Functional characterization of candidate tumor suppressor genes localized in the critical chromosomal region 13q14

INAUGURAL-DISSERTATION

for the degree of Doctor rerum naturalium (Dr. rer. nat.)

Faculty of Natural Sciences

University of Ulm

2008

Presented by Leticia Serra Barrionuevo

born in Málaga, Spain.

Dean of the Faculty of Natural Sciences
University of Ulm
Prof. Dr. Klaus-Dieter Spindler

First Reviewer:
Prof. Dr. Thomas Wirth
Head of the division of Physiological Chemistry
from the Faculty of Natural Sciences
University of Ulm

Second Reviewer:

Date of the conferral of the doctorate:

To my family, for giving me more than words could possibly describe

“La realidad, pues, se ofrece en perspectivas individuales.”
“Reality, thus, presents itself in each individual perspective.”

José Ortega y Gasset

“Damit das Mögliche entsteht, muss immer wieder das Unmögliche versucht werden.“
„In order for what is possible to emerge, the impossible must be tried over and over again.”

Hermann Hesse

Index:

Abbreviations.....i

1. Introduction.....1

1.1. Molecular basis of cancer..... 1

1.2. Chronic Lymphocytic Leukemia 3

1.2.1. The biology of B-Cells..... 3

1.2.2. The biology of B-Cell Chronic Lymphocytic Leukemia (B-CLL) 6

1.2.3. Genetic abnormalities in B-CLL..... 7

1.2.4. The importance of the deleted region at 13q14 for the pathomechanism of B-CLL 10

1.2.5. The candidate genes in 13q14 11

1.3. Protein coding genes of 13q14.....14

1.3.1. Identifying candidates from the protein coding genes of 13q14..... 15

1.4. Aims of this work 24

2. Material and Methods 25

2.1. Material 25

2.2. Methods 35

3. Results 61

3.1. Expression of recombinant fusion proteins in bacteria 62

3.1.1. Expression of recombinant RFP2 alters bacterial growth 65

3.1.2. Purification of recombinant 13q14 candidate tumor suppressor proteins 67

3.2. Production of specific antibodies for 13q14 candidate proteins.....70

3.2.1. Polyclonal antibodies from Guinea pigs 70

3.2.2. Monoclonal antibodies from Mice 73

3.2.3. Expression of 13q14 genes in cell lines 76

3.3. Intracellular localization of 13q14 proteins 78

3.3.1. Recombinant protein detection of 13q14 genes..... 78

3.3.2. Immuno-detection of endogenous proteins from 13q14..... 82

3.4. Identification of interaction partners of 13q14 proteins.....86

3.4.1. Interaction partners for 13q14 candidate proteins..... 87

3.4.2. Validated interaction partners from the yeast-two-hybrid screen 90

4. Discussion 99

4.1. Subcellular localization of 13q14 proteins 99

4.1.1. Cytoplasmic 13q14 proteins 100

4.1.2. 13q14 proteins localized to the nucleus 101

4.2. Antibody detection and expression levels of endogenous 13q14 proteins .103
4.3. The yeast-two-hybrid screen approach 104
4.3.1. Novel interactions identified in pull-down experiments..... 106
4.3.2. Common interaction partners of 13q14 proteins 107
4.4. Cellular processes involving 13q14 proteins 107
Involvement of 13q14 genes in the..... 108
4.4.1. regulation of translation 108
4.4.2. Involvement of 13q14 proteins in regulation of gene expression 111
4.4.3. Additional interaction partners of 13q14 proteins 113
4.4.4. Lessons learned from binding partners of 13q14 proteins 115
4.5. Concluding remarks 118

5. Appendix.....119

6. Bibliography 137

Summary.....149

Zusammenfassung 151

Publications.....153

Curriculum vitae.....155

Acknowledgements 157

Abbreviations

A	adenine
α	anti- (antibody)
AD	activation domain
BD	binding domain
B-CLL	B-cell chronic lymphocytic leukemia
bp	base pairs
β-ME	beta-mercapto ethanol
BSA	bovine serum albumin
C	cytosine
CD	cluster of differentiation
cDNA	complementary DNA
CLL	chronic lymphocytic leukemia
CT	carboxy-terminus
DNA	desoxy-ribonucleic acid
ddNTP	dideoxy-nucleotides tri-phosphate mix
Dn	donkey
dNTP	desoxy-nucleotides tri-phosphate mix
dsRNA	double stranded RNA
E	glutamic acid
E.coli	*Escherichia coli*
EF	elongation factor
e.g.	“for example” (latin: “exempli gratia”)
ELISA	enzyme linked immunosorbent assay
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum-associated degradation
Fow	forward primer
G	guanine
g	unit of acceleration
GEFs	guanine nucleotide exchange factors
GFP	green fluorescent protein
Gp	guinea pig
GR	glucocorticoid receptor
HA	haemmagglutinin
His	histidine
HMTase	histone methyl transferase
HRP	horseradish peroxidise
hr(s)	hour(s)
Ig	immunoglobulin
i.e.	“that is” (latin: “id est”)
IPTG	isopropyl β-D thiogalactosidase
KD	kilo-daltons
Leu	leucine
mA	mili-Amper
mRNA	messenger RNA
min	minute(s)
Ms	mouse
Ni	nickel ions
NLSs	nuclear localization signals

NR	nuclear receptor
NT	amine-terminus
OD	optical density
O/N	over night
ORF	open reading frame
PA	protein A
PCR	polymerase chain reaction
PD	pull down
pH	latin: potentia hydrogenii (potential of hydrogen)
PPIs	protein-protein interactions
RA	retinoic acid
Rb	rabbit
Rev	reverse primer
RNA	ribonucleic acid
RT	room temperature
S	serine
SAP	shrimp alkaline phosphatise
S.cerevisiae	*Saccaromices cerevisiae*
SDS-PAGE	sodium-dodecyl-sulfate polyacrylamide gel electrophoresis
sec	second(s)
Sh	sheep
T	thymine
Trp	tryptophan
TSG	tumor suppressor gene
UPR	unfolded protein response
Ura	uracyl
UTR	un-translated region
UV light	ultraviolet light
v/v	volume per volume
w/v	weight per volume
Y2H	yeast-two-hybrid

Genes:

