UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

GENETIC ABERRATIONS IN PATIENTS WITH LEUKAEMIA AND THEIR IMPACT ON PROGNOSIS

by

Sebastian Kreil

Thesis for the degree of Doctor of Philosophy

April 2012

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE

Human Genetics

Doctor of Philosophy

GENETIC ABERRATIONS IN PATIENTS WITH LEUKAEMIA AND THEIR IMPACT ON PROGNOSIS

by Sebastian Kreil

Leukaemia is believed to arise through a multistep accumulation of genetic changes that result in deregulation of cell growth, differentiation, and programmed cell death. Many of these acquired changes are important indicators of prognosis and increasingly are being used to direct therapy. Inherited genetic factors may also influence the course of disease and response to treatment. The work of this thesis focuses on (i) the prognostic effect of derivative chromosome 9 deletions associated with the *BCR-ABL1* fusion gene in chronic myeloid leukaemia (CML) and (ii) the frequency of acquired point mutations in myeloid and lymphoid disorders, and (iii) a single nucleotide polymorphism that is associated with STAT3 expression and response of CML to therapy with interferon *alpha*.

Deletions within the derivative chromosome 9, or der(9), of the translocation t(9;22)(q34;q11), are seen in approximately 10-15 % of CML patients and have been associated with a poor prognosis, however no studies have been performed in the context of a randomised clinical trial. A DNA-based deletion screen was

developed and 339 chronic phase patients treated with IFN as first line therapy in three controlled German studies with a median observation time of seven years were investigated. Deletions were detected in pre-treatment DNA samples of 59 out of 339 patients in total (17 %). Of these, 21 spanned the *ABL1-BCR* junction and 38 were centromeric only (n=20) or telomeric only (n=18) of the breakpoint. There was no significant difference in overall survival between deleted and non-deleted patients. Patients with breakpoint-spanning deletions had poorer survival compared to patients without deletions (4.7 versus 7.8 years; P=0.003) but this was not significant when censored at allogeneic stem cell transplantation (n=129) or imatinib (n=62) treatment in first chronic phase (P=0.078). Unexpectedly, deletions that did not span the breakpoint were associated with improved survival compared to cases without deletions (P=0.001). Multiple Cox regression analysis indicated that deletion status (P=0.007), age (P=0.018) and spleen enlargement (P<0.001) were significant independent indicators of survival and confirmed that only deletions spanning the *ABL1-BCR* breakpoint were associated with an adverse prognosis (P=0.039).

(ii) Point mutations are one of the most common genetic alterations of the genome that are linked with the evolution of cancer. They may affect the amino acid sequence by changing single nucleotide of the coding triplets (codons). Three different populations of patients were investigated in this study who presented with either (a) myeloid malignancies (n=471), (b) lymphoid malignancies (n=145), or (c) polycythaemia vera that were negative for the *JAK2*^{V617F} point mutation (n=78). Mutation analyses were performed for genes that are known to be involved in the intracellular signalling process in haematopoietic cells or that are altered in their expression level in other malignancies. Screening involved all coding exons of *CSF1R*, *JAK2*, and *VHL;* the catalytic domains of *CSK*, *FES*, *SYK*, *STYK1*, and *TIE1*; the kinase domains of *EPOR*, *BTK,* and *BMX*; homologous regions for the *JAK2*^{V617F} and other known *JAK2* mutations in *JAK1*, *JAK3*, and *TYK2*; *FLT3* exon 20; and the two micro RNAs *hsa-mir-221* and *hsa-mir-222*. As a result, a total of six new nucleotide alterations were found: (a) four novel changes were detected in patients with myeloid disorder, i.e. *CSF1R*^{L40V}, *CSF1R*^{R106W}, *CSF1R*^{G747R},

CSF1R^{R753Q}, plus the known *JAK2*^{V617F} mutation; (b) a single patient with CLL presented with a novel *TYK2*^{A659P} change; (c) no changes were found in *JAK2*^{V617F}-negative patients. Further investigations are required to assess the significance of these changes.

(iii) Interferon *alpha* (IFN) induces variable responses in CML, with 8-30 % of early chronic phase cases achieving a complete cytogenetic response. We hypothesized that polymorphic difference in genes encoding IFN signal transduction components might account for different patient responses. A total of 174 IFN-treated patients have been studied, of whom 79 achieved less than 35 % Philadelphia-chromosome (Ph) positive metaphases (responders) and 95 failed to show any cytogenetic response (more than 95 % Ph-positive metaphases; non-responders). In detail, 17 single nucleotide polymorphisms (SNPs) within the genes *IFNAR1, IFNAR2, JAK1, TYK2, STAT1, STAT3* and *STAT5a/b* were compared between the two groups. A significant difference was found for rs6503691, a SNP tightly linked to *STAT5a*, *STAT5b* and *STAT3* (minor allele frequency 0.16 for non-responders; 0.06 for responders, P=0.007). Levels of *STAT3* mRNA correlated with rs6503691 genotype (P<0.001) as assessed by real time quantitative PCR and therefore we concluded that rs6503691 is associated with the *STAT3* expression levels and response of CML patients to IFN.

Declaration of Originality

I declare that this thesis is a result of my own work carried out between October 2004 and October 2007 in the Wessex Regional Genetics Laboratory, Salisbury District Hospital, Salisbury, United Kingdom. I carried out all the work apart from elements that are clearly specified below. I have acknowledged all sources and clearly attributed their published works. Parts of this work have been published:

Kreil, S., Pfirrmann, M., Haferlach, C.*, et al.* (2007) Heterogeneous prognostic impact of derivative chromosome 9 deletions in chronic myelogenous leukemia. *Blood,* 110, 1283-1290.

Kreil, S., Waghorn, K., Ernst, T.*, et al.* (2010) A polymorphism associated with STAT3 expression and response of chronic myeloid leukemia to interferon alpha. *Haematologica,* 95, 148-152.

Chase, A., Schultheis, B., **Kreil, S***., et al.* (2009) Imatinib sensitivity as a consequence of a CSF1R-Y571D mutation and CSF1/CSF1R signaling abnormalities in the cell line GDM1. *Leukemia.,* 23, 358-364.

Kreil, S., Hochhaus, A., Cross, N.C.*, et al.* (2009) A high-throughput candidate gene mutation screen in lymphoproliferative and myeloproliferative neoplasias. *Leuk.Res.,* 33, e168-e169.

Sebastian Kreil

Acknowledgements

Many thanks go to my supervisor Nick Cross who gave me the great opportunity to realise this research project at his laboratory. He guided me through the field of molecular genetics and has always been a great support throughout the whole time and beyond.

Furthermore I thank my fellow lab colleagues, namely Andy Chase, Helen White, Francis Grand, Chris Mattocks, David Bunyan, Kathy Waghorn, Jo Score, Amy Jones, Claire-Hidalgo Curtis, Vicky Hall, Gemma Watkins, Jonathan Callaway, and all the others who were not only a great help during the time but also became close friends.

I also thank Andreas Hochhaus for his never-ending support from the beginning throughout my whole career. He led me the way to haematology and encouraged me in my interest in science and molecular genetics.

Many thanks also go to Markus Pfirrmann in Munich. He carried out some of the more complex statistical analyses, especially the univariate and multivariate analyses. His support gave me a more profound knowledge of statistics.

Furthermore I thank my family, first of all my parents without whom I had never reached this point. Their support continued and their trust in all I did never came to an end. Thanks to the family also includes my brothers Philipp and Alexander and their partners. Finally I thank my partner Markus who never lost patience when listening to my explanations about leukaemia and mutations.

Abbreviations

%	per cent
°C	degree Celsius
A	adenine
a.k.a.	also known as
ABL1	Abelson murine leukaemia oncogene homolog 1
ABI1	ABL interactor1
AKT1	*v-akt* murine thymoma viral oncogenes homolog1
ADP	adenosine diphosphate
ALL	acute lymphoblastic leukaemia
AML	acute myeloid leukaemia
AML1	synonym of RUNX1
AMP	adenosine monophosphate
aMPN	atypical myeloproliferative neoplasm
APL	acute promyelocytic leukaemia
Ara-C	cytosine arabinoside
ARMS	amplification refractory mutation system
ASS1	argininosuccinate synthetase1
b	base
BAD	BCL2 antagonist
BCL	B-cell lymphoma gene family member
BCR	breakpoint cluster region
BM	bone marrow
BMX	bone marrow kinase, X-linked
bp	base pair
BRCA	breast cancer gene family member
BTK	Bruton agammaglobulinaemia tyrosine kinase
C	cytidine
CBFB	core-binding factor, *beta* subunit
CCR	complete cytogenetic response
cDNA	complementary deoxyribonucleic acid
CFU	colony-forming unit

CGH	comparative genomic hybridisation
Chr.	chromosome
CMD	chronic myeloid disorder
CML	chronic myeloid leukaemia
CLL	chronic lymphocytic leukaemia
CMML	chronic myelomonocytic leukaemia
cMPN	chronic myeloproliferative neoplasia
CNV	copy number variation
CRKL	*v-crk* avian sarcoma virus CT10 oncogene homologous
CSF1R	colony-stimulating factor 1 receptor
CSK	cytoplasmic tyrosine kinase
CSCE	conformation-sensitive capillary electrophoresis
Da	Dalton
dATP	deoxyadenosinetriphosphate
dCTP	deoxycytidinetriphosphate
der	derivative
df	degrees of freedom
dGTP	deoxyguanosinetriphosphate
DHPLC	denaturing high-performance liquid chromatography
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
ds	double stranded
dsDNA	double stranded deoxyribonucleic acid
dTTP	deoxythymidine triphosphate
e.g.	*exempli gratia* (lat.); for example
EGF	epidermal growth factor
EDTA	ethylenediaminetetraacetic acid
ELN	European LeukemiaNet
EPO	erythropoietin
EPOR	erythropoietin receptor
ERBB2; HER2/NEU	erythroblastic leukaemia oncogene homolog 2
ET	essential thrombocythaemia
et al.	*et alii* (lat.); and others

ETV6	a.k.a. TEL; ETS variant gene 6
EXOSC2	exosome component 2
f	femto
FAB	French-American-British
FES	v-FES feline sarcoma viral/v-FPS Fujinami avian sarcoma viral oncogene homolog
FMS	synonym of CSF1R
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FIP1L1	FIP1-like 1
FLJ31568	synonym for ZDHHC8P1
g	gram
G	guanine
G-CSF	granulocyte colony-stimulating factor
gDNA	genomic DNA
GM-CSF	granulocyte/ macrophage colony-stimulating factor
GSTT1	glutathione S-transferase, theta 1
GvH	graft versus host
GvL	graft versus leukaemia
GWAS	genomewide association study
HES	hypereosinophilic syndrome
Hg	*hydrargyrum* (lat.); mercury
HGF	haematopoietic growth factor
HIF	hypoxia inducible factor
hr	hour
HRM	high resolution melting curve analysis
HSC	haematopoietic stem cell
hSNF5/INI1	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily B, member 1
HU	hydroxyurea
i.e.	*id est* (lat.); that is
IFN	interferon
IFNAR1	human interferon *alpha*-receptor 1

IGF1	insulin like growth factor 1
IGLL1	immunoglobulin lambda-like polypeptide 1
IM	imatinib mesylate (Novartis Pharma, Basel, Switzerland)
IMF	a.k.a. PMF; idiopathic myelofibrosis
IBE	Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (München, Germany)
ITD	intragenic tandem duplication
JAK	Janus kinase
JMML	juvenile myelomonocytic leukaemia
JUN	*v-jun* avian sarcoma virus 17 oncogene homolog
K562	BCR-ABL-positive patient derived cell line (Lozzio and Lozzio 1975)
kb	kilo bases
KIT	a.k.a. SCFR; stem cell factor receptor
L	litre
LD	linkage disequilibrium
LOH	loss of heterozygosity
m	milli
µ	micro
M	molar
MAPH	multiplex amplifiable probe hybridisation
MAF	minor allele frequency
MAPK	mitogen-activated protein kinase gene family member
Mb	mega base
MC3	Ph-positive leukaemic cell line (Okabe*, et al.* 1995)
MCR	major cytogenetic response
MDM2	mouse double minute 2 homolog
MDS	myelodysplastic syndrome
min	minute
MLH1	homolog of *E. coli* mutL 1
MLL	myeloid/lymphoid or mixed leukaemia gene

MLPA	multiplex ligation-dependent probe amplification
MMM	myelofibrosis with myeloid metaplasia
MPN	myeloproliferative neoplasm
MPL	myeloproliferative leukaemia virus oncogene
mRNA	messenger ribonucleic acid
MSH2	mutS homolog 2
MYH11	myosin, heavy chain 11, smooth muscle
MYK; EPHB4	ephrin receptor EPHB4
NDNL2	necdin-like gene 2
NDP	nucleotide diphosphate
NF-κB	nuclear factor kappa B
NK-cell	natural killer cell
NMP	nucleotide monophosphate
NOS	not otherwise classified
NTP	nucleotide triphosphate
p	pico
PAG	phosphoprotein associated with glycosphingolipid-enriched microdomains
PB	peripheral blood
PCR	polymerase chain reaction
PDGFA	platelet-derived growth factor *alpha*
PDGFB	platelet-derived growth factor *beta*
PDGFRA	platelet-derived growth factor receptor *alpha*
Ph	Philadelphia chromosome, also Ph1
PI3	phosphatidyl 3-kinase gene family member
PMF	primary myelofibrosis
PML	inducer of acute promyelocytic leukaemia
PPi	pyrophosphate
PRDM12	PR domain containing 12
PTGES	prostaglandin E synthase
PV	polycythaemia vera
PVSG	Polycythaemia Vera Study Group
RAF	*v-raf* murine leukaemia viral oncogene homolog

RARA	retinoic acid receptor *alpha*
RAS	rat sarcoma viral oncogene homolog
REAL	Revised European-American Lymphoma
RNA	ribonucleic acid
rpm	rounds per minute
RQ-PCR	real-time quantitative polymerase chain reaction
RT	room temperature
RT-PCR	reverse transcriptase polymerase chain reaction
RUNX1	RUNT-related transcription factor 1
SCT	stem cell transplantation
sec	second
SLE	systemic lupus erythematosus
SH3	SRC homology 3 domain
SMARCB1	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily B, member 1
SNP	single nucleotide polymorphism
SOCS	suppressor of cytokine signalling
STI571	synonym of imatinib mesylate, used in early studies
STYK1	serine/threonine/tyrosine kinase 1
SYK	protein-tyrosine kinase SYK, spleen tyrosine kinase
SRC	sarcoma oncogene
ss	single stranded
STAT	signal transducer and activator of transcription
SYP	synaptophysin
T	thymidine
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
TGFB	transforming growth factor *beta*
TIE 1	tyrosine kinase with immunoglobulin and EGF factor homology domains 1
TKI	tyrosine kinase inhibitor
TPO	thrombopoietin
Tris	Tris (hydroxymethyl) amino-methane

TYK2	tyrosine kinase 2
U	unit
UBR1	ubiquitin-protein ligase E3 component N-recognin 1
UK	United Kingdom
U.S.A.	United States of America
UV	ultra violet
VHL	von Hippel-Lindau gene
vs.	versus
WBC	white blood cell
WHO	World Health Organisation
ZAP70	Zeta-chain-associated protein kinase 70
ZDHHC8P1	zinc finger, DHHC-type containing 8 pseudo gene 1

Nomenclature

The recommendations for nomenclature in the Genetic Nomenclature Guide (1998)* have been followed. Names of genes and RNA transcripts are given in italics, proteins are in normal type. Human genes and products are fully capitalised; mouse genes have capital initial letters only, proteins are capitalised, e.g.:

Human gene: *CDX2*

Human protein: CDX2

Mouse gene: *Cdx2*

Mouse protein: CDX2

* Genetic Nomenclature Guide. Trends Genet 14: S1-S49, 1998.

Contents

Index of Tables

Index of Figures

1 Introduction

1.1 Haematopoiesis

The process of the production of different blood cells, such as erythrocytes, granulocytes, platelets, and lymphocytes from the haematopoietic stem cell (HSC) is called haematopoiesis. Each day, about one billion leukocytes, two billion erythrocytes, and nearly one billion platelets are required to sustain the haematopoietic system in an adult (Kaushansky 2006a). The system is a hierarchy of cells in which multipotent HSCs give rise to lineage-committed progenitor cells, which divide to generate the mature blood cells. HSC are rare, occurring with a frequency of one in 10^4 to 10^5 total nucleated bone marrow cells. They are defined as predominantly quiescent cells that are capable of self-renewal. By their proliferative capacity they give rise to separate lineages of differentiated progeny (Martin-Rendon and Watt 2003). The quiescence of stem cells is mainly facilitated by the microenvironment of the matrix and the stromal cells of the bone marrow. The process of self-renewal describes the ability of the HSC to divide into at least one other HSC, with the other one proceeding to form lineage-committed progeny. A self-renewal probability of 0.5 is required to sustain the haematopoietic system (Gordon and Blackett 1998, Gordon*, et al.* 1999). Upon commitment to proliferation, the pluripotent HSC may differentiate into lymphoid or myeloid restricted progenitors and subsequently into one of the eight distinguishable mature cell lineages. A schematic of haematopoietic differentiation is shown in figure 1. With increasing lineage commitment, the characteristic of pluripotency is lost. Finally, mature haematological cells are unable to divide and undergo apoptosis after their cell-characteristic lifespan (Bellantuono 2004).

Multipotent myeloid and lymphoid progenitor cells are derived from multipotent stem cells. Named after their characteristic growth pattern in cell cultures, multipotent myeloid progenitors are described as colony-forming units of granulocytes, erythrocytes, megakaryocytes, and macrophages (CFU-GEMM). Further differentiation of the myeloid progenitors results in six myeloid lineages, i.e. neutrophils (CFU-G), eosinophils (CFU-Eo), basophils (CFU-Baso), erythrocytes

(CFU-E), megakaryocytes and platelets (CFU-Meg), and macrophages (CFU-M). Multipotent lymphoid progenitors differentiate to B-lymphocytes and T-lymphocytes (Bellantuono 2004).

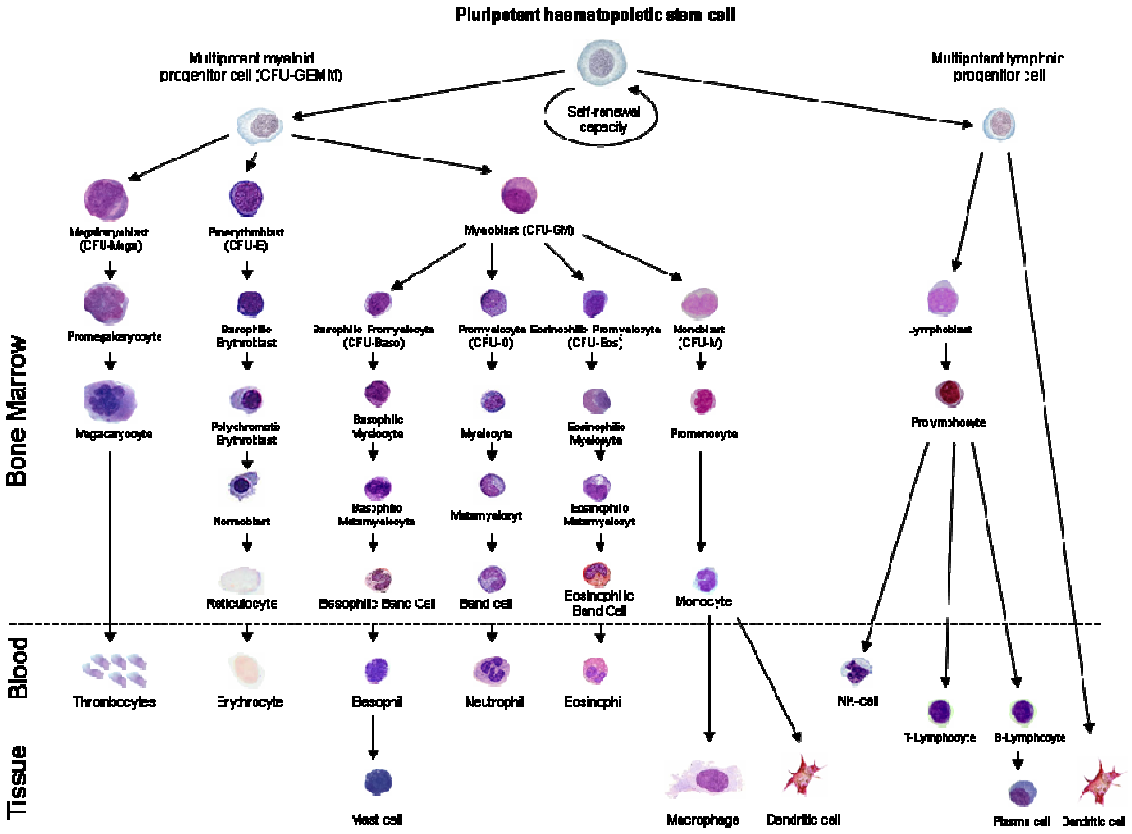


Figure 1: Schematic of haematopoiesis

The pluripotent haematopoietic stem cell has self-renewal capacity. Upon commitment, it divides and produces one of two major multipotent progenitor cells for the myeloid or lymphoid lineages. Genetic mutation on the stem cell level may affect several cell lineages depending on the origin of the alteration. Abbreviations are explained in the main text.

Erythrocytes are anucleate, round, biconcave cells that contain haemoglobin, the major molecule for the transport of the respiratory gases oxygen and carbon dioxide. The number of erythrocytes in the human blood is usually a constant 4.5 x 10^6 cells per micro litre. The cells have a lifespan of up to 120 days in circulation. Granulocytes are characterised by their specific granules and are classified as

neutrophils, eosinophils, and basophils. They have the ability to migrate into many tissues and play a key role in inflammation and phagocytosis. Neutrophils are round cells with a single lobulated nucleus (3 to 4 lobes) and contribute to the majority (50-70 %) of white blood cells (leukocytes) with an absolute number of 7,500 cells per microlitre whole blood. They have a lifespan of a few days up to one week. Eosinophils are characterised as round cells with a bi-lobed nucleus and azurophilic granules. Their contribution to total leukocytes in the blood is small (1-4 %) and their lifespan is less than two weeks. Basophils have a large nucleus and prominent, specific histamine-containing granules. They are the least numerous of all blood cells (0.5-1 %) and survive up to 1.5 years. Closely related to basophils are mast cells with similar morphology and function in phagocytosis and allergy response. Platelets are small, anucleate cell fragments that derive from large, multi-nucleate megakaryocytes in the bone marrow. Their key property is the ability to provide haemostasis by adherence, aggregation, and coagulation reactions. The range of the platelet count in blood is 2-4.5 x 10^5 platelets per microlitre. Lymphocytes are mononucleate cells that mediate highly specific immunity against microorganisms and many kinds of foreign antigens. Their proportion in the blood is 20-25 % of all leukocytes. B-lymphocytes have their origin of maturation in the bone marrow and confer immunity through the production of specific, soluble antibodies as part of the humeral response. Their lifespan is variable, usually of a few months. T-lymphocytes, however, mature in the thymus and interact directly with their target by cell-cell contact. They may survive for many years (Greer*, et al.* 2003).

The differentiation, maturation, and survival of haematopoietic cells are tightly regulated by a number of growth factors, cytokines, and transcription factors. Under conditions of increased need, such as inflammation or hypoxia, the rate of cell production may increase ten-fold (Kaushansky 2006a). Haematopoietic growth factors (HGFs) are acidic glycoproteins that bind to specific cell surface receptors. Their regulation is mainly determined by the physiological need of each blood cell type. Granulocyte colony-stimulating factor (G-CSF) stimulates the growth of the neutrophil population and promotes their survival. The gene is expressed in different tissues, e.g. endothelial cells, macrophages, and fibroblasts.

Erythropoietin (EPO) is mainly produced in the kidney and regulates the erythrocyte population. Reduced oxygenation of the juxtatubular interstitial cells of the renal cortex encourages an increase of EPO expression. Thrombopoietin (TPO) is produced by liver and kidney tissue and stimulates platelet production in the bone marrow. All HGFs interact with their target cell via membrane receptors that transfer the signal into the cell to transcription within the nucleus. HGF production is regulated at several molecular and cellular levels, including gene transcription, messenger ribonucleic acid (mRNA) stability, protein translation and receptor mediated uptake and destruction (Kaushansky 2006b). Several subtypes of special cytokines, i.e. interleukins, are responsible for proliferation and differentiation of B- and T-lymphocytes as well as their close relatives, the natural killer cells (NK-cells) and dendritic cells.

1.2 Haematological Malignancies

The first published description of a case of haematological malignancy in the medical literature dates to 1827, when a French physician named Alfred-Armand-Louis-Marie Velpeau described a 63-year-old florist who developed an illness characterised by fever, weakness, urinary stones, and substantial enlargement of the liver and spleen. Velpeau noted that the blood of this patient had a consistency 'like gruel', and speculated that the appearance of the blood was due to white corpuscles (Velpeau 1827). Subsequently in 1845, two scientists reported at nearly the same time a clinical and pathological phenomenon that later became known as leukaemia. John Hughes Bennett (Edinburgh, Scotland, United Kingdom) was the first who described the clinical condition of a patient with massive splenomegaly associated with leukocytosis, two characteristic features of chronic myeloid leukaemia (CML) (Bennett 1845). The German physician and scientist Rudolf Virchow (Berlin, Germany) provided a more detailed pathological picture just a few weeks later. He used the term 'Weisses Blut' in his publication to describe the overwhelming amount of white or 'colourless' blood cells in the patient's blood, even though staining methods were unknown in those days (Virchow 1845). Virchow suggested the term 'Leukaemie' in 1847 (Virchow 1847, Virchow 1856).

However, it was only some decades later that a general appreciation of the term took place in the scientific world (Geary 2000).

Today, leukaemia describes a wide range of clonal haematological disorders with heterogeneous features. It embraces acute and chronic forms of neoplasms of the myeloid and/or the lymphoid cell lineage. The World Health Organisation (WHO) provided the latest classification as a diagnostic tool that has been developed in conjunction with the American Society of Hematology (ASH) and European colleagues. The revised system includes not only morphological characteristics, but also genomic, immunophenotypic, biologic, and clinical features. Several different classifications for disorders of the myeloid lineage [e.g. the French-American-British (FAB) Cooperative Group classifications of acute myeloid leukaemia (AML) and myelodysplastic syndrome (MDS), guidelines of the Polycythaemia Vera Study Group (PVSG) and for myeloproliferative disorders] were adapted and included as well as the widely used Revised European-American Lymphoma (REAL) classification for lymphoid diseases. Since new genetic and molecular aberrations continue to be discovered that are of diagnostic importance, the classification of haematologic diseases is under continual review and updated regularly (Swerdlow 2008).

Leukaemia in general is a rare disease and occurs worldwide in about 15 of 100,000 of the population. Males are up to twice as often affected as females. The diversity of leukaemia is determined by the origin of the malignant cell clone from different cell types. Defined myeloid and lymphoid lineages entities are summarised in table 1.

Research in recent years has revealed an increasing number of genomic alterations. In some cases these have provided explanations of disease pathogenesis as well as leading to dramatic changes in therapeutic opportunities to the benefit of patients. However, a complete explanation of underlying molecular mechanisms that lead to leukaemia is still lacking. Also, the diagnostic value of possible molecular variations still needs to be determined for many entities.

Table 1: Myeloid and lymphoid cell lineages involved in leukaemia *

Myeloid cell lineage

Acute Myeloid Leukaemia (AML)
- AML with recurrent genetic abnormalities
  - AML with t(8;21)(q22;q22), (*AML1-ETO*)
  - AML with inv(16)(p13q22) or t(16;16)(p13;q22), (*CBFβ-MYH11*)
  - Acute promyelocytic leukaemia with t(15;17)(q22;q12), (*PML-RARα*)
  - AML with t(9;11)(p22;q23), (*MLLT3-MLL*)
  - AML with t(6;9)(p23;q34), (*DEK-NUP2314*)
  - AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2), (*RPN1-EVI1*)
  - AML (megakaryoblastic) with t(1;22)(p13;q13), (*RBM15-MKL1*)
  - AML with mutated *NPM1*
  - AML with mutated *CEBPA*
- Acute myeloid leukaemia with myelodysplasia-related changes
- Therapy-related myeloid neoplasms
- Acute myeloid leukaemia, not otherwise categorised
  - AML with minimal differentiation
  - AML without maturation
  - AML with maturation
  - Acute myelomonocytic leukaemia
  - Acute monoblastic and monocytic leukaemia
  - Acute erythroid leukaemia
  - Acute megakaryoblastic leukaemia
  - Acute basophilic leukaemia
  - Acute panmyelosis with myelofibrosis
- Myeloid sarcoma
- Myeloid proliferations related to Down syndrome
  - Transient abnormal myelopoiesis
  - Myeloid leukaemia associated with Down syndrome

Myelodysplastic syndromes
- Refractory cytopenia with unilineage dysplasia
  - Refractory anaemia
  - Refractory neutropenia
  - Refractory thrombocytopenia
- Refractory anaemia with ringed sideroblasts
- Refractory cytopenia with multilineage dysplasia
- Refractory anaemia with excess blasts
- MDS associated with isolated del(5q)
- Myelodysplastic syndrome, unclassifiable
- Childhood myelodysplastic syndrome
  - Refractory cytopenia of childhood

Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)
- Chronic myelomonocytic leukaemia
- Atypical chronic myeloid leukaemia, *BCR-ABL1* negative
- Juvenile myelomonocytic leukaemia
- Myelodysplastic/myeloproliferative neoplasm, unclassifiable
- Refractory anaemia with ring sideroblasts associated with marked thrombocytosis

Chronic myeloproliferative neoplasms
- Chronic myelogenous leukaemia, *BCR-ABL1* positive
- Chronic neutrophilic leukaemia
- Polycythaemia vera
- Primary myelofibrosis
- Essential thrombocythaemia
- Chronic eosinophilic leukaemia, NOS
- Chronic myeloproliferative disease, unclassifiable
- Mastocytosis
  - Cutaneous mastocytosis
  - Systemic mastocytosis
  - Mast cell leukaemia
  - Mast cell sarcoma
  - Extracutaneous mastocytoma

Myeloid and lymphoid neoplasms with eosinophilia and abnormalities of *PDGFRA*, *PDGFRB* or *FGFR1*

Lymphoid cell lineage

Precursor lymphoid neoplasms
- B-lymphoblastic leukaemia/lymphoma
  - B-lymphoblastic leukaemia/lymphoma, NOS
  - B-lymphoblastic leukaemia/lymphoma with recurrent genetic abnormalities
    - B-lymphoblastic leukaemia/lymphoma with t(9;22)(q34;q11.2), (*BCR-ABL1*)
    - B-lymphoblastic leukaemia/lymphoma with t(v;11q23), (*MLL* rearranged)
    - B-lymphoblastic leukaemia/lymphoma with t(12;21)(p13;q22), (*ETV6-RUNX1*)
    - B-lymphoblastic leukaemia/lymphoma with hyperdiploidy
    - B-lymphoblastic leukaemia/lymphoma with hypodiploidy (hypodiploid ALL)
    - B-lymphoblastic leukaemia/lymphoma with t(5;14)(q31;q32), (*IL3-IGH*)
    - B-lymphoblastic leukaemia/lymphoma with t(1;19)(q23;q13.3), (*TCF3-PBX1*)
- T-lymphoblastic leukaemia/lymphoma

Mature B-cell neoplasms
- Chronic lymphocytic leukaemia/small lymphocytic lymphoma
- B-cell prolymphocytic leukaemia
- Splenic marginal zone lymphoma
- Hairy cell leukaemia
- Lymphoplasmacytic lymphoma
- Heavy chain disease
- Plasma cell myeloma
- Solitary plasmacytoma of bone
- Extraosseous plasmacytoma
- Extranodal marginal zone lymphoma of MALT type
- Nodal marginal zone lymphoma
- Follicular lymphoma
- Mantle-cell lymphoma
- Diffuse large B-cell lymphoma
- Burkitt's lymphoma/Burkitt cell leukaemia

Mature T-cell and NK-cell neoplasia
- T-cell prolymphocytic leukaemia
- T-cell granular lymphocytic leukaemia
- Aggressive NK-cell leukaemia
- Adult T-cell lymphoma/leukaemia
- Extranodal NK/T-cell lymphoma, nasal type
- Enteropathy-associated T-cell lymphoma
- Hepatosplenic T-cell lymphoma
- Subcutaneous panniculitis-like T-cell lymphoma
- Mycosis fungoides/ Sezary syndrome
- Primary cutaneous gamma-delta T-cell lymphoma
- Peripheral T-cell lymphoma, NOS
- Angioimmunoblastic T-cell lymphoma
- Anaplastic large-cell lymphoma, *ALK* positive

Hodgkin's lymphoma (Hodgkin's disease)
- Nodular lymphocyte predominant Hodgkin's lymphoma
- Classical Hodgkin's lymphoma
  - Nodular sclerosis classical Hodgkin's lymphoma
  - Lymphocyte-rich classical Hodgkin's lymphoma
  - Mixed cellularity classical Hodgkin's lymphoma
  - Lymphocyte-depleted classical Hodgkin's lymphoma

* adapted after (Swerdlow 2008)

1.2.1 Clonality of haematologic disorders

In contrast to the polyclonal growth of normal haematopoietic cells following extrinsic stimulation, neoplasia arises from mutations in single cells and thus all progeny are clonal. The aetiology for most haematological malignancies is not fully understood, but in some cases it may be driven by known mutagens such as high dose radiation or chemical substances such as benzene. Following an initial mutation and clonal outgrowth, other acquired changes may also accumulate that result in more aggressive subclones, e.g. additional cytogenetic aberrations in advanced disease. Clonal populations of cells can be identified by a number of different independent approaches, including analyses of karyotype, gene rearrangements, deletions or point mutations, X-linked polymorphisms, and integration of virus into the genome (Gilliland*, et al.* 1991).

1.2.2 Molecular pathogenetic mechanisms associated with leukaemic transformation

Until the discovery of a chromosomal aberration in patients with chronic leukaemia by Nowell and Hungerford in 1960 (later named the Philadelphia (Ph) chromosome), sub-cellular mechanisms of the origin of leukaemia were obscure (Nowell and Hungerford 1960). The realisation that the Ph-chromosome was the smaller derivative of a reciprocal translocation between chromosomes 9 and 22 by Rowley in 1973 (Rowley 1973), the discovery of the fusion gene *BCR-ABL1* (de Klein*, et al.* 1982) and murine studies that clearly demonstrated that *BCR-ABL1* is the primary driver of the disease process have served as a paradigm for understanding the molecular pathogenesis of haematological malignancies (Daley and Baltimore 1988). Several mechanisms leading to the abnormal expansion of a neoplastic cell clone have since been proposed. However, current models do not fully explain cancer evolution. Multi-step mutational mechanisms in association with epigenetic effects are proposed but not yet completely understood. Enhanced cell proliferation combined with elevated mutation rates and impaired genomic stability both are believed to be important factors in the development of cancer.

There are two gene categories considered to be main targets for mutations in cancer evolution, i.e. oncogenes and tumour suppressor genes. The latter have lost their normal function of inhibiting events that may lead to cancer. However, most proto-oncogenes (i.e. the normal unmutated counterparts of oncogenes) are normally involved in the regulation of differentiation and cell growth. Oncogenes are activated by different mechanisms in tumourigenesis. The family of oncogenes embraces HGFs (e.g., *PDGFB*), cell surface receptors (e.g., *ERBB*, *CSF1R*), components of an intracellular signal transduction pathway (e.g., *ABL*, *RAS*), transcription factors (e.g., *MYC*, *JUN*), and genes involved in cell cycle and damage control processes (e.g., *MDM2*). There are four known mechanisms for activation of proto-oncogenes, i.e. (i) gene amplification, (ii) point mutation, (iii) chromosomal rearrangements creating a chimaeric gene, and (iv) intragenic tandem duplications. In the end, it is thought that the normal cell cycle is interrupted and proliferation, impaired differentiation, and loss of apoptotic stimulation will lead to an expanded cell clone that overpowers normal cell growth (Kelly and Gilliland 2002).

1.2.2.1 Activated protein kinases in leukaemia

Protein kinases are molecules with enzymatic activity that catalyse cellular reactions by transferring phosphate groups to other proteins. A signal is transmitted by a cascade of different proteins and is thereby enhanced or decreased. In most cases, the signal's origin is extra-cellular and reaches the cell as a ligand to a specific receptor, e.g. a receptor tyrosine kinase. The bound ligand leads receptor homo- or hetero-dimerization resulting in conformational changes of the catalytic domains and full activation through auto-phosphorylation. For receptor proteins without intrinsic kinase activity, such as the erythropoietin receptor, a receptor-associated tyrosine kinase, e.g. a Janus kinase (JAK), is activated by a conformational change of the receptor. The activated complex acts as the starting point for the signal cascade of several pathways, e.g. the JAK/STAT, the PI3-Kinase/AKT, or the RAS/MAP-Kinase pathway. The activated pathways may influence cell structure, gene transcription, cell proliferation, and survival.

Protein tyrosine kinases are tightly regulated by other small molecules or proteins, e.g. phosphatases, suppressors of cytokine signalling, and other negative regulators. Aberrations of the signal cascade may lead to uncontrolled proliferation, cell cycle deregulation, abrogation of apoptosis and in some instances a block of differentiation. A schematic of the JAK/STAT signalling pathway is shown as an example in figure 2.

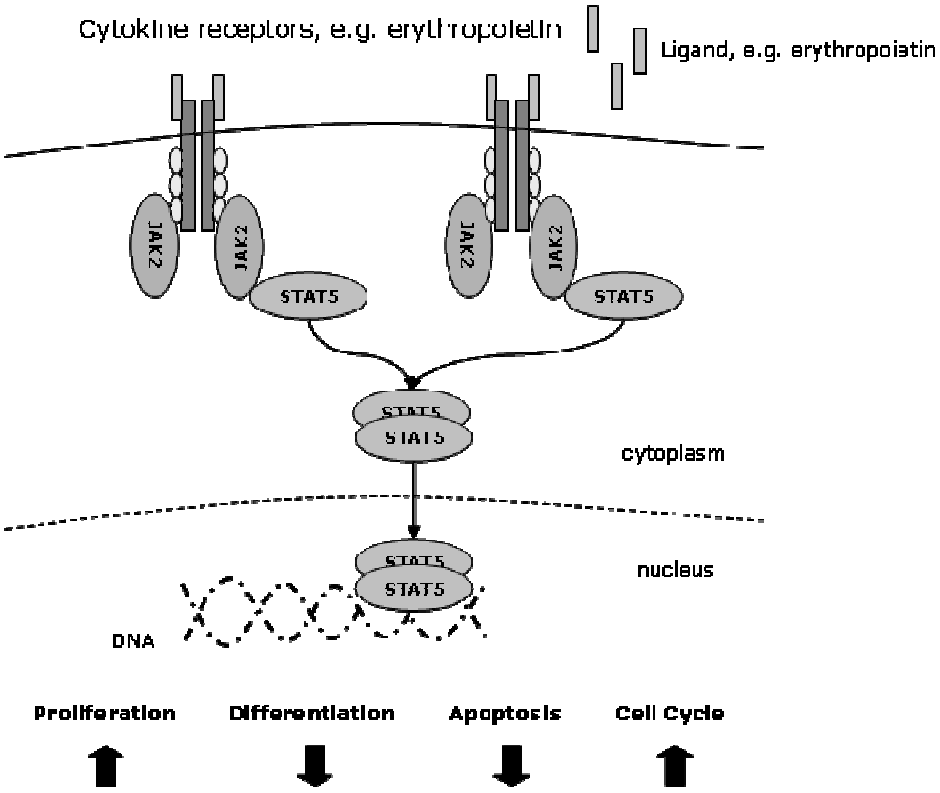


Figure 2: Intracellular JAK/STAT signalling pathway

The schematic shows the JAK/STAT pathway as an example for the intracellular signalling process. The binding of a receptor-specific ligand, e.g. erythropoietin, leads to dimerisation of the receptor. This activates JAK2 and initiates a cascade by phosphorylation, which involves the receptor, JAK2, and downstream signalling components, e.g. the signal transducer and activator of transcription proteins (STATs). Phosphorylated STAT molecules move to the nucleus where they act as direct transcription factors, thereby up regulating target genes with consequent effects on proliferation, differentiation, apoptosis, and cell cycle.

1.2.2.1.1 Gene amplification

Genomic amplification of kinase genes has been observed in many cancers but has not yet been shown to play an important role in leukaemia. For example, amplification of the *ERBB2* (*HER2*/*NEU*) gene in a proportion of breast cancer patients leads to an over-expression of the cell surface receptor *ERBB2* that has been successfully targeted therapeutically by the monoclonal antibody trastuzumab (Piccart-Gebhart*, et al.* 2005). Gene amplification can be studied by comparative genomic hybridisation (CGH), either standard CGH or array-CGH, and these techniques have identified prognostically or functionally significant amplifications in leukaemia that do not involve kinases, e.g. RUNX1 (alternatively named AML1) and MYB in B-ALL and T-ALL, respectively (Lahortiga*, et al.* 2007, Moorman*, et al.* 2007).

1.2.2.1.2 Point mutations

Although the accuracy of DNA polymerase and editing enzymes is very high, mutations may arise at low frequency during the replication process. The mutation frequency may be increased by exposure to mutagens from the external environment such as toxins or radiation. There are three main classes of mutations, i.e. base substitutions (point mutations), deletions, and insertions (Strachan and Read 2004). In coding DNA, point mutations may be synonymous (silent), i.e. no amino acid change occurs, or non-synonymous, i.e. with a change of the gene product by alteration of the amino acid sequence. Insertions and deletions may lead to a shift in the translational reading frame (frame shift deletion) that may lead to a premature termination codon resulting in either a truncated protein or mRNA degradation through nonsense-mediated decay. Mutation in non-coding DNA sequences may also cause impaired gene expression as a result of changes to promoter or regulatory regions; or may alter splicing.

There are a number of acquired, activating point mutations in kinase genes that are associated with the evolution of leukaemic sub-entities. For example, mutations of

the proto-oncogene *KIT* have been found in most patients with systemic mastocytosis and some cases of acute myeloid leukaemia (Pardanani, *et al.* 2006, Schnittger, *et al.* 2006). Mutation of *KIT* exons 9, 11, 13, and 17 were also found in patients with gastrointestinal stromal cell tumour (Hirota, *et al.* 1998, Lux, *et al.* 2000, Rubin, *et al.* 2001). The discovery of the *JAK2*^{V617F} point mutation in patients with myeloproliferative neoplasms such as polycythaemia vera (PV), essential thrombocythaemia (ET), primary myelofibrosis (PMF), and other related disorders has shed new light on the pathogenetic mechanisms of the evolution of leukaemia (Baxter, *et al.* 2005, James, *et al.* 2005, Jones, *et al.* 2005, Kralovics, *et al.* 2005, Levine, *et al.* 2005b).

1.2.2.1.3 Polymorphic variations

Small DNA sequence variations in a population with a prevalence of more than one per cent are called single nucleotide polymorphisms (SNPs). SNPs are a frequent human genomic variation (Human Genome Sequencing 2004, Komura, *et al.* 2006) with the average nucleotide diversity between individuals having been estimated to be about 0.08 % (Przeworski, *et al.* 2000). Based on the number of three billion nucleotides in the human genome, the expected number of SNPs was two to three million. However, the Human Genome Project published a much higher number of estimated SNPs in 2005 (Hinds, *et al.* 2005). They estimated the presence of seven to ten million SNPs with a minor allele frequency (MAF) of at least 5 %. Another four million SNPs have been documented with an MAF of one to 5 % (Crawford, *et al.* 2005) and more recently the 1000 genomes project identified 15 million SNPs, most of which were previously undescribed (2010).

The replacement of cytosine by thymine is the most frequently observed alteration. SNPs occur in both coding and non-coding genome regions, while those not affecting the amino acid sequence are the majority. However, the latter may impair promoter regions, other regulatory regions, or microRNA sequences and are therefore maybe of importance. Although more than 99 % of SNPs are observed with a certain frequency in the population, some may vary and indicate an impact

on human disease or drug response. Several groups, e.g. 'The Human Genome Project', 'The International HapMap Project' and 'The 1000 genomes project', started to identify the human genomic sequence including the frequency of SNPs in different populations around the world. Their observations were published, openly accessible (www.ornl.gov, www.hapmap.org, www.1000genomes.org) and have been used as a basis for whole genome association studies for a variety of common diseases, yielding new loci associated with disease predisposition, including cancer (Chung and Chanock 2011, Hosking*, et al.* 2011, Visscher*, et al.* 2012).

In addition to SNPs, analysis using array comparative genomic hybridisation has revealed a hitherto unsuspected variation in copy number of DNA sequences between individuals. These copy number variants (CNVs) are very common and disease susceptibility associated with CNVs are just beginning to emerge, e.g. individuals with an amplification of the gene *CCL3L1* were less likely to become infected with the human immunodeficiency virus (HIV), or to progress to the acquired immunodeficiency syndrome (AIDS) once they were infected (Gonzalez*, et al.* 2005). So far, however, there is no clear connection between inherited CNVs and cancer, although some CNVs contain cancer-associated genes and it is possible that a connection may emerge (Fanciulli*, et al.* 2010).

1.2.2.1.4 Chromosomal rearrangements

The translocation t(9;22) was the first consistent chromosomal rearrangement to be associated with leukaemia and indeed any human malignancy (Nowell and Hungerford 1960, Rowley 1973).

As described below in more detail, later research revealed that translocation fused the major part of the tyrosine kinase *ABL1* to a gene with dimerisation ability (*BCR*), resulting in constitutive activation of the kinase (Ben-Neriah*, et al.* 1986, Shtivelman*, et al.* 1985, Stam*, et al.* 1985). Following this pioneering work, more than 300 gene fusions are recognised up to now in cancer and many other tyrosine kinase fusions have been identified, particularly in atypical myeloproliferative

neoplasms (Mitelman*, et al.* 2007). An exemplary list of those known in atypical myeloproliferative neoplasms is given in table 2.

Table 2: Examples of tyrosine kinase fusion genes in myeloproliferative neoplasms

Karyotype	Fusion gene	References
t(9;22)(q34;q11)	*BCR–ABL1*	(Bartram*, et al.* 1983, de Klein*, et al.* 1982)
t(9;12)(q34;p13)	*ETV6-ABL1*	(Andreasson*, et al.* 1997)
t(8;13)(p11;q12)	*ZNF198-FGFR1*	(Xiao*, et al.* 1998)
t(6;8)(q27;p11)	*FGFR1OP-FGFR1*	(Popovici*, et al.* 1999)
t(8;9)(p11;q33)	*CEP110-FGFR1*	(Guasch*, et al.* 2000)
t(8;22)(p11;q22)	*BCR-FGFR1*	(Demiroglu*, et al.* 2001)
t(7;8)(q34;q11)	*TIF1-FGFR1 (TRIM24-FGFR1)*	(Belloni*, et al.* 2005)
ins(12;8)(p11;p11p21)	*FGFR1OP2-FGFR1*	(Grand*, et al.* 2004a)
t(8;17)(p11;q25)	*MYO18A-FGFR1 (TIAF1-FGFR1)*	(Walz*, et al.* 2005)
t(8;19)(p12;q13)	*HERVK-FGFR1*	(Guasch*, et al.* 2003)
t(12;13)(p13;q12)	*ETV6-FLT3*	(Vu*, et al.* 2006)
t(9;12)(p24;p13)	*ETV6-JAK2*	(Peeters*, et al.* 1997)
t(8;9)(p21;p24)	*PCM1-JAK2*	(Reiter*, et al.* 2005)
t(9;12)(q22;p12)	*ETV6-SYK*	(Kuno*, et al.* 2001)
del(4q)	*FIP1L1-PDGFRA*	(Cools*, et al.* 2003)
t(4;22)(p11;q11)	*BCR-PDGFRA*	(Baxter*, et al.* 2002)
t(3;4;10;13)(p21;q12;?;q?)	*KIF5B-PDGFRA*	(Score*, et al.* 2006)
t(2;4)(p24;q12)	*STRN-PDGFRA*	(Curtis*, et al.* 2007)
t(4;12)(q2?3;p1?2)	*ETV6-PDGFRA*	(Curtis*, et al.* 2007)
t(5;12)(q33;p13)	*ETV6-PDGFRB*	(Golub*, et al.* 1994)
t(5;7)(q33;q11)	*HIP1-PDGFRB*	(Ross*, et al.* 1998)
t(5;10)(q33;q21)	*H4-PDGFRB (CCDC6-PDGFRB)*	(Kulkarni*, et al.* 2000)
t(5;17)(q33;p13)	*RABEP1-PDGFRB (RABPT5-PDGFRB)*	(Magnusson*, et al.* 2001)
t(1;5)(q23;q33)	*PDE4DIP-PDGFRB*	(Wilkinson*, et al.* 2003)
t(5;17)(q33;p11)	*HCMOGT-PDGFRB (SPECC1-PDGFRB)*	(Morerio*, et al.* 2004)
t(5;14)(q33;q24)	*NIN-PDGFRB*	(Vizmanos*, et al.* 2004)
t(5;15)(q33;q22)	*TP53BP1-PDGFRB*	(Grand*, et al.* 2004b)
t(5;14)(q33;q32)	*KIAA1509-PDGFRB*	(Levine*, et al.* 2005c)
t(5;14)(q33;q32)	*CEV14-PDGFRB (TRIP11-PDGFRB)*	(Abe*, et al.* 1997)

1.2.3 Chronic myeloid disorders

Myeloid disorders are categorised into either acute myeloid leukaemias (AML) or chronic myeloid disorders (CMD). William Dameshek first classified the chronic myeloproliferative disorders in 1951 as an entity of haematologic diseases with characteristic clinical and morphologic features (Damashek 1951). He distinguished between chronic myeloid leukaemia (CML), polycythaemia vera (PV), essential thrombocythaemia (ET), myelofibrosis with myeloid metaplasia (MMM), and erythroleukaemia. The latter later became re-categorised as a sub-entity of acute leukaemias. The former are referred to as 'classical' myeloproliferative neoplasms (cMPN). Several other clinicopathologic entities that resemble the cMPN in terms of both phenotype and biology have subsequently been described and operationally assigned the term 'atypical' myeloproliferative neoplasm (aMPN) (table 3).

