Aus der Medizinischen Klinik mit Schwerpunkt

Hämatologie und Onkologie

der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Sequence Analysis of the NUMB Gene in Chronic Myeloid Leukaemia Patients

zur Erlangung des akademischen Grades

Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Christian Oberender

aus Merseburg

Datum der Promotion: 12.09.2014

List of Abbreviations

ABL-1	Abelson murine leukemia viral oncogene homolog 1
AD	advanced disease
Allo-SCT	allogeneic haematopoietic stem-cell transplantation
AML	acute myeloid leukaemia
AP	accelerated phase
BCR	breakpoint cluster region
BLAST	basic local alignment search tool
BM	bone marrow
bp	base pairs
BP	blast phase
BP-L	lymphoid blast phase
BP-M	myeloid blast phase
CCyR	complete cytogenetic response
CHR	complete haematologic response
CLP	common lymphoid progenitor cells
CML	chronic myeloid / myelogenous leukaemia
CMP	common myeloid progenitor cells
CMR	complete molecular response
CP	chronic phase
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
GMP	granulocyte macrophage progenitor
GTC	guanidium isothiocyanate
GUSβ	β-glucuronidase
HHT	homoharringtonine
HHR	homologous recombination repair
HOXA9	homeobox gene A9
HSC	haematopoietic stem cell
HU	hydroxyurea
IFN	Interferon-α
kb	kilobases
KD	BCR-ABL1 kinase domain
LMPP	lymphoid myeloid primed progenitor cells

LSC	leukaemic stem cell
LSK	$Lin^{-}Sca\text{-}1^{+}c\text{-}Kit^{+}$ haematopoietic cells
LT-HSC	long-term haematopoietic stem cells
MCyR	major cytogenetic response
MEP	megakaryocyte erythroid progenitor cells
MLL-1	Mixed lineage leukemia-1
MMR	major molecular response
MSI1	Musashi-1
MSI2	Musashi-2
NA	not analysed
NC	normal control
NHEJ	non-homologous end-joining
NIC	NOTCH intracellular domain
PB	peripheral blood
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PP2A	protein phosphatase 2A
qPCR	real-time quantitative polymerase chain reaction
RCLB	red cell lysis buffer
ROS	reactive oxygen species
rpm	rotations per minute
RRM	RNA recognition motif
RT	reverse transcription
RT-PCR	reverse transcription polymerase chain reaction
SETBP1	SET binding protein-1
SNP	single nucleotide polymorphism
SNV	single nucleotide variation
ST-HSC	short-term haematopoietic stem cells
TGF-β	Transforming growth factor β
TKI	tyrosine kinase inhibitor
UTR	untranslated region
WBC	white blood cells
WHO	world health organization
wt	wild type

List of Figures

List of Tables

Abstract

Better understanding of the transformation of chronic myeloid leukemia (CML) from chronic phase (CP) to the invariably fatal blast phase (BP) is of critical importance for the clinical management of patients with CML. However, the mechanisms responsible for triggering disease progression have eluded investigators' efforts. Recently, our group verified reported data showing increased levels of Musashi-2 (MSI2) transcripts in patients with CML in BP compared to those in CP, implying a role for MSI2 in CML transformation[1-3].

The Musashi gene family is reported to control critical cell fate decisions by binding to target mRNAs, including the *NUMB* mRNA, thereby inhibiting translation[4,5]. Unregulated increased expression of MSI2 results in the dysfunction of NUMB-Notch signalling, leading to haematopoietic stem cell (HSC) proliferation, impaired myeloid differentiation and worse clinical prognosis in CML[2,3]. Therefore, we hypothesized that mutations mapping to the *NUMB* gene may perturb this signalling pathway and thereby influence CML transformation.

I tested this notion by directly sequencing the entire *NUMB* transcript in 22 patients with CML of whom 10 were in CP and 12 were in BP. Archived RNA extracted from peripheral blood from subjects with CML was reverse transcribed to cDNA and the entire *NUMB* gene transcript was amplified. The amplified products were subjected to Sanger sequencing.

For the 22 patients with CML, the *NUMB* gene transcript sequence was determined to be identical to the published wild type sequence, apart from two previously reported single nucleotide polymorphisms (SNP) mapping to the 3'-UTR: rs11625196 (C/G) and rs7202 (C/T) [6,7].

I observed no significant difference in the distribution of the genotypes of the two SNP between that reported for normal healthy individuals and the CML patients, nor between the different disease phases. However, rs7202 genotype had significant influence on the mortality rate of patients in BP – an observation which was fortuitously biased by different treatment modalities. The software tools Mfold and SNPfold predicted a negligible effect of the two SNP on the secondary structure of *NUMB* mRNA.

In a summary, these observations suggest that NUMB, which regulates Notch, Hedgehog and p53 signalling, is not the primary cause of CML evolution[8-11]. However, it would be prudent to confirm this finding in a study with greater number of CML CP and BP patients and/or using a sequencing method with higher sensitivity, such as deep-gene sequencing.

Abstrakt

Ein besseres Verständnis der molekularbiologischen Vorgänge, die zur Transformation der chronischen myeloischen Leukämie (CML) von der relativ indolenten chronischen Phase (CP) zur fatalen Blastenkrise führen, ist von entscheidender Bedeutung für das klinische Management von CML-Patienten. Unsere Arbeitsgruppe konnte vorher publizierte Daten bestätigen, die eine höhere Expression von Musashi-2 (MSI2) in der Blastenkrise im Vergleich zur CP zeigten. Eine Funktion von MSI2 im Rahmen der Transformation der CML wird diskutiert[1-3]. Die Mitglieder der Musashi-Genfamilie gelten als Regulatoren von Zellteilung und Zelldifferenzierung unreifer Zellen und agieren, indem sie die Translation bestimmter mRNAs wie der *NUMB* mRNA inhibieren[4,5]. Eine Dysregulation von MSI2 führt zu einer Dysfunktion des NUMB-Notch-Signalweges und daraufhin zu einer verstärkten Proliferation hämatopoietischer Stammzellen (HSC), einer eingeschränkten myeloischen Differenzierung und einer schlechteren Prognose[2,3]. Wir vermuteten, dass Mutationen im *NUMB*-Gen den NUMB-Notch-Signalweg deregulieren und zur CML-Transformation beitragen können.

Diese Vermutung testete ich, indem ich *NUMB* cDNA von 22 CML-Patienten, davon 10 in CP und 12 in BP, nach der Sanger-Kettenabbruchmethode sequenzierte. Dafür wurde RNA aus dem peripheren Blut von CML-Patienten extrahiert und revers transkribiert zu cDNA. Das gesamte Transkript des *NUMB*-Genes wurde mittels PCR amplifiziert und daraufhin sequenziert.

Unter den 22 CML-Patienten fanden sich keine Abweichungen der *NUMB* cDNA-Basensequenz im Vergleich zur publizierten Wildtyp-Sequenz, abgesehen von zwei Einzelnukleotid-polymorphismen (SNP) in der 3‘-untranslatierten Region: rs11625196 (C/G) und rs7202 (C/T) [6,7]. Ich konnte keine signifikanten Unterschiede im Auftreten der SNP-Genotypen zwischen gesunden Kontrollpersonen und CML-Patienten oder zwischen CP- und Blastenkrise-Patienten beobachten. Allerdings zeigte sich ein signifikanter Einfluss des rs7202-Genotyps auf die Mortalität von Blastenkrise-Patienten. Diese Beobachtung ist jedoch am ehesten auf ungleiche Therapieregime zurückzuführen. Mfold- und SNPfold-Software sagen einen vernachlässigbaren Effekt der beiden SNPs auf die räumliche *NUMB*-mRNA-Struktur vorher.

Zusammengefasst weisen die Beobachtungen dieser Studie daraufhin, dass *NUMB*-Mutationen nicht als die primäre Ursache der CML-Transformation anzusehen sind. Dennoch wäre es gerechtfertigt, die Erkenntnisse dieser Studie mit einer größeren Anzahl an CML-Patienten oder auch einer sensitiveren Next-Generation-Sequenzierungsmethode zu überprüfen.

1. Introduction

1.1 Chronic myeloid leukaemia

Chronic myeloid leukaemia (CML) is a rare malignancy with a worldwide annual incidence of 1-2 per 100,000 and is often diagnosed later in life, between 40 and 60 years of age, although it is not uncommon among younger individuals[12]. There is a slight predominance of the disease among males with reported male to female ratios of 1.3-1.8[(13)].

CML is a clonal myeloproliferative disorder of a pluripotent haematopoietic stem cell (HSC) associated with a balanced reciprocal chromosomal translocation, the Philadelphia translocation t(9;22)(q34;q11), involving chromosomes 9 and 22[(14,15)]. This translocation juxtaposes the *BCR* gene on chromosome 22 to the *ABL1* gene on chromosome 9, leading to the formation of the *BCR-ABL1* fusion gene, which encodes a chimeric onco-protein with constitutive tyrosine kinase activity[16]. This up-regulated tyrosine kinase activity is generally accepted to be the causal agent for increased proliferation, inhibited apoptosis and altered adhesion properties of haematopoietic cells in the bone marrow (BM) compartment observed in CML patients. As a consequence there is uncontrolled expansion of the myeloid lineage[17].

1.1.1 Chronic phase

The clinical course of CML is frequently classified into three distinct phases – chronic, accelerated and blastic phase. Up to 90% of patients are diagnosed in the relatively indolent chronic phase (CP)[12]. Remarkably, a significant proportion of patients is asymptomatic at presentation and is diagnosed fortuitously when a routine medical examination shows an elevated white blood cell (WBC) count. Often patients present with common findings such as fatigue, splenomegaly, night sweats, weight loss and anaemia[12,18].

A CML patient in CP usually presents with neutrophilic leukocytosis without significant dysplasia but left-shift and <15% blasts in the peripheral blood (PB)[12,18]. Elevated numbers of platelets, basophils and eosinophils are not uncommon. Bone marrow (BM) biopsy reveals hypercellularity due to expanded myeloid series, majority of which includes maturating granulocytes and <15% blasts[12,18]. The spleen may be enlarged as a consequence of the granulocytic infiltration of the red pulp cords[12]. Diagnosis of CML is confirmed by demonstrating expression of the *BCR-ABL1* transcripts as recommended by WHO[12].

1.1.2 Clinical management of newly diagnosed CP CML

Treatment of CML has been revolutionised since the start of the 21st century by the introduction of targeted therapy, namely by the tyrosine kinase inhibitor (TKI) imatinib mesylate (IM)[19]. IM induces complete cytogenetic remission (CCyR), defined as the absence of Philadelphia chromosome in at least 20 metaphases prepared from bone marrow aspirate biopsy, in up to 75% of newly diagnosed CML CP patients within 18 months[18,20,21]. Furthermore, 40-50% of patients achieve major molecular response (MMR) within the first 18 months[20,21], which is defined as 3 log reduction in tumour load. These data represent unprecedented success in managing CML patients with IM such that it is now one of the recommended first-line therapies, thereby replacing interferon-α, cytarabine and hydroxyurea[18,19,21]. Since the advent of IM, the life expectancy of CML patients has improved enormously (Figure 1). 15% of patients are reported to have achieved complete molecular remission (CMR) after 36 months, which is defined as failure to detect *BCR-ABL1* transcripts by quantitative real-time PCR and is equivalent to a 4.5-5 log reduction[22]. For a select few patients who achieved CMR, the IM treatment has been terminated without the re-appearance of *BCR-ABL1* fusion gene in peripheral blood samples analysed[23]. Relapse, i.e. detection of *BCR-ABL1* transcripts, was observed generally within the first 6 months of stopping IM therapy[23].

However, a significant number of patients either fail to respond or have only a transient response to IM. Of the mechanisms reported to be cause of resistance, the detection of kinase domain (KD) mutations are the most common cause in the clinics. The KD mutations impede binding of TKI to the BCR-ABL1 protein and therefore the latter remains active[24]. Approved TKIs for treating CML in the clinic now include the 2nd and 3rd generation drugs nilotinib, dasatinib, bosutinib and ponatinib. These, along with IM, have unique activity profiles against the various KD mutations, such that a resistance to IM due to a KD mutation may be overcome by switching to a different 2nd- or 3rd-generation TKI drug, with the aim of reversing disease progression and restoring molecular response. Nilotinib and dasatinib have meanwhile been approved for the first-line treatment of CML.

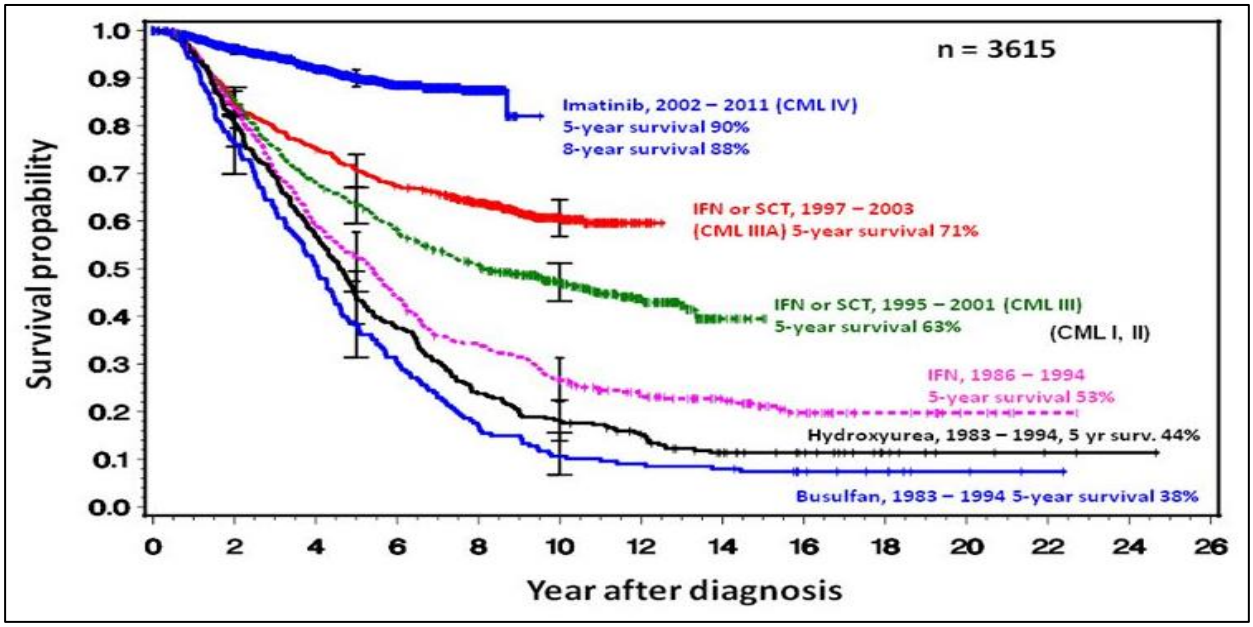

Figure 1. CML treatment efficacy in chronic phase 1983-2011.
Survival probability as shown in the German CML trials I – IV has climbed steeply over the past three decades, culminating in the advent of tyrosine kinase inhibitors, which have replaced all former treatment regimes.
Taken from: Hochhaus A. Lecture "The Evolution of CML treatment". CML-GOLS 09-03-2012.

Nevertheless, there are BCR-ABL1-independent mechanisms of TKI resistance, e.g. the activation of alternative signalling pathways, such as altered wnt/β-catenin, hedgehog or Alox-5 signalling[25]. Importantly, IM is reported not to eradicate leukaemic stem cells (LSC) and therefore potentially the disease reservoir persists from which relapse may develop[26,27].

However, an increasing number of patients who achieve complete molecular remission over a period of 2 years have their IM therapy interrupted. While some of these patients relapse, usually within 6 months of stopping IM treatment, a significant number continue to have undetectable disease by real-time PCR[23]. These observations argue against IM being ineffective against LSC or it may mean the Philadelphia positive stem cells are exhausted over time following long periods of TKI therapy.

1.1.3 Advanced disease

Without a continuous and life-long therapy, patients invariably progress from chronic phase to fatal blast phase (BP) either abruptly or through an intermittent accelerated phase (AP). These two phases of disease progression are often grouped together as “advanced disease” (AD). Patients progressing to AD have worse prognosis, with severe anaemia, thrombocytopenia or marked splenic enlargement[12].

The criteria defining AP, the disease phase between CP and BP, are ambiguous. However, AP is suspected in patients with persistently high or increasing leucocytosis, splenomegaly, thrombocytosis or thrombocytopenia[12,18].

As defined by the World Health Organization (WHO), progression to BP is recorded when the proportion of blasts in PB or BM is higher than 20% or when extramedullary blast proliferation, so-called chloroma, is present. Around 70% of CML CP patients progress to myeloid BP, while 20-30% transform into lymphoid BP[12]. The underlying mechanisms determining which lineage, i.e. myeloid or lymphoid, the CML patient in CP progresses to are unclear. The detection of additional cytogenetic abnormalities such as an additional Philadelphia chromosome, trisomy 8, trisomy 19 or isochromosome 17q is associated with disease progression[28]. In addition, mutations in genes encoding proto-oncogenes or tumour-suppressors are not uncommon in advanced-stage CML patients[28]. Such mutations were detected in >75% of BP patients by deep sequencing[29].

The DNA-damaging function of BCR-ABL1 tyrosine kinase has been extensively studied. These studies show that this novel fusion gene produces relatively high levels of reactive oxygen species (ROS) which are reported to cause oxidative stress and mutagenesis[30,31]. Equally, several publications show BCR-ABL1-dependent mechanisms to promote unfaithful DNA repair. BCR-ABL1 is reported to provoke incorrect non-homologous end-joining (NHEJ) and erroneous homologous recombination repair (HRR) mechanisms after DNA double-strand breaks[30,32-34].

The precise mechanisms of blast transformation are ill-defined. A variety of individual genetic alterations have been reported by investigators but it is not clear how and if these observations are related to each other. Given the complexity of the data reported, it is likely that several distinct chains of genetic hits probably lead to clonal evolution and blast transformation. Recently, investigators reported increased expression of the RNA-binding protein Musashi-2

(MSI2) in CML BP and acute myeloid leukaemia (AML)[2,3,35]. These observations were consistent with a putative role for MSI2 as regulator of haematopoietic cell fate decisions.

1.2 The Musashi-NUMB-Notch signalling pathway

1.2.1 The Musashi family

The evolutionarily conserved Musashi family includes the two RNA-binding proteins Musashi-1 (MSI1) and Musashi-2 (MSI2). Both contain two RNA recognition motifs (RRM) showing high sequence homology between MSI1 and MSI2 – the first RRM exhibits 81% and the second RRM 93% identity on the amino acid level, respectively. Remarkably, there is a 100% identity within the ribonucleoprotein consensus sequence octapeptide (RGFGFVTF)[36]. In vitro studies showed that both proteins have similar RNA binding specifity[37].