ACTC	Actin, alpha cardiac muscle 1
C11orf2	chromosome 11 open reading frame 2
C13orf1	chromosome 13 open reading frame 1
CRMP1	collapsin response mediator protein 1
EED	embryonic ectoderm development
eEF1A	eukaryotic translation elongation factor 1 alpha
eEF1G	eukaryotic translation elongation factor 1 gamma
GNB2	guanine nucleotide binding protein (G protein), beta polypeptide 2
KPNA3	karyopherin alpha 3
RCBTB1	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1
RFP2	ret finger protein 2
RPAIN	RPA interacting protein
RXRA	retinoid X receptor, alpha
SETDB2	SET-domain, bifurcated 2

Metric units:

m	moles
M	molar (moles/liter)
mM	mili-molar (10^{-3} M)
g	gram
mg	mili-gram (10^{-3} gr)
µg	micro-gram (10^{-6} gr)
ng	nano-gram (10^{-9} gr)
l (L)	liter
ml	mili-liter (10^{-3} l)
µl	micro-liter (10^{-6} l)

1. Introduction

1.1. Molecular basis of cancer

Cancer arises from the disruption of normal cellular function that results in uncontrolled cell proliferation and/or apoptosis. Normal cells give rise to normal tissue development and maintenance by following an ordered growth, division, differentiation and apoptosis program. Each cell in an organism contains the same genetic information, but different cell types and functions arise from the differential regulation of that information. As a result, specific molecules will be present in different cell types or at different time points and will confer specific functions. In contrast, malignant cells have altered cell functions due to genomic aberrations. These genomic aberrations can be of genetic or epigenetic nature, but they ultimately lead to uncontrolled cell division, resistance to undergo programmed cell death (apoptosis) or the ability of spreading (metastasize) to other tissues. The process of cancer development is thought to be an evolutionary course of action leading to progressive transformation of normal cells through the acquisition of alterations in the genome. The changes affect essential cellular pathways and, depending on the nature of the processes being involved, the genes altered can be classified into two categories: proto-oncogenes and tumor suppressor genes.

Proto-oncogenes

Proto-oncogenes are normal genes that can become oncogenic due to mutations or increased gene expression. Their gene products usually promote cell growth and division, and in cancer they are altered through “gain of function” mutations that increase their activity turning them into oncogenes (Falco & Giordano 2006). Proto-oncogenes code for highly variable types of molecules, like transcription factors that stimulate the expression of other genes, signal transduction molecules that stimulate cell division, or cell cycle regulators that move the cell through the cell cycle. A typical example of a proto-oncogene is *Ras* (Weinberg 2007). *Ras* codes for a protein which functions as a molecular switch. It can be found in two states, an inactive that is bound to GDP and an active one that is bound to GTP. The switching from one state to the other is regulated by upstream signals, which control the transition by favouring the exchange of GDP for GTP. In the active GTP-bound state, the Ras

protein has an effect on its downstream effector molecules, which will be altered (usually activated) leading to increased cell proliferation. In many types of cancer, Ras activity is modified by point mutations that hold Ras in a permanently activated state, thus conferring transformed cells with increased proliferative capacity.

Tumor suppressor genes

Tumor suppressor genes are involved in many different functions like control of cell division, growth, DNA-repair mechanisms and apoptosis. Tumor suppressor genes are inactivated by “loss of function” mutations (Falco & Giordano 2006), resulting in a failure to respond to signals that would lead to quiescence or apoptosis. Thus, cancer cells, which have lost these restraints, are able to proliferate independently of suppressing signals (Hanahan & Weinberg 2000). An example of a tumor suppressor gene is *p53* (Lane 1992). *p53* codes for a transcription factor that monitors a number of functions in the cell, with a central role in DNA-damage. A wide variety of signals ranging from metabolic disorder to genomic alterations can induce p53 activation. Upon damage, sensor molecules activate the p53 protein, which in turn enters the nucleus, binds to target genes and regulates their expression. As a result, different pathways are activated to either induce cell-cycle arrest or apoptosis (Lane 1992). Therefore, cells with a loss of p53 activity are able to continue active proliferation under circumstances that normally would cause a proliferation arrest or even apoptotic death. In addition, loss of p53 activity leads to chromosomal instability, a state characterized by numerous chromosomal abnormalities and the emergence of highly abnormal karyotypes (Knudson 2001). The importance of p53 for proper cellular function is illustrated by the fact that it is disrupted in roughly 50% of all human tumors (Carter et al. 2004). For example, in Chronic Lymphocytic Leukemia (CLL), p53 inactivation is strongly associated with adverse clinical outcome and drug resistance (Döhner et al. 1995), but only occurs in 10-15% of patients. This suggests that in the remaining CLL cases mechanisms other than the p53 pathway overcome programmed cell death. Further studies on the molecular consequences of genetic abnormalities in CLL, e.g. at the protein level, will help to elucidate the pathomechanism resulting in this type of cancer.

1.2. Chronic Lymphocytic Leukemia

Leukemia is a type of malignant proliferation characterized by abnormal blood-cell counts in the patients. The term “leukemia” means “white blood” (Greek: *λευκός*, "white"; *αίμα*, "blood") and reflects the finding that lymphocytes are found in abnormally high numbers in the blood of patients presenting with these malignancies. Hematological malignancies are classified according to the cell lineage they originate from. Thus, lymphoblastic or lymphocytic malignancies occur in progenitor or stem cells that will give rise to lymphocytes. In contrast, myeloid or myelogeneous neoplasms derive from progenitor or stem cells which will give rise to red cells, some types of white cells and platelets[1]. The two former types can be further subdivided in different sub-classes and sub-groups according to different surface markers, global gene expression and other signatures of the malignant cells (Jaffe & Stein 2001). Among all types of leukemia, Chronic Lymphocytic Leukemia is the most common leukemia in adults in Western countries, accounting for ca. 70% of all lymphoid leukemias[2]. The overall incidence of CLL is around 3.9/100,000 per year, and the male to female ratio is approximately 2:1 ([3]). It affects mainly elderly people, with a median age at diagnosis of 72 years[4] (Jaffe & Stein 2001). The dominant type of CLL is B-Cell Chronic Lymphocytic Leukemia (B-CLL), which is derived from B-cells and constitutes over 95% of all CLL cases. It has a highly variable clinical course, with some patients dying within a few months after diagnosis and others surviving for 20 years or more after the disease has been diagnosed (Ripolles et al. 2006).

1.2.1. The biology of B-Cells

In humans, the development of B-cells starts in the bone marrow from lymphocytic progenitor cells (Figure 1.1). These progenitor cells originate from hematopoietic stem cells, and differentiate into specific cell types by undergoing changes in gene expression. In turn, differential gene expression results in the production of specific molecules that confer different properties to different cell types.

