemopathy targets the centrosome. Cancer Res 65:7231–7240.

Demiroglu A, Steer EJ, Heath C, Taylor K, Bentley M, Allen SL, Koduru P, Brody JP, Hawson G, Rodwell R, Doody ML, Carnicero F, Reiter A, Goldman JM, Melo JV, Cross NC. 2001. The t(8;22) in chronic myeloid leukemia fuses BCR to FGFR1: Transforming activity and specific inhibition of FGFR1 fusion proteins. Blood 98:3778–3783.

Galietta A, Gunby RH, Redaelli S, Stano P, Carniti C, Bachi A, Tucker PW, Tartari CJ, Huang CJ, Colombo E, Pulford K, Puttini M, Piazza RG, Ruchatz H, Villa A, Donella-Deana A, Marin O, Perrotti D, Gambacorti-Passerini C. 2007. NPM/ALK binds and phosphorylates the RNA/DNA binding protein PSF in anaplastic large cell lymphoma. Blood 110:2600–2609.

Golub TR. 1997. TEL gene rearrangements in myeloid malignancy. Hematol Oncol Clin North Am 11:1207–1220.

Grand EK, Grand FH, Chase AJ, Ross FM, Corcoran MM, Oscier DG, Cross NC. 2004. Identification of a novel gene, FGFR1OP2, fused to FGFR1 in 8p11 myeloproliferative syndrome. Genes Chromosomes Cancer 40:78–83.

Graux C, Cools J, Melotte C, Quentmeier H, Ferrando A, Levine R, Vermeesch JR, Stul M, Dutta B, Boeckx N, Bosly A, Heimann P, Uyttebroeck A, Mentens N, Somers R, MacLeod RA, Drexler HG, Look AT, Gilliland DG, Michaux L, Vandenberghe P, Wlodarska I, Marynen P, Hagemeijer A. 2004. Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. Nat Genet 36:1084–1089.

Groffen J, Stephenson JR, Heisterkamp N, de Klein A, Bartram CR, Grosveld G. 1984. Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell 36:93–99.

Guasch G, Mack GJ, Popovici C, Dastugue N, Birnbaum D, Rattner JB, Pebusque MJ. 2000. FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33). Blood 95:1788–1796.

Guasch G, Popovici C, Mugneret F, Chaffanet M, Pontarotti P, Birnbaum D, Pébusque MJ. 2003. Endogenous retroviral sequence is fused to FGFR1 kinase in the 8p12 stem-cell myeloproliferative disorder with t(8;19)(p12;q13.3). Blood 101:286–288.

Handa N, Kukimoto-Niino M, Akasaka R, Kishishita S, Murayama K, Terada T, Inoue M, Kigawa T, Kose S, Imamoto N, Tanaka A, Hayashizaki Y, Shirouzu M, Yokoyama S. 2006. The crystal structure of mouse Nup35 reveals atypical RNP motifs and novel homodimerization of the RRM domain. J Mol Biol 363:114–124.

Hoch RV, Soriano P. 2006. Context-specific requirements for Fgfr1 signaling through Frs2 and Frs3 during mouse development. Development 133:663–673.

Kasyapa CS, Kunapuli P, Cowell JK. 2005. Mass spectroscopy identifies the splicing-associated proteins, PSF, hnRNP H3, hnRNP A2/B1, and TLS/FUS as interacting partners of the ZNF198 protein associated with rearrangement in myeloproliferative disease. Exp Cell Res 309:78–85.

Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. 2004. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. EMBO J 23:616–626.

Liang S, Lutz CS. 2006. p54nrb is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. RNA 12:111–121.

Lupas A, Van Dyke M, Stock J. 1997. Predicting coiled coils from protein sequences. Science 252:1162–1164.

Lutz CS, Cooke C, O'Connor JP, Kobayashi R, Alwine JC. 1998. The snRNP-free U1A (SF-A) complex(es): Identification of the largest subunit as PSF, the polypyrimidine-tract binding protein-associated splicing factor. RNA 4:1493–1499.

Popovici C, Zhang B, Gregoire MJ, Jonveaux P, Lafage-Pochitaloff M, Birnbaum D, Pebusque MJ. 1999. The t(6;8)(q27;p11) translocation in a stem cell myeloproliferative disorder fuses a novel gene, FOP, to fibroBLAST growth factor receptor 1. Blood 93:1381–1389.

Preker PJ, Lingner J, Minvielle-Sebastia L, Keller W. 1995. The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. Cell. 81:379–389.

Rappsilber J, Ryder U, Lamond AI, Mann M. 2002. Large-scale proteomic analysis of the human spliceosome. Genome Res 12:1231–1245.

Roumiantsev S, de Aos IE, Varticovski L, Ilaria RL, Van Etten RA. 2001. The src homology 2 domain of Bcr/Abl is required for efficient induction of chronic myeloid leukemia-like disease in mice but not for lymphoid leukemogenesis or activation of phosphatidylinositol 3-kinase. Blood 97:4–13.

Roumiantsev S, Krause DS, Neumann CA, Dimitri CA, Asiedu F, Cross NC, Van Etten RA. 2004. Distinct stem cell myeloproliferative/T lymphoma syndromes induced by ZNF198-FGFR1 and

BCR-FGFR1 fusion genes from 8p11 translocations. Cancer Cell 5:287–298.

Ruegsegger U, Blank D, Keller W. 1998. Human pre-mRNA cleavage factor I-m is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. Mol Cell 1:243–253.

Shav-Tal Y, Zipori D. 2002. PSF and p54(nrb)/NonO—Multi-functional nuclear proteins. FEBS Lett 531:109–114 (review).

Simpson PJ, Monie TP, Szendroi A, Davydova N, Tyzack JK, Conte MR, Read CM, Cary PD, Svergun DI, Konarev PV, Curry S, Matthews S. 2004. Structure and RNA interactions of the N-terminal RRM domains of PTB. Structure 12:1631–1643.

Sohal J, Chase A, Mould S, Corcoran M, Oscier D, Iqbal S, Parker S, Welborn J, Harris RI, Martinelli G, Montefusco V, Sinclair P, Wilkins BS, van den Berg H, Vanstraelen D, Goldman JM, Cross NC. 2001. Identification of four new translocations involving FGFR1 in myeloid disorders. Genes Chromosomes Cancer 32:155–163.

Walz C, Chase A, Schoch C, Weisser A, Schlegel F, Hochhaus A, Fuchs R, Schmitt-Graff A, Hehlmann R, Cross NC, Reiter A. 2005. The t(8;17)(p11;q23) in the 8p11 myeloproliferative syndrome fuses MYO18A to FGFR1. Leukemia 19:1005–1009.

Xiao S, Nalabolu SR, Aster JC, Ma J, Abruzzo L, Jaffe ES, Stone R, Weissman SM, Hudson TJ, Fletcher JA. 1998. FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. Nat Genet 18:84–87.

Zhou Z, Sim J, Griffith J, Reed R. Purification and electron microscopic visualization of functional human spliceosomes. 2002. Proc Natl Acad Sci USA 99:12203–12207.

Frequent *CBL* mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms

*Francis H. Grand,[1] *Claire E. Hidalgo-Curtis,[1] Thomas Ernst,[1] Katerina Zoi,[2] Christine Zoi,[2] Carolann McGuire,[3] Sebastian Kreil,[1] Amy Jones,[1] Joannah Score,[1] Georgia Metzgeroth,[4] David Oscier,[5] Andrew Hall,[6] Christian Brandts,[7] Hubert Serve,[7] Andreas Reiter,[4] Andrew J. Chase,[1] and Nicholas C. P. Cross[1]

[1]Wessex Regional Genetics Laboratory, Salisbury, and Human Genetics Division, School of Medicine, University of Southampton, Southampton, United Kingdom; [2]Haematology Research Laboratory, Biomedical Research Foundation, Academy of Athens, Athens, Greece; [3]Inflammation, Infection and Repair Division, School of Medicine, University of Southampton, Southampton, United Kingdom; [4]III Medizinische Universitätsklinik, Fakultät für Klinische Medizin Mannheim der Universität Heidelberg, Mannheim, Germany; [5]Department of Haematology, Royal Bournemouth Hospital, Bournemouth, United Kingdom; [6]Northern Institute for Cancer Research, Newcastle upon Tyne, United Kingdom; and [7]Department of Medicine, Hematology & Oncology, University of Frankfurt, Frankfurt, Germany

Recent evidence has demonstrated that acquired uniparental disomy (aUPD) is a novel mechanism by which pathogenetic mutations in cancer may be reduced to homozygosity. To help identify novel mutations in myeloproliferative neoplasms (MPNs), we performed a genome-wide single nucleotide polymorphism (SNP) screen to identify aUPD in 58 patients with atypical chronic myeloid leukemia (aCML; n = 30), *JAK2* mutation–negative myelofibrosis (MF; n = 18), or *JAK2* mutation–negative polycythemia vera (PV; n = 10). Stretches of homozygous, copy neutral SNP calls greater than 20Mb were seen in 10 (33%) aCML and 1 (6%) MF, but were absent in PV. In total, 7 different chromosomes were involved with 7q and 11q each affected in 10% of aCML cases. *CBL* mutations were identified in all 3 cases with 11q aUPD and analysis of 574 additional MPNs revealed a total of 27 *CBL* variants in 26 patients with aCML, myelofibrosis or chronic myelomonocytic leukemia. Most variants were missense substitutions in the RING or linker domains that abrogated CBL ubiquitin ligase activity and conferred a proliferative advantage to 32D cells overexpressing FLT3. We conclude that acquired, transforming *CBL* mutations are a novel and widespread pathogenetic abnormality in morphologically related, clinically aggressive MPNs. (Blood. 2009;113: 6182-6192)