The widespread Musashi proteins are involved in asymmetric cell division, germ and somatic stem cell function and cell fate determination in somatic tissues[38]. However, their expression patterns differ between tissues.

MSI1 activity was found to be enriched in stem and progenitor cells of the embryonic and postnatal central nervous system, the gastrointestinal system, the mammary gland and the skin[39-46]. Increased MSI1 expression correlates with advanced disease stages and worse prognosis in colon and breast cancer[47,48]. Similarly, MSI1 overexpression is observed in urothelial, esophagus and cervix carcinoma and in tumours of the central nervous system, such as glioma, medulloblastoma and ependymoma[49-54].

In contrast, MSI2 is expressed ubiquitously[37]. Several tissues show coexistent MSI1 and MSI2 activity, e.g. the central nervous system[40]. Importantly, MSI1 expression in hematopoietic stem (HSC) and progenitor cells is negligible in comparison to MSI2 expression[3].

1.2.2 Molecular biological function of Musashi proteins and NUMB

Much of the data on how Musashi proteins function was gleaned from MSI1 studies. MSI1 is reported to recognise and bind to a consensus sequence ((G/A)U_nAGU, n=1-3 bases) in the 3'-untranslated region (3'-UTR) of mRNAs and to compete with the eukaryotic initiation factor 4G (eIF4G) for the poly(A) binding protein (PABP), hence hampering translation initiation[4,5,55]. It is to be assumed that mRNAs from a number of genes are potential MSI1 targets. This includes genes involved in cell cycle, cell proliferation, cell differentiation, apoptosis and ubiquitination of proteins[56]. Well-characterised mRNA targets are NUMB and CDKN1A (p21, CIP1, WAF1)[4,57]. However, there are many more mRNAs which include the 3'-UTR consensus sequence recognised by MSI1[56]. Based on these observations, a number of investigators have implied that MSI1 might work as a master regulator of genes involved in proliferation and differentiation of stem and progenitor cells.

The Musashi target NUMB is reported to possess an important role as an inhibitor of the cell cycle-driving and differentiation-constraining Notch and Hedgehog pathways (summarised in Figure 2). NUMB protein is believed to enhance ubiquitination of NOTCH-1 receptor, NOTCH intracellular domain (NIC) and downstream effector of Hedgehog pathway GLI1 and to inhibit ubiquitination of p53[8-11]. The product of the *CDKN1A* gene (*CIP1, WAF1*), p21, works as a cyclin-dependent kinase inhibitor that decreases the activity of cyclin/cyclin dependent kinase activity and is regulated by p53 on the level of transactivation of transcription[58,59]. NUMB protein-mediated stabilization of p53 protein and MSI1 protein-mediated posttranscriptional repression of *NUMB* and *p21* (*CIP1*, *WAF1*) mRNA reveal a strong relation between MSI1 function and activity of p53 and p21. High MSI1 expression putatively leads to lower NUMB and p21 protein levels and increased p53 proteasomal degradation, thus promoting cell cycle progression and inhibiting apoptosis.

Notch signalling is involved in neural development and haematopoiesis[60]. It is often dysregulated in different types of malignancies, such as T-acute lymphoblastic leukaemia/lymphoma[61]. Finally, Notch activation leads to increased transcription of its downstream effectors *HES1* and *TRIB2*, which are believed to promote symmetric division of stem cells and progenitor cells and to inhibit their differentiation[60].

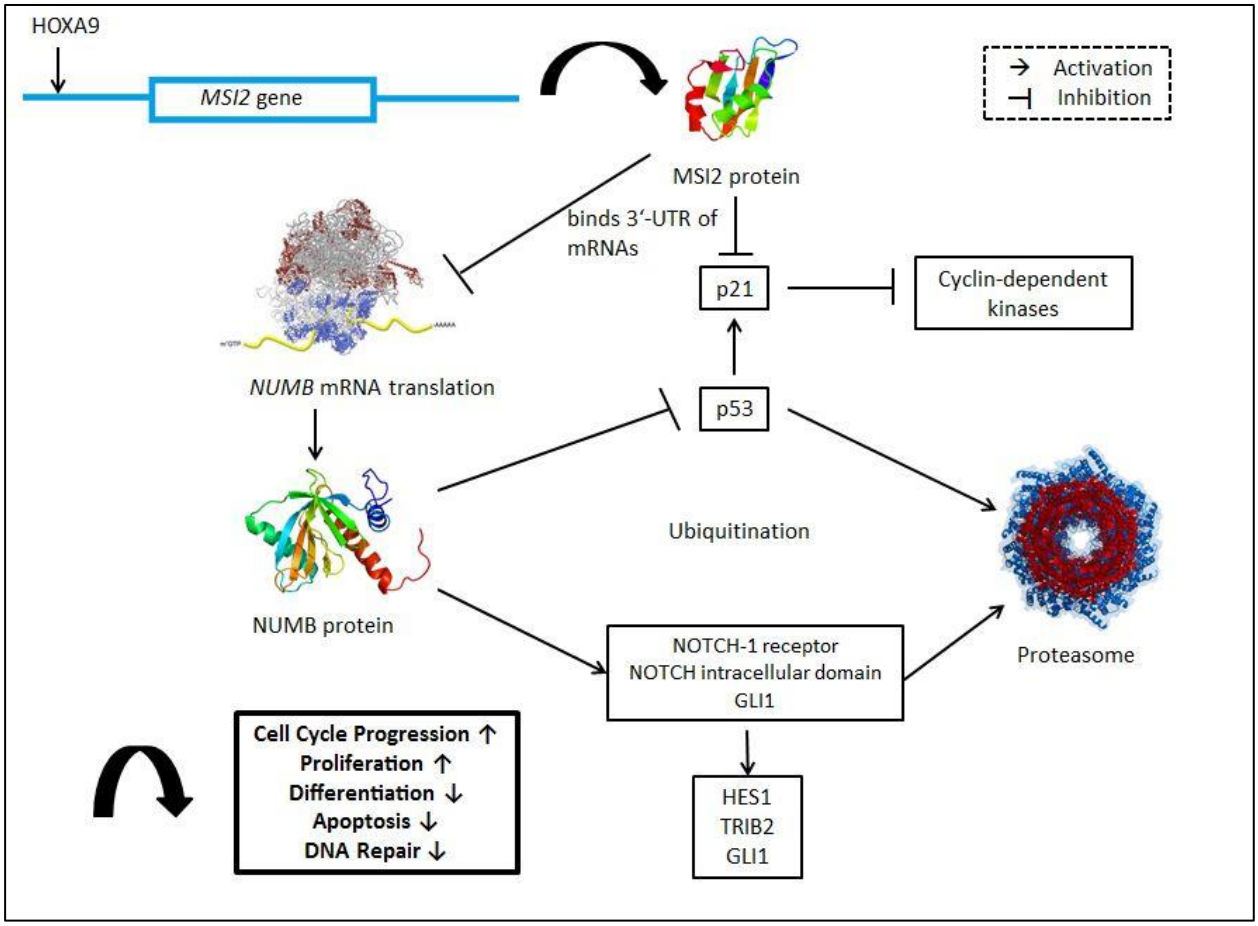

Figure 2. Summary of published molecular interactions between MSI2, NUMB, Notch, Hedgehog and p53 signalling pathways.
The putative consequences of MSI2 overexpression in a haematopoietic stem cell are illustrated. HOXA9 is reported to be an activator of *MSI2* gene transcription[2]. The four images were taken from web pages[62-65].

1.2.3 Assumed physiological role of Musashi-2 in the haematopoietic system

Investigators have reported detailed studies in which the effects of *MSI2* knockdown and induced *MSI2* overexpression were closely monitored in mice[3,66]. These authors quantified *MSI2* expression in haematopoietic cells and found that *MSI2* expression decreases rapidly from haematopoietic stem cells (HSC) to intermediate progenitors and mature cells[66]. Long-term (LT-HSC) and short term (ST-HSC) HSC as well as lymphoid myeloid primed progenitors (LMPP) showed the highest percentage of *MSI2*-expressing cells. In contrast, common lymphoid (CLP), common myeloid (CMP), granulocyte macrophage (GMP) and megakaryocyte-erythroid (MEP) progenitors displayed lower *MSI2* expression, whereas differentiated cells had no detectable expression[66].

MSI2-null mice were fully viable for more than a year but severe defects in primitive precursors with a decrease in the number and proliferation capacity of ST-HSC and LMPP were detectable[66]. In addition, the number of circulating leukocytes was markedly decreased[66]. Interestingly, the decrease in the number of LT-HSC was not statistically significant. The observed effects after *MSI2* knockout were not completely related to strengthened differentiation. Cell cycle studies displayed a significant decrease in the number of cells in S-G2-M phase and an insignificant increase in apoptosis[66]. Additionally, *MSI2*-defective mice showed smaller spleens and thymi and their *MSI2*-defective HSC a dramatically impaired competitive repopulation capacity in bone marrow transplantation experiments[66].

Kharas et al. (2010) reported effects of induced *MSI2* overexpression in mice[3]. At the acute stage these investigators observed that overexpression of MSI2 caused impaired differentiation of early myeloid, megakaryocytic and granulocytic lineages, as well as a quantitative expansion of total hematopoietic stem and progenitor cells. LT-HSC expanded in absolute numbers but ST-HSC expanded preferentially. At the same time cell numbers decreased in quiescent G0 phase and increased in G1 and S-G2-M phases. As a consequence of MSI2 overexpression Ras, mitogene-activated protein kinase (Mapk), cyclin D1 and Myc activity were found to be increased[3]. Predominantly asymmetric partitioning of NUMB in MSI2 overexpressing haematopoietic stem and progenitor cells undergoing cell division could be observed[3,67].

These observations suggest that augmented MSI2 activity stimulates HSC to proliferate and partly differentiate to progenitor cells but inhibits further differentiation to more mature haematopoietic cells.

1.2.4 Pathophysiological role in advanced-stage CML

Ito et al. (2010) used two mouse models resembling CML chronic phase and CML myeloid blast phase and compared gene expression levels[2,68]. They observed significantly lower *NUMB* expression levels in BP and that exogenous expression of *NUMB* in BP mouse models weakened propagation of leukaemia, showing a more differentiated, less infiltrative phenotype[2]. Besides, they discerned that NUMB function depends at least in part on p53 because they observed impotence of NUMB to affect leukaemia cell growth upon p53 knockdown. *NOTCH* expression was found to be increased in BP and *NOTCH* knockdown to have NUMB-similar effects. As a consequence of *MSI2* knockdown in BP mouse models, leukaemia was more differentiated and less able to propagate disease, with an increase in *NUMB* expression[2].

Kharas et al. (2010) reported similar data using inducible *MSI2* expression systems *ex-vivo* and *in-vivo*[3]. *MSI2* induction in *BCR-ABL1* expressing HSC caused a rapid and lethal myeloproliferative CML BP-like disease in mice[3].

Both groups assessed if there was a correlation between *MSI2* expression and clinical status of CML patients. Ito et al. (2010) compared gene expression in patient samples enriched for $CD34^+$ cells and found significantly higher *MSI2* expression in BP than in CP patients and down-regulation of *NUMB* in the majority of BP patients. Patients after allogeneic stem cell transplantation (Allo-SCT) who had the highest *MSI2* expression levels had significantly higher disease relapse and death risk[2]. These observations were mirrored by Kharas et al. (2010) findings[3].

Our group examined whether *MSI2* expression levels might be a suitable marker to monitor patients' clinical status and risk for disease progression. Our results presented at the congress of the American Society of Hematology (ASH) 2011 and at the ESH-iCMLf congress 2012[1,69] confirmed increased *MSI2* expression in PB samples of BP patients compared to samples of CP patients and greater mortality of acute myeloid leukaemia (AML) patients with high *MSI2* expression[1]. However, a retrospective longitudinal study of MSI2 expression measurements in PB RNA samples showed that while MSI2 expression kinetics reflected those seen with *BCR-ABL1* transcript numbers, the increase in the former did not precede increase in fusion gene mRNA levels[69]. These data suggest that serial *MSI2* mRNA quantification in bulk WBC is not more informative than longitudinal *BCR-ABL1* transcript measurement.

1.3 Subject of the dissertation

Genetic lesions in addition to BCR-ABL1 that may affect a variety of genes and cytogenetic aberrations are very common features in advanced disease stages of CML. These observations can at least partly be explained by the mentioned BCR-ABL1-provoked direct and indirect mechanisms which may lead to an accumulation of DNA damage. The initial balanced translocation t(9;22) may arise as a consequence of incorrect repair of two DNA double-strand breaks[70]. If this is proven, it would suggest that subsequent genetic instability in CML patients arises due to an inherent predisposition, mutagenic function of BCR-ABL1 and other undefined mechanisms.

From a stochastic point of view, the whole genome of cells expressing *BCR-ABL1* is to a varying extent susceptible to acquisition of additional mutations. The consequences of mutations within a cell will invariably depend on the affected gene, the cell involved, i.e. stem cell, early progenitor cell or mature differentiated cell, and the mutation, i.e. the effect it has on gene expression or gene function. Whereas mutations which inactivate house-keeping enzymes might lead to cell death, mutations affecting cell cycle or apoptosis regulating genes might confer evolutional selection advantages to a cell. Mutations which switch-on proto-oncogenes or switch-off tumour suppressor genes belong to the latter group.

The *NUMB* gene studies suggest that it has properties of a tumour suppressor and as such, mutations could have fatal consequences. It is possible that mutations affecting the protein-coding sequence of *NUMB* could impair the protein's abilities to inhibit Notch and Hedgehog signalling. Similarly, mutations within the 3'-UTR of *NUMB* might alter the interaction between MSI2 protein and *NUMB* mRNA, either leading to an attenuated or an enhanced interaction.

As far as I could ascertain, no data had been published where the entire *NUMB* transcript had been sequenced in advanced-stage CML patients. Therefore, given our group's MSI2 findings we decided to exclude the possibility of NUMB mutations that might affect the NUMB-Notch signalling pathway. Thus, I screened the entire NUMB transcript, including the 3'-UTR, by direct sequencing in 22 CML patients, 10 patients in CP and 12 in BP, for NUMB cDNA mutations, and gathered substantial clinical data.

2. Methods and Materials

2.1 Patients

In total, 22 CML patients were selected for this study from among the 65 included in a separate project to assess the clinical significance of *MSI2* gene expression levels[1] (Kaeda et al. (2011)). Of these 22 patients, 10 had the highest *MSI2* mRNA levels of the CML patients in chronic phase (CP) studied by Kaeda et al. The remaining 12 were selected because they had progressed to blast phase (BP), irrespective of the *MSI2* expression levels.

The clinical characteristics of the 22 CML patients are summarised in Table 1. The *NUMB* gene transcript was sequenced for all the 22 CML patients by dideoxy chain termination sequencing reactions, hereafter referred to as Sanger sequencing.

2.1.1 Patients in blast phase

The 12 patients in BP (3 females, 9 males) had a median age at diagnosis of 51.5 years (range 20-75 years)(Table 1). Of the 12 patients, nine had progressed to myeloid blast transformation and three to lymphoid blast transformation. The patients in AD had been treated with different modalities as detailed in Table 1. Briefly, all of them had been treated with one or more TKI and 5 of them had at some point undergone allogeneic stem cell transplantation (Allo-SCT). Three of them had been administered hydroxyurea (HU) and one Interferon-α (IFN) upfront for several years. Nine of those twelve patients had died after a median survival time of 38 months after first diagnosis (range 13-179 months). Three patients are alive after successful Allo-SCT, with a median survival of 69 months (range 31-70 months, August 2012). Nine patients had been screened for kinase domain (KD) mutations and best response to therapy data was available in nine of the twelve AD cases and is detailed in Table 1.

2.1.2 Patients in chronic phase

The control cohort group included 10 patients in CP (6 females and 4 males) with a median age of 50.5 years at diagnosis (range 20-73 years). The clinical characteristics are detailed in Table 1. Of the 10 patients, one died 73 months after diagnosis from undefined causes. The remaining nine patients are alive with a median survival of 93 months (range 43-162 months).

These ten patients had been administered one or more TKI, two had IFN and three had HU upfront for more than one year. None of them had undergone Allo-SCT. The 10 patients' response to therapy at the time of the sample collection for this study is detailed in Table 1. In general, seven out of these ten patients achieved MMR as best response, one CCyR and two no more than CHR. The KD mutations found in the patients screened by direct sequencing are detailed in Table 1. The KD mutation screening was performed as part of routine investigation in subjects who experienced or were suspected of developing resistance to TKI. In addition, some of the patients were screened as part of systematic screening for KD mutations to determine the frequency of KD mutations in CML patients irrespective of treatment modality as part of a separate on-going study. The findings for this study will be submitted to a peer-reviewed scientific journal for publications in due course. For the patients included in the study, the clinical data was collated by me as part of my CML studies. For nine patients KD mutation status was available. Seven were wild type, one had the F317L and one the H396R mutation.

2.2 Samples and routine follow-up studies.

CML patients are seen in the Charité Campus Virchow-Klinikum outpatients' Haematology-Oncology clinic every three months for routine follow-up studies. The follow-up studies include routine blood analysis, i.e. full peripheral blood picture and leukocyte differential. Furthermore, patients' response to therapeutic agents is monitored at molecular level by quantifying *BCR-ABL1* transcripts. In addition to these, bone marrow biopsy is performed at regular intervals for cytogenetic analysis if clinically indicated and to document complete cytogenetic remission. These biopsies are performed to assess the number of Philadelphia chromosome positive metaphases and to check for clonal evolution that might indicate disease progression. A bone marrow aspirate smear is also prepared and stained to examine the cellularity and cellular morphology. All these follow-up studies are critical for routine monitoring of patients' response to treatment at haematological, cytogenetic and molecular level. In keeping with the declaration of Helsinki, informed signed consent for collection of blood specimens for routine clinical laboratory and research studies is obtained when the patients present for the first time in the clinic. As well as the studies listed above, an additional 10 mL of peripheral blood and/or 1-2 mL bone marrow aspirate is often collected into anticoagulant EDTA from patients of interest for research studies.