Myeloproliferative neoplasms (MPN) are pathologically distinct from the myelodysplastic syndrome (MDS) whose hallmark is a single-, bi-or tri-lineage dysplasia and cytopenia rather than excess proliferation. Although each family member is characterised by increased numbers of specific blood cells, such as granulocytes in chronic myeloid leukaemia (CML), red cells in PV, and platelets in ET, all share clinical features, including basophilia, hyperproliferative marrow, slow progression, active haematopoiesis in the liver and spleen, and propensity for evolution to acute leukaemia.

The most recent major revision of the WHO classification focused more on molecular characteristics than on morphological features. CML, for example, was defined as '*BCR-ABL1* positive' disease while other chronic myeloid disorders are referred to as '*BCR-ABL1*-negative' myeloproliferative neoplasms (Vardiman*, et al.* 2009). It is assumed that molecular alterations are responsible for most neoplastic features MPN, although their presence has been defined in only a few cases so far. Specific mutations have been identified in tyrosine kinases of some patients with eosinophilic leukaemia (*PDGFRA* or *PDGFRB*), systemic mastocytosis (*KIT*), or 8p11 myeloproliferative syndrome (*FGFR1*) (Reiter*, et al.* 1998, Tan*, et al.* 1990).

Table 3: Semi-molecular definition of myeloproliferative neoplasm (Tefferi 2006a)

	Main categories	Subcategories
Classical MPN	BCR-ABL1-positive	Chronic myeloid leukaemia (CML)
	BCR-ABL1 negative	Polycythaemia vera (100 % $JAK2^{V617F}$-positive) Essential thrombocythaemia (50 % $JAK2^{V617F}$-positive) Primary myelofibrosis (50 % $JAK2^{V617F}$-positive)
Atypical MPN	Chronic myelomonocytic leukaemia	
	Juvenile myelomonocytic leukaemia (frequent *PTP11*, *NF1*, and *RAS* mutations)	
	Chronic neutrophilic leukaemia (20 % JAK^{V617F}-positive)	
	Chronic eosinophilic leukaemia/ eosinophilic MPN	PDGFRA-rearranged (e.g., FIP1L1-PDGFRA) PDGFRB-rearranged (e.g., TEL/ETV6-PDGFRB) FGFR1-rearranged (e.g., ZNF198/FIM/RAMP-FGFR1); a.k.a. 8p11 myeloproliferative syndrome Molecularly undefined
	Hypereosinophilic syndrome	
	Chronic basophilic leukaemia	
	Systemic mastocytosis	PDGFRA-rearranged (e.g., FIP1L1-PDGFRA) KIT-mutated (e.g., KITD816V) Molecularly undefined
	Unclassified MPN (20 % JAK^{V617F}-positive)	Mixed/overlap myelodysplastic syndrome/MPN CML-like, but BCR-ABL1-negative

1.2.3.1 Chronic myeloid leukaemia (CML)

CML is often referred to as the paradigm of a myeloproliferative disorder. The constitutive activation of a tyrosine kinase causes a malignant transformation resulting in the clonal expansion of the leukaemic cell clone. While this molecular hallmark has been proven for chronic phase CML, a more heterogeneous pattern is associated with other myeloproliferative disorders (Vardiman*, et al.* 2009).

1.2.3.1.1 Clinical characteristics

CML is a clonal haematopoietic disorder that is phenotypically characterised by (i) leukocytosis (98 %), anaemia (62 %), and thrombocytosis (48 %) in the peripheral blood, (ii) splenomegaly (72 %) and (iii) marked myeloid hyperplasia in the bone marrow (Hehlmann*, et al.* 1994). CML accounts for 15 % of all cases of leukaemia with an incidence of 1.6-2 patients per 100,000 persons per year. There is a slight male predominance of 1.4 to 1. The median age of patients at presentation (in clinical trials) is 54 years with a wide range from early childhood onwards and an increasing incidence in elderly patients. The clinical course is usually tri-phasic. The criteria of CML phases vary between WHO and international study groups, however they are generally characterised by clinical and laboratory findings and other abnormalities as listed in table 4 (Baccarani*, et al.* 2006).

Table 4: Comparison of diagnostic criteria for accelerated phase CML *

* adapted from (Baccarani*, et al.* 2006)

WHO criteria	Other criteria used in most studies
10-19 % blast cells in blood or bone marrow	15-29 % blast cells in blood or bone marrow; >30 % blast cells plus promyelocytes in blood or bone marrow, with blast cells less than 30 %
>20 % basophils in blood	>20 % basophils in blood
Persistent thrombocytopenia (platelet count less than $100x10^9$/L) unrelated to therapy	Persistent thrombocytopenia (platelet count less $100x10^9$/L) unrelated to therapy
Thrombocytosis (platelet count greater than $1,000x10^9$/L) unresponsive to therapy	not included
Increasing spleen size and increasing WBC count unresponsive to therapy	not included
Cytogenetic evidence of clonal evolution (the appearance of additional genetic abnormalities that were not present at the time of diagnosis)	not included

Most patients present for the first time at an early stage, i.e. in chronic phase. Chronic phase lasts, without appropriate treatment, for three to five years. This is

followed by a shorter, more aggressive accelerated phase of few months duration, and finally leading to the terminating blastic phase which may be of myeloid or lymphoid phenotype but which is almost universally fatal. The individual duration of each phase varies widely (Sawyers 1999).

1.2.3.1.2 Cytogenetic and molecular characteristics

Early cytogenetic studies in 1960 by Nowell and Hungerford described an aberrant chromosome of the G group as a characteristic of CML, which was later identified as a derivative chromosome 22 and termed the Philadelphia chromosome (Ph) after the location of its discovery (Nowell and Hungerford 1960). A decade later, it emerged that the altered chromosome 22, or 22q-, is in fact one product of a reciprocal translocation between the long arms of chromosomes 9 and 22. The abnormality then was designated t(9;22)(q34;q11) by Janet Rowley (Rowley 1973). The translocation of genetic material leads to the fusion of parts of the human Abelson murine leukaemia viral oncogene homolog 1 (*ABL1;* or previously named *ABL)* proto-oncogene on chromosome 9 to a region on chromosome 22, named the breakpoint cluster region (*BCR*) (Shtivelman*, et al.* 1985). *ABL1* is the human homolog of the v-*abl* oncogene that is carried by the murine leukaemia virus (A-MuLV) (Abelson and Rabstein 1970). *ABL1* encodes a non-receptor tyrosine kinase that is ubiquitously expressed as a 145 kDa protein and is functionally involved in the activation of transcription factors such as NF-κB (Deininger*, et al.* 2000). Wild type ABL1 exists in both the cytoplasm and the nucleus, and is able to shuttle between them; BCR-ABL1, however, is only located in the cytoplasm (Goldman and Melo 2003). The 160 kDa BCR protein is also ubiquitously expressed and normally plays a role in macrophage activation.

By definition, the *BCR-ABL1* fusion gene is present in all newly diagnosed CML patients. About 5 % of CML patients lack the characteristic t(9;22) or complex variants, and instead have an occult *BCR-ABL1* fusion, either on the normal appearing chromosome 22 or on chromosome 9 (Goldman and Melo 2003). Patients that have a clinical picture of CML but no *BCR-ABL1* fusion are usually

referred to as atypical CML (aCML) or another atypical subtype of MPN. The constitutive activation of other tyrosine kinases other than *ABL1* (e.g., *JAK2*, *PDGFRB*, *FGFR1*, or others as shown in table 2) can give rise to a CML-like syndrome in these patients (Baxter, *et al.* 2002, Reiter, *et al.* 1998, Reiter, *et al.* 2005, Steer and Cross 2002). Although these abnormalities are infrequent, they are often strong prognostic and therapeutic indicators.

The breakpoint within *ABL1* may occur over a region of >200 kb either upstream of exon 1b, or downstream of exon 1a, or between the two. A schematic of both genes, *ABL1* and *BCR*, is shown in figure 3 A.

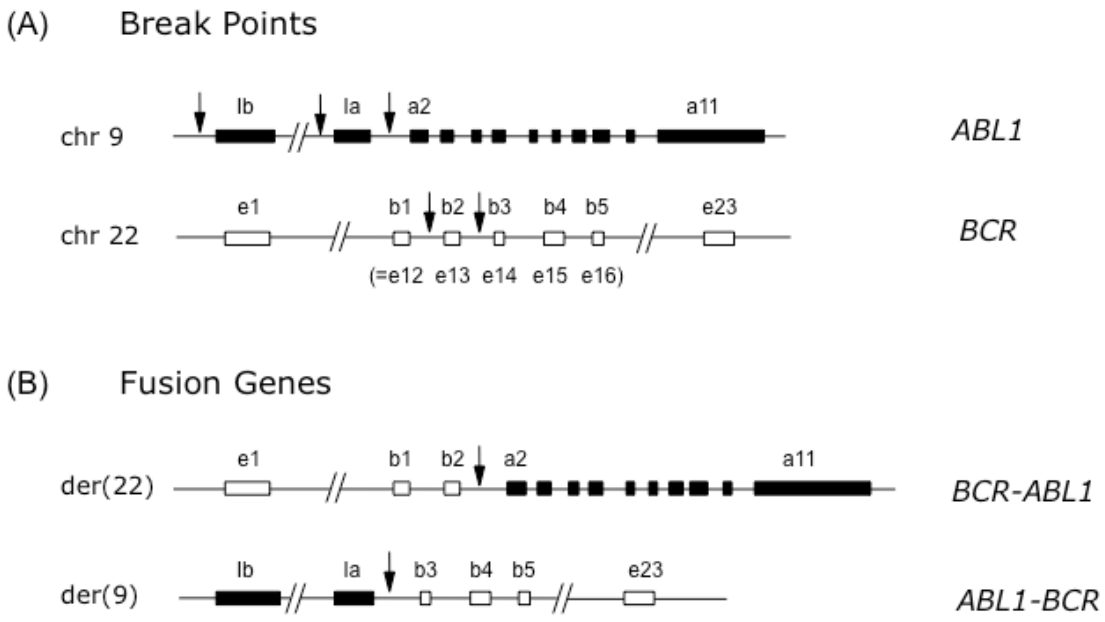


Figure 3: Schematic of the *BCR-ABL1* translocation in CML

(A) The *ABL1* gene on chromosome 9 and the *BCR* gene on chromosome 22 (not to scale) with arrows indicating common breakpoints. Selected exon numbers are shown. (B) An example for a b2a2 *BCR-ABL1* fusion gene with the corresponding reciprocal Ia-b3 *ABL1-BCR* fusion.

However, splicing of *BCR-ABL1* pre-mRNA usually fuses *ABL1* sequences of exon a2 to *BCR*, as schematically shown in figure 3 B. In contrast, on chromosome 22, breakpoints are more variable, and are clustered in three main regions. Most of

CML patients, but only one third of Ph-positive ALL patients, express the p210 BCR-ABL1 protein. The breakpoints are located in the 5.8 kb major breakpoint cluster region (M-bcr) between exon e12 and exon e13, but are alternatively spliced at *BCR* exon e13 and exon e14 (historically known as b2 and b3, respectively) (Daley*, et al.* 1990). Occasional patients with CML, but the majority of Ph-positive ALL, are characterised by the expression of the p190 BCR-ABL1 protein. The smaller protein results from a more upstream breakpoint located within *BCR* intron 1 designated as the minor breakpoint cluster region (m-bcr). Phenotypically, CML patients presenting with p190 BCR-ABL1 are characterised by a more pronounced monocytosis (Melo*, et al.* 1994). A small proportion of CML cases and also Ph-positive chronic neutrophilic leukaemia have been found to express a large p230 BCR-ABL1 protein, resulting from a breakpoint located in the micro breakpoint cluster region (μ-bcr) near *BCR* exon e19. p230-positive cells are less growth factor independent than their p210- or p190-BCR-ABL1 counterparts and the clinical course associated with this fusion is often but not always relatively mild (Wilson*, et al.* 1997). Low-level expression of p190 *BCR-ABL1* has also been detected in most patients with mainly p210 *BCR-ABL1*, probably due to alternative splicing of the primary mRNA transcript (van Rhee*, et al.* 1996). Occasional patients harbour other *BCR-ABL1* fusion transcripts than those arising from M-bcr (i.e. b2a2 and b3a2), m-bcr (e1a2) or μ-bcr (e19a2) (e.g. b2a3, b3a3, e1a3, e6a2) (Hochhaus*, et al.* 1996, Liu*, et al.* 2003, Melo 1997, Roman*, et al.* 2001). While *BCR-ABL1* is expressed by definition in all CML patients, mRNA transcribed from the reciprocal *ABL1-BCR* gene on the derivative chromosome 9 is only detectable in 70 % of cases (Melo*, et al.* 1993).

The initiating event of the t(9;22) translocation remains uncertain. There was an increased incidence of leukaemia reported after the detonation of the nuclear bombs in Hiroshima and Nagasaki, Japan, in 1945. It has been suggested that many of these cases had CML (Ichimaru*, et al.* 1991, Preston*, et al.* 1994). Ionising radiation has been shown to induce translocations *in vivo* and *in vitro* (Corso*, et al.* 1995, Deininger*, et al.* 1998). The physically short distance between *ABL1* and *BCR* in the interphase nucleus may play a role by providing more opportunities for illegitimate recombination following DNA breaks (Deininger*, et al.* 1998, Kozubek*,*

et al. 1997, Neves, *et al.* 1999). However, the appearance of *BCR-ABL1* in a haematologic cell alone does not necessarily lead to CML, as the fusion transcript is also detectable at a low level in a healthy population that will never develop leukaemia (Biernaux, *et al.* 1995).

1.2.3.1.3 Pathophysiological characteristics

The fundamental transforming event in CML is believed to be the acquisition of the *BCR-ABL1* fusion, as evidenced by the fact that *BCR-ABL1* alone produces a CML-like disease when expressed in the mouse haematopoietic system (Daley and Baltimore 1988, Daley, *et al.* 1990). The deregulation of tyrosine kinase activity seems to play a pivotal role in myeloproliferative disorders, as it is also seen in *BCR-ABL1* negative MPN in the form of the $JAK2^{V617F}$ mutation. In chronic phase CML, cell differentiation is hardly impaired. The increased leukocyte count is characterised by the appearance of myeloid progenitor cells in conjunction with elevated numbers of fully functional neutrophils in the peripheral blood. The leukaemic effect results from (i) inhibition of apoptosis accompanied with increased proliferation, at least at the stem cell level, and impaired growth factor dependency, (ii) mitogenic activation, and (iii) altered cell adhesion to stromal cells and the extracellular matrix (Bedi, *et al.* 1994, Gordon, *et al.* 1987, Puil, *et al.* 1994).

BCR-ABL1 is detectable in haematopoietic stem cells, early progenitor cells and differentiated myeloid cells as well as B-lymphocytes (Haferlach, *et al.* 1997, Martin, *et al.* 1980, Weber-Matthiesen, *et al.* 1992). However, the involvement of T-lymphocytes has been controversial, with the consensus probably being that a small proportion of T-cells are Ph-positive in some cases (Allouche, *et al.* 1985, Bartram, *et al.* 1987, MacKinney, *et al.* 1993). Under certain conditions, that are so far unknown in detail, the acquisition of the Ph chromosome by a single, pluripotent haematopoietic stem cell may give rise to a malignant cell clone that leads to leukaemia. The ABL1 tyrosine kinase activity under normal physiological conditions is tightly regulated, mainly by the inhibitory effect of its own SRC

homology 3 (SH3) domain and N-terminal myristoylation modification. Activation may be triggered by binding of the ABL interactor protein (ABI) 1 or 2 to the ABL SH3 domain, but also by internal interaction (Cicchetti, *et al.* 1992, Dai and Pendergast 1995, Mayer and Baltimore 1994, Shi, *et al.* 1995). The activation of ABL1 may be initiated by oxidative stress to the cell and followed by dissociation of the PAG protein from ABL1 (Wen and Van Etten 1997). In *BCR-ABL1* positive CML, a coiled-coil motif close to the amino terminal end of BCR induces dimerisation followed by autophosphorylation of the tyrosine kinase fusion protein. The auto-inhibitory effect of the SH3 domain is lost in this scenario (McWhirter, *et al.* 1993). The same effect may be achieved by fusion of *ABL1* to a partner other than *BCR*, e.g. *TEL* (*ETV6*) (Golub, *et al.* 1996).

Several signal transduction pathways are activated by BCR-ABL1. Adapter molecules (CRKL, $p62^{DOK}$) (Carpino, *et al.* 1997, Oda, *et al.* 1994), proteins associated with the organisation of the cytoskeleton and the cell membrane (paxillin and talin) (Salgia, *et al.* 1995a, Salgia, *et al.* 1995b, Salgia, *et al.* 1995c), and proteins with catalytic function (FES, SYP) are a few of the BCR-ABL1 substrates. The mechanism of signal transduction after activation is not yet fully understood. Mitogenic, proliferative, pro- or anti-apoptotic signals may be released simultaneously by several pathways and, therefore, potentially act against each other. Depending on the host cell and other, so far unknown, factors, the balance between them may vary. Two of the most prominent pathways activated by BCR-ABL1 are the RAS/MAPK/ERK pathway (Deininger, *et al.* 2000, Raitano, *et al.* 1997) and the JAK/STAT pathway, although several others are also important (Ilaria and Van Etten 1996). While the role of the signal transducers and activator of transcription (STAT) proteins 1 and 5 has been well established in CML, involvement of JAK2 remains controversial (Hantschel, *et al.* 2012). Interestingly, phosphorylation of STAT6 has been observed without exception in p190 BCR-ABL1 positive cell-lines and, therefore, linked with a lymphoid phenotype (Ilaria and Van Etten 1996).

Both, the RAS/MAPK/ERK pathway and the JAK/STAT pathways are tightly involved in the signal transduction cascades induced by growth factors.

Constitutive activation by a tyrosine kinase explains the characteristic growth factor hypersensitivity of chronic phase CML. A third pathway, the phosphatidylinositol-3 (PI3) kinase pathway, mediates an anti-apoptotic signal by blocking the effect of the pro-apoptotic BAD protein with a constitutive enhancement of the anti-apoptotic effect of BCL_{XL} (Horita*, et al.* 2000). The ABL1 kinase-inhibiting effect of proteins ABI1 and ABI2 may be negated by proteasome-mediated degradation of these molecules (Dai*, et al.* 1998). The altered adhesion of CML cells to stromal cells and extracellular matrix within the bone marrow is believed to be an important mechanism by which the malignant cells escape normal inhibition of cell proliferation (Deininger*, et al.* 2000). This adhesive defect is believed to result from up-regulation of β–integrin in BCR-ABL1 positive progenitors (Jiang*, et al.* 2000).

1.2.3.1.4 Therapeutic options

For a long time, treatment options for CML were few and largely palliative. Conventional chemotherapy with the alkylating agent busulfan was in use for decades, but it controls symptoms for only a few months (Sawyers 1999). Hydroxyurea (HU) treatment was slightly more effective with prolonged survival and less toxic side effects (Hehlmann*, et al.* 1993). In 1981, interferon (IFN)-based therapies were introduced. For the first time, haematologic, cytogenetic and, in very few patients, molecular remissions were achieved. Results were improved by combination with cytosine arabinoside (Ara-C). The achievement of a cytogenetic or molecular remission was associated with prolonged survival and better prognosis (Allan*, et al.* 1995, Chronic Myeloid Leukemia Trialists' Collaborative 1997, Guilhot*, et al.* 1997, Hehlmann*, et al.* 2003, Hehlmann*, et al.* 1994, The Italian Cooperative Study Group on Chronic Myeloid Leukaemia 1994,, Ohnishi*, et al.* 1995). However, few patients reached a sustained completed molecular remission. A total eradication of the host haematologic tissue and its replacement with an allogeneic graft by transplantation is considered to be the only therapeutic option with a 'curative' effect, i.e. no evidence for both Ph and a *BCR-ABL1* transcript. The approach is challenged by a high transplant-related mortality, graft versus host

(GvH) side effects, and relapses (Hehlmann, *et al.* 2007). Indeed, a superior effect of an IFN-*alpha*-based drug treatment compared to allogeneic stem cell transplantation with regard to overall survival (Hehlmann, *et al.* 2007).

The impact of complete cytogenetic response on survival of IFN-treated patients was established in a meta-analysis of published studies (Bonifazi, *et al.* 2000). The introduction of the selective tyrosine kinase inhibitor imatinib to the treatment of CML in 1998 has provided a new therapeutic option, which has rapidly become the standard of care. In early chronic phase, imatinib induces remarkably high rates of haematologic, cytogenetic and, in many cases, molecular remission (Druker, *et al.* 2006). Advantages were also seen in late chronic phase after failure of previous treatments, e.g. IFN, and accelerated and blastic phase CML, although the superior effect declines with advanced disease stages. Relapses occur mainly due to resistance to imatinib by means of mutations of secondary mutations within the *BCR-ABL1* tyrosine kinase domain, clonal evolution with additional cytogenetic aberrations followed by disease progression, altered activity of influx or efflux transporters and other mechanisms that appear to be independent of BCR-ABL1 (Gorre, *et al.* 2001, Hochhaus, *et al.* 2001, Hochhaus, *et al.* 2002, Hughes, *et al.* 2006, Khorashad, *et al.* 2006, Lahaye, *et al.* 2005). So far, only a few cases with sustained Ph and *BCR-ABL1* negativity after imatinib withdrawal have been reported. In this situation, relapses occur often and are most likely due to BCR-ABL1 activation in previously quiescent stem cells (Mahon, *et al.* 2010, Rousselot, *et al.* 2007).

As for all malignancies, considerable effort has been devoted to defining significant prognostic indicators that may be used to guide therapy in CML. The Sokal score is based on standard clinical parameters and has been used for many years to define high, intermediate, and low risk patients (Sokal, *et al.* 1984). A modification of this, known as the Hasford or EURO score, was specifically developed for patients treated with IFN (Hasford, *et al.* 1998) and subsequently the EUTOS score for patients treated with imatinib (Hasford, *et al.* 2011). However molecular markers have also been suggested to have prognostic value, most notably the finding of heterogeneous deletions at the *ABL1-BCR* fusion point on the reciprocal 9q+ chromosome. These deletions are seen in approximately 15 % of CML cases and

have been associated with adverse outcome (Hasford, *et al.* 2011, Huntly, *et al.* 2001, Kolomietz, *et al.* 2001, Sinclair, *et al.* 2000).

1.2.3.2 Polycythaemia vera (PV)

The first extensive description of polycythaemia (rubra) vera (PV) by Vaquez dates back to 1892 (Vaquez 1892). A few years later, PV was further characterised by Osler as combining cyanosis, polycythaemia and moderate splenomegaly. Following the observation of similar clinical symptoms of four other haematological entities, Damashek included PV into the syndrome complex of myeloproliferative disorders in 1951 (Damashek 1951). The incidence of PV in an investigated northern European population is about 1.5 patients per 100,000 (Kutti and Ridell 2001). PV is an acquired multipotential stem cell defect with an excessive clonal production of erythrocytes and variable overproduction of leukocytes and platelets. Characteristically, all myeloid cell lineages present with a clonal pattern, i.e. reticulocytes, granulocytes, and platelets. Occasionally, B-lymphocytes are clonal as well. The expansion of the erythrocyte clone is growth factor independent and erythroid precursors present with a characteristic hypersensitivity to EPO that is used as a diagnostic tool to distinguish PV from familial congenital polycythaemia (Tefferi 2006b). Loss of heterozygosity (LOH) of chromosome 9p was observed as a common recurrent abnormality in PV patients (Kralovics, *et al.* 2002). This finding led to the recent discovery of the *JAK2*^{V617F} point mutation in >90 % of PV patients with a high proportion of them showing homozygosity of the mutation (Baxter, *et al.* 2005, James, *et al.* 2005, Kralovics, *et al.* 2005, Levine, *et al.* 2005b). Homozygosity results from mitotic recombination and subsequent selection for *JAK2* mutant copies in a process known as acquired uniparental disomy (Jones, *et al.* 2005). The mutation is mainly detectable in cells of the myeloid lineage, but occasionally also in B- and T-lymphocytes and in lympho-myeloid progenitors (Delhommeau, *et al.* 2007, Ishii, *et al.* 2006, Larsen, *et al.* 2007). The *JAK2* mutation prevents erythroid progenitor cells from undergoing apoptosis, and it promotes proliferation (Baxter, *et al.* 2005, James, *et al.* 2005). The mutation occurs in the pseudo-kinase domain of JAK2 that harbours auto-

regulatory properties (James*, et al.* 2005). Several pathways, including JAK2/STAT5, PI3-Kinase/AKT, and SRC, are involved in the signalling cascade that is activated by mutant JAK2 (Heinrich*, et al.* 1998, James*, et al.* 2005). The homozygous variant is associated with longer disease duration and more complications, e.g. fibrosis, haemorrhage, thrombosis, and the need for cytoreductive therapy (Tefferi*, et al.* 2005). Clonality studies indicate that *JAK2*^{V617F} may not be the initial event in the development of PV, although this remains controversial (Cross 2011, Kralovics*, et al.* 2006). Other mutations of *JAK2* upstream of V617F were detected in a few PV patients that were negative for V617F. The mutations embrace point mutations (i.e. H538QK539L, K539L), deletions (N542-E543del), and insertions (F537-K539delinsL) between amino acids 537 and 543 (Campbell*, et al.* 2006a, Scott*, et al.* 2007).

1.2.3.3 Essential Thrombocythaemia (ET)

In his work on the myeloproliferative disorders, William Dameshek included a syndrome with a characteristic thrombocytosis of the peripheral blood and a marked myeloid metaplasia of the bone marrow called thrombocythaemia (Damashek 1951). Twenty years earlier, Epstein already described a case with 'haemorrhagic thrombocythaemia'. But it was only much later that it became clear that primary or essential thrombocythaemia (ET) is an independent clinical entity (Harrison 2002).

The incidence of ET was reported as 1.5 in 100,000 in a Swedish study of 2001 (Kutti and Ridell 2001). The latest diagnostic criteria basically represent those proposed by the Polycythaemia Vera Study Group (PVSG) and involve an elevated platelet count of more than 600,000 per microlitre whole blood and exclusion of other diagnoses indicated by reactive thrombocytosis, the Philadelphia chromosome, or the *BCR-ABL1* translocation, collagen fibrosis of the bone marrow, or an elevated haematocrit (Harrison and Green 2006). Thrombosis or haemorrhagic events are the main complications associated with ET. Thrombosis is associated with an increased baseline leukocyte count and more likely to occur

in elderly patients over the age of 60 with a previous event of thrombosis (Carobbio, *et al.* 2007, Cortelazzo, *et al.* 1990). As ET is a clonal haematologic disorder, thrombopoietin-independent megakaryocyte colony assays and X-chromosome inactivation pattern studies have been proposed as additional diagnostic tools but so far have not proved their potential in the routine laboratory (Fialkow, *et al.* 1981). The clonal character of ET has been challenged by several studies, as nearly 50 % of patients demonstrate polyclonality using standard clonality tests. Clinically, they are less likely to suffer from thrombotic complications than others (Garsa, *et al.* 2005).

The molecular event leading to ET is only partially understood. Cytogenetic alterations have only been detected in the minority of cases with 95 % of cases being normal. The stimulating cytokine for megakaryocyte growth is thrombopoietin (TPO), which has been demonstrated to be elevated in most ET patients, but also in other MPN. In contrast to PV and primary myelofibrosis (PMF) patients, the Myeloproliferative Leukaemia Virus-Oncogene (*MPL*, i.e. the thrombopoietin receptor) is normally expressed in ET patients (Moliterno, *et al.* 2004). However, a mutation of the *MPL* gene has only been described for patients with familial ET, but not in sporadic cases (Ding, *et al.* 2004). The acquired V617F mutation of *JAK2* has been detected in more than 50 % of ET patients (Baxter, *et al.* 2005, James, *et al.* 2005, Levine, *et al.* 2005a). In many cases the mutation is seen at a relatively low level and often in cases with 'polyclonal' disease, indicating that ET is indeed a clonal disorder but the size of the clone is variable. The clinical impact of the mutation for ET patients is not clear and thus far an association of mutation status and risk for a thrombotic event has not been made (Carobbio, *et al.* 2007). Biologically and clinically, $JAK2^{V617F}$ positive ET more closely resembles PV that $JAK2^{V617F}$ negative ET, which has led to the suggestion that $JAK2^{V617F}$ positive PV and ET are basically the same disease (Campbell, *et al.* 2005).

1.2.3.4 Primary Myelofibrosis (PMF)

PMF, previously known as idiopathic myelofibrosis, myelofibrosis with myeloid metaplasia or osteomyelofibrosis, was first described at about the same time as PV and later linked with PV, ET, and CML as another form of chronic myeloid disorder. The incidence of PMF is below that of PV and ET with an estimated 0.4 cases in 100,000 (Kutti and Ridell 2001, Mesa*, et al.* 2006). The median age is 67 years. The characteristic clinical feature is a bone marrow fibrosis in association with an abnormal megakaryocyte proliferation and maturation, splenomegaly due to extramedullary haematopoiesis, and anaemia. The peripheral blood smear shows circulating immature granulocytes and erythroblasts. Clonality studies showed different clones of a tri-lineage proliferation of stem cell origin, i.e. granulocytes, erythrocytes, and megakaryocytes. In a few patients, B- or T-lymphocytes are also involved (Chagraoui*, et al.* 2006). The myelofibrosis, however, results from the polyclonal proliferation of stromal cells, i.e. fibroblasts, derived from bone marrow (Takai*, et al.* 1998). The characteristic collagen fibrosis, osteogenesis and osteosclerosis are driven by an increased cytokine release, i.e. PDGFA, FGF, and TGFB1, of the abnormal megakaryocyte and monocyte clones (Bousse-Kerdiles and Martyre 2001, Martyre*, et al.* 1997, Rameshwar*, et al.* 1998). Dysmegakaryopoiesis may be induced by increased levels of growth factors, e.g. thrombopoietin (TPO). Recently, an intrinsic activation of the growth factor pathway by the acquired *JAK2*^{V617F} mutation has been detected in 30-50 % of PMF patients (Baxter*, et al.* 2005, James*, et al.* 2005, Jones*, et al.* 2005, Kralovics*, et al.* 2005, Levine*, et al.* 2005b), many of whom are homozygous. Clinically, the mutation is associated with a poorer survival (Campbell*, et al.* 2006b). One of the pathogenetic key features is the elevated expression of *TGFB1*, as it not only promotes myelofibrosis, but also osteosclerosis (Martyre*, et al.* 1997).

1.3 Aims of the project

The aims of this project are to explore the incidence and impact of both acquired and constitutional genetic differences in chronic myeloid disorders. The work described here focuses on three distinct aspects of these diseases:

- What is the impact of deletions of the derivative chromosome 9 on the prognosis of patients with CML?

- Are components of signalling pathways commonly deregulated in atypical myeloproliferative disorders and, if so what is their significance?

- Is a genetic variability of genes involved in the interferon-signalling pathway the reason for a heterogeneous response to interferon-*alpha* in first-line therapy of CML patients?

2 Patients, materials and methods

2.1 Patients

Whole blood samples or bone marrow specimen of patients with clinical diagnoses of either CML or chronic MPN were sent to the laboratories in Mannheim, Germany, and Salisbury, UK, for molecular analysis. In the case of CML, analysis included expression and quantity of the *BCR-ABL1* transcript and in some cases the expression of the *ABL1-BCR* transcript. The analysis has been carried out in context of international multi centre trials and was approved by the local Ethics Committee of the Medical Faculty Mannheim of the Heidelberg University. Specimens from patients with chronic MPN and other disorders were sent to the laboratories in either Mannheim or Salisbury for individual testing for expression of known fusion genes such as *BCR-ABL1, FIP1L1-PDGFRA* or a range of other abnormalities. Further investigation was approved by the Salisbury and South Wiltshire Research Ethics Committee in the approved study entitled 'Mechanisms and consequences of tyrosine kinase activation in chronic myeloproliferative disorders and related conditions' (LREC study number O5Q2008/6), either following informed consent or anonymisation.

2.2 Material

2.2.1 Equipment and material

2 µl Pipette	Anachem, Luton, UK
10 µl Pipette	Anachem, Luton, UK
20 µl Pipette	Anachem, Luton, UK
200 µl Pipette	Anachem, Luton, UK
1000 µl Pipette	Anachem, Luton, UK
1000 µl Biosphere Tips, plugged	Sarstedt Ltd., Leicester, UK
96 well SEQ plate	Millipore Ltd., Watford, UKABI Prism
ABI Prism Genetic Analyzer 3730	Applied Biosystems, Warrington, UK
ART 20P Pipette Tips, plugged	Fisher Scientific, Loughborough, UK
ART 20E Pipette Tips, plugged	Fisher Scientific, Loughborough, UK
ART 200 Pipette Tips, plugged	Fisher Scientific, Loughborough, UK
Centrifuge Thermo IEC Micromax RF	Thermo Scientific, Basingstoke, UK
Centrifuge tubes, 50ml	Sarstedt Ltd., Leicester, UK
Class II Biological Safety Cabinet	Nuaire, Caerphilly, UK
Collection tube, 2ml	Qiagen Ltd, Crawley, UK
Corbett Rotor Gene 6500 PCR machine	Corbett Research Ltd., Cambridge, UK
Genetic Analyzer 3100	Applied Biosystems, Warrington, UK
Glass flask	Fisher Scientific, Loughborough, UK
Gel tank	Fisher Scientific, Loughborough, UK
Micro tube with cap, 200 µl	Sarstedt Ltd., Leicester, UK
Micro tube with cap, 0.5ml	Sarstedt Ltd., Leicester, UK
Micro tube with cap, 1.5ml	Sarstedt Ltd., Leicester, UK
Micro tube with cap, 2ml	Sarstedt Ltd., Leicester, UK
Microsoft Office Excel 2003 SP2	Microsoft Corp., Richmond, U.S.A.
Microwave oven EM-SL50S	Sanyo Europe Ltd., Watford, UK
MSI Minishaker	IKA Works, Inc., Wilmington, U.S:A.
Mutation Surveyor software	Soft Genetics, State College, PA, U.S.A.
NanoDrop ND-1000 Spectrophotometer	NanoDrop Technol., Wilmington, U.S.A.
Neubauer haemocytometer	Fisher Scientific, Loughborough, UK
OLIGO 5 software	Mol. Biol. Insights Inc., Cascade, U.S.A.
PCR plates 96 well half skirt	Thistle Scientific, Glasgow, UK

Prism version 5.00 for Windows	GraphPad Software, San Diego, U.S.A.
QIAamp Spin Column	Qiagen Ltd, Crawley, UK
Tetrad automated thermocycler with heated lid	MJ Research, Waltham, MA, U.S.A.
Tube caps, 0.1ml for Corbett Rotor Gene	Corbett Research Ltd, Cambridge, UK
UV transilluminator	Syngene, Cambridge, UK
Vortex-Genie 2	VWR International, Leicestershire, UK

2.2.2 Reagents

2-Mercaptoethanol (β-Mercaptoethanol)	Sigma-Aldrich Ltd., Gillingham, UK
6 x Orange Loading Dye	Promega, Southampton, UK
Acetic Acid	Sigma-Aldrich Ltd., Gillingham, UK
Agarose	Bioline, London, UK
AmpliTaq Gold DNA polymerase	Applied Biosystems, Warrington, UK
AmpliTaq Gold PCR buffer II	Applied Biosystems, Warrington, UK
BigDye Terminator v1.1 sequencing reaction kit	Applied Biosystems, Warrington, UK
Boric Acid	Sigma-Aldrich Ltd., Gillingham, UK
Buffer AL	Qiagen Ltd, Crawley, UK
Buffer ATL	Qiagen Ltd, Crawley, UK
Buffer AW1	Qiagen Ltd, Crawley, UK
Buffer AW2	Qiagen Ltd, Crawley, UK
Buffer AE	Qiagen Ltd, Crawley, UK
Buffer P1	Qiagen Ltd, Crawley, UK
Buffer P2	Qiagen Ltd, Crawley, UK
Buffer N3	Qiagen Ltd, Crawley, UK
Buffer PB	Qiagen Ltd, Crawley, UK
Buffer PE	Qiagen Ltd, Crawley, UK
Buffer EB	Qiagen Ltd, Crawley, UK
Deoxynucleotide triphosphates (dNTPs, i.e. a set of dATP, dGTP, dCTP, dTTP)	Promega, Southampton, UK
EDTA (ethylene-diamine-tetra-acetic acid)	Qiagen Ltd, Crawley, UK
Ethanol	VWR International Ltd., Lutterworth, UK
Ethidium bromide	Sigma-Aldrich Ltd., Gillingham, UK
Exonuclease I	New England Biolabs; Ipswich, U.S.A.
GeneScan-ROX 500	Applied Biosystems, Warrington, UK
Guanidine thiocyanate (GTC)	Sigma-Aldrich Ltd., Gillingham, UK
HCl (hydrochloric acid)	Sigma-Aldrich Ltd., Gillingham, UK
HiDi-Formamide	Applied Biosystems, Warrington, UK
Injection Solution	Millipore Ltd., Watford, UK
KCl (potassium chloride)	Sigma-Aldrich Ltd., Gillingham, UK
LCGreen Plus	Cadama Medical Ltd, Stourbridge, UK
Loading buffer, 1kb plus	Invitrogen Ltd, Paisley, UK

$MgCl_2$ (magnesium chloride)	Applied Biosystems, Warrington, UK
MLPA reaction kit	MRC Holland BV, Amsterdam, NL
MLPA Ligase-65 buffer A and B	MRC Holland BV, NL
MLPA Ligase-65	MRC Holland BV, NL
Montage SEQ_{96} Sequencing Reaction Cleanup	Millipore Ltd., Watford, UK
N-Lauroylsarcosine sodium salt	Sigma-Aldrich Ltd., Gillingham, UK
NaCl (sodium chloride)	Sigma-Aldrich Ltd., Gillingham, UK
Phosphate Buffer Saline tablets	Sigma-Aldrich Ltd., Gillingham, UK
POP6-Polymer	Applied Biosystems, Warrington, UK
Proteinase K (0.6U/ml)	Qiagen Ltd, Crawley, UK
Proteinase K (20mg/ml)	Promega, Southampton, UK
QIAamp Blood DNA Mini Kit	Qiagen Ltd, Crawley, UK
QIAprep Spin Miniprep Kit	Qiagen Ltd, Crawley, UK
Red Blood Cell Lysing Buffer	Sigma-Aldrich Ltd., Gillingham, UK
Salsa Enzyme Dilution Buffer	MRC Holland BV, NL
Salsa MLPA Buffer	MRC Holland BV, NL
Salsa PCR Buffer	MRC Holland BV, NL
Salsa PCR Primer Mix (incl. dNTP)	MRC Holland BV, NL
Salsa Polymerase	MRC Holland BV, NL
SDS (sodium dodecyl sulfate)	Sigma-Aldrich Ltd., Gillingham, UK
Shrimp alkaline phosphatase	Roche Diagnostics, Burgess Hill, UK
Sodium Citrate	Sigma-Aldrich Ltd., Gillingham, UK
Tris base	Sigma-Aldrich Ltd., Gillingham, UK
Titanium Taq DNA polymerase reaction kit	Clontech, Saint-Germain-en-Laye, France
Trypan Blue	Sigma-Aldrich Ltd., Gillingham, UK
Water, purified	Invitrogen Ltd, Paisley, UK

2.2.3 Recipes

Tris-EDTA buffer (TE buffer; 1x)

10 mMTris-HCl, pH 7.5
1 mM EDTA

Tris-Borate EDTA buffer (TBE buffer; 10x)

1M Tris base
1M Boric Acid
20mM EDTA
adjusted to pH 8.3

Tris-HCl (1 M)

121.1 g Tris base
add purified water up to 1000 ml
adjusted to pH 8

Phosphate Buffered Saline (PBS)

1 tablet
dissolved in 200 µl purified water

Guanidine thiocyanate (GTC) solution

4M Guanidiumthiocyanate
5mM EDTA
25mM Sodium citrate pH7.0
0.5 % N-Lauroylsarcosine sodium salt
7 µl β-Mercaptoethanol per 1 ml of GTC was
added fresh prior to use.

2.3 Methods

2.3.1 Sample preparation

2.3.1.1 Red cell lysis and storage options

For an optimal molecular diagnostic analysis, a volume of 10 to 40 ml PB or 1 to 4 ml BM in either preservative-free heparin or ethylenediaminetetraacetic acid (EDTA) was required. White blood cell counts were performed using the Neubauer haemocytometer (Fisher Scientific). PB or BM was diluted 1 in 10 in a 2 % acetic acid solution (Sigma) to lyse the red blood cells, and separated cells were diluted in trypan blue (Sigma) to differentiate alive and dead cells.

The number of 1 x 10^8 cells or a maximum volume of 15 ml whole blood was transferred to a clean 50 ml centrifuge tube (Sarstedt). Ice-cold Red Blood Cell Lysing Buffer (Sigma) was added up to a total volume of 50 ml. The suspension was thoroughly mixed by inverting and left on ice for ten minutes, with intermittent shaking. After centrifugation at 1,000 x g at a centrifuge 'Thermo IEC Micromax RF' (Thermo Scientific), the pellet of the remaining white blood cells was washed in a volume of 20 ml PBS (Sigma) and spun again at 3,000 x g for ten minutes. Optionally, the washing step was repeated to clean the sample from remaining red cell fragments. After washing, the pellet was either directly stored at minus 70°C, or it was lysed in a final volume of up to 2.5 ml GTC solution (Sigma) by passing several times through a needle for later DNA or RNA extraction and stored at minus 20°C.

2.3.1.2 DNA extraction

DNA extraction may be performed from any source containing traces of nucleic acid. Based upon the ethanol precipitation protocol, several commercial kits are available and well established.

2.3.1.2.1 Isolation of genomic DNA using proteinase K

Cell pellets were resuspended in 1 ml of PBS per 1 x 10^7 cells and transferred to a clean 1.5 ml micro tube. 100 µg/ml of Proteinase K solution (Sigma) and 30 µl of 10 % SDS (Sigma) was added to digest cell proteins and remove cell lipids. The sample was thoroughly mixed and incubated overnight at 37°C. To precipitate non-DNA components such as protein and other cell debris, 300 µl of a 6M NaCl-solution (Sigma) was added and mixed by inverting the tube. Centrifuging at 5,000 x g for 30 minutes at room temperature left a supernatant containing DNA which was collected in a clean 2 ml micro tube. The equivalent volume of 99 % ethanol (VWR International Ltd.) was added. The sample was mixed for 15 to 30 seconds using an MSI Minishaker (IKA Works, Inc.). Precipitated DNA became visible as a white 'cotton-like' cloud, which was collected using a sterile needle, and transferred to a clean 1.5 ml micro tube. The pellet was washed in 1 ml of 70 % ethanol (VWR International Ltd.) to remove residual salts. After air-drying, the pellet was dissolved in 50 to 500 µl of 1 x TE buffer, depending on the yield of DNA. The average length of an isolated DNA strand is usually 100 – 150 kb. Optionally, the sample was stored at minus 20°C for 30 to 60 minutes to encourage the precipitation procedure. Then it was spun again at 5,000 x g for 20 minutes, and the resulting pellet was dissolved in 50 µl of 1 x TE buffer. DNA samples were kept either at 4°C for medium-term use or at minus 20°C for long-term storage.

2.3.1.2.2 Genomic DNA extraction using QIAamp Blood DNA Mini Kit

As an alternative method, the QIAamp Blood DNA Mini Kit by Qiagen (Qiagen Ltd., Crawley, UK) allows DNA extraction of several sources in less time than by ethanol precipitation as described above. The estimated yield after extraction out of 200 µl whole blood is 3 to 12 µg of genomic DNA in fragments of up to 50 kb length. 20 µl of Proteinase K (0.6 U/ml) was added to 200 µl whole blood in a clean 1.5 ml micro tube (Sarstedt). 200 µl of AL Buffer (Qiagen) was added and mixed by pulse-vortexing for 15 seconds at a vortexer (Vortex-Genie 2 by VWR International Ltd.). After incubation for ten minutes at 56°C, 200 µl of 99 % ethanol (VWR

International Ltd.) was added and thoroughly mixed by vortexing for 15 seconds. The sample mix was transferred to a QIAamp Spin Column (Qiagen) in a 2 ml collection tube (Sarstedt) and centrifuged at 6,000 x g for one minute. The filtrate was discarded and the spin column transferred to a clean 2 ml collection tube. By applying 500 µl of Buffer AW1 (Qiagen) and centrifuging for 1 minute at 6,000 x g, the spin column was washed. After discarding the filtrate and transferring the column to a clean 2 ml collection tube, the column was washed with 500 µl of Buffer AW2 (Qiagen) and centrifuged at 20,000 x g for three minutes. Optionally, the column was centrifuged again at 20,000 x g for one minute to allow drying. The column was transferred to a clean 1.5 ml micro tube and 100 - 200 µl AE Buffer (Qiagen) were applied. Centrifugation at 6,000 x g for one minute eluted the bound DNA. For maximal yield, the AE Buffer was allowed to incubate in the column for at least five minutes before spinning.

As a third method, DNA extraction was performed from samples stored in GTC solution. Hereby, 200 µl of the GTC-stored sample was transferred to a clean 2 ml micro tube. After adding 200 µl of AL Buffer (Qiagen) and incubating at 56°C, 200 µl of 96 % ethanol (VWR International Ltd.) was added, and the mixture was transferred to a QIAamp Spin Column (Qiagen) in a 2 ml collection tube. The following steps were identical to the DNA extraction procedure described in the QIAamp Blood DNA Mini Kit protocol above.

2.3.1.2.3 OD spectrography

DNA concentration was determined by measuring absorbance at 260 nm in a NanoDrop Spectrophotometer (NanoDrop Technologies). For nucleic acid quantification, the adjusted Beer-Lambert equation was used.

$$c = (A*e)/b$$

c is nucleic acid concentration in ng/µl A is absorbance in U

e is extinction coefficient b is path length in cm (for NanoDrop 0.1 cm)

A minimum volume of 1 µl of DNA solution was required for the process. The procedure followed instructions provided by the manufacturer. The purity of DNA samples was ensured by determining the 260/280 nm ratio, which ideally should be about 1.8. Samples that were more than 0.1-0.2 units away from this figure tended to amplify poorly and if possible were re-extracted.

5.3.1.3 RNA extraction and cDNA synthesis

Total peripheral blood leukocytes from patient samples had been stored in guanidine thiocyanate (GTC) solution at -20°C for several years in Mannheim as part of the German clinical trial sample collection process. GTC solution is regularly used in molecular biology laboratories to lyse cells for subsequent RNA and DNA extractions.

For RNA extraction, the RNeasy mini kit by Qiagen was used. To economise with regard to sample usage, 100 µl of GTC lysates were mixed with 250 µl of RLT buffer (part of RNeasy Mini Kit, Qiagen) in a 1.5ml micro tube with cap. 350 µl of 70 % ethanol was added and mixed by pipetting. The mix then was applied to an RNeasy mini column in a 2 ml collection tube (both part of RNeasy Mini Kit) and centrifuged at 8,000 x g (10,000 rpm) for 30 seconds. The flow-through was discarded. 700 µl of buffer RW1 (part of RNeasy mini kit) was applied and centrifuged at 8,000 x g for 30 seconds. The flow-through again was discarded. 500 µl of buffer RPE (part of RNeasy mini kit) was applied and centrifuged at 8,000 x g for 30 seconds. The flow-through was discarded. 500 µl of buffer RPE (part of RNeasy mini kit) applied and centrifuged at 8,000 x g for 2 minutes. The whole collection tube with flow-through then was discarded. The RNeasy mini column was transferred to a new 2 ml collection tube (part of RNeasy Mini Kit) and centrifuged at maximum speed for one minute. The whole collection tube with flow-through was again discarded. The RNeasy mini column was transferred to a new 1.5 ml micro tube with cap. 30 µl of RNase-free water (part of RNeasy Mini Kit) was applied and kept on bench for 2 minutes. Then, the micro tube containing the RNeasy mini column was centrifuged at 8,000 x g for one minute. The mini

column was discarded and the flow through (typically about 25 µl containing 1-5 µg of total RNA)was placed on ice, immediately, to avoid degradation.