Introduction

Myeloproliferative neoplasms (MPNs) are clonal hematopoietic stem cell disorders characterized by overproliferation of one or more myeloid cell lineages in the bone marrow and increased numbers of mature and immature myeloid cells in the peripheral blood. Excess proliferation is frequently associated with splenomegaly and cardiovascular complications as well as increased risk of transformation to acute leukemia. MPNs are categorized into subtypes based on specific morphologic, hematologic, and laboratory parameters, the best characterized being the 4 so-called classic MPNs: polycythemia vera (PV), essential thrombocythemia (ET), primary myelofibrosis (MF), and chronic myeloid leukemia (CML).[1] In addition, several atypical MPNs are recognized, some of which show both dysplastic and proliferative features, such as atypical, *BCR-ABL* negative CML (aCML).[2]

MPNs are associated with acquired, activating mutations or gene fusions of tyrosine kinases, abnormalities that are believed to be critical drivers of excess proliferation as a result of deregulated or constitutive signaling.[3] The 2 most prominent examples are *BCR-ABL* in CML[4] and the V617F *JAK2* mutation in PV, ET, and MF,[5-8] but more than 40 variant tyrosine kinase fusions have been identified in MPNs as well as other mutations in *JAK2* and *FLT3*.[2] Activating mutations have been described in components that signal upstream (eg, *MPL*) or downstream (eg, *NRAS*) of tyrosine kinases[9,10]; however, the molecular pathogenesis of the majority of atypical MPNs and approximately 50% of ET and MF cases remains obscure.

V617F *JAK2* was initially identified by several different routes, one of which was based on the observation that many PV patients show evidence of acquired uniparental disomy (aUPD) at chromosome 9p.[11,12] Regions of aUPD exhibit loss of heterozygosity (LOH) compared with constitutional DNA without change of copy number and arise by mitotic recombination followed by selection for one of the products. After the initial observations in PV, it has emerged that aUPD is common in both hematologic and epithelial malignancies, and is associated with known oncogenic changes in a variety of genes within the affected regions.[13] In this study we set out to determine whether aUPD characterizes MPNs of unknown molecular etiology and, if so, whether it could be used as a tool to help identify novel driver mutations.

Methods

Patients

Peripheral blood or bone marrow samples were received from patients diagnosed with an MPN or other hematologic malignancy according to

Submitted December 13, 2008; accepted April 7, 2009. Prepublished online as *Blood* First Edition paper, April 22, 2009; DOI 10.1182/blood-2008-12-194548.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

*F.H.G. and C.E.H.-C. contributed equally to this work.

standard morphologic, hematologic, and laboratory criteria. Clinical data were available from a subset of these cases. The study was approved by the Internal Review Boards from participating institutions and informed consent was obtained in accordance with the Declaration of Helsinki.

SNP array analysis

DNA labeling and hybridization to Affymetrix 50k *Xba*I chips was performed at the Deutsches Ressourcenzentrum für Genomforschung (RZPD, Berlin, Germany). Raw data were imported into the Affymetrix GeneChip Operating Software, analyzed using Affymetrix (High Wycombe, United Kingdom) GeneChip Genotyping Analysis Software (GTYPE 4.1) and copy number analysis tool (CNAT). Data were exported to custom-designed spreadsheets that display loss of heterozygosity and copy number changes in ideogram format. Data were also analyzed using Affymetrix Genotyping Console (version 2.1). Overall, a median of 98.2% (range, 91.5%-99.6%) of SNPs gave readable calls.

Mutation analysis

Detection of mutations by high resolution melting (HRM) analysis was performed as described[14] using a Rotor-Gene 6000 (Corbett Life Sciences, St Neots, United Kingdom). Direct sequencing of polymerase chain reaction (PCR) products was performed by standard techniques on an ABI 3130 Genetic Analyzer (Applied Biosystems, Warrington, United Kingdom) and analyzed using Mutation Surveyor (SoftGenetics, State College, PA). *CBL* mutations are numbered relative to the ATG (A = 1) in the Ensembl cDNA sequence ENSG00000110395. A dropping factor (relative intensity drop of the wild-type allele peak relative to that seen in a concurrently run normal sample) of 60% or more was considered indicative of a biallelic mutation whereas a dropping factor of less than 60% was considered as monoallelic. Primer sequences for *CBL* exon 8 were CBL_i7f (5′-tgtggtttcactttaaaccctgga-3′) and CBL_i8r (5′-gccaggccacccttgtatc-3′) and for *CBL* exon 9 were CBL_i8f (5′-ggcctggcttttgggttagg-3′) and CBL_i9r (5′-cacaatggattttgccagtctcc-3′). *CBL* reverse-transcriptase PCR (RT-PCR) was performed on random hexamer–reverse transcribed RNA using primers CBLe7F: 5′-gaatcctgatctgactggcttatg-3′ and CBLe10R 5′-gctgcggcagaaggtcaagt-3′. Other primer sequences are available on request.

Microsatellite analysis

DNA samples were amplified with a series of fluorescently labeled primer pairs flanking highly polymorphic microsatellite markers on chromosome 11q as described.[10] aUPD was scored if 4 or more consecutive markers encompassing *CBL* were homozygous without change in allelic copy number ($P < .05$ of this occurring in the absence of aUPD based on published rates of heterozygosity). Copy number at 11q was estimated by multiplex ligation probe amplification (MLPA) using the Marfan probeset 1 (P065; MRC Holland, Amsterdam, The Netherlands), a kit that contains four 11q control probes including one at 11q23.3.

Constructs

pMY-wtCBL and pMY-CBLR420Q express wild-type and R420Q CBL, respectively, along with green fluorescent protein (GFP).[15] Other *CBL* mutant constructs were derived from pMY-wtCBL using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). pAL/FLT3 and pCMV–HA ubiquitin have also been described previously.[15]

Cell lines

The IL-3–dependent murine myeloid cell 32Dcl3 (32D) was purchased from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (Braunschweig, Germany) and was maintained in 90% RPMI 1640 medium with 10% fetal bovine serum (FBS) and 10% WEHI-3B conditioned medium (WEHI). All cells were maintained in a humidified incubator at 37°C and 5% carbon dioxide. 32D cells were stably transduced by electroporation with wt-FLT3 in a pAL vector as previously described[16] and maintained in 15 ng/mL blastocidin (Invitrogen, Paisley United Kingdom). The Platinum-E (Plat-E) retrovirus packaging cell line[17] was transfected using the calcium phosphate method with 5 μg CBL construct DNA: 10^6 32D/wtFLT3 cells were cocultivated with the Plat-E cells for 48 hours in a volume of 2 mL (50% RPMI, 50% DMEM, 10% FBS, 5% WEHI). GFP positive cells were flow sorted to highest possible purity using a BD FACSAria cell sorter (Becton Dickinson, Stanford, CA). If insufficient cells were obtained for experiments after flow sorting, the cells were maintained in culture and resorted to greater than 95% purity.

Growth factor withdrawal assays

Cell growth was compared with wt-FLT3, vector-only transfected cells and the parental 32D line in stably transfected cells. Cells (10^5/mL) were grown over 3 days in triplicate for each *CBL* mutant and a wt-*CBL* control in 96-well plates. Cultures were analyzed daily for proliferation using the Aqueous One Solution Cell Proliferation Assay kit (MTS assay; Promega, Southampton, United Kingdom). All experiments were performed at least 3 times independently and results were analyzed using GraphPad Prism 4 software (GraphPad Software, San Diego, CA).

Ubiquitin ligase activity

The ubiquitin ligase activity of *CBL* mutants in comparison to wild-type *CBL* was determined by transiently transfecting 10^7 32D/wtFLT3 cells expressing wild-type or *CBL* mutants by electroporation with 20 μg pCMV-HA ubiquitin. Cells were maintained in culture with WEHI for 12 hours after transfection before being washed, serum deprived overnight, exposed to FLT3 ligand for 10 minutes, and lysed in buffer as described.[18] Lysates were incubated with a FLT3-specific polyclonal antibody overnight at 4°C. Protein G agarose beads (100 μL) were added and incubated at 4°C for 4 hours. The beads were washed 4 times in ice-cold radioimmunoprecipitation assay (RIPA) lysis buffer. The supernatant was mixed with sodium dodecyl sulphate (SDS) sample loading buffer before immunoblot analyses with antibodies against HA-ubiquitin (sc-805; Santa Cruz Biotechnology, Santa Cruz, CA), CBL (2111C3a; Abcam, Cambridge, United Kingdom), and FLT3 (sc-479; Cell Signaling Technology, Danvers, MA).

Results

SNP array analysis

We performed genome-wide SNP analysis on leukocyte DNA extracted from 58 MPD patients: aCML (n = 30), V617F *JAK2* negative MF (n = 18), and *JAK2* mutation negative PV/idiopathic erythrocytosis (n = 10). As comparators, we also analyzed cases of CML in transformation to blast crisis (n = 30, of which 20 were myeloid and 10 lymphoid) and chronic lymphocytic leukemia (CLL; n = 20). We conservatively defined likely regions of aUPD as at least 20 Mb of absent or significantly reduced heterozygous SNP calls (ie, < 5% SNPs called as heterozygous) running to a telomere without change in copy number (Figure 1A). Such large regions are uncommon in healthy individuals.[19] Strikingly, we found that aUPD was common in aCML with 10 (33%) of cases affected. In contrast, aUPD was only seen in a single case of MF and was not seen in other patient subgroups. In total, 7 different chromosomes were affected, with abnormalities at 7q and 11q each seen recurrently in 10% of aCML cases. The affected regions ranged in size from 29 Mb to a complete copy of chromosome 13 (Table 1).