Ref	Disease Phase	Dead / Alive	Sex	Age at Diagnosis (years)	Survival (months)	Best Response		KD Mutation	Therapy
						Achieved by patient	At Date of Sample		
1	BP-M	A	F	52	>69	CMR	-	NA	Imatinib-Nilotinib-Allo-SCT
2	BP-M	D	M	59	179	CHR	-	wt	Myleran-HU-Imatinib
3	BP-L	A	M	51	>70	CMR	-	wt	Imatinib-Dasatinib -Allo-SCT
4	BP-M	D	M	20	96	NA	-	G250E	HU-Imatinib-Nilotinib-Bafetinib-HHT
5	BP-M	A	M	56	>31	CMR	-	NA	Imatinib-Allo-SCT
6	BP-L	D	M	75	13	No CHR	-	L248V	Imatinib-Bafetinib -Dasatinib
7	BP-M	D	F	53	114	CCyR	-	wt	IFN-Imatinib-IFN-Nilotinib-Dasatinib
8	BP-M	D	M	46	10	No CHR	-	E255V	Imatinib-Nilotinib-Allo-SCT
9	BP-M	D	M	22	38	No CHR	-	W478R	Imatinib
10	BP-L	D	M	48	32	NA	-	T315I	Imatinib-Dasatinib-Allo-SCT(x2)
11	BP-M	D	M	46	15	No CHR	-	T315I	Imatinib-Nilotinib
12	BP-M	D	F	66	48	NA	-	wt	HU-Imatinib-Nilotinib
13	CP	A	F	45	>129	MMR	MMR	wt	IFN-Imatinib
14	CP	A	M	48	>43	CHR	CHR	wt	Imatinib-Nilotinib
15	CP	A	F	50	>49	MMR	diagnostic	wt	Imatinib
16	CP	A	F	62	>111	MMR	CCyR	H396R	HU-Imatinib-Bafetinib-Nilotinib-Dasatinib
17	CP	A	M	51	>93	MMR	MMR	wt	Imatinib
18	CP	A	F	61	>113	CCyR	CHR	F317L	HU-Imatinib-Nilotinib-Dasatinib-Ponatinib
19	CP	A	F	20	>162	MMR	MMR	wt	HU+IFN-Imatinib
20	CP	A	F	73	>51	MMR	MMR	wt	Imatinib
21	CP	A	M	46	>48	MMR	MMR	NA	Nilotinib
22	CP	D	M	60	73	CHR	CHR	wt	Imatinib-Nilotinib-Ponatinib

Table 1. Characteristics of 22 CML patients involved in this study.

BP-M: myeloid blastic phase; BP-L: lymphoid blastic phase; CP: chronic phase; D: dead; A: alive; F: female; M: male; CHR: complete haematological response; CCyR: complete cytogenetic response; MMR: major molecular response; CMR: complete molecular response; KD: BCR-ABL1 kinase domain; NA: not analysed; wt: wild type; HHT: homoharringtonine; therapy in chronological order.

2.3 Sample processing

The peripheral blood (10mL) was centrifuged for 10 minutes at 2.5k rpm and the buffy coat layer, approximately 1 to 2 mL immediately above the red cells, collected using a sterile disposable Pasteur pipette and transferred into a 50mL polypropylene tube. The contaminating red cells were removed by re-suspending the buffy coat in 30-40 mL of ice-cold in-house prepared red cell lysis buffer (RCLB: 155mM NH_4Cl, 12mM $NaHCO_3$, 0.1mM EDTA, pH 7.3) and incubated on ice for 30 minutes. The suspension was subsequently centrifuged for 10 minutes at 2.5k rpm and the supernatant discarded. The lysis with RCLB was repeated once more to obtain cellular pellet that was free of contaminating red cells and haem. The cellular pellet was then washed once in phosphate-buffered saline (PBS) to exclude RCLB. The cells were subsequently lysed with 1 mL of guanidinium isothiocyanate (GTC) containing β-mercaptoethanol (7.1µL β-mercaptoethanol per mL of GTC). β-mercaptoethanol was added to GTC just prior to use. The cellular pellet was disrupted and the high molecular weight DNA sheared mechanically by repeatedly aspirating through needle and syringe. The resulting GTC lysates were transferred to 1.5 mL micro-centrifuge tubes and stored at -20°C until required for RNA extraction.

2.4 RNA extraction

The GTC samples were thawed and 350 µL aliquot subjected to ion exchange column chromatography using commercially available RNeasy Minikit (*Qiagen, Hilden, Germany*) to extract total RNA as per manufacturer's recommendations. Briefly, the nucleic acids were bound to the column and the DNA and the protein eluted by using buffers (RW1 and RPE) included in the kit by centrifuging the columns provided by the manufacturer in a micro-centrifuge at 10k rpm for 15 seconds. Finally, the RNA was eluted into 25 µL water and the eluate collected and transferred into another clean sterile 1.5 mL micro-centrifuge tube.

2.5 cDNA synthesis

The total RNA extracted from the GTC lysates was immediately subjected to cDNA synthesis[71]. This was achieved by first heating the whole of the 25 µL eluate containing the RNA for 10 minutes at 65°C in a heating dry-block to linearise the mRNA, i.e. disrupt any secondary RNA structure. The 1.5 mL microcentrifuge tubes were rapidly transferred to the micro-centrifuge and subjected to pulse spin at maximum of 14k rpm to collect all the contents to the bottom of

the tube. The tubes were then immediately put on ice. To this, 21 µL of in-house prepared cDNA synthesis mix (428µL 5X RT-Buffer (*Invitrogen®, Carlsbad, CA, USA*), 21.5µL dithiothreitol (DTT), 85.5µL of 25mM dNTPs, 45µL of 5mg/mL random hexamers, 419.5µL sterile water) was added and subsequently incubated at 37°C for 2 hours. The reverse transcriptase (RT) and RNase inhibitor (RNasin, *Promega®, Fitchburg, WI, USA*) were added to the cDNA synthesis mix just prior to use. The reaction tube was then centrifuged using a 14k pulse spin, i.e. maximum speed for 1-2 seconds, to collect any condensation and bring all the contents down to the bottom of the tube. The tubes were then incubated at 65°C for 10 minutes to inactivate the RT. The resulting cDNA was stored at -20°C until required.

2.6 Quantitative real-time PCR

Our group has confirmed that there is a differential expression of *MSI2* expression levels in CML patients in CP and BP, suggesting the mRNA levels of this asymmetric cell division regulator were prognostic[1] (Kaeda et al. (2011)). These data were generated by quantifying *MSI2* transcripts in 65 CML patients (54 in CP and 11 in advanced disease, i.e. accelerated phase or blast crisis). Briefly, the cDNA was subjected to quantification of *MSI2* and *BCR-ABL1* transcripts and the transcripts of the endogenous control gene *ß-glucuronidase* (*GUS*ß*)* by quantitative real-time PCR using StepOnePlus™ instrument (*Applied Biosystems™, Foster City, CA, USA*)[72]. The instrument default PCR conditions were used. These were: 50°C-2min; 95°C-10min; then 50 cycles at 95°C-15sec, 60°C-1min. *MSI2* and *BCR-ABL1* expression levels were expressed as % ratio of the control gene. The three genes were assayed in separate reactions using Taqman chemistry. The assays included standard curves using serially diluted plasmids, ranging from $1x10^3$ to $1x10^6$ copies of the target gene. The samples were assayed in duplicate, while the standards were assayed in triplicate. Only those samples with ≥ 5500 *GUS*β transcripts were evaluated. Those samples with <5500 *GUS*β copies were considered to be sub-optimum with respect to quality, integrity and quantity of mRNA in the original sample. In summary, we found significantly higher ($p<0.0001$) *MSI2/GUS*β values in advanced disease (median 6.7%, range 1.3-22.9%) in comparison to chronic phase (median 2.2%, range 0.2-6.3%). The data were reported at the American Society of Haematology, San Diego, USA in December 2011 as a poster presentation (Kaeda et al. (2011)).

2.7 *NUMB* gene transcript amplification

Figure 3 shows the location of the *NUMB* gene on the long arm of chromosome 14, in the transition area between chromosomal bands q24.2 and q24.3.

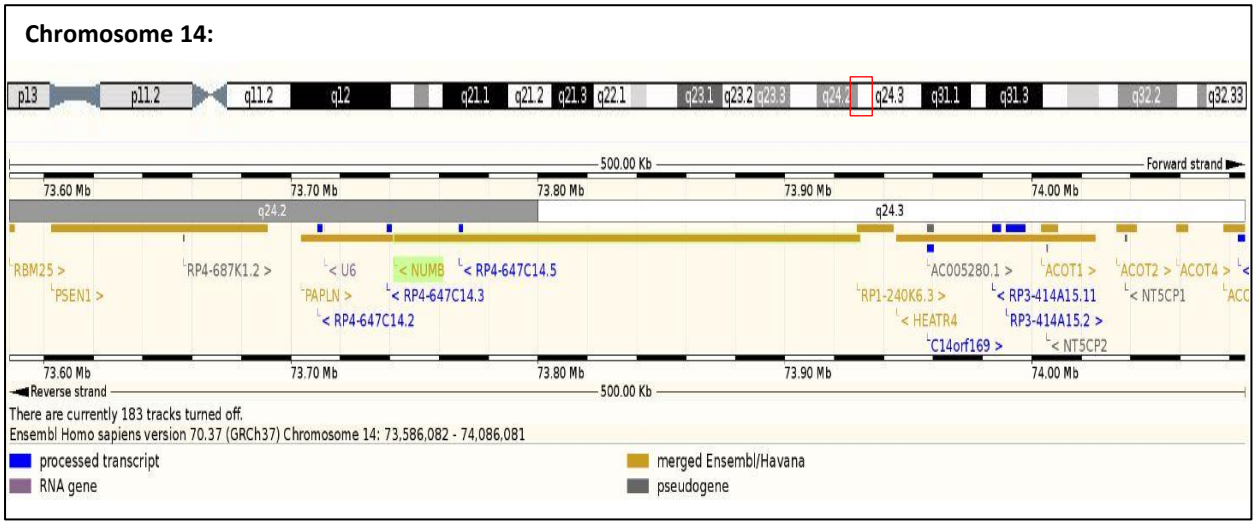

Figure 3. Location of *NUMB* gene on chromosome 14.
Human *NUMB* gene maps to the long arm of chromosome 14 in the region of chromosomal bands q24.2-q24.3 (red frame).
From Ensembl database:
http://Jan2013.archive.ensembl.org/Homo_sapiens/Location/Overview/region?db=core;g=ENSG00000133961;r=14:73586082-74086081

2.7.1 Splice Variants of *NUMB* gene

Alternatively spliced *NUMB* gene transcripts have been desribed[73]. However, the role and function of the different transcripts reported in man are unclear. At the time of the present study, Ensembl Database listed seven different protein-coding transcript types of which two (ENST00000355058 and ENST00000359560) had a 3'-untranslated region including the putative MSI2-binding site sequence *5'-TAGGTT GTAGT TTTTT-3'* and were in contrast to transcript type ENST00000356296 not discarded in later releases of the Ensembl Database[74]. ENST00000355058 is the longest transcript of *NUMB* gene and includes 13 exons, whereas ENST00000359560 misses the 33 base-long third exon of the coding region and is otherwise identical to ENST00000355058[(74)].

Because there were no publications available which could unambiguously point out the expression pattern of the two splice variants in the haematopoietic system, I decided to design

primers to amplify both of them and to see whether I could obtain more than one amplicon following PCR.

Meanwhile, the number of Ensembl-listed protein-coding splice variants of *NUMB* gene has expanded to twenty[75]. Seven out of these twenty splice variants included the 3'-UTR putative MSI2-binding site and were contained within the PCR amplified products, if expressed in cells of interest. The primers used and the expected amplicon sizes in base pair (bp) length for the different splice variants are listed in Table 3.

Hereafter, I will focus on and refer only to the longest *NUMB* gene transcript ENST00000355058.

2.7.2 Polymerase chain reaction amplification of *NUMB* gene transcript

The entire length of the *NUMB* transcript includes the 1.95kb coding region, 1.36kb 3'-untranslated region (3'-UTR) and the 0.3kb 5'-untranslated region (5'-UTR) (as shown in Figure 4). The coding region includes 10 exons. I used the NCBI/Primer BLAST software (http://www.ncbi.nlm.nih.gov/tools/primer-blast) to design the *NUMB* specific primers and the primers were synthesized by *TIB MOLBIOL® Syntheselabor GmbH, Berlin/Germany*. The primer sequences and their relative positions are detailed in Figure 4 and Table 2. All PCR amplifications were carried out using Biometra T3 Thermocycler (*Analytik Jena, Jena/Germany*)[76].

	Primer sequence	Direction	Location	Application
1	5'-TCCAGAGGCCAGTCGTCCACA	Forward	Exon 1	PCR+Sequencing
2	5'-CCACTGTCACTGGTTTGGTCATCGG	Reverse	Exon 8	PCR+Sequencing
3	5'-ACAGCTTGCTCGCCAAGGCTC	Forward	Exon 7	PCR+Sequencing
4	5'-CGCTACCCCCTGCTCCCTGT	Reverse	3'-UTR	PCR+Sequencing
5	5'-ACGGTTCTGCAGCTTTCAAT	Forward	Exon 10	PCR+Sequencing
6	5'-GCTTCTACCATGAACATTTATTTTTG	Reverse	3'-UTR	PCR+Sequencing
7	5'-ACAAGTCCAGGCCCAGAGTA	Forward	3'-UTR	Sequencing
8	5'-AAAACCCACCAAGACCCATA	Reverse	3'-UTR	Sequencing

Table 2. List of primers and their relative positions along the *NUMB* transcript.

The *NUMB* transcript was amplified in three fragments, A, B and C as detailed in Figure 4. Fragments A and B, 1.0 kb and 1.2 kb in length, respectively, contained the entire coding region. Fragment C contained the entire 3'-untranslated region. To sequence the NUMB transcript 50 bases 3' of the start codon to 1.36 kb downstream of the stop codon, two primers in addition to the oligonucleotides used in PCR were designed, see Figure 4. Sequencing of the entire NUMB transcript from each patient included in the study was achieved in eight separate reactions, i.e. four reactions each with forward and reverse primers.

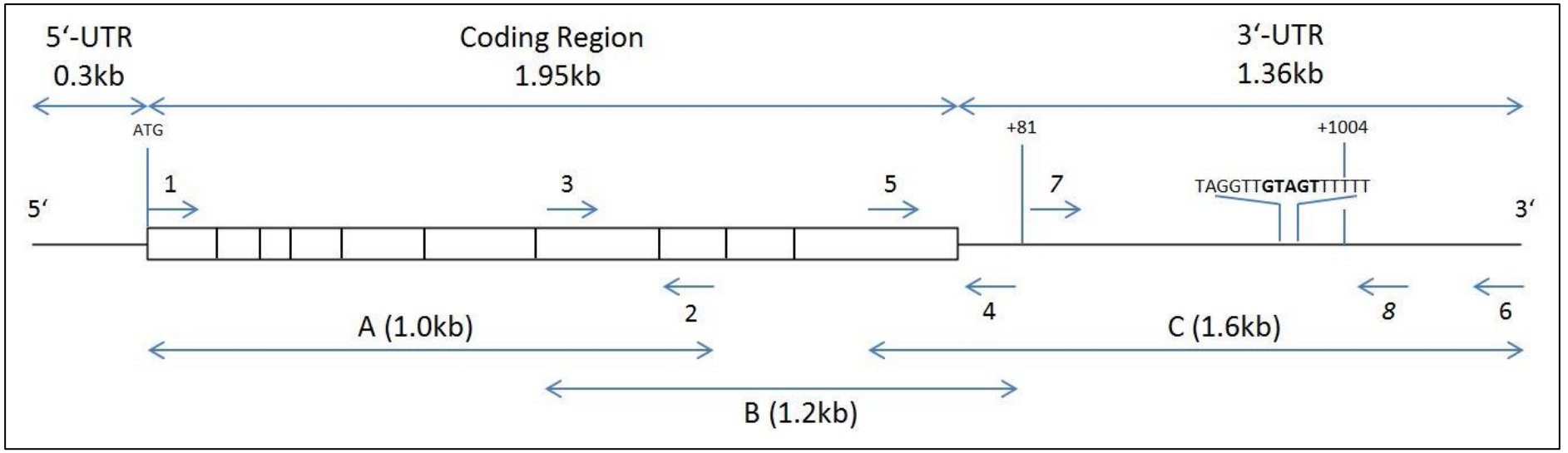

Figure 4. Amplification of *NUMB* gene transcript.

A schematic representation of *NUMB* gene transcript ENST00000355058 is illustrated (not to scale). The 10 exons of the coding region are indicated by the open boxes and the relative positions of the 8 forward and reverse primers are represented by the arrows.

The length of the three PCR products, A (primers 1 and 2), B (primers 3 and 4) and C (primers 5 and 6) are indicated in parentheses. These primers were also used to prime the respective sequencing reactions. Primers 7 and 8 were used only to prime the Sanger sequencing reaction. MSI2 protein is reported to bind to the 3‘-UTR and to inhibit translation of *NUMB* mRNA by recognising the highlighted consensus sequence 5‘-TAGGTT*GTAGT*TTTTT-3‘. The relative positions of two single nucleotide polymorphisms in the 3’-UTR, rs11625196 and rs7202, are shown +81 and +1004 nucleotides from the stop codon TAA, respectively. In this figure the nucleobase thymine is used to demonstrate *NUMB* cDNA and could be interchanged with uracil to demonstrate *NUMB* mRNA.

Primer sequences (5‘-3‘) designed for this study:

1: TCCAGAGGCCAGTCGTCCACA; 2: CCACTGTCACTGGTTTGGTCATCGG; 3: ACAGCTTGCTCGCCAAGGCTC; 4: CGCTACCCCCTGCTCCCTGT;

5: ACGGTTCTGCAGCTTTCAAT; 6: 5’-GCTTCTACCATGAACATTTATTTTTG; 7: 5’-ACAAGTCCAGGCCCAGAGTA; 8: 5’-AAAACCCACCAAGACCCATA.

	Transcript Type	A	B	C
a) Expected amplicons	ENST00000355058	1021	1204	1570
	ENST00000359560	988	1204	1570
b) Retrospectively expected amplicons	ENST00000554546	988	1060	1570
	ENST00000557597	988	1204	1570
	ENST00000555238	1021	1204	1570
	ENST00000356296	1021	1060	1570
	ENST00000556772	-	1204	1570
c) Observed unexpected amplicons	ENST00000560335	727	-	-
	ENST00000555738	693	-	-

Table 3. The expected PCR products.
The expected sizes (in bases) of the amplicons using the primers listed in Table 2 are detailed here.
a) at the time the experimental study was designed[74]. b) retrospectively when the number of splice variants had expanded meanwhile[75]. c) Two transcript types which gave rise to an additional unexpected 0.7 kb band after PCR amplification of fragment A.
Fragments A,B,C as described in Figure 4, fragment size in bases. This table includes all transcript types whose 3'-UTR was expected to be amplified. In addition, an observed 0.7 kb band upon amplification of fragment A could be related to two different splice variants (c). The occurrence of additional fragments of different size was possible but could not be observed.