For cDNA synthesis, the 1-2 µg RNA solution in a micro tube was denatured by 65°C in a hot block and quick-chilled on ice. A list of reagents as shown in table 5 was added while the micro tube remained on ice.

Table 5: List of reagents for cDNA synthesis

Tris pH 8.3	50mM
KCl	75mM
$MgCl_2$	3mM
DTT	1mM
dATP, dCTP, dTTP and dGTP	1mM, each
Random $pd(N)_6$hexamers	100 µg/µl
Murine Moloney Leukaemia Virus (MMLV) reverse transcriptase	14,000 U
RNAse inhibitor	1,400 U/ml

The micro tube containing the mix was placed in a 37°C hot block (alternatively, a Tetrad thermocycler) and incubated for at least 120 minutes. The enzymatic reaction was stopped by placing the tubes containing the mix on a 65°C hot block. The cDNA was stored at -20°C for several months.

2.3.2 Standard polymerase chain reaction

The polymerase chain reaction (PCR) is a rapid and easy to perform approach for amplifying DNA fragments. Its advantages embrace high sensitivity and robustness. A clear disadvantage of PCR is the limited extend of the size of the amplification products and the number of mutations introduced during the amplification process, particularly for long templates. Several techniques such as long-range PCR or use of proofreading polymerase have been developed to overcome the problem.

Before starting a PCR process, information about the target sequence is needed in order to design specific oligonucleotide primers. These are the start points for the simultaneous extension in each direction of a double stranded DNA. Primer design is a critical point and should be performed carefully. Primers are usually about 20 to 25 nucleotides long with a GC content of 40 to 60 %. The melting temperature of both primers should not differ by more than 5°C as the annealing temperature in an optimised reaction is dependant on it. Internal repeats or inverted sequences more than three base pairs (bp) in length should be avoided as they may impair the allele-specific amplification.

To design the optimal primer pair for a PCR reaction and to calculate the specific melting temperature, several calculations have been proposed. A comfortable way to determine the melting temperature is by using the Wallace rule (Wallace*, et al.* 1979):

$$T_M = 2 * (A+T) + 4 * (G+C)$$

TM is the melting temperature
A, T, G, and C are the number of occurrences of each nucleotide

In practice, an optimal design of primers was typically achieved by the use of a specific software program such as the OLIGO 5 software (Molecular Biology Insights Inc). The PCR was basically conducted by a heat-stable DNA polymerase, e.g. AmpliTaq Gold DNA Polymerase (1-5 U; Applied Biosystems), and consisted of a number of cycles of three successive reactions. The four deoxynucleotides dATP, dCTP, dGTP, dTTP (200µM each; Promega), 1 x PCR Buffer II (10mM Tris-HCl pH8.3, 50mM KCl; Applied Biosystems), $MgCl_2$ (1.5mM) and primer oligonucleotides (1µM; Applied Biosystems) completed the list of reagents. The mass of the template DNA was usually one picogram to one microgram per reaction.

The following is an example of the typical conditions of a single cycle in a standard PCR:

Denaturation	94°C	for 1 minute
Primer annealing	60°C	for 1 minute
DNA synthesis	72°C	for 1 minute

Commonly, there is a single prolonged denaturation step for up to ten minutes to ensure complete separation of the template DNA strands. The actual PCR consists of a sequence of 25 to 35 cycles. During the first cycle, primers localised on either strand are extended starting at the 5' position and proceeding towards 3' along the DNA template. In the following cycles, the procedure is repeated with products serving as templates in the next cycle, respectively. From the second cycle on, there are more and more templates of short length, as the sequence flanked by both primers is amplified exponentially. The final product typically contains a ten million or more fold amplification of the target sequence. Cycle conditions for special approaches are explained in detail elsewhere. A stringent policy for setting up amplification reactions and handling the PCR products was in place, both being kept spatially strictly separated. The preparation of the reaction mix was performed under semi-sterile conditions in a class II safety cabinet (Nuaire) using irradiated and plugged pipette tips. All reagents were kept chilled during the process. PCRs were performed on a PTC-225 Tetrad Thermal Cycler (MJ Research).

2.3.3 Real-time polymerase chain reaction

Real-time PCR is a quantitative method that uses a fluorescence detecting thermocycler machine to amplify specific nucleic acid sequences and simultaneously measure their copy number. There are two major research applications: (i) quantification of gene expression and (ii) mutation screening and confirmation of single nucleotide polymorphisms. For quantification of gene expression the specimen analysed is messenger ribonucleic acid (mRNA) reverse

transcribed into cDNA. PCRs were performed on the Corbett Rotor-Gene 6000 (Corbett Life Science) (Foroni*, et al.* 2011).

2.3.4 Gel electrophoresis

Nucleic acid and proteins are charged and able to migrate in an electric field. Matrices such as agarose or polyacrylamide gels allow size separation of different fragments in one approach. The matrix is basically homogenous with pores that allow molecules to passage following the electric gradient. Optimal setting of the electric field and/or the matrix concentration allows both, the visualisation of small fragments and their quantification. Agarose gels of low concentration such as 0.8 % are most suitable for longer fragments such as plasmid DNA with several kilo bases length or whole genomic DNA, whereas gels of 2 to 3 % agarose allow the visualisation of small fragments of a few hundred bases.

The following example of a 2 % agarose gel electrophoresis was mostly used in routine laboratory approaches unless stated otherwise. In a 500 ml glass flask (Fisher Scientific), 1 g of agarose (Bioline) was completely dissolved in a volume of 50 ml TBE buffer, supported by boiling in a microwave oven (Sanyo). After cooling down to about 50°C, 25 µg ethidium bromide (Sigma) was added. Ethidium bromide is a stain that intercalates into nucleic acid allowing the visualisation of DNA fragments under UV light. The stained gel was poured into a gel tank with a suitable comb (Fisher Scientific) and allowed to cool until it became solid. The PCR product was mixed with a solution of 6 x Orange Loading Dye (Promega) and loaded onto the gel floating in TBE buffer. Control ladders (e.g., 1 kb Plus Ladder; Invitrogen) were run alongside the PCR products to attain an estimation of the PCR product length. The fragments were allowed to migrate in an electric field of 100 V for approximately 30 minutes. The bands were visualised in a UV transilluminator (Syngene).

2.3.5 Multiplex ligation-dependent probe amplification

Multiplex ligation-dependent probe amplification (MLPA) was invented in 2002 as an application for the detection of duplications and deletions of genomic DNA sequences in the *BRCA1*, *MSH2,* and *MLH1* genes, the detection of trisomy 21 in Down's syndrome, the characterisation of chromosomal aberrations, and SNP detection (Schouten*, et al.* 2002). For this approach, an individual probe set was designed and MLPA reaction kits by MRC Holland (MRC Holland BV) were used.

2.3.5.1 MLPA probe design

MLPA probes hybridise with the target sequence and will be amplified by using universal MLPA primers. Therefore, probes contained gene specific sequences as well as universal MLPA primer sequences. To use several probe sets at the same time, stuffer sequences of individual length were introduced to allow up to ten different target sequences to be distinguished. Probes were designed manually and manufactured by Biomers.net GmbH (Ulm, Germany).

2.3.5.2 MLPA reaction

In this approach, only genomic DNA (gDNA) was used, even though the use of complementary DNA (cDNA) is also possible. The MLPA reaction required a thermal cycler with a heated lid (e.g., PTC-225 Tetrad Thermal Cycler by MJ Research). DNA samples were prepared that contained a standard concentration of 25 ng of genomic DNA (gDNA) per µl solution. 2 µl of gDNA were diluted in 3 µl of 1 x TE buffer in a 200 µl micro tube. The DNA was denatured for five minutes at 95°C. After adding 1.5 µl of the probe mix and 1.5 µl of the Salsa MLPA Buffer (1.5M KCl, 300mM Tris-HCl pH 8.5, 1mM EDTA; MRC Holland BV) to the DNA, the solution was incubated overnight for at least 16 hours at 60°C to allow the probes to hybridise with their target sequence. The ligation mix was prepared containing 3 µl of each, Ligase-65 Buffer A and B (2.6mM $MgCl_2$, 5mM Tris-HCl pH 8.5, 0.013 %

non-ionic detergents, 0.2mM NAD; MRC Holland BV) diluted in 25 µl dH_2O, before 1U Ligase-65 enzyme (MRC Holland BV) was added. The ligation mix was added to the reaction mix and incubated for 15 minutes at 54°C. Inactivation of the ligation process was achieved by heating up to 98°C for five minutes. The annealed and ligated probes were amplified by PCR with universal Salsa MLPA primers (MRC Holland BV; Salsa MLPA Forward primer sequence labelled for later analysis: 6-FAM-GGGTTCCCTAAGGGTTGGA; Salsa MLPA Reverse primer sequence: GTGCCAGCAAGATCCAATCTAGA). The forward primer was labelled with carboxyfluorescin (6-FAM), a fluorescent dye with an excitation of 492 nm and an emission of 517 nm, respectively. For the PCR mix, 2 µl Salsa PCR Buffer, 1 µl Salsa Enzyme Dilution Buffer, 1 µl Salsa PCR Primers (10 pmol) and 1.25 U Salsa Polymerase (all MRC Holland BV) were made up to a final volume of 15.75 µl with dH_2O in a 200 µl micro tube. 5 µl of the ligation reaction were added to the PCR mix. PCR was performed on a PTC-225 Tetrad Thermal Cycler (MJ Research).

PCR conditions:	35 cycles	95°C	for 30 seconds
		60°C	for 30 seconds
		72°C	for 1 minute
	1 cycle	72°C	for 20 minutes

2.3.5.3 Data analysis by capillary electrophoresis

The amplified MLPA probes were analyzed by capillary electrophoresis on an ABI Prism Genetic Analyzer 3100 (Applied Biosystems). PCR products were prepared by diluting 1 µl DNA in 8.9 µl HiDi-Formamide (Applied Biosystems) and 0.1 µl of GeneScan-ROX 500 (Applied Biosystems) was added as a size standard. The following run characteristics were applied: 36 mm capillaries, POP-6 polymer, run temperature 60°C, capillary filling volume 184, pre-run voltage 15 kV, pre run time 180 sec, injection voltage 1 kV, injection time 12 sec, run voltage 10 kV, data delay time 1 sec, run time 2100 sec. Analysis was performed following instructions of the manufacturer (MRC Holland BV).

2.3.6 Mutation analysis

There are several established methods for detecting unknown mutations. Conformation-sensitive capillary electrophoresis (CSCE) and high-resolution melt curve analysis (HRM) were used in this project and will be described in detail below.

2.3.6.1 Conformation-sensitive capillary electrophoresis (CSCE)

During electrophoresis, homoduplex and heteroduplex double stranded DNA (dsDNA) alleles demonstrate different mobility characteristics. Therefore, variations of fragment size or conformation changes resulting from point mutations can be distinguished. Conformation-sensitive capillary electrophoresis (CSCE) enables homoduplex and heteroduplex alleles to be differentiated by their characteristic fluorescence emission and differences in migration through a polymer matrix (Thomas*, et al.* 2001). An amplification product consisting of alleles that are absolute identical in both, forward and reverse strand sequences, will form only one type of homoduplex on denaturation and re-annealing. Fluorescence detection after capillary electrophoresis will show a single peak as shown in figure 4. In case of a heterozygous amplification product, there are four possibilities. A point mutation in one allele, but not in the other may, after denaturation and re-annealing, lead to either two mismatch combinations (wild type forward/mutant reverse and mutant forward/wild type reverse) plus two homozygous combinations (wild type and mutant). However, very often only two or three of these four predicted peaks are detected. This is mainly due to mobility variations of incomplete resolved duplex molecules (see a schematic for the CSCE principle in figure 4).

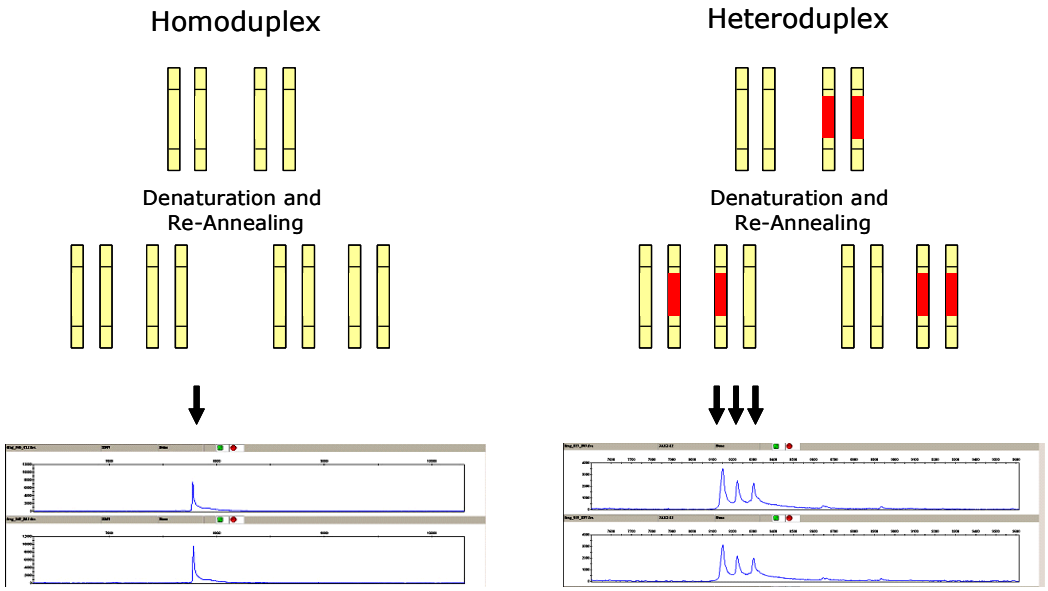


Figure 4: Schematic of the CSCE principle

Conformation-sensitive capillary electrophoresis of homozygous alleles result in a single peak as all alleles are identical (shown left) while heterozygous alleles lead to multiple signals as shown in the right hand panel.

Target sequence specific primers of 22 to 26 nucleotide length were designed with reference to the published Ensembl database (www.ensembl.org) and checked for cross homologous sequences by running a BLAST search (www.ncbi.nlm.nih.gov/blast/). Both forward and reverse primers were equipped with a universal nucleotide tag of a specific sequence that was complementary to the universal CSCE (M13) forward and reverse primer sequences. To choose optimal primer pairs and to avoid hairpin formation after adding universal primer tags, the OLIGO 5 software (Molecular Biology Insights Inc.) was used. A typical 10 µl PCR reaction mix contained ABI AmpliTaq Gold PCR Master Mix (Applied Biosystems), 8 % of Glycerol (Sigma), gene specific forward probe (100 fmol) and reverse probe (1 pmol), both obtained from Biomers.net GmbH (Ulm, Germany), 150 fmol of the universal CSCE primer mix (M13 forward primer sequence FAM-labelled: FAM-TGTAAAACGACGGCCAGT, M13 reverse primer sequence: CAGGAAACAGCTATGACC), and purified water. Optimal PCR was performed with 20 to 30 ng DNA. A summary of PCR conditions is given in table 6. PCR was performed on a PTC-225 Tetrad Thermal Cycler (MJ Research).

Table 6: PCR conditions for CSCE

The standard polymerase chain reaction (PCR) for CSCE was performed in a single tube on a Tetrad Thermal Cycler. The PCR product was diluted and analysed in an ABI DNA Analyzer 3730.

	°C	Time (sec)	
Initial denaturation	95	10:00	
Denaturation	95	00:05	
Annealing	60	00:30	30 cycles
Elongation	72	01:00	
Denaturation	95	00:05	
Annealing	55	00:30	10 cycles
Elongation	72	01:00	
Additional Elongation	72	05:00	
Denaturation before Heteroduplex formation	96	10:00	
Heteroduplex formation	96	00:20	
	-0.5	/cycle	80 cycles

The amplified product was diluted 1:200 with purified water before analysis on a 48-capillary DNA Analyzer 3730 (Applied Biosystems). The ABI Data Collection Software was used to control the DNA Analyzer, while the ABI Prism Genemapper software (all Applied Biosystems) provided the final tool for the analysis.

2.3.6.2 High-resolution melt curve analysis

An alternative procedure for mutation analysis is the high-resolution melt analysis (HRM). HRM was used as an alternative to CSCE in the latter stages of the project simply because access to the ABI3730 was limited due to heavy diagnostic use. In a closed tube procedure, PCR products and a saturating dsDNA dye (e.g., LC-Green Plus by Cadama Medical Ltd.) are processed without any additional handling procedures. The HRM process requires a previous PCR in the presence of DNA binding dyes with an ability to distinguish double stranded (ds) from single stranded (ss) DNA. During the PCR process of strand extension, dye molecules are continuously built into the dsDNA resulting in the emission of a fluorescent signal

that is different from ssDNA. In contrast to traditional SYBR Green dyes, LC-Green dye has been developed that can be used at saturating concentrations without inhibiting the amplification reaction. The use of LC-Green dye leads to an increased sensitivity and specificity. When dsDNA is melted, both strands dissociate at a temperature that is characteristic of the nucleotide sequence. Therefore, homozygote and heterozygote allele differences can be distinguished from WT nucleotide sequences. During the strand dissociation process, dye is released that leads to a decreased fluorescence signal (Wittwer, *et al.* 2003).

Target sequence specific primers of 18 to 26 nucleotide length were designed with reference to the published Ensembl database (www.ensembl.org) and checked for cross homologous sequences by running a BLAST search (www.ncbi.nlm.nih.gov/blast). The reaction conditions for a typical PCR/HRM assay are summarised in table 7.

Table 7: Amplification reaction for high-resolution melt curve (HRM) analysis

Reagent	Manufacturer/ distributor	Final concentration	Volume
Titanium Taq Polymerase	Clontech (Saint-Germain-en-Laye, France)	0.2 units	0.04µl
10X Titanium Taq polymerase buffer	Clontech (Saint-Germain-en-Laye, France)	1X	1µl
10X LCGreen Plus dye	Cardama Ltd. (London, UK)	1X	1µl
Forward primer	biomers.net (Ulm, Germany)	200nM	0.2µl
Reverse primer	biomers.net (Ulm, Germany)	100nM	0.1µl
Deoxynuclotide mix	Sigma-Aldrich (Gillingham, UK)	250nM	0.1µl
template DNA		20-30ng	1µl
dH_20			6.56µl
		Final volume	**10µl**

The optimal annealing temperature for each template was determined by performing a gradient PCR. Both PCR amplification and HRM were performed with the Corbett Rotor-Gene 6000 (Corbett Life Science). The PCR conditions were for each cycle 15 seconds of denaturation at 95°C, 20 seconds of annealing at the determined temperature, and 20 seconds of elongation at 72°C. A total of 45 cycles

were performed that were flanked by an initial denaturation step of five minutes length at 95°C and a final elongation step at 72°C for five minutes. In a real-time manner, dye fluorescence was analysed at the end of each cycle. Before starting the HRM process, two initial steps of five seconds at 95°C and 30 seconds at 50°C were performed to allow dissociation and re-assembling of dsDNA. In a gradient temperature rise by 0.1°C each two seconds, the HRM process started with 70°C and ended with 95°C. Fluorescence was measured after each melting step.

2.3.7 Direct genomic sequencing

Fred Sanger established the standard method for direct genomic sequencing in 1975 (Sanger*, et al.* 1977). The enzymatic assay is based upon amplification of single stranded DNA that acts as a template for making a new complementary DNA strand *in vitro*. The DNA strand is extended by a polymerase and in addition to the four known deoxynucleotide triphosphates (dNTPs), there are four additional dideoxynucleotides (ddNTPs) in the reaction that differ from the dNTPs by lacking a hydroxyl group and which serve as a base-specific chain terminator. As the concentration of the ddNTPs is much lower than the dNTPs, they will be incorporated by chance at a low frequency. Each ddNTP is labelled with a different fluorochrome. The resulting reaction products can then be separated by electrophoresis and the terminating nucleotide of each fragment identified by emitting light of its specific fluorochrome. The combination of all resulting DNA fragments will provide the DNA sequence structure of the sample. In case of a homogenous sample with identical DNA fragments the sequence will show no variation. However, alterations of the nucleotide follow up will be detected according their frequency (Sanger*, et al.* 1977). For most sequencing reactions, clean up of the DNA template, e.g. the PCR product, is recommended to remove dNTPs and primers. Therefore, a 30 minute enzymatic reaction (Exo/Sap) with *Exonuclease I* (New England Biolabs) and *Shrimp alkaline phosphatase* (Roche Diagnostics) was performed using a PTC-225 Tetrad Thermal Cycler.

Exo/Sap reaction mix:

Exonuclease I (20,000U/ml*)*		0.4 µl
Shrimp alkaline phosphatase	(1,000U/ml)	1.6 µl
DNA template (e.g. PCR product)		5 µl

Reaction conditions:

1 cycle of	37°C	for 15 minutes
	80°C	for 15 minutes

The sequencing reaction was performed by using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). Sequencing reactions for both, forward and reverse strand, were performed separately. Sequencing reaction conditions were applied as shown below. The thermal reaction was performed on the PTC-225 Tetrad Thermal Cycler.

Sequencing reaction mix:

BigDye Terminator v1.1 Sequencing buffer 5x	2 µl
BigDye Terminator v1.1 Sequencing reaction mix	0.5 µl
Template-specific primer 10µM	1 µl
DNA template	3-10 ng
Purified water up to the final volume of 10 µl	

Sequencing cycle conditions:

1 cycle	96°C	for 1 minute
25 cycles	96°C	for 30 second
	50°C	for 15 seconds
	60°C	for 2 minutes

In a final purification reaction, excess of salt, dNTPs, and primers was removed from the mix using the vacuum-based Montage Seq_{96} Sequencing Reaction Cleanup Kit (Millipore). The sequencing reaction mix was added to 30 µl Injection Solution (Millipore) and transferred to the 96 well SEQ plate (Millipore). A constant vacuum

of 23 to 25 mm Hg (hydrargyrum) was applied by the Millipore Vacuum Manifold until the solution had completely removed from the bottom of the well. The process was repeated by adding another 30 µl of Injection Solution and applying the vacuum. In a final procedure, the template was resuspended in 20 µl Injection Solution. 15 µl was transferred to a 96 well PCR plate and 5 µl of HiDi-Formamide (Applied Biosystems) was added and well mixed.

Capillary electrophoresis and data analysis were performed on an ABI Prism Genetic Analyzer 3100 (Applied Biosystems) with 36 mm capillaries and the POP-6 polymer. Run module 'RapidSeq36_POP6DefaultModule' applied. For dye set E the mobility file 'DT3100POP6(BD)v2.mob' was used. The raw data were analysed with the Mutation Surveyor software (Soft Genetics LLC).

2.3.8 Pyrosequencing

Pyrosequencing is a method that enables rapid, accurate, quantitative analysis of DNA sequence variation. Mostafa Ronaghi developed the technology in 1996 (Ronaghi*, et al.* 1996). In an automated high-throughput process, the analysis of SNPs and other genetic variation such as point mutations, insertions, and deletions as well as methylation status analysis is possible. Up to 500 samples per hour or a single SNP in six seconds may be analysed. Further information about the technology is available from the manufacturer's web site at www.pyrosequencing.com.

In a first step, a standard PCR using specific biotinylated pyrosequencing primers was performed. Primers were designed by using the PSQ Assay Design software version 1.0.6 (Biotage AB). All PCR forward and reverse primers as well as sequencing primer were obtained from biomers.net (biomers.net GmbH). Primers used for genotyping are listed in table 8. DNA amplification was performed in a 50 µl reaction volume with 15 pmol of each forward and reverse PCR primer, 0.2mM dNTPs (Promega), 1.5mM $MgCl_2$, 1 X Buffer II (Applied Biosystems), 1 U AmpliTaq Gold (Applied Biosystems) using 20 ng genomic DNA (see table 9 for details).

Table 8: List of SNPs and primers used for (i) PCR amplification and (ii) sequencing reaction

SNP	Forward Primer	Reverse Primer	Sequencing Primer
	ATGCCTATATGTGTGTGTATGAGT	Biotin-CTTGGCAGAGAAGTCTAATTTCAC	TTTCTTTGCACATCTATTT
rs6503695	Biotin-GTGAGCCTGAGCATATTAAAGTGA	GGCAGGGATGTAGATCCAAACT	GCACAGATGTCCAGC
rs2293152	Biotin-TGTTCCCCTGTGATTCAGA	TGCTCCCTCAGGGTCTGT	TCTCCCTAGCCCTCTC
rs2293154	TAGCTGTCAAGGGAAGAACAGTGG	Biotin-TGATGTGAGCAGGAGGGAGAC	CCCAAGTGGGGTGGA
rs16967611	Biotin-GTCAAGTCAGTCATTCTGCAACAT	CCCCTTTTTGTTATGTGTTTTACA	TTGTCAGCATTGAACC
rs6503691	Biotin-GGCTCACAAAGAAGCAAACAAG	GCAGTGATTTGGCTGTGACATAA	GGCTGTGACATAAAGATTAA
rs9900213	GATCTGACCTCCACCAGAGTATTC	Biotin-CCACCTCCTCATTAGTAGGAAAAC	GGAGAAACTAGAACGTAACC
rs17500235	CCCTCTTGGGCTACTTAAATGAC	Biotin-ATGGAAATCCAGGTAAGCA	TTCACGTACCCAGAA
rs17591972	Biotin-AGCAAACAAGGGAACTAACCA	CACCGAACCAAAACTAAATACAGA	TTTCAAATAACAAAATAAA
rs2304256	ATGTCCCGGAAGTCACAGAAGTAG	Biotin-TGGGTGTTCAGGGTTCTA	CTGCCGGCTGGCCGA
rs12720356	Biotin-AGAAGCAGGCAGGTT	TCCAAGCCATGGGCACAT	CCACGTACTCTGTCACC
rs2991269	TTGCCTAGACAGCACCGTAATG	Biotin-TGGTCCATGGAAATGTGTGTAC	GGTCACTGAGCTTGATGAA
rs310227	Biotin-TGGCAGGGTGGAACTGAAAAG	ATGAAGGAAGAAGGGCCCTACAG	TACAGGGGGAGGATG
rs310229	Biotin-CTTCTGCACCTCGATCTGAAAGT	ATGCCTGACCCCATTTTAAACA	GTAATTTCTGGTCATTTCC
rs2850015	Biotin-GAGAGCTAAGAGGGGCAGCG	GATCCCACCAGTTACATGTTCG	CACACCGCCCCTCTG
rs2257167	Biotin-TGGTCCTCCAGAAGTACATTTAGA	TTCTACACCTGAAGAGTTTTTCCA	CTGAAGAGTTTTTCCAGAT
rs7279064	Biotin-AGATGCTTTTGAGCCAGAATG	GGCATCACAGCTTGCTTCTATA	AACCAAATTAAGTGATCTGA

PCR started with an initial denaturation step at 94°C for 15 minutes. The next step consisted of 50 cycles with denaturation at 94°C for 60 seconds, annealing at 58°C for 60 seconds and elongation at 72°C for 60 seconds. In a third step, an additional elongation at 72°C was performed for 7 minutes. PCR was performed using a PTC0225 DNA Engine Tetrad (Bio-Rad).

Table 9: List of reagents and reaction condition for a Pyrosequencing PCR reaction

Reagent	Condition	Volume
AmpliTaq Gold PCR reaction buffer	10X	3 µl
AmpliTaq $MgCl_2$	25 µM	3 µl
dNTP	25 µM	0.5 µl
Primer F	10 µM	2.5 µl
Primer R	10 µM	2.5 µl
Taq polymerase	5 U/µl	0.33 µl
DNA	25 ng/µl	1 µl
dH_2O		33.17 µl
Final total volume		50 µl

Next, the PCR products were purified from substrates and fragments and are processed to single stranded sequencing products. The PyroMark Q96 Vacuum Prep Workstation (Biotage AB) provides a platform for a quick and reliable processing workflow. The workstation consists of a vacuum preparation tool with replaceable filter probes and a worktable. PCR products were stored short-term in a 96 well micro titre plate. A mix of 3 µl of Strepdavidin Sepharose™ High Performance beads (Amersham), 37 µl of binding buffer (Biotage AB), 20 µl of PCR product, and 20 µl of purified water was transferred to a 96 well micro titre plate. The mix was incubated by agitation for five to ten minutes at room temperature at a Variomag Monoshaker (Camlab). This process leads to immobilisation of the biotinylated PCR products on the sepharose beads. The beads were captured by the vacuum preparation tool for five to ten seconds and flushed with a 70 % ethanol solution for five seconds. Next, the beads were flushed with 0.2M sodium hydroxide for five seconds and then with 10mM Tris acetate buffer for five

seconds. The vacuum was turned off for ten seconds and the beads were washed with purified water for 15 seconds. On a separate PSQ 96 well plate low (Biotage AB), 5 µl of sequencing primer designed with PSQ Assay Design software version 1.0.6 (Biotage AB) and manufactured by biomers.net (biomers.net GmbH) was added to 40 µl of annealing buffer (20mM TrisAcetate, 2mM $MgAc_2$ pH 7.6; Biotage AB). The washed beads were shaken out into the mix on the PSQ 96 well plate low. Afterwards, the solution was incubated at 80°C and allowed to cool down for two minutes.

The 96 well plate was transferred to the PSQ 96 MA pyrosequencing facility (Biotage AB) and subsequent processes were automated and run by the pyrosequencing software. A sequencing primer was hybridised to a single-stranded DNA fragment and individual deoxynucleotide triphosphates (dNTP) were added to the reaction in a pre-determined dispensation order. DNA polymerase catalyses the incorporation of the dNTPs into the DNA strand if complementary to the template (figure 5).

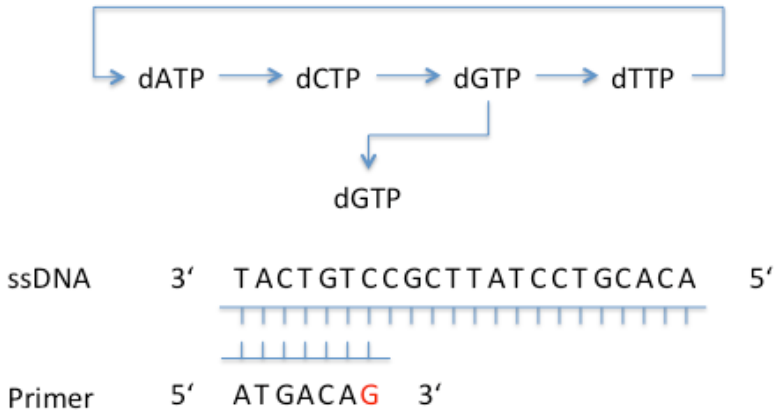


Figure 5: Incorporation of nucleotides in the Pyrosequencing reaction

Nucleotides are presented to the reaction in a strict order to allow the extension of the primer strand in presence of polymerase.

Abbreviation: single strand DNA (ssDNA)

Each incorporation event was accompanied by the release of pyrophosphate (PPi) in a quantity that is equimolar to the amount of the incorporated nucleotide (figure 6).

Figure 6: Sequencing reaction accompanied by the release of pyrophosphate
Abbreviation: pyrophosphate (PPi).

Adenosine triphosphate ATP sulphurylase quantitatively converts PPi to ATP in the presence of adenosine 5' phosphosulphate (figure 7).

Figure 7: ATP sulphurylase converts PPi to ATP in the presence of APS
Abbreviations: adenosine tri phosphate (ATP), adenosine 5' phosphosulphate (APS), pyrophosphate (PPi).

In an oxidative reaction, luciferin reacts with the released ATP to produce oxyluciferin and a flash of light. A camera detects the emitted light. The intensity of the light is related to the number of incorporated nucleotides (figure 8), i.e. two type A nucleotides will produce twice the light intensity of one type A nucleotide.

Figure 8: Luciferase-mediated conversion of luciferin to oxyluciferin in the presence of ATP resulting in light emission

Unincorporated nucleotides or free ATP might lead to a false positive result. Therefore, both products need to be eliminated from the reaction. Apyrase, a nucleotide-degrading enzyme, continuously degrades unincorporated nucleotides and ATP (figure 9).

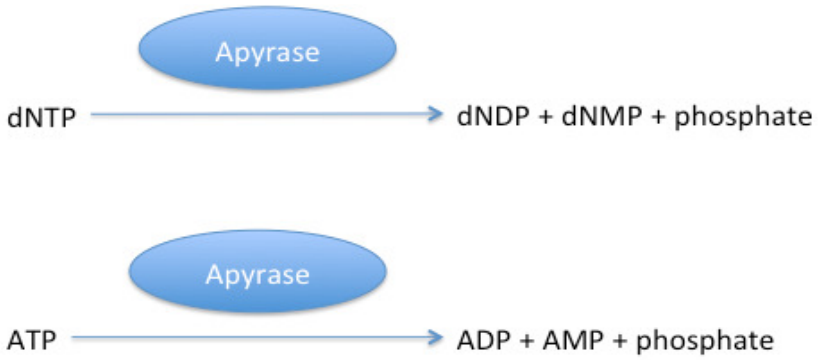


Figure 9: Degradation of unincorporated nucleotides (above) and ATP (below) by apyrase
Abbreviations: adenosine-mono-phosphate (AMP), adenosine-di-phosphate (ADP), adenosine-tri-phosphate (ATP), nucleotide-mono-phosphate (NMP), nucleotide-di-phosphate (NDP), nucleotide-tri-phosphate (NTP).

Pyrosequencing reactions were performed according to the manufacturer's instructions using the PSQ 96 SNP Reagent Kit (Biotage AB) that contained the enzyme and substrate mixture and nucleotides. Assays were performed using the nucleotide dispensation orders shown in table 10. The sample genotype was determined using the Allele Frequency Quantification (AQ) function in the PSQ96MA SNP software version 2.1 (Biotage AB). For the Pyrosequencing assays, the sequence to analyse is immediately 3' to the sequencing primer-binding site on

the biotinylated strand. The dispensation order of the nucleotides includes several reference peaks where no signal should be observed. The dispensation orders used were those determined by the software in the Simplex SNP entry function.

Table 10: Dispensation orders and sequence analysed for the 17 SNPs

The sequence to analyse is located immediately 3' to the sequencing primer binding site on the biotinylated strand. The mutation is highlighted either with (brackets) or in **bold font** in the dispensation order, respectively. The dispensation order of the nucleotides includes several reference peaks, where no signal should be observed, these are shown in *italics*.

Gene	SNP	Genotype	Sequence to analyse	Dispensation order
STAT1	rs1467199	C/G	GTC (C/G) TTCTTGTCACAAT *	*C*GT**CG***A*TCT
STAT3	rs6503695	C/T	(A/G) ACTATTTTCTT	*CG***AG**CTATC
STAT3	rs2293152	C/G	CC (C/G) GCAGCCAGAGGCCCTTT	*T***CG***T*CAGCA
STAT5A	rs2293154	A/G	A (G/A) CCACGCCCCATG	*T***AG***T*CACGC
STAT5B	rs16967611	A/G	(C/T) TGCCTGTGCCCTAGCATGTTGCAGA	*G***CT***A*GCTGT
STAT5B	rs6503691	C/T	TG (A/G) CGGATTGTCTTCCCA	*C***TGA***T*CGAT
STAT5B	rs9900213	G/T	TG (G/T) TCCAGTGTCACCT	*C*T**GT***G*CAGT
STAT5B	rs17500235	G/T	(A/C) CAAGACTTC	*G***AC***T*AGACT
STAT5B	rs17591972	A/G	A (A/G) CTGAACTGAA	*T***AG***A*CTGAC
TYK2	rs2304256	A/C	ATG (A/C) TATC	*G*ATG**AC***G*TA
TYK2	rs12720356	G/T	CG (A/C) TGCCTT	*T*CG**AC***A*TGC
JAK1	rs2991269	C/T	(C/T) GGGCCACACTCACTGTC	*A***CT***A*GCACA
JAK1	rs310227	A/G	(C/T) AGGGGTCTGGCTACTGG	*G***CT***C*AGTCT
JAK1	rs310229	A/G	(C/T) GAATAGCAGGTGCAGGGTGCCCAGA	*A***CT***C*GATAG
IFNAR1	rs2850015	C/T	C (A/G) CACGCGCTGCCCCTCTTAGCTCTCA	*T*C**AG***T*CACG
IFNAR1	rs2257167	C/G	AA (C/G) TAAGCTATATGTAAAGCTTAAACCA	*TA***CG***C*TAGC
IFNAR2	rs7279064	G/T	(A/C) GATGAAGGCATTCTGGCTCAAAAGC	*T***AC***T*GATGA

3 Prognostic significance of der(9) deletions in chronic myeloid leukaemia (CML)

3.1 Introduction

In 1983, Bartram *et al.* described the fusion of *ABL1* sequences on the long arm of chromosome 9 to an unknown gene on the long arm of chromosome 22 (Bartram, *et al.* 1983), later named the 'breakpoint cluster region' or *BCR* gene (Groffen, *et al.* 1984, Heisterkamp, *et al.* 1985). The translocation is reciprocal and it forms a second fusion gene on the derivative chromosome 9 [der(9)]. However, the reciprocal fusion gene product, the ABL1-BCR protein, is only expressed in 70 % of patients (Melo, *et al.* 1993).

De Klein *et al.* observed by Southern blot analysis and cloning that the fusion of the *ABL1* gene to the *BCR* gene on der(9) in a single case was accompanied by an 88 nucleotide deletion within *BCR* (de Klein, *et al.* 1986). Larger *BCR* deletions of up to 10 kb on the der(9) were described by Popenoe *et al.* (Popenoe, *et al.* 1986). *BCR* deletions were also found in the CML cell line K562 (Grosveld, *et al.* 1986). Shtalrid *et al.* analysed the *BCR* breakpoint and confirmed the occurrence of deletions (Shtalrid, *et al.* 1988). Grand *et al.* found that the deletions downstream of *BCR* were larger than had been previously suspected and encompassed the hSNF5/INI1 gene (Grand, *et al.* 1999). Herens *et al.* described deletions centromeric of *ABL1* up to the *ASS* gene in 10 % of patients and suggested that deletions were associated with a shorter duration of the chronic phase (Herens, *et al.* 2000). Dewald *et al.* described deletions of der(9) on both sides of the breakpoint in 19 % of CML patients (Dewald, *et al.* 1999). Sinclair *et al.* also found not only deletions of *BCR* and further telomeric located chromosome 22 sequences on the der(9), but also large deletions of *ABL1* sequences upstream of the breakpoint (Sinclair, *et al.* 2000). Furthermore, deletions were also observed on additional partner chromosomes involved in variant translocations. Deletions were of variable length up to several mega bases long and spanned the *ABL1-BCR* breakpoint. Fifteen of 56 patients (27 %) studied had deletions by metaphase fluorescence *in situ* hybridisation (FISH) and these cases had a relatively poor prognosis after

treatment with IFN-*alpha* and or chemotherapy. Deletions were confirmed by microsatellite PCR (Sinclair, *et al.* 2000). Kolomietz *et al.* associated the appearance of submicroscopic deletions in 9 % of their CML patients with an impaired outcome after allogeneic stem cell transplantation (Kolomietz, *et al.* 2001). However, Lundan *et al.* showed a prevalence of 15 % of deletions in their study on 110 transplanted patients but failed to confirm a worse prognosis (Lundan, *et al.* 2005). Kolomietz *et al.* also described deletions at the reciprocal breakpoint in other haematologic malignancies with chromosomal rearrangements, including *MLL* gene translocations and inv(16) (Kolomietz, *et al.* 2001). Bacher *et al.* later reported a similar prevalence (2–9 %) of submicroscopic deletions in haematologic malignancies with reciprocal translocations to *BCR-ABL1* (CML) in patients with *CBFB-MYH11* (acute lymphoblastic leukaemia; ALL), or *PML-RARA* (acute promyelocytic leukaemia; APL) (Bacher, *et al.* 2005). Specchia *et al.* investigated the frequency of der(9) deletions in *BCR-ABL1*-positive ALL and found four of 45 patients (9 %) deleted, similar to CML (Specchia, *et al.* 2003). In childhood ALL, Robinson *et al.* reported about 3 of 27 cases (11 %) with der(9) deletions (Robinson, *et al.* 2005) but other groups have failed to confirm this (Lee, *et al.* 2003). Cohen *et al.* found deletions proximal to *ABL1* in 14 of 94 patients (15 %). They were also associated with a poor prognosis in IFN treated patients. In a gene expression assay based on DNA micro array technology, several genes with different expression pattern were identified that are involved in cell adhesion and migration (Cohen, *et al.* 2001). Another study by Gonzalez *et al.* on 57 patients revealed deletions in 33 % of patients. However, no difference in the clinical course of both patient groups was seen (Gonzalez, *et al.* 2001). No evidence for an adverse prognostic association has also been observed in several other small studies (Dewald, *et al.* 1999, Shtalrid, *et al.* 1988).

In a large study of 253 patients with CML, Huntly *et al.* identified deletions in 39 patients (15 %) and associated deletion status with a poor prognosis under IFN therapy. Importantly, the poor prognosis was due primarily to an increased rate of progression to advanced phase disease. It was suggested that deletions were formed at the time of translocation and might be therefore closely linked to the chromosomal rearrangement itself (Huntly, *et al.* 2001). Other studies confirmed

the appearance of deletions as an early event (Grand, *et al.* 1999, Sinclair, *et al.* 2000). They were more prevalent in patients with variant translocations (Grand, *et al.* 1999, Huntly, *et al.* 2001). Huntly's group also compared the prognostic strength of deletion status with both Sokal and Hasford scores and defined deletion status as superior to them. However, for both scores, patients were grouped as high risk and non-high risk which may have weakened the statistical significance of both scoring systems and supported the strength of the deletion status (Huntly, *et al.* 2001). Huntly's observations have never been confirmed in a large study.

The advent of imatinib as the standard therapy in CML led to a study analysing the impact of this new compound in the context of der(9) deletion status. Huntly *et al.* examined 397 patients treated with imatinib and whilst no difference in overall survival was seen, those with deletions progressed earlier than all others (Huntly, *et al.* 2003b). In another study, Quintas-Cardama *et al.* found 33 of 352 patients (9 %) treated with imatinib harbouring deletions of der(9). Their prognosis regarding progression and survival did not differ from all others without deletion (Quintas-Cardama, *et al.* 2005).

A prominent feature of der(9) deletions is their variability in length. While early studies reported small deletions of few kilo bases (de Klein, *et al.* 1986, Popenoe, *et al.* 1986), it is now known that they may extend up to several mega bases (Huntly, *et al.* 2001, Sinclair, *et al.* 2000). Storlazzi *et al.* mapped the der(9) and found deleted regions that embrace minimal loci containing the *ASS* gene on chromosome 9 and *IGLL1* on chromosome 22 sequences. Deletions occurred in ten out of 71 patients (14 %) (Storlazzi, *et al.* 2002a). Reid *et al.* investigated the presence of deletions in the patient derived Ph-positive leukaemic cell line MC3 and several patient samples by the DNA based multiplex amplifiable probe hybridisation (MAPH) assay (Okabe, *et al.* 1995, Reid, *et al.* 2003b). Fourouclas *et al.* correlated the size of deletions with the patient's outcome in 69 cases. Large deletions extending 1.4 mega bases were associated with a worse prognosis than others. The hypothesis of the deletion of a tumour suppressor was stressed but not confirmed (Fourouclas, *et al.* 2006). Douet-Guilbert *et al.* mapped deletions upstream of *ABL1*

of 2-5 mega bases in length. Deletions telemetric of the breakpoint were shorter, up to 500 kilo bases (Douet-Guilbert*, et al.* 2006).

The molecular basis for the purported relatively poor prognosis associated with deletions of the der(9) has so far not been identified. Several hypotheses, discussed below, have been proposed: (i) haploinsufficiency of a tumour suppressor gene, (ii) loss of *ABL1-BCR* expression, (iii) genomic instability, and (iv) enhanced *BCR-ABL1* expression.

(i) There are about 50 genes flanking the *ABL1-BCR* breakpoint with known function that are lost in many cases with deletions. Grand *et al.* found the *hSNF5/INI1* (*SMARCB1*) gene downstream of *BCR* was deleted in nine of 25 patients (36 %). *SMARCB1* mutations are observed in solid tumours and, therefore, the possibility of the deletion of a candidate tumour suppressor gene was examined but no mutations in the remaining allele were found (Grand*, et al.* 1999). In a study by Kolomietz *et al.*, deletions were observed by real-time PCR and FISH in 48 of 250 patients (19 %). The putative tumour suppressor gene *PRDM12* was proposed as a candidate since it lay within the deleted region in all patients (Kolomietz*, et al.* 2003). Specchia *et al.* found the genes *PTGES*, *SMARCB1* and *GSTT1* within the deleted region in almost all cases with der(9) deletion (Specchia*, et al.* 2004). However, as yet there is no further evidence that implicates the involvement of these genes in the impaired outcome of patients with deletions.

(ii) Der(9) deletions that are detectable by FISH using BAC probes (by far the most commonly used approach to detect deletions) must usually involve the loss of some or all *ABL1* and *BCR* exons and thus would be expected to result in loss of *ABL1-BCR* expression. However loss of *ABL1-BCR* expression may also be caused by smaller deletions or other abnormalities that would not be detected by FISH and thus the frequency of loss of *ABL1-BCR* expression is expected to be higher than the frequency of deletions. Nevertheless, it may be that loss of *ABL1-BCR* expression may be the important consequence of der(9) deletions.

Before the possible significance of der(9) deletions was widely known, Melo *et al.* reported a 70 % prevalence of *ABL1-BCR* expression in CML patients (Melo, *et al.* 1993) but no difference was observed in the clinical course of either *ABL1-BCR* positive or negative cases. Subsequently, de la Fuente *et al.* confirmed that *ABL1-BCR* is expressed in about two thirds of CML patients but only 46 % of cases that lacked *ABL1-BCR* had large der(9) deletions (de la Fuente, *et al.* 2001). Furthermore, an impaired response to IFN therapy was not observed in *ABL1-BCR* negative patients (Melo, *et al.* 1996). Huntly *et al.* confirmed the lack of deletion in half of the patients with absent *ABL1-BCR* expression (Huntly, *et al.* 2002). Loncarevic *et al.* characterised the lack of *ABL1-BCR* expression as mainly associated with (i) deletion of der(9) sequences and or (ii) variant chromosomal translocations (Loncarevic, *et al.* 2002). Albano *et al.* investigated cases with variant translocations and found deletions on the third partner chromosome (e.g., chromosomes 4, 6 and 13) (Albano, *et al.* 2003). Reid *et al.* demonstrated a higher frequency of deletions in patients with variant Philadelphia chromosome (40 %) compared to those with a classical translocation (14 %). They found, in contrast to previous studies, that variant translocations were associated with a significantly worse prognosis. The rise of a variant translocation was proposed to be a multistep process that may lead to loss of genetic material each time a breakage occurs (Anelli, *et al.* 2004, Reid, *et al.* 2003a).

(iii) As a third hypothesis, it was proposed that deleted sequences are of no direct relevance but simply reflect an underlying genomic instability, and it is this instability that results in an increased propensity for disease progression and hence, inferior survival. Although this is an attractive proposition, there are no data to support it and no increased frequency of additional cytogenetic changes was observed by several groups (Huntly, *et al.* 2002, Silly, *et al.* 1994). Nevertheless, it is possible that there is a generalised increase in mutations rates that has not yet been detected, however it is unclear how this could lead to relatively large der(9) deletions.

(iv) Loss of genomic material on the der(9) may be accompanied by micro-deletions on the Philadelphia chromosome. Variations of the *BCR-ABL1* expression

may lead to a different CML phenotype. Therefore, Huntly *et al.* performed a real-time quantitative RT-PCR but found no differences in *BCR-ABL1* transcript levels in patients with or without deletions of der(9) (Huntly*, et al.* 2003a, Huntly*, et al.* 2002).

To summarise, der(9) deletions are seen in 9 to 15 % of CML cases and have been associated with an adverse prognosis in several studies. However other studies have not detected any prognostic association and, importantly, no study so far has addressed the issue systematically in a randomised clinical trial. Ideally such an analysis would be performed prospectively and indeed this is taking place for some ongoing studies with imatinib. However these studies will take many years to mature and they cannot address the possibility that the phenomenon might be associated with specific treatment modalities. This study aimed to reassess the significance of der(9) deletions using DNA stored as part of CML clinical trials in Germany for which there was a relatively long follow up. Since the largest studies described above suggested an adverse prognosis for IFN-treated patients, it was decided to focus solely on patients treated with IFN-based regimens.