Association between aUPD and known acquired oncogenic mutations

Although we did not have constitutional DNA to formally confirm that the regions of homozygosity were acquired, in 1 individual we found a marked reduction in the proportion of 11q aUPD positive cells on comparison of mononuclear cells and granulocytes,

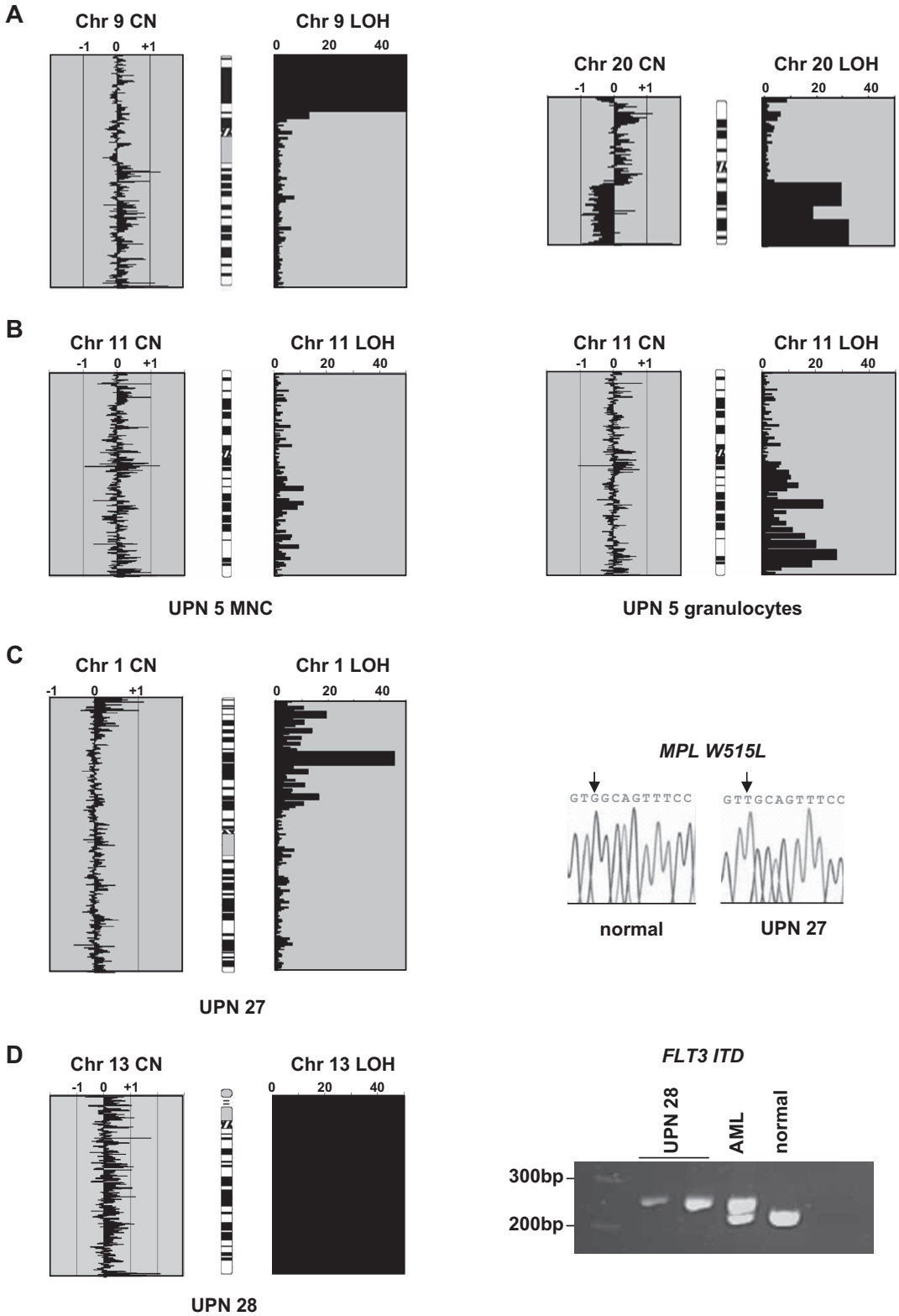


Figure 1. Single nucleotide polymorphism array results and initial mutation analysis. (A) Array analysis of 2 control patients showing copy number (CN) and loss of heterozygosity (LOH) outputs. The x-axis for the CN plots show (copy number)$^{-2}$ as determined by CNAT, with values approaching or exceeding -1 indicating deletions and greater than $+1$ indicating amplifications. The x-axis for LOH shows $\log_{10}$ *P* values (ie, 20 indicates 10^{-20}). The 2 panels on the left are from a polycythemia vera (PV) case with a homozygous V617F *JAK2* mutation. There is no gross copy number change (individual datapoints spread around the zero line), but a large block of homozygous SNP calls at 9p indicative of aUPD. The 2 panels on the right are from a chronic myeloid leukemia (CML) blast crisis case and show LOH accompanied by a 20q deletion. (B) For case UPN 5, a large block of 11q aUPD is apparent on analysis of granulocytes but almost completely absent from mononuclear cells (MNCs) extracted from the same sample. (C) Chromosome 1p aUPD in case UPN 27 is associated with a biallelic G>T *MPL* mutation that is predicted to result in a W515L substitution. (D) Whole chromosome 13 isodisomy in case UPN 28 associated with a homozygous *FLT3* ITD. Two controls are shown: an acute myeloid leukemia (AML) patient previously known to have a heterozygous *FLT3* ITD and a healthy individual.

Table 1. Chromosomal regions and cases affected by aUPD

Disease category	No. with aUPD / no. analyzed	Region affected	Size of affected region; Mb
aCML	10/30	7q11.21-qter	95
		7q11.22-qter	87
		7q22.3-qter	53
		11q12.1-qter	73
		11q13.1-qter	69
		11q13.5-qter	57
		13pter-qter	114
		17cen-qter	55
		20q11.22-qter	30
		21q21.1-qter	29
MF	1/18	1pter-1p21.3	95
PV	0/10		
CML-BC	0/30		
CLL	0/20		

aUPD indicates acquired uniparental disomy; aCML, atypical chronic myeloid leukemia; MF, myelofibrosis; PV, polycythemia vera; CML-BC, CML in transformation to blast crisis; and CLL, chronic lymphocytic leukemia.

consistent with an acquired abnormality (Figure 1B). Some of the affected regions harbor genes known to be relevant to hematologic malignancies and therefore we assessed their mutational status. Sequencing of the *MPL* gene in the case UPN 27 with 1p aUPD revealed a homozygous G>T change (Figure 1C) that results in the previously reported W515L substitution.[9] Case UPN 28 with whole chromosome 13 UPD had a homozygous *FLT3* internal tandem duplication (ITD; Figure 1D). Candidate genes for the abnormalities at 17q and 21q include *NF1* and *RUNX1*, respectively; however, these were not investigated.

11q aUPD is associated with *CBL* mutations

Next, we focused our analysis on the 2 recurrent regions of aUPD at 7q and 11q. Because of the strong association between MPNs and aberrant tyrosine kinase signaling, we analyzed genes encoding tyrosine kinases and associated signaling components. The minimally affected region at chromosome 7q contains 3 tyrosine kinases: *MET*, *EPHA1*, and *EPHB6*. We fully sequenced the coding exons of these genes in the 3 cases with 7q aUPD and 2 normal controls but found no abnormalities. Similarly, we found no sequence changes in exons 11 through 17 of *BRAF*, a gene encoding a downstream component of RAS signaling that is mutated in malignant melanoma and other cancers.[20] There are no tyrosine kinase genes on chromosome 11q, and although there are several genes encoding signal transduction components, *CBL* stood out as a known oncogene that negatively regulates tyrosine kinase signaling.[21] Sequencing of *CBL* in the three 11q aUPD cases revealed that UPN 2 and UPN 5 each had a T>C change at nucleotide 1277 in exon 8 that predicts a L380P substitution. Case UPN 3 had a C>G missense change in exon 9 at nucleotide 1387 that predicts a P417A change (Figure 2). Consistent with the array results, the mutations in cases UPN 3 and UPN 5 were predominant with the residual wild-type alleles only weakly visible. For case UPN 2 both the mutation and level of aUPD was less prominent, presumably due to a greater background of normal cells.

Prevalence and nature of *CBL* mutations

To determine the prevalence of *CBL* mutations, we sequenced exons 8 and 9 (that encode amino acids 366 through 409 and 410 through 477, respectively) in an additional 574 MPN cases. A total of 27 sequence variants were identified in 26 patients, of whom 3 had MF, 10 had chronic myelomonocytic leukemia (CMML), 12 had aCML and 1 had hypereosinophilic syndrome/chronic eosinophilic leukemia (HES/CEL; Tables 2 and 3). To ensure that we were not missing mutations elsewhere in the gene we analyzed all coding exons in 70 *CBL* mutation negative aCML or unclassified, CML-like MPN cases by high resolution melt (HRM) analysis but failed to detect any further sequence changes apart from known polymorphisms. Similarly, no mutations in *CBLB* exons 9 or 10 or *CBLC* exons 7 and 8 (the exons that correspond to *CBL* exons 8 and 9) were identified. The activity of *BCR-ABL* is known to increase during progression of CML from chronic phase to blast crisis[22] and we speculated that abrogation of CBL activity might be at least partially responsible for this increase. HRM analysis of *CBL* exons 8 and 9 in blast crisis cases (n = 31), however, failed to reveal any mutations.