2.7.3 PCR conditions

PCR was performed in final volume of 30µL, which included 1µL cDNA sample, 3µL 10xPCR Rxn Buffer (Invitrogen®, final concentration x1), 4.8µL 10mM dNTP mix (Invitrogen®, 400nmol/L each), 1.1µL 50mM $MgCl_2$ (Invitrogen®, 1.8mmol/L), 0.5µL forward primer (80nmol/L), 0.5µL reverse primer (80nmol/L), 0.25µL Taq polymerase (Invitrogen®, 0.04units/µL). To this RNase/DNase free sterile water was added to a final volume of 30µL. The thermocycling conditions used to amplify the 3 amplicons are listed below.

For fragments A and B:

94°C-1:00; 35 cycles 98°C-0:10, 60°C-0:30, 72°C-0:50; 72°C-5:00.

For fragment C:

94°C-1:00; 35 cycles 98°C-0:10, 59°C-0:30, 72°C-1:30; 72°C-5:00.

I checked the success of PCR by subjecting the PCR products to a gel electrophoresis (2% agarose, as shown in Figure 5).

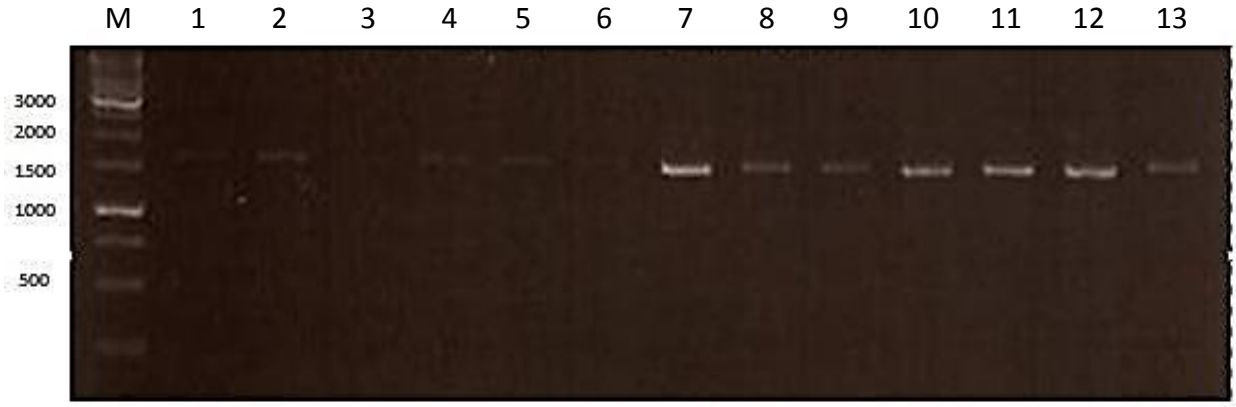

Figure 5. Agarose gel electrophoresis showing amplification of *NUMB* gene transcript 3'-UTR.
The primers 5 and 6 were used to amplify 3'-UTR of the *NUMB* gene, yielding a single amplicon of 1.6kb in length. The gel with PCR fragments of the expected size is shown for 13 CML patients in CP or BP.
M: Molecular marker. Fragment size in bases.

2.7.4 PCR products

The primers used to amplify fragments B and C using the conditions listed above yielded a single amplicon of the expected length, 1.2 and 1.6 kb, respectively. However, primers 1 and 2 which were used to amplify fragment A yielded an amplicon of 0.7 kb in length in addition to the expected PCR product of 1.0 kb size. The amplicon of 0.7 kb size corresponded to 2 (ENST00000560335 and ENST00000555738) of 20 protein-coding *NUMB* transcripts listed in Ensembl database[75]. These two transcripts included a shorter 3'-UTR of only 29 bases and excluded the MSI2-recognition sequence. Therefore, these 2 smaller transcripts were excluded from further analysis. Fragment A was isolated using a commercially available agarose gel purification kit (QIAquick® Gel Extraction Kit (*Qiagen, Hilden, Germany*)). The fragment A was sliced from the agarose gel, the gel piece weighed and transferred into a 1.5 mL micro-centrifuge tube. To this, the threefold volume of QG Buffer included in the kit was added and the tubes were incubated at 50°C until the gel had completely dissolved. Subsequently, isopropanol was added to the suspension in a 1:1 mass of gel: volume of isopropanol ratio and

the entire contents were applied to the QIAquick® column and centrifuged for 1 minute at 13k rpm. The eluate was discarded and the columns washed with Buffer PE included in the kit. The column was centrifuged again for 1 minute at 13k rpm and the eluate discarded. The DNA was eluted using 50µL of Buffer EB (10mmol/L Tris-HCl, pH 8.5) and centrifugation for 1 min at 13k rpm.

PCR fragments B and C were purified using ExoSAP-IT® (*Affymetrix, Santa Clara, CA, USA*)[77]. 4µL of the ExoSAP-IT were added to 10µL aliquot of the PCR products and incubated at 37°C for 20 minutes and then at 80°C for 15 minutes to inactivate the Exonuclease I and recombinant Shrimp Alkaline Phosphatase included in the ExoSAP-IT reagent.

2.8 Sanger sequencing reactions

The PCR products were subjected to direct Sanger sequencing[78]. This was performed using commercially available BigDye Terminator v3.1 Cycle sequencing kit (*Applied Biosystems, Foster City, CA, USA*) as per manufacturer's recommendations. The BigDye sequencing reagent includes the four chain-terminating dideoxynucleotides, i.e. G, T, A and C labelled with different fluorochromes, black, red, green and blue, respectively.

The sequencing reaction included a 5µL aliquot of the purified PCR, 8µL of the BigDye reagent and 1µL of the appropriate primer. The final volume was made up to 20µL with RNase/DNase free sterile water. The mixture was then subjected to 40 cycles of sequencing reaction conditions which were as follows: Initial incubation at 95°C for 30 seconds, followed by 54°C for 20 seconds and finally at 60°C for 4 minutes.

The unincorporated dye terminators were removed after the sequencing reaction using DyeEx 2.0 Spin Kit (*Qiagen, Hilden, Germany*). The sequencing reaction was loaded on to resin columns provided with the kit and the newly synthesized single strand DNA labelled with the respective dideoxynucleotide was eluted with RNase/DNase-free sterile water, while the unincorporated dyes remained bound to the resin column. The eluted sequencing products were subjected to capillary electrophoresis, which was performed by the Department of Genetics at Charité Campus Virchow-Klinikum using the 3730xl sequencer (*Applied Biosystems, Foster City, CA, USA*)[79].

2.9 Sequencing data analysis

I analysed the DNA sequences for each patient using Chromas 2.33 software (*Technelysium Pty Ltd, Australia*). The patient sequences were compared with the reported wild type sequence using the freely available Basic Local Alignment Search Tool (*BLAST, National Center for Biotechnology Information, GenBank+EMBL+DDBJ+PDB sequences*)[80]. The frequencies of the reported single nucleotide polymorphisms (SNP) in the *NUMB* gene for the European population were obtained from HapMap CEU database[81].

The web servers Mfold and SNPfold were used to predict if two distinct SNP had potential consequences on the secondary structure of *NUMB* mRNA[82,83]. Mfold is software which computes several secondary structures of RNA ordered by minimum free energy. In contrast, SNPfold software computes the ensemble and probability of all possible structures of an RNA molecule by calculating the probability of base-pairing for every possible base-pair in a partition function. In addition, SNPfold algorithm calculates the extent to which wild type RNA and SNP-bearing RNA differ and shows the regions of the RNA where base-pairing likelihoods are most dissimilar. Web servers were accessed between January 10th and 20th, 2013. I entered the cDNA sequence of the longest *NUMB* transcript ENST00000355058 (3593 bases) and its SNP variants for all studies on RNA secondary structure.

2.10 Apportionment of work

The work steps of sample processing, RNA extraction, cDNA synthesis and quantitative real-time PCR studies were carried out by Jaspal S. Kaeda, PhD.

PCR amplification of *NUMB* gene transcript, gel electrophoresis, sequencing studies and analysis were performed by me.

In addition, I collected the relevant CML patient data for this and other projects.

3. Results

3.1 *NUMB* sequencing detects two single nucleotide polymorphisms within 3'-untranslated region

For the 22 CML patients, the *NUMB* gene transcript sequence was determined to be identical to the published wild type sequence for the CML patients, apart from two previously reported single nucleotide polymorphisms (SNP, see Figure 6) – rs11625196 (C/G) and rs7202 (C/T)[6,7].

The two SNP mapped to the 3'-UTR – rs11625196 (C/G) 81 bases and rs7202 (C/T) 1004 bases from the stop codon TAA, respectively.

rs11625196 is referred to as "+81 C/G" and rs7202 to as "+1004 C/T" (see Figure 4 and Table 5).

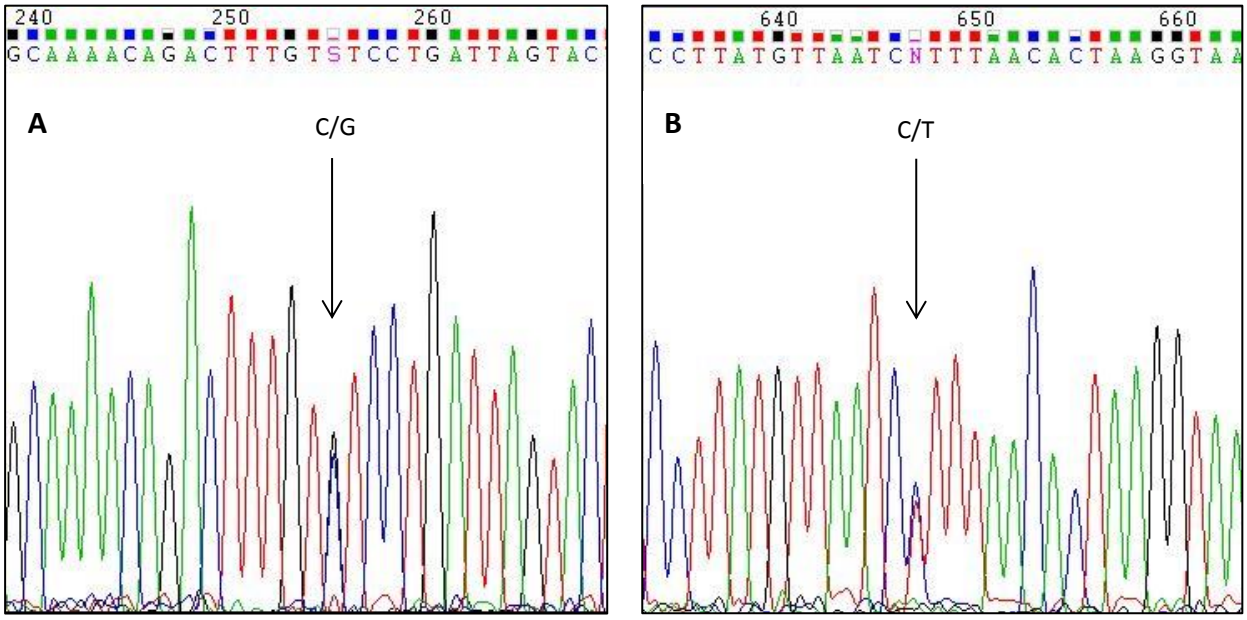

Figure 6. Single nucleotide polymorphisms.
A: SNP rs11625196 (+81 C/G) **B: SNP rs7202 (+1004 C/T)**
A) Sequencing result showing co-occurrence of bases cytosine and guanine (C/G, see arrow) – encoded by the letter "S".
B) Sequencing result showing co-occurrence of bases cytosine and thymine (C/T, see arrow) – in this example encoded by the letter "N", otherwise "Y".
A: adenine, C: cytosine, G: guanine, T: thymine.

3.2 Frequencies of the two SNP in CML cohort and European population

The relative frequencies of these two SNP were not significantly different from the reported frequencies for the healthy European population (Table 4), recorded in HapMap CEU database[6,7]. Furthermore the frequencies for these SNP were similar for BP and CP patients. The sequencing data with respect to +81 SNP and +1004 SNP are summarised in Table 4.

Briefly, the frequencies reported by HapMap CEU database for the +81 SNP among 120 normal controls (NC) are 72.3% CC, 25.5% CG and 2.3% GG. For the +81 SNP, "C" is referred to as the "major allele" and "G" as the "minor allele". Statistical analysis showed no significant difference between the observed frequencies in our CML cohort and the reported frequencies for the NC ($p=0.5718$, Freeman-Halton extension of Fisher's exact test) nor was any significant difference revealed between NC and CP ($p=0.3129$). Similarly the frequency difference between NC and BP ($p>0.9999$) or between CP and BP ($p=0.801$) was not significant either.

The data for the +1004 SNP are detailed in Table 4. As with +81 SNP, the frequencies observed for the different genotypes at position +1004 in the CML cohort were not statistically different from those reported by the HapMap CEU database for 226 normal controls (54.6% CC, 38.6% CT, 6.8% TT). For the +1004 SNP, "C" is referred to as the "major allele" and "T" as the "minor allele". Statistical analysis showed no significant difference between the observed frequencies in our CML cohort and the reported frequencies for the NC ($p=0.5969$, Freeman-Halton extension of Fisher's exact test) nor was any significant difference revealed between NC and CP ($p=0.8755$). Similarly the frequency difference between NC and BP ($p=0.5635$) or between CP and BP ($p>0.9999$) was not significant either.

		+ 81 C/G			+ 1004 C/T		
Samples	**N**	**C/C**	**C/G**	**G/G**	**C/C**	**C/T**	**T/T**
NC	120 | 226	87 (72%)	31 (26%)	2 (2%)	123 (55%)	87 (39%)	16 (7%)
BP	12	9 (75%)	3 (25%)	0 (0%)	5 (42%)	6 (50%)	1 (8%)
CP	10	7 (70%)	2 (20%)	1 (10%)	5 (50%)	5 (50%)	0 (0%)

Table 4. The observed SNP frequencies.
The frequencies for the SNP observed in our study are listed for the healthy control group of European origin (NC) and the CML patients included in the study.
No significant differences in the distribution of the two SNP could be found, neither between NC and CML cohort nor between BP and CP patients.

Ref	Disease Phase	Dead/ Alive	Sex	Age at Diagnosis (years)	Survival (months)	Best Response		KD Mutation	Therapy	rs11625196 (+81C/G)	rs7202 (+1004C/T)
						Achieved by patient	At date of sample				
1	BP-M	A	F	52	>69	CMR	-	NA	Imatinib-Nilotinib-Allo-SCT	CG	CC
2	BP-M	D	M	59	179	CHR	-	Wt	Myleran-HU-Imatinib	CC	CT
3	BP-L	A	M	51	>70	CMR	-	Wt	Imatinib-Dasatinib -Allo-SCT	CG	CC
4	BP-M	D	M	20	96	NA	-	G250E	HU-Imatinib-Nilotinib-Bafetinib-HHT	CC	CT
5	BP-M	A	M	56	>31	CMR	-	NA	Imatinib-Allo-SCT	CC	CC
6	BP-L	D	M	75	13	No CHR	-	L248V	Imatinib-Bafetinib –Dasatinib	CC	CT
7	BP-M	D	F	53	114	CCyR	-	Wt	IFN-Imatinib-IFN-Nilotinib-Dasatinib	CG	CC
8	BP-M	D	M	46	10	No CHR	-	E255V	Imatinib-Nilotinib-Allo-SCT	CC	CC
9	BP-M	D	M	22	38	No CHR	-	W478R	Imatinib	CC	CT
10	BP-L	D	M	48	32	NA	-	T315I	Imatinib-Dasatinib-Allo-SCT(x2)	CC	CT
11	BP-M	D	M	46	15	No CHR	-	T315I	Imatinib-Nilotinib	CC	CT
12	BP-M	D	F	66	48	NA	-	Wt	HU-Imatinib-Nilotinib	CC	TT
13	CP	A	F	45	>129	MMR	MMR	Wt	IFN-Imatinib	CC	CT
14	CP	A	M	48	>43	CHR	CHR	Wt	Imatinib-Nilotinib	CG	CC
15	CP	A	F	50	>49	MMR	diagn.	Wt	Imatinib	CC	CT
16	CP	A	F	62	>111	MMR	CCyR	H396R	HU-Imatinib-Bafetinib-Nilotinib-Dasatinib	CC	CT
17	CP	A	M	51	>93	MMR	MMR	Wt	Imatinib	GG	CC
18	CP	A	F	61	>113	CCyR	CHR	F317L	HU-Imatinib-Nilotinib-Dasatinib-Ponatinib	CC	CC
19	CP	A	F	20	>162	MMR	MMR	Wt	HU+IFN-Imatinib	CC	CC
20	CP	A	F	73	>51	MMR	MMR	Wt	Imatinib	CC	CT
21	CP	A	M	46	>48	MMR	MMR	NA	Nilotinib	CC	CC
22	CP	D	M	60	73	CHR	CHR	Wt	Imatinib-Nilotinib-Ponatinib	CG	CT

Table 5. Characteristics and observed single nucleotide polymorphisms of 22 CML patients involved in this study.

BP-M: myeloid blastic phase; BP-L: lymphoid blastic phase; CP: chronic phase; D: dead; A: alive; F: female; M: male; CHR: complete haematological response; CCyR: complete cytogenetic response; MMR: major molecular response; CMR: complete molecular response; KD: BCR-ABL1 kinase domain; NA: not analysed; wt: wild type; diagn.: at first diagnosis; HHT: homoharringtonine; therapy in chronological order.

3.3 +81 C/G genotype and clinical correlation

16 patients with CC genotype at position +81 had a median survival time of 48.5 months (range 10-179 months), of these 8 had died (median survival 35 months (10-179 months))(Figure 7). Of the 16 patients, 9 were in BP and had a median survival of 32.0 months (10-179 months). Only one of these 9 BP patients is alive and was among the three patients who had undergone allogeneic stem cell transplantation. The other 7 patients in CP are alive with median survival of 111 months (48-162 months).

Five patients had the heterozygous CG genotype of SNP +81 C/G with 70.0 months (43-114 months) median survival. Of these 5 patients, 2 died with survival of 114 and 73 months, one following progression to BP and the other while in CP, cause of death unclear. The remaining 2 BP patients are alive post allogeneic stem cell transplantation. The remaining CP patient with GG genotype is alive after 93.0 months.