[1] http://www.leukemia-lymphoma.org/
[2] http://atlasgeneticsoncology.org/Anomalies/CLL.html
[3] http://seer.cancer.gov/statfacts/html/clyl.html
[4] http://seer.cancer.gov/statfacts/html/clyl.html

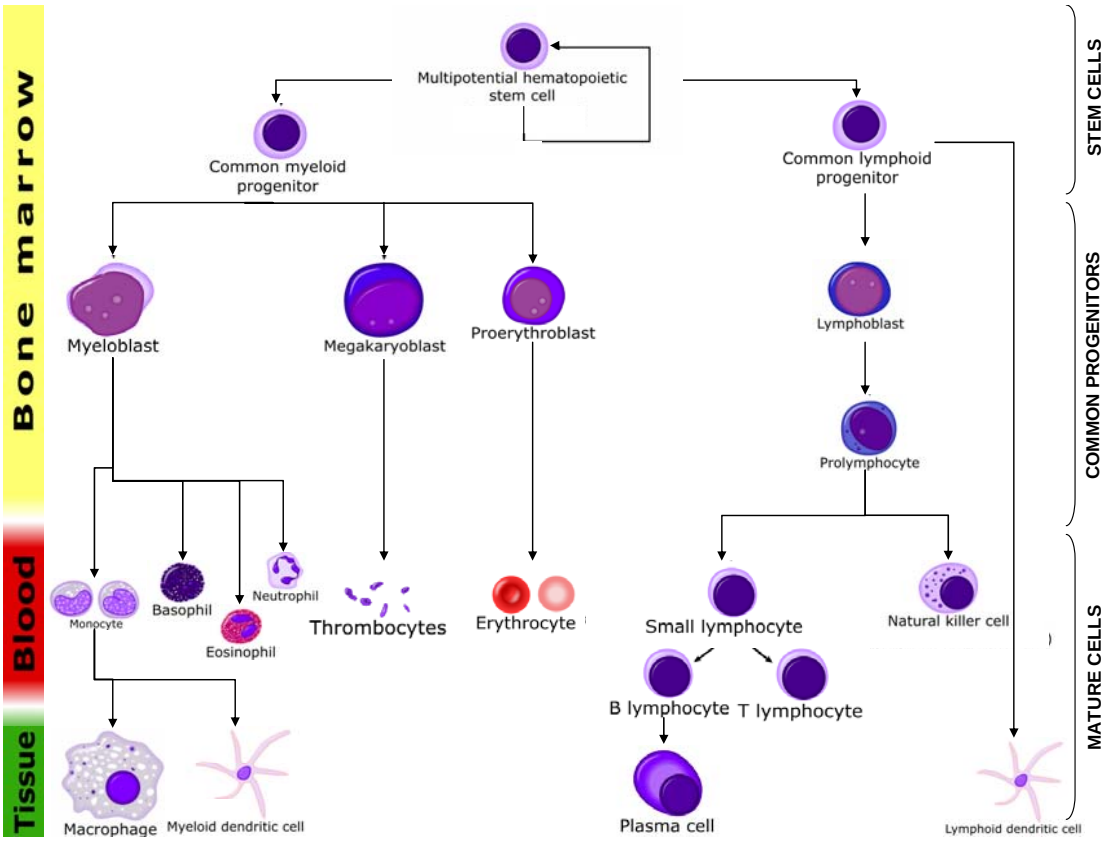


Figure 1.1 Schematic picture of blood cells development from their bone marrow progenitors to fully differentiated cells (http://en.wikipedia.org/wiki/Image:Hematopoiesis_%28human%29_diagram.png#file**).**

An important step in the differentiation process of B-cells is the assembly of genes for the variable regions of heavy and light immunoglobulin (Ig) chains, a process called V(D)J recombination (Chiorazzi et al. 2005). After recombination has occurred, the B-cells are endowed with functional Ig genes, which, together with other molecules, constitute the B-cell receptor (BCR). Only cells expressing functional BCR molecules survive and become mature naïve B-cells, i.e. not yet exposed to any antigen, that leave the bone marrow and can participate in immune responses (Küppers et al. 1999). Mature naïve B-lymphocytes are activated, when the surface BCRs encounter a corresponding antigen. As a result of binding the antigen, the BCR triggers cytoplasmic signalling cascades, which exert different responses depending on the differentiation stage of the B-lymphocyte (Rajewsky 1996). This activation process is often dependent on T-cells and takes place within the lymph nodes in structures called “germinal centres”, where B-lymphocytes activated by T-cells undergo cell proliferation and further genetic changes in their maturation process (Figure 1.2). Maturation of B-cells can also occur independently from T-cells and outside the germinal centres, in the marginal zones around lymphoid follicles (Küppers 2005).

Changes leading to B-cell maturation comprise modifications in Ig genes by somatic hypermutations and class switch recombination.

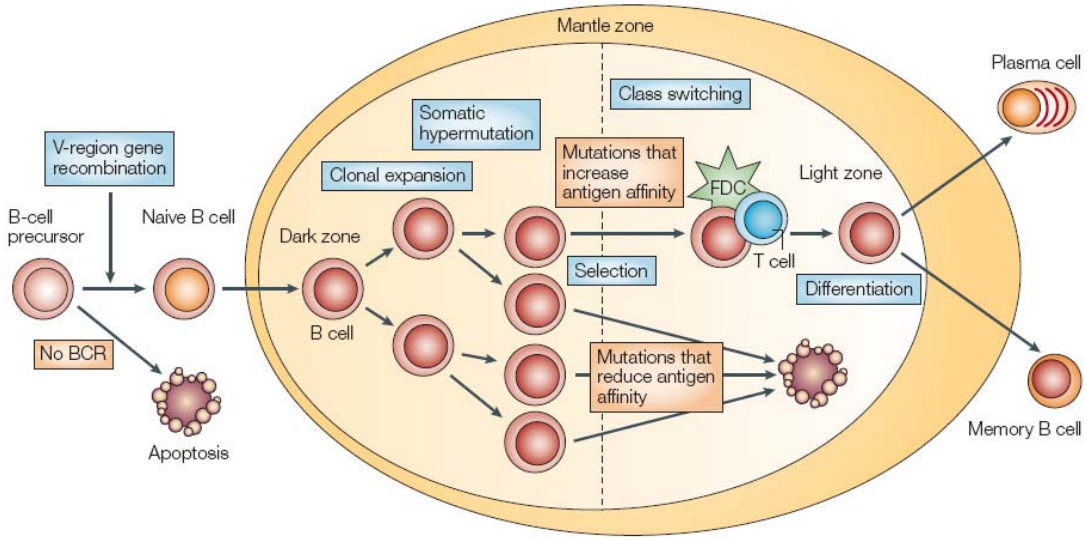


Figure 1.2 Schematic picture of the B-cell differentiation in the germinal centre of lymph nodes (Küppers, 2005).