Of the 27 *CBL* variants, 21 (78%) were missense substitutions affecting 12 amino acids (15 different substitutions), 5 (19%) were candidate splicing abnormalities, and 1 (3%) produced a stop codon (Table 3). Most of the changes were novel, although 4 cases had R420Q as previously identified in a single case of AML.[15] Apart from N454D, all residues affected by missense mutations were completely conserved in *CBL* orthologues as well as the 2 other human *CBL* family members (Figure 3). The mutation to a stop codon was found in a patient (UPN 9) who also had the Y371H variant; PCR across the affected regions and sequencing indicated that the 2 changes were on different alleles.

CBL mutations are acquired and rarely seen in conjunction with other known mutations

For 2 cases (UPN 5 and UPN 14) we were able to compare the mutational status of leukocytes and T cells or buccal epithelia cells. In both cases the L380P mutation was only present in leukocytes and was therefore acquired (Figure 4A). A third case (UPN 15) originally presented with V617F *JAK2* negative ET in 1989 but progressed to MF in 2004. Analysis of DNA extracted from sequential bone marrow slides indicated that the R420Q mutation was absent at the ET phase but acquired on transformation (Figure 4B). The 3 *CBL*-positive MF cases tested negative for *MPL* 515 mutations and all *CBL* mutation positive cases were negative for cytogenetic indicators of tyrosine kinase fusions, *FLT3* ITD, activating *NRAS* mutations and V617F *JAK2*, with the exception of 1 case (UPN 14) with MF who tested positive for both L380P *CBL* and V617F *JAK2*. Analysis of retrospective bone marrow slides for this case indicated the presence of V617F *JAK2* at low level initially but absence of L380P *CBL*. The *CBL* mutation was detectable 7 years later and completely displaced V617F after a further 3 years, corresponding with a significant rise is peripheral leukocyte counts (Figure 4C). A further case (UPN 21) presented with systemic mastocytosis before evolving to *CBL* mutation-positive CMML but retrospective samples were unavailable.

Splicing abnormalities

RNA was available for 3 of the 5 cases with potential splicing mutations. Patients UPN 18 and UPN 20 both exhibited double bands by RT-PCR, sequencing of which showed one to be the normal product and the smaller band to result from the complete deletion of exon 8 in each case (Figure 4D), similar to that described in the cell line MOLM-13.[23] Patient UPN 19 showed a single band by RT-PCR that was entirely normal on sequence analysis indicating that the exon 8+4 C>T change does not alter splicing. It seems likely, therefore, that the variant identified in case UPN 6 is also of no consequence. The pathogenicity of the intronic variant seen in case UPN 13 remains unknown.

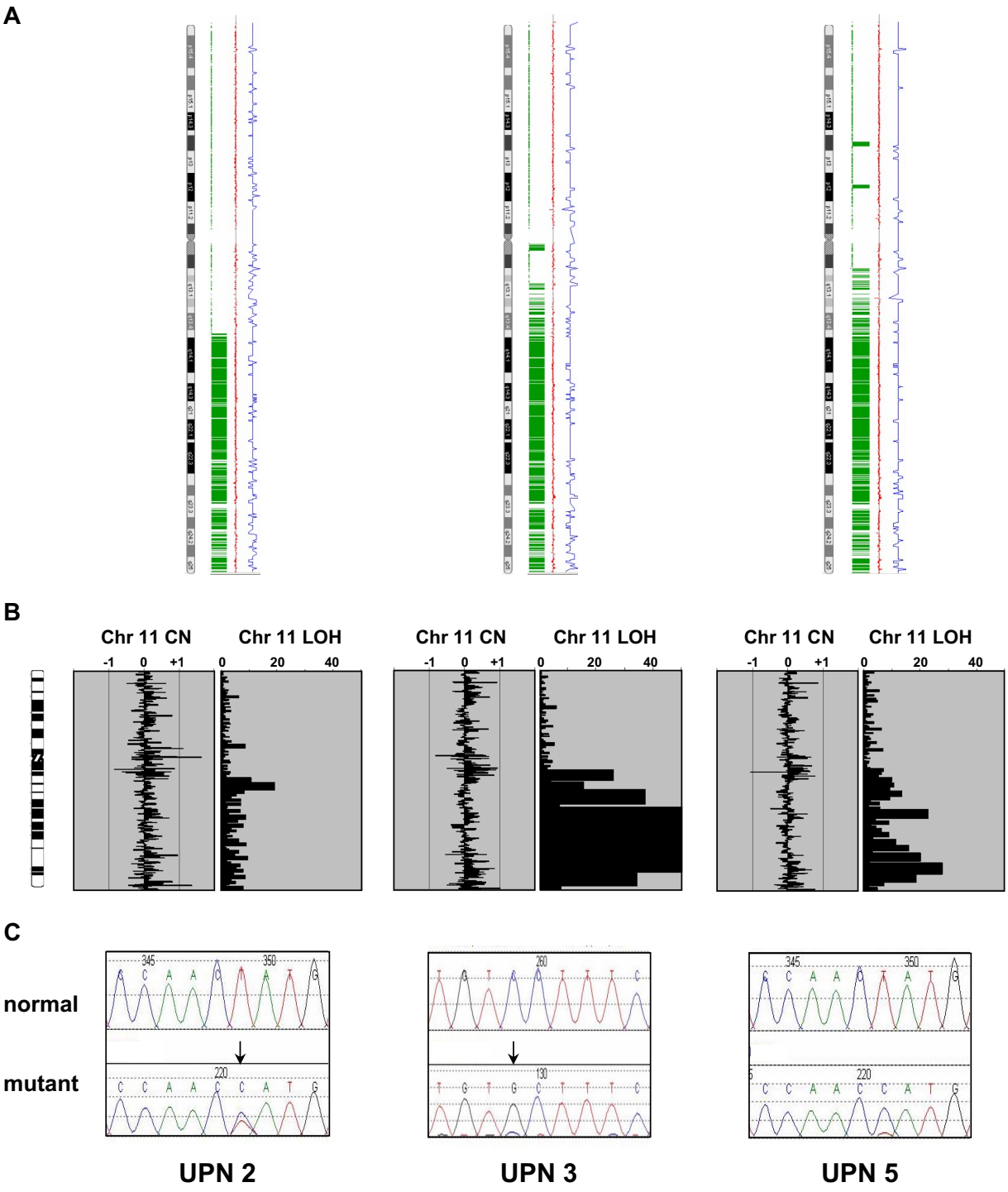


Figure 2. *CBL* mutations in the 3 cases with 11q aUPD. (A) Genotyping Console output indicating homozygous SNP calls in green, copy number state as estimated by a Hidden Markov Model in blue, and smoothed copy number log(2) ratio values in red. (B) Ideograms showing lack of large copy number changes and blocks of homozygosity at 11q. Abbreviations are as in Figure 1. (C) Sequence changes indicated by ↓ in each case compared with normal controls. UPN 2 and UPN 5 each have the c.1139T>C: L380P mutation whereas UPN 3 has c.1249C>G: P417A.

The association between *CBL* mutations and UPD

To examine the association between *CBL* mutations and aUPD in more detail we performed microsatellite analysis on all mutated cases using 9 polymorphic markers spanning chromosome 11q13-qter. Of the 26 cases, 11 displayed patterns that indicated significant tracts of homozygosity, including all 3 with 11q aUPD detected by SNP arrays. Case UPN 9 with 2 mutations was heterozygous for all markers. Of the 11 cases with biallelic *CBL* mutations as judged by sequence analysis, 5 showed complete homozygosity at all microsatellite loci tested and 6 had at least 4 consecutive homozygous calls encompassing the location of the *CBL* gene at 11q23.3 (Figure 5). We performed multiplex ligation probe amplification (MLPA) analysis to determine whether the observed homozygosity was a consequence of a deletion but no copy number changes were detected (data not shown) and therefore it is likely that homozygosity arose by aUPD. The number of

Table 2. Cases analyzed for *CBL* mutations

Disease category	Cases analyzed	Cases with *CBL* variants
PV	74	0
MF	53	3 (6%)
ET	24	0
SM	60	0
CMML	78	10 (13%)
aCML	152	12 (8%)
HES/CEL	96	1 (1%)
CNL	9	0
CML-BC	31	0
Total	577	

PV indicates polycythemia vera; MF, myelofibrosis; ET, essential thrombocythemia; SM, systemic mastocytosis; CMML, chronic myelomonocytic leukemia; aCML, atypical chronic myeloid leukemia; HES/CEL, hypereosinophilic syndrome/chronic eosinophilic leukemia; CNL, chronic neutrophilic leukemia; and CML-BC, CML in transformation to blast crisis.

mutated alleles estimated by sequencing was concordant with the microsatellite analysis in 20 cases (Figure 5). Of those that were discordant, UPN 6 and UPN 25 had significant tracts of 11q homozygosity but had a monoallelic *CBL* mutation whereas UPN 10 and UPN 16 did not show 11q homozygosity but had clear biallelic *CBL* mutations. The reason for these discrepant cases is unclear, but it is possible that the aUPD seen in UPN 6 and UPN 25 targeted another gene at 11q before acquisition of the *CBL* mutation. Homozygosity for the *CBL* mutations in the absence of 11q aUPD in UPN 10 and UPN 16 might be explained by a subcytogenetic deletion of 1 allele.