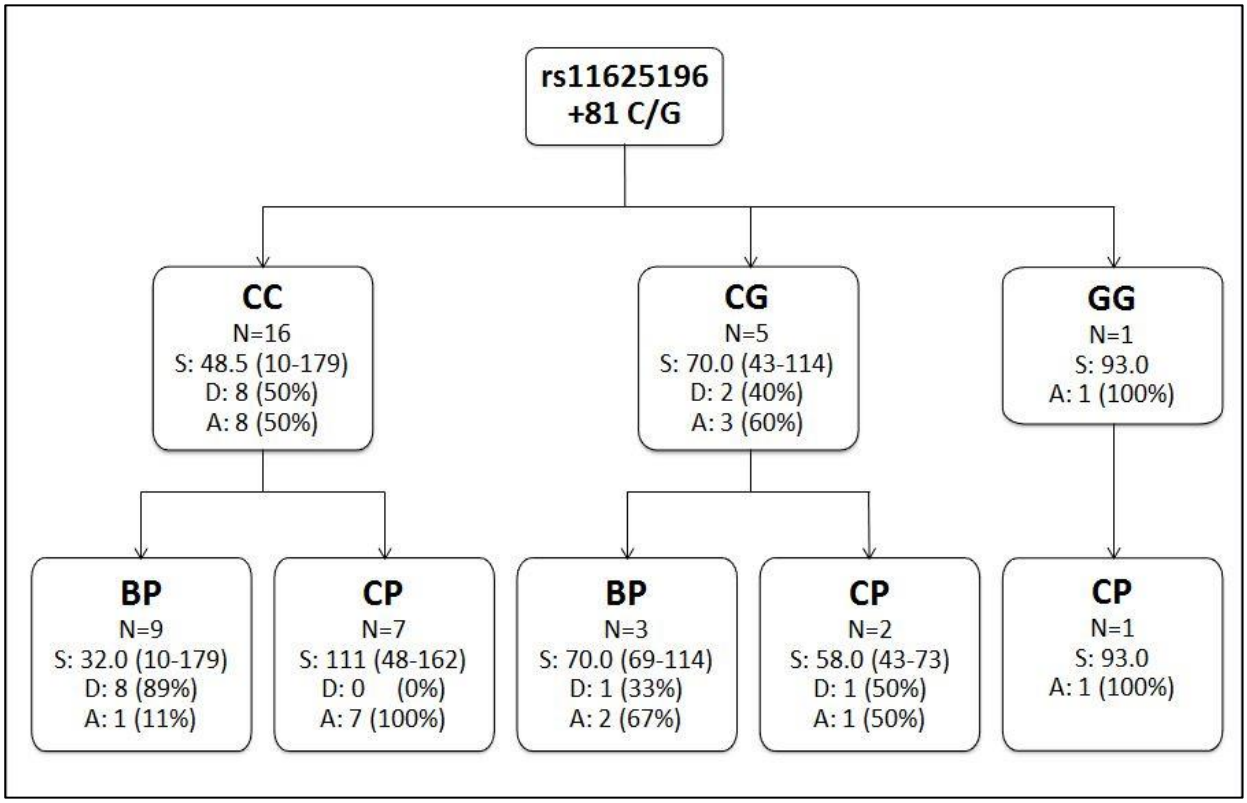

Figure 7. rs11625196 (+81 C/G) – occurrence of genotypes, survival in months & mortality.
S: survival time in months (range in months). D: number of dead patients. A: number of living patients.

Kruskal-Wallis-Test did not reveal a significant difference (p=0.116) in survival times between BP patients with CC genotype when compared with those with CG genotype. Similarly, Fisher's exact test did not reveal a significant difference when comparing mortality of 16 patients with CC genotype versus those with CG and GG genotype (p=0.299). Furthermore, no significant difference (p=0.123) was observed when comparing mortality among BP patients with CC genotype (n=9) versus those with CG genotype (n=3) when analysed with Fisher's exact test.

3.4 +1004 C/T genotype outcome analysis

10 patients with +1004 CC genotype had a median survival time of 69.5 months (range 10-162 months)(Figure 8). Two of these 10 died with survival times of 10 and 114 months (Figure 8). Of these 10 patients, 5 were in BP with a median survival of 69 months (10-114 months) and 2 out of 5 had died. Of this group, 4 underwent allogeneic stem cell transplantation and 3 of these are alive. The other 5 patients in CP are alive with a median survival of 95.0 months (43-162 months).

The median survival was determined to be 49 months (13-179 months) for 11 patients who were heterozygous (C/T) for +1004 SNP. Six of these 11 subjects were in BP with median survival of 35.0 months (13-179 months). Significantly, all 6 died, including one who had undergone Allo-SCT. Five of the 11 patients were in CP with median survival of 51 months (45-111 months). One patient in CP died for reasons unclear. In addition, the one patient with homozygous TT genotype died following progression to BP.

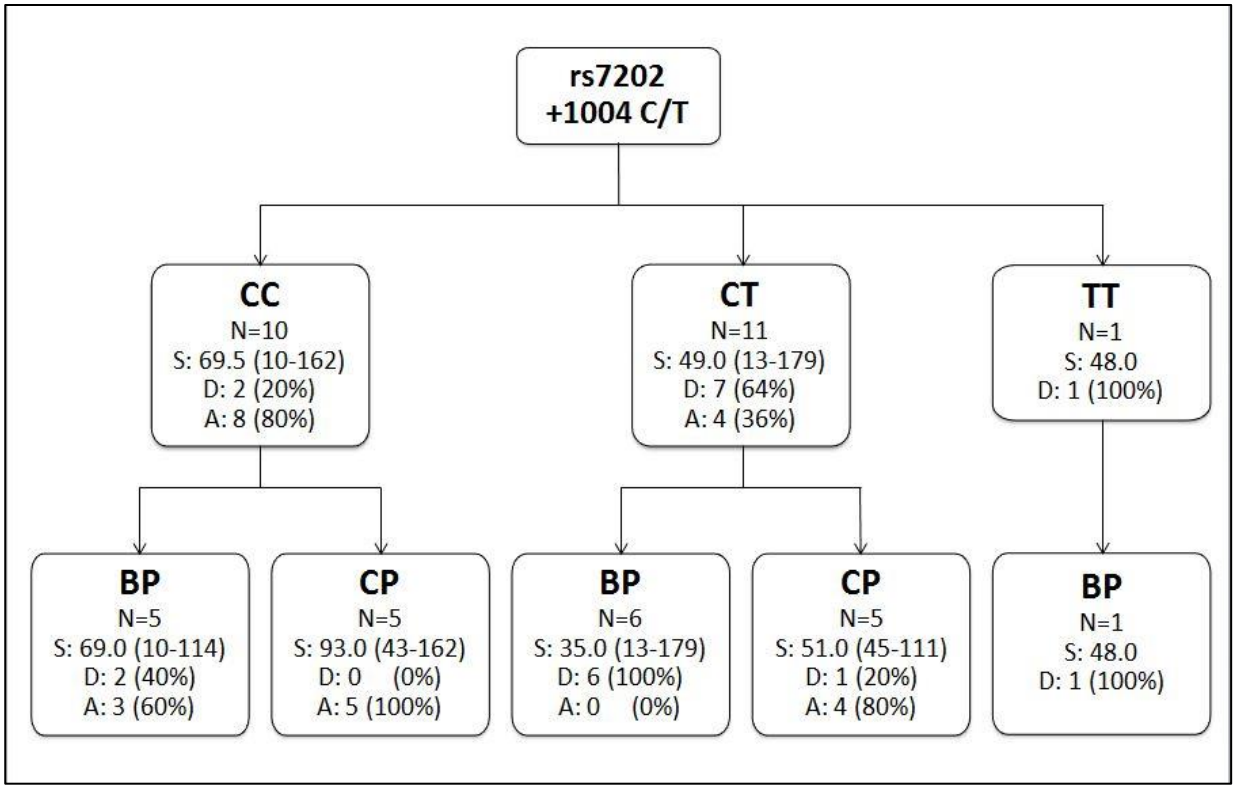

Figure 8. rs7202 (+1004 C/T) – occurrence of genotypes, survival times & mortality.
S: survival in months (range in months). D: proportion of dead patients. A: proportion of living patients.

Among the patients in BP, it was noted that survival rates were different for those who were homozygous, CC, when compared with those that were heterozygous, CT, or homozygous, TT, at position +1004 (Figure 9). Three of the 5 BP patients with CC genotype (BP-CC) are alive following Allo-SCT. In contrast, the 7 patients with CT (n=6) or TT (n=1) genotype died.

Fisher's exact test revealed the mortality frequency of the BP-CT group not to be significantly higher than the mortality among the BP-CC group (p=0.0606).
However, a comparison of mortality among the BP-CC group versus the BP-CT and BP-TT groups showed a significantly lower frequency among the CC group (Fisher's exact test, p=0.0455).

In contrast to mortality rates, survival times of BP patients harbouring the homozygous CC genotype were not significantly longer (Kruskal-Wallis test, p=0.573) than of patients with CT and TT genotype.

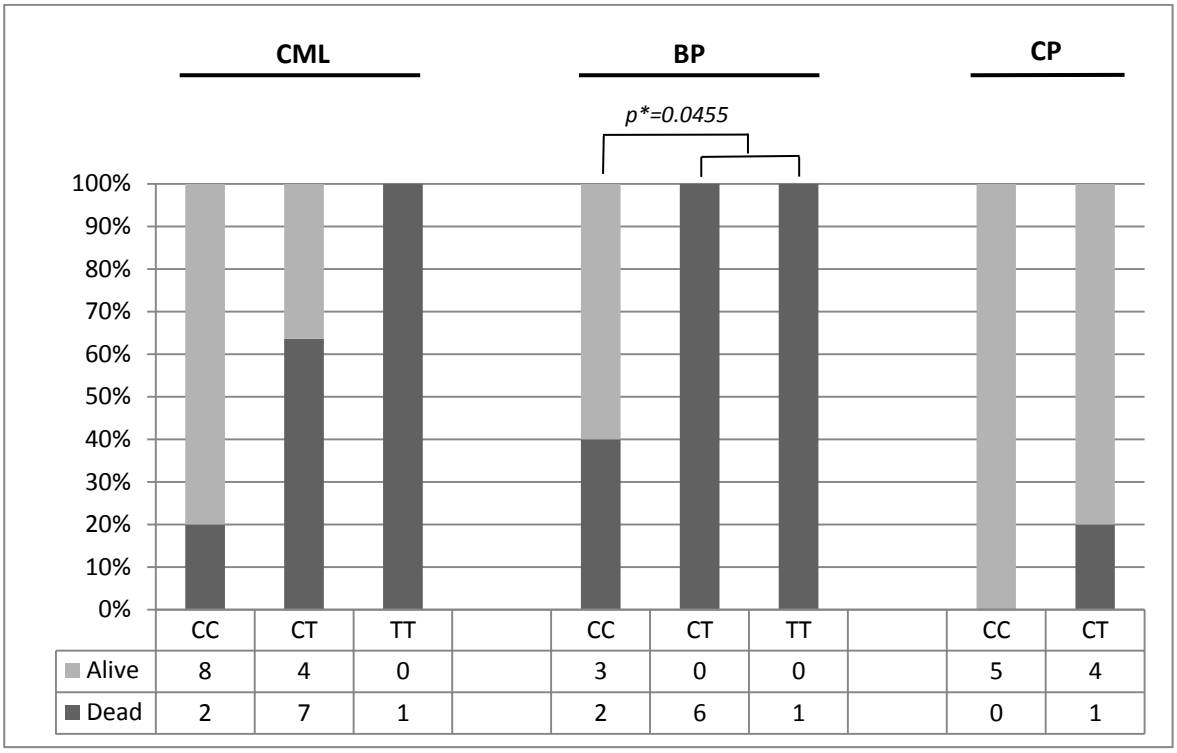

Figure 9. Comparison of mortality rates among the different genotypes of +1004 C/T – overall CML, BP and CP cohorts.

CML BP patients carrying homozygous CC genotype of SNP +1004 C/T have a significantly lower mortality than carriers of CT or TT genotype (p*=0.0455).

3.5 Bioinformatic estimation of SNP-induced conformational changes of *NUMB* mRNA

3.5.1 Simulation of *NUMB* mRNA secondary structure

To assess if the secondary structures of *NUMB* mRNA and *NUMB* mRNA 3'-UTR were influenced by the different genotypes of the two SNP +81 C/G and +1004 C/T, I carried out RNA secondary structure analysis using Mfold online web server (*RNA Institute, University of Albany, NY, USA*)[82]. I subjected the entire *NUMB* mRNA (3.6 kb) sequence to Mfold analysis to compare the predicted RNA secondary structures for the two alleles. The predicted structures were then screened by visual evaluation for any structural shifts that might enhance or attenuate MSI2 binding in the 3'-UTR of the *NUMB* mRNA.

Mfold analysis suggested 50 putative RNA secondary structures. These 50 possibilities were scored according to minimum free energy (ΔG) values. The predicted structures with smaller ΔG values, i.e. more negative, are considered to be more likely. The ΔG values of the first (most probable) and the last (least probable) proposed structure differed only by 3 to 5 per cent. These data imply all the predicted structures were equally possible.

The predicted structures of *NUMB* mRNA carrying major alleles of both SNP (+81 C & +1004 C, referred to as "wild type") and *NUMB* mRNA including +1004 C/T minor allele (+81 C & +1004 T) were completely identical to the order of the structures and the secondary structures as such (Figure 10). In contrast, when the minor allele of +81 SNP (+81 G & +1004 C) was analysed, only 13 of the 50 predicted mRNA secondary structures were identical to the wild type sequence structure. These data imply that a significant change in the mRNA secondary structure (Figure 11) is expected with the +81 SNP minor allele. However, the 3'-UTR secondary structure, which includes the putative MSI2 binding site, appeared to be unaffected in 76% (38 of 50) when compared to the structure associated with the wild type sequence.

The analysis showed that of the 50 predicted RNA secondary structures for minor allele of +81 SNP (+81 G & +1004 C) and the two minor alleles (+81 G & +1004 T) 45 were identical (Figure 12). These observations imply that the minor allele of +1004 SNP has little or no impact on the *NUMB* mRNA secondary structure. However, the software suggests that the minor allele of +81 SNP has a pronounced influence on global *NUMB* mRNA secondary structure, but putatively not around the 3'-UTR domain that includes the MSI2 binding site.

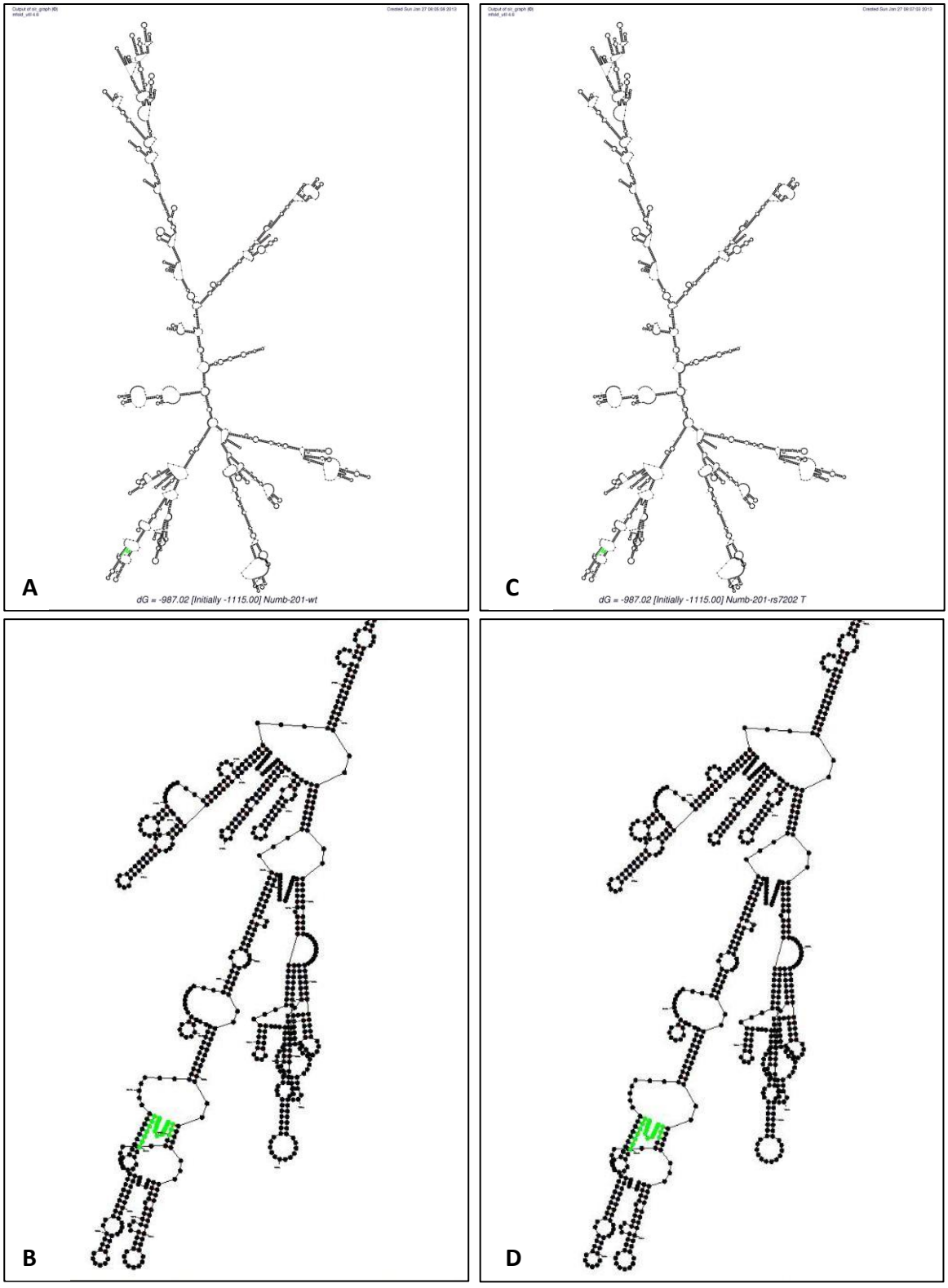

Figure 10. Comparison of *NUMB* mRNA secondary structures with minimal ΔG (A+C) and enlarged images of 3'-UTR (B+D). A+B: Wild type sequence. C+D: +1004 C/T minor allele +1004 T. Reported MSI2-binding site UAGGUU GUAGU UUUUU highlighted in green. +1004 T appears to have relatively little impact on *NUMB* mRNA secondary structure.