3.2 Patients

From 1983 until 2001, newly diagnosed patients were enrolled in three CML trials of the German CML study group. Of a total of 1435 patients that were entered into these three trials, 843 were randomised to receive either IFN as a single treatment option or in combination with HU or Ara-C. The randomisation arm for IFN in the CML I trial was opened following a protocol amendment in 1986. Therefore, only 133 out of 513 patients were randomised to this arm (Hehlmann*, et al.* 1993). From 1991 until 1994 a total of 340 patients were recruited to CML trial II and 226 of them were randomised to receive an IFN/HU combination treatment (Hehlmann*, et al.* 2003). In the CML III trial, a total of 621 patients were randomised of whom 484 patients received IFN based therapy (Hehlmann*, et al.* 2007). DNA samples prior to treatment were not collected systematically for these trials, but were stored at different time periods for a number of ongoing research

projects. In total pre-treatment DNA was available from 339 patients enrolled in either CML I (n=13), CML II (n=79) or CML III trial (n=247). A slight male excess (58 %) was observed as in other large studies before (Guilhot*, et al.* 1997, The Italian Cooperative Study Group on Chronic Myeloid Leukaemia 1994). The median age of the study population was 49 years (range, 10-83) and the follow up time was a median of 7.1 years (range, 0.3-16). The median treatment duration with IFN was 17 months (range, 1-115). 286 patients were evaluable for cytogenetic analysis. A major cytogenetic response (i.e., ≤35 % Ph-positive metaphases) was achieved by 76 patients (22 %), while 35 patients (10 %) achieved a complete response. At the time of analysis, 158 patients (47 %) were still alive. A summary of patient characteristics is given in table 11.

3.3 Statistical analysis

Survival probabilities were estimated by the Kaplan-Meier method. For comparison of survival probabilities between different groups, the log rank test was applied. To investigate the distribution of baseline values between groups, univariate tests were performed by using the Mann-Whitney, Fisher's exact, or $\chi 2$ tests, as appropriate. Results were initially analysed by myself using GraphPad Prism version 5.00 for Windows (GraphPad Software; San Diego, CA, U.S.A.). Subsequently the analysis was refined and examined in more detail by Dr. Markus Pfirrmann (IBE, LMU; Munich, Germany).

3.4 Development of an MLPA assay to detect der(9) deletions

In most studies, der(9) deletions were detected by fluorescence *in-situ* hybridisation (FISH) (Albano*, et al.* 2003, Cohen*, et al.* 2001, Grand*, et al.* 1999, Huntly*, et al.* 2002, Huntly*, et al.* 2001, Kim*, et al.* 2005, Kolomietz*, et al.* 2001, Kolomietz*, et al.* 2003, Lee*, et al.* 2003, Sinclair*, et al.* 2000, Specchia*, et al.* 2004,

Storlazzi*, et al.* 2002b). For this study, however, fixed cells were generally not available and instead I sought to develop a methodology based on genomic DNA.

Table 11: Summary of patients characteristic screened for der(9) deletions in CML

	Total	no 9q+ del	9q+ del	large 9q+ del	small 9q+ del
Patients analysed; n=	339	280	59	21	38
Age in years (range)	49 (10-83)	50 (10-83)	49 (11-72)	56 (21-71)	45 (11-65)
Gender (% male)	58	59	54	52	55
Clinical characteristics					
Spleen size (cm below costal margin)	2 (0-30)	2 (0-30)	3 (0-18)	0 (0-9)	5 (0-18)
WBC count (range; x10^9/l)	107 (6-650)	106 (6-650)	117 (20-494)	90 (22-494)	121 (20-396)
PB blasts (range; %)	1 (0-22)	1 (0-22)	1 (0-10)	0 (0-6)	2 (0-10)
PB eosinophils (range; %)	2 (0-51)	2 (0-51)	2 (0-9)	2 (0-6)	3 (0-9)
PB basophils (range; %)	3 (0-40)	3 (0-40)	3 (0-14)	2 (0-7)	4 (0-14)
Platelet count (range; x10^9/l)	397 (49-3490)	398 (49-3490)	397 (124-1986)	397 (124-1109)	376 (184-1986)
Hemoglobin (range; g/dl)	12 (6-16)	12 (6-16)	12 (7-16)	13 (8-15)	12 (7-16)
Ph chromosome positive)	97	97	94	100	90
additional aberrations (%)	1	1	0	0	0
BCR-ABL transcript					
evaluable	173	150	27	7	20
b2a2 (%)	61 (35)	52 (35)	10 (37)	4 (57)	6 (30)
b3a2 (%)	80 (46)	68 (45)	14 (52)	3 (43)	11 (55)
b2a2/b3a2 (%)	31 (18)	29 (19)	3(11)		3 (15)
others	1 (1)	1 (1)			
Median observation time in years (range)	7.1 (0.3-16)	6.9 (0.3-16)	8.0 (0.5-16)	5.7 (0.5-9.8)	8.4 (0.5-16)
Alive (%)	158 (47)	125 (45)	33 (56)	3 (14)	30 (79)
Median time on IFN-*alpha* (months)	17 (1-115)	18 (1-115)	13 (1-106)	14 (4-81)	13 (1-106)
Cytogenetic response					
evaluable	286	239	47	19	28
CR (%)	76 (22)	66 (24)	10 (21)	5 (26)	5 (18)
CCR (%)	35 (10)	30 (13)	5 (11)	2 (11)	3 (11)
Subsequent therapy					
Imatinib (%)	62 (18)	49 (18)	13 (22)	4 (19)	9 (24)
thereof in 1st chronic phase	49 (14)	39 (14)	10 (17)	2 (10)	8 (21)
Allogeneic stem cell transplantation (%)	129 (38)	109 (39)	20 (34)	6 (29)	14 (37)
thereof in 1st chronic phase	105 (31)	87 (31)	18 (31)	6 (29)	12 (32)
Sokal risk groups					
evaluable	338	279	59	21	38
low (%)	133 (39)	108 (39)	25 (42)	10 (48)	15 (39)
intermediate (%)	112 (33)	91 (33)	21 (36)	9 (43)	12 (32)
high (%)	93 (28)	80 (28)	13 (22)	2 (9)	11 (29)
Hasford risk groups					
evaluable	338	279	59	21	38
low (%)	132 (39)	106 (38)	26 (44)	11 (52)	15 (39)
intermediate (%)	166 (49)	142 (51)	24 (41)	10 (48)	14 (37)
high (%)		31 (11)		0	

Detection of deletions using genomic DNA has been performed in previous studies using real time PCR (Kolomietz*, et al.* 2003) or multiplex amplifiable probe

hybridisation (Reid, *et al.* 2003b) but both these techniques are considered to be cumbersome in this context and difficult to perform reliably. However, MLPA has rapidly become an established technique for the reliable detection of copy number DNA copy number changes (Bunyan, *et al.* 2004, Schouten, *et al.* 2002, Vorstman, *et al.* 2006) and therefore I focused on the development of an MLPA der(9) deletion test. I designed six probes to span the region on der(9) that is known to be deleted in most cases using commercially available FISH probes (figure 10).

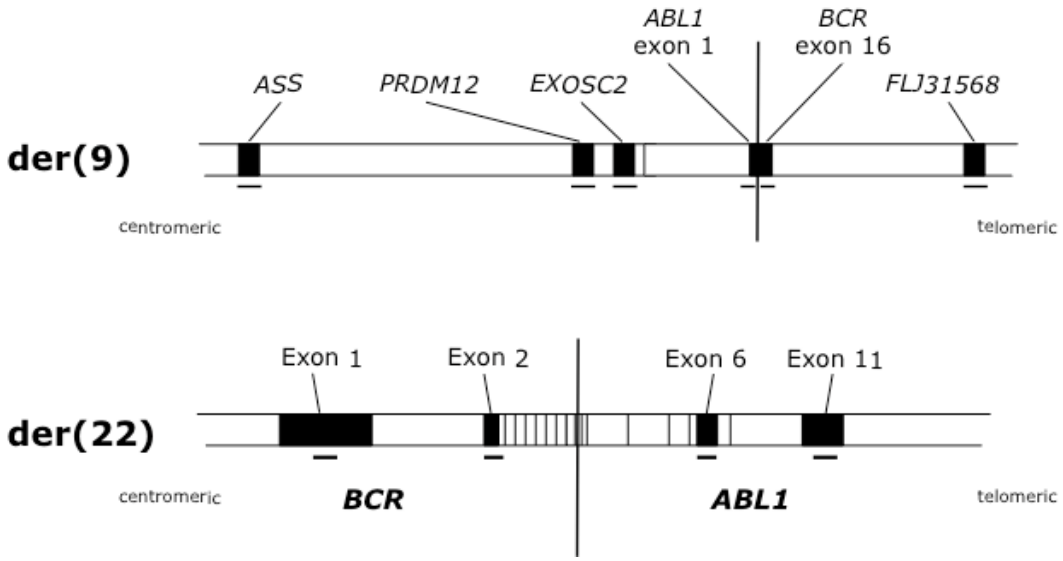


Figure 10: Map of MLPA probes for der(9) deletion detection

MLPA probes were designed to detect deletions on the derivative chromosome 9 (upper schematic) compared to the derivative chromosome 22 (lower schematic) and chromosome 15 as a control (not shown). The location of the probe pairs is indicated by a black underline.

Depending on the precise positions of the breakpoints, these probes are expected to span a region of 360-500 kb. As controls I designed four probes within *BCR* and *ABL1* that are retained on the der(22) and, in addition, I designed two additional control probes from chromosome 15 (table 12). All samples were analysed initially with the ten chromosome 9 and 22 probes (*BCR* exon 1, *ABL1* exon 1b, *BCR* exon 16, *FLJ31568* exon 2, *PRDM12* exon 12, *ASS* exon 4, *BCR* exon 2, *ABL1* exon 6, *ABL1* exon 11, *and EXOSC2* exon 7).

Table 12: MLPA probe sequences

Chr	Gene	Exon	Amplicon	5' MLPA oligonucleotide probe sequences
der(9)	*ABL1*	1b	95bp	*GGG TTC CCT AAG GGT TGG* AGT **GCT GCA TTT TAT CAA AGG AGC AGG**
der(9)	*BCR*	16	100bp	*GGG TTC CCT AAG GGT TGG* ACG TTG **CAA GAC GAA GAT CCC CAA GGA GGA**
der(9)	*FLJ31568*	2	105bp	*GG GTT CCC TAA GGG TTG GAC* GTG ATG **CAA CCA CAA TCG CAG CCT GTC CTA**
der(9)	*PRDM12*	2	110bp	*GG GTT CCC TAA GGG TTG GAC* CAT GCA TGC CAC **GTG GAC ATC TGC AAG AAC AACAAC**
der(9)	*ASS*	4	115bp	*G GGT TCC CTA AGG GTT GGA* CCG CCG GAT CGG **TGC CTA CTT CTT CCT TCT GGG CTC**
der(9)	*EXOSC2*	7	135bp	*GGG TTC CCT AAG GGT TGG* ACG GCA TAT GCC GCT TTA AGA GAT ATA **GTT TTG GTC CAG GTT TCC CCC TCC**
der(22)	*BCR*	1	90bp	*GGG TTC CCT AAG GGT TGG* **ATA CCA GAG CAT CTA CGT CGG GGG**
der(22)	*BCR*	2	120bp	*G GGT TCC CTA AGG GTT GGA* GCA TGC GAC AGC TGC **ACC AAG ATG GGC TGC CCT ACA TTG**
der(22)	*ABL1*	6	125bp	*GGG TTC CCT AAG GGT TGG* AGT CGA TCT TCC AGC TGC **CCC CCG TTC TAT ATC ATC ACT GAG**
der(22)	*ABL1*	11	130bp	*G GGT TCC CTA AGG GTT GGA* CGT AGG CTA GGG ATG TAC GCG **CTG AAT GAA GAT GAG CGC CTT CTC**
15	*UBR1*	2	105bp	*GGG TTC CCT AAG GGT TGG* ACG GCA TAC **CAC AGT GGA GCA TTT CAG CTT TGT**
15	*NDNL2*	1	110bp	*GG GTT CCC TAA GGG TTG GAC* GGC ATA TGC **CAG TAC GTC TTC GGG TAT AAG CTG**

Chr	Gene	exon	Amplicon	3' MLPA oligonucleotide probe sequences
der(9)	*ABL1*	1b	95bp	**GAA GAA GGA ATC ATC GAG GCA TGG** GCG *TCT AGA TTG GAT CTT GCT GGC AC*
der(9)	*BCR*	16	100bp	**CGG CGA GAG CAC GGA CAG ACT CAT** GAG CC*T CTA GAT TGG ATC TTG CTG GCA C*
der(9)	*FLJ31568*	2	105bp	**CAA TAA TGT GCT CAA CCC TGG CTC** GAT CGC TC*T CTA GAT TGG ATC TTG CTG GCA C*
der(9)	*PRDM12*	2	110bp	**CTC ATG TGG GAG GTA CGC GCG CCG** TAT CTA G*TC TAG ATT GGA TCT TGC TGG CAC*
der(9)	*ASS*	4	115bp	**CTC TTC CCG TAG GTG TTC ATT GAG** CGT AGT CAT GAC C*TC TAG ATT GGA TCT TGC TGG CAC*
der(9)	*EXOSC2*	7	135bp	**CTG GTG AAA CGG CAG AAG ACC CAC** GAT GGA ACT ATA CAT ACG C*TC TAG ATT GGA TCT TGC TGG CAC*
der(22)	*BCR*	1	90bp	**CAT GAT GGA AGG GGA GGG CAA GGG** C*TC TAG ATT GGA TCT TGC TGG CAC*
der(22)	*BCR*	2	120bp	**ATG ACT CGC CCT CCT CAT CGC CCT** CGA TCG CCA GCA TGC *TCT AGA TTG GAT CTT GCT GGC AC*
der(22)	*ABL1*	6	125bp	**TTC ATG ACC TAC GGG AAC CTC CTG** CGT ACT ACA GCT GAG CTC *TC TAG ATT GGA TCT TGC TGG CAC*
der(22)	*ABL1*	11	130bp	**CCC AAA GAC AAA AAG ACC AAC TTG** GAT GCG ATC GCA TAC AGC C*TC TAG ATT GGA TCT TGC TGG CAC*
15	*UBR1*	2	105bp	**GGG AGG GTT TTC AAA AGT GGA GAG** GGT GAT G*TC TAG ATT GGA TCT TGC TGG CAC*
15	*NDNL2*	1	110bp	**GTG GAA CTT GAA CCC AAG AGC AAC** GGT GAT CGT C*TC TAG ATT GGA TCT TGC TGG CAC*

Tag sequences are in *italics*, stuffer sequences in plain type and sequences corresponding to the region of interest are in **bold**.

Abbreviations: chromosome (Chr), multiplex ligation-dependent probe amplification (MLPA)

In MLPA, pairs of adjacent probes are designed to specific regions and hybridised to denatured genomic DNA. Only those probes that hybridise correctly can be ligated together and amplified. Each probe pair is designed to give a distinct size product after PCR amplification with a common set of primers and thus as long as conditions are established to ensure that hybridisation goes to completion, the relative final peak height reflects the amount of the genomic DNA target (Schouten*, et al.* 2002). Since the efficiency with which each fragment is amplified is not identical, the actual peak height for each target may vary but once conditions are established the relative peak height is constant.

Peak heights from each patient were exported to an Excel spreadsheet (Microsoft Office Excel 2003 SP2 software), which was designed to assess the ratios of each test peak relative to all other peaks for that individual. Ratios of test peaks to control peaks and control peaks to other control peaks in each patient sample were compared to the same ratios obtained for two normal individuals who were included in each run. For normal sequences a dosage quotient of 1.0 is expected; if a deletion or duplication is present the dosage quotient should be 0.5 and 1.5, respectively. Dr. David Bunyan at the Wessex Regional Genetics Laboratory previously established a dosage quotient of <0.7 as a reliable cut off for scoring deletions and thus the spreadsheet highlights any value below 0.7 in red and any reciprocal values above 1.3 in green (Bunyan*, et al.* 2004).

Table 13 illustrates the data calculation table in case of no deletion (A), *ABL1-BCR* deletion (B), *ABL* deletion only (C), and *BCR* deletion only (D). Since this probe set would not distinguish between a der(9) deletion and a gain of the Ph chromosome, samples were retested with a second set that included the two chromosome 15 control probes (*BCR* exon 1, *SMARCB1* exon 2, *UBR1* exon 2, *NDNL2*, *BCR* exon 2, *ABL1* exon 6, *ABL1* exon 11). The two additional control probes were selected from random genes from a chromosome that is rarely abnormal in CML.

Table 13: Calculation tables for multiplex ligation-dependent probe amplification (MLPA) analysis

Examples of MLPA results (A) in a case with no deletion, (B) with a der(9) deletion encompassing the *ABL1-BCR* fusion point, (C) with a der(9) deletion on the *ABL1* side only, and (D) with a der(9) deletion on the BCR side only.

A

PEAK AREA		
Exons	Control	Sample
BCR 1	87882	37664
ABL 1	93026	39039
BCR 16	101140	39186
FLJ31568-2	121629	43097
PRDM12-12	95414	42244
ASS 4	105032	37943
BCR 2	49122	21013
ABL 6	82122	32650
ABL 11	76544	29742
EXOSC2-7	49316	21493

Exons									
BCR 1	ABL 1	BCR 16	FLJ31568-2	PRDM12-2	ASS 4	BCR 2	ABL 6	ABL 11	EXOSC2-7
1	1,021	1,106161463	1,209529891	0,967996	1,186359955	1,001877191	1,077960485	1,102979634	0,9833708
0,979191936	1	1,083144384	1,184361915	0,9478539	1,1616741	0,981030066	1,055530214	1,080028763	0,9629088
0,904027154	0,923	1	1,093447866	0,8750947	1,072501614	0,905724186	0,97450555	0,99712354	0,8889939
0,826767497	0,844	0,914538344	1	0,8003077	0,98084385	0,828319497	0,891222692	0,911907711	0,813019
1,033062083	1,055	1,142733466	1,249519469	1	1,225583486	1,035001338	1,113600104	1,139446439	1,0158831
0,842914493	0,861	0,932399529	1,019530275	0,8159379	1	0,844496805	0,908628516	0,92971752	0,8288975
0,998126326	1,019	1,104088878	1,207263627	0,9661823	1,184137103	1	1,075940739	1,10091301	0,9815283
0,927677789	0,947	1,026161421	1,122054015	0,8979884	1,10055978	0,929419218	1	1,023209709	0,9122513
0,906635054	0,926	1,002884758	1,096602198	0,8776191	1,075595521	0,908336981	0,977316763	1	0,8915585
1,016910392	1,039	1,124867087	1,229983516	0,9843652	1,206421767	1,018819327	1,09618922	1,121631452	1

B

PEAK AREA		
Exons	Control	Sample
BCR 1	87882	119869
ABL 1	93026	67286
BCR 16	101140	66795
FLJ31568-2	121629	78012
PRDM12-12	95414	73301
ASS 4	105032	62217
BCR 2	49122	64446
ABL 6	82122	94496
ABL 11	76544	89194
EXOSC2-7	49316	34610

Exons									
BCR 1	ABL 1	BCR 16	FLJ31568-2	PRDM12-2	ASS 4	BCR 2	ABL 6	ABL 11	EXOSC2-7
1	1,886	2,065313317	2,126584648	1,7754529	2,302605403	1,03964968	1,185367574	1,170529769	1,9435387
0,530289971	1	1,095214939	1,127706511	0,9415049	1,221048552	0,551315799	0,628588536	0,620720197	1,0306391
0,484188037	0,913	1	1,029666845	0,8596531	1,114893989	0,503385937	0,573940798	0,566756511	0,9410382
0,470237571	0,887	0,971187918	1	0,8348847	1,082771572	0,48888234	0,557404369	0,550427076	0,9139249
0,563236571	1,062	1,163259992	1,197770246	1	1,296911572	0,585568721	0,667642368	0,659285174	1,0946721
0,434290651	0,819	0,896946266	0,923555832	0,7710626	1	0,451510137	0,514794056	0,508350136	0,8440607
0,961862461	1,814	1,986547351	2,045481944	1,7077415	2,2147897	1	1,140160572	1,125888645	1,8694169
0,843620175	1,591	1,742339983	1,794029713	1,4978079	1,942524373	0,877069445	1	0,987482529	1,6396085
0,854314026	1,611	1,764426136	1,816771093	1,5167943	1,967148093	0,888187304	1,012676145	1	1,6603924
0,514525384	0,97	1,062656127	1,094181782	0,9135156	1,184748928	0,53492615	0,609901705	0,602267278	1

C

PEAK AREA		
Exons	Control	Sample
BCR 1	87882	117901
ABL 1	93026	62791
BCR 16	101140	117608
FLJ31568-2	121629	77317
PRDM12-12	95414	69163
ASS 4	105032	60021
BCR 2	49122	61133
ABL 6	82122	95885
ABL 11	76544	79898
EXOSC2-7	49316	31197

Exons									
BCR 1	ABL 1	BCR 16	FLJ31568-2	PRDM12-2	ASS 4	BCR 2	ABL 6	ABL 11	EXOSC2-7
1	1,988	1,153728554	2,110472504	1,8507844	2,347664134	1,077997836	1,149016861	1,285265358	2,1207651
0,503124534	1	0,580469142	1,061830495	0,9311751	1,181167424	0,542367159	0,578098573	0,646648535	1,067009
0,866755006	1,723	1	1,829262607	1,6041767	2,03484964	0,93436002	0,995916116	1,114010183	1,8381838
0,473827543	0,942	0,546668366	1	0,8769526	1,112387927	0,510785065	0,544435836	0,608994126	1,0048769
0,54031144	1,074	0,623372737	1,140312438	1	1,26846979	0,582454563	0,620826955	0,694443577	1,1458736
0,425955308	0,847	0,491436802	0,898966966	0,7883515	1	0,4591789	0,489429831	0,547465602	0,9033512
0,927645647	1,844	1,070251272	1,957770632	1,7168721	2,177800415	1	1,06588049	1,192270815	1,9673185
0,870309248	1,73	1,00410063	1,836763737	1,6107548	2,043193806	0,938191485	1	1,118578327	1,8457215
0,778049446	1,546	0,897657863	1,642051963	1,4400018	1,826598779	0,838735619	0,893991932	1	1,6500601
0,471527941	0,937	0,544015249	0,995146754	0,8726966	1,106989235	0,5083061	0,541793554	0,606038528	1

D

PEAK AREA		
Exons	Control	Sample
BCR 1	87882	99591
ABL 1	93026	112462
BCR 16	101140	55618
FLJ31568-2	121629	58723
PRDM12-12	95414	111782
ASS 4	105032	100640
BCR 2	49122	50450
ABL 6	82122	79455
ABL 11	76544	76479
EXOSC2-7	49316	51537

Exons									
BCR 1	ABL 1	BCR 16	FLJ31568-2	PRDM12-2	ASS 4	BCR 2	ABL 6	ABL 11	EXOSC2-7
1	0,937	2,060761551	2,347194414	0,9672982	1,182690664	1,103405213	1,171273849	1,134198619	1,0843984
1,066795793	1	2,198411752	2,503977125	1,0319097	1,261689424	1,177108039	1,249510014	1,209958315	1,1568317
0,485257501	0,455	1	1,138993695	0,4693887	0,573909516	0,535435656	0,568369421	0,550378387	0,5262125
0,42604055	0,399	0,877967985	1	0,4121083	0,503874181	0,470095364	0,499010155	0,483214604	0,4619977
1,033807341	0,969	2,130430418	2,426546815	1	1,22267429	1,140708409	1,210871503	1,172542858	1,121059
0,845529631	0,793	1,742434953	1,984622426	0,8178793	1	0,932961802	0,990346745	0,958998539	0,916891
0,906285368	0,85	1,867638041	2,127227954	0,8766482	1,071855244	1	1,061508352	1,027907613	0,9827744
0,853771303	0,8	1,759419074	2,003967233	0,8258515	1,009747349	0,942055706	1	0,968346233	0,9258282
0,88167979	0,826	1,816931811	2,069473877	0,8528473	1,042754456	0,972850076	1,032688481	1	0,9560922
0,922170296	0,864	1,90037309	2,164512968	0,8920137	1,0906422	1,017527512	1,080113952	1,045924277	1

3.4.1 Sensitivity and validation of MLPA

CML blood samples contain a variable background of normal Ph-chromosome negative cells. Therefore, sensitivity of the MLPA assay to detect der(9) deletions was initially established by using CML patient DNA known to have der(9) deletions by FISH in normal DNA. The deletion was detected in 100 %, 80 %, and 60 %

dilutions of DNA from the two deleted CML patients, but was not detected at lower dilutions (table 14 and figure 11). MLPA is therefore capable of detecting deletions in the great majority of pre-treatment CML samples, which has been previously shown by quantitative Southern blot analysis to harbour a median of 84 % (range, 64-100 %) cells derived from the malignant clone (Reiter, *et al.* 1997).

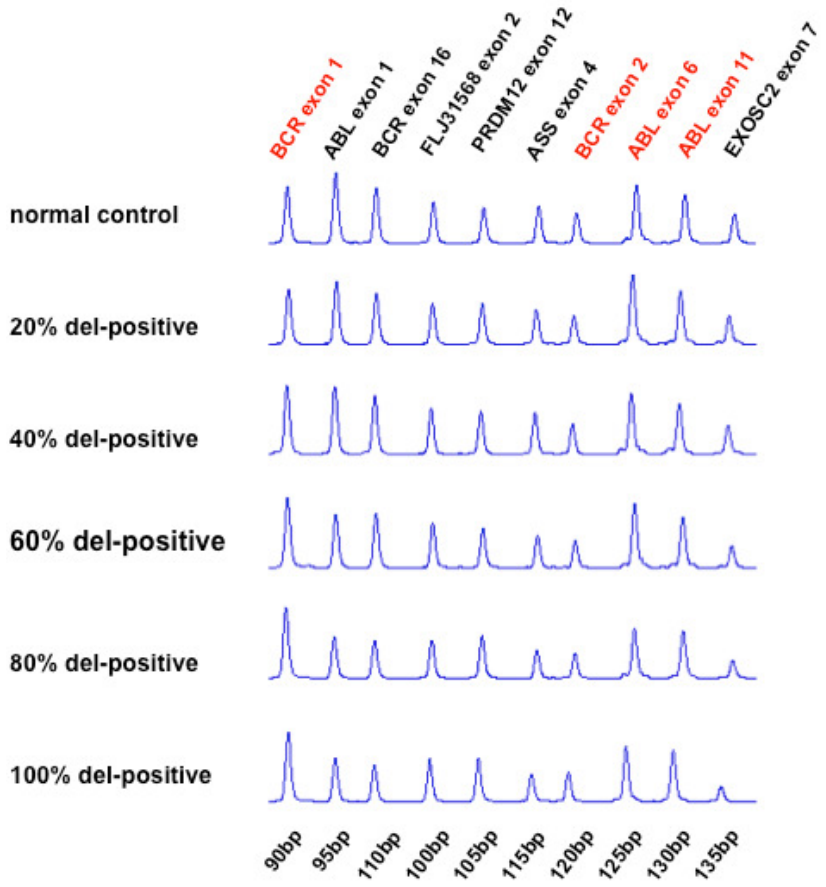


Figure 11: Sensitivity of MLPA

Example of MLPA traces of CML DNA known to harbour a large der(9) deletion diluted in non-deleted DNA. The relative ratio of der(9) to control der(22) probes reduces as the proportion of patient DNA increases (control probes are labelled in red).

To validate the MLPA assay, 20 samples from healthy individuals and 18 samples from CML patients with der(9) deletions as determined by FISH (kindly provided by Dr Claudia Haferlach, Munich, Germany) were tested. No deletions were detected in the healthy individuals: of the 480 individual measurements (six test probes relative to four control probes for the 20 individuals), the mean dosage

quotient was 1.001 (standard deviation 0.07) and only a single measurement was outside the range 0.7-1.3. In contrast, deletions were detected in 17 of the 18 CML samples. The CML case that was not detected had only 16 % der(9) positive cells, and thus was well below the sensitivity of detection.

Table 14: Sensitivity of MLPA

The ability of MLPA to detect deletions when the proportion of patient DNA known to harbour deletions is 60 % or higher. Data were derived from analysis of triplicate replicates of the traces shown in figure 11. '+' indicates not deleted (dosage quotient $\geq$0.7), 'del' indicates deleted (dosage quotient < 0.7). At ≥60 % patient DNA (marked in **bold**), all six markers are correctly scored as deleted whereas at lower dilutions deletions are often missed.

Deletion Proportion	ASS	PRDM12	EXOSC2	ABL exon 1	BCR exon 16	FLJ31568
100%	del	del	del	del	del	del
80%	del	del	del	del	del	del
60%	**del**	**del**	**del**	**del**	**del**	**del**
40%	del	+	del	del	+	+
20%	+	+	del	+	+	+
0%	+	+	+	+	+	+

3.5 Der(9) deletions in the patient cohort

Having validated the MPLA assay, I then analysed the patient study group and detected der(9) deletions in 59 of the 339 patients (17 %). Deletions encompassed both chromosome 9 and chromosome 22 derived sequences in 21 cases (36 %), or were either upstream only (n=20) or downstream only (n=18) of the *ABL1-BCR* fusion point (figure 12). The median age of the 59 patients with deletions was 49 years (range 11-72) with 54 % being male. According to the Hasford score, 26 patients were low risk, 24 were intermediate risk, and 9 were high risk. Therapy subsequent to IFN included imatinib (n=13) and allogeneic stem cell transplantation (n=20). After a median observation time of 8 years (range, 0.5-16), 33 (56 %) of the der(9) deleted cases were still alive.

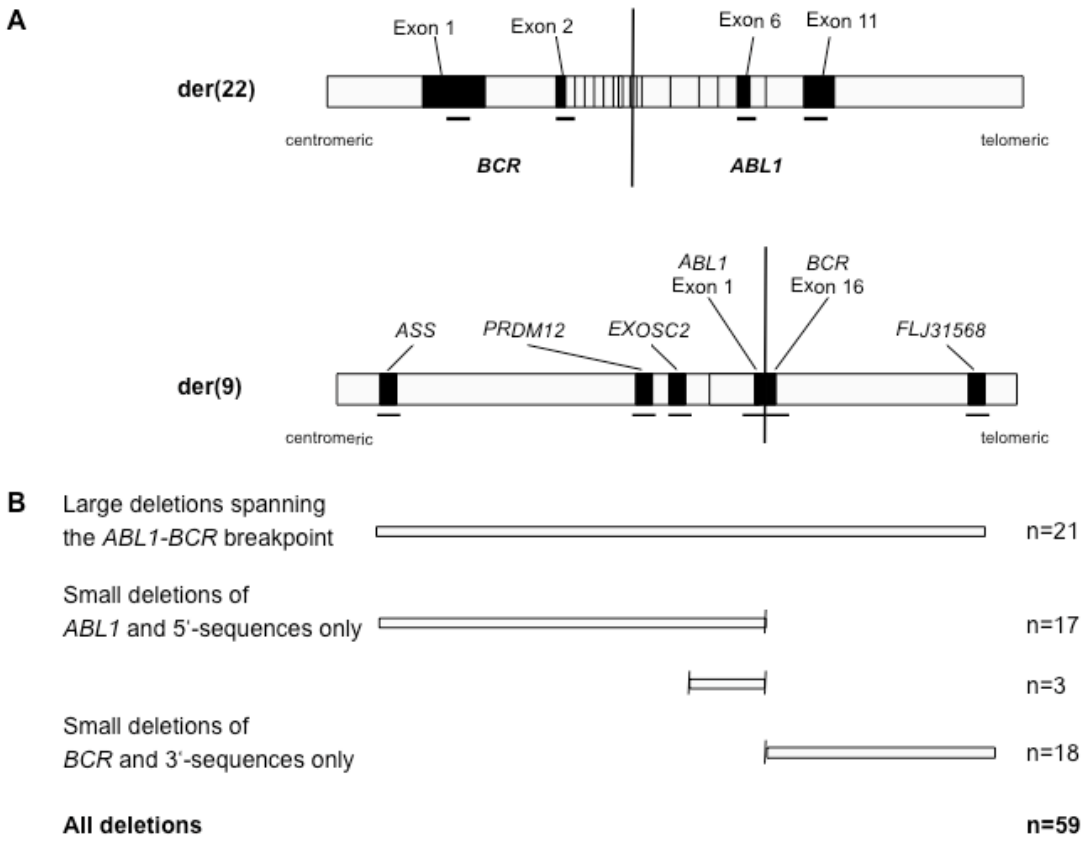


Figure 12: Summary of the MLPA assay to detect der(9) deletions

Top panel: map showing positions of the der(9) and der(22) probes (not to scale). Probes on the der(9) were designed to detect deletions whereas der(22) probes acted as controls. Bottom panel: summary of patient results.

Of the 47 deleted cases for which cytogenetic data were available, 10 (21 %) achieved a major cytogenetic response (MCR) and 5 (11 %) achieved a complete cytogenetic response (CCR). These proportions were somewhat lower than that seen for the 239 evaluable cases that did not have deletions, of whom 66 (28 %) achieved MCR and 30 (13 %) achieved CCR. However, no significant difference was seen between the overall survival of the 59 deleted cases compared to the 280 non-deleted patients (P=0.25; median survival: 9.8 versus 7.8 years; figure 13). If chronic phase patients were censored at the time of switchover to imatinib or allogeneic stem cell transplantation (SCT), the difference between both groups further declined (P=0.55; median survival 8.4 vs. 6.8 years; figure 14). Furthermore, no impact of deletion status on survival was observed for the 105 patients who underwent allogeneic SCT in first chronic phase (P=0.61; figure 15). In this group the median survival was not reached. The five-year survival for the

18 patients who harboured deletions was 0.67 compared to 0.63 for the 87 non-deleted cases.

3.5.1 *ABL1-BCR* breakpoint-spanning deletions are correlated with impaired survival

To determine if the position or size of the deletion might be important, I considered those individuals with deletions that spanned the *ABL1-BCR* junction separately from those that had deletions on the *ABL1* or *BCR* sides only. Of the 21 cases with breakpoint spanning deletions, 20 were missing all six markers and one lacked all markers apart from *FLJ31568*. Overall, the survival for these 21 patients was significantly worse than for the non-deleted patients (P=0.003; median survival 4.7 years versus 7.8 years; figure 16). Censorship at the time of imatinib or allogeneic SCT in chronic phase leads to a loss of power and significance (P=0.078; median survival: 6.1 vs. 6.8 years), but the trend is still apparent (figure 17). Despite this adverse association, five of the 19 evaluable patients with junction spanning deletions achieved MCR and two achieved CCR on IFN-therapy. The survival for the six patients with *ABL1-BCR* spanning deletions who underwent allogeneic SCT in first chronic phase was significantly shorter compared to the 87 non-deleted patients (P=0.030; figure 18).

3.5.2 Deletions of either *ABL1* or *BCR* sequences only

Of the 20 patients with deletions only on the *AB1L* side of the der(9) fusion point, 17 included all four markers and three included only *ABL1* exon 1b and *EXOSC2*. By definition, all 18 cases with deletions on the *BCR* side only included both markers since I only scored cases as being deletion positive if they lacked two consecutive markers. Whilst no difference in survival was seen between cases with deletions either upstream only (*ABL1* side: four patients died; eight-year survival = 0.80) or downstream only (*BCR* side: four patients died: eight-year survival = 0.78) of the breakpoint, unexpectedly I found that these deletions were associated with a superior survival when compared with the 280 cases without deletions (P=0.029

for *BCR* deletion only; P=0.014 for *ABL1* deletion only; P=0.001, all 38 grouped together; figure 19). As before, when patients in chronic phase were censored at the time of IM or SCT treatment, none of the comparisons retained significance (figure 20). The superior impact of the one-sided deletions was also not detectable with regard to outcome after allogeneic SCT in first chronic phase (87 without deletions vs. 12 patients with one-sided deletions; P=0.055; five-year survival: 0.63 and 0.92; figure 21). Of the 28 cases with one-sided deletions and cytogenetic data available, 5 patients (18 %) achieved MCR and 3 patients (11 %) achieved CCR.

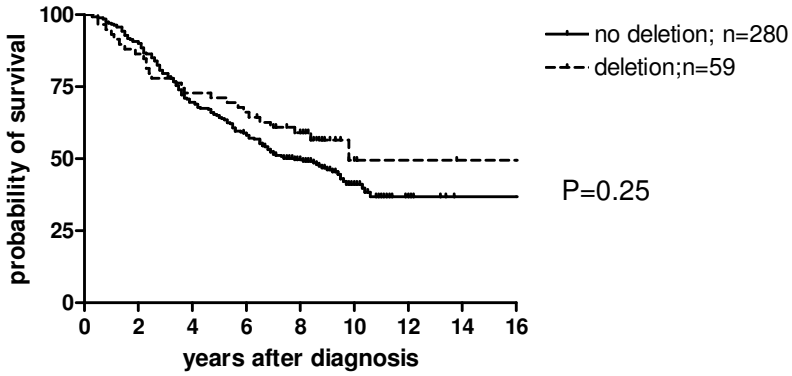


Figure 13: Deletions of der(9) – overall survival

Kaplan-Meier survival plots for the 59 patients with deletions compared to the 280 patients without deletions.

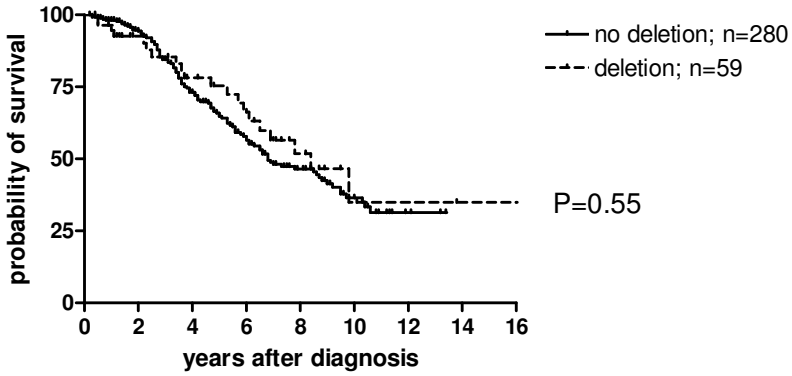


Figure 14: Deletions of der(9) – censored overall survival

Kaplan-Meier survival plots for the 59 patients with deletions compared to the 280 patients without deletions. Patients were censored at time of switch to imatinib or allogeneic stem cell transplantation in first chronic phase.

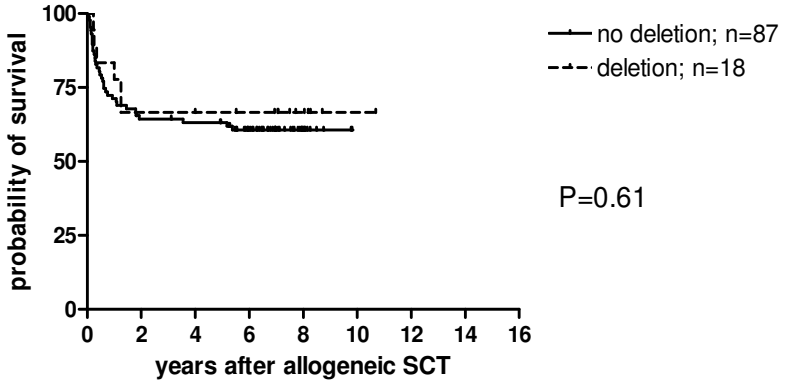


Figure 15: Deletions of der(9) – survival after allogeneic stem cell transplantation (SCT)
Kaplan-Meier survival plots for the 59 patients with deletions compared to the 280 patients without deletions.

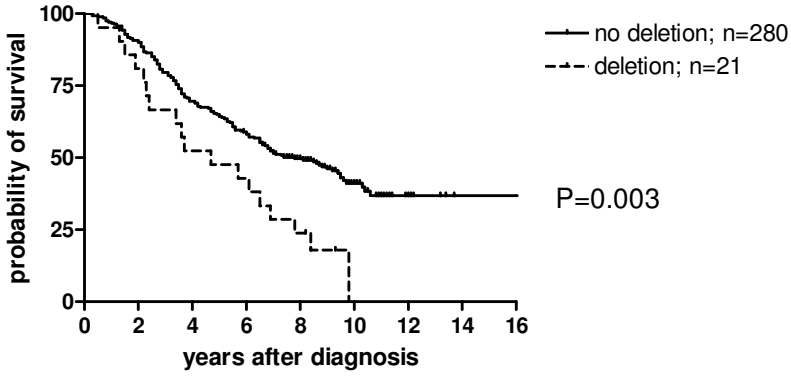


Figure 16: *ABL1-BCR* deletions – overall survival
Kaplan-Meier survival plots for the 21 patients with *ABL1-BCR* spanning deletions compared to the 280 patients without deletions.

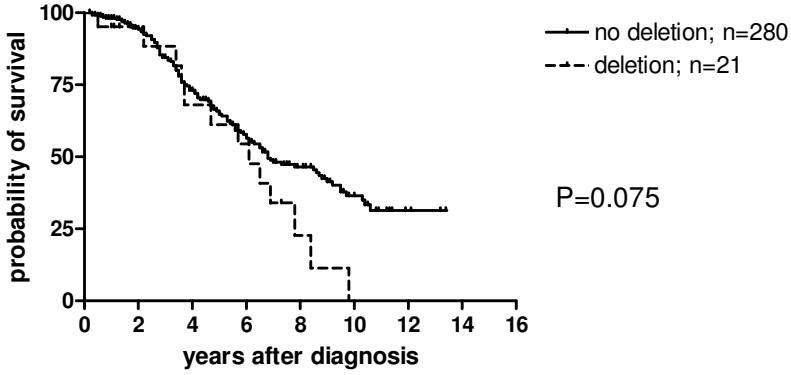


Figure 17: *ABL1-BCR* deletions – censored overall survival
Kaplan-Meier survival plots for the 21 patients with *ABL1-BCR* spanning deletions compared to the 280 patients without deletions. Patients were censored at time of switch to imatinib or allogeneic stem cell transplantation in first chronic phase.

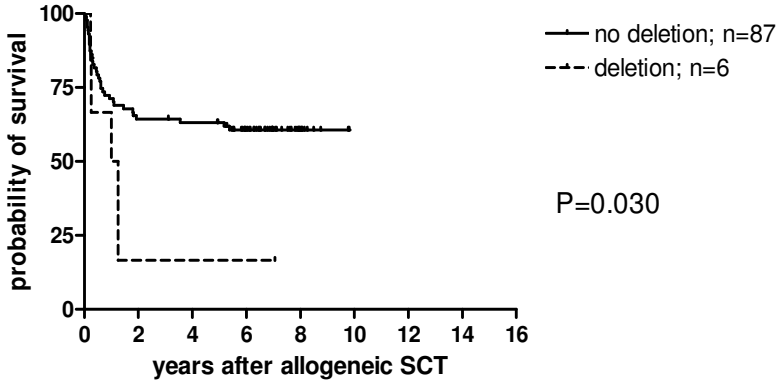


Figure 18: *ABL1-BCR* spanning deletions – survival after allogeneic stem cell transplantation (SCT)

Kaplan-Meier survival plots for the 6 patients with deletions compared to the 87 patients without deletions.

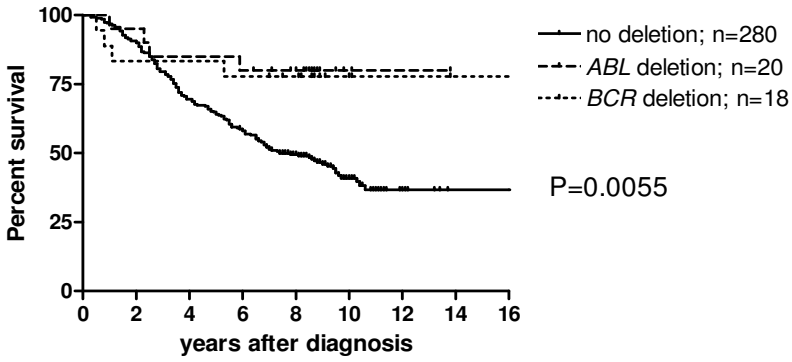


Figure 19: *ABL1* or *BCR* deletions only – overall survival

Kaplan-Meier survival plots for the 20 patients with *ABL1* deletions and 18 patients with *BCR* deletions compared to the 280 patients without any deletion.

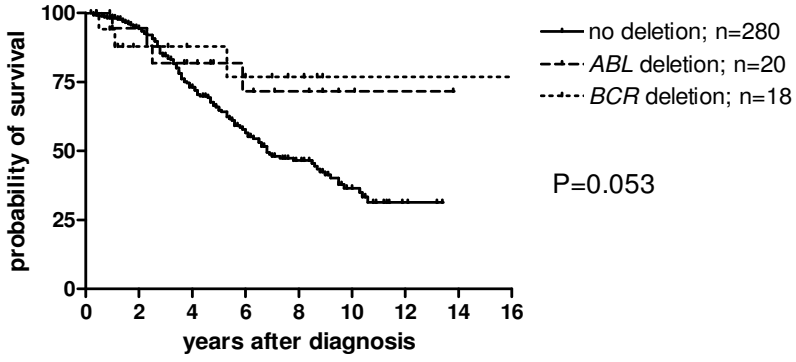


Figure 20: *ABL1* or *BCR* deletions only – censored overall survival

Kaplan-Meier survival plots for the 20 patients with *ABL1* deletions and 18 patients with *BCR* deletions compared to the 280 patients without any deletion. Patients were censored at time of switch to imatinib or allogeneic stem cell transplantation in first chronic phase.

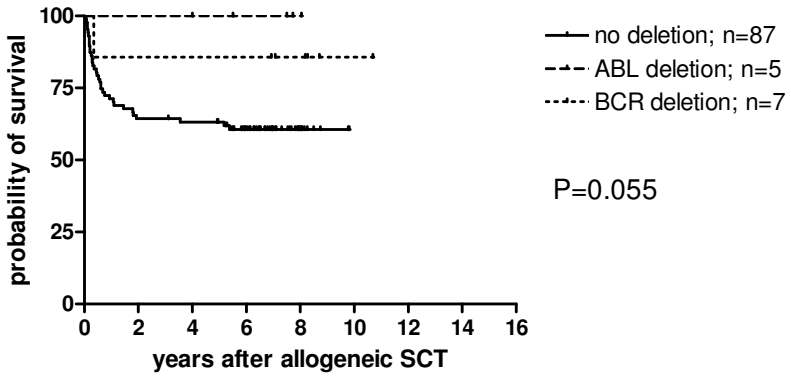


Figure 21: *ABL1* or *BCR* deletions only – survival after allogeneic stem cell transplantation (SCT)

Kaplan-Meier survival plots for the 20 patients with *ABL1* deletions and 18 patients with *BCR* deletions compared to the 280 patients without any deletion.

3.5.3 Comparison between deletion status and haematologic parameters

To determine if deletion status correlated with standard clinical and haematological parameters (age, spleen enlargement, white blood cell count, haemoglobin, and the proportions of blasts, eosinophils, basophils and platelets in PB), univariate analysis was performed for patients with (i) deletions that spanned the *ABL1-BCR* breakpoint (ii) deletions upstream or downstream of *ABL1* or *BCR* or (iii) patients without deletions. This analysis revealed that spleen enlargement was larger in patients with deletions not spanning the breakpoint as compared to either patients with spanning *ABL1-BCR* deletions or no deletions (P=0.002). Patients with *ABL1* or *BCR* deletions only were younger (P=0.030) and showed a higher proportion of blasts than others (P=0.004). All other parameters showed no significant differences (table 15).

Table 15: Prognostic risk scores and haematologic parameters at diagnosis according to 9q+ deletion status

Spleen size is larger in patients with small deletions as compared to patients without deletion and patients with large deletions.
Patients with small deletions are younger and show a higher proportion of blasts than patient with large deletions. There is a trend towards higher basophil counts in patients with small deletions compared to patients with large deletions. All other differences are not significant.