Table 3. List of *CBL* sequence variants

Patient identifier	Diagnosis	Exon 8 amplicon	Exon 9 amplicon
UPN 1	aCML	c.1127C>T: S376F	N
UPN 2	aCML	c.1139T>C: L380P	N
UPN 3	aCML	N	c.1249C>G: P417A
UPN 4	aCML	N	c.1259G>A: R420Q
UPN 5	aCML	c.1139T>C: L380P	N
UPN 6	aCML	int+5 G>A	N
UPN 7	aCML	c.1224G>C: W408C	N
UPN 8	aCML	N	c.1254C>A: F418L
UPN 9	aCML	c.1111T>C: Y371H	c.1384C>T R462X
UPN 10	aCML	N	c.1249C>G: P417A
UPN 11	aCML	N	c.1259G>T: R420L
UPN 12	aCML	N	c.1259G>A R420Q
UPN 13	CEL	N	int-2 A>G
UPN 14	MF	c.1139T>C: L380P	N
UPN 15	MF	N	c.1259G>A: R420Q
UPN 16	MF	N	c.1259G>A: R420Q
UPN 17	CMML	c.1142G>A: C381Y	N
UPN 18	CMML	c.1227–1227+4 del ggtac	N
UPN 19	CMML	int+4 C>T	N
UPN 20	CMML	int-1 G>C	N
UPN 21	CMML	c.1186T>G: C396G	N
UPN 22	CMML	N	c.1250C>T P417L
UPN 23	CMML	c.1151G>A: C384Y	N
UPN 24	CMML	N	c.1360A>G: N454D
UPN 25	CMML	c.1150T>C: C384R	N
UPN 26	CMML	c.1192C>T: H398Y	N

UPN indicates unique patient number; aCML, atypical chronic myeloid leukemia; CEL, chronic eosinophilic leukemia; MF, myelofibrosis; CMML, chronic myelomonocytic leukemia; and N, normal.

Correlation between *CBL* mutations and clinical features

We compared clinical and hematologic features in patients with or without *CBL* mutations, restricting the analysis to cases with aCML, MF and CMML because these were the subgroups in which *CBL* mutations were most commonly found. Patients with *CBL* mutations had a shorter overall survival and progression-free survival compared with mutation negative cases (overall survival: 33 months vs 39 months; progression-free survival: 22 months vs 32 months) but the differences were not significant (Figure 6). Similarly there was no significant difference in gender distribution, age and standard hematologic parameters between mutation positive and mutation negative cases and no differences emerged when individual disease entities were considered.

Transforming activities of *CBL* mutants

We determined whether selected CBL variants had oncogenic activity by testing whether they could transform the interleukin-3 (IL-3)–dependent cell line 32D to growth factor independence in conjunction with overexpression of wild-type FLT3, as has been described previously for R420Q.[15] We found that S376F, H398Y, P417A, and R420Q alone each conferred a degree of growth factor independence to 32D cells, but this effect was considerably enhanced in 32Ds that overexpressed wild-type FLT3. In contrast, wild-type CBL conferred no growth advantage either with or without FLT3 overexpression. The N454D change was not transforming, indicating that it is probably a rare polymorphism or a pathogenetically unimportant passenger mutation (Figure 7). Reproducible differences in survival between the different mutants were seen, as illustrated in Figure 8A.

Abrogation of E3 ubiquitin ligase activity

It has been shown that FLT3 physically associates with CBL and that, upon ligand stimulation, FLT3 is rapidly ubiquitinylated by wild-type but not R420Q CBL.[15] We found that transforming activity correlated directly with the ubiquitin ligase activity of the CBL variants, as assayed by their ability to transfer HA-tagged ubiquitin to FLT3 after stimulation with FLT3 ligand. The S376F, H398Y, P417A and R420Q mutants showed loss of ubiquitin ligase activity, whereas N454D was comparable to wild-type CBL (Figure 8B).

Discussion

The initial aim of our study was to identify large regions of aUPD and therefore we undertook a relatively low resolution genome scan using 50k SNP arrays. Because we did not have constitutional DNA available from most of our cases, we were not able to unambiguously determine whether the extended blocks of homozygosity that we observed were a consequence of aUPD, constitutional UPD or autozygosity due to consaguinity, however we used a very conservative definition of candidate aUPD regions (at least 20 Mb homozygous SNPs calls running to a telomere) and therefore expected that most or all would be acquired. We found that aUPD was relatively common in aCML, but was uncommon in other MPNs that were negative for known mutations, plus also CML blast crisis and CLL. Six distinct regions of aUPD were identified in aCML, indicating substantial genetic heterogeneity in the genesis of this disorder. However, we identified 2 recurrent regions of aUPD at 7q and 11q that were each seen in 10% of cases indicating the presence of common molecular abnormalities.

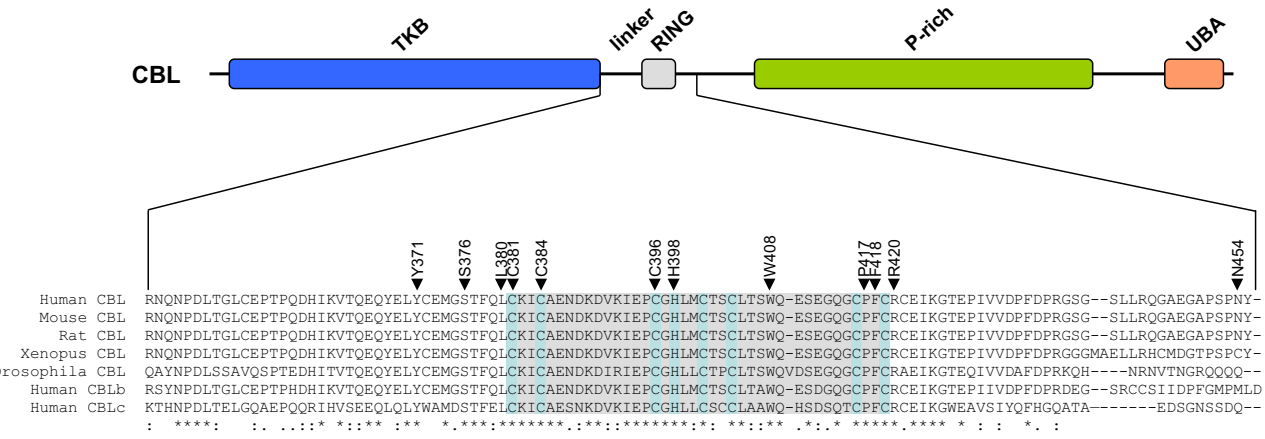


Figure 3. Location of missense mutations. (Top panel) Schematic representation of the CBL protein showing the tyrosine kinase binding (TKB), linker, RING, proline-rich (P-rich), and ubiquitin-associated/leucine zipper (UBA/LZ) domains. (Bottom panel) ClustalW alignment of the entire linker and RING domains plus part of the distal sequence. The RING domain is shaded in gray with the defining C and H residues shaded pale blue. Residues affected by missense mutations found in this study are indicated by ▼.

Our mutation screening focused primarily on tyrosine kinases and associated signal transduction components because of the known association between deregulated tyrosine kinase signaling and MPNs.[3] Although we failed thus far to identify any abnormality on chromosome 7q, we found missense *CBL* mutations in all three 11q aUPD cases. Screening of further MPNs revealed a total of 27 *CBL* variants in 26 patients who had been diagnosed with aCML (n = 12; 8%), MF (n = 3, 6%), CMML (n = 10, 13%), or