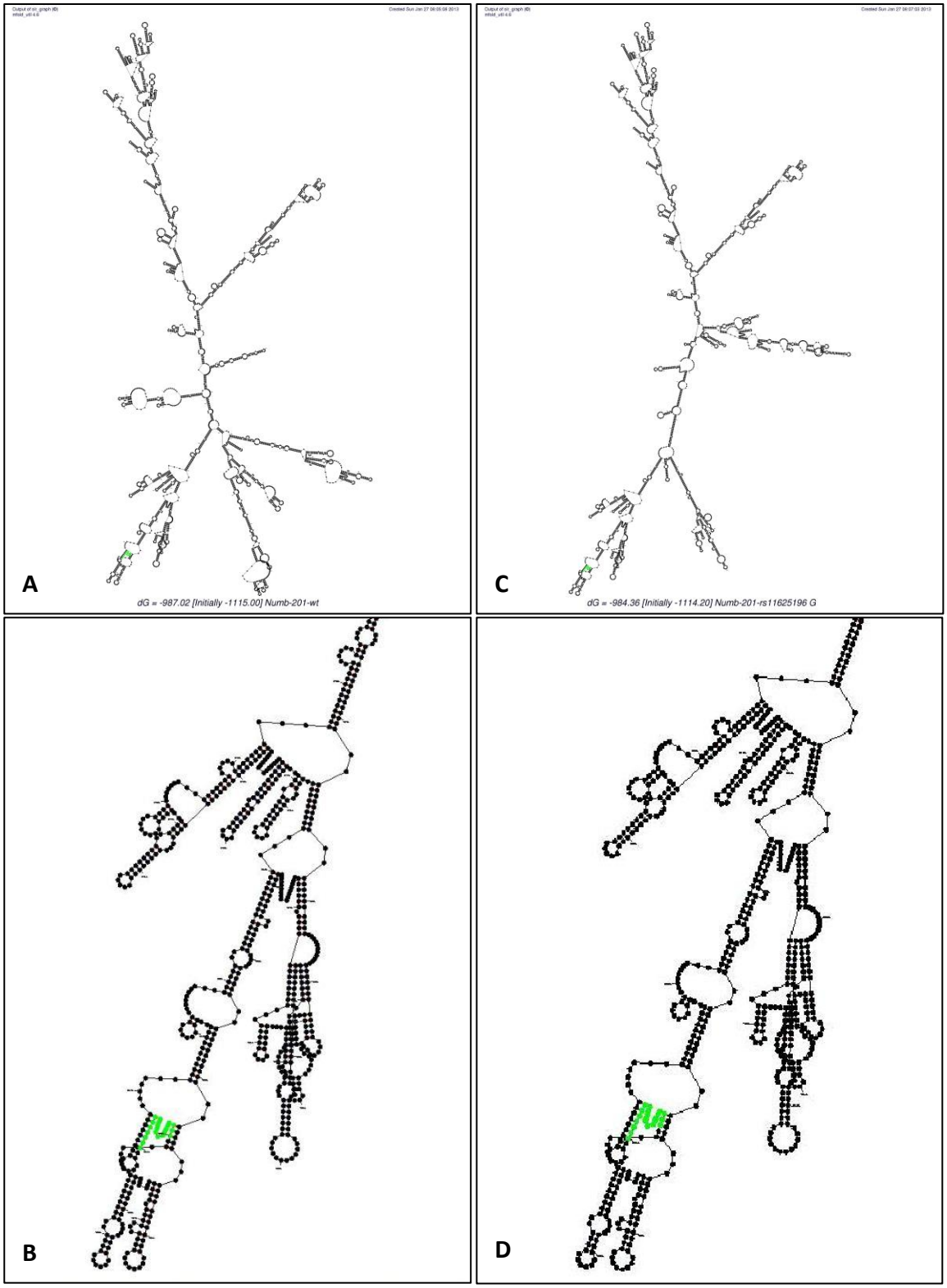

Figure 11. Comparison of *NUMB* mRNA secondary structures with minimal ΔG (A+C) and enlarged images of 3'-UTR (B+D). A+B: Wild type sequence. C+D: +81 C/G minor allele +81 G. Reported MSI2-binding site UAGGUU GUAGU UUUUU highlighted in green.
+81 G putatively has a relatively strong effect on global mRNA secondary structure but leaves 3'-UTR and MSI2-binding site unaltered.

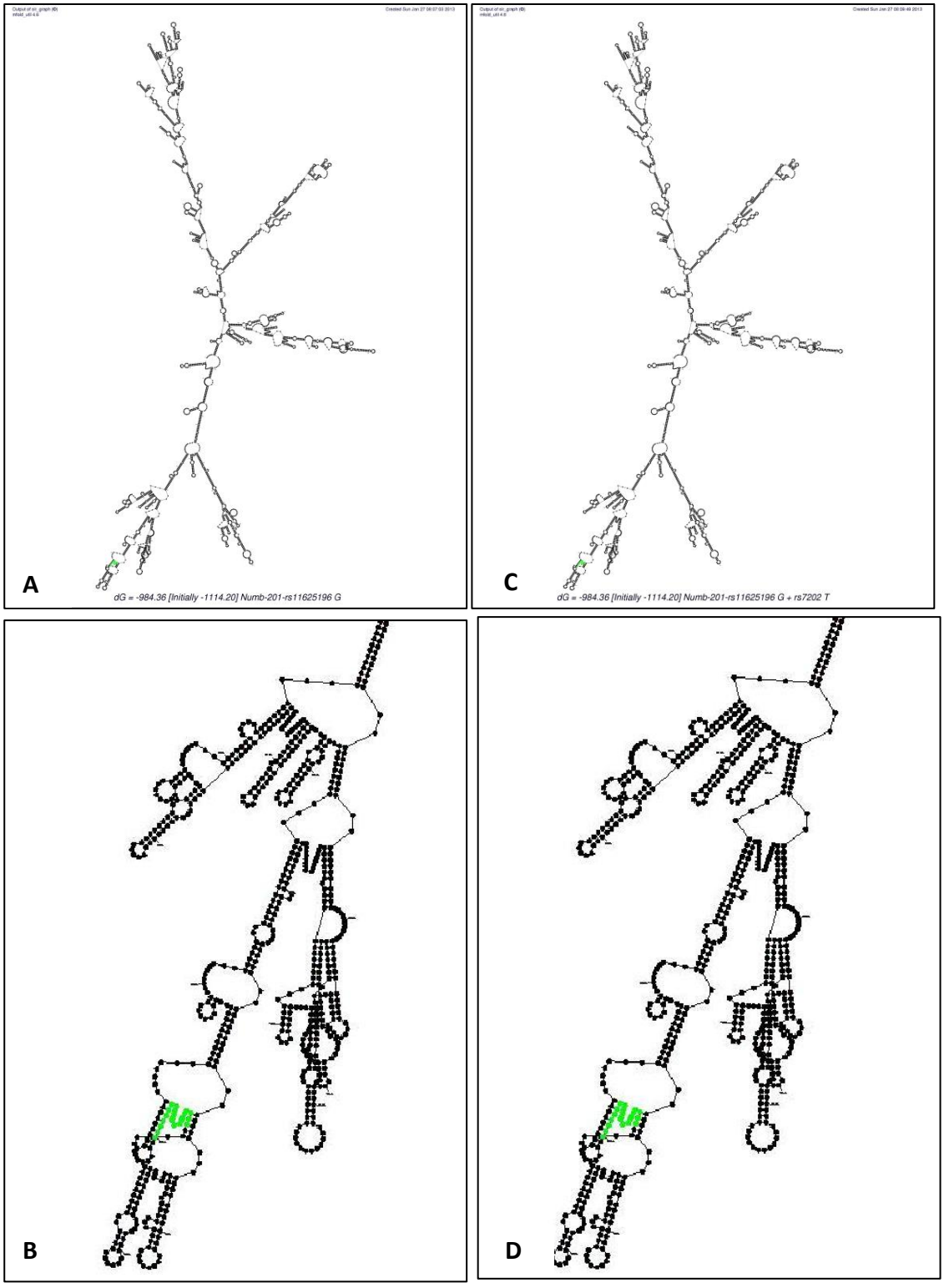

Figure 12. Comparison of *NUMB* mRNA secondary structures with minimal ΔG (A+C) and enlarged images of 3'-UTR (B+D). A+B: +81 G and +1004 C. C+D: +81 G and +1004 T. Reported MSI2-binding site UAGGUU GUAGU UUUUU highlighted in green. +1004 T does not appear to lead to additional structural shifts in comparison to +81 G alone.

3.5.2 Prediction of SNP-induced conformational changes of *NUMB* mRNA

To assess the reliability of data obtained with Mfold software I subjected the sequence data to the SNPfold software programme. This on-line analysis algorithm estimates the potential consequences of SNP on RNA secondary structure. It also compares and estimates the probabilities of base-pairing between the wild type RNA and the RNA with the minor allele of a SNP.

SNPfold algorithm calculates the correlation coefficient, rank and p-value for each data. The correlation coefficient quantifies the level of predicted change by the SNP, so that values close to 1 represent minimally altered secondary structure. For the *NUMB* sequences, SNPfold algorithm identified the correlation coefficient for 6000 different simulated mutations within the RNA strand and subsequently compared them to the correlation coefficient for my CML patients' sequences. The correlation coefficient for my *NUMB* data was ranked among the 6000 values and the p-value calculated through division of rank by 6000. The calculated SNPfold data is represented graphically as a line chart showing the allocation for each of the RNA nucleotide positions with respect to average changes of the base-pairing probabilities (the partition function) between wild type and SNP-carrying RNA, based on the values of the average change in partition function. This line chart illustrates in which domains of the two RNA sequences and by how much the two RNA secondary structures differ from each other.

3.5.2.1 Comparison of *NUMB* wild type mRNA and mRNA with +1004 T SNP

Figure 13 illustrates the line chart comparing *NUMB* wild type mRNA (+81 C & +1004 C) and RNA with the minor SNP variant +1004 T. The minor allele of this SNP appears to have minimal impact on mRNA secondary structure. This is consistent with 0.999989 correlation coefficient, rank of 5880/6000 and p-value of 0.9800. The average change in partition function was estimated to be zero between the nucleotides 0 and approximately 2700 and between approximately 3400 and 3593. Furthermore, the average change of base-pairing probabilities between residues 2700 and 3400 was determined to be ≤0.012, which is negligible. Hence, the potency of +1004 SNP minor allele to alter *NUMB* mRNA secondary structure significantly is minimal.

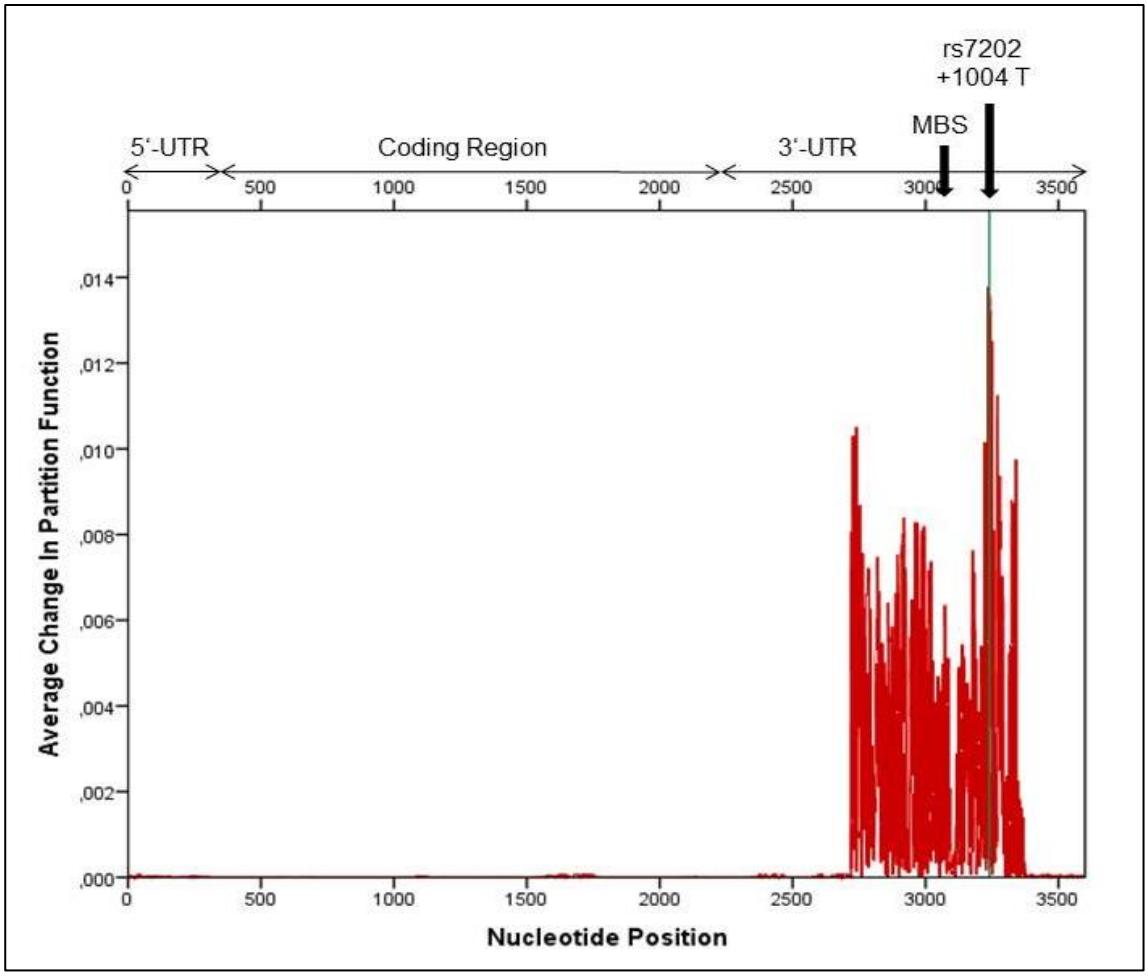

Figure 13. Line chart representing the average change in partition function column sum per residue in the RNA strand. Comparison of *NUMB* wild type mRNA (+81 C & + 1004 C) and mRNA with +1004 T. Arrows: MBS – MSI2-binding site (nucleotide position 2980-3000), +1004 T – nucleotide position 3141.

3.5.2.2 Comparison of *NUMB* wild type mRNA and mRNA with +81 G SNP

In contrast to SNP +1004 C/T, the potency of +81 SNP minor variant to induce changes in *NUMB* mRNA secondary structure is predicted to be relatively more pronounced. The correlation coefficient was calculated to be 0.997985, with 4978/6000 rank and p-value of 0.8296. Consistent with Mfold software predictions for my data, shifts in secondary structure tend to be localized to the coding region and nucleotides 2250-2400 of 3'-UTR but do not include the major segments of 3'-UTR.

Figure 14 shows the corresponding line chart which illustrates that the likelihood of an effect on the secondary structure is highest in 3 distinct regions of the RNA sequence – approximately between residues 1750 and 1900, 2100 and 2200 and between 2250 and 2400. These three

regions represent 3'-sections of the coding region and marginal 5'-sections of 3'-UTR, respectively. In these sections, average changes in the partition function range from 0.01 up to 0.06. Probability of variation from the structure for the wild type sequence is the least between residues 0 and 1750 and between 1900 and 2100. Importantly, average change in partition function is negligible between residues 2400 and 3593, representing the core part of 3'-UTR containing MSI2-binding site.

These observations are consistent with the minor variant of +81 SNP having a distinct impact on *NUMB* mRNA secondary structure, but not on the major functional segment of the corresponding 3'-UTR. There is no evidence to suggest that the interaction between MSI2 protein and *NUMB* mRNA is altered by the +81 G SNP.

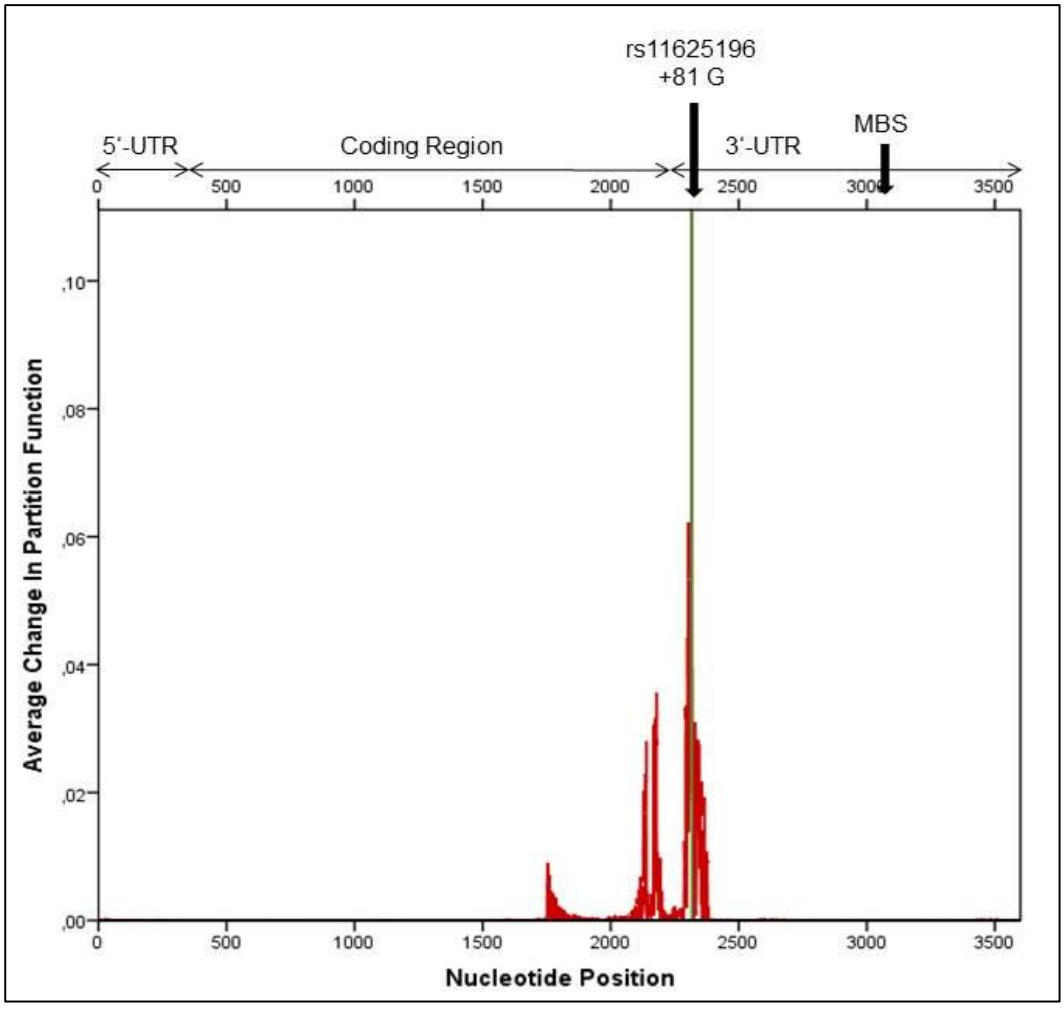

Figure 14. Line chart representing the average change in partition function column sum per residue in the RNA strand. Comparison of *NUMB* wild type mRNA (+81 C & +1004 C) and mRNA with +81 G and +1004 C. Arrows: MBS – MSI2-binding site (nucleotide position 2980-3000), +81 G – nucleotide position 2217.