		Total		**No deletion**		**p**	**Small deletions**		**p**	**Large deletions**		**Total deletions**	
		n=348		n=289			n=38			n=21		n=59	
		median	range	median	range		median	range		median	range	median	range
Hasford Score	% low risk	38.8		37.6			40.5			52.4		44.8	
Sokal Score	% low risk	38.7		38.0			39.5			47.6		42.4	
Age	years	51	10-83	51	10-83		45	11-65	0.030	56	21-72	49	11-72
Spleen size	cm below costal margin	2	0-30	2	0-30	0.021	5	0-18	0.0023	0	0-9	3	0-18
WBC count	/nl	104	6-650	98.7	6-650		120.8	19.6-396		90.3	21.5-494	117	19.6-494
Blast PB	%	1	0-22	1	0-22		1.5	0-10	0.0038	0	0-6	1	0-10
Eosinophils PB	%	2	0-51	2	0-51		4	0-14		2	0-7	3	0-14
Basophils PB	%	3	0-40	3	0-40		4	0-14	0.10	2	0-7	3	0-14
Platelets	/nl	384	49-3490	380	49-3490		376	184-1986		397	124-1109	397	124-1986
Haemoglobin	g/dl	12.0	6.0-16.3	12.0	6.0-16.3		11.7	7.4-16.3		12.6	7.8-15.4	11.8	7.4-16.3

3.5.4 Independent influence of deletion status on survival probabilities

Cox regression analysis was used to determine if deletion status had an independent influence on survival. In order to attribute survival probabilities to IFN treatment, survival times were censored at the start of imatinib therapy or the date of allogeneic SCT for patients still in first chronic phase. The aim was to establish the prognostic value of deletion status in relation to the Hasford score, which was specifically developed and validated to differentiate three prognostic risk groups with regard to survival following treatment with IFN (Baccarani*, et al.* 2002, Bonifazi*, et al.* 2000, Hasford*, et al.* 1998, Kluin-Nelemans*, et al.* 2004, Pfirrmann and Hasford 2001).

Of 338 patients with deletion status and Hasford score available, 131 had died at the time of analysis. The low-risk group contained 132 patients (median survival: 8.6 years), the intermediate-risk group 166 (median survival: 6.8 years), and the high-risk group 40 patients (median survival: 5.6 years). The survival probabilities between the three prognostic groups were significantly different (P=0.049). As prognostic factors in a common Cox model, deletion status and the Hasford score kept their statistical significance (P=0.007 and P=0.011, respectively), which hints that both are an independent prognostic influence on survival (model A, table 16). By comparison, deletion status and Sokal score yielded slightly higher P-values (P=0.016 and P=0.028, respectively).

The Hasford score employs age, spleen enlargement, platelet count, and proportions of blasts, eosinophils, and basophils as covariates (Hasford*, et al.* 1998). The independent impact of these variables was assessed along with deletion status, age, haemoglobin, white blood cell count, and gender. This analysis yielded a final model (model B, table 17) with three independent variables: deletion status, age, and spleen enlargement with P= 0.007, P=0.018, and P<0.001, respectively. The P value for deletion status reflects both the adverse risk associated with breakpoint spanning deletions (P=0.026) and the beneficial effect of one-sided deletions (P=0.039).

Table 16: Multiple Cox regression – Model A

Deletion status and Hasford score as independent prognostic factors for survival probabilities [a]

Model A	n / died [b]	Estimation of coefficient β	Standard deviation of estimated β	Wald's chi-square statistic	P-value	Hazard ratio
Deletion status	338 / 131			10.084 (2 df) [c]	0.007	
no	279 / 111	Baseline	-	-	-	-
one side of breakpoint	38 / 7	-0.896	0.394	5.163	0.023	0.408
whole breakpoint	21 / 13	0.620	0.297	4.343	0.037	1.859
Hasford score	338 / 131			8.966 (2 df) [c]	0.011	
low	132 / 36	-0.279	0.204	1.876	0.171	0.756
intermediate	166 / 73	Baseline	-	-	-	-
high	40 / 22	0.546	0.249	4.818	0.028	1.727

[a] Survival times of patients with imatinib or allogeneic stem cell transplantation in first chronic phase were censored.

[b] For one patient, the Hasford score was not evaluable.

[c] df: degrees of freedom.

Table 17: Multiple Cox regression - Model B

Best prognostic model for survival probabilities [a]

Model B	n / died [b]	Estimation of coefficient β	Standard deviation of estimated β	Wald's chi-square statistic	P-value	Hazard ratio
Deletion status	335 / 130			9.784 (2 df) [c]	0.007	
no	276 / 110	Baseline	-	-	-	-
one side of breakpoint	38 / 7	-0.870	0.391	4.944	0.026	0.419
whole breakpoint	21 / 13	0.623	0.302	4.251	0.039	1.864
Age (in fully completed years)	335 / 130	0.018	0.007	5.627 (1 df) [c]	0.018	1.018
Spleen enlargement (in cm)	335 / 130	0.058	0.016	13.519 (1 df)[c]	<0.001	1.059

[a] Survival times of patients with imatinib or allogeneic stem cell transplantation in first chronic phase were censored.

[b] For the candidate variables deletion status, age, spleen enlargement, blasts, basophils, eosinophils, platelet count, haemoglobin, white blood cell count, and gender, 335 patients with complete cases were available of whom 130 patients died.

[c] df: degrees of freedom.

3.6 Discussion

Deletions of downstream *BCR* sequences at the der(9) fusion junction were first identified in a minority of CML patients by Southern blot analysis (Popenoe*, et al.* 1986). Development of FISH probes revealed that the deletions were larger than had been previously suspected and frequently included loss of *SMARCB1*, 500 kb telomeric of *BCR* (Grand*, et al.* 1999). Subsequently it was shown that the deletions were variable in size, often encompassed several mega bases and usually included both chromosome 9 and chromosome 22 derived sequences, i.e. they spanned the reciprocal *ABL1-BCR* junction (Sinclair*, et al.* 2000). This study was also the first to suggest that the deletions may be associated with a poor prognosis, a hypothesis that was expanded in a larger analysis from the same group in which the median survival of the 14 % of cases with deletions (most of which spanned the *ABL1-BCR* junction) was significantly shorter than those without deletions (Huntly*, et al.* 2001). The adverse prognosis was confirmed in an independent large study in which deleted patients had a shorter duration of chronic phase, inferior survival, and increased probability of relapse after SCT (Kolomietz*, et al.* 2001). A more recent study of IFN-treated cases however failed to detect any difference in clinical outcome between deleted and non-deleted cases, although deleted patients were found to present with significantly lower haemoglobin levels and higher leukocyte counts (Yoong*, et al.* 2005). Following imatinib therapy, der(9) deletions have not yet been reported to be a strong prognostic indicator. Huntly *et al.* found no difference in survival between patients with and without deletions, although deleted cases had more rapid disease progression and exhibited poorer haematologic and cytogenetic responses (Huntly*, et al.* 2003b). Quintas-Cardama *et al*., however, found no influence of der(9) deletions with regard to response, survival, or response duration (Quintas-Cardama*, et al.* 2005). It seems that much longer follow up will be required to determine whether deletions do or do not have prognostic value for imatinib-treated patients.

In my study I sought to (i) determine if deletion status genuinely did predict a poor prognosis for IFN-treated cases and (ii) determine the relationship of

deletion status to the Hasford and Sokal risk scores. I developed and validated a novel DNA-based MLPA deletion assay and investigated 339 patients enrolled over 15 years in three trials of the German CML Study Group with an observation time up to 16 years. This is the largest group of IFN-treated cases to be analysed for the impact of deletions. A similar proportion of deleted cases was found compared to other studies but my series is unusual in that only a relatively small proportion of deletions (36 % of deleted cases; 6 % of all cases) spanned the *ABL1-BCR* breakpoint. This compares to Huntly *et al*., who found that 12 % of deletions spanned the *ABL1-BCR* breakpoint and only 2 % and 1 % were restricted to the *ABL1* or *BCR* sides only, respectively (Huntly*, et al.* 2003b). Whether this difference is due to the different technical approaches (MLPA vs. FISH) that were employed is not clear, but in view of the validation of MLPA that was undertaken and the fact that the MLPA probe set was specifically designed to cover the region examined by FISH, it seems likely that technical reasons may only partially account for the differences.

Although no difference in survival between deleted and non-deleted cases was seen, more detailed analysis indicated that deletions that spanned the breakpoint were a significant indicator of inferior prognosis (P=0.039; table 17). Unexpectedly, I found that deletions on one side of the breakpoint only were associated with improved survival (P=0.026; table 17), resulting in an overall multivariate P value of 0.007 for deletion status. Improved survival was seen for both deletions on the *BCR* side only and on the *ABL1* side only. Since other published series only had a small proportion of cases with one-sided deletions, it is possible that any beneficial effect would have escaped notice. The biological explanation for a beneficial effect of one-sided deletions is not immediately obvious. Current models suggest that the poor prognosis associated with deletions is likely to be due to heterozygous loss of one or more loci rather than loss of *ABL1-BCR* expression, general genomic instability or effects on *BCR-ABL1* (Fourouclas*, et al.* 2006, Huntly*, et al.* 2002, Melo*, et al.* 1996). My findings are consistent with the hypothesis that at least two loci may be targeted, one on each side of the fusion junction. However it is not clear why deletion of just one of these loci (or other loci) may have the opposite effect and the possibility that this

result is in some way artefactual should not be dismissed. For example, a P value of 0.026 is expected to arise by chance approximately 1 in 38 times and although this is unlikely to be the case in any individual study, inevitably there must be large numbers of results in the literature with seemingly significant results that have in fact arisen by chance.

Currently, the Hasford score is the best predictor of outcome for CML patients treated with IFN and outperforms the Sokal score in this context (Hasford*, et al.* 1998, Pfirrmann and Hasford 2001). Information was available to calculate both scores for all but one of the cases in my study group and overall there was no difference in clinical baseline data between deleted and non-deleted cases. It is noteworthy, however, that there were relatively few Hasford or Sokal high risk patients (table 11) in the breakpoint spanning deletion group and a relative excess of high risk cases in patients with one-sided deletions. This is consistent with risk score and deletion status being independent prognostic variables.

To examine the relationship between the Hasford score and deletion status in more detail their contribution as prognostic factors in a common Cox model was examined. Both kept their statistical significance, suggesting they are independent prognostic influences on survival (table 16). When individual components of the Hasford score were considered independently, age and spleen enlargement were the most important prognostic variables in addition to deletion status (table 17). My findings thus confirm that at diagnosis, deletion status provides independent, statistically significant prognostic information with respect survival probability. However only deletions that span the der(9) breakpoint are an adverse risk factor and this needs to be taken into account in future studies and also for clinical management of individual cases. Although I used MLPA to define the deletion size, it should also be straightforward to use FISH to determine if any deletion is breakpoint spanning or one sided.

4 Nucleotide sequence alterations in haematologic malignancies

4.1 Introduction

The fate of a single cell is mainly determined by its response to extracellular but also intracellular stimuli and consequent effects on proliferation, differentiation, cell cycle, and apoptosis. Activation of a network of signalling molecules usually results in transmission of signals to the nucleus, activation of transcription factors and protein synthesis. Important regulators of the signalling cascade are tyrosine kinases and serine/threonine kinases, which phosphorylate tyrosine, and serine/threonine respectively; the activity of which is countered by dephosporylation by phosphatases and other negative regulatory feedback mechanisms. These signals are fed through multiple signal transduction pathways that affect gene transcription through the regulation of transcription factors as well as influencing the ability of specific mRNAs to be translated into proteins. Recently it has emerged that this process may be modified by small RNA fragments of 19-24 bp in length, so called micro RNAs that are encoded in genomic DNA but not translated into proteins (Ambros 2004, Bartel 2004, Havelange and Garzon 2010, Huang and Mo 2012).

As described above, constitutive activation of tyrosine kinase signalling pathways has emerged as a common theme in the pathogenesis of MPN as well as other malignancies. Initially, activated tyrosine kinases were found in CML and a range of atypical MPN but the finding of the *JAK2*^{V617F} point mutation indicated that tyrosine kinase activation was widespread in these disorders. Despite this significant advance, no causative mutations can be found in roughly 50 % of MPN cases. Our research group is interested in identifying new molecular abnormalities in MPN and as part of this I developed a large mutation screen for abnormalities of specific tyrosine kinases and related signalling proteins. In addition, I explored the hypothesis that other JAK family members may be mutated in lymphoproliferative disorders. Three populations of patients were studied that were diagnosed with either (i) myeloproliferative disorders (n=471), (ii) lymphoproliferative disorders (n=145), or (iii) polycythaemia vera

(PV) that was negative for *JAK2*^{V617F} (n=78). Fifteen different genes and two micro RNAs were selected for point mutation screening in these populations.

4.2 Patients and methods

4.2.1 Patients

(i) Myeloproliferative disorders. Between 1999 and 2006, a total of 471 samples from patients with myeloid malignancies were collected either in the research laboratory Mannheim, Germany, or the Wessex Regional Genetics Laboratory, Salisbury, UK. The population included patients with acute myeloid leukaemia (AML; n=8), unclassified atypical myeloproliferative disease (aMPN; n=284), chronic myelomonocytic leukaemia (CMML; n=7), chronic neutrophilic leukaemia (CNL; n=3), essential thrombocytosis (ET; n=28), hypereosinophilic syndrome (HES; n=20), primary myelofibrosis (PMF; n=35), myelodysplastic syndrome (MDS; n=1), polycythaemia vera (PV; n=65), and systemic mastocytosis (SM; n=20).

(ii) Lymphoproliferative disorders. A total of 145 patients with chronic lymphocytic leukaemia (CLL; n=48), acute lymphoblastic leukaemia (ALL; n=32), non-Hodgkin lymphoma (NHL; n=24), multiple myeloma (MM; n=23), and mantle cell lymphoma (MCL; n=18) were included in the study.

(iii) *JAK2*^{V617F} negative PV. Patients presented with the clinical diagnosis of PV but tested negative for the *JAK2*^{V617F} point mutation using sensitive tests (Jones, *et al.* 2005, Jones, *et al.* 2006).

4.2.2 Methods

Mutation screening was performed by two different methods, i.e. conformation-sensitive capillary electrophoresis (CSCE) and high-resolution melting curve

analysis (HRM), both of which are explained above. Primer sets were designed for both assays according to the prevailing recommendations for either test. Confirmation of possible base changes was achieved by direct genomic sequencing of the PCR products. Genes were selected for analysis on the basis that they are known or suspected to play an important role in haematopoietic development. Genes that were considered particularly good candidates were screened in full whereas other were only partially screened, focusing on regions where mutations are most likely to be found.

(i) In myeloid disorders, nine genes that regulate haematopoiesis or contribute to haematologic malignancies were screened for mutations (table 18). All coding exons of *JAK2* and *CSF1R* were analysed i.e. 23 exons in *JAK2* (exons 3 to 25), and 21 exons in *CSF1R* (exons 2 to 22). For other genes, only sequences encoding the catalytic domain of those tyrosine kinases were analysed; i.e. *CSK* (exon 13), *FES* (exon 17), *SYK* (exon 11), *STYK1* (exon 8), and *TIE1* (exon 18). *JAK1* and *JAK3*, both member of the Janus kinase family, were screened for mutations just in the exon that encodes the regions that corresponds to residue V617 in *JAK2*.

Table 18: CSCE and HRM primer for mutation screening in myeloid leukaemia

Gene	Exon	Assay	Primer A (forward)	Primer B (reverse)
CSK	13	CSCE	*TGTAAAACGACGGCCAGT*TCTGAGGGGACATGTGGCTG	*CAGGAAACAGCTATGACC*AGGACTCCTGCCCACCTCTGT
FES	17	CSCE	*CAGGAAACAGCTATGACC*GCGACTGTTGGCCAAATGAGC	*TGTAAAACGACGGCCAGT*AGCGTGAGTGCACCCTCCCT
CSF1R	2	CSCE	*TGTAAAACGACGGCCAGT*GGAAAACAAGACAAACAGCCAGTGC	*CAGGAAACAGCTATGACC*CATCACACCCCAACAAAGTCCC
CSF1R	3	CSCE	*CAGGAAACAGCTATGACC*ACCTCACTTCTGTCTGCTGTCCC	*TGTAAAACGACGGCCAGT*TCCAGGTCCTTGCTCATAGCC
CSF1R	4	CSCE	*TGTAAAACGACGGCCAGT*TGAGCTTCCTTGTCACCTATGTCCA	*CAGGAAACAGCTATGACC*GGCTCTCTGTCCCCACTCTTCAG
CSF1R	5	CSCE	*TGTAAAACGACGGCCAGT*GATTTGGGATGGGGCAGTG	*CAGGAAACAGCTATGACC*GATGGCCAGACTTCACCTTGTG
CSF1R	6	CSCE	*CAGGAAACAGCTATGACC*GAACCCAGATCTTTCTAATCCCTAA	*TGTAAAACGACGGCCAGT*CTTCCTCCTCAAAACCCTTCT
CSF1R	7	CSCE	*CAGGAAACAGCTATGACC*GGATGACAAAATGGACAAATAAA	*TGTAAAACGACGGCCAGT*CTCAAGGTCATACACCAAGGTT
CSF1R	8	CSCE	*CAGGAAACAGCTATGACC*GGGCTCACCAACACCACCTGATTCT	*TGTAAAACGACGGCCAGT*CCAGTCAACCCCATCCCTCCAG
CSF1R	9	CSCE	*TGTAAAACGACGGCCAGT*CCTGGAGGGATGGGGTTGACT	*CAGGAAACAGCTATGACC*CACCTGGCAAAAGCAGAATGTGG
CSF1R	10	CSCE	*TGTAAAACGACGGCCAGT*CCTGAAGCTCTGAGCCACATTCTGC	*CAGGAAACAGCTATGACC*AGGGGAGGAGCCGCCTAAAGGA
CSF1R	11	CSCE	*TGTAAAACGACGGCCAGT*CATTTATTGAGCACCCACTGTGTTC	*CAGGAAACAGCTATGACC*GGACTGACTGAGGAGGAGGAAGG
CSF1R	12	CSCE	*TGTAAAACGACGGCCAGT*AAACAAGGTCCAGGAACTGC	*CAGGAAACAGCTATGACC*AAAGGGCCTCTGTCCAAG
CSF1R	13	CSCE	*CAGGAAACAGCTATGACC*ACTTGGACAGAGGCCCTTTGGTGC	*TGTAAAACGACGGCCAGT*TACTGTCCCCCAGGCTTCAACAGGTT
CSF1R	14	CSCE	*TGTAAAACGACGGCCAGT*TGTGTCTGCATTTTTCCTTCT	*CAGGAAACAGCTATGACC*CGATGCTCCCTATCAGTCTCT
CSF1R	15	CSCE	*TGTAAAACGACGGCCAGT*CTATGCTGAGTCCAAGTCCAGT	*CAGGAAACAGCTATGACC*CCCAGAACACGGTAGGTG
CSF1R	16/17	CSCE	*TGTAAAACGACGGCCAGT*GAGCAGTGCAGTGATGATGA	*CAGGAAACAGCTATGACC*ATACTCTGTCTCCTCCCCTACC
CSF1R	18/19	CSCE	*CAGGAAACAGCTATGACC*TAGGAGAAGGCCCAAGACTAAC	*TGTAAAACGACGGCCAGT*CCAACCCTCGCTGTGTC
CSF1R	20	CSCE	*TGTAAAACGACGGCCAGT*TCACACCTTCCTAAAGCCTAAG	*CAGGAAACAGCTATGACC*TCTGAGAACATCCCCCATAG
CSF1R	21	CSCE	*TGTAAAACGACGGCCAGT*CAGAATGAGGGTGATGAGAGTA	*CAGGAAACAGCTATGACC*GAGGGACCAGAAACTGTCAG
CSF1R	22	CSCE	*TGTAAAACGACGGCCAGT*CAGACCCACCTTCCAGCAGAT	*CAGGAAACAGCTATGACC*GAGTTTGTATGTTCTCCCCGTGTC
JAK1	13	CSCE	*CAGGAAACAGCTATGACC*GCCAAGGCCAGAGGATTGAT	*TGTAAAACGACGGCCAGT*TCCCAGGTCTAGCACAGTGTAGG
JAK2	3	HRM	GCCCAGCCCATTTGTAACTT	AGCATAATTAAAACAATCCA
JAK2	4	HRM	TTAATAATTCCTTTCTCTGC	TTGATAATTTCTCCCATTT
JAK2	5	HRM	TTGTATTTGAACTATTTGGA	TGACTTATTAAACAGCAAGA
JAK2	6	HRM	AATGAAACTTACGATGAGAT	TGGGTAAACACAGAAAATAT
JAK2	7	HRM	TGTTTTTCTGTATGTGCTTT	CTTCCAGGTTCTTTTACTTC
JAK2	7	HRM	TATAAATCTGGAAACTCTGC	TCAAATGGGAGAAGTGCA
JAK2	8	HRM	TGTTTTAAAATGGCTCTGTA	ACAATCAAAATTCCTACCAG
JAK2	9	HRM	GGAGTTGACTTTCTAAAAGG	TGTAAATACAGCATTTACT
JAK2	10	HRM	ATAATATTATGGTGCTTGAT	AATAAATATGGAGAGGAAAT
JAK2	11	HRM	CTATCCCTCCCTTTCTTTAT	AACAATCCATGACCAGTAAT
JAK2	12	HRM	GCAATTCATACTTTCAGTGT	AAAATAATTCCAATGTCACA
JAK2	13	CSCE	*CAGGAAACAGCTATGACC*CTTCGTTCTCCATCTTTACTCATTCT	*TGTAAAACGACGGCCAGT*TGAGAGCACATCTTTAAACAGCATAA
JAK2	14	CSCE	*TGTAAAACGACGGCCAGT*GACAACAGTCAAACAACAAT	*CAGGAAACAGCTATGACC*CTATATAAACAAAAACAGATGCTC
JAK2	15	CSCE	*CAGGAAACAGCTATGACC*ACTTAAGCCTTATTATTATTACTT	*TGTAAAACGACGGCCAGT*ATCCAATTACAGATTTATTTT
JAK2	16	CSCE	*TGTAAAACGACGGCCAGT*TGTGGACTGATATTTGAAT	*CAGGAAACAGCTATGACC*CCAAGTAAAGCTTAGTAACAT
JAK2	17	CSCE	*CAGGAAACAGCTATGACC*TTACTTGTGAATTATTTAACCCTACT	*TGTAAAACGACGGCCAGT*AAAGAAATATGAAAGTTCAGAGATAG
JAK2	18	CSCE	*CAGGAAACAGCTATGACC*ATTTTAAAAAGAAGGTTGGTGTG	*TGTAAAACGACGGCCAGT*ATGAAAATAATTCTCAGTCAAGGAG
JAK2	19	HRM	GTCCAGAGAATGTTATTTGC	CGCTCTTAAGTTTTTCCTAA
JAK2	20	HRM	TTGTAATTTGCCTTGAAAAC	GCCAAACACATACCTTGAAT
JAK2	21	HRM	TAAAACATTATTTCCACCTT	ACGTCTAGATGATTTCCTAA
JAK2	22	CSCE	*TGTAAAACGACGGCCAGT*CGTGTTGAAGTAGACATTAGGAAATCATCT	*CAGGAAACAGCTATGACC*TGAAGTCTGTGCTCTAAAATTTCATTTATCAT
JAK2	23	HRM	GGTAAAATCAAGAGTCCACA	AATCAAACGAACAAACAAAA
JAK2	24	HRM	TGAGAAAGAATTTTGCTACA	TGAACTGAAGTCCTTGAGTA
JAK2	25	HRM	TGAAAATTAATGTCTTCCAC	ATAATAGTCCACAGCAATGT
JAK3	13	CSCE	*CAGGAAACAGCTATGACC*TTGGGATTATTGGAGTGGAAGAAAC	*TGTAAAACGACGGCCAGT*CAGAACAGAGGTGGGAAGAACAGC
STYK1	8	CSCE	*TGTAAAACGACGGCCAGT*CTTGTGTTTTTGAAGATTCC	*CAGGAAACAGCTATGACC*CCCAAGTCATGAAGAAAGTAA
SYK	11	CSCE	*TGTAAAACGACGGCCAGT*GTTGTTTGTTGTGAAATGGGTCA	*CAGGAAACAGCTATGACC*TTTTTACCACCTGCTGCTCTGAATA
TIE	18/19	CSCE	*CAGGAAACAGCTATGACC*CTCATCCCCAGTCTCTCCTGA	*TGTAAAACGACGGCCAGT*CCCTCCTGTGCCCTCTCA

Tag sequences are in *italics*, and sequences corresponding to the region of interest are in plain type.

A

JAK1	634	A**F**F**E**A**AS**M**M**RQV**S**HK**H**IVYLY**GVC**V 658
JAK2	595	S**F**F**E**A**AS**M**M**SKL**S**HK**H**LVLNY**GVC**V 619
JAK3	568	S**F**L**E**A**AS**L**M**SQV**S**YR**H**LVLLH**GVC**M 592
TYK2	654	A**F**Y**E**T**AS**L**M**SQV**S**HT**H**LAFVH**GVC**V 678

B

JAK1	572	**Q**LS**F**DRILKKDL 583
JAK2	534	**Q**MV**F**HKIRNEDL 545
JAK3	510	**Q**MT**F**HKIPADSL 521
TYK2	578	**Q**LS**F**HRVDQKEI 589

Figure 22: JAK2 homologous amino acid sequence in Janus kinase family members
(A) Homologous sequences (conserved amino acids in **bold** letters) in the *JAK2*^{V617} region (underlined) in the four JAK family members *JAK1*, *JAK2*, *JAK3* and *TYK2*. (B) Homology between JAK family members in the region affected by *JAK2* exon 12 mutations (Scott, *et al.* 2007).

(ii) Mutation screening in lymphoid malignancies focused on *JAK2* exon 14 (which encodes residue V617) and the homologous regions of *JAK1* (i.e. exon 13), *JAK3* (exon 13), and *TYK2* (exon 12). In addition, *JAK2* exon 12 has recently been shown to be mutated in V617F-negative PV (Scott, *et al.* 2007) and the corresponding region of *JAK1* (exon 12) and *JAK3* (exon 11), respectively, were also analysed. Amino acid conservations in these regions are represented in figure 22 and a list of all primers for these assays is shown in table 19.

Table 19: HRM primer for TK screening in lymphoid malignancies

Gene	Exon	Assay	Primer A (forward)	Primer B (reverse)
JAK1	12	HRM	TCGGATCTTGGGTGCTGTGC	CCCCAAGCTGCTCCATCGTG
	13	HRM	GCCAAGGCCAGAGGATTGAT	TCCCAGGTCTAGCACAGTGTAGG
JAK2	12	HRM	GCAATTCATACTTTCAGTGT	AAAATAATTCCAATGTCACA
	14	HRM	GACAACAGTCAAACAACAAT	CTATATAAACAAAAACAGATGCTC
JAK3	11	HRM	TTTTCACCTCTGATTTCTGG	CATCTGAGTCTCTGGACCC
	13	HRM	TTGGGATTATTGGAGTGGAAGAAAC	CAGAACAGAGGTGGGAAGAACAGC
TYK2	10	HRM	CGCCCTCCATGACTTGAT	ACCGCCATGGTGAAAGTTAG
	12	HRM	GGGAGGGGTCGGCTTTGC	GGGTTGGGGTACAGATCAGGGAG

(iii) The population of red blood cells (erythrocytes) is mainly regulated by the cytokine erythropoietin (EPO) and its corresponding receptor pathway. Erythropoietin is formed in response to hypoxia detected in suprarenal cells. The von Hippel-Lindau gene (*VHL*) codes for a protein that is a component of the protein complex that is involved in the ubiquitination and degradation of hypoxia-inducible-factor (HIF) which plays a central role in the regulation of gene expression by oxygen. Mutations of *VHL* may lead to an accumulation of the downstream product EPO and has been shown to be the cause of many familial cases of polycythaemia (Ang*, et al.* 2002, Cario*, et al.* 2005, Pastore*, et al.* 2003). Because of its important role in the cytokine response mechanism, the complete *VHL* gene (exons 1 to 3) was therefore screened for mutations in the *JAK2*^{V617F} negative PV cohort.

Another obvious candidate in this context is the erythropoietin receptor (EPOR) itself, which is known to be mutated in some families with inherited erythrocytosis. I screened sequences encoding the intracellular domain of the receptor where mutations have been reported previously.

The post receptor-signalling cascade is mainly mediated by the JAK/STAT signalling pathway. It was decided not to screen *JAK2* for mutations since this had already been reported as having been fully screened in V617F negative PV and no mutations had been found (Baxter*, et al.* 2005, Levine*, et al.* 2005b), although subsequently some cases were identified with exon 12 mutations (Scott*, et al.* 2007). Instead, the study focused on an analysis of the regions encoding the catalytic domains of the BMX and BTK non-receptor tyrosine kinases both of which have been reported to play an important role in erythropoiesis or myelopoiesis (Saharinen*, et al.* 1997) (von Lindern*, et al.* 2004) Furthermore the genes encoding these kinases are both located on the X-chromosome, which could potentially have a bearing on the observation that there is a significant excess of males on *JAK2*^{V617F} negative PV (Jones*, et al.* 2005, Vizmanos*, et al.* 2006).

Finally, two micro RNAs, *hsa-mir-221* and *hsa-mir-222,* have been reported to inhibit normal erythropoiesis via down-modulation of the KIT receptor (Felli*, et al.* 2005). Loss of these micro RNAs might therefore result in increased erythropoiesis and I therefore screen the corresponding genes for sequence changes. Furthermore, both genes are also located on the X-chromosome. The KIT receptor itself was not considered a likely target of mutations in PV since previous screens had not identified any mutations and imatinib, which inhibits KIT, has limited effects in most PV patients (Silver 2003).

Table 20: CSCE primer for TK screening in *JAK2*^{V617F}-negative PV patients

Gene	Exon	Assay	Primer A (forward)	Primer B (reverse)
BMX	9	CSCE	*CAGGAAACAGCTATGACC*GCATGGCTACGCACTTGGATTATT	*TGTAAAACGACGGCCAGT*TACCAGCTCATTGTCACCTGTCATCT
BMX	10	CSCE	*CAGGAAACAGCTATGACC*TAATCATTCTCTGGAGCAGTTAG	*TGTAAAACGACGGCCAGT*TATAATTAGCAGGAAGGTTTCTATG
BMX	11	CSCE	*CAGGAAACAGCTATGACC*TAAAGAGATTGAGTGAGAAACCTG	*TGTAAAACGACGGCCAGT*TATCCCAAAATGTGATTGTGA
BMX	12	CSCE	*CAGGAAACAGCTATGACC*ACTGACTTCCCTCTCCAAAAT	*TGTAAAACGACGGCCAGT*GCTAAAAATTATGATAGGCAAAAG
BMX	13	CSCE	*CAGGAAACAGCTATGACC*ATTACAACATTCTGAACTCTTTGC	*TGTAAAACGACGGCCAGT*CTTCCTGTCTTTCCTCCTTT
BMX	14	CSCE	*TGTAAAACGACGGCCAGT*ACCACATCACTAAGGGTTTTT	*CAGGAAACAGCTATGACC*TACAGTATGTGGACTTTTGCTCTA
BMX	15	CSCE	*TGTAAAACGACGGCCAGT*TCAGTACATGTAGAAAATGGAGTGT	*CAGGAAACAGCTATGACC*ACAAGTGAGCAAAGTAGAAGTGAT
BMX	16	CSCE	*CAGGAAACAGCTATGACC*GAATTCTACTCAAAAGGAGGAAAAT	*TGTAAAACGACGGCCAGT*ACAAATTATGATGGTGTTCAAATGTG
BMX	17	CSCE	*TGTAAAACGACGGCCAGT*ACATTGAAAGGAAACAAAAGTTC	*CAGGAAACAGCTATGACC*TAATACCAAACCCAAGAAAATAAAT
BMX	18	CSCE	*TGTAAAACGACGGCCAGT*ACTCTGACCCATAGCCTTCCTGAC	*CAGGAAACAGCTATGACC*ACGCCTGCTCTGTACCAGTTGCTTTCC
BMX	19	CSCE	*TGTAAAACGACGGCCAGT*GATTATTATAGTGAGCGTCCTTG	*CAGGAAACAGCTATGACC*TAAATTACATTATGCCTTGCTACTT
BTK	10	CSCE	*TGTAAAACGACGGCCAGT*GGGAAAGGGATACAGTGTGCTAT	*CAGGAAACAGCTATGACC*CCGTCTCAGAAAAACAAAAACAAA
BTK	11	CSCE	*TGTAAAACGACGGCCAGT*AGAGCACCACTTCCTCCTACAGACA	*CAGGAAACAGCTATGACC*GGACGGGCACAGCATCAAG
BTK	12	CSCE	*CAGGAAACAGCTATGACC*ACCTTTGTGCCCAAGTTACTGACT	*TGTAAAACGACGGCCAGT*ACCTGCATTGCTTATCCTGGTG
BTK	13	CSCE	*TGTAAAACGACGGCCAGT*TCCTACACCACACCAACAGCAT	*CAGGAAACAGCTATGACC*AAAAGACCCAAGACAACGACAAAG
BTK	14	CSCE	*CAGGAAACAGCTATGACC*CAAAGAATCACACCAAGACTT	*TGTAAAACGACGGCCAGT*AGCAAATAGATTGAGAGTTGAGTT
BTK	15	CSCE	*CAGGAAACAGCTATGACC*TATTATGCAGAGCCAAAAGGAGAA	*TGTAAAACGACGGCCAGT*TGCTACTTCCACCCCATCAG
BTK	16	CSCE	*CAGGAAACAGCTATGACC*ACCCTGGCTTCATTCTACTGG	*TGTAAAACGACGGCCAGT*CATGTTTCATACTGTGCTATTTTT
BTK	17	CSCE	*TGTAAAACGACGGCCAGT*CTAATGCAACAAGTCCTGAATCC	*CAGGAAACAGCTATGACC*CATTGCATTTCTTATCCTTTGAGC
BTK	18	CSCE	*TGTAAAACGACGGCCAGT*TCCTGCTATCCAAAAAGACTG	*CAGGAAACAGCTATGACC*GGCAAGTAGATTCAAGGAAATAAT
BTK	19	CSCE	*TGTAAAACGACGGCCAGT*ATTTAATAAACATTTGCTGCTTAC	*CAGGAAACAGCTATGACC*GGAGAAGAGAAGTAGAACCAAG
EPOR	8/1	CSCE	*TGTAAAACGACGGCCAGT*CACAAGGGTAACTTCCAGGTAGGTG	*CAGGAAACAGCTATGACC*AAAGCAGATGAGCAGGAGGATG
EPOR	8/2	CSCE	*CAGGAAACAGCTATGACC*GGCAGTGTGGACATAGTGGC	*TGTAAAACGACGGCCAGT*CTCTGAGTCATATTGGATCCCTGAT
FLT3	20	CSCE	*TGTAAAACGACGGCCAGT*CACTCCAGGATAATACACATCACAG	*CAGGAAACAGCTATGACC*ATAACGACACAACACAAAATAGCC
hsa-mir 221		CSCE	*TGTAAAACGACGGCCAGT*ATAGCATTTCTGACTGTTGGTT	*CAGGAAACAGCTATGACC*TCTATTCAATGATAAACTCCACTG
hsa-mir 222		CSCE	*CAGGAAACAGCTATGACC*TAAGTAAATAAATCCAGAGGTTGTT	*TGTAAAACGACGGCCAGT*TGTGTAATTCAAGGTAAAGTTTTCA
VHL	1	CSCE	*CAGGAAACAGCTATGACC*TTTGCCCGGGTGGTCTGGAT	*TGTAAAACGACGGCCAGT*GGCGGTAGAGGGGCTTCA
VHL	2	CSCE	*TGTAAAACGACGGCCAGT*GCTCTTTAACAACCTTTGCTT	*CAGGAAACAGCTATGACC*GGCTTAATTTTTCAAGTGGTCTAT
VHL	3	CSCE	*CAGGAAACAGCTATGACC*TCAGGTAGTTGTTGGCAAAGC	*TGTAAAACGACGGCCAGT*TGAAACTAAGGAAGGAACCAGTC

Tag sequences are in Italics and sequences corresponding to the region on interest are in plain type.

A summary of genes and exons that were screened and primers used for the genomic amplification is provided in table 20. As the sex status of some tested samples was not known and some of the genes were on the X-chromosome, the PCR products for all samples were mixed with approximately equal amounts of amplicons from normal control DNA in order to ensure heteroduplex formation if any sequence change was present. Due to the large size of *EPOR* exon 8, this sequence was screened in two fragments.

4.3 Results

4.3.1 Six new nucleotide changes in myeloid and lymphoid leukaemias

In total, approximately 27 kilo bases of sequence, 12 kilo bases of those coding for being translated to a protein, was screened for mutations in the 471 patients. After exclusion of known SNPs, six novel non-synonymous sequence changes were identified in five exons of three different genes. A summary of the changes that were detected is given in table 21. There were four various base changes of *CSF1R* in aMPN patients, several cases with *JAK2*^{V617F} and a single case with lymphoid leukaemia with a change within *TYK2*. Unfortunately constitutional material was not available from any of these cases to determine if the sequence changes were somatically acquired.

Table 21: Summary of non-synonymous changes detected in myeloid and lymphoid malignancies

Gene	Exon	Population	Base change	Amino acid substitution	known SNP
JAK2	14	aMPD	**G-T 1789**	**V617F**	no
CSF1R	3	aMPD	**T-G 118**	**L40V**	no
	4	aMPD	**C-T 316**	**R106W**	no
	17	aMPD	**G-A 2239**	**G747R**	no
	17	aMPD	**G-A 2258**	**R753Q**	no
TYK2	12	lymphoid leukaemias	**G-C 1975**	**A659P**	no

Abbreviations: aMPN - atypical myeloproliferative disorder, SNP - single nucleotide polymorphism.

There were 19 different sequence changes found that were either known SNPs (n=15) or changes that did not affect the amino acid sequence i.e. synonymous changes (n=4) and therefore unlikely to be acquired changes that contribute to disease pathogenesis (table 22).

Table 22: Summary of SNPs and nucleotide changes other than non-synonymous changes

Gene	Exon	Population	Base change	Amino acid substitution	known SNP	SNP genotype
TIE1	18	aMPD	A-G 2973	L991L	rs1199039	T/C
SYK	11	aMPD	C-T 1521	Y507Y	rs2306041	G/A
	11	aMPD	C-T 1539	F513F	rs34231418	C/T
	11	aMPD	T-C 1545	L515L	rs2306040	A/G
FES	17	aMPD	C-T 2127	A709A	rs2227989	C/T
JAK2	6	aMPD	C-T 489	H163H	rs10429491	C/T
	9	aMPD	**C-G 1177**	**L393V**	rs35844880	C/G
	19	aMPD	G-A 2490	L830L	rs2230724	G/A
CSF1R	3	aMPD	T-C 85	P28P	rs216123	A/G
	3	aMPD	C-T 282	S94S	no	no
	5	aMPD	C-T 726	T242T	rs2228422	C/T
	8	aMPD	**A-G 1085**	**H362R**	rs10079250	T/C
	9	aMPD	**G-A 1237**	**G413S**	rs34951517	C/T
	19	aMPD	C-G 2535	L845L	no	no
	21	aMPD	C-T 2709	F903F	no	no
	21	aMPD	**G-C 2760**	**E920D**	rs34030164	C/G
VHL	1	V617F negative PV	**C-T 241**	**P81Y**	rs5030806	C/T
JAK1	13	lymphoid leukaemias	C-T 1977	R659R	rs17127063	G/A

Base changes that affected the amino acid sequence are highlighted in bold. Abbreviations: aMPN - atypical myeloproliferative disorder, SNP - single nucleotide polymorphism.

4.3.1.1 Four different nucleotide changes of *CSF1R*

The receptor proto-oncogene *CSF1R* is encoded by 23 different exons although only 21 of them (i.e. exon 2 to 22) are coding. Only coding exons were screened for abnormalities.

Four different base changes were found in the coding region of *CSF1R*. None of them has been described before. The nucleotide change T to G in codon 40 leads to an amino acid substitution of leucine by valine. The sequencing reaction

showed a heterozygous alteration. The base change was been found in a single patient with unclassified atypical MPN (figure 23).

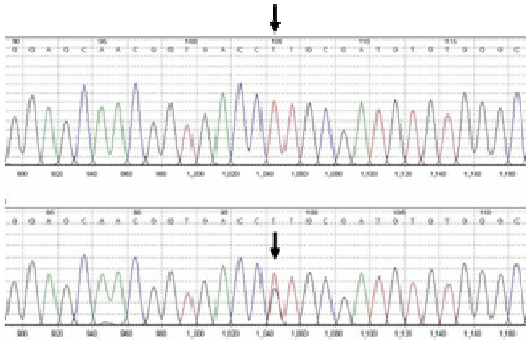


Figure 23: Nucleotide change of *CSF1R* exon 3

The base change T to G leads to the amino acid substitution L40V. The lower trace presents the heterozygous sequence; the wild type sequence is shown above.

A second base change of *CSF1R* is the nucleotide alteration C to T. The event affects the amino acid sequence by changing arginine to tryptophan at position 106 (figure 24).

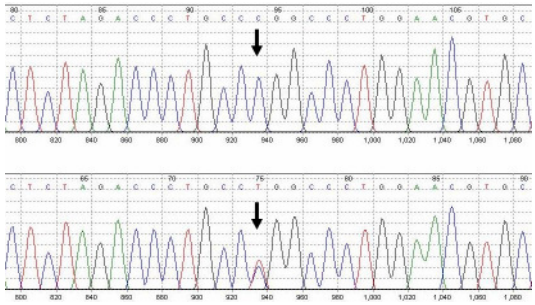
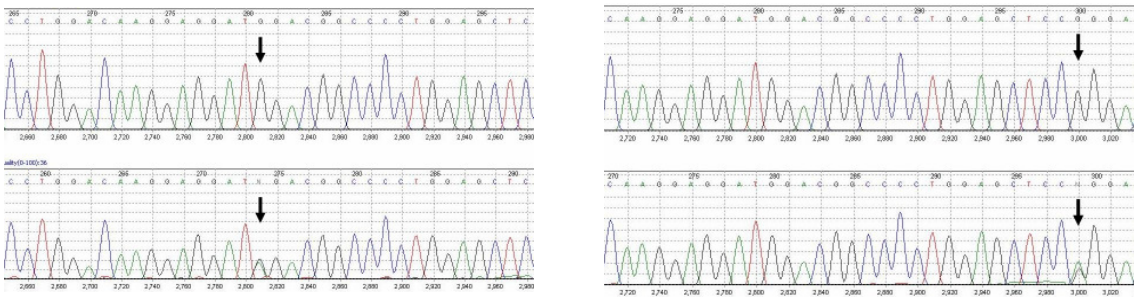


Figure 24: Nucleotide change of *CSF1R* exon 4

The base change C to T was found in a single patient with aMPN. The change affects the amino acid sequence by changing arginine to tryptophan at position 106. The alteration is shown in the lower trace.

A third base change was detected further downstream in *CSF1R* exon 17. The amino acid substitution G747R is caused by the nucleotide change G to A. The change was found in two patient samples, one with the clinical characteristics of systemic mastocytosis, the other one presenting with an unspecified MPN (figure 25 A).

The fourth amino acid substitution was of glutamine for arginine at codon 753 was caused by the base change G to A. The change was found in a sample of a patient who has been diagnosed with hypereosinophilic syndrome (figure 25 B).



A B

Figure 25: Nucleotide change of *CSF1R* exon 17

Two different nucleotide changes were detected in *CSF1R* exon 17. (A) The change G to A leads to the amino acid substitution G747R. The heterozygous alteration is shown in the lower lane. (B) Another change G to A affects the amino acid arginine to glutamine at codon 753. The sequence is shown in the lower lane.

4.3.1.2 Confirmation of the *JAK2*^{V617F} mutation

The point mutation G to T of *JAK2* was detected in 69 of 409 samples (17 %) of patients with myeloid malignancies. An example of detection of this mutation by HRM and subsequent sequence analysis is shown in figure 26 and figure 27. No other mutations in JAK2 exon 14 were detected.

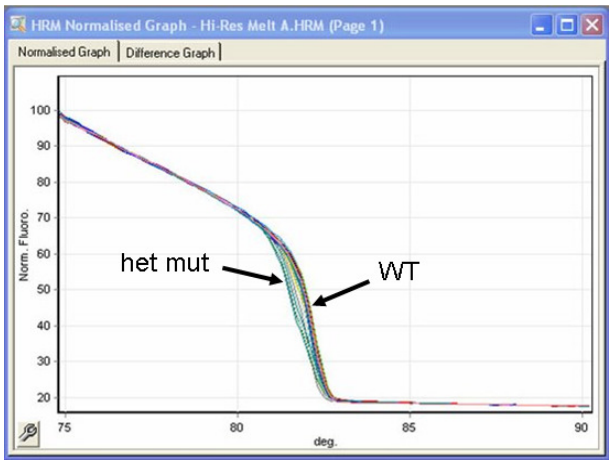


Figure 26: HRM of *JAK2* exon 14

Screen shot of HRM analysis of *JAK2* exon 14 with a homozygous wild type (WT) population and several samples that are heterozygous mutant (het mut) for the G/T nucleotide exchange at position 1789. The mutation causes the amino acid change V617F.

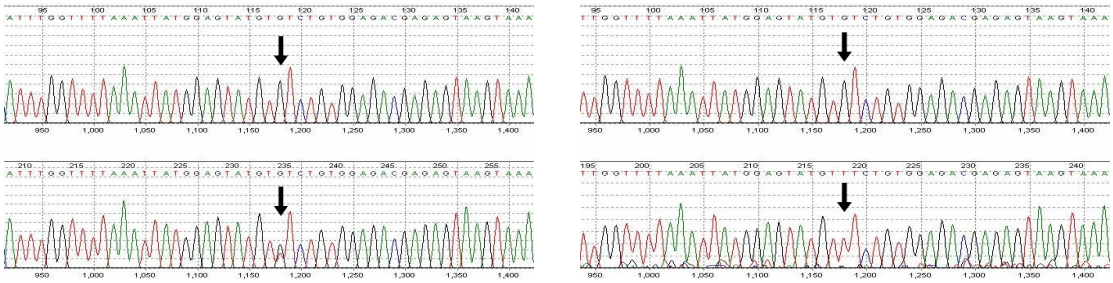


Figure 27: Mutation of JAK2 exon 14

The mutation *JAK2*^{V617F} is caused by the somatic mutation G/T in exon 14. (A) The traces show an example of the heterozygous mutation while in (B) the homozygous nucleotide change is seen.

4.3.1.3 A nucleotide change of *TYK2* in CLL

A heterozygous nucleotide change of G to C was found in *TYK2* exon 12 (figure 28) in a CLL patient sample. The alteration has not been described before. The base change affects the amino acid sequence by substituting proline for alanine at position 659.

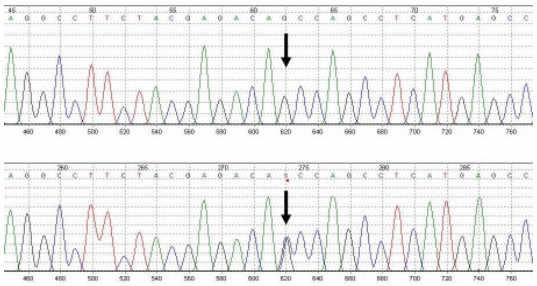


Figure 28: A single base change of *TYK2* exon 12

The heterozygous nucleotide change G to C was detected in a single patient with CLL. The sequence alteration affects the amino acid sequence by a substitution of proline for alanine (A659P).

4.3.1.4 No genomic alterations in *JAK2*^{V617F}-negative PV

No base changes were detected in five genes and two micro RNA sequences when screening the 78 patients with a clinically confirmed diagnosis of PV that were negative for the common *JAK2*^{V617F} point mutation.

4.4 Discussion

In recent years activated tyrosine kinases have emerged as key events in the pathogenesis of leukaemia and especially MPN. Tyrosine kinases can be activated by several mechanisms, i.e. activation by (i) gene amplification, (ii) point mutations, (iii) chromosomal rearrangements creating a chimaeric gene, and (iv) intragenic tandem duplications (ITD) (Kelly and Gilliland 2002). Gene amplification and chromosomal rearrangements would generally be detected by conventional cytogenetic analysis, or more high-resolution techniques such as array comparative genomic hybridisation. Detection of point mutations or ITDs requires more extensive and targeted screening.

4.4.1 Techniques to detect sequence variations

Several techniques have therefore been developed to detect polymorphisms or acquired sequence changes within the genome. Direct DNA sequencing was developed in 1977 and the Sanger dideoxy chain termination methodology has since been greatly improved and become the gold standard for sequence analysis (Sanger*, et al.* 1977). Its main advantage is the accurate detection of any nucleotide alteration in a DNA sequence up to 600-900 base pairs. Despite the technique being widely used, there are limitations due to relatively high costs, and the inability to detect mutations that are present in >20 % of DNA molecules, which may be particularly important for acquired abnormalities where there may be a high or unknown background of normal cells. Nevertheless, high throughput sequencing has been used successfully for large-scale analysis of the kinome (i.e., all kinase genes) in cancer and a diverse range of abnormalities has been found (Greenman*, et al.* 2007). Subsequent to the work I performed as part of this study, new technologies have emerged that are dramatically transforming the capacity of mutation screens. Indeed it is expected that will be possible to resequence an entire human genome for $1,000 within the next year or two (Bayley 2010).

In a relatively small laboratory such as ours, the most cost effective methodology was to use a pre-screening strategy to search for mutations and then standard DNA sequencing to examine any aberrant amplicons. Heteroduplex analysis electrophoretic mobility assays such as conformation-sensitive capillary electrophoresis (CSCE), denaturing high-performance liquid chromatography (DHPLC) and high-resolution melt analysis (HRM) combine both a high throughput platform with a high sensitivity of mutation detection and thus are ideally suited for mutation screens in cancer. Other techniques such as amplification-refractory mutation system (ARMS) and real time PCR can be used sensitive techniques to detect specific, known mutations (Taylor and Taylor 2004).