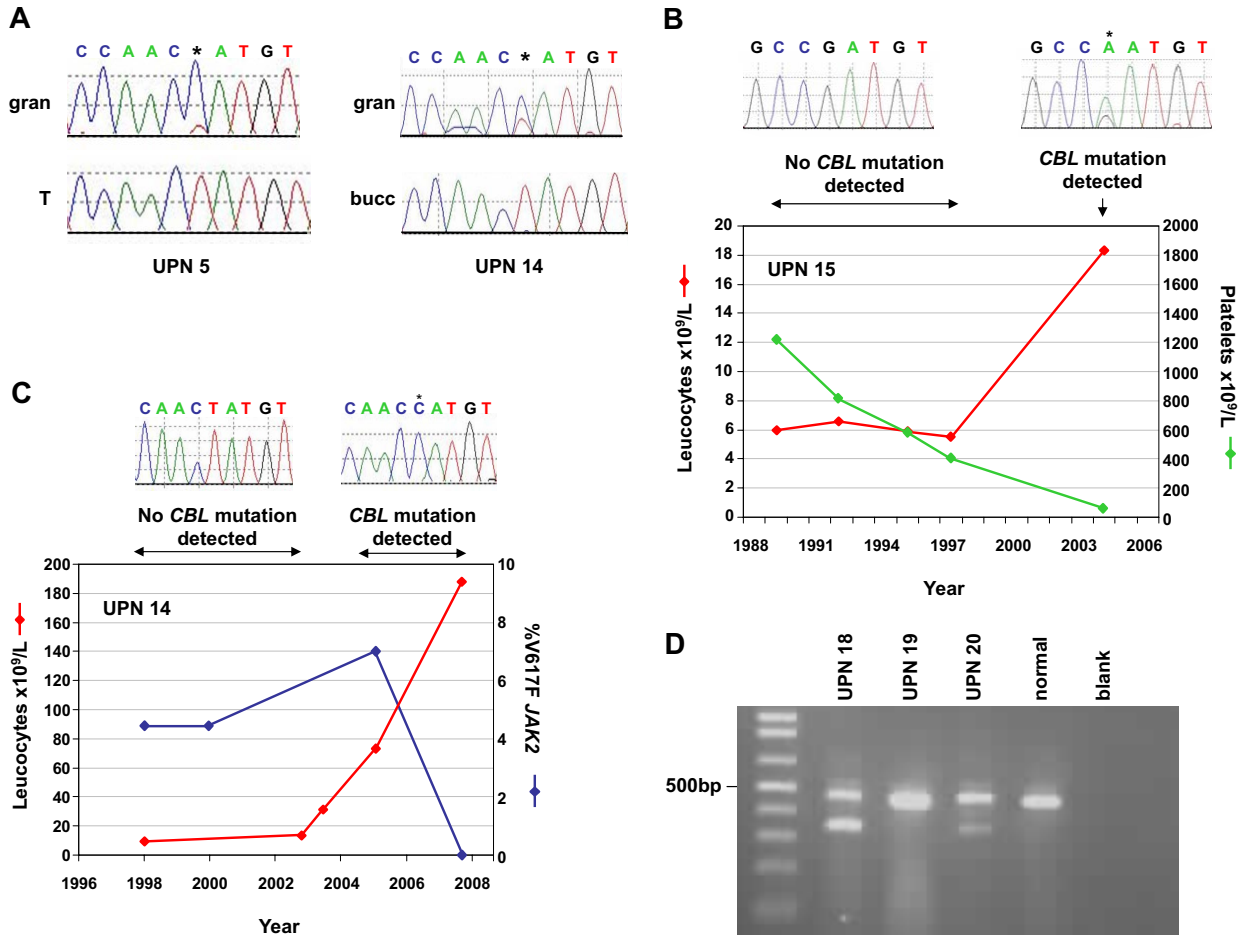
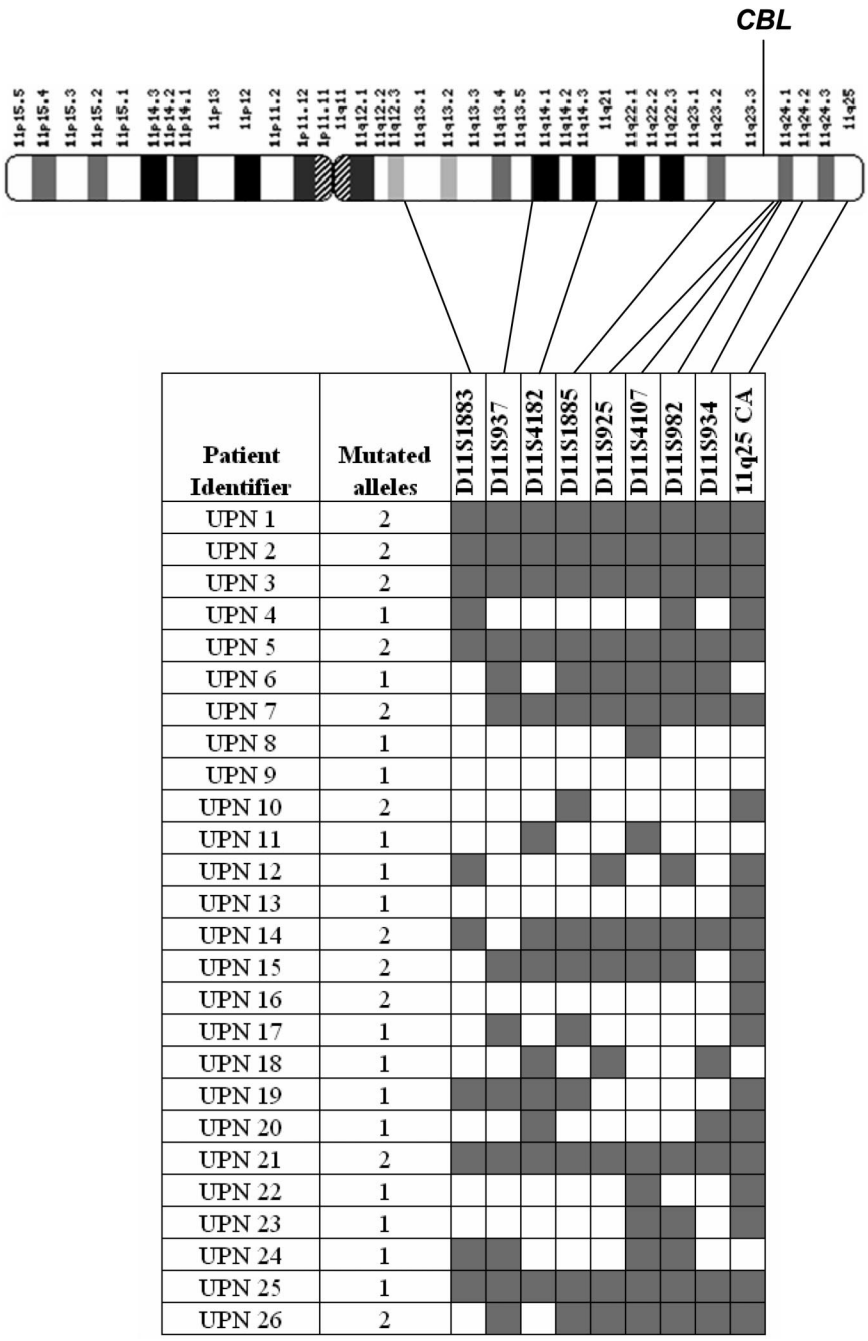


Figure 4. Acquisition of *CBL* mutations. (A) Comparison of granulocytes (gran) with buccal epithelial cells (bucc) or T cells (T) demonstrates that the L380P mutations in cases UPN 5 and 14 are acquired. (B) Sequential leukocyte and platelet counts for case, who presented in 1989 with essential thrombocythemia (ET) but progressed to myelofibrosis (MF) in 2004. Homozygous *CBL* R420Q, indicated by *, was detected on transformation but was undetectable in previous specimens. (C) Clinical course of case UPN 14, who presented with MF in 1998 but experienced elevated leukocyte counts in 2005, concomitant with the appearance of the *CBL* L380P mutation. The *JAK2* V617F mutation, which had been present at low level since diagnosis, was still detectable when *CBL* L380P first appeared but was undetectable in the 2008 sample, when the *CBL* mutated clone predominated (* on sequence trace). (D) RT-PCR analysis. All cases and controls show the expected product from *CBL* exons 7-10. Cases UPN 18 and UPN 20 show smaller bands that result from complete deletion of exon 8.



Patient Identifier	Mutated alleles	D11S1883	D11S937	D11S4182	D11S1885	D11S925	D11S4107	D11S982	D11S934	11q25 CA
UPN 1	2	■	■	■	■	■	■	■	■	■
UPN 2	2	■	■	■	■	■	■	■	■	■
UPN 3	2	■	■	■	■	■	■	■	■	■
UPN 4	1	■	□	□	□	□	□	■	□	■
UPN 5	2	■	■	■	■	■	■	■	■	■
UPN 6	1	□	■	□	■	■	■	■	■	□
UPN 7	2	□	■	■	■	■	■	■	■	■
UPN 8	1	□	□	□	□	□	■	□	□	□
UPN 9	1	□	□	□	□	□	□	□	□	□
UPN 10	2	□	□	□	■	□	□	□	□	■
UPN 11	1	□	□	■	□	□	■	□	□	□
UPN 12	1	■	□	□	□	■	□	■	□	■
UPN 13	1	□	□	□	□	□	□	□	□	■
UPN 14	2	■	□	■	■	■	■	■	■	■
UPN 15	2	□	■	■	■	■	■	■	□	■
UPN 16	2	□	□	□	□	□	□	□	□	■
UPN 17	1	□	■	□	■	□	□	□	□	■
UPN 18	1	□	□	■	□	■	□	□	■	□
UPN 19	1	■	■	■	■	□	□	□	□	■
UPN 20	1	□	□	■	□	□	□	□	■	■
UPN 21	2	■	■	■	■	■	■	■	■	■
UPN 22	1	□	□	□	□	□	■	□	□	■
UPN 23	1	□	□	□	□	□	■	■	□	■
UPN 24	1	■	■	□	□	□	■	■	□	□
UPN 25	1	■	■	■	■	■	■	■	■	■
UPN 26	2	□	■	□	■	■	■	■	■	■

Figure 5. Microsatellite analysis of *CBL* mutation positive cases. The number of *CBL* mutated alleles (1 monoallelic; 2 biallelic) was estimated by sequence analysis as described in "Experimental procedures." Microsatellites were scored as homozygous (■) or heterozygous (□) for each case and their positions on chromosome 11q are indicated.

HES/CEL (n = 1, 1%); that is, morphologically and clinically related diseases that generally exhibit a poor prognosis. We were unable to discern, however, any specific clinical or prognostic features specifically associated with *CBL* mutated cases.

Casitas B-lineage lymphoma (CBL) is a well characterized protein that plays both positive and negative regulatory roles in tyrosine kinase signaling. In its positive role, CBL binds to activated receptor tyrosine kinases via its N-terminal tyrosine kinase binding (TKB) domain and serves as an adaptor by recruiting downstream signal transduction components such as SHP2 and P13K. However the RING domain of CBL has E3 ligase activity and ubiquitinylates activated receptor tyrosine kinases on lysine residues, a signal that triggers internalization of the receptor/ligand complex and subsequent recycling or degradation.[21,24,25] CBL was originally identified after the characterization of v-Cbl, the transforming component of the Cas NS-1 retrovirus, and thus its association with neoplasia is well established[26]; however, it is only very recently that *CBL* mutations were first identified in a human malignancy, specifically occasional cases of AML.[15,23,27] We have found that *CBL* mutations are much more common in MPNs, and previous reports of 11q aUPD in MDS[28] suggest that *CBL* mutations will also prove to be common in this disease. Indeed, 7 of 12 MDS cases with aUPD at 11q were recently shown to harbor CBL mutations.[29] There are 2 other human *CBL* family members, *CBLB* and *CBLC*, but we did not identify aUPD in the regions containing these genes and did not detect any mutations of the linker/RING domain of *CBLB*.