3.5.2.3 Comparison of *NUMB* wild type mRNA and mRNA with +81 G and +1004 T

When the minor alleles +81 G and +1004 T are in *cis*, the additive effects on the partition function are shown in Figure 15. The line chart is similar to the data illustrated in Figure 13 and Figure 14. In contrast to the line chart showing the impact of +81 G alone (Figure 14), the small impact of +1004 T on residues 2500 to 3400 was added. These findings are reflected by statistical analysis data, with correlation coefficient of 0.997974, rank of 4975 and p-value of 0.8291.

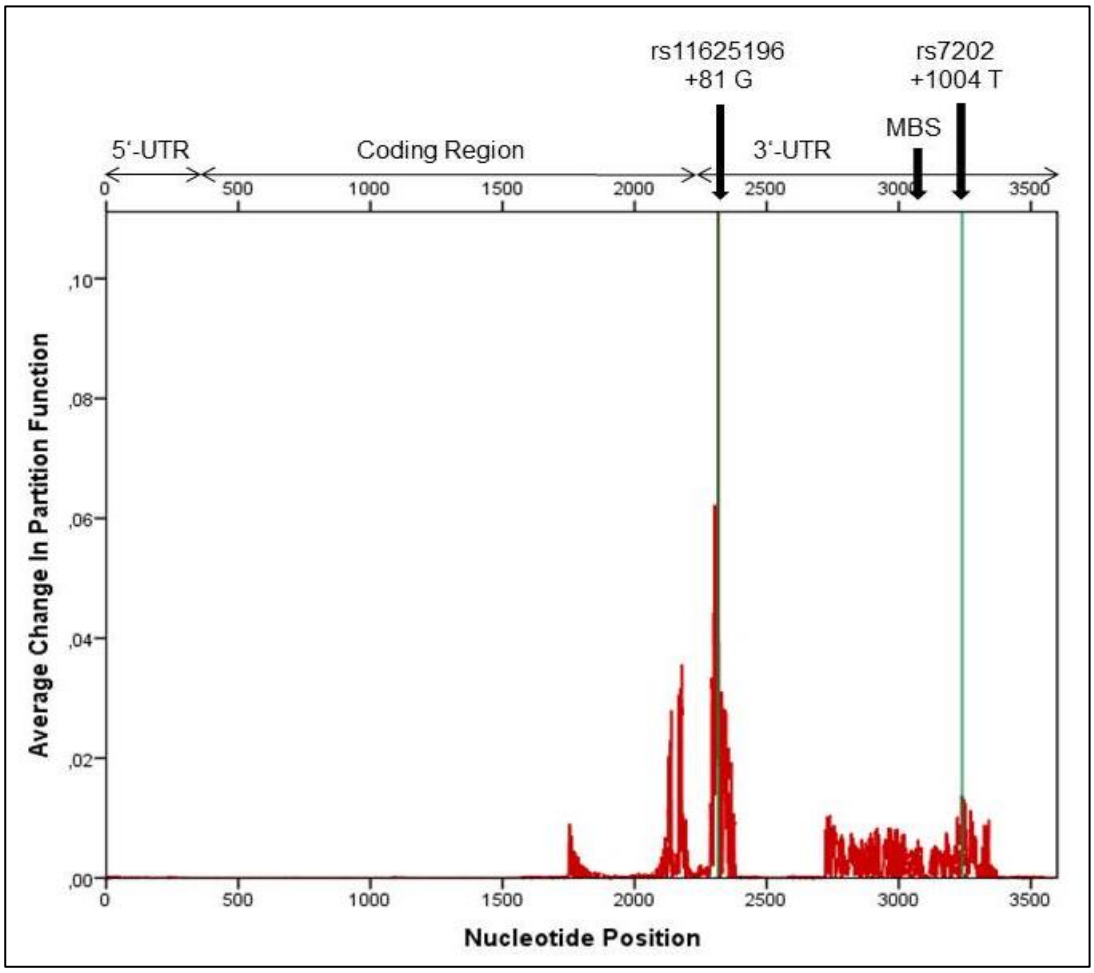

Figure 15. Line chart representing the average change in partition function column sum per residue in the RNA strand. Comparison of *NUMB* wild type mRNA (+81 C & +1004 C) and mRNA with both minor SNP variants +81 G and +1004 T. Arrows: MBS – MSI2-binding site (nucleotide position 2980-3000), +81 G – nucleotide position 2217, +1004 T – nucleotide position 3141.

4. Discussion

4.1 Role of NUMB and the Musashi2-NUMB-Notch signalling pathway in advanced-stage CML

The events triggering arrested differentiation and a more aggressive disease in chronic myeloid leukaemia (CML) patients are unclear. Dysregulation of *MSI2* has been suggested as a causal event in the transformation of chronic phase (CP), a relatively indolent disease phase, to blast phase (BP), which is usually fatal. The Musashi gene family, regulated by *HOXA9*, has been shown to control critical cell fate decisions by binding to the 3'-untranslated region of target mRNAs, thereby inhibiting translation. This results in dysfunction of the regulatory pathway, leading to haematopoietic stem cell (HSC) proliferation, impaired myeloid differentiation and worse clinical prognosis in CML and AML[1-3]. These events are reported to be regulated by MSI2 via NUMB-Notch signalling pathway.

Indeed, *NUMB* expression was found to be diminished while *NOTCH* levels were increased, i.e. there is an inverse relationship between *NUMB* and *NOTCH* expression[2].

These observations imply that a dysregulation of NUMB-Notch signalling leads to disease progression. The underlying cause of the dysregulation is likely to vary from patient to patient, consistent with heterogeneity of CML. It is conceivable that mutations within coding region or SNP within 3'-UTR might impair NUMB protein's function as an inhibitor of Notch and Hedgehog signalling or might enhance or attenuate MSI2 protein mediated inhibition of *NUMB* translation. An enhancement of this interaction and increased NUMB decay could be a cause for blast phase. However, this study found no evidence to support the notion that mutated *NUMB* might be the primary cause of CML disease progression in some patients.

Ito et al. (2010) found that the homeodomain-containing protein HOXA9 worked as a transcriptional activator of *MSI2* gene. The authors concluded this from observations that the fusion gene *NUP98-HOXA9 in-vivo* induced an upregulation of *HOXA9*, which correlated with increased expression of *MSI2*[(2)]. Furthermore, HOXA9 is reported to up-regulate genes which trigger proliferation and to inhibit expression of genes involved in myeloid differentiation[84]. *NUP98-HOXA9* has been observed in a sub-set of AML patients and CML patients who progressed to myeloid BP as a consequence of the acquisition of the additional cytogenetic aberration t(7;11)(p15;p15)[85]. Interestingly, in the majority of CML-BP patients, *HOXA9*

expression was found to be elevated which may be the causal agent of the elevated *MSI2* levels in BP[2].

4.2 Interpretation of the results

The present study aimed to assess whether the mutated *NUMB* is involved in CML disease progression. This question was addressed by subjecting cDNA synthesised from nuclear pellet isolated from 12 CML patients in BP and 10 CML patients in CP to Sanger dideoxy chain termination sequencing.

This study observed no mutations in the sequenced *NUMB* transcripts of 22 patients screened. These data suggest that mutations mapping to the *NUMB* gene are very rare in CML patients and therefore are unlikely to be directly involved in blastic transformation. Furthermore, as far as it can be ascertained, no *NUMB* mutations have been documented. These observations suggest NUMB protein which regulates Notch, Hedgehog and p53 signalling is not the primary cause of CML evolution[8-11].

While no mutations were detected, I did observe two SNP within the 3'-UTR region of *NUMB* gene transcript. This was not entirely unexpected, as the minor alleles of these two SNP, from among the 20 SNP mapping to the 3'-UTR listed by Ensembl, are reported to occur with frequencies of 26.1% (rs7202) and 15.0% (rs11625196)[6,7]. I failed to detect any of the other reported SNP probably because of the low frequencies with which they occur, <2% for 16 and <4% for the remaining 2 cases[86].

The frequencies for SNP rs7202 +1004 C/T and SNP rs11625196 +81 C/G do not statistically differ between patients in CP, BP and normal healthy controls, which implies they do not influence CML disease progression.

Analysis of the CML patients' clinical data – survival times and mortality – supported this notion with regards to +81 SNP. But, for the +1004 SNP, the mortality rate in BP was found to be significantly lower in patients homozygous for CC, in contrast to those who were either heterozygous (CT) or homozygous for TT. However, of the 5 patients with CC genotype, 4 had undergone Allo-SCT and of these 3 are alive. In contrast, only 1 patient of the 7 with CT or TT genotype had Allo-SCT and subsequently died. These observed differences between CC, CT and TT genotype are likely to be biased by the small number of patients and the different treatment modalities.

As it is generally accepted that single nucleotide variations (SNV) within the mRNA 3'-UTR can

cause dysregulation of translation and disease, e.g. by affecting polyadenylation signal or secondary structure, I investigated the possibility that the two *NUMB* 3'-UTR SNP I detected altered the secondary structure of mRNA and as a consequence modulated the interaction of Musashi-2 protein and *NUMB* transcripts[87].

To test this notion I examined the *NUMB* transcript sequences I obtained using web based Mfold and SNPfold software. These programmes predicted small changes in the *NUMB* mRNA secondary structure as a consequence of the two SNP detected by Sanger sequencing. In particular, the +1004 genotype had relatively weak influence on the secondary structure of 3'-UTR and negligible influence on the coding region and 5'-UTR (Figure 13). The +81 genotype had moderate influence on the secondary structure in the immediate surrounding region, but no or negligible influence on the remaining 5'-UTR, coding sequence region and the far removed 3'-UTR (Figure 14). It was estimated that a steric change within the MSI2-binding site for both SNP was <0.01.

In summary, the combination of the observed frequencies of the two SNP and their modest potency to induce steric NUMB mRNA alterations would suggest that the two variations neither account for the statistically different mortality rates in blast phase nor possess a disease-driving function. However, given the small sample population size, it is difficult to draw any definite conclusions. A study with a greater number of CML CP and BP patients is required to test the influence of the 2 SNP identified in this study.

4.3 Critical appraisal of the results

4.3.1 Experimental design

While the function and the expression patterns of the reported *NUMB* splice variants are equivocal, this study and the data obtained raise some questions. For example, in contrast to Ensembl database, NCBI database lists only four different *NUMB* splice variants[88]. More importantly, the cDNA sequences do not correlate with Ensembl-listed splice variants. This is illustrated by the fact that it is not possible to assign the unexpected band obtained upon amplification of fragment A (see Figure 4), 0.7kb in length, to any of the four NCBI Gene-listed variants. PCR amplification of fragments B and C yielded only one band. The DNA sequence data obtained were compared with published reference sequences using the basic local

alignment search tool (BLAST), which utilises DNA sequences of NCBI Gene database as reference.

As a consequence, it is unclear which splice variants I amplified by PCR and sequenced, and whether I failed to amplify transcript types that might have been of interest for my study (see Table 3). However, I am confident that I amplified and sequenced the transcripts that were among the first to be reported. In retrospect it would have been prudent to sequence the 5'-UTR also, rather than limiting the sequencing to the coding region and in particular the 3'-UTR. The latter was of particular interest because MSI2 is reported to bind to the 3'-UTR and thereby regulate *NUMB* expression at post transcript level.

Sanger dideoxy chain termination sequencing is a reliable and robust method with sensitivity of approximately 20-25% and until recently the method of choice for sequencing. However, in recent year a more rapid sequencing with higher level of sensitivity, up to 0.1%, is now possible with second generation sequencing, i.e. deep sequencing. Thus, it might be prudent to subject the samples I studied to this new technology to exclude mutations that might be present at a much lower level than the 20-25% detection limit of Sanger sequencing. However, equally a driver mutation is likely to be present in >20% of the CML clone in patients in BP, if it is responsible for the transformation.

4.3.2 Analysis of the results

The Mfold software I used to visually compare approximately fifty predicted secondary structures as consequence of the detected SNP is an excellent qualitative simulation software tool. The second software I used to process the NUMB sequencing data, SNPfold, is more quantitative and therefore less subjective, when assessing the effects of SNP on mRNA secondary structure. Nevertheless, it is unclear as to how reliable these secondary structure predictions reflect the secondary structures that may arise *in-vivo*. However, given these limitations, the software tools do suggest that the probability that the SNP rs7202 and rs11625196 alter NUMB RNA secondary structure is very low.

4.4 Possible future research focus areas

The assumption that in a certain proportion of CML BP patients elevated *HOXA9* levels lead to up-regulated *MSI2* raises the question as to the mechanisms causing overexpression of the former.

An up-regulated *HOXA9* might be explained by recent acute myeloid leukaemia studies. Several potential signaling pathways have been described which could result in the activation of *HOXA* genes such as *HOXA9* (see Figure 16) and subsequent leukaemic stem cell (LSC) immortalization[84,89-92].

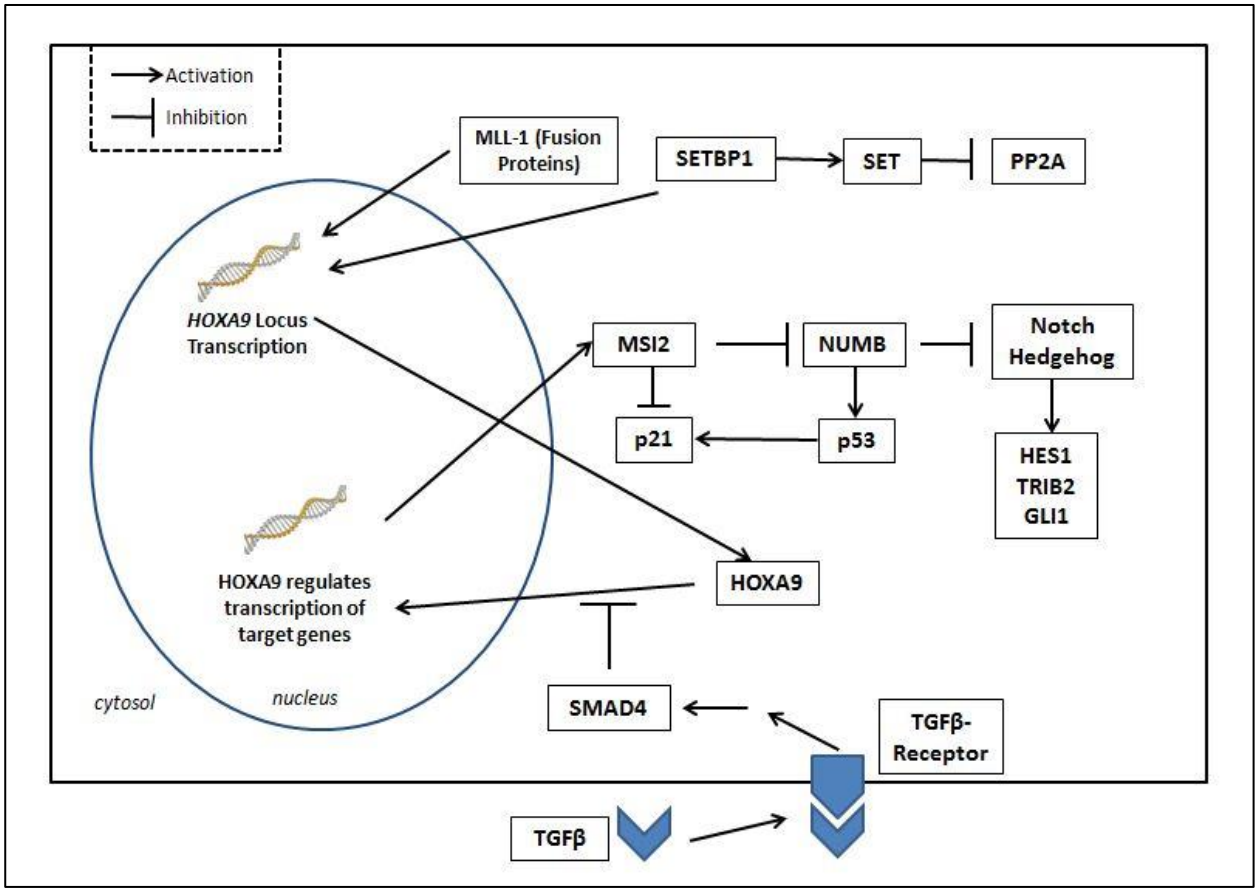

Figure 16. Compilation of recently published molecular interactions which might regulate *HOXA9* expression and contribute to CML LSC immortalization.

HOXA9 transcription is reported to be activated by MLL-1 histone methyltransferase and SETBP1, HOXA9 to induce expression of *MSI2*, SMAD-4 to impede transfer of cytosolic HOXA9 to the nucleus. As a stabilizer of SET, SETBP1 is believed to suppress activity of tumour suppressor PP2A.

SET binding protein-1 (SETBP1) and mixed lineage leukemia-1 (MLL-1) proteins are assumed to work as activators of *HOXA9* transcription[89,90]. Also, SETBP1 has been reported to stabilize SET, a potent inhibitor of tumour suppressor PP2A (protein phosphatase 2A)[93,94]. Moreover, acting together with BCR-ABL1, SETBP1 was shown to transform primary myeloid progenitors *in-vivo* to induce fatal myeloid BP-like leukaemia[89]. Concomitant *SETBP1* and *HOXA9* overexpression was observed in 15% of CML AD patients and *SETBP1* overexpression in more than 27% of AML patients[89,94]. Furthermore, MLL-1, a histone-methyltransferase, is reported to activate *HOXA9* transcription by specific histone methylation of the *HOXA9* gene locus[90]. This finding could explain the leukaemogenic effect of cytogenetic MLL-1 rearrangements which are related to a subgroup of AML cases[95]. Other investigators have reported a TGF-β signalling downstream effector, SMAD-4, a tumour suppressor that forms a complex with HOXA9 in the cytosol. The formation of the complex impedes translocation of HOXA9 to the nucleus[91]. In addition, other authors emphasize the function of Micro RNA-181 (*miR-181*) as a suppressor of *HOX* gene signalling[92].

In summary, several mechanisms have been described which might explain activation of *HOXA9* signalling in myeloid leukaemia. These mechanisms may also lead to *HOXA9* overexpression, thereby increasing *MSI2* levels. This sequence of events down-regulates *NUMB* and activates Notch. But as CML is highly heterogeneous, this chain of events most probably only contributes to CML blastic transformation in a subset of patients.