4.4.2 Sequence variations in haematological malignancies

There are 90 tyrosine kinases in the human genome, the majority of which are large genes consisting of >20 exons. Comprehensive screening of these genes is therefore a very significant undertaking and has only been achieved by major sequencing centres (Greenman*, et al.* 2007). However, there are concerns that sequencing is relatively insensitive and may miss mutations when there is a significant background of non-malignant cells. My analysis therefore used more sensitive CSCE and HRM screens for a limited number of candidate genes that are known to play important roles in haematopoietic development. In addition to activated tyrosine kinases, other genes and two micro RNAs have been screened. Three different populations of patients that were diagnosed with either (i) a disorder of the myeloid cell lineage (n=471), (ii) a lymphoid malignancy (n=145), or (iii) the classical MPN entity polycythaemia vera (PV) but negative for the widely common *JAK2*^{V617F} point mutation (n=78) were investigated for point mutations of several genes. The list of genes embraces nine genes screened for myeloid malignancies (i.e. *CSK*, *FES*, *CSF1R*, *JAK1*, *JAK2*, *JAK3*, *SYK*, *STYK1*, and *TIE1*; table 18), four genes for lymphoid malignancies (i.e. *JAK1*, *JAK2*, *JAK3*, and *TYK2*; table 18), and five genes plus two micro RNAs for *JAK2*^{V617F}-negative PV patients (i.e. *EPOR*, *VHL*, *BMX*, *FLT3*, *BTK*, *hsa-mir-221*, and *hsa-mir-222*; table 20).

4.4.2.1 Myeloid disorders

4.4.2.1.1 Sequence variations of *JAK2*

The *JAK2* gene, a member of the Janus kinase family, has become one of the most interesting genes in myeloproliferative diseases in the last few years. The detection of an acquired uniparental disomy of chromosome 9 in PV patients led several research groups to focus on that particular genomic region (Kralovics*, et al.* 2002). Within a short period, several groups published their data of an acquired somatic gain of function-mutation of *JAK2* that is detectable in various

myeloid disorders, mainly PV, ET and PMF, but not in CML and healthy donors, and rarely in other entities (Baxter, *et al.* 2005, James, *et al.* 2005, Jones, *et al.* 2005, Kralovics, *et al.* 2005, Levine, *et al.* 2005b). The nucleotide exchange G to T in codon 617 leads to the substitution of the amino acid phenylalanine for valine. The mutation results in constitutive tyrosine phosphorylation activity that promotes cytokine hypersensitivity and induces erythrocytosis in a mouse model (James, *et al.* 2005). The highest prevalence of the mutation has been shown in PV patients (65-97 %) with a substantial number of cases with homozygosity (about 25 %). The polymorphism is less frequent in patients with ET and PMF (up to 50 % of cases) (Baxter, *et al.* 2005, James, *et al.* 2005, Jones, *et al.* 2005, Kralovics, *et al.* 2005, Levine, *et al.* 2005b). Homozygosity is rare in the latter populations. Individual cases have been reported for patients with HES (Jones, *et al.* 2005), AML (Lee, *et al.* 2005, Levine, *et al.* 2005a), CMML (Levine, *et al.* 2005a), and JMML (Tono, *et al.* 2005). Subsequently, mutations of *JAK2* upstream of V617F were detected in a few patients that were negative for the more common alteration. These mutations embrace point mutations (i.e. H538QK539L, K539L), deletions (N542-E543del), and insertions (F537-K539delinsL) in the amino acid region between position 537 and position 543 of exon 12 (Scott, *et al.* 2007).

The JAK2 protein is an essential mediator in the signalling cascade of cytokine receptors, e.g. EPOR or GCSFR (Parganas, *et al.* 1998). Saltzman *et al.* demonstrated that JAK2 phosphorylates STAT1, STAT2, STAT3, STAT4, and STAT5 (STAT5A and STAT5B), but not STAT6 (Saltzman, *et al.* 1998). STAT5 is activated in human haematologic malignancies. Their critical role in the pathogenesis of myelo- and lymphoproliferative diseases has been determined for the first time in *TEL/JAK2* translocation studies (Schwaller, *et al.* 2000). Mouse models confirmed the importance of JAK2 in the cytokine signalling mechanism in erythropoiesis (Neubauer, *et al.* 1998).

It is remarkable that the $JAK2^{V617F}$ mutation is so widespread in MPN but other point mutations are extremely rare. The project initially focused on *JAK2* exon 22 in patients with aMPN, since this encodes the catalytic domain, a region which is

a target of activating mutations for other tyrosine kinases such as KIT and FLT3 (Tan, *et al.* 1990, Yamamoto, *et al.* 2001). Subsequently the screen was extended to all 23 coding *JAK2* exons (i.e., exon 3 to exon 25). A total of three different base changes were identified in addition to *JAK2*^{V617F} but all were known SNPs. Subsequent to my study, *JAK2* mutations that predominantly affecting residue R683, were identified in ALL associated with Down syndrome as well as high risk childhood ALL (Bercovich, *et al.* 2008, Mullighan, *et al.* 2009). Unfortunately this study did not include these groups of patients.

4.4.2.1.2 Sequence variations of *CSF1R*

Screening of the *CSF1R* proto-oncogene (alternatively named *FMS* or *c-fms*) showed four new nucleotide changes in *CSF1R* exon 3 (L40V), exon 4 (R106W), and exon 17 (G747R and R753Q, respectively). More base changes were either known polymorphisms or they did not alter the amino acid sequence. *CSF1R* is a very plausible candidate for involvement in myeloid disorders. The gene is located on chromosome 5 (q33.2-q33.3) and encoded the receptor for macrophage colony stimulating factor (M-CSF), a critical cytokine for myelopoiesis (Le Beau, *et al.* 1986). Structurally, CSF1R is highly related to KIT and FLT3, both of which suffer activating mutations in haematological malignancies. Indeed, screening of 110 patients with myelodysplastic syndrome or leukaemia revealed activating mutations within codon 969 of *CSF1R* in 14 (12.7 %) patients, and within codon 301 in 2 (1.8 %) patients, respectively (Ridge, *et al.* 1990). *CSF1R* mutations were prevalent in chronic myelomonocytic leukaemia (20 %) and FAB type M4 acute myeloblastic leukaemia (23 %), both of which are characterised by monocytic differentiation (Ridge, *et al.* 1990). However the high incidence of activating *CSF1R* mutations was not been confirmed in subsequent studies, possibly because the methods employed in the initial analysis were specifically aimed at these mutations and may have detected sequence changes in small subclones.

4.4.2.1.3 Other sequence variations

Five other known polymorphisms were found by screening aMPN patient samples. All nucleotide alterations were synonymous SNPs. The polymorphism A to G (rs1199039) in codon 991 of the protein receptor tyrosine kinase *TIE1* was detected with a frequency of 40 % in the population of patients with aMPN. This is equivalent to the reported allele frequency by the HapMap project (www.hapmap.org). The *TIE1* gene is located in 1p34-p33 and was first mapped in 1992 by Partanen *et al.* (Partanen, *et al.* 1992). Expression of the *TIE1* gene is rare and limited to endothelial cells and some myeloid leukaemia cell lines that present with erythroid and megakaryoblastoid characteristics (Verstovsek, *et al.* 2001, Verstovsek, *et al.* 2002).

The screening of the non-receptor spleen tyrosine kinase gene (*SYK*) exon 11 only revealed known synonymous SNPs. *SYK*^{Y507Y}, *SYK*^{F513F}, and *SYK*^{L515L}. SYK was first described as a protein predominantly expressed in the spleen and thymus (Zioncheck, *et al.* 1988). SYK is a protein-tyrosine kinase that is widely expressed in haematopoietic cells, in epithelial cells, but also in benign and malign breast tissue cells. There it has been determined as a modulator of epithelial cell growth and a potential tumour suppressor (Coopman, *et al.* 2000). As a further functional role, SYK is involved in haematopoietic and endothelial cell growth and survival as well as in the ERK signalling pathway (Inatome, *et al.* 2001). The protein is furthermore important for integrin signalling in neutrophils (Mocsai, *et al.* 2002). Toyabe *et al.* determined that a subpopulation of T cells can express high levels of *SYK* and partially compensate for loss of T-cell functions in patients with deficiency of *ZAP70* (Colucci, *et al.* 2002, Toyabe, *et al.* 2001).

The exon coding for the catalytic domain of the oncogene *FES*, i.e. exon 17, was screened in 425 patients with aMPN. The synonymous SNP C to T (rs2227989) was detected in three patients with PV (n=2) and unclassified MPN (n=1). *FES* is located on chromosome 15 (q26.1) (Dalla-Favera, *et al.* 1982) and encodes a non-receptor kinase with tumour forming potential (Cooper 1982, Sodroski, *et al.* 1984). Animal models have shown that mutations of *FES* result in reduced

numbers of myeloid progenitors and circulating mature myeloid cells and increased circulating erythrocytes (Zirngibl, *et al.* 2002). However it appears that *FES* can also act as a tumour suppressor rather than an oncogene in model systems, although there is no clear evidence as yet to implicate this gene in human malignancies (Sangrar, *et al.* 2005).

4.4.2.2 Lymphoid disorders

A novel base change within *TYK2* was revealed when screening patients with lymphoid malignancies for mutations in Janus kinase family member genes. The alteration of *TYK2* was found in a patient with CLL (*TYK2*^{A659P}) *TYK2* has been mapped to chromosome 19 (p13.2) (Trask, *et al.* 1993) and has been proposed as a key regulator in the control of B-lymphoid tumours as well as playing an important role in transducing IFN signals (Stoiber, *et al.* 2004).

4.4.2.3 *JAK2*^{V617F}-negative PV

The screening of *JAK2*^{V617F}-negative patients for point mutations in seven genes did not reveal any new mutations. The investigation only revealed the known non-synonymous SNP in the *VHL* gene (rs5030806) in a single patient. The von Hippel-Lindau gene (*VHL)* is located on chromosome 3 (p26-p25) (Latif, *et al.* 1993) and is a well-characterised tumour suppressor. Mutations at several sites of *VHL* cause a loss of function that results in the development of a neoplastic syndrome that includes several entities as haemangioblastoma, pheochromocytoma, and renal cell carcinoma (Kanno, *et al.* 1994, Kenck, *et al.* 1996, Oberstrass, *et al.* 1996). Polymorphisms of *VHL* were also detected in a Russian population with polycythaemia and another population with congenital erythrocytosis (Ang, *et al.* 2002, Cario, *et al.* 2005, Pastore, *et al.* 2003).

4.4.3 Conclusion

In summary, there were 24 different changes of the wild type nucleotide sequence detected by screening patients with myeloid malignancies (n=21), lymphoid malignancies (n=2), and *JAK2*^{V617F}-negative PV patients (n=1). Five of the base changes predict new nucleotide substitutions that have not been described before. These sequence changes do not correspond to other known activating tyrosine kinase mutations and currently their significance is unclear. At the time of writing no constitutional material is available from these cases to determine if these sequence changes are inherited or acquired, but hopefully this material will be available for at least some of the cases. If a mutation was found to be inherited it is extremely unlikely that it is relevant pathogenetically.

4.4.4 Functional analyses of nucleotide changes

To definitively determine if mutations are activating or not they need to be analysed using functional assays. For two of the *CSF1R* mutations that have being identified in this study (i.e., R106W and R753Q), this work was undertaken at the WRGL by Dr Andrew Chase. In a first step, these nucleotide changes were introduced into appropriate expression vectors, which were transfected into the murine interleukin 3-dependent pro-B cell line Ba/F3. Activating mutants are expected to result in interleukin-3 independent growth of the transfected cells and increased phosphorylation of downstream signalling such as STAT5 and ERK. Unfortunately neither of the mutations was transforming, indicating that they were probably rare polymorphisms or passenger mutations without pathogenic significance (Chase*, et al.* 2009).

5 A polymorphism associated with *STAT3* expression and response of chronic myeloid leukaemia to interferon *alpha*

5.1 Introduction

5.1.1 The role of Interferon in cancer

In 1957, Isaacs and Lindeman used the term 'interferon' for the first time to describe a recently discovered factor that is involved in viral interference. They recognized an interference effect of heat-inactivated influenza virus and live virus in chicken eggs and named the effect 'The Interferon' (Isaacs and Lindenmann 1957). Many clinical studies have subsequently demonstrated that interferons have potent therapeutic effect in malignant and non-malignant diseases. Antitumor effects were shown for renal cell carcinoma, melanoma, Kaposi's sarcoma, and several hematologic malignancies such as follicular lymphoma and leukaemia (Jonasch and Haluska 2001). Interferon was established as a therapeutic option in CML before tyrosine kinase inhibitors were established in the first decade of the 21st century (Hehlmann*, et al.* 2003, The Italian Cooperative Study Group on Chronic Myeloid Leukaemia 1994, Talpaz, *et al.* 1983) However, the molecular basis of its action still remains obscure.

Interferon's are glycoproteins which are produced and released by leukocytes, fibroblasts or lymphocytes in response to a viral stimulus. It is member of the cytokine family. There are three classes of interferon involved in the human immune response cascade with the division based on the binding of each interferon to a specific cell surface receptor. Interferon-*alpha* (IFN-α) is a type I-interferon, or 'leukocyte-interferon', that is predominantly produced by B-lymphocytes as an innate response to viral or bacterial nucleic acid. Second, there is interferon-*beta* (IFN-β) that is produced by fibroblasts and which is also member of the type I-interferon class. Third, interferon-*gamma* (IFN-γ) is a type

II-interferon that is mainly a product of stimulated T-lymphocytes and also known as the 'immune-interferon'. The specific cell surface interacting molecule of type II interferons is the interferon-*gamma* receptor.

Within these three classes, there are about seven distinct interferons and more than ten subtypes known in humans. Pestka and colleagues published a first description of the purification of interferon in 1978 (Rubinstein, *et al.* 1978). Later on, the interferon genes were cloned for the first time allowing laboratory tests to confirm the functional role of these proteins and their antiviral ability. More members of the interferon family were isolated in the following years.

The *IFNA* and *IFNB* genes are located on the long arm of chromosome 9 (9q) in close proximity to each other (Shows, *et al.* 1982). The cytokines are released after being in contact with viral particles as nucleic acid or unique parts of a bacterial cell. In addition, interferon may also be activated by several mitogens or other cytokines such as interleukins or tumour necrosis factor (Shows, *et al.* 1982).

5.1.2 Interferon downstream signalling

The specific receptor for type I interferon is the type I interferon receptor. The type I interferon receptor is composed of two different chains, the human interferon *alpha*-receptor type 1 (IFNAR1) and type 2 (IFNAR2). There are two different forms of IFNAR2: IFNAR2c and IFNAR2b, which are products of the same gene *IFNAR2* resulting from different RNA splicing. The binding of interferon or another cytokine to the specific cell surface receptor activates an intracellular signalling cascade. The receptor itself does not harbour any enzymatic activity.

The non-receptor tyrosine kinases of the Janus kinase family are critical for downstream signalling. Janus is the name of the two-faced Roman god of

doorways; Janus kinases proteins were named after this god because of their internal structure two nearly identical kinase domains. One of these is the actual kinase domain that transfers phosphate groups from ATP and thus provides the required energy for signalling. The other one has a regulatory function without any enzymatic activity and is termed the 'pseudokinase' or JAK homology (JH2) domain (figure 29). Mutations in this region lead to over proliferation of myeloid cells and have been detected in the majority of patients with polycythaemia vera and less frequently in patients with either essential thrombocytosis or primary myelofibrosis (Baxter, *et al.* 2005, James, *et al.* 2005, Jones, *et al.* 2005, Kralovics, *et al.* 2005, Levine, *et al.* 2005b). Janus kinases are intracellular non-receptor tyrosine kinases. Members of the Janus kinase family enable the transduction of cytokine-mediated signals from the cytokine receptor to downstream proteins as part of a signalling pathway. The principal recipient of that signal within the JAK-STAT pathway is the STAT-complex.

Figure 29: Schematic of the Janus kinase protein with both kinase and pseudokinase domains

The STAT complex is a family of several signal transducer and activator of transcription (STAT) factors that regulates the expression of genes important for an adequate immune response, as well being involved more widely in the control of cellular functions. Several extracellular signalling proteins such as cytokines, growth factors, and hormones may activate these transcription factors. This leads to several physiological cell processes, such as differentiation, proliferation, apoptosis, and angiogenesis (Calo, *et al.* 2003). Aberrant activation, however, may activate different pathological events such as cell transformation and oncogenesis.

The Janus kinase family comprises JAK1, JAK2, JAK3 and TYK2 that bind non-covalently to specific receptors. The cytokine binding leads to Janus kinase apposition and transphosphorylation on tyrosine residues. Thereby, the intrinsic catalytic activity is released (Levy and Darnell 2002). In addition, phosphotyrosine binding sites for Src-homology-2 (SH2) domains of STAT proteins are created that enable STAT proteins to bind to the JAK-receptor complex and become activated. The STAT family comprises seven members: STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, and STAT6 (Darnell 1997). STAT2, STAT4, and STAT6 are activated specifically by a small subset of cytokines, i.e. IFN-*alpha*, IL-6, IL-12, and IL-13, respectively. In contrast, STAT1, STAT3, STA5A, and STA5B can be activated not only by cytokines, but also by growth factors, i.e. EGF, PDGF, insulin, IGF-1, and others. All STATs share conserved domains that play important roles in protein-protein interaction, DNA binding, and transactivation (schematic of STAT protein in figure 30). STAT1, STAT3, STAT4, STAT5A, and STAT5B form homodimers. Heterodimers, however, are formed either by STAT1 and STAT2, or STAT1 and STAT3, respectively, depending on the upstream activation signal and cell type. Activated dimers migrate to the cell nucleus (figure 31) where they act as transcription factor by interacting with promotor regions of their target genes and/or other transcription factors (Brivanlou and Darnell 2002, Levy and Darnell 2002, O'Shea*, et al.* 2002).

Figure 30: Schematic of the STAT protein

Abbreviations: amino group (NH_2), carboxy group (COOH), Src-homology-2 domain (SH2)

Regulation of the JAK-STAT pathway occurs at several levels. Phosphatases oppose phosphorylation and lead to inactivation of kinases. Suppressors of cytokine signalling (SOCS) proteins inhibit phosphorylation of JAKs and STATs by competing with phosphotyrosine binding sites. In the nucleus, protein

inhibitors of activated STATs (PIAS) interact with STATs and prevent transcriptional factors from being activated (Krebs and Hilton 2001, Shuai 2006).

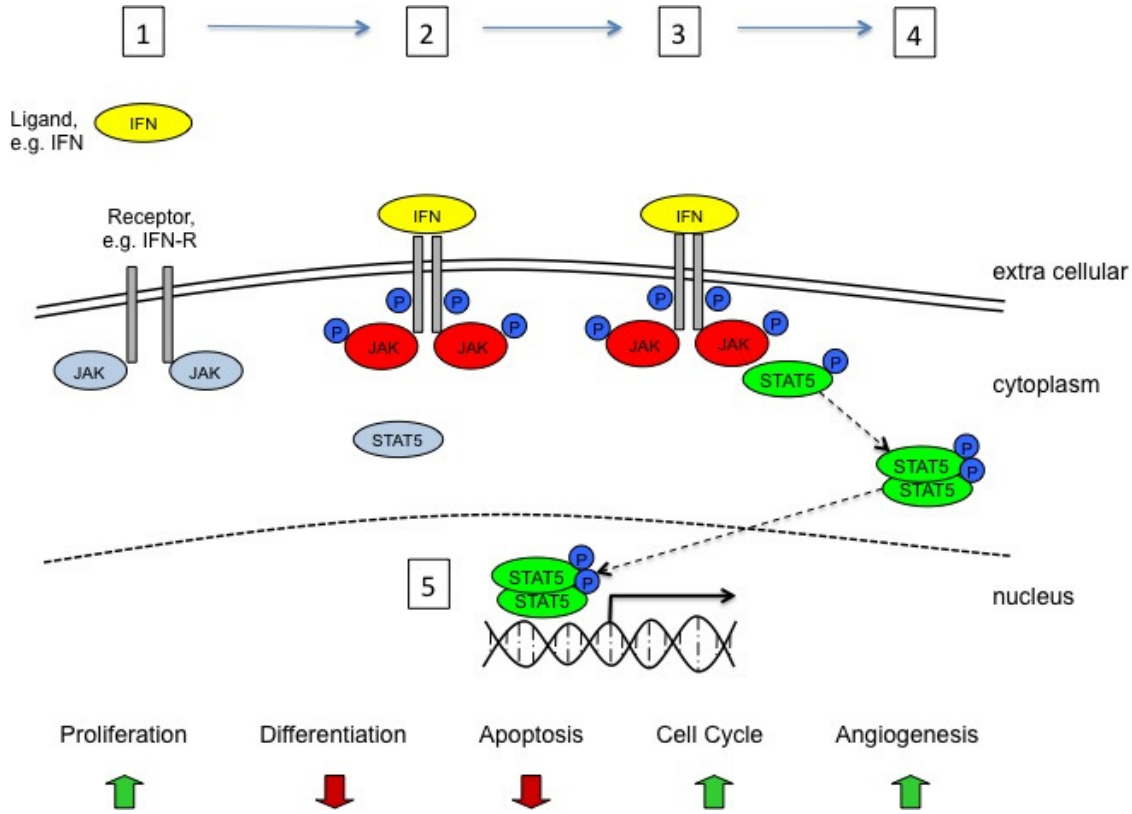


Figure 31: Schematic of the JAK-STAT signalling cascade

Janus kinases become activated through autophosphorylation after binding of a ligand (e.g., IFN) to its specific cell surface receptor, e.g., IFN-R (1,2). STAT molecules are activated by phosphorylation (3). Dimerised STAT complexes (3) migrate to the nucleus and act as transcription factors. Thereby, regulation of several functional cell processes occurs.

Abbreviations: interferon (IFN), interferon receptor (IFN-R), Janus kinase (JAK), phosphate (P), signal transducer and activator of transcription 5 (STAT5).

5.1.3 The role of STATs in tumourigenesis

Tumourigenesis describes the process of the development of a neoplasia. Activation of STAT proteins has been observed in several human malignancies including leukaemia and lymphoma, but also solid tumours. Studies in human tumour cell lines showed that three principal activated STAT proteins are strongly associated with tumour development and tumour progression, i.e.

STAT1, STAT3, and STAT5. They contribute to tumourigenesis through their intimate connection to growth factor signalling, apoptosis, and angiogenesis. This is mainly due to upregulation of genes encoding (i) inhibitors of apoptosis (e.g., Mcl-1 or Bcl-x), (ii) cell cycle regulators (e.g., cyclins D1/D2 or c-Myc), and (iii) inducers of angiogenesis (e.g., VEGF) (Calo*, et al.* 2003, Silva 2004).

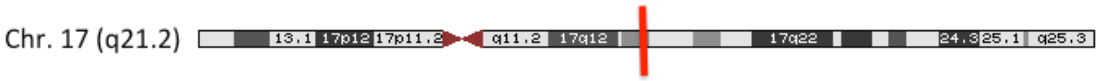
STAT1 is essential for the host's tumour cell control. It is considered to be associated with apoptosis and growth arrest. In mammary cells, an association between STAT1 activation and BRCA1 has been described. BRCA1 function often is lost in familial and other forms of breast cancer (Ouchi*, et al.* 2000).

STAT3 is required to maintain a transformed phenotype. Work with a spontaneously dimerising mutant form of STAT3, known as STAT3-C, demonstrated that tyrosine phosphorylation *per se* was not required for fibroblast transformation (Bromberg*, et al.* 1999). A naturally occurring mutant, however, has not been described so far but constitutive activation of STAT3 has been described in cases of Hodgkin's lymphoma, primary prostate cancer and large granular lymphocyte leukaemia, suggesting abnormalities further upstream in the signalling cascade (Epling-Burnette*, et al.* 2001, Gao*, et al.* 2001, Kube*, et al.* 2001). STAT3 is also activated in many MPN although a direct connection with the presence of *JAK2* mutations has not been made (Steimle*, et al.* 2007).

STAT5A/B is constitutively activated mainly in haematological neoplasia such as leukaemia or lymphoma. That has been demonstrated in detail for acute myeloid leukaemia, acute lymphoblastic leukaemia and erythroleukaemia (Bromberg 2002, Schwaller*, et al.* 2000)

The genes coding for *STAT3*, *STAT5A* and *STAT5B* are located in close proximity to each other on the long arm of chromosome 17 at location q21.2. A schematic is shown in figure 32.

A



B

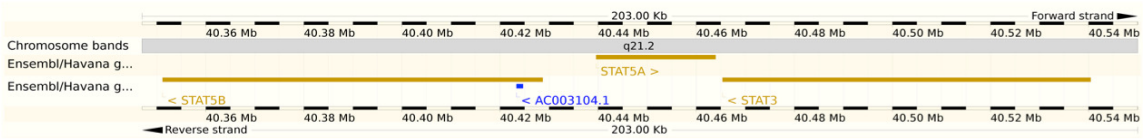


Figure 32: Schematic of *STAT* gene locus

The localisation of the three *STAT* genes, *STAT3*, *STAT5a* and *STAT5b*, in close proximity to each other on the long arm of chromosome 17 at position q21.2 is shown.

Abbreviation: signal transducer and activator of transcription (STAT)

5.1.4 Single nucleotide polymorphisms

The occurrence of DNA sequence variation at a single nucleotide between members of a species is called a single nucleotide polymorphism (SNP). There are nearly 9.5 million SNPs known in the 3 billion base pairs comprising human genome (http://www.ncbi.nlm.nih.gov/projects/SNP/). SNPs are found approximately every 100 to 300 base pairs and usually consist of two variants called alleles that occur at a particular frequency in a defined population. Within a population, the lowest allele frequency at a certain SNP is named the minor allele frequency (MAF). These allele frequencies may vary between ethnic groups or geographical regions. Furthermore, DNA sequence variations may also influence the development of disease or the response to drugs or chemicals. The nucleotide variation may lead to a different amino acid sequence when falling in a polypeptide coding region. This leads either to a missense mutation resulting in a different amino acid or a nonsense mutation with stop codon, respectively. A SNP that leads to a nucleotide change resulting in the same polypeptide is named synonymous or silent mutation. The great majority of SNPs, however, fall outside coding regions and are probably of no functional consequence.

Genetic variation is often linked together in groups known as haplotypes that may extend for several tens or even hundreds of kilo bases. Haplotypes occur because new genetic variation arises through mutation on a particular genetic background. This background gradually gets broken up over timescales of many thousands of years by the process of recombination, which occurs during meiosis. In regions of the genome where recombination is high, genetic variation will be broken up and randomly assorted relatively quickly but in regions where recombination is low the ancestral genetic structure will be preserved in the form of large haplotype blocks. The fact that genetic variation is not random in these regions is known as linkage disequilibrium (LD), and regions with low recombination are said to have high LD. Since SNPs in regions of high LD are tightly linked, genetic variation in these regions can usually be captured by genotyping a small number of polymorphisms known as tagSNPs. The analysis of a tagSNP may identify a specific haplotype without the need of genotyping the whole region.

There have been several large-scale efforts to establish SNP maps to catalogue and describe the location of all known SNPs. The aim is to provide all kind of information regarding SNPs. Data are published online by the Cancer Genome Anatomy Project (cgap.nci.nih.gov) or the International HapMap Project (www.hapmap.org), for example. The latter is a collaboration between scientists in Canada, China, Japan, Nigeria, the United Kingdom, and the Unites States of America. It aimed to establish a detailed haplotype map of the human genome to describe common patterns of SNPs seen in human DNA populations. Genetic information of different populations around the world have been collected, compared with each other and published online. The information contains genetic variations and allele frequencies among different ethnic or local populations and the data are free to download from the Internet. Recently, many more SNPs have been identified through the 1000 genomes project (www.1000genomes.org), which aims to exploit next generation sequencing technologies to fully catalogue genetic variation between individuals and populations. This information is important as SNPs may act as markers to

identify genes possibly involved in susceptibility to disease, severity of disease and drug responsiveness.

5.2 Aims of this study

As detailed above, CML patients show a heterogeneous response to interferon *alpha*, with some achieving a significant reduction in the number of Ph-positive metaphases but others showing minimal response. The reason for this variation is unknown, but it was hypothesised that at least part of the variation may be genetic. This study therefore set out to compare genotypes at candidate genes involved in interferon signalling between cases that showed a good response to IFN, compared to cases that showed a poor response.

5.3 Study design and methods

5.3.1 Patient samples

In this project, pre-treatment genomic DNA (gDNA) samples from *BCR-ABL*-positive CML patients receiving IFN as first-line therapy were investigated. Patients were enrolled in three consecutive German CML studies of the German CML study group between 1986 and 1999, i.e. CML studies I, II, and III. Internal Review Boards of the participating institutions approved the studies and informed consent was obtained according to the Declaration of Helsinki (2000).

Patients were selected and analysed according to (i) the availability of DNA, and (ii) their cytogenetic response to IFN therapy. Cytogenetic response was defined corresponding the suggested criteria by the expert panel of the European Leukemia Net (www.leukemia-net.org) and outlined in detail in table 23 (Baccarani*, et al.* 2009).

Table 23: Criteria defined by the European LeukemiaNet (ELN) for cytogenetic response in CML patients (Baccarani*, et al.* 2009)

Cytogenetic response	Philadelphia chromosome-positive
Complete	0 %
Partial	1-35 %
Minor	36-65 %
Minimal	66-95 %
None	>95 %

In total, 174 patients were analysed (table 24). Of these, 79 patients were cytogenetically responsive to IFN-based therapy (complete response, n=37; major response, n=42) whilst 95 patients did not respond (>95 % Philadelphia chromosome-positive metaphases remaining on therapy). Patients who had a partial or minor response were not analysed.

The median age of all patients was 49.9 years (IFN responders: 49.1 years, IFN non-responders: 50.5 years, respectively). 59 % of all patients were male (IFN responders: 62 %, IFN non-responders: 57 %, respectively). The median time on IFN therapy was 26 months (IFN responder: 39 months, IFN non-responder 22 months, respectively). At the date of analysis in May 2007, 99 patients (57 %) were still alive, with the majority of these being in the IFN responder group (n=57). Major subsequent therapy regimens were either imatinib mesylate (n=48) or allogeneic SCT (n=65). Both, Sokal and Hasford clinical prognostic scores were determined for 166 patients. Additional chromosomal abnormalities were observed in 18 % of all patients (IFN responders: 19 %, IFN non-responders: 16 %, respectively). Submicroscopic chromosome 9q-deletions were analysed in all available DNA samples (n=174) and were tested positive in 4 % of all samples (IFN responders: 2 %, IFN non-responders: 6 %, respectively) (Kreil*, et al.* 2007).

Table 24: Summary of clinical parameters

Clinical characteristics [all data given as median (range) or n (%)]	Total	Non-responders	Responders
	n=174	n=95 (55 %)	n=79 (45 %)
Age in years	50 (10-83)	51 (17-83)	49 (10-76)
Sex (% male)	103 (59)	54 (57)	49 (62)
White blood cell count (*10^9/l)	84 (6-475)	97 (19-475)	66 (6-471)
Blasts (% in peripheral blood)	1 (0-15)	1 (0-15)	0 (0-9)
Eosinophils (% in peripheral blood)	2 (0-10)	2 (0-10)	2 (0-9)
Basophils (% in peripheral blood)	3 (0-40)	3 (0-20)	3 (0-40)
Platelets (x 10^9/l)	371 (86-2343)	407 (126-2343)	356 (86-2020)
Haemoglobin (g/dl)	12.3 (6.9-18.8)	12.0 (6.9-18.8)	12.7 (8.0-16.2)
Spleen enlargement (cm)	1 (0-30)	2 (0-30)	0 (0-17)
Sokal risk group (%)			
low	76 (46)	34 (37)	42 (58)
intermediate	54 (33)	34 (37)	20 (27)
high	36 (21)	25 (26)	11 (15)
Hasford risk group (%)			
low	72 (43)	37 (40)	35 (48)
intermediate	79 (48)	44 (47)	35 (48)
high	15 (9)	12 (13)	3 (4)
Additional chromosomal abnormality (% yes)	27 (18)	15 (19)	12 (16)
Chromosome 9q+ deletion (% yes)	7 (4)	2 (2)	5 (6)

5.3.2 Methods

5.3.2.1 SNP genotyping by pyrosequencing

SNPs that are either within, or in close proximity to nine genes involved in the signalling cascade of IFN were analysed in this study, i.e. *IFNAR1*, *IFNAR2*, *JAK1*, *JAK3*, *TYK2*, *STAT1*, *STAT3,* and *STAT5a/b*. SNPs were selected either on the basis of published data that indicated an association with one or more human diseases, or as tagged SNPs with minor allele frequencies (MAF) of more than 0.2 according the International HapMap Project (release 21) as published on their website (www.hapmap.org). In total, 17 SNPs were selected for analysis. A list of SNPs and primers used for amplification of the genomic region of these SNPs as well as the sequencing primers for the pyrosequencing reaction is given in detail in table 8.

Genomic DNA from the two groups of patients was diluted in water to a standard concentration of 25ng per micro litre. In a PCR reaction, DNA was amplified as described above. Running 10 µl of the PCR product on an agarose gel in a gel electrophoresis proved a successful amplification. Pyrosequencing reactions were performed according to the manufacturer's instructions using the PSQ 96 SNP Reagent Kit (Biotage AB) that contained the enzyme and substrate mixture and nucleotides. Assays were performed using the nucleotide dispensation orders shown in table 10. The sample genotype was determined using the Allele Frequency Quantification (AQ) function in the PSQ96MA SNP software version 2.1 (Biotage AB) as detailed in Chapter 2.

5.3.2.2 Gene expression analysis

Total peripheral blood leukocytes from patient samples had been stored in guanidine thiocyanate (GTC) solution at -20°C for several years in Mannheim as part of the German clinical trial sample collection process. GTC solution is

regularly used in molecular biology laboratories to lyse cells for subsequent RNA and DNA extractions.

Reverse transcriptase real-time PCR (RQ-PCR) was performed to quantify *STAT3*, *STAT5a* and *STAT5b* expression relative to GUSB expression and as an internal control for cDNA quality and quantity. STAT3 expression was determined by using the custom designed Perfect Probe Gene Detection Kit (PrimerDesign, Southampton, UK)
(sense primer: 5'-GAAGGAGGCGTCACTTTCAC-3';
antisense primer: 5'-CTGCTGCTTTGTGTATGGTTC-3';
probe 5'FAM-CTCTTACCGCTGATGTCCTTCTCCACCCAGGTAAGAG-DABCYL3').

STAT5a and *STAT5b* expression was determined using off the shelf TaqMan® Gene Expression Assays (Applied Biosystems; STAT5a: catalogue number 4331182; STAT5b: catalogue number 4331182). After demonstrating equal amplification efficiencies for each target, samples were tested in triplicate and mean *STAT* levels were normalised to *GUSB* and compared using the $2^{-\Delta ct}$ method (Livak and Schmittgen 2001).

5.3.3 Statistical analysis

To investigate the distribution of baseline values such as age, sex, haematological parameters, spleen size, additional chromosomal abnormalities, chromosome 9q-deletions, and both Hasford and Sokal risk scores between the two groups (responders and non-responders), univariate tests were performed by using the Mann-Whitney, Fisher's exact test, or $\chi 2$ tests, as appropriate. The possible independent influence of SNP genotype was assessed by multiple Cox regression analysis using SAS version 9.1.3 (SAS Institute Inc., Cary, NC, U.S.A.). Real time PCR results were compared to genotype by Kruskal-Wallis analysis.

5.4 Results

5.4.1 Identification of a single SNP in *STAT5b* associated with response to IFN

Initially, genotyping of 12 SNPs was performed and the allele frequency between responders and non-responders was compared. Only one SNP (rs6503691) in exon 1 of the STAT5b gene showed a significant difference of genotypes between both groups with a minor allele frequency of 0.16 for non-responders and 0.06 for responders (P=0.0066, Odds ratio 0.36, 95 % confidence intervals 0.17-0.76). A pyrogram showing both variants of the variable genotype of SNP rs6503691 is shown in figure 33. The significance of the SNP was marginally lost when a Bonferroni correction for multiple testing was applied (P=0.08), however it is generally recognised that this correction is highly conservative and may lead to genuine associations being discarded. It was decided therefore to investigate the apparent association in more detail.

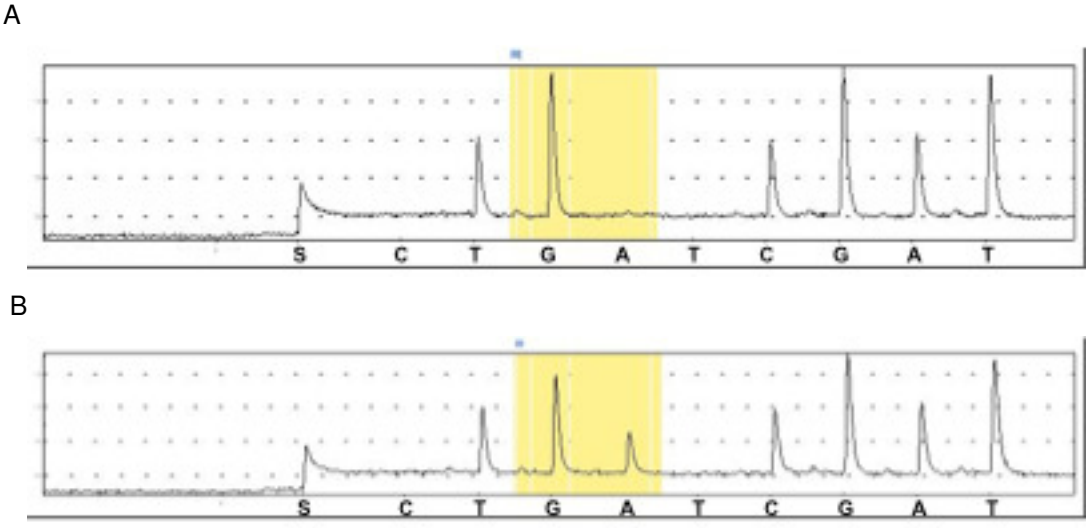


Figure 33: Pyrosequencing pyrogram view of SNP rs6503691

The schematic shows the pyrogram of SNP rs6503691 with the sequence to analyse TG(A/G)CGGATTGTCTTCCCA. The dispensation order of the sequence is CTGATCGAT. Relevant peaks in this schematic are highlighted with a yellow background. Panel A shows an identical genotype in both alleles (i.e. TGGCGGATT), while a different genotype is seen in

schematic B (i.e. TGACGGATT). The dispensation order of the nucleotides includes several reference peaks, where no signal is observed, i.e. C or T (see table 10 for details).

Five further SNPs in close proximity to rs6503691 were subsequently analysed, i.e. rs6503695, rs16967611, rs9900213, rs17500235, and rs17591972. These SNPs were selected according to the HapMap database to capture more that 80 % of the genetic variation in this region. However, all of them failed to reveal any significant variation between responders and non-responders. A summary of all genotyping results is shown in table 24.

The single positively identified SNP rs6503691 is located in a genomic region with strong linkage disequilibrium spanning approximately 80 kb. Inspection of the HapMap data revealed that the area at chromosome 17q21 not only embraces the 5' end of *STAT5b* but also the entire *STAT5a* gene as well as the 3' end of *STAT3*. A schematic taken from the HapMap data file (www.hapmap.org) demonstrates the association and is shown in figure 34.

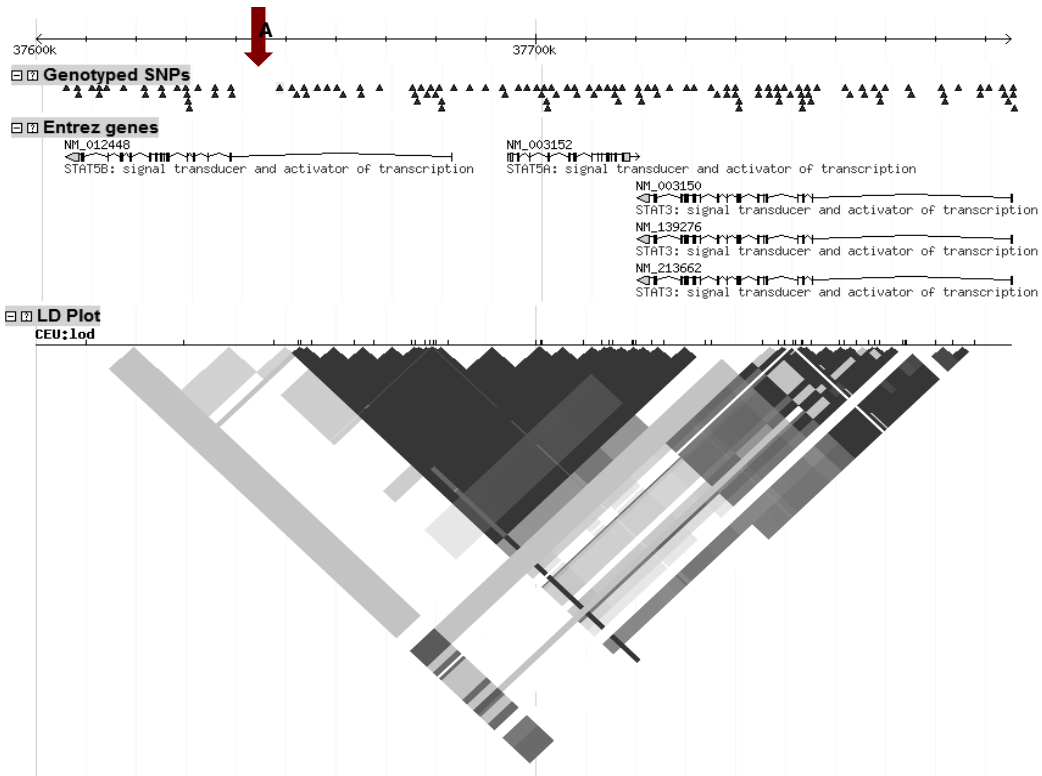


Figure 34: Schematic of high linkage disequilibrium on chromosome 17q21

The SNP rs6503691 is located in an area of high linkage disequilibrium including the genes *STAT3*, *STAT5a* and *STAT5b*. (source: www.hapmap.org). An arrow indicates the location of SNP rs6503961.

The local genetic structure suggest the possibility that rs6503691 might be linked to other variants that influence the expression of *STAT3*, *STAT5a* or *STAT5b*. To test this possibility, the expression level of these genes were analysed by RQ-PCR in 245 pre-treatment samples from *BCR-ABL1*-positive CML patients and the results compared to the genotype of rs6503691. It was found that *STAT3* expression was strongly related to rs6503691 genotype as shown in figure 35 A. There were no significant differences seen for *STAT5a* and *STAT5b*, as shown in figures 35 B and 35 C.

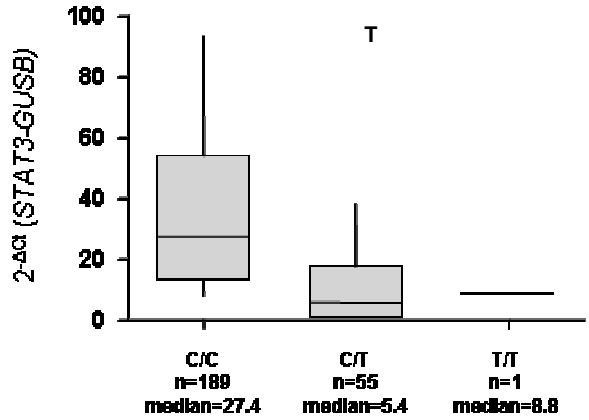
The impact of rs6503691 was evaluated in more detail by taking other prognostic factors into account. On univariate analysis, the leukocyte count, per cent blasts, spleen size, Sokal score, and rs6503691 genotype were all significantly associated with response (table 26). On multivariate analysis, however, rs6503691 genotype fell marginally below the level of significance (P=0.056; table 27).

Table 25: Results of genotyping

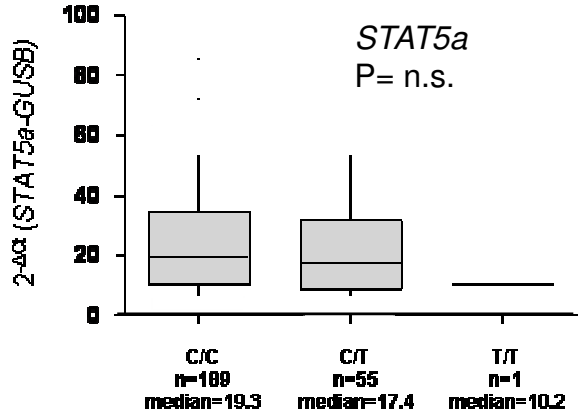
SNP rs6503691 in exon 1 of the *STAT5b* showed a significant difference (Fisher's exact test) in allele frequency between IFN responders and IFN non-responders.

Gene	***IFNAR1***	***IFNAR1***	***IFNAR2***	***JAK1***	***JAK1***	***JAK1***	***STAT1***	***STAT3***	***STAT3***	***STAT5A***	***STAT5B***	***STAT5B***	***STAT5B***	***STAT5B***	***STAT5B***	***TYK2***	***TYK2***
SNP	**rs2850015**	**rs2257167**	**rs7279064**	**rs2991269**	**rs310227**	**rs310229**	**rs1467199**	**rs6503695**	**rs2293152**	**rs2293154**	**rs16967611**	**rs6503691**	**rs9900213**	**rs17500235**	**rs17591972**	**rs2304256**	**rs12720356**
Genotype	**C/T**	**C/G**	**G/T**	**C/T**	**A/G**	**A/G**	**C/G**	**C/T**	**C/G**	**A/G**	**A/G**	**C/T**	**G/T**	**G/T**	**A/G**	**A/C**	**G/T**
IFN responder; n=	73	72	74	73	72	64	79	79	79	79	79	79	79	77	77	74	74
Genotype A/A; n=	42	3	8	6	5	1	66	9	2	4	42	70	59	0	65	4	1
Genotype A/B; n=	26	22	33	19	17	8	7	34	28	14	29	8	18	10	11	28	10
Genotype B/B; n=	5	47	33	48	50	55	6	36	49	61	8	1	2	67	1	42	63
Allele A; n=	110	28	49	31	27	10	139	52	32	22	113	149	136	10	141	36	12
Allele B; n=	36	116	99	115	117	118	19	106	126	136	45	10	22	144	13	112	136
Frequency allele A [%]	75.3	19.4	33.1	21.2	18.8	7.8	88.0	32.9	20.3	13.9	71.5	93.7	86.1	6.5	91.6	24.3	8.1
Frequency allele B [%]	24.7	80.6	66.9	78.8	81.3	92.2	12.0	67.1	79.7	86.1	28.5	6.3	13.9	93.5	8.4	75.7	91.9
IFN non-responder; n=	84	84	84	84	83	80	95	95	95	95	95	95	100	93	83	84	83
Genotype A/A; n=	42	1	6	5	1	2	68	12	0	7	42	70	73	0	62	9	1
Genotype A/B; n=	35	27	43	28	21	13	21	47	42	23	43	20	22	7	19	35	14
Genotype B/B; n=	7	56	35	51	61	65	6	36	53	65	10	5	5	86	2	40	68
Allele A; n=	119	29	55	38	23	17	157	71	42	37	127	160	168	7	143	53	16
Allele B; n=	49	139	113	130	143	143	33	119	148	153	63	30	32	179	23	115	150
Frequency allele A [%]	70.8	17.3	32.7	22.6	13.9	10.6	82.6	37.4	22.1	19.5	66.8	84.2	84.0	3.8	86.1	31.5	9.6
Frequency allele B [%]	29.2	82.7	67.3	77.4	86.1	89.4	17.4	62.6	77.9	80.5	33.2	15.8	16.0	96.2	13.9	68.5	90.4
P-Value	0.38	0.66	1.00	0.79	0.28	0.54	0.18	0.43	0.70	0.20	0.35	0.0066	0.66	0.32	0.16	0.17	0.69
Odds ratio	0.79	0.86	0.98	1.08	0.70	1.40	0.65	1.22	1.12	1.50	0.80	0.36	0.85	0.56	0.57	1.43	1.21
95 % confidence interval	0.48-1.31	0.49-1.54	0.61-1.57	0.63-1.86	0.38-1.28	0.62-3.18	0.35-1.20	0.78-1.90	0.67-1.88	0.84-2.66	0.50-1.27	0.17-0.76	0.47-1.53	0.21-1.52	0.28-1.18	0.87-2.36	0.55-2.65

A



B



C

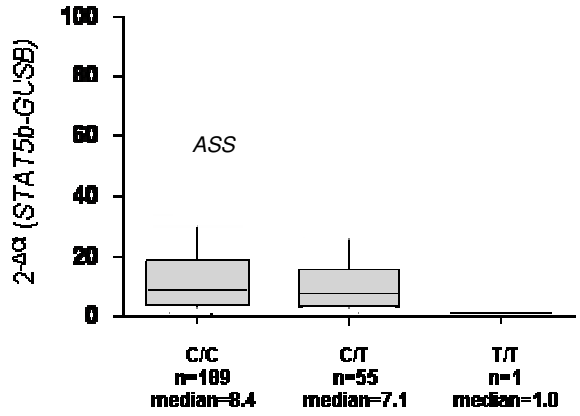


Figure 35: Expression analyses for *STAT3*, *STAT5a* and *STAT5B*.