We identified a total of 15 different *CBL* missense mutations affecting 12 residues (Table 2; Figure 3). Mutations of some of these residues (eg, Y371, C381, H398 and W408) have been

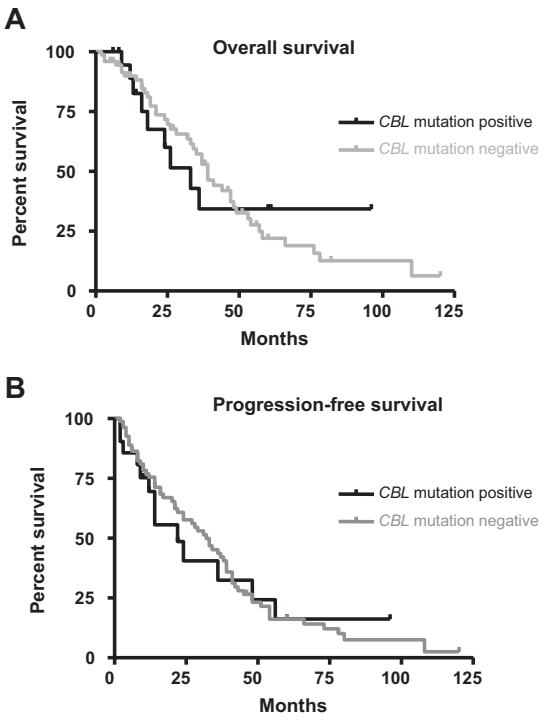


Figure 6. Clinical significance of CBL mutations. Overall survival (A) and progression-free survival (B) for atypical chronic myeloid leukemia (aCML), chronic myelomonocytic leukemia (CMML), and MF cases with (n = 19) or without (n = 87) *CBL* mutations. The differences between mutation positive and negative cases was not significant as determined by Kaplan-Meier analysis.

analyzed previously in a study that concluded that loss of E3 ubiquitin ligase activity and concomitant impairment of EGFR internalization by RING finger mutations was insufficient for oncogenicity, as assayed in NIH3T3 cells. However mutations of key residues within the α-helix of the linker region abolished ubiquitin ligase activity and were transforming.[30] Although the precise amino acid substitutions we identified were different from those studied, we found that mutations in both the RING finger and linker regions conferred autonomous growth to 32D/wt-FLT3 cells, consistent with a causal relationship to the MPN phenotype. The fact that one mutation (N454D) was not transforming emphasizes the importance of functional analysis to distinguish driver mutations from irrelevant passengers or infrequent polymorphisms.[31] Notably, this mutation was only weakly conserved between species, in contrast to all other missense mutations that affected highly conserved residues and are therefore likely to be functionally significant.

Enhanced autonomous growth in the presence of CBL mutations and overexpression of wild-type FLT3 may simply reflect the paucity of receptor tyrosine kinase gene expression in 32D cells or, alternatively, may indicate that the mutants are only weakly transforming and require the cooperation of other events to give rise to clinically manifest disease. Consistent with the latter hypothesis, we identified 2 cases in which *CBL* mutations were acquired during progression of a preexisting MPN (Figure 4). In both cases the identity of the cooperating change is unknown: although one presented with a (relatively low level) V617F *JAK2* mutation, this disappeared on progression to *CBL* L380P positive disease, indicating that the 2 mutations must have arisen independently in different clones. This is reminiscent of the observation that the leukemic blasts of approximately half of V617F positive MPNs that evolve to AML are negative for the *JAK2* mutation.[32,33] It remains to be

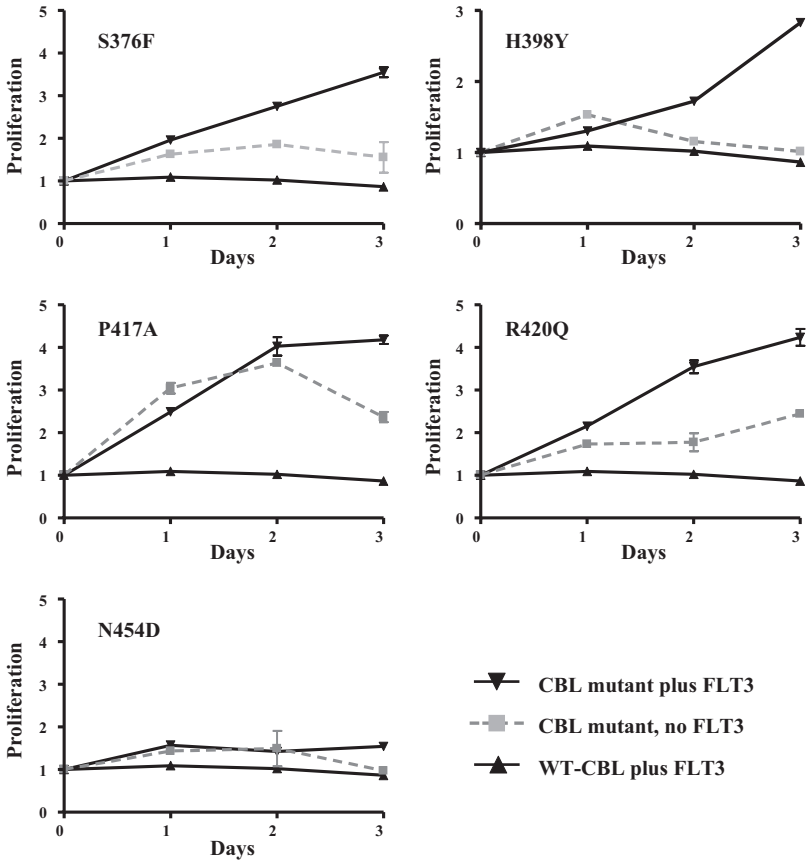


Figure 7. Transforming activity of *CBL* mutants. 32D cells or 32Ds overexpressing wild-type FLT3 were transfected with wild-type or mutant *CBL* constructs and assayed for IL-3 independent growth over 3 days. Results shown are the mean of at least 3 independent experiments, each of which were performed in triplicate. The y-axis shows proliferation in arbitrary units as determined by MTS assay (Promega) and the x-axis time in days.

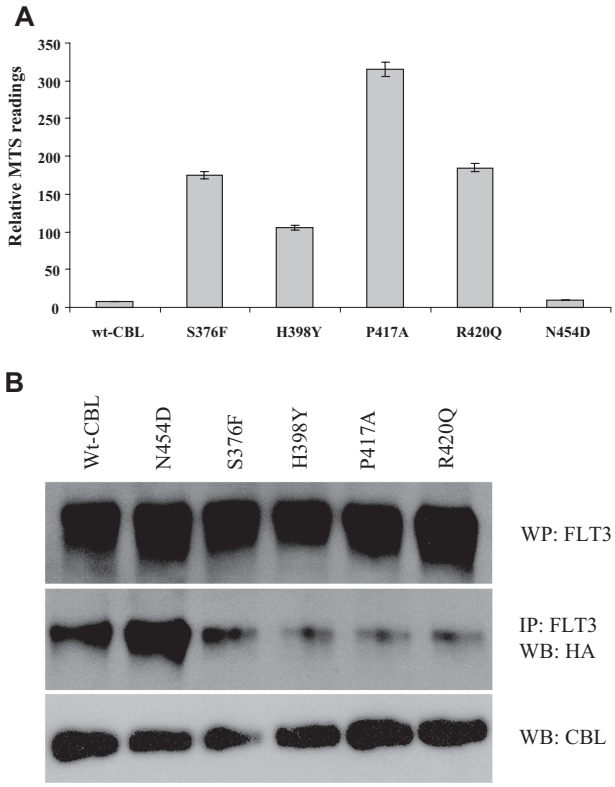


Figure 8. Survival and abrogation of E3 ubiquitin ligase activity. (A) Relative survival/proliferation of 32D cells expressing wild-type or mutant CBL constructs 48 hours after growth factor withdrawal as determined by MTS assay. (B) 32D cells expressing wild-type FLT3 and CBL constructs as indicated were transfected with HA-tagged ubiquitin and stimulated with FLT3 ligand. After lysis, FLT3 was immunoprecipitated, blotted, and probed with anti-HA. Total blots for FLT3 AND CBL are also shown.

established, however, if *CBL* mutations are usually secondary events or whether the 2 cases we identified are exceptional.

Further evidence for the requirement for additional cooperating genetic events comes from published mouse studies in which neither complete *CBL* knockout nor heterozygous knockin of the RING finger inactivating mutant C379A (mouse C379 is equivalent to human C381) resulted in evidence of an MPN or other malignancy.[34,35] However, closer analysis revealed subtle hemopoietic perturbations that clearly relate to the pathogenesis of MPNs. Specifically, hemopoietic stem cells (HSCs) from these animals were increased in number, hyperresponsive to thrombopoietin and more potent than wild-type HSCs in repopulating the hemopoietic system. These effects were associated with increased STAT5 accumulation and phosphorylation.[34,35] Hyperresponsiveness to growth factors and stem cell involvement are 2 of the cardinal features of MPNs. Furthermore, the association between STAT5 activation and MPNs is well documented.[36,37] While it is possible that overt hematologic disease might have been induced by a homozygous knockin mutant, these findings suggest that *CBL* mutations may not be solely responsible for a full MPN phenotype.