Somatic hypermutations introduce mutations in the variable-region of Ig genes at high rates in order to improve specificity and binding of the BCR to antigen. Changes in the BCR sequence may generate cells with either increased affinity for the antigen, which will be positively selected in secondary lymphoid organs, or cells with lower affinity for the antigen, which will undergo apoptosis. Through class switch recombination the constant-region of the heavy chains are switched to other classes of Ig with different effector functions resulting in B-cells with different cellular functions (Küppers et al. 1999). In the germinal centers B-cells undergo clonal expansion, and some, after somatic hypermutations and class switching, mature into plasma cells or memory B-cells (Küppers 2005). The development of B-lymphocytes can be traced by monitoring the appearance/disappearance of appropriate surface markers that change with the differentiation stage of the B-cell (Figure 1.3) (Küppers et al. 1999). As an example, naïve circulating B-cells are CD27 negative and do not present somatic mutations, while differentiated plasma cells carry somatic mutations in their BCR genes and are CD27 positive. These markers help to characterize the differentiation stage of the B-cells and provide objective criteria for the classification of lymphoid malignancies, as these relate to the cell type from which the neoplasm probably originated. To this respect, it is not clear, which is the cell of origin of CLL, because CLL-cells present several properties of different B-cell subtypes.

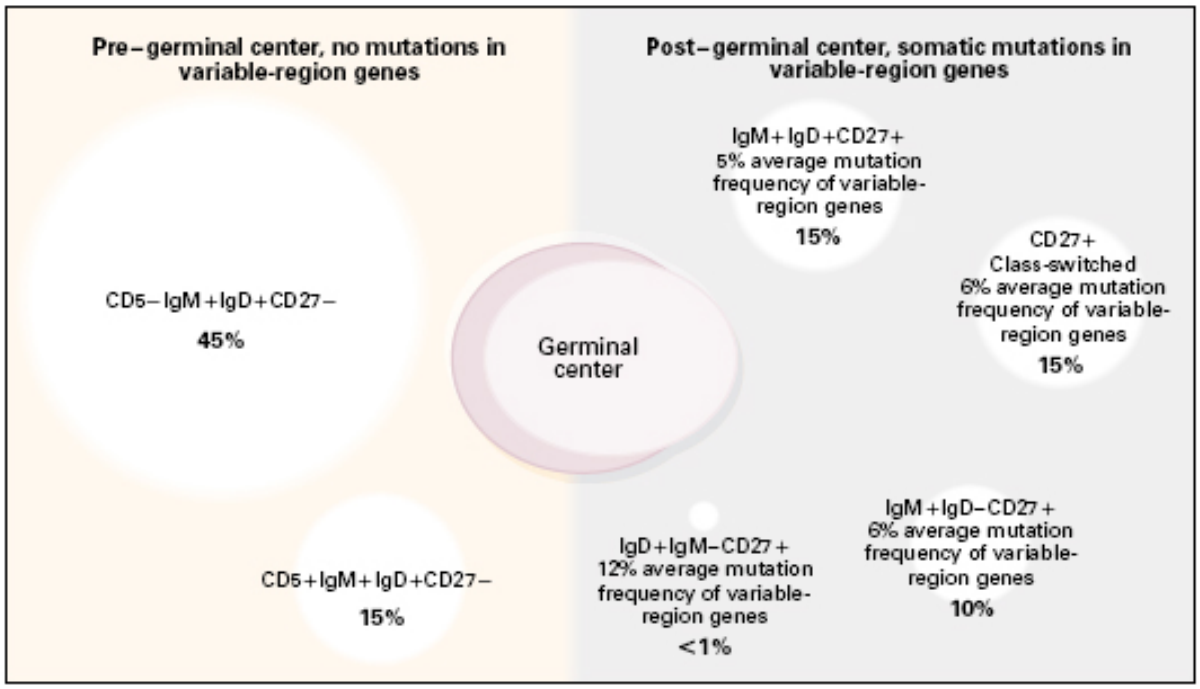


Figure 1.3 Composition of the B-cell repertoire in peripheral blood in humans (Küppers, 1999). Depicted in the picture are different B-cell subgroups based on their B-cell marker phenotype together with their occurrence in peripheral blood. In subpopulations from the post-germinal center the average frequency of mutations of the heavy-chain variable-region genes is also indicated.

1.2.2. The biology of B-Cell Chronic Lymphocytic Leukemia (B-CLL)

B-CLL is characterized by a neoplastic accumulation of abnormal, functionally incompetent B-lymphocytes, mainly in blood and the bone marrow. In B-CLL, B-lymphocytes express an unusual combination of surface markers (e.g. CD19, CD5, and CD23) and present reduced levels of membrane IgM, IgD and CD79b, a phenotype characteristic of mature, activated B-lymphocytes (Chiorazzi et al. 2005). In addition, B-CLL cells within one individual are clonal, i.e. stem from one single cell, which can be inferred from the deviation from the random usage of one of the mutually exclusive antibody light chains, kappa or lambda (which, in the case of B-CLL, is the same for the entire population of the abnormal B cells). Classification of B-CLL into clinical subgroups can be performed according to different parameters, such as the mutation state of the variable heavy-chain (V_H) genes, with unmutated genes correlating with more aggressive B-CLL course (Hamblin et al. 1999), or ZAP-70 expression, with ZAP-70 over-expression correlating with shorter overall survival and unmutated V_H genes (Zanotti et al. 2007). Different parameters can also be combined to better predict the outcome of the disease (Damle et al. 1999). An interesting aspect in the biology of B-CLL is to which extent the abnormal B-cells depend on BCR activation for survival. In normal B-cells, the BCR plays a crucial role

in survival, with BCR-negative cells rapidly undergoing apoptosis in mice (Kraus et al. 2004, Lam et al. 1997). The activation of the BCR triggers a number of signaling events that are transmitted to the nucleus and that are responsible for survival. These signaling cascades have been proven to play a crucial role in lymphoid malignancies (Küppers et al. 1999), but signaling cascades circumventing the BCR have been proposed as well (Chiorazzi et al. 2005). This later finding indicates that, in addition to the pathways concerning the BCR, there could be further alternative signaling mechanisms involved in the development of B-CLL that should be investigated.