The expression of *STAT3* (figure 35 A), but not *STAT5A* (figure 35 B) and *STAT5B* (figure 35 C), was associated with a specific SNP rs6503691 genotype. (P=<0.0001; Kruskal-Wallis test). n.s = not significant.

Table 26. Clinical characteristics and univariate analysis of responders and non-responders

Clinical characteristics [data given as median (range) or n (%)]	Total	Non-responders	Responders	P-value*
	n=174	n=95 (55 %)	n=79 (45 %)	
Age in years	50 (10-83)	51 (17-83)	49 (10-76)	n. s.
Sex (% male)	103 (59)	54 (57)	49 (62)	n. s.
White blood cell count (*10^9/l)	84 (6-475)	97 (19-475)	66 (6-471)	0.021
Blasts (% in peripheral blood)	1 (0-15)	1 (0-15)	0 (0-9)	< 0.001
Eosinophils (% in peripheral blood)	2 (0-10)	2 (0-10)	2 (0-9)	n. s.
Basophils (% in peripheral blood)	3 (0-40)	3 (0-20)	3 (0-40)	n. s.
Platelets (*10^9/l)	371 (86-2343)	407 (126-2343)	356 (86-2020)	n. s.
Haemoglobin (g/dl)	12.3 (6.9-18.8)	12.0 (6.9-18.8)	12.7 (8.0-16.2)	0.012
Spleen enlargement (cm below costal margin)	1 (0-30)	2 (0-30)	0 (0-17)	0.009
Sokal risk group (%)				0.033
– low	76 (46)	34 (37)	42 (58)	
– intermediate	54 (33)	34 (37)	20 (27)	
– high	36 (21)	25 (26)	11 (15)	
Hasford risk group (%)				n. s.
– low	72 (43)	37 (40)	35 (48)	
– intermediate	79 (48)	44 (47)	35 (48)	
– high	15 (9)	12 (13)	3 (4)	
Additional chromosome abnormality (% yes)	27 (18)	15 (19)	12 (16)	n. s.
Chromosome 9q+ deletion (% yes)	7 (4)	2 (2)	5 (6)	n. s.
rs6503691 (% genotype A/A)	140 (80)	70 (74)	70 (89)	**0.0066**

*p-value when testing the hypothesis of equal distribution between patients with and without response (Fisher- or Mann-Whitney U test, as appropriate)

Table 27. Multiple logistic regression analysis

Clinical characteristics in final model (n=174)	n (%)	Estimated coefficient	Standard deviation	P-value	Estimated odds ratio	95 % confidence interval lower limit	upper limit
Blasts (% > 0 % in peripheral blood)	101 (58)	-1.246	0.340	< 0.001	0.288	0.148	0.560
Platelets (% > $600*10^9$/l)	41 (24)	-0.924	0.401	0.021	0.397	0.181	0.872
rs6503691 (% genotype A/A)	140 (80)	0.816	0.427	0.056	2.262	0.980	5.221

*cases belonging to the indicated groups were coded by 1, otherwise by 0.

5.5 Discussion

Treatment of CML patients has changed dramatically over the past ten years. The advent of tyrosine kinase inhibitors (TKI) has led to a significant improvement of cytogenetic and molecular response to treatment. Major cytogenetic responses to imatinib in newly diagnosed patients have reached rates of 87 % (O'Brien and Deininger 2003). Major molecular responses after twelve months of therapy were seen in 57-59 % of patients (Hehlmann, *et al.* 2011, Hughes, *et al.* 2003). The implementation of second-generation tyrosine kinase inhibitors has led to significantly higher and faster response rates in newly diagnosed patients, although it is not yet clear if this will ultimately translate into improved survival (Kantarjian, *et al.* 2010, Saglio, *et al.* 2010). Despite these major advances some patients remain refractory to therapy and some relapse after having achieved a good initial response. Combination therapies might overcome the lack of sustained response in these cases and the use of IFN in combination with tyrosine kinase inhibitors was considered as a possibility early in the evolution of targeted therapy (Burchert, *et al.* 2003). Indeed, a favourable outcome of patients treated with IFN plus imatinib has been shown recently in two large randomised trials in France and Germany (Hehlmann, *et al.* 2011, Preudhomme, *et al.* 2010).

Interest has been fuelled further by pre-clinical studies that demonstrated activation of dormant haematopoietic stem cells *in vivo* by the administration of IFN (Essers, *et al.* 2009). It is generally believed that tyrosine kinase inhibitors are ineffective against haematopoietic stem cells (at least in most cases) and thus the combination with IFN might provide the potential to eradicate the disease completely (Nicholson and Holyoake 2009). Indeed, disease eradication is implied by the finding that discontinuation of imatinib in some cases that have achieved complete molecular remission does not always lead to relapse (Mahon, *et al.* 2010). It has been suggested that sustained remission may be achieved by sequential treatment with IFN after discontinuation of a combination therapy of IFN plus imatinib (Burchert, *et al.* 2010). Therefore, variation in response to IFN

therapy still remains of interest, even after ten years of successful treatment with tyrosine kinase inhibitors.

In this study, genomic differences of 17 SNPs were analysed and compared with response to IFN as monotherapy. Since the number of cases available for analysis was limited, it was not possible to do a more extensive genomewide association study (GWAS) since any real association would have been drowned out by background noise: GWAS studies typically require analysis of several hundred or preferably several thousand cases to robustly identify genetic associations. Instead, genes that are involved in the IFN signalling cascade were selected for analysis and, within these, SNPs were selected that (i) had either been previously associated with human disease or (ii) were designed to capture the maximum amount of genetic variation based on HapMap data. This approach has the advantage that the problem of multiple testing is minimised and the focus of the analysis is on biologically plausible genes; the disadvantage is that potentially important associations in other genes would not be detected.

A significant difference at a single SNP, rs6503691, located in the gene coding for *STAT5b* was identified by univariate analysis. The strong linkage disequilibrium in this genomic region of chromosome 17q21 meant that the SNP could be acting as a marker for variation that might affect any of the three genes in the haplotype block. Indeed analysis by RQ-PCR identified that variation in the expression of *STAT3* was significantly linked to rs6503691 genotype. *STAT3* is also a core member of the IFN signalling pathway and STAT3 activation has been described in patients with Hodgkin's lymphoma, primary prostate cancer, and large granular lymphocyte leukaemia (Epling-Burnette*, et al.* 2001, Gao*, et al.* 2001, Kube*, et al.* 2001). It is notable that rs6503691 has been reported to be associated with the risk of development of breast cancer. However, that finding has not yet been replicated, so far (Vaclavicek*, et al.* 2007). Furthermore, a nearby polymorphism has been reported to be linked to both *STAT3* mRNA levels and the response of metastatic renal cell carcinoma to IFN (Ito*, et al.* 2007). BCR-ABL1 is known to activate STAT3 (Coppo*, et al.* 2006) and elevated expression of SOCS3, a known STAT3 target, confers IFN resistance to CML cells

(Sakai*, et al.* 2002). Taken together, these results indicate that polymorphic differences in *STAT3* expression levels may be a determinant of response to IFN in CML, and that the marginal lack of statistical significance on univariate analysis after Bonferroni correction and also on multivariate analysis may have been due at least in part to limited sample numbers.

A weakness of this study was that a second independent series of cases was not available to validate the impact of rs6503691 and therefore further prospective studies are needed. In this context, the impact of rs6503691 and also genetic variation in other immune-regulatory genes will be investigated as a scientific sub-project of the German CML V study (alternative study title TIGER, i.e. 'Tasigna and Interferon *alpha* evaluation conducted by the German CML Study group'), which will compare the outcome of CML patients treated with nilotinib and a combination therapy of nilotinib with interferon.

6 Conclusion

The pathology of a haematological neoplasia is characterised by the uninhibited clonal expansion of an aberrant mutated cell. The aetiology in most entities still is not fully understood. Mutagens such as high dose radiation as well as chemical substances may drive the development of cancer. Other genetic changes may aggravate the neoplastic transformation and result in a more aggressive phenotype (Gilliland*, et al.* 1991). Several molecular pathogenetic mechanisms have been described that are associated with leukaemic transformation in haematological malignancies. The discovery of the typical chromosomal aberration in CML patients, the t(9;22) translocation, as well as the underlying molecular alteration in chronic myeloid leukaemia still is a paradigm for understanding the molecular pathogenesis in haematology (Daley and Baltimore 1988, de Klein*, et al.* 1982, Nowell and Hungerford 1960, Rowley 1973). Several other mechanisms leading to the abnormal expansion of a neoplastic cell clone in haematological malignancies have subsequently been proposed. However, current knowledge does not fully explain the evolution of cancer. Multistep mutational mechanisms in association with epigenetic effects are proposed but are not completely understood. Enhanced mutation rates and impaired genomic stability are believed to be associated with the development of cancer and the key genetic targets of mutations are proto-oncogenes and tumour suppressor genes. Known mechanisms for activation of proto-oncogenes are (i) gene amplification, (ii) point mutation, (iii) chromosomal rearrangements creating a chimaeric gene and (iv) intragenic tandem-duplications. As a consequence of such changes, the normal cell cycle may be interrupted. Impaired regulation of differentiation, proliferation, and apoptosis leads to an expanded cell clone, often with displacement of the normal haematopoietic population. The leukaemic effect may be enhanced by altered interactions between stem/progenitor cells and stromal cells or the extracellular matrix (Bedi*, et al.* 1994, Gordon*, et al.* 1987, Kelly and Gilliland 2002, Puil*, et al.* 1994), or by abnormalities in the stromal cells themselves (Fujisaki*, et al.* 2011).

The prognosis of haematological malignancies is of primary interest for both patients as well as physicians. Until relatively recently the natural course of leukaemia usually led to rapid death and the outcome was seen as an inevitable destiny. Gradually, empirical treatment was supplemented more and more by a rising knowledge of anatomy and pathophysiology. New prognostic indicators were revealed by the implementation of routine diagnostic testing procedures like the simple assessment of blood counts and blood chemistry measurements. The development of chemotherapy and ultimately targeted therapy has dramatically changed the clinical course for many leukaemia patients, and many can now be cured or at least experience a substantial extension to their lifespan. Nevertheless there is still considerable heterogeneity in clinical course for individual patients, even for very well defined entities such as CML. As a result, more and more individualised prognostic scores were developed that are based not only upon clinical parameters but also technical examination results. In CML, for example, the Sokal score is based on standard clinical parameters and defines high, intermediate, and low risk patients (Sokal*, et al.* 1984). A modification of this, generally known as the Hasford or Euro score, was specifically developed for patients treated with IFN (Hasford*, et al.* 1998). The latter scoring system has been adapted recently for patients treated with imatinib and has been called the EUTOS score (Hasford*, et al.* 2011). Despite considerable advances in the understanding of the molecular biology of CML, it is striking that the score is based on simple, 'old-fashioned' haematological parameters, i.e. basophils (in percentage of the whole blood cell count), and spleen size, measured in centimetres below the costal margin (Hasford*, et al.* 2011).

In other leukaemias such as AML, molecular and cytogenetic markers have clearly been shown to provide prognostic value. The impact of chromosomal rearrangements, gene expression, and point mutations have been studied in a large number of patients, and is being further facilitated by the introduction of high-throughput technologies. In myeloid leukaemias, for example, these studies led to a new WHO classification that was introduced first in 2001 (Harris*, et al.* 1999, Vardiman*, et al.* 2002) and last revised in 2008 (Vardiman*, et al.* 2009). For the

first time, cytogenetic and molecular features of haematological malignancies were considered as markers for a general classification of these entities. The prognostic relevance of these alterations led to a new standard tool in diagnosis for many physicians. However, the availability of different and developing therapeutic options mean that any prognostic relevance has to be continually re-evaluated in the light of new treatments. In addition, the finding of more and more molecular abnormalities means that those diseases such as AML can be broken up into smaller and smaller subtypes. This presents a problem because of the very large number of cases that are required in clinical trials to tease out prognostically significant combinations of mutations.

Although CML is often referred to as a paradigm for our understanding of malignancy, BCR-ABL1 is the only known molecular abnormality in chronic phase disease and thus, in contrast to AML, there are no molecular subtypes for which prognosis can be determined. Furthermore the therapeutic options for CML are very different and have changed dramatically over the past decade. Until the 1990s, hydroxyurea and IFN-based regimens remained gold standard for treatment of the majority of CML patients. Allogeneic haematopoietic stem-cell transplantation was the only curative option (Silver, *et al.* 1999) but was only available to the minority of cases who were young enough to withstand the treatment and who had an HLA-matched sibling or unrelated bone marrow donor. The situation fundamentally changed when the tyrosine kinase inhibitor imatinib was introduced first in 1996 (Druker, *et al.* 1996). It took only six years until the superiority of imatinib over IFN was proven (Sawyers, *et al.* 2002) and the era of so-called 'targeted therapies' had begun. Now in 2012, in many countries (but not yet the UK), three different BCR-ABL1 inhibitors have been approved for the first-line treatment of CML patients in first chronic phase, i.e. imatinib (EMA approval in 2001), dasatinib (EMA approval in 2005), and nilotinib (EMA approval in 2007) (Druker, *et al.* 2006, O'Brien, *et al.* 2003).

Since treatment with tyrosine kinase inhibitors is so effective, it has been difficult to evaluate incremental improvements in therapy using the traditionally accepted

endpoint of overall survival. Because of this, the achievement of a major or complete cytogenetic response has been the main surrogate endpoint in clinical trials. The 60-months follow-up analysis of the IRIS trial (i.e., International Randomized IFN vs. STI571 trial for first-line treatment in CML patients) presented a cumulative rate of complete cytogenetic response of 87 %, although this headline figure masks the fact that some patients progressed, changed therapy or dropped out of the study. Nevertheless, only 7 % progressed to an advanced stage of disease after five years treatment (Druker, *et al.* 2006) and most cases remained in cytogenetic remission, thus negating the fears of many pessimists who expected most patients to relapse with drug resistant disease.

Since complete cytogenetic response could now be reached in the majority of patients, molecular markers and molecular response rates came into focus as new surrogate endpoints to distinguish between available treatment options (Hughes, *et al.* 2003). For example, the ENESTnd study ('Evaluating Nilotinib Efficacy and Safety in Clinical Trials of Newly Diagnosed Ph+ CML Patients') used the achievement of a major molecular response (MMR), defined as a *BCR-ABL1* transcript level of 0.1 % or less according the International Scale, as a primary endpoint. After 24 months on treatment, 71 % of nilotinib-treated patients, but only 44 % of imatinib-treated patients reached MMR, leading to the claim that nilotinib is superior to imatinib (Saglio, *et al.* 2010). Similar data have been demonstrated for dasatinib, although the primary endpoint in the Dasision trial ('Dasatinib versus Imatinib Study in Treatment-Naïve CML Patients') was cytogenetic response (Kantarjian, *et al.* 2010). It is not known if these second generation tyrosine kinase inhibitors have a positive influence on survival, although in the ENESTnd trial much has been made of the fact that fewer patients on nilotinib progressed to advanced phase compared to the imatinib arm (2 cases progressed on nilotinib 300 mg versus 11 cases on imatinib; P=0.0095) (Saglio, *et al.* 2010).

The prognostic classification scores by introduced by Sokal and Hasford have been shown to have predictive value for imatinib treated patients, however

achievement of a complete cytogenetic response overcomes the pre-treatment risk scores (Druker, *et al.* 2006). The ELN (European LeukemiaNet) recommendations include the analysis of cytogenetic response as well as molecular response in the evaluation workflow to predict the overall response to imatinib therapy in early chronic phase. The proposals were published first in 2006, and were updated in 2009 (Baccarani, *et al.* 2009, Baccarani, *et al.* 2006). However, latest data suggest that the current recommendations are not applicable for CML patients treated with second-generation tyrosine kinase inhibitors like nilotinib and dasatinib (Jabbour, *et al.* 2011). Early achievement of complete cytogenetic response and deeper molecular responses in these patients demands the definition of new milestones. It has been suggested that a complete cytogenetic response at only 3 months after initiation of therapy should be considered to be as an optimal response, but revised guidelines are unlikely to be published before the end of 2012 at the earliest.

Optimal response to treatment with imatinib according the current ELN recommendations is considered to be the achievement of a complete cytogenetic response at month 12, and MMR at month 18 after start of treatment, respectively (Baccarani, *et al.* 2009). However, complete eradication of the malignant cell clone is rare, even in patients who have undetectable disease as determined by standard real time PCR analysis of peripheral blood leukocytes (Hughes, *et al.* 2003). Despite a sustained clinical response, imatinib-insensitive primitive CD34-positive precursor cells remain detectable by sensitive methods (Graham, *et al.* 2002). Patients who discontinue therapy after complete molecular remission, i.e. *BCR-ABL1* transcript level below detection threshold, generally relapse after a short period of time (Cortes, *et al.* 2004, Mahon, *et al.* 2010). Therefore, continuous imatinib therapy is still strongly recommended to prevent the possibility of relapse and disease progression. On the other hand the effect of a permanent intake of tyrosine kinase inhibitors is unknown and there are continued concerns about long-term safety and tolerability.

To overcome the problem, a major research focus is on targeting the CML stem cell. Allogeneic haematopoietic stem cell transplantation still is the only proven curative therapeutic strategy (Saussele, *et al.* 2010) with the curative effect driven by the graft versus leukaemia effect (GvL) (Kolb, *et al.* 1995), an immunological attack against residual leukaemia cells. The molecular basis for this attack is believed to be principally due to minor histocompatibility antigens arising from a variety of polymorphic differences between the donor and the host; however the possibility that BCR-ABL1 itself might serve as a true leukaemia-specific antigen continues to be a focus of interest. Outside the context of transplantation, IFN, but not imatinib, has been suggested to initiate a BCR-ABL1 specific T-cell response (Burchert, *et al.* 2003). Therefore, the combination of both therapies, IFN and imatinib, represents a promising therapeutic approach. More recently Essers *et al.* demonstrated in a mouse-model that IFN overcomes a dormancy of *BCR-ABL1* positive haematopoietic stem cells. In response to treatment of mice with IFN, HSCs efficiently exit G(0) and enter an active cell cycle. Thereby they become accessible to therapeutic drugs such as tyrosine kinase inhibitors (Essers, *et al.* 2009). The effectiveness of IFN in combination with imatinib has been demonstrated independently *in vivo* in clinical studies, e.g. by Burchert *et al.* where 20 patients treated with imatinib and IFN where followed-up after discontinuation of imatinib and IFN maintenance therapy (Burchert, *et al.* 2010). The majority of patients remained in a sustained remission. Large clinical trials as the French Spirit study ('STI571 Prospective International Randomised Trial'), the Scandinavian study by Simonsson *et al.*, and the German CML IV study confirmed that in a randomised setting (Hehlmann, *et al.* 2011, Preudhomme, *et al.* 2010, Simonsson, *et al.* 2011). Results of these studies suggest a deeper and more durable response in patients treated with the combination therapy compared with monotherapy. A large prospective randomised trial will be initiated later in 2012 by the German CML study group (The German CML-V study) comparing nilotinib monotherapy with a combination of nilotinib and interferon as baseline therapy followed by an IFN maintenance regimen after reaching MMR. The current development in CML clinical research is thus leading to a renaissance in IFN after a decade that was characterised by a wide lack of interest in that therapeutic option.

More widely, IFN has been shown to be effective in other myeloproliferative neoplasms such as polycythaemia vera or myelofibrosis where, in contrast to JAK2 targeted therapy (Verstovsek*, et al.* 2010), the JAK2 mutated clone is often substantially reduced is size (Kiladjian*, et al.* 2006).

The rediscovery of IFN again raises the question 'why do some CML patients respond so well to treatment while others fail'? Looking back, the prognostic value of sub-microscopic deletions on the derivative chromosome 9 [der(9)] in IFN-treated patients has been a subject of controversy for many years. Der(9) deletions are seen in 9-15 % of CML cases and have been associated with an adverse prognosis in several, but not all studies. No study had addressed the issue systematically a randomised controlled clinical trial.

The first part of the work described in this thesis addresses the issue. Ideally such an analysis would be performed prospectively and indeed follow up studies are taking place for some ongoing studies with imatinib. However these studies will take many years to mature and they cannot address the possibility that any prognostic influence might be associated with specific treatment modalities. I aimed to reassess the significance of der(9) deletions using DNA stored as part of CML clinical trials in Germany for which there was a relatively long follow up. Since the largest studies described above suggested an adverse prognosis for IFN-treated patients, I decided to focus solely on patients treated with IFN-based regimens. I sought to (i) determine if deletion status genuinely did predict a poor prognosis for IFN-treated cases and (ii) determine the relationship of deletion status to the Hasford and Sokal risk scores. I developed and validated a novel DNA-based MLPA deletion assay and investigated 339 patients enrolled over 15 years in three trials of the German CML Study Group with an observation time up to 16 years. This is the largest group of IFN-treated cases to be analysed for the impact of deletions. Overall, a similar proportion of deleted cases was found compared to other studies but my series is unusual in that only a relatively small proportion of deletions (36 % of deleted cases; 6 % of all cases) spanned the *ABL1-BCR* breakpoint. The reason for this is unclear. Although no difference in survival

between deleted and non-deleted cases was seen, more detailed analysis indicated that deletions that spanned the breakpoint were a significant indicator of inferior prognosis. Unexpectedly, I found that deletions on one side of the breakpoint only were associated with improved survival, resulting in an overall multivariate P value of 0.007 for deletion status. Improved survival was seen for both deletions on the *BCR* side only and on the *ABL1* side only. Biologically, this is difficult to rationalise but the results are consistent with the hypothesis that at least two separate loci may be targeted by the deletions, one on each side of the fusion junction. However, it is not clear how deletion of both these loci could result in an adverse prognosis, whereas deletion of just one has the opposite effect.

In the other major part of this thesis, the impact of genomic differences of 17 SNPs were analysed and compared with response to IFN as monotherapy. Since the number of cases available for analysis was limited, it was not possible to do a more extensive unbiased genomewide association study (GWAS) since any real association would have been drowned out by background noise: GWAS studies typically require analysis of several hundred or preferably several thousands of cases to robustly identify genetic associations. Instead, genes that are known to be involved in the IFN signalling cascade were selected for analysis and, within these, SNPs were selected that (i) had either been previously associated with human disease or (ii) were designed to capture the maximum amount of genetic variation based on HapMap data. This approach has the advantage that the problem of multiple testing is minimised and the focus of the analysis is on biologically plausible genes; the disadvantage is that potentially important associations in other genes would not be detected.

A significant difference at a single SNP, rs6503691, located in the gene coding for *STAT5b* was identified by univariate analysis. The strong linkage disequilibrium in this genomic region of chromosome 17q21 meant that the SNP could be acting as a marker for variation that might affect any of the three genes in the haplotype block. Indeed analysis by RQ-PCR identified that variation in the expression of *STAT3* was significantly linked to rs6503691 genotype. STAT3 is also a core

member of the IFN signalling pathway and STAT3 activation has been described in patients with Hodgkin's lymphoma, primary prostate cancer, and large granular lymphocyte leukaemia (Epling-Burnette*, et al.* 2001, Gao*, et al.* 2001, Kuno*, et al.* 2001). It is notable that the same SNP has been reported to be associated with the risk of development of breast cancer. However, that has not yet been replicated, so far (Vaclavicek*, et al.* 2007). Furthermore, a nearby polymorphism was reported to be linked to both *STAT3* mRNA levels and the response of metastatic renal cell carcinoma to IFN (Ito*, et al.* 2007). BCR-ABL1 is known to activate STAT3 (Coppo*, et al.* 2006) and elevated expression of SOCS3, a known STAT3 target, confers IFN resistance to CML cells (Sakai*, et al.* 2002). Taken together, my results suggest that polymorphic differences in *STAT3* expression levels may be a determinant of response to IFN in CML, and that the marginal lack of statistical significance on univariate analysis after Bonferroni correction and also on multivariate analysis may have been due at least in part to limited sample numbers.

A weakness of this study was that a second independent series of cases was not available to validate the impact of rs6503691 and therefore further prospective studies are needed. In this context, the impact of rs6503691 and also genetic variation in other immune-regulatory genes will be investigated as a scientific sub-project of the German CML V Study. It is now widely recognised that a critical part of clinical trials should be the collection of DNA samples for evaluation of known or not yet discovered prognostic markers. The true value of these markers will only be discovered by building up large collaborative cohort studies.

Constitutive activation of tyrosine kinase signalling pathways has emerged as a common theme in the pathogenesis of MPN as well as other malignancies. The power of the chimaeric *BCR-ABL1* fusion gene in activating a tyrosine kinase has been demonstrated in detail in CML, and in a wider sense this fusion can be considered as a very powerful biomarker for likely response to imatinib (de Klein*, et al.* 1986). Subsequently, activated signalling pathways have been demonstrated in a wide range of classic and atypical MPN. In particular, the finding of the *JAK2*^{V617F} point mutation indicated that tyrosine kinase activation was widespread

in these disorders. However, in about 50 % of MPN cases, no causative signalling mutations have been found to date. I was interested in identifying new molecular abnormalities in MPN that might be amenable to targeted therapy and as part of this I developed a large mutation screen for abnormalities of specific tyrosine kinases and related signalling proteins. In addition, I explored the hypothesis that other JAK family members may be mutated in lymphoproliferative disorders.

In this part of the study, I aimed to implement a cost effective methodology by using a pre-screening strategy to search for mutations and then standard DNA sequencing to examine any aberrant amplicons. Heteroduplex analysis electrophoretic mobility shift assays such as conformation-sensitive capillary electrophoresis (CSCE), denaturing high-performance liquid chromatography (DHPLC) and high resolution melt analysis (HRM) combine both a high throughput platform with a high sensitivity of mutation detection and thus are ideally suited for mutation screens in cancer. Other techniques such as amplification refractory mutation system (ARMS) and real time PCR can be used sensitive techniques to detect specific, known mutations (Taylor and Taylor 2004).

Now five years after I performed my analysis, the landscape of mutation screening has been completely transformed by next generation sequencing techniques. At the time, however, comprehensive screening of the 90 known tyrosine kinase genes was a very significant undertaking to identify molecular abnormalities and had only been achieved by major sequencing centres. Furthermore, there were concerns that sequencing is relatively insensitive and may miss mutations when there is a significant background of non-malignant cells. This is a particular problem for normal karyotype MPN since the normal background may be substantial and morphological analysis alone is not able to provide an estimate of the clone size.

My analysis therefore used more sensitive CSCE and HRM screens for a limited number of candidate genes that are known to play important roles in haematopoietic development. In addition to activated tyrosine kinases, other genes

and two micro RNAs were screened. Three different populations of patients that were diagnosed with either (i) a disorder of the myeloid cell lineage (n=471), (ii) a lymphoid malignancy (n=145), or (iii) the classical MPN entity polycythaemia vera (PV) but negative for the widely common *JAK2*^{V617F} point mutation (n=78) were investigated for point mutations of several genes. The list of genes embraces nine genes screened for myeloid malignancies (i.e., *CSK*, *FES*, *CSF1R*, *JAK1*, *JAK2*, *JAK3*, *SYK*, *STYK1*, and *TIE1*), four genes for lymphoid malignancies (i.e., *JAK1*, *JAK2*, *JAK3*, and *TYK2*), and five genes plus two micro RNAs for *JAK2*^{V617F}-negative PV patients (i.e., *EPOR*, *VHL*, *BMX*, *FLT3*, *BTK*, *hsa-mir-221*, and *hsa-mir-222*).

In summary, there were 24 different changes of the wild type nucleotide sequence detected by screening patients with myeloid malignancies (n=21), lymphoid malignancies (n=2), and *JAK2*^{V617F}-negative PV patients (n=1). Five of the base changes predict new nucleotide substitutions that have not been described before, i.e., 1975G>C in *TYK2,* and 118T>G, 316C>T, 2239G>A, 2258G>A, all four in *CSF1R.* These sequence changes did not correspond to other known activating tyrosine kinase mutations and their significance was unclear. No constitutional material was available from these cases to determine if these sequence changes were inherited or acquired and functional analysis of two of the changes (316C>T; R106W and 2258G>A; R753Q) demonstrated that they were not transforming. It seems clear therefore that none of the genes I studied are frequently mutated.

Overall possibly the most promising results in this thesis for future investigation is the potential impact of genetic polymorphisms in the IFN signalling pathways in CML. Within the perspective of new therapeutic options for IFN in combination with tyrosine kinase inhibitors, future work will be able to focus on this issue in a prospective manner. In the currently planned German CML V study), which aims to recruit at least 626 patients over a period of five years, a sub-analysis will focus on the impact of the *STAT3*-polymorphism on cytogenetic response in patients treated with IFN in combination with nilotinib. In parallel, patients with nilotinib will be analysed as a control group. As the control group is not exposed to external IFN, the effect of an endogenous IFN stimulus only will be detectable. Further studies

will include the investigation of polymorphisms located in other known genes to be involved in the IFN signalling cascade, i.e. *CTLA4, PTPN22, and PD1*. A polymorphism in the *CTLA4* gene has been associated with a strong disposition to the occurrence of paraneoplastic myasthenia gravis in thymoma patients (Chuang, *et al.* 2005) and a similar effect was seen for a polymorphism in the protein tyrosine phosphatase, non-receptor type 22 (*PTPN22*) gene (Chuang, *et al.* 2009).

The results of this and upcoming work may be translated into clinical practice in future. As T-cell responses seem to play an important role in BCR-ABL1 positive CML patients, the investigation of immune-pathological mechanisms remains an important area for future investigation. In addition to the combination of tyrosine kinase inhibitors with IFN, immunosuppression by BCR-ABL1 peptide vaccination (Bocchia, *et al.* 2010) is under active investigation as a therapeutic option to induce a higher rate of complete molecular remission in CML patients. The possibility that further biomarkers may be identified and validated to help therapeutic decisions in CML remains an important goal that should be realised in the future through large scale carefully planned studies.

References

Abe, A., Emi, N., Tanimoto, M., *et al.* (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.

Abelson, H.T. & Rabstein, L.S. (1970) Lymphosarcoma: virus-induced thymic-independent disease in mice. *Cancer Res.,* **30,** 2213-2222.

Albano, F., Specchia, G., Anelli, L., *et al.* (2003) Genomic deletions on other chromosomes involved in variant t(9;22) chronic myeloid leukemia cases. *Genes Chromosomes Cancer,* **36,** 353-360.

Allan, N.C. & Richards, S.M. & Shepherd, P.C. (1995) UK Medical Research Council randomised, multicentre trial of interferon-alpha n1 for chronic myeloid leukaemia: improved survival irrespective of cytogenetic response. The UK Medical Research Council's Working Parties for Therapeutic Trials in Adult Leukaemia. *Lancet,* **345,** 1392-1397.

Allouche, M., Bourinbaiar, A., Georgoulias, V., *et al.* (1985) T cell lineage involvement in lymphoid blast crisis of chronic myeloid leukemia. *Blood,* **66,** 1155-1161.

Ambros, V. (2004) The functions of animal microRNAs. *Nature,* **431,** 350-355.

Andreasson, P., Johansson, B., Carlsson, M., *et al.* (1997) BCR/ABL-negative chronic myeloid leukemia with ETV6/ABL fusion. *Genes Chromosomes Cancer,* **20,** 299-304.

Anelli, L., Albano, F., Zagaria, A., *et al.* (2004) A chronic myelocytic leukemia case bearing deletions on the three chromosomes involved in a variant t(9;22;11). *Cancer Genet.Cytogenet.,* **148,** 137-140.

Ang, S.O., Chen, H., Gordeuk, V.R., *et al.* (2002) Endemic polycythemia in Russia: mutation in the VHL gene. *Blood Cells Mol.Dis.,* **28,** 57-62.

Association, W.M. (2000) World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.,* **20;284,** 3043-3045.

Baccarani, M., Cortes, J., Pane, F., *et al.* (2009) Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *Journal of Clinical Oncology,* **27,** 6041-6051.

Baccarani, M., Rosti, G., de Vivo, A., *et al.* (2002) A randomized study of interferon-alpha versus interferon-alpha and low-dose arabinosyl cytosine in chronic myeloid leukemia. *Blood,* **99,** 1527-1535.

Baccarani, M., Saglio, G., Goldman, J., *et al.* (2006) Evolving concepts in the management of chronic myeloid leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood,* **108,** 1809-1820.

Bacher, U., Schnittger, S., Kern, W.*, et al.* (2005) The incidence of submicroscopic deletions in reciprocal translocations is similar in acute myeloid leukemia, BCR-ABL positive acute lymphoblastic leukemia, and chronic myeloid leukemia. *Haematologica,* **90,** 558-559.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell,* **116,** 281-297.

Bartram, C.R., de Klein, A., Hagemeijer, A.*, et al.* (1983) Translocation of c-ab1 oncogene correlates with the presence of a Philadelphia chromosome in chronic myelocytic leukaemia. *Nature,* **306,** 277-280.

Bartram, C.R. & Janssen, J.W. & Becher, R. (1987) Persistence of CML despite deletion of rearranged bcr/c-abl sequences. *Haematol.Blood Transfus.,* **31,** 145-148.

Baxter, E.J., Hochhaus, A., Bolufer, P.*, et al.* (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., Scott, L.M., Campbell, P.J.*, et al.* (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet,* **365,** 1054-1061.

Bayley, H. (2010) Nanotechnology: Holes with an edge. *Nature,* **467,** 164-165.

Bedi, A., Zehnbauer, B.A., Barber, J.P.*, et al.* (1994) Inhibition of apoptosis by BCR-ABL in chronic myeloid leukemia. *Blood,* **83,** 2038-2044.

Bellantuono, I. (2004) Haemopoietic stem cells. *Int J Biochem Cell Biol.,* **36,** 607-620.

Belloni, E., Trubia, M., Gasparini, P.*, et al.* (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.

Ben-Neriah, Y., Daley, G.Q., Mes-Masson, A.M.*, et al.* (1986) The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. *Science,* **233,** 212-214.

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.

Bercovich, D., Ganmore, I., Scott, L.M.*, et al.* (2008) Mutations of JAK2 in acute lymphoblastic leukaemias associated with Down's syndrome. *Lancet,* **372,** 1484-1492.

Biernaux, C., Loos, M., Sels, A.*, et al.* (1995) Detection of major bcr-abl gene expression at a very low level in blood cells of some healthy individuals. *Blood,* **86,** 3118-3122.

Bocchia, M., Defina, M., Aprile, L.*, et al.* (2010) Complete molecular response in CML after p210 BCR-ABL1-derived peptide vaccination. *Nature reviews. Clinical oncology,* **7,** 600-603.

Bonifazi, F., de Vivo, A., Rosti, G.*, et al.* (2000) Testing Sokal's and the new prognostic score for chronic myeloid leukaemia treated with alpha-interferon. Italian Cooperative Study Group on Chronic Myeloid Leukaemia. *Br.J Haematol.,* **111,** 587-595.

Bousse-Kerdiles, M.C. & Martyre, M.C. (2001) Involvement of the fibrogenic cytokines, TGF-beta and bFGF, in the pathogenesis of idiopathic myelofibrosis. *Pathol Biol (Paris),* **49,** 153-157.

Brivanlou, A.H. & Darnell, J.E., Jr. (2002) Signal transduction and the control of gene expression. *Science,* **295,** 813-818.

Bromberg, J. (2002) Stat proteins and oncogenesis. *J Clin Invest,* **109,** 1139-1142.

Bromberg, J.F., Wrzeszczynska, M.H., Devgan, G.*, et al.* (1999) Stat3 as an oncogene. *Cell,* **98,** 295-303.

Bunyan, D.J., Eccles, D.M., Sillibourne, J.*, et al.* (2004) Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. *Br.J Cancer.,* **91,** 1155-1159.

Burchert, A., Muller, M.C., Kostrewa, P.*, et al.* (2010) Sustained molecular response with interferon alfa maintenance after induction therapy with imatinib plus interferon alfa in patients with chronic myeloid leukemia. *Journal of Clinical Oncology,* **28,** 1429-1435.

Burchert, A., Wolfl, S., Schmidt, M.*, et al.* (2003) Interferon-alpha , but not the ABL-kinase inhibitor imatinib (STI571), induces expression of myeloblastin and a specific T-cell response in chronic myeloid leukemia. *Blood,* **101,** 259-264.

Calo, V., Migliavacca, M., Bazan, V.*, et al.* (2003) STAT proteins: from normal control of cellular events to tumorigenesis. *Journal of cellular physiology,* **197,** 157-168.

Campbell, P.J., Baxter, E.J., Beer, P.A.*, et al.* (2006a) Mutation of JAK2 in the myeloproliferative disorders: timing, clonality studies, cytogenetic associations, and role in leukemic transformation. *Blood,* **108,** 3548-3555.

Campbell, P.J., Griesshammer, M., Dohner, K.*, et al.* (2006b) V617F mutation in JAK2 is associated with poorer survival in idiopathic myelofibrosis. *Blood,* **107,** 2098-2100.

Campbell, P.J., Scott, L.M., Buck, G.*, et al.* (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.

Cario, H., Schwarz, K., Jorch, N.*, et al.* (2005) Mutations in the von Hippel-Lindau (VHL) tumor suppressor gene and VHL-haplotype analysis in patients with presumable congenital erythrocytosis. *Haematologica,* **90,** 19-24.

Carobbio, A., Finazzi, G., Guerini, V.*, et al.* (2007) Leukocytosis is a risk factor for thrombosis in essential thrombocythemia: interaction with treatment, standard risk factors, and Jak2 mutation status. *Blood,* **109,** 2310-2313.

Carpino, N., Wisniewski, D., Strife, A., *et al.* (1997) p62(dok): a constitutively tyrosine-phosphorylated, GAP-associated protein in chronic myelogenous leukemia progenitor cells. *Cell,* **88,** 197-204.

Chagraoui, H. & Wendling, F. & Vainchenker, W. (2006) Pathogenesis of myelofibrosis with myeloid metaplasia: Insight from mouse models. *Best Practice & Research Clinical Haematology,* **19,** 399-412.

Chase, A., Schultheis, B., Kreil, S., *et al.* (2009) Imatinib sensitivity as a consequence of a CSF1R-Y571D mutation and CSF1/CSF1R signaling abnormalities in the cell line GDM1. *Leukemia.,* **23,** 358-364.

Chronic Myeloid Leukemia Trialists' Collaborative, G. (1997) Interferon alfa versus chemotherapy for chronic myeloid leukemia: a meta-analysis of seven randomized trials. *JNCI Cancer Spectrum,* **89,** 1616-1620.

Chuang, W.Y., Strobel, P., Belharazem, D., *et al.* (2009) The PTPN22gain-of-function+1858T(+) genotypes correlate with low IL-2 expression in thymomas and predispose to myasthenia gravis. *Genes and immunity,* **10,** 667-672.

Chuang, W.Y., Strobel, P., Gold, R., *et al.* (2005) A CTLA4high genotype is associated with myasthenia gravis in thymoma patients. *Annals of neurology,* **58,** 644-648.

Chung, C.C. & Chanock, S.J. (2011) Current status of genome-wide association studies in cancer. *Human genetics,* **130,** 59-78.

Cicchetti, P., Mayer, B.J., Thiel, G., *et al.* (1992) Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science,* **257,** 803-806.

Cohen, N., Rozenfeld-Granot, G., Hardan, I., *et al.* (2001) Subgroup of patients with Philadelphia-positive chronic myelogenous leukemia characterized by a deletion of 9q proximal to ABL gene: expression profiling, resistance to interferon therapy, and poor prognosis. *Cancer Genet.Cytogenet.,* **128,** 114-119.

Colucci, F., Schweighoffer, E., Tomasello, E., *et al.* (2002) Natural cytotoxicity uncoupled from the Syk and ZAP-70 intracellular kinases. *Nat Immunol,* **3,** 288-294.

Consortium, G.P. (2010) A map of human genome variation from population-scale sequencing. *Nature,* **467,** 1061-1073.

Cools, J., DeAngelo, D.J., Gotlib, J., *et al.* (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.

Cooper, G.M. (1982) Cellular transforming genes. *Science,* **217,** 801-806.

Coopman, P.J., Do, M.T., Barth, M., *et al.* (2000) The Syk tyrosine kinase suppresses malignant growth of human breast cancer cells. *Nature,* **406,** 742-747.

Coppo, P., Flamant, S., De, M.V., *et al.* (2006) BCR-ABL activates STAT3 via JAK and MEK pathways in human cells. *Br.J Haematol.,* **134,** 171-179.

Corso, A., Lazzarino, M., Morra, E.*, et al.* (1995) Chronic myelogenous leukemia and exposure to ionizing radiation--a retrospective study of 443 patients. *Ann.Hematol.,* **70,** 79-82.

Cortelazzo, S., Viero, P., Finazzi, G.*, et al.* (1990) Incidence and risk factors for thrombotic complications in a historical cohort of 100 patients with essential thrombocythemia. *Journal of Clinical Oncology,* **8,** 556-562.

Cortes, J. & O'Brien, S. & Kantarjian, H. (2004) Discontinuation of imatinib therapy after achieving a molecular response. *Blood,* **104,** 2204-2205.

Crawford, D.C. & Akey, D.T. & Nickerson, D.A. (2005) The patterns of natural variation in human genes. *Annu.Rev Genomics Hum.Genet.,* **6:287-312.,** 287-312.

Cross, N.C. (2011) Genetic and epigenetic complexity in myeloproliferative neoplasms. *Hematology (Am Soc.Hematol Educ.Program.),* **2011,** 208-214.

Curtis, C.E., Grand, F.H., Musto, P.*, et al.* (2007) Two novel imatinib-responsive PDGFRA fusion genes in chronic eosinophilic leukaemia. *Br.J.Haematol.,* **138,** 77-81.

Dai, Z. & Pendergast, A.M. (1995) Abi-2, a novel SH3-containing protein interacts with the c-Abl tyrosine kinase and modulates c-Abl transforming activity. *Genes Dev.,* **9,** 2569-2582.

Dai, Z., Quackenbush, R.C., Courtney, K.D.*, et al.* (1998) Oncogenic Abl and Src tyrosine kinases elicit the ubiquitin-dependent degradation of target proteins through a Ras-independent pathway. *Genes Dev.,* **12,** 1415-1424.

Daley, G.Q. & Baltimore, D. (1988) Transformation of an interleukin 3-dependent hematopoietic cell line by the chronic myelogenous leukemia-specific P210bcr/abl protein. *Proceedings of the National Academy of Sciences,* **85,** 9312-9316.

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.

Dalla-Favera, R., Franchini, G., Martinotti, S.*, et al.* (1982) Chromosomal assignment of the human homologues of feline sarcoma virus and avian myeloblastosis virus onc genes. *Proceedings of the National Academy of Sciences,* **79,** 4714-4717.

Damashek, W. (1951) Editorial: Some Speculations on the Myeloproliferative Syndromes. *Blood,* **6,** 372-375.

Darnell, J.E., Jr. (1997) STATs and gene regulation. *Science,* **277,** 1630-1635.

de Klein, A., van Agthoven, T., Groffen, C.*, et al.* (1986) Molecular analysis of both translocation products of a Philadelphia-positive CML patient. *Nucleic Acids Res.,* **14,** 7071-7082.

de Klein, A., van Kessel, A.G., Grosveld, G.*, et al.* (1982) A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature,* **300,** 765-767.

de la Fuente, J., Merx, K., Steer, E.J.*, et al.* (2001) ABL-BCR expression does not correlate with deletions on the derivative chromosome 9 or survival in chronic myeloid leukemia. *Blood,* **98,** 2879-2880.

Deininger, M.W., Bose, S., Gora-Tybor, J.*, et al.* (1998) Selective induction of leukemia-associated fusion genes by high-dose ionizing radiation. *Cancer Res.,* **58,** 421-425.

Deininger, M.W.N. & Goldman, J.M. & Melo, J.V. (2000) The molecular biology of chronic myeloid leukemia. *Blood,* **96,** 3343-3356.

Delhommeau, F., Dupont, S., Tonetti, C.*, et al.* (2007) Evidence that the JAK2 G1849T (V617F) mutation occurs in a lymphomyeloid progenitor in polycythemia vera and idiopathic myelofibrosis. *Blood,* **109,** 71-77.

Demiroglu, A., Steer, E.J., Heath, C.*, et al.* (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.

Dewald, G.W. & Wyatt, W.A. & Silver, R.T. (1999) Atypical BCR and ABL D-FISH patterns in chronic myeloid leukemia and their possible role in therapy. *Leukemia and Lymphoma,* **34,** 481-491.

Ding, J., Komatsu, H., Wakita, A.*, et al.* (2004) Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. *Blood,* **103,** 4198-4200.

Douet-Guilbert, N., Morel, F., Quemener, S.*, et al.* (2006) Deletion size characterization of der(9) deletions in Philadelphia-positive chronic myeloid leukemia. *Cancer Genet.Cytogenet.,* **170,** 89-92.

Druker, B.J., Guilhot, F., O'Brien, S.G.*, et al.* (2006) Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *New England Journal of Medicine,* **355,** 2408-2417.

Druker, B.J., Tamura, S., Buchdunger, E.*, et al.* (1996) Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med,* **2,** 561-566.

Epling-Burnette, P.K., Liu, J.H., Catlett-Falcone, R.*, et al.* (2001) Inhibition of STAT3 signaling leads to apoptosis of leukemic large granular lymphocytes and decreased Mcl-1 expression. *J Clin Invest,* **107,** 351-362.

Essers, M.A., Offner, S., Blanco-Bose, W.E.*, et al.* (2009) IFNalpha activates dormant haematopoietic stem cells in vivo. *Nature,* **458,** 904-908.

Fanciulli, M. & Petretto, E. & Aitman, T.J. (2010) Gene copy number variation and common human disease. *Clin.Genet.,* **77,** 201-213.

Felli, N., Fontana, L., Pelosi, E.*, et al.* (2005) MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. *Proceedings of the National Academy of Sciences,* **102,** 18081-18086.

Fialkow, P.J., Faguet, G.B., Jacobson, R.J., *et al.* (1981) Evidence that essential thrombocythemia is a clonal disorder with origin in a multipotent stem cell. *Blood,* **58,** 916-919.

Foroni, L., Wilson, G., Gerrard, G., *et al.* (2011) Guidelines for the measurement of BCR-ABL1 transcripts in chronic myeloid leukaemia. *Br.J.Haematol.,* **153,** 12.

Fourouclas, N., Campbell, P.J., Bench, A.J., *et al.* (2006) Size matters: the prognostic implications of large and small deletions of the derivative 9 chromosome in chronic myeloid leukemia. *Haematologica,* **91,** 952-955.

Fujisaki, J., Wu, J., Carlson, A.L., *et al.* (2011) In vivo imaging of Treg cells providing immune privilege to the haematopoietic stem-cell niche. *Nature,* **474,** 216-219.

Gao, B., Shen, X., Kunos, G., *et al.* (2001) Constitutive activation of JAK-STAT3 signaling by BRCA1 in human prostate cancer cells. *FEBS letters,* **488,** 179-184.

Garsa, A.A. & McLeod, H.L. & Marsh, S. (2005) CYP3A4 and CYP3A5 genotyping by Pyrosequencing. *BMC.Med Genet.,* **6:19.,** 19.

Geary, C.G. (2000) The story of chronic myeloid leukaemia. Historical review. *Br.J.Haematol.,* **110,** 2-11.

Gilliland, D.G. & Blanchard, K.L. & Bunn, H.F. (1991) Clonality in acquired hematologic disorders. *Annu.Rev Med.,* **42,** 491-506.

Goldman, J.M. & Melo, J.V. (2003) Chronic Myeloid Leukemia -- Advances in Biology and New Approaches to Treatment. *New England Journal of Medicine,* **349,** 1451-1464.

Golub, T.R., Barker, G.F., Lovett, M., *et al.* (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., Goga, A., Barker, G.F., *et al.* (1996) Oligomerization of the ABL tyrosine kinase by the Ets protein TEL in human leukemia. *Mol.Cell Biol,* **16,** 4107-4116.

Gonzalez, E., Kulkarni, H., Bolivar, H., *et al.* (2005) The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science,* **307,** 1434-1440.

Gonzalez, F.A., Anguita, E., Mora, A., *et al.* (2001) Deletion of BCR region 3-prime; in chronic myelogenous leukemia. *Cancer Genet.Cytogenet.,* **130,** 68-74.

Gordon, M.Y. & Blackett, N.M. (1998) Reconstruction of the Hematopoietic System After Stem Cell Transplantation. *Cell Transplantation,* **7,** 339-344.