The nature of any cooperating abnormalities remains largely obscure. Our results show that overexpression of a tyrosine kinase can cooperate with *CBL* mutants to promote transformation, however we failed to identify any kinase that was consistently overexpressed in mutation positive cases by tyrosine kinase expression profiling (data not shown). One *CBL* mutation positive case, however, also tested positive for 7q UPD by SNP array, suggesting that these 2 abnormalities may complement each other. A *CBL* mutation resulting in skipping of exon 8 has also been identified in a murine model of NUP98-HOXD13 on progression from MDS to AML,[38] further corroborating the theory that acute leukemia arises from complementary mutations, one of which inhibits differentiation, and a second (in this case *CBL*) that enhances proliferation or inhibits apoptosis.

The fact that many *CBL* mutations are associated with aUPD strongly suggests that clones with homozygous mutations have a selective advantage over those that are heterozygous. This contrasts with the prevailing view that transforming *CBL* RING finger and linker region mutants act in a dominant negative fashion. Although it is conceivable that another, unknown locus on 11q is the target of aUPD, the finding of 1 case with a Y371H mutation and a stop codon on opposite alleles strongly suggests that homozygous *CBL* mutations do indeed confer a direct selective advantage. Because we did not screen the entire gene for mutations, it is possible that other cases with heterozygous mutations may also be functionally homozygous due to premature stop codons on the other allele. Alternatively, heterozygosity in these cases may only be apparent due to variable proportions of background normal cells that are not part of the malignant clone. Although the nontransforming N454D mutant was heterozygous, 2 cases with heterozygous R420Q were observed and thus absence of homozygosity cannot be used to infer the presence of an irrelevant, nontransforming sequence variant.

In summary, we have found that oncogenic *CBL* mutations are acquired in a subset of hematologically related, poor prognosis MPNs. Our findings further strengthen the notion that MPNs are primarily "tyrosine kinopathies," that is, diseases caused by aberrant activation of proliferation and survival pathways as a consequence of mutations that either directly or indirectly promote excess tyrosine kinase signaling. In view of the fact that tyrosine kinases are such good drug targets, the possibility that *CBL*-mutated cases might be amenable to therapeutic inhibition is clearly attractive.

Acknowledgments

This study was supported by Leukaemia Research (London, United Kingdom), Dr Mildred Scheel Stiftung für Krebsforschung (Bonn, Germany), the Lady Tata Memorial Trust (London, United Kingdom), Deutsche Krebshilfe eV (Bonn, Germany), and the Deutsche José Carreras Leukämie-Stiftung eV-DJCLS R06/02 (Munich, Germany).

Authorship

Contribution: F.H.G. and C.E.H.-C. designed the study and performed or assisted with the laboratory work and data analysis; T.E. and K.Z. performed or assisted with the laboratory work and data analysis; C.Z. provided clinical data and samples; C.M., S.K., A.J., and J.S. performed or assisted with the laboratory work and data analysis; G.M. and D.O. provided clinical data and samples; A.H. performed or assisted with the laboratory work and data analysis; C.B. and H.S. provided essential reagents and advice; A.R. provided clinical data and samples; and A.J.C. and N.C.P.C. performed or assisted with the laboratory work and data analysis. All authors contributed to the design of the study and the final manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Prof N. C. P. Cross, Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury SP2 8BJ, United Kingdom; e-mail: ncpc@soton.ac.uk.

References

1. Tefferi A, Vardiman JW. Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms. Leukemia. 2008;22:14-22.
2. Reiter A, Walz C, Cross NCP. Tyrosine kinases as therapeutic targets in BCR-ABL negative chronic myeloproliferative disorders. Current Drug Targets. 2007;8:205-216.
3. De Keersmaecker K, Cools J. Chronic myeloproliferative disorders: a tyrosine kinase tale. Leukemia. 2006;20:200-205.
4. Shtivelman E, Lifshitz B, Gale RP, Canaani E. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature. 1985;315:550-554.
5. James C, Ugo V, Le Couedic JP, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature. 2005;434:1144-1148.
6. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med. 2005;352:1779-1790.
7. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005;7:387-397.
8. Baxter EJ, Scott LM, Campbell PJ, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005;365:1054-1061.
9. Pikman Y, Lee BH, Mercher T, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. PLoS Med. 2006;3:e270.
10. Jones AV, Kreil S, Zoi K, et al. Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. Blood. 2005;106:2162-2168.
11. Kralovics R, Guan Y, Prchal JT. Acquired uniparental disomy of chromosome 9p is a frequent stem cell defect in polycythemia vera. Exp Hematol. 2002;30:229-236.
12. Yamamoto G, Nannya Y, Kato M, et al. Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. Am J Hum Genet. 2007;81:114-126.
13. Fitzgibbon J, Smith LL, Raghavan M, et al. Association between acquired uniparental disomy and homozygous gene mutation in acute myeloid leukemias. Cancer Res. 2005;65:9152-9154.
14. White HE, Hall VJ, Cross NCP. Methylation-sensitive high-resolution melting-curve analysis of the SNRPN gene as a diagnostic screen for Prader-Willi and Angelman syndromes. Clin Chem. 2007;53:1960-1962.
15. Sargin B, Choudhary C, Crosetto N, et al. Flt3-dependent transformation by inactivating c-Cbl mutations in AML. Blood. 2007;110:1004-1012.
16. Grand FH, Iqbal S, Zhang L, et al. A constitutively active SPTBN1-FLT3 fusion in atypical chronic myeloid leukemia is sensitive to tyrosine kinase inhibitors and immunotherapy. Exp Hematol. 2007;35:1723-1727.
17. Morita S, Kojima T, Kitamura T. Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther. 2000;7:1063-1066.
18. George P, Bali P, Cohen P, et al. Cotreatment with 17-allylamino-demethoxygeldanamycin and FLT-3 kinase inhibitor PKC412 is highly effective against human acute myelogenous leukemia cells with mutant FLT-3. Cancer Res. 2004;64:3645-3652.
19. Simon-Sanchez J, Scholz S, Fung HC, et al. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. Hum Mol Genet. 2007;16:1-14.
20. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. Nature. 2002;417:949-954.
21. Thien CB, Langdon WY. Negative regulation of PTK signalling by Cbl proteins. Growth Factors. 2005;23:161-167.
22. Melo JV, Barnes DJ. Chronic myeloid leukaemia as a model of disease evolution in human cancer. Nat Rev Cancer. 2007;7:441-453.
23. Caligiuri MA, Briesewitz R, Yu J, et al. Novel c-CBL and CBL-b ubiquitin ligase mutations in human acute myeloid leukemia. Blood. 2007;110:1022-1024.
24. Swaminathan G, Tsygankov AY. The Cbl family proteins: ring leaders in regulation of cell signaling. J. Cell Physiol. 2006;209:21-43.
25. Schmidt MH, Dikic I. The Cbl interactome and its functions. Nat Rev Mol Cell Biol. 2005;6:907-918.
26. Langdon WY, Hartley JW, Klinken SP, Ruscetti SK, Morse HC III. v-cbl, an oncogene from a dual-recombinant murine retrovirus that induces early B-lineage lymphomas. Proc Natl Acad Sci U S A. 1989;86:1168-1172.
27. Abbas S, Rotmans G, Lowenberg B, Valk PJ. Exon 8 splice site mutations in the gene encoding the E3-ligase CBL are associated with core binding factor acute myeloid leukemias. Haematologica. 2008;93:1595-1597.
28. Gondek LP, Tiu R, O'Keefe CL, et al. Chromosomal lesions and uniparental disomy detected by SNP arrays in MDS, MDS/MPD, and MDS-derived AML. Blood. 2008;111:1534-1542.
29. Dunbar AJ, Gondek LP, O'Keefe CL, et al. 250K single nucleotide polymorphism array karyotyping identifies acquired uniparental disomy and homozygous mutations, including novel missense substitutions of c-Cbl, in myeloid malignancies. Cancer Res. 2008;68:10349-10357.
30. Thien CB, Walker F, Langdon WY. RING finger mutations that abolish c-Cbl-directed polyubiquitination and downregulation of the EGF receptor are insufficient for cell transformation. Mol Cell. 2001;7:355-365.
31. Frohling S, Scholl C, Levine RL, et al. Identification of driver and passenger mutations of FLT3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. Cancer Cell. 2007;12:501-513.
32. Campbell PJ, Baxter EJ, Beer PA, et al. Mutation of JAK2 in the myeloproliferative disorders: timing, clonality studies, cytogenetic associations, and role in leukemic transformation. Blood. 2006;108:3548-3555.
33. Theocharides A, Boissinot M, Girodon F, et al. Leukemic blasts in transformed JAK2-V617F-positive myeloproliferative disorders are frequently negative for the JAK2-V617F mutation. Blood. 2007;110:375-379.
34. Thien CB, Blystad FD, Zhan Y, et al. Loss of c-Cbl RING finger function results in high-intensity TCR signaling and thymic deletion. EMBO J. 2005;24:3807-3819.
35. Rathinam C, Thien CB, Langdon WY, Gu H, Flavell RA. The E3 ubiquitin ligase c-Cbl restricts development and functions of hematopoietic stem cells. Genes Dev. 2008;22:992-997.
36. Cain JA, Xiang Z, O'Neal J, et al. Myeloproliferative disease induced by TEL-PDGFRB displays dynamic range sensitivity to Stat5 gene dosage. Blood. 2007;109:3906-3914.
37. Kotecha N, Flores NJ, Irish JM, et al. Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. Cancer Cell. 2008;14:335-343.
38. Slape C, Liu LY, Beachy S, Aplan PD. Leukemic transformation in mice expressing a NUP98-HOXD13 transgene is accompanied by spontaneous mutations in Nras, Kras, and Cbl. Blood. 2008;112:2017-2019.