1.2.3. Genetic abnormalities in B-CLL

Genetic abnormalities in B-CLL comprise different types of alterations like duplications and deletions. To delineate genetic abnormalities in B-CLL, different approaches have been taken. By conventional cytogenetics, clonal chromosomal aberrations are detected in 40 to 50% of cases, the most common being trisomy 12 and deletions in 13q14 (Juliusson et al. 1990). But conventional cytogenetics require dividing cells in order to produce metaphase spreads, and this approach has been hindered by the low mitotic activity of the tumor cells in B-CLL *in vitro* (Stilgenbauer et al. 1998). However, using interphase cytogenetics such as fluorescence *in situ* hybridization (FISH), it is possible to detect genetic abnormalities in interphase nuclei as well. In a large study of 325 B-CLL patients, FISH was successfully used to assess the frequency and clinical relevance of genomic aberrations in B-CLL (Döhner et al. 2000). In their results, Döhner et al. could show that 82% of the B-CLL patients presented genomic aberrations, with the most common being deletion of 13q14, followed by deletion of 11q22 and trisomy of chromosome 12 (Table 1.1). In addition, a correlation between certain genetic abnormalities and prognosis was found. Studies showed that patients with 17p13 and 11q22 deletions had the shortest survival and more advanced disease, while patients with 13q14 deletion had the longest survival and better prognosis (Figure 1.4) (Döhner et al. 2000, Stilgenbauer et al. 1998). The later finding, together with the high frequency of the 13q14 deletion, suggest that this aberration could be a primary event in the development of B-CLL, possibly predisposing B-cell clones to undergo additional alterations in the course of the disease. While the identity of the genes involved in the 13q14 deletion is still unclear, for the other most common deletions at 11q22 and 17p13, it is likely that the

genes ATM (in 11q) (Schaffner et al. 1999) and p53 (in 17p) (Gaidano & Balleruni 1991) are causative for the poor outcome of the disease, as they both regulate apoptosis.

Aberration	Patients (%)
13q deletion	55
11q deletion	18
12q trisomy	16
17p deletion	7
6q deletion	6
8q trisomy	5
t(14q32)	4
3q trisomy	3
Clonal abnormalities	82
Normal karyotype	18

Table 1.1 Percentage of patients with CLL presenting the different genomic aberrations (Döhner et al. 2000). A cohort of 325 patients was assayed with fluorescence in situ hybridisation. 175 patients presented one aberration, 67 presented two and 26 presented more than two aberrations.

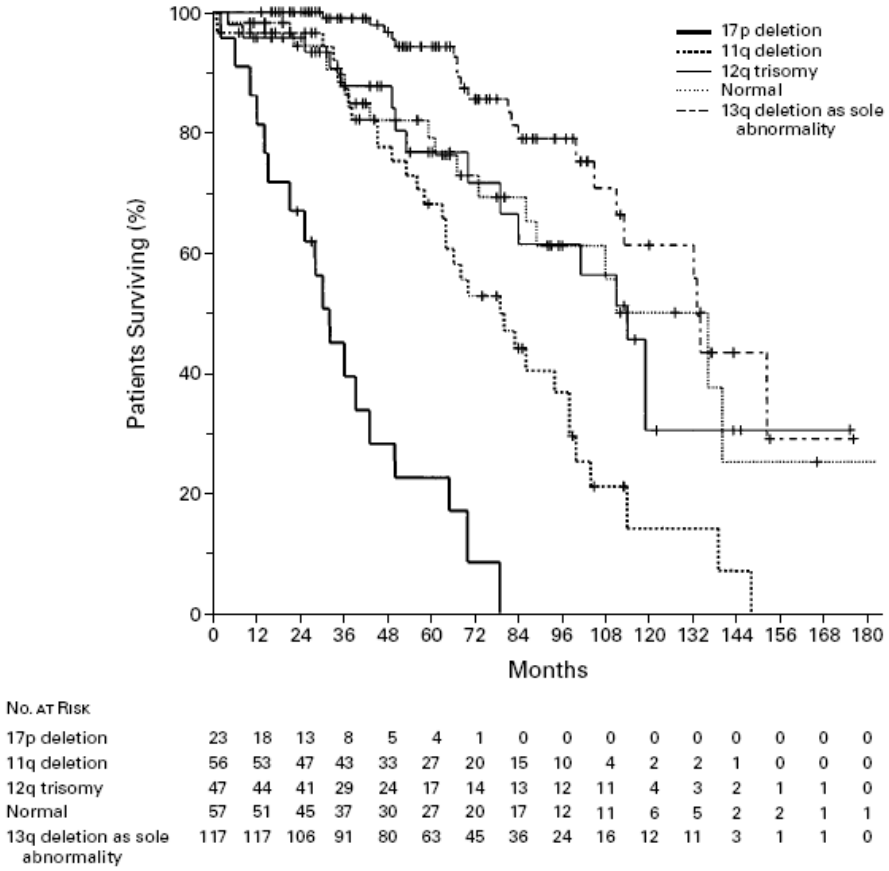


Figure 1.4 Kaplan-Meier plot of survival of CLL patients with different genetic abnormalities (Döhner et al. 2000). Patients with 17p deletion show the shortest survival times while patients with 13q deletion show the longest survival. Number of patients with each genetic abnormality at each time point is indicated at the bottom.

Other genomic features that can also be related to outcome of the disease include V_H gene mutations. Based on the presence or absence of V_H gene mutations, B-CLL patients can be further divided into two groups of different prognostic outcome, with unmutated V_H showing worst prognosis and shorter overall survival (Figure 1.5) (Hamblin et al. 1999).