Gordon, M.Y., Dazzi, F., Marley, S.B., *et al.* (1999) Cell biology of CML cells. *Leukemia.,* **13,** S65-S71.

Gordon, M.Y., Dowding, C.R., Riley, G.P., *et al.* (1987) Altered adhesive interactions with marrow stroma of haematopoietic progenitor cells in chronic myeloid leukaemia. *Nature,* **328,** 342-344.

Gorre, M.E., Mohammed, M., Ellwood, K.*, et al.* (2001) Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science,* **293,** 876-880.

Graham, S.M., Jorgensen, H.G., Allan, E.*, et al.* (2002) Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 in vitro. *Blood,* **99,** 319-325.

Grand, E.K., Chase, A.J., Heath, C.*, et al.* (2004a) Targeting FGFR3 in multiple myeloma: inhibition of t(4;14)-positive cells by SU5402 and PD173074. *Leukemia.,* **18,** 962-966.

Grand, E.K., Grand, F.H., Chase, A.J.*, et al.* (2004b) Identification of a novel gene, FGFR1OP2, fused to FGFR1 in 8p11 myeloproliferative syndrome. *Genes Chromosomes Cancer,* **40,** 78-83.

Grand, F., Kulkarni, S., Chase, A.*, et al.* (1999) Frequent deletion of hSNF5/INI1, a component of the SWI/SNF complex, in chronic myeloid leukemia. *Cancer Res.,* **59,** 3870-3874.

Greenman, C., Stephens, P., Smith, R.*, et al.* (2007) Patterns of somatic mutation in human cancer genomes. *Nature,* **446,** 153-158.

Greer, J.P. & Foerster, J. & Lukens, J.N. (2003) *Wintrobe's Clinical Hematology.* Lippincott Williams & Wilkins.

Groffen, J., Stephenson, J.R., Heisterkamp, N.*, et al.* (1984) The human c-abl oncogene in the Philadelphia translocation. *J Cell Physiol Suppl,* **3,** 179-191.

Grosveld, G., Verwoerd, T., van Agthoven, T.*, et al.* (1986) The chronic myelocytic cell line K562 contains a breakpoint in bcr and produces a chimeric bcr/c-abl transcript. *Mol.Cell Biol.,* **6,** 607-616.

Guasch, G., Mack, G.J., Popovici, C.*, et al.* (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.*, et al.* (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.

Guilhot, F., Chastang, C., Michallet, M.*, et al.* (1997) Interferon Alfa-2b Combined with Cytarabine versus Interferon Alone in Chronic Myelogenous Leukemia. *New England Journal of Medicine,* **337,** 223-229.

Haferlach, T., Winkemann, M., Nickenig, C.*, et al.* (1997) Which compartments are involved in Philadelphia-chromosome positive chronic myeloid leukaemia? An answer at the single cell level by combining May-Grunwald-Giemsa staining and fluorescence in situ hybridization techniques. *Br.J Haematol.,* **97,** 99-106.

Hantschel, O., Warsch, W., Eckelhart, E.*, et al.* (2012) BCR-ABL uncouples canonical JAK2-STAT5 signaling in chronic myeloid leukemia. *Nature chemical biology,* **8,** 285-293.

Harris, N.L., Jaffe, E.S., Diebold, J.*, et al.* (1999) World Health Organization Classification of Neoplastic Diseases of the Hematopoietic and Lymphoid Tissues: Report of the Clinical Advisory Committee MeetingAirlie House, Virginia, November 1997. *Journal of Clinical Oncology,* **17,** 3835-3849.

Harrison, C.N. (2002) Current trends in essential thrombocythaemia. *Br.J.Haematol.,* **117,** 796-808.

Harrison, C.N. & Green, A.R. (2006) Essential thrombocythaemia. *Best Practice & Research Clinical Haematology,* **19,** 439-453.

Hasford, J., Baccarani, M., Hoffmann, V.*, et al.* (2011) Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. *Blood,* **118,** 686-692.

Hasford, J., Pfirrmann, M., Hehlmann, R.*, et al.* (1998) A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa. Writing Committee for the Collaborative CML Prognostic Factors Project Group. *J Natl Cancer Inst.,* **90,** 850-858.

Havelange, V. & Garzon, R. (2010) MicroRNAs: emerging key regulators of hematopoiesis. *Am.J Hematol.,* **85,** 935-942.

Hehlmann, R., Berger, U., Pfirrmann, M.*, et al.* (2007) Drug treatment is superior to allografting as first-line therapy in chronic myeloid leukemia. *Blood,* **109,** 4686-4692.

Hehlmann, R., Berger, U., Pfirrmann, M.*, et al.* (2003) Randomized comparison of interferon alpha and hydroxyurea with hydroxyurea monotherapy in chronic myeloid leukemia (CML-study II): prolongation of survival by the combination of interferon alpha and hydroxyurea. *Leukemia.,* **17,** 1529-1537.

Hehlmann, R., Heimpel, H., Hasford, J.*, et al.* (1993) Randomized comparison of busulfan and hydroxyurea in chronic myelogenous leukemia: prolongation of survival by hydroxyurea. The German CML Study Group. *Blood,* **82,** 398-407.

Hehlmann, R., Heimpel, H., Hasford, J.*, et al.* (1994) Randomized comparison of interferon-alpha with busulfan and hydroxyurea in chronic myelogenous leukemia. The German CML Study Group. *Blood,* **84,** 4064-4077.

Hehlmann, R., Lauseker, M., Jung-Munkwitz, S.*, et al.* (2011) Tolerability-adapted imatinib 800 mg/d versus 400 mg/d versus 400 mg/d plus interferon-alpha in newly diagnosed chronic myeloid leukemia. *Journal of Clinical Oncology,* **29,** 1634-1642.

Heinrich, P.C., Behrmann, I., Muller-Newen, G.*, et al.* (1998) Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem.J,* **334 (Pt 2),** 297-314.

Heisterkamp, N., Stam, K., Groffen, J.*, et al.* (1985) Structural organization of the bcr gene and its role in the Ph' translocation. *Nature,* **315,** 758-761.

Herens, C., Tassin, F., Lemaire, V., *et al.* (2000) Deletion of the 5'-ABL region: a recurrent anomaly detected by fluorescence in situ hybridization in about 10% of Philadelphia-positive chronic myeloid leukaemia patients. *Br.J.Haematol.,* **110,** 214-216.

Hinds, D.A., Stuve, L.L., Nilsen, G.B., *et al.* (2005) Whole-genome patterns of common DNA variation in three human populations. *Science,* **307,** 1072-1079.

Hirota, S., Isozaki, K., Moriyama, Y., *et al.* (1998) Gain-of-Function Mutations of c-kit in Human Gastrointestinal Stromal Tumors. *Science,* **279,** 577-580.

Hochhaus, A., Kreil, S., Corbin, A., *et al.* (2001) Roots of clinical resistance to STI-571 cancer therapy. *Science,* **293,** 2163.

Hochhaus, A., Kreil, S., Corbin, A.S., *et al.* (2002) Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy. *Leukemia.,* **16,** 2190-2196.

Hochhaus, A., Reiter, A., Skladny, H., *et al.* (1996) A novel BCR-ABL fusion gene (e6a2) in a patient with Philadelphia chromosome-negative chronic myelogenous leukemia. *Blood,* **88,** 2236-2240.

Horita, M., Andreu, E.J., Benito, A., *et al.* (2000) Blockade of the Bcr-Abl kinase activity induces apoptosis of chronic myelogenous leukemia cells by suppressing signal transducer and activator of transcription 5-dependent expression of Bcl-xL. *J Exp Med,* **191,** 977-984.

Hosking, F.J. & Dobbins, S.E. & Houlston, R.S. (2011) Genome-wide association studies for detecting cancer susceptibility. *British medical bulletin,* **97,** 27-46.

Huang, J. & Mo, Y.Y. (2012) Role of microRNAs in leukemia stem cells. *Frontiers in bioscience,* **4,** 799-809.

Hughes, T., Deininger, M., Hochhaus, A., *et al.* (2006) Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood,* **108,** 28-37.

Hughes, T.P., Kaeda, J., Branford, S., *et al.* (2003) Frequency of major molecular responses to imatinib or interferon alfa plus cytarabine in newly diagnosed chronic myeloid leukemia. *New England Journal of Medicine,* **349,** 1423-1432.

Human Genome Sequencing, C. (2004) Finishing the euchromatic sequence of the human genome. *Nature,* **431,** 931-945.

Huntly, B.J.P. & Bench, A. & Green, A.R. (2003a) Double jeopardy from a single translocation: deletions of the derivative chromosome 9 in chronic myeloid leukemia. *Blood,* **102,** 1160-1168.

Huntly, B.J.P., Bench, A.J., Delabesse, E., *et al.* (2002) Derivative chromosome 9 deletions in chronic myeloid leukemia: poor prognosis is not associated with loss of ABL-BCR expression, elevated BCR-ABL levels, or karyotypic instability. *Blood,* **99,** 4547-4553.

Huntly, B.J.P., Guilhot, F., Reid, A.G.*, et al.* (2003b) Imatinib improves but may not fully reverse the poor prognosis of patients with CML with derivative chromosome 9 deletions. *Blood,* **102,** 2205-2212.

Huntly, B.J.P., Reid, A.G., Bench, A.J.*, et al.* (2001) Deletions of the derivative chromosome 9 occur at the time of the Philadelphia translocation and provide a powerful and independent prognostic indicator in chronic myeloid leukemia. *Blood,* **98,** 1732-1738.

Ichimaru, M., Tomonaga, M., Amenomori, T.*, et al.* (1991) Atomic bomb and leukemia. *J Radiat.Res (Tokyo),* **32 Suppl,** 162-167.

Ilaria, R.L., Jr. & Van Etten, R.A. (1996) P210 and P190(BCR/ABL) induce the tyrosine phosphorylation and DNA binding activity of multiple specific STAT family members. *J Biol Chem,* **271,** 31704-31710.

Inatome, R., Yanagi, S., Takano, T.*, et al.* (2001) A critical role for Syk in endothelial cell proliferation and migration. *Biochem.Biophys.Res Commun.,* **286,** 195-199.

Isaacs, A. & Lindenmann, J. (1957) Virus interference. I. The interferon. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society,* **147,** 258-267.

Ishii, T., Bruno, E., Hoffman, R.*, et al.* (2006) Involvement of various hematopoietic-cell lineages by the JAK2V617F mutation in polycythemia vera. *Blood,* **108,** 3128-3134.

Ito, N., Eto, M., Nakamura, E.*, et al.* (2007) STAT3 Polymorphism Predicts Interferon-Alfa Response in Patients With Metastatic Renal Cell Carcinoma. *Journal of Clinical Oncology,* **25,** 2785-2791.

Jabbour, E., Kantarjian, H.M., O'Brien, S.*, et al.* (2011) Front-line therapy with second-generation tyrosine kinase inhibitors in patients with early chronic phase chronic myeloid leukemia: what is the optimal response? *Journal of Clinical Oncology,* **29,** 4260-4265.

James, C., Ugo, V., Le Couedic, J.P.*, et al.* (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature,* **434,** 1144-1148.

Jiang, Y. & Zhao, R.C.H. & Verfaillie, C.M. (2000) Abnormal integrin-mediated regulation of chronic myelogenous leukemia CD34+ cell proliferation: BCR/ABL up-regulates the cyclin-dependent kinase inhibitor, p27Kip, which is relocated to the cell cytoplasm and incapable of regulating cdk2 activity. *Proceedings of the National Academy of Sciences,* **97,** 10538-10543.

Jonasch, E. & Haluska, F.G. (2001) Interferon in oncological practice: review of interferon biology, clinical applications, and toxicities. *The oncologist,* **6,** 34-55.

Jones, A.V., Kreil, S., Zoi, K.*, et al.* (2005) Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. *Blood,* **106,** 2162-2168.

Jones, A.V., Silver, R.T., Waghorn, K.*, et al.* (2006) Minimal molecular response in polycythemia vera patients treated with imatinib or interferon alpha. *Blood,* **107,** 3339-3341.

Kanno, H., Kondo, K., Ito, S.*, et al.* (1994) Somatic mutations of the von Hippel-Lindau tumor suppressor gene in sporadic central nervous system hemangioblastomas. *Cancer Res.,* **54,** 4845-4847.

Kantarjian, H., Shah, N.P., Hochhaus, A.*, et al.* (2010) Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. *New England Journal of Medicine,* **362,** 2260-2270.

Kaushansky, K. (2006a) Hematopoietic growth factors, signaling and the chronic myeloproliferative disorders. *Cytokine & Growth Factor Reviews,* **17,** 8.

Kaushansky, K. (2006b) Lineage-Specific Hematopoietic Growth Factors. *New England Journal of Medicine,* **354,** 2034-2045.

Kelly, L.M. & Gilliland, D.G. (2002) Genetics of Myeloid Leukemias. *Annu.Rev Genomics Hum.Genet.,* **3,** 179-198.

Kenck, C., Wilhelm, M., Bugert, P.*, et al.* (1996) Mutation of the VHL gene is associated exclusively with the development of non-papillary renal cell carcinomas. *J Pathol,* **179,** 157-161.

Khorashad, J.S., Anand, M., Marin, D.*, et al.* (2006) The presence of a BCR-ABL mutant allele in CML does not always explain clinical resistance to imatinib. *Leukemia.,* **20,** 658-663.

Kiladjian, J.J., Cassinat, B., Turlure, P.*, et al.* (2006) High molecular response rate of polycythemia vera patients treated with pegylated interferon {alpha}-2a. *Blood,* **108,** 2037-2040.

Kim, Y.R., Cho, H.I., Yoon, S.S.*, et al.* (2005) Interpretation of submicroscopic deletions of the BCR or ABL gene should not depend on extra signal-FISH: problems in interpretation of submicroscopic deletion of the BCR or ABL gene with extra signal-FISH. *Genes Chromosomes Cancer,* **43,** 37-44.

Kluin-Nelemans, H.C., Buck, G., le Cessie, S.*, et al.* (2004) Randomized comparison of low-dose versus high-dose interferon-alfa in chronic myeloid leukemia: prospective collaboration of 3 joint trials by the MRC and HOVON groups. *Blood,* **103,** 4408-4415.

Kolb, H.J., Schattenberg, A., Goldman, J.M.*, et al.* (1995) Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients. *Blood,* **86,** 2041-2050.

Kolomietz, E., Al Maghrabi, J., Brennan, S.*, et al.* (2001) Primary chromosomal rearrangements of leukemia are frequently accompanied by extensive submicroscopic deletions and may lead to altered prognosis. *Blood,* **97,** 3581-3588.

Kolomietz, E., Marrano, P., Yee, K.*, et al.* (2003) Quantitative PCR identifies a minimal deleted region of 120 kb extending from the Philadelphia chromosome ABL translocation breakpoint in chronic myeloid leukemia with poor outcome. *Leukemia.,* **17,** 1313-1323.

Komura, D., Shen, F., Ishikawa, S.*, et al.* (2006) Genome-wide detection of human copy number variations using high-density DNA oligonucleotide arrays. *Genome Research,* **16,** 1575-1584.

Kozubek, S., Ryznar, L., Kozubek, M.*, et al.* (1997) Distribution of ABL and BCR Genes in Cell Nuclei of Normal and Irradiated Lymphocytes. *Blood,* **89,** 4537-4545.

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.*, et al.* (2005) A Gain-of-Function Mutation of JAK2 in Myeloproliferative Disorders. *New England Journal of Medicine,* **352,** 1779-1790.

Kralovics, R., Teo, S.S., Li, S.*, et al.* (2006) Acquisition of the V617F mutation of JAK2 is a late genetic event in a subset of patients with myeloproliferative disorders. *Blood,* **108,** 1377-1380.

Krebs, D.L. & Hilton, D.J. (2001) SOCS proteins: negative regulators of cytokine signaling. *Stem Cells,* **19,** 378-387.

Kreil, S., Pfirrmann, M., Haferlach, C.*, et al.* (2007) Heterogeneous prognostic impact of derivative chromosome 9 deletions in chronic myelogenous leukemia. *Blood,* **110,** 1283-1290.

Kube, D., Holtick, U., Vockerodt, M.*, et al.* (2001) STAT3 is constitutively activated in Hodgkin cell lines. *Blood,* **98,** 762-770.

Kulkarni, S., Heath, C., Parker, S.*, et al.* (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.*, et al.* (2001) Constitutive kinase activation of the TEL-Syk fusion gene in myelodysplastic syndrome with t(9;12)(q22;p12). *Blood,* **97,** 1050-1055.

Kutti, J. & Ridell, B. (2001) Epidemiology of the myeloproliferative disorders: essential thrombocythaemia, polycythaemia vera and idiopathic myelofibrosis. *Pathol Biol.(Paris).* **49,** 164-166.

Lahaye, T., Riehm, B., Berger, U.*, et al.* (2005) Response and resistance in 300 patients with BCR-ABL-positive leukemias treated with imatinib in a single center: a 4.5-year follow-up. *Cancer,* **103,** 1659-1669.

Lahortiga, I., De Keersmaecker, K., Van Vlierberghe, P.*, et al.* (2007) Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. *Nat. Genet.,* **39,** 593-595.

Larsen, T.S., Christensen, J.H., Hasselbalch, H.C.*, et al.* (2007) The JAK2 V617F mutation involves B- and T-lymphocyte lineages in a subgroup of patients with Philadelphia-chromosome negative chronic myeloproliferative disorders. *Br.J.Haematol.,* **136,** 745-751.

Latif, F., Tory, K., Gnarra, J.*, et al.* (1993) Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science,* **260,** 1317-1320.

Le Beau, M.M., Westbrook, C.A., Diaz, M.O.*, et al.* (1986) Evidence for the involvement of GM-CSF and FMS in the deletion (5q) in myeloid disorders. *Science,* **231,** 984-987.

Lee, D.S., Lee, Y.S., Yun, Y.S.*, et al.* (2003) A study on the incidence of ABL gene deletion on derivative chromosome 9 in chronic myelogenous leukemia by interphase fluorescence in situ hybridization and its association with disease progression. *Genes Chromosomes Cancer,* **37,** 291-299.

Lee, J.W., Kim, Y.G., Soung, Y.H.*, et al.* (2005) The JAK2 V617F mutation in de novo acute myelogenous leukemias. *Oncogene,* **25,** 3.

Levine, R.L., Loriaux, M., Huntly, B.J.*, et al.* (2005a) The JAK2V617F activating mutation occurs in chronic myelomonocytic leukemia and acute myeloid leukemia, but not in acute lymphoblastic leukemia or chronic lymphocytic leukemia. *Blood,* **106,** 3377-3379.

Levine, R.L., Wadleigh, M., Cools, J.*, et al.* (2005b) 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.*, et al.* (2005c) KIAA1509 is a novel PDGFRB fusion partner in imatinib-responsive myeloproliferative disease associated with a t(5;14)(q33;q32). *Leukemia.,* **19,** 27-30.

Levy, D.E. & Darnell, J.E., Jr. (2002) Stats: transcriptional control and biological impact. *Nature reviews. Molecular cell biology,* **3,** 651-662.

Liu, L.G., Tanaka, H., Ito, K.*, et al.* (2003) Chronic myelogenous leukemia with e13a3 (b2a3) type of BCR-ABL transcript having a DNA breakpoint between ABL exons a2 and a3. *Am.J Hematol.,* **74,** 268-272.

Livak, K.J. & Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.,* **25,** 402-408.

Loncarevic, I.F., Romer, J., Starke, H.*, et al.* (2002) Heterogenic molecular basis for loss of ABL1-BCR transcription: deletions in der(9)t(9;22) and variants of standard t(9;22) in BCR-ABL1-positive chronic myeloid leukemia. *Genes Chromosomes.Cancer,* **34,** 193-200.

Lozzio, C.B. & Lozzio, B.B. (1975) Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood,* **45,** 321-334.

Lundan, T., Volin, L., Ruutu, T.*, et al.* (2005) Allogeneic stem cell transplantation reverses the poor prognosis of CML patients with deletions in derivative chromosome 9. *Leukemia.,* **19,** 138-140.

Lux, M.L., Rubin, B.P., Biase, T.L.*, et al.* (2000) KIT Extracellular and Kinase Domain Mutations in Gastrointestinal Stromal Tumors. *American Journal of Pathology,* **156,** 791-795.

MacKinney, A.A., Jr., Clark, S.S., Borcherding, W.*, et al.* (1993) Simultaneous demonstration of the Philadelphia chromosome in T, B, and myeloid cells. *Am.J Hematol.,* **44,** 48-52.

Magnusson, M.K., Meade, K.E., Brown, K.E.*, et al.* (2001) Rabaptin-5 is a novel fusion partner to platelet-derived growth factor beta receptor in chronic myelomonocytic leukemia. *Blood,* **98,** 2518-2525.

Mahon, F.X., Rea, D., Guilhot, J.*, et al.* (2010) Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. *Lancet Oncology,* **11,** 1029-1035.

Martin, P.J., Najfeld, V., Hansen, J.A.*, et al.* (1980) Involvement of the B-lymphoid system in chronic myelogenous leukaemia. *Nature,* **287,** 49-50.

Martin-Rendon, E. & Watt, S.M. (2003) Exploitation of stem cell plasticity. *Transfusion Medicine,* **13,** 325-349.

Martyre, M.C., Bousse-Kerdiles, M.C., Romquin, N.*, et al.* (1997) Elevated levels of basic fibroblast growth factor in megakaryocytes and platelets from patients with idiopathic myelofibrosis. *Br.J Haematol.,* **97,** 441-448.

Mayer, B.J. & Baltimore, D. (1994) Mutagenic analysis of the roles of SH2 and SH3 domains in regulation of the Abl tyrosine kinase. *Mol.Cell Biol,* **14,** 2883-2894.

McWhirter, J.R. & Galasso, D.L. & Wang, J.Y. (1993) A coiled-coil oligomerization domain of Bcr is essential for the transforming function of Bcr-Abl oncoproteins. *Mol.Cell Biol,* **13,** 7587-7595.

Melo, J.V. (1997) BCR-ABL gene variants. *Baillieres Clin Haematol.,* **10,** 203-222.

Melo, J.V., Gordon, D.E., Cross, N.C.*, et al.* (1993) The ABL-BCR fusion gene is expressed in chronic myeloid leukemia. *Blood,* **81,** 158-165.

Melo, J.V., Hochhaus, A., Yan, X.H.*, et al.* (1996) Lack of correlation between ABL-BCR expression and response to interferon-alpha in chronic myeloid leukaemia. *Br.J Haematol.,* **92,** 684-686.

Melo, J.V., Myint, H., Galton, D.A.*, et al.* (1994) P190BCR-ABL chronic myeloid leukaemia: the missing link with chronic myelomonocytic leukaemia? *Leukemia.,* **8,** 208-211.

Mesa, R.A., Barosi, G., Cervantes, F.*, et al.* (2006) Myelofibrosis with myeloid metaplasia: Disease overview and non-transplant treatment options. *Best Practice & Research Clinical Haematology,* **19,** 495-517.

Mitelman, F. & Johansson, B. & Mertens, F. (2007) The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer,* **7,** 233-245.

Mocsai, A., Zhou, M., Meng, F.*, et al.* (2002) Syk is required for integrin signaling in neutrophils. *Immunity.,* **16,** 547-558.

Moliterno, A.R., Williams, D.M., Gutierrez-Alamillo, L.I.*, et al.* (2004) Mpl Baltimore: a thrombopoietin receptor polymorphism associated with thrombocytosis. *Proceedings of the National Academy of Sciences,* **101,** 11444-11447.

Moorman, A.V., Richards, S.M., Robinson, H.M.*, et al.* (2007) Prognosis of children with acute lymphoblastic leukemia (ALL) and intrachromosomal amplification of chromosome 21 (iAMP21). *Blood,* **109,** 2327-2330.

Morerio, C., Acquila, M., Rosanda, C.*, et al.* (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.

Mullighan, C.G., Zhang, J., Harvey, R.C.*, et al.* (2009) JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences,* **106,** 9414-9418.

Neubauer, H., Cumano, A., Muller, M.*, et al.* (1998) Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. *Cell,* **93,** 397-409.

Neves, H., Ramos, C., da Silva, M.G.*, et al.* (1999) The Nuclear Topography of ABL, BCR, PML, and RARalpha Genes: Evidence for Gene Proximity in Specific Phases of the Cell Cycle and Stages of Hematopoietic Differentiation. *Blood,* **93,** 1197-1207.

Nicholson, E. & Holyoake, T. (2009) The chronic myeloid leukemia stem cell. *Clinical lymphoma & myeloma,* **9 Suppl 4,** S376-381.

Nowell, P.C. & Hungerford, D.A. (1960) Chromosome studies on normal and leukemic human leukocytes. *J Natl Cancer Inst.,* **25:85-109.,** 85-109.

O'Brien, S. & Deininger, M. (2003) Imatinib in patients with newly diagnosed chronic-phase chronic myeloid leukemia. *Semin Hematol,* **40,** 26-30.

O'Brien, S.G., Guilhot, F., Larson, R.A.*, et al.* (2003) Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *New England Journal of Medicine,* **348,** 994-1004.

O'Shea, J.J. & Gadina, M. & Schreiber, R.D. (2002) Cytokine signaling in 2002: new surprises in the Jak/Stat pathway. *Cell,* **109 Suppl,** S121-131.

Oberstrass, J., Reifenberger, G., Reifenberger, J.*, et al.* (1996) Mutation of the Von Hippel-Lindau tumour suppressor gene in capillary haemangioblastomas of the central nervous system. *J Pathol,* **179,** 151-156.

Oda, T., Heaney, C., Hagopian, J.R.*, et al.* (1994) Crkl is the major tyrosine-phosphorylated protein in neutrophils from patients with chronic myelogenous leukemia. *J Biol Chem,* **269,** 22925-22928.

Ohnishi, K., Ohno, R., Tomonaga, M.*, et al.* (1995) A randomized trial comparing interferon-alpha with busulfan for newly diagnosed chronic myelogenous leukemia in chronic phase. *Blood,* **86,** 906-916.

Okabe, M., Kunieda, Y., Nakane, S.*, et al.* (1995) Establishment and characterization of a new Ph1-positive chronic myeloid leukemia cell line MC3 with trilineage phenotype and an altered p53 gene. *Leukemia and Lymphoma,* **16,** 493-503.

Ouchi, T., Lee, S.W., Ouchi, M.*, et al.* (2000) Collaboration of signal transducer and activator of transcription 1 (STAT1) and BRCA1 in differential regulation of IFN-gamma target genes. *Proceedings of the National Academy of Sciences,* **97,** 5208-5213.

Pardanani, A. & Akin, C. & Valent, P. (2006) Pathogenesis, clinical features, and treatment advances in mastocytosis. *Best Practice & Research Clinical Haematology,* **19,** 595-615.

Parganas, E., Wang, D., Stravopodis, D.*, et al.* (1998) Jak2 Is Essential for Signaling through a Variety of Cytokine Receptors. *Cell,* **93,** 385-395.

Partanen, J., Armstrong, E., Makela, T.P.*, et al.* (1992) A novel endothelial cell surface receptor tyrosine kinase with extracellular epidermal growth factor homology domains. *Mol.Cell Biol,* **12,** 1698-1707.

Pastore, Y.D., Jelinek, J., Ang, S.*, et al.* (2003) Mutations in the VHL gene in sporadic apparently congenital polycythemia. *Blood,* **101,** 1591-1595.

Peeters, P., Raynaud, S.D., Cools, J.*, et al.* (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.

Pfirrmann, M. & Hasford, J. (2001) Testing Sokal's and the new prognostic score for chronic myeloid leukaemia treated with alpha-interferon: comments. *Br.J Haematol.,* **114,** 241-243.

Piccart-Gebhart, M.J., Procter, M., Leyland-Jones, B.*, et al.* (2005) Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *New England Journal of Medicine,* **353,** 1659-1672.

Popenoe, D.W., Schaefer-Rego, K., Mears, J.G.*, et al.* (1986) Frequent and extensive deletion during the 9,22 translocation in CML. *Blood,* **68,** 1123-1128.

Popovici, C., Zhang, B., Gregoire, M.J.*, et al.* (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.

Preston, D.L., Kusumi, S., Tomonaga, M.*, et al.* (1994) Cancer incidence in atomic bomb survivors. Part III. Leukemia, lymphoma and multiple myeloma, 1950-1987. *Radiat.Res,* **137,** S68-S97.

Preudhomme, C., Guilhot, J., Nicolini, F.E.*, et al.* (2010) Imatinib plus peginterferon alfa-2a in chronic myeloid leukemia. *New England Journal of Medicine,* **363,** 2511-2521.

Przeworski, M. & Hudson, R.R. & Di Rienzo, A. (2000) Adjusting the focus on human variation. *Trends Genet,* **16,** 296-302.

Puil, L., Liu, J., Gish, G.*, et al.* (1994) Bcr-Abl oncoproteins bind directly to activators of the Ras signalling pathway. *EMBO J,* **13,** 764-773.

Quintas-Cardama, A., Kantarjian, H., Talpaz, M., *et al.* (2005) Imatinib mesylate therapy may overcome the poor prognostic significance of deletions of derivative chromosome 9 in patients with chronic myelogenous leukemia. *Blood,* **105,** 2281-2286.

Raitano, A.B. & Whang, Y.E. & Sawyers, C.L. (1997) Signal transduction by wild-type and leukemogenic Abl proteins. *Biochim.Biophys.Acta,* **1333,** F201-F216.

Rameshwar, P., Chang, V.T., Thacker, U.F., *et al.* (1998) Systemic transforming growth factor-beta in patients with bone marrow fibrosis--pathophysiological implications. *Am.J Hematol.,* **59,** 133-142.

Reid, A.G., Huntly, B.J.P., Grace, C., *et al.* (2003a) Survival implications of molecular heterogeneity in variant Philadelphia-positive chronic myeloid leukaemia. *Br.J.Haematol.,* **121,** 419-427.

Reid, A.G. & Tarpey, P.S. & Nacheva, E.P. (2003b) High-resolution analysis of acquired genomic imbalances in bone marrow samples from chronic myeloid leukemia patients by use of multiple short DNA probes. *Genes Chromosomes Cancer,* **37,** 282-290.

Reiter, A., Skladny, H., Hochhaus, A., *et al.* (1997) Molecular response of CML patients treated with interferon-alpha monitored by quantitative Southern blot analysis. German chronic myeloid leukaemia (CML) Study Group. *Br.J Haematol.,* **97,** 86-93.

Reiter, A., Sohal, J., Kulkarni, S., *et al.* (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., Watmore, A., *et al.* (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.

Ridge, S.A., Worwood, M., Oscier, D., *et al.* (1990) FMS mutations in myelodysplastic, leukemic, and normal subjects. *Proceedings of the National Academy of Sciences,* **87,** 1377-1380.

Robinson, H.M., Martineau, M., Harris, R.L., *et al.* (2005) Derivative chromosome 9 deletions are a significant feature of childhood Philadelphia chromosome positive acute lymphoblastic leukaemia. *Leukemia.,* **19,** 564-571.

Roman, J., Jimenez, A., Barrios, M., *et al.* (2001) E1A3 as a unique, naturally occurring BCR-ABL transcript in an indolent case of chronic myeloid leukaemia. *Br.J.Haematol.,* **114,** 635-637.

Ronaghi, M., Karamohamed, S., Pettersson, B., *et al.* (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal.Biochem.,* **242,** 84-89.

Ross, T.S., Bernard, O.A., Berger, R., *et al.* (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.

Rousselot, P., Huguet, F., Rea, D.*, et al.* (2007) Imatinib mesylate discontinuation in patients with chronic myelogenous leukemia in complete molecular remission for more than 2 years. *Blood,* **109,** 58-60.

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.

Rubin, B.P., Singer, S., Tsao, C.*, et al.* (2001) KIT Activation Is a Ubiquitous Feature of Gastrointestinal Stromal Tumors. *Cancer Res.,* **61,** 8118-8121.

Rubinstein, M., Rubinstein, S., Familletti, P.C.*, et al.* (1978) Human leukocyte interferon purified to homogeneity. *Science,* **202,** 1289-1290.

Saglio, G., Kim, D.W., Issaragrisil, S.*, et al.* (2010) Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *New England Journal of Medicine,* **362,** 2251-2259.

Saharinen, P., Ekman, N., Sarvas, K.*, et al.* (1997) The Bmx Tyrosine Kinase Induces Activation of the Stat Signaling Pathway, Which Is Specifically Inhibited by Protein Kinase Cdelta. *Blood,* **90,** 4341-4353.

Sakai, I., Takeuchi, K., Yamauchi, H.*, et al.* (2002) Constitutive expression of SOCS3 confers resistance to IFN-alpha in chronic myelogenous leukemia cells. *Blood,* **100,** 2926-2931.

Salgia, R., Brunkhorst, B., Pisick, E.*, et al.* (1995a) Increased tyrosine phosphorylation of focal adhesion proteins in myeloid cell lines expressing p210BCR/ABL. *Oncogene,* **11,** 1149-1155.

Salgia, R., Li, J.L., Lo, S.H.*, et al.* (1995b) Molecular cloning of human paxillin, a focal adhesion protein phosphorylated by P210BCR/ABL. *J Biol Chem,* **270,** 5039-5047.

Salgia, R., Uemura, N., Okuda, K.*, et al.* (1995c) CRKL links p210BCR/ABL with paxillin in chronic myelogenous leukemia cells. *J Biol Chem,* **270,** 29145-29150.

Saltzman, A., Stone, M., Franks, C.*, et al.* (1998) Cloning and characterization of human Jak-2 kinase: high mRNA expression in immune cells and muscle tissue. *Biochem.Biophys.Res Commun.,* **246,** 627-633.

Sanger, F. & Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences,* **74,** 5463-5467.

Sangrar, W., Zirgnibl, R.A., Gao, Y.*, et al.* (2005) An identity crisis for fps/fes: oncogene or tumor suppressor? *Cancer Res.,* **65,** 3518-3522.

Saussele, S., Lauseker, M., Gratwohl, A.*, et al.* (2010) Allogeneic hematopoietic stem cell transplantation (allo SCT) for chronic myeloid leukemia in the imatinib era: evaluation of its impact within a subgroup of the randomized German CML Study IV. *Blood,* **115,** 1880-1885.

Sawyers, C.L. (1999) Chronic Myeloid Leukemia. *New England Journal of Medicine,* **340,** 1330-1340.

Sawyers, C.L., Hochhaus, A., Feldman, E.*, et al.* (2002) Imatinib induces hematologic and cytogenetic responses in patients with chronic myelogenous leukemia in myeloid blast crisis: results of a phase II study. *Blood,* **99,** 3530-3539.

Schnittger, S., Kohl, T.M., Haferlach, T.*, et al.* (2006) KIT-D816 mutations in AML1-ETO-positive AML are associated with impaired event-free and overall survival. *Blood,* **107,** 1791-1799.

Schouten, J.P., McElgunn, C.J., Waaijer, R.*, et al.* (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.,* **30,** e57.

Schwaller, J., Parganas, E., Wang, D.*, et al.* (2000) Stat5 is essential for the myelo- and lymphoproliferative disease induced by TEL/JAK2. *Mol.Cell,* **6,** 693-704.

Score, J., Curtis, C., Waghorn, K.*, et al.* (2006) Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia. *Leukemia.,* **20,** 827-832.

Scott, L.M., Tong, W., Levine, R.L.*, et al.* (2007) JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *New England Journal of Medicine,* **356,** 459-468.

Shi, Y. & Alin, K. & Goff, S.P. (1995) Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal portion of the Abl protein, suppresses v-abl transforming activity. *Genes Dev.,* **9,** 2583-2597.

Shows, T.B., Sakaguchi, A.Y., Naylor, S.L.*, et al.* (1982) Clustering of leukocyte and fibroblast interferon genes of human chromosome 9. *Science,* **218,** 373-374.

Shtalrid, M., Talpaz, M., Blick, M.*, et al.* (1988) Philadelphia-negative chronic myelogenous leukemia with breakpoint cluster region rearrangement: molecular analysis, clinical characteristics, and response to therapy. *Journal of Clinical Oncology,* **6,** 1569-1575.

Shtivelman, E., Lifshitz, B., Gale, R.P.*, et al.* (1985) Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature,* **315,** 550-554.

Shuai, K. (2006) Regulation of cytokine signaling pathways by PIAS proteins. *Cell research,* **16,** 196-202.

Silly, H., Chase, A., Mills, K.I.*, et al.* (1994) No evidence for microsatellite instability or consistent loss of heterozygosity at selected loci in chronic myeloid leukaemia blast crisis. *Leukemia.,* **8,** 1923-1928.

Silva, C.M. (2004) Role of STATs as downstream signal transducers in Src family kinase-mediated tumorigenesis. *Oncogene,* **23,** 8017-8023.

Silver, R.T. (2003) Imatinib mesylate (Gleevec(TM)) reduces phlebotomy requirements in polycythemia vera. *Leukemia.,* **17,** 1186-1187.

Silver, R.T., Woolf, S.H., Hehlmann, R.*, et al.* (1999) An evidence-based analysis of the effect of busulfan, hydroxyurea, interferon, and allogeneic bone marrow transplantation in treating the chronic phase of chronic myeloid leukemia: developed for the American Society of Hematology. *Blood,* **94,** 1517-1536.

Simonsson, B., Gedde-Dahl, T., Markevarn, B.*, et al.* (2011) Combination of pegylated IFN-alpha2b with imatinib increases molecular response rates in patients with low- or intermediate-risk chronic myeloid leukemia. *Blood,* **118,** 3228-3235.

Sinclair, P.B., Nacheva, E.P., Leversha, M.*, et al.* (2000) Large deletions at the t(9;22) breakpoint are common and may identify a poor-prognosis subgroup of patients with chronic myeloid leukemia. *Blood,* **95,** 738-743.

Sodroski, J.G. & Goh, W.C. & Haseltine, W.A. (1984) Transforming potential of a human protooncogene (c-fps/fes) locus. *Proceedings of the National Academy of Sciences,* **81,** 3039-3043.

Sokal, J.E., Cox, E.B., Baccarani, M.*, et al.* (1984) Prognostic discrimination in "good-risk" chronic granulocytic leukemia. *Blood,* **63,** 789-799.

Specchia, G., Albano, F., Anelli, L.*, et al.* (2004) Derivative chromosome 9 deletions in chronic myeloid leukemia are associated with loss of tumor suppressor genes. *Leukemia and Lymphoma,* **45,** 689-694.

Specchia, G., Albano, F., Anelli, L.*, et al.* (2003) Deletions on der(9) chromosome in adult Ph-positive acute lymphoblastic leukemia occur with a frequency similar to that observed in chronic myeloid leukemia. *Leukemia.,* **17,** 528-531.

Stam, K., Heisterkamp, N., Grosveld, G.*, et al.* (1985) Evidence of a new chimeric bcr/c-abl mRNA in patients with chronic myelocytic leukemia and the Philadelphia chromosome. *New England Journal of Medicine,* **313,** 1429-1433.

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.

Steimle, C., Lehmann, U., Temerinac, S.*, et al.* (2007) Biomarker analysis in polycythemia vera under interferon-alpha treatment: clonality, EEC, PRV-1, and JAK2 V617F. *Ann.Hematol.,* **86,** 239.

Stoiber, D., Kovacic, B., Schuster, C.*, et al.* (2004) TYK2 is a key regulator of the surveillance of B lymphoid tumors. *J Clin Invest,* **114,** 1650-1658.

Storlazzi, C.T., Anelli, L., Surace, C.*, et al.* (2002a) Molecular cytogenetic characterization of a novel additional chromosomal aberration in blast crisis of a Ph-positive chronic myeloid leukemia. *Cancer Genet.Cytogenet.,* **134,** 109-113.

Storlazzi, C.T., Specchia, G., Anelli, L.*, et al.* (2002b) Breakpoint characterization of der(9) deletions in chronic myeloid leukemia patients. *Genes Chromosomes Cancer,* **35,** 271-276.

Strachan, T. & Read, A.P. (2004) Instability of the human genome: mutation and DNA repair. In: *Hum.Mol.Genet* (ed. by Strachan, T. & Read, A.P.), Vol. 3 pp. 315-349. Garland Science, a member of the Taylor & Francis Group, New York.

Swerdlow, S.H.C., E.; Harris, N.L.; Jaffe, E.S.; Pileri, S.A.; Stein, H.; Thiele, J.; Vardiman, J.W. (2008) *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissue* IARC, Lyon.

Takai, H., Kanematsu, M., Yano, K.*, et al.* (1998) Transforming growth factor-beta stimulates the production of osteoprotegerin/osteoclastogenesis inhibitory factor by bone marrow stromal cells. *J Biol Chem,* **273,** 27091-27096.

Talpaz, M., McCredie, K.B., Mavligit, G.M.*, et al.* (1983) Leukocyte interferon-induced myeloid cytoreduction in chronic myelogenous leukemia. *Blood,* **62,** 689-692.

Tan, J.C., Nocka, K., Ray, P.*, et al.* (1990) The dominant W42 spotting phenotype results from a missense mutation in the c-kit receptor kinase. *Science,* **247,** 209-212.

Taylor, C.F. & Taylor, G.R. (2004) Current and emerging techniques for diagnostic mutation detection: an overview of methods for mutation detection. *Methods Mol.Med,* **92,** 9-44.

Tefferi, A. (2006a) Classification, Diagnosis and Management of Myeloproliferative Disorders in the JAK2V617F Era. *Hematology (Am Soc.Hematol Educ.Program.),* **2006,** 240-245.

Tefferi, A. (2006b) The diagnosis of polycythemia vera: New tests and old dictums. *Best Practice & Research Clinical Haematology,* **19,** 455-469.

Tefferi, A., Lasho, T.L., Schwager, S.M.*, et al.* (2005) The JAK2V617F tyrosine kinase mutation in myelofibrosis with myeloid metaplasia: lineage specificity and clinical correlates. *Br.J.Haematol.,* **131,** 320-328.

The Italian Cooperative Study Group on Chronic Myeloid Leukaemia (1994) Interferon Alfa-2a as Compared with Conventional Chemotherapy for the Treatment of Chronic Myeloid Leukemia. *New England Journal of Medicine,* **330,** 820-825.

Thomas, G.A. & Williams, D.L. & Soper, S.A. (2001) Capillary Electrophoresis-based Heteroduplex Analysis with a Universal Heteroduplex Generator for Detection of Point Mutations Associated with Rifampin Resistance in Tuberculosis. *Clinical Chemistry,* **47,** 1195-1203.

Tono, C., Xu, G., Toki, T.*, et al.* (2005) JAK2 Val617Phe activating tyrosine kinase mutation in juvenile myelomonocytic leukemia. *Leukemia.,* **19,** 1843-1844.

Toyabe, S., Watanabe, A., Harada, W.*, et al.* (2001) Specific immunoglobulin E responses in ZAP-70-deficient patients are mediated by Syk-dependent T-cell receptor signalling. *Immunology,* **103,** 164-171.

Trask, B., Fertitta, A., Christensen, M.*, et al.* (1993) Fluorescence in situ hybridization mapping of human chromosome 19: cytogenetic band location of 540 cosmids and 70 genes or DNA markers. *Genomics,* **15,** 133-145.

Vaclavicek, A., Bermejo, J.L., Schmutzler, R.K., *et al.* (2007) Polymorphisms in the Janus kinase 2 (JAK)/signal transducer and activator of transcription (STAT) genes: putative association of the STAT gene region with familial breast cancer. *Endocrine-Related Cancer,* **14,** 267-277.

van Rhee, F., Hochhaus, A., Lin, F., *et al.* (1996) p190 BCR-ABL mRNA is expressed at low levels in p210-positive chronic myeloid and acute lymphoblastic leukemias. *Blood,* **87,** 5213-5217.

Vaquez, H. (1892) Sur une forme speciale de cyanose s'accompagnant d'hyperglobulie excessive et persitante. *C R Soc Biol,* **44,** 384-388.

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., *et al.* (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood,* **114,** 937-951.

Velpeau, A.A. (1827) Sur le resorption du pusaet sur l'alteration du sang dans les maladies clinique de persecution nememant. Premier observation. *Rev.Med.,* **2,** 216.

Verstovsek, S., Estey, E., Manshouri, T., *et al.* (2001) High expression of the receptor tyrosine kinase Tie-1 in acute myeloid leukemia and myelodysplastic syndrome. *Leukemia and Lymphoma,* **42,** 511-516.

Verstovsek, S., Kantarjian, H., Manshouri, T., *et al.* (2002) Prognostic significance of Tie-1 protein expression in patients with early chronic phase chronic myeloid leukemia. *Cancer,* **94,** 1517-1521.

Verstovsek, S., Kantarjian, H., Mesa, R.A., *et al.* (2010) Safety and efficacy of INCB018424, a JAK1 and JAK2 inhibitor, in myelofibrosis. *New England Journal of Medicine,* **363,** 1117-1127.

Virchow, R. (1845) Weisses Blut. *Frorieps Notizen,* **36,** 151-156.

Virchow, R. (1847) Weisses Blut und Milztumoren. *Medicale Zeitung,* **16,** 9-15.

Virchow, R. (1856) Die Leukaemie. In: *Gesammelte Abhandlungen Zur Wissenschaftlichen Medizin*, pp. 190-210. Meidinger, Frankfurt.

Visscher, P.M., Brown, M.A., McCarthy, M.I., *et al.* (2012) Five years of GWAS discovery. *Am.J Hum.Genet.,* **90,** 7-24.

Vizmanos, J.L., Hernandez, R., Vidal, M.J., *et al.* (2004) Clinical variability of patients with the t(6;8)(q27;p12) and FGFR1OP-FGFR1 fusion: two further cases. *Hematol.J,* **5,** 534-537.

Vizmanos, J.L., Ormazabal, C., Larrayoz, M.J., *et al.* (2006) JAK2 V617F mutation in classic chronic myeloproliferative diseases: a report on a series of 349 patients. *Leukemia.,* **20,** 534-535.

von Lindern, M. & Schmidt, U. & Beug, H. (2004) Control of erythropoiesis by erythropoietin and stem cell factor: a novel role for Bruton's tyrosine kinase. *Cell Cycle,* **3,** 876-879.

Vorstman, J.A., Jalali, G.R., Rappaport, E.F.*, et al.* (2006) MLPA: a rapid, reliable, and sensitive method for detection and analysis of abnormalities of 22q. *Hum.Mutat.,* **27,** 814-821.

Vu, H.A., Xinh, P.T., Masuda, M.*, et al.* (2006) FLT3 is fused to ETV6 in a myeloproliferative disorder with hypereosinophilia and a t(12;13)(p13;q12) translocation. *Leukemia.,* **20,** 1414-1421.

Wallace, R.B., Shaffer, J., Murphy, R.F.*, et al.* (1979) Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch. *Nucleic Acids Res.,* **6,** 3543-3557.

Walz, C., Chase, A., Schoch, C.*, et al.* (2005) The t(8;17)(p11;q23) in the 8p11 myeloproliferative syndrome fuses MYO18A to FGFR1. *Leukemia.,* **19,** 1005-1009.

Weber-Matthiesen, K., Winkemann, M., Muller-Hermelink, A.*, et al.* (1992) Simultaneous fluorescence immunophenotyping and interphase cytogenetics: a contribution to the characterization of tumor cells. *J Histochem.Cytochem.,* **40,** 171-175.

Wen, S.T. & Van Etten, R.A. (1997) The PAG gene product, a stress-induced protein with antioxidant properties, is an Abl SH3-binding protein and a physiological inhibitor of c-Abl tyrosine kinase activity. *Genes Dev.,* **11,** 2456-2467.

Wilkinson, K., Velloso, E.R., Lopes, L.F.*, et al.* (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.

Wilson, G., Frost, L., Goodeve, A.*, et al.* (1997) BCR-ABL Transcript With an e19a2 (c3a2) Junction in Classical Chronic Myeloid Leukemia. *Blood,* **89,** 3064.

Wittwer, C.T., Reed, G.H., Gundry, C.N.*, et al.* (2003) High-Resolution Genotyping by Amplicon Melting Analysis Using LCGreen. *Clinical Chemistry,* **49,** 853-860.

Xiao, S., Nalabolu, S.R., Aster, J.C.*, et al.* (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.*, et al.* (2001) Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood,* **97,** 2434-2439.

Yoong, Y., VanDeWalker, T.J., Carlson, R.O.*, et al.* (2005) Clinical correlates of submicroscopic deletions involving the ABL-BCR translocation region in chronic myeloid leukemia. *Eur.J Haematol.,* **74,** 124-127.

Zioncheck, T.F., Harrison, M.L., Isaacson, C.C.*, et al.* (1988) Generation of an active protein-tyrosine kinase from lymphocytes by proteolysis. *J Biol Chem,* **263,** 19195-19202.

Zirngibl, R.A. & Senis, Y. & Greer, P.A. (2002) Enhanced endotoxin sensitivity in fps/fes-null mice with minimal defects in hematopoietic homeostasis. *Mol.Cell Biol,* **22,** 2472-2486.