Finally, the precise mechanism or mechanisms leading to CML disease progression remain unclear. Studies specifically addressing the role of HOXA9 and MSI2 in haematopoietic cells may lead to a better understanding of how this pathway interacts with BCR-ABL1 resulting in CML transformation. A better understanding may also identify suitable therapeutic targets, such as HOXA9, MSI2, SETBP1 or NOTCH. Finally, longitudinal studies involving quantification of *MSI2* expression in $CD34^+$ enriched samples may help to assess if it might be a useful prognostic marker to identify patients at risk of disease progression. This would enable clinicians to intervene early in the course of the disease, before patients become refractory to further treatment.

5. Table of References

1. Kaeda JS, Mills KI, Kharas MG, et al. Increased *MSI2* expression Is Associated with Aggressive CML and AML. Blood. 2011; 118 (Supplement). Abstract 2516.2011.
2. Ito T, Kwon HY, Zimdahl B, et al. Regulation of myeloid leukaemia by the cell-fate determinant Musashi. Nature 2010;466:765-8.
3. Kharas MG, Lengner CJ, Al-Shahrour F, et al. Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. Nat Med 2010;16:903-8.
4. Imai T, Tokunaga A, Yoshida T, et al. The neural RNA-binding protein Musashi1 translationally regulates mammalian numb gene expression by interacting with its mRNA. Mol Cell Biol 2001;21:3888-900.
5. Kawahara H, Imai T, Imataka H, Tsujimoto M, Matsumoto K, Okano H. Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. J Cell Biol 2008;181:639-53.
6. Reference SNP(refSNP) Cluster Report: rs11625196. NCBI. (Accessed 2012-01-16, 2012, at http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs11625196.)
7. Reference SNP(refSNP) Cluster Report: rs7202. NCBI. (Accessed 2012-01-16, 2012, at http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs7202.)
8. McGill MA, Dho SE, Weinmaster G, McGlade CJ. Numb regulates post-endocytic trafficking and degradation of Notch1. J Biol Chem 2009;284:26427-38.
9. McGill MA, McGlade CJ. Mammalian numb proteins promote Notch1 receptor ubiquitination and degradation of the Notch1 intracellular domain. J Biol Chem 2003;278:23196-203.
10. Di Marcotullio L, Ferretti E, Greco A, et al. Numb is a suppressor of Hedgehog signalling and targets Gli1 for Itch-dependent ubiquitination. Nat Cell Biol 2006;8:1415-23.
11. Colaluca IN, Tosoni D, Nuciforo P, et al. NUMB controls p53 tumour suppressor activity. Nature 2008;451:76-80.
12. Vardiman JW, Melo JV, Baccarani M, Thiele J. Chronic myelogenous leukaemia, BCR-ABL1 positive. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th ed2008:32-7.
13. Rohrbacher M, Hasford J. Epidemiology of chronic myeloid leukaemia (CML). Best Pract Res Clin Haematol 2009;22:295-302.
14. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature 1973;243:290-3.
15. Nowell PC, Hungerford, D. A. A minute chromosome in human chronic granulocytic leukemia. Science 132: 14971960.
16. Shtivelman E, Lifshitz B, Gale RP, Canaani E. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature 1985;315:550-4.
17. Leukemia, chronic myeloid; CML. (Accessed 2013-02-16, 2013, at http://omim.org/entry/608232.)
18. Baccarani M, Deininger MW, Rosti G, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. Blood 2013;122:872-84.
19. Druker BJ, Talpaz M, Resta DJ, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med 2001;344:1031-7.

20. Ferdinand R, Mitchell SA, Batson S, Tumur I. Treatments for chronic myeloid leukemia: a qualitative systematic review. J Blood Med 2012;3:51-76.
21. Hochhaus A, O'Brien SG, Guilhot F, et al. Six-year follow-up of patients receiving imatinib for the first-line treatment of chronic myeloid leukemia. Leukemia 2009;23:1054-61.
22. Larson RA, Hochhaus A, Hughes TP, et al. Nilotinib vs imatinib in patients with newly diagnosed Philadelphia chromosome-positive chronic myeloid leukemia in chronic phase: ENESTnd 3-year follow-up. Leukemia 2012;26:2197-203.
23. Mahon FX, Réa D, Guilhot J, et al. 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 Oncol 2010;11:1029-35.
24. Ernst T, Hochhaus A. Chronic myeloid leukemia: clinical impact of BCR-ABL1 mutations and other lesions associated with disease progression. Semin Oncol 2012;39:58-66.
25. Chen Y, Peng C, Sullivan C, Li D, Li S. Critical molecular pathways in cancer stem cells of chronic myeloid leukemia. Leukemia 2010;24:1545-54.
26. Corbin AS, Agarwal A, Loriaux M, Cortes J, Deininger MW, Druker BJ. Human chronic myeloid leukemia stem cells are insensitive to imatinib despite inhibition of BCR-ABL activity. J Clin Invest 2011;121:396-409.
27. Chomel JC, Turhan AG. Chronic myeloid leukemia stem cells in the era of targeted therapies: resistance, persistence and long-term dormancy. Oncotarget 2011;2:713-27.
28. Skorski T. Genetic mechanisms of chronic myeloid leukemia blastic transformation. Curr Hematol Malig Rep 2012;7:87-93.
29. Grossmann V, Kohlmann A, Zenger M, et al. A deep-sequencing study of chronic myeloid leukemia patients in blast crisis (BC-CML) detects mutations in 76.9% of cases. Leukemia 2011;25:557-60.
30. Nowicki MO, Falinski R, Koptyra M, et al. BCR/ABL oncogenic kinase promotes unfaithful repair of the reactive oxygen species-dependent DNA double-strand breaks. Blood 2004;104:3746-53.
31. Koptyra M, Falinski R, Nowicki MO, et al. BCR/ABL kinase induces self-mutagenesis via reactive oxygen species to encode imatinib resistance. Blood 2006;108:319-27.
32. Gaymes TJ, Mufti GJ, Rassool FV. Myeloid leukemias have increased activity of the nonhomologous end-joining pathway and concomitant DNA misrepair that is dependent on the Ku70/86 heterodimer. Cancer Res 2002;62:2791-7.
33. Brady N, Gaymes TJ, Cheung M, Mufti GJ, Rassool FV. Increased error-prone NHEJ activity in myeloid leukemias is associated with DNA damage at sites that recruit key nonhomologous end-joining proteins. Cancer Res 2003;63:1798-805.
34. Slupianek A, Nowicki MO, Koptyra M, Skorski T. BCR/ABL modifies the kinetics and fidelity of DNA double-strand breaks repair in hematopoietic cells. DNA Repair (Amst) 2006;5:243-50.
35. Byers RJ, Currie T, Tholouli E, Rodig SJ, Kutok JL. MSI2 protein expression predicts unfavorable outcome in acute myeloid leukemia. Blood 2011.
36. Bandziulis RJ, Swanson MS, Dreyfuss G. RNA-binding proteins as developmental regulators. Genes Dev 1989;3:431-7.
37. Sakakibara S, Nakamura Y, Satoh H, Okano H. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci 2001;21:8091-107.

38. Okano H, Kawahara H, Toriya M, Nakao K, Shibata S, Imai T. Function of RNA-binding protein Musashi-1 in stem cells. Exp Cell Res 2005;306:349-56.
39. Sugiyama-Nakagiri Y, Akiyama M, Shibata S, Okano H, Shimizu H. Expression of RNA-binding protein Musashi in hair follicle development and hair cycle progression. Am J Pathol 2006;168:80-92.
40. Sakakibara S, Nakamura Y, Yoshida T, et al. RNA-binding protein Musashi family: roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. Proc Natl Acad Sci U S A 2002;99:15194-9.
41. Sakakibara S, Imai T, Hamaguchi K, et al. Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. Dev Biol 1996;176:230-42.
42. Sakakibara S, Okano H. Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. J Neurosci 1997;17:8300-12.
43. Akasaka Y, Saikawa Y, Fujita K, et al. Expression of a candidate marker for progenitor cells, Musashi-1, in the proliferative regions of human antrum and its decreased expression in intestinal metaplasia. Histopathology 2005;47:348-56.
44. Kayahara T, Sawada M, Takaishi S, et al. Candidate markers for stem and early progenitor cells, Musashi-1 and Hes1, are expressed in crypt base columnar cells of mouse small intestine. FEBS Lett 2003;535:131-5.
45. Potten CS, Booth C, Tudor GL, et al. Identification of a putative intestinal stem cell and early lineage marker; musashi-1. Differentiation 2003;71:28-41.
46. Clarke RB, Spence K, Anderson E, Howell A, Okano H, Potten CS. A putative human breast stem cell population is enriched for steroid receptor-positive cells. Dev Biol 2005;277:443-56.
47. Li D, Peng X, Yan D, et al. Msi-1 is a predictor of survival and a novel therapeutic target in colon cancer. Ann Surg Oncol 2011;18:2074-83.
48. Wang XY, Penalva LO, Yuan H, et al. Musashi1 regulates breast tumor cell proliferation and is a prognostic indicator of poor survival. Mol Cancer 2010;9:221.
49. Nakano A, Kanemura Y, Mori K, et al. Expression of the Neural RNA-binding protein Musashi1 in pediatric brain tumors. Pediatr Neurosurg 2007;43:279-84.
50. Nikpour P, Baygi ME, Steinhoff C, et al. The RNA binding protein Musashi1 regulates apoptosis, gene expression and stress granule formation in urothelial carcinoma cells. J Cell Mol Med 2011;15:1210-24.
51. Bobryshev YV, Freeman AK, Botelho NK, Tran D, Levert-Mignon AJ, Lord RV. Expression of the putative stem cell marker Musashi-1 in Barrett's esophagus and esophageal adenocarcinoma. Dis Esophagus 2010;23:580-9.
52. Ye F, Zhou C, Cheng Q, Shen J, Chen H. Stem-cell-abundant proteins Nanog, Nucleostemin and Musashi1 are highly expressed in malignant cervical epithelial cells. BMC Cancer 2008;8:108.
53. Kanemura Y, Mori K, Sakakibara S, et al. Musashi1, an evolutionarily conserved neural RNA-binding protein, is a versatile marker of human glioma cells in determining their cellular origin, malignancy, and proliferative activity. Differentiation 2001;68:141-52.
54. Toda M, Iizuka Y, Yu W, et al. Expression of the neural RNA-binding protein Musashi1 in human gliomas. Glia 2001;34:1-7.
55. Ohyama T, Nagata T, Tsuda K, et al. Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. Nucleic Acids Res 2012;40:3218-31.

56. de Sousa Abreu R, Sanchez-Diaz PC, Vogel C, et al. Genomic analyses of musashi1 downstream targets show a strong association with cancer-related processes. J Biol Chem 2009;284:12125-35.
57. Battelli C, Nikopoulos GN, Mitchell JG, Verdi JM. The RNA-binding protein Musashi-1 regulates neural development through the translational repression of p21WAF-1. Mol Cell Neurosci 2006;31:85-96.
58. Choisy-Rossi C, Reisdorf P, Yonish-Rouach E. Mechanisms of p53-induced apoptosis: in search of genes which are regulated during p53-mediated cell death. Toxicol Lett 1998;102-103:491-6.
59. Gartel AL, Serfas MS, Tyner AL. p21--negative regulator of the cell cycle. Proc Soc Exp Biol Med 1996;213:138-49.
60. Roy M, Pear WS, Aster JC. The multifaceted role of Notch in cancer. Curr Opin Genet Dev 2007;17:52-9.
61. Weng AP, Ferrando AA, Lee W, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science 2004;306:269-71.
62. Msi2 protein image. (Accessed 2013-02-19, at http://en.wikipedia.org/wiki/File:Musashi2_protein_in_homolog_2_in_Homo_sapiens.png.)
63. NUMB protein image. 2013-02-19, at http://en.wikipedia.org/wiki/File:Protein_NUMB_PDB_1wj1.png.)
64. Proteasome image. (Accessed 2013-02-19, at http://en.wikipedia.org/wiki/File:Proteaosome_1fnt_top.png.)
65. 3D structure of a ribosome from *T. thermophilus* at 2,8 Å resolution. (Accessed 2013-02-19, at http://www.crcl.fr/311-Projects.crcl.aspx?language=en-GB.)
66. de Andrés-Aguayo L, Varas F, Kallin EM, et al. Musashi 2 is a regulator of the HSC compartment identified by a retroviral insertion screen and knockout mice. Blood 2011;118:554-64.
67. Wu M, Kwon HY, Rattis F, et al. Imaging hematopoietic precursor division in real time. Cell Stem Cell 2007;1:541-54.
68. Dash AB, Williams IR, Kutok JL, et al. A murine model of CML blast crisis induced by cooperation between BCR/ABL and NUP98/HOXA9. Proc Natl Acad Sci U S A 2002;99:7622-7.
69. Kaeda J, Neuman D, Oberender C, et al. Longitudinal study to assess the clinical significance of MSI2 in CML patients - Abstract and poster presentation. ESH-iCMLf Meeting Baltimore 2012.
70. Koptyra M, Cramer K, Slupianek A, Richardson C, Skorski T. BCR/ABL promotes accumulation of chromosomal aberrations induced by oxidative and genotoxic stress. Leukemia 2008;22:1969-72.
71. Mocharla H, Mocharla R, Hodes ME. Coupled reverse transcription-polymerase chain reaction (RT-PCR) as a sensitive and rapid method for isozyme genotyping. Gene 1990;93:271-5.
72. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. Genome Res 1996;6:986-94.
73. Verdi JM, Bashirullah A, Goldhawk DE, et al. Distinct human NUMB isoforms regulate differentiation vs. proliferation in the neuronal lineage. Proc Natl Acad Sci U S A 1999;96:10472-6.
74. NUMB Splice Variants (Release 63). (Accessed 2011-08-01, 2011, at http://Jun2011.archive.ensembl.org/Homo_sapiens/Gene/Splice?g=ENSG00000133961.)

75. NUMB Splice Variants (Release 70). 2013. (Accessed 2013-01-20, 2013, at http://Jan2013.archive.ensembl.org/Homo_sapiens/Gene/Splice?g=ENSG00000133961.)
76. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol 1986;51 Pt 1:263-73.
77. Bell J. A simple way to treat PCR products prior to sequencing using ExoSAP-IT. Biotechniques 2008;44:834.
78. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977;74:5463-7.
79. Swerdlow H, Gesteland R. Capillary gel electrophoresis for rapid, high resolution DNA sequencing. Nucleic Acids Res 1990;18:1415-9.
80. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403-10.
81. Consortium IH. A haplotype map of the human genome. Nature 2005;437:1299-320.
82. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406-15.
83. Halvorsen M, Martin JS, Broadaway S, Laederach A. Disease-associated mutations that alter the RNA structural ensemble. PLoS Genet 2010;6:e1001074.
84. Huang Y, Sitwala K, Bronstein J, et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood 2012;119:388-98.
85. Ahuja HG, Popplewell L, Tcheurekdjian L, Slovak ML. NUP98 gene rearrangements and the clonal evolution of chronic myelogenous leukemia. Genes Chromosomes Cancer 2001;30:410-5.
86. Reference SNP(refSNP) Cluster Report: rs75236173. NCBI. (Accessed 2013-02-15, 2013, at http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs75236173.)
87. Chatterjee S, Pal JK. Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. Biol Cell 2009;101:251-62.
88. NUMB numb homolog (Drosophila) [*Homo sapiens*]. 2013. (Accessed 2013-02-15, 2013, at http://www.ncbi.nlm.nih.gov/gene/8650.)
89. Oakley K, Han Y, Vishwakarma BA, et al. Setbp1 promotes the self-renewal of murine myeloid progenitors via activation of Hoxa9 and Hoxa10. Blood 2012;119:6099-108.
90. Milne TA, Kim J, Wang GG, et al. Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. Mol Cell 2010;38:853-63.
91. Quéré R, Karlsson G, Hertwig F, et al. Smad4 binds Hoxa9 in the cytoplasm and protects primitive hematopoietic cells against nuclear activation by Hoxa9 and leukemia transformation. Blood 2011;117:5918-30.
92. Li Z, Huang H, Li Y, et al. Up-regulation of a HOXA-PBX3 homeobox-gene signature following down-regulation of miR-181 is associated with adverse prognosis in patients with cytogenetically abnormal AML. Blood 2012;119:2314-24.
93. Li M, Makkinje A, Damuni Z. The myeloid leukemia-associated protein SET is a potent inhibitor of protein phosphatase 2A. J Biol Chem 1996;271:11059-62.
94. Cristóbal I, Blanco FJ, Garcia-Orti L, et al. SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. Blood 2010;115:615-25.
95. Ayton PM, Cleary ML. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. Oncogene 2001;20:5695-707.

6. Eidesstattliche Versicherung

„Ich, Christian Oberender, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Sequence Analysis of the NUMB Gene in Chronic Myeloid Leukaemia Patients“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.
Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -*www.**icmje**.org*) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o) und werden von mir verantwortet.
Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Berlin, den 15.12.2013

Anteilserklärung an erfolgten Publikationen

Christian Oberender hatte Anteil an den folgenden Publikationen:

Publikation 1:
Christian Oberender, Jaspal Kaeda, Bärbel Pawlaczyk-Peter, Peter Daniel, Renate Arnold, Bernd Dörken & Philipp le Coutre. Sequencing of NUMB transcripts in chronic myeloid leukemia detects two single nucleotide polymorphisms. Leukemia & Lymphoma, February 2013; 54(2):421-422.
Beitrag im Einzelnen:
Diese Publikation beschreibt die angewandten Methoden und die erzielten Ergebnisse meiner experimentellen Arbeit.

Publikation 2:
Jaspal S. Kaeda, Ken I. Mills, Michael G. Kharas, Giuseppe Saglio, Michaela Schwarz, Christian Oberender, Philipp D. le Coutre.
Increased MSI2 expression is associated with aggressive CML and AML. Blood (ASH Annual Meeting Abstracts). 2011; 118 (Supplement). Abstract 2516. 2011.
Beitrag im Einzelnen:
Sämtliche klinische Daten der in dieser Studie eingeschlossenen Patienten wurden von mir aus Ambulanz- und Archivakten herausgesucht und übersichtlich zusammengestellt.

Berlin, den 15. Dezember 2013.

Christian Oberender

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Danksagung

Für die Bereitstellung des Themas, die umfangreiche Betreuung bei meiner experimentellen Arbeit und die Korrektur meiner Monographie möchte ich mich herzlichst bei Jaspal Kaeda, PhD, und PD Dr. Philipp le Coutre bedanken.

Mein umfassender Dank gilt ebenso meinen Eltern und Großeltern, die mich während des gesamten Studiums tatkräftig unterstützt haben.