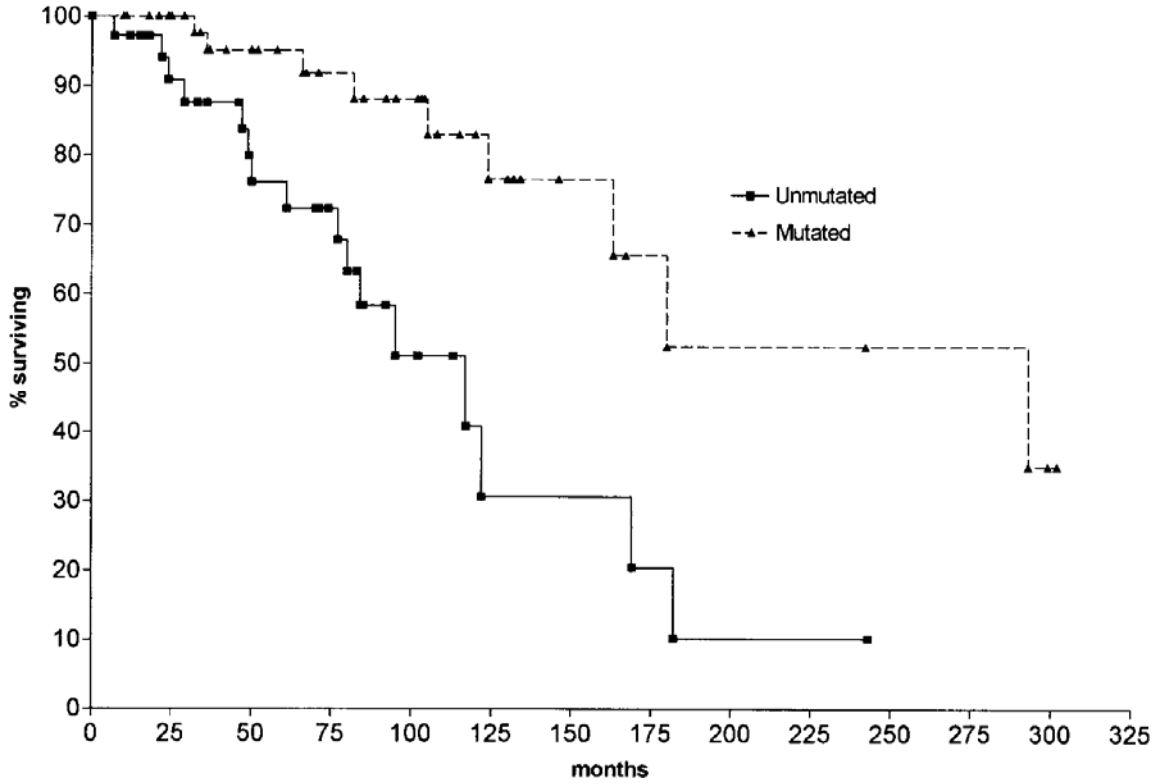


Figure 1.5 Kaplan-Meier plot showing survival curves of CLL patients with mutated and unmutated V_H genes (Hamblin et al 1999). The median survival times for unmutated vs. mutated CLL were 117 vs. 293 months respectively.

In a complementary study, further genetic features were related to the V_H mutation status (Stilgenbauer et al. 2002). It was shown that, although the overall incidence of aberrations was not different between the V_H mutated and unmutated subgroups, the incidence of unmutated V_H was higher in patients with high risk aberrations like deletion of 17p and 11q, while a higher frequency of mutated V_H was seen in patients with favorable aberrations like deletion 13q (Table 1.2). Thus, there are significant correlations between the genetic abnormalities found in the patients and a) phenotypic characteristics of the malignant cells, and b) the course of the disease. These correlations are not only central in the clinic for prognosis and therapeutic decisions, but also indicate that there must be genes in the regions of the genomic aberrations that play a major role in the development of the disease. To this respect, some of the genes associated with the malignant phenotype have been identified for

certain genomic aberrations. Nevertheless, for the most common abnormality found in CLL patients, 13q14 deletion, the gene or genes playing a role in the development of the disease are still unknown.

Percentage of CLL patients with mutated vs. unmutated V_H genes

Aberration	*V_H mutated (homology <98%) n = 132 (44%)*	*V_H unmutated (homology ≥98%) n = 168 (56%)*	*P value[a]*
Clonal aberrations	80%	84%	0.37
13q deletion	65%	48%	0.004
13q deletion single	50%	26%	<0.001
Trisomy 12	15%	19%	0.44
11q deletion	4%	27%	<0.001
17p deletion	3%	10%	0.03
17p or 11q deletion	7%	35%	<0.001

Table 1.2 Percentage of CLL patients with mutated versus unmutated V_H genes (Stilgenbauer et al. 2002). Blue arrows indicate aberrations where the percentage of patients with unmutated V_H is significantly higher than that of patients with mutated V_H genes. Orange arrow indicates the aberration where the percentage of patients with mutated V_H is significantly higher than that of patients with unmutated V_H genes.

1.2.4. The importance of the deleted region at 13q14 for the pathomechanism of B-CLL

Deletion of the critical region in chromosomal band 13q14 is the most frequent genetic aberration found in B-CLL. Delineation of the minimally deleted region in several studies led to different results, but consistently the markers D13S273 and D13S294 distal to the RB1 locus were deleted (Figure 1.6). Indeed, Stilgenbauer et al. could narrow down the minimally deleted region to 400Kb between markers D13S273 and D13S25 (Stilgenbauer et al. 1998b). This minimally deleted region was in agreement with that found by others (Kalachikov et al. 1997). Discontinuous deletions, as also found in other studies (Bullrich et al. 1996), could be detected, although at a very low rate (one of 39 B-CLL samples tested by Stilgenbauer et al.). In addition, evolutionary conservation of the region supports an important cellular role for 13q14. In one study, high overall sequence and nucleotide conservation as well as gene order and content has been found between the 13q14 human region and a syntenic region on mouse chromosome 14 (Kapanazade et al. 2000). In another study a common ancestral origin for the deleted regions on 11q22 and a sub-region

within 13q14 was found on chromosome 9 in the zebrafish and on chromosome 1 in the chicken. Because these two regions are commonly deleted in B-CLL, it is hypothesized that they might be functionally linked (Auer et al. 2007). These studies concerning cross-species DNA-sequence comparisons help to identify conserved sequences likely to encode important biological functions, and further highlight the biological significance of this region. In the search for the pathomechanism of B-CLL, characterization of the genetic aberrations associated with the disease was of crucial importance. The delineation of the deleted region in 13q14 narrowed down the number of candidate genes that might be involved in the development of B-CLL. In addition, the evolutionary data further support the hypothesis of a gene or genes of vital relevance being present in that region.

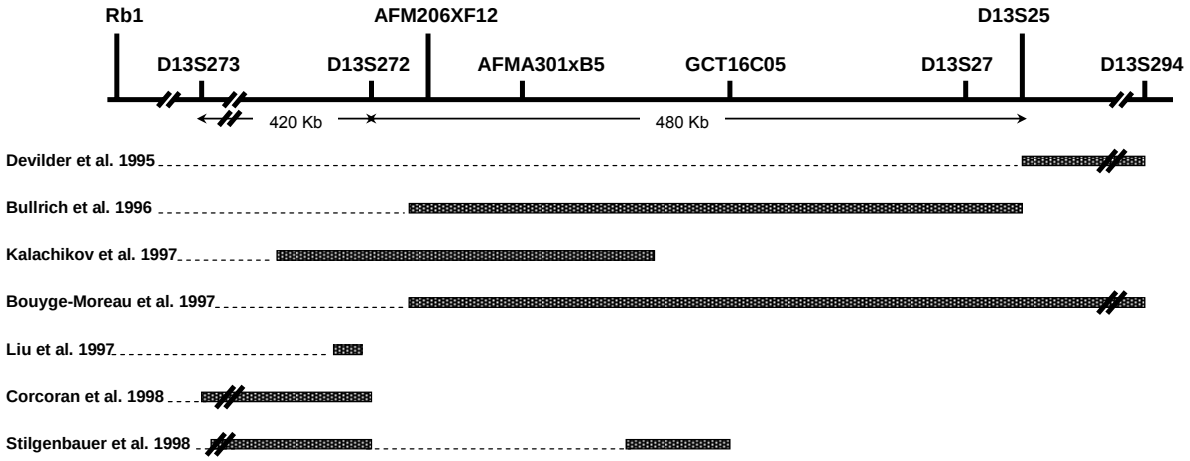


Figure 1.6 Schematic representation of the minimally deleted region in 13q14 in B-CLL as described by different groups. Distances are not drawn to scale. Figure based on Wolf et al. 2001.

1.2.5. The candidate genes in 13q14

Since more than a decade, several groups try to identify the tumor suppressor mechanism underling the 13q14 region, and to date the mechanism is still not clear. Many genes localized in the 13q14 region, both non-coding and protein-coding, have been proposed to be the candidate tumor suppressor gene (TSG) of 13q14 (Figure 1.7) (Baranova et al. 2004, Calin et al. 2002, Hammarsund et al. 2004, Kapanadze et al. 2000, Mertens et al. 2000, Migliazza 2001, Rowntree et al. 2002, van Everdink et al. 2003, Wolf et al. 2001).

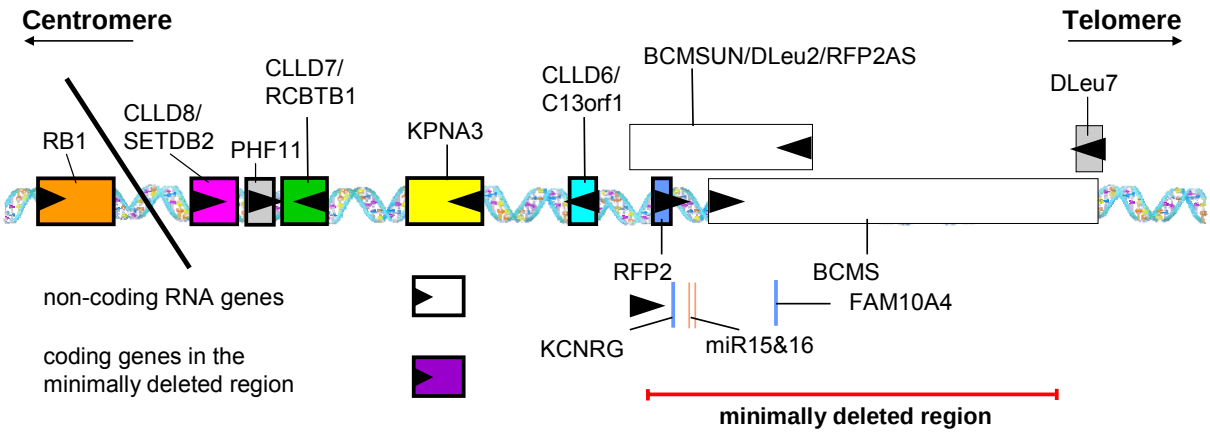


Figure 1.7 Schematic representation of the critical region in 13q14. Depicted in the figure are: the minimally deleted region (according to Stilgenbauer et al. 1998) (red bar), coding genes (coloured squares and blue bars), non-coding genes (white squares), and micro-RNAs (orange bars). Arrow-heads indicate the direction of gene transcription. Gene distances are not drawn to scale.

At first, the close vicinity of the *RB1* locus and the deleted region suggested that *RB1* might be the TSG involved in the pathogenesis of B-CLL. Mutational screening showed that in the majority of B-CLL, at least one of the copies of *RB1* was intact, pointing to a tumor suppressor mechanism in an adjacent locus (Liu et al. 1993). In subsequent studies, it further became clear that other markers distal to *RB1* (like D13S319 and D13S25) were more frequently deleted than *RB1* itself, pointing to a possible tumor suppressor mechanism involving another gene or genes in the region distal to *RB1* (Stilgenbauer et al. 1995, van Everdink et al. 2003). Interestingly, there are several non-coding RNA genes in the minimally deleted region that have therefore been proposed to be involved in the pathomechanism of the critical region (*Dleu1/BCMS* and *Dleu2/BCMSUN*) (Kapanadze et al. 2000, Migliazza 2001, Rowntree et al. 2002, Wolf et al. 2001). The other non-coding genes of the region, the two micro-RNA genes *miR15a* and *miR16-1*, were also reported to play a pathogenic role in B-CLL (Calin et al. 2002, Cimmino et al. 2005). Nevertheless, the pathomechanism for the disease is not yet fully understood. In addition to the non-coding genes, some of the protein-coding genes of 13q14 have been proposed as the most likely candidates for the TSG in B-CLL (van Everdink et al. 2003). The protein coding genes are candidates for the tumor suppressor mechanism because they are evolutionary highly conserved and are down-regulated in patients (Baranova et al. 2004, Hammarsund et al. 2004, Mertens et al. 2002, van Everdink et al. 2003). Recently, we could show that the 13q14 region is monoallelically expressed due to epigenetic regulation of several candidate genes (Mertens et al. 2006). This kind of regulation could explain both the lack of genetic aberrations in both gene copies and

the high level of down-regulation detected for some of the candidate genes upon genetic loss of a single gene copy (Mertens et al. 2002) (Figure 1.8).

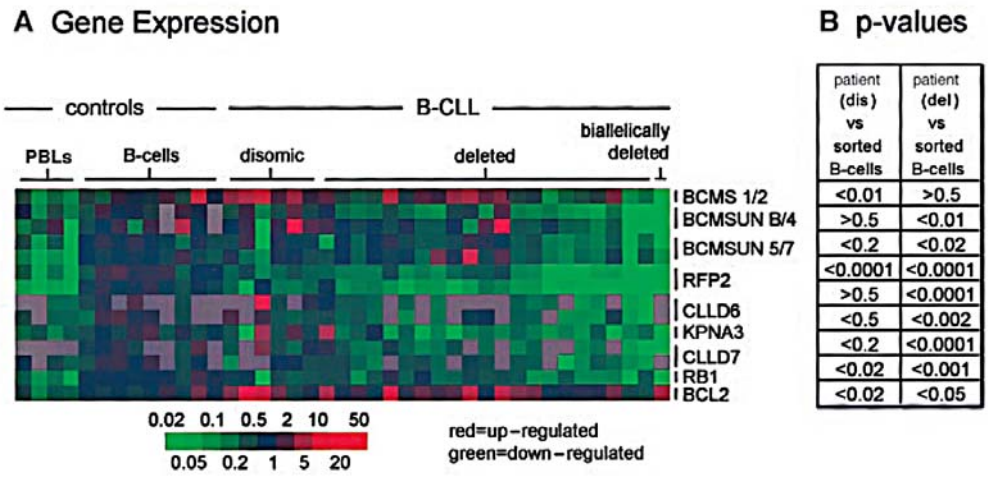


Figure 1.8 Real-time PCR data from expression analysis of 13q14 candidate genes (Mertens et al. 2002). Expression was analyzed for B-CLL patients and compared to that of peripheral blood lymphocytes (PBLs) and sorted B-cells of healthy donors (B-cells). Three house-keeping genes were used to normalize gene expression. Sorted B-cells from healthy donors were compared to patients with both copies of the region (disomic) and to patients with loss of one (deleted) or both (biallelically deleted) copies of the region. Three genes, RFP2, CLLD6 (C13orf1) and CLLD7 (RCBTB1) were the most significantly down-regulated in the deleted patients (B).

With epigenetic inactivation of one copy already in non-malignant cells, deletion of the single active chromosome copy of the critical region in malignant cells would lead to complete inactivation of the TSG in 13q14. Also, in patients with retention of both gene copies, gene function could be lost by epigenetic inactivation of the two copies (Figure 1.8). However, whereas some of the molecular intricacies of the regulatory mechanism have been elucidated, the identity and function of the candidate TSG in the region is still unclear. While some studies have concentrated on the possible functions and implications of the non-coding genes, to date no comprehensive functional study has been performed with the protein products of the candidate genes of the region. Such studies are critical in clarifying the role of 13q14 in the pathomechanism of B-CLL.

1.3. Protein coding genes of 13q14

While chromosome 13 is one of the human chromosomes with the lowest number of genes, the critical region in 13q14 is gene dense (see Figure 1.9).

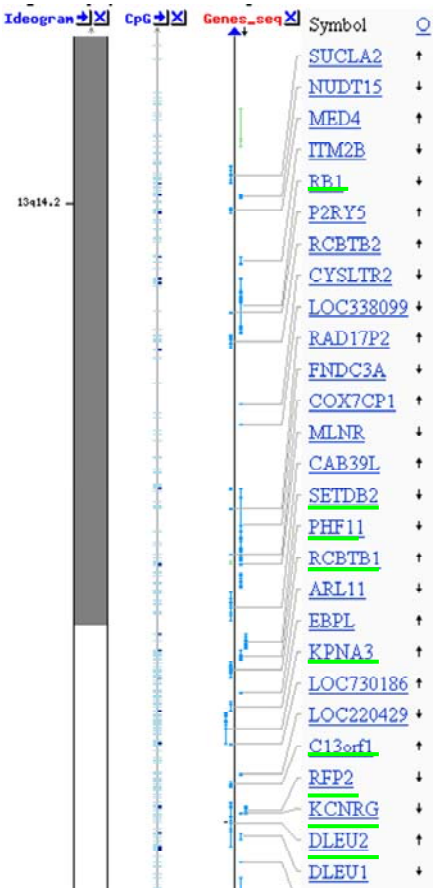


Figure 1.9 Physical map of 13q14 (region 47,400K-49,600K). Indicated are CpG islands ("CpG"), position of the genes ("genes_seq"), the gene symbols ("symbol") and the direction of transcription ("O"). Genes depicted in figure 1.7 are underlined in green. http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?taxid=9606&chr=13

As proteins are the molecules ultimately determining the phenotype of each cell, to understand the pathomechanism of 13q14 the candidate genes coding for proteins should also be functionally characterized. First hints towards the function of proteins of unknown function are the analysis of homology to well characterized proteins. Then, proteins with the most interesting inferred functions can be selected for further biochemical analysis and once the functions of the proteins have been identified, additional experiments can be designed to determine the role of the proteins in the development of malignancy, as in the case of 13q14 proteins and the development of CLL.

1.3.1. Identifying candidates from the protein coding genes of 13q14

The function of the 13q14 candidate tumor suppressor proteins has mainly been inferred from homology to proteins or domains with known function. To decide which proteins from this region are the most likely candidates for the development of malignancy and select them for further analysis, two aspects were considered in this work: their genomic localization within or in close vicinity to the 13q14 minimally deleted region defined by Stilgenbauer et al. (Stilgenbauer et al. 1998); and the putative assigned function based on homology to known proteins. According to these criteria, the proteins C13orf1, RFP2, SETDB2, RCBTB1 and KPNA3 were selected for further analysis.

1.3.1.1. C13orf1 — a SPRY domain-containing protein

C13orf1 (also known as *CLLD6*) is a gene adjacent to the minimally deleted region (Figure 1.7) and codes for a protein of 196 amino acids. It has an unusually high conservation, with a nucleotide sequence similarity between mouse and human of 97% (Baranova et al. 2004), being the average conservation of orthologs between these two species 82-85%. This high conservation at the amino acid level for C13orf1 highlights its functional importance. The only homology domain found for the C13orf1 protein is a SPRY domain (Figure 1.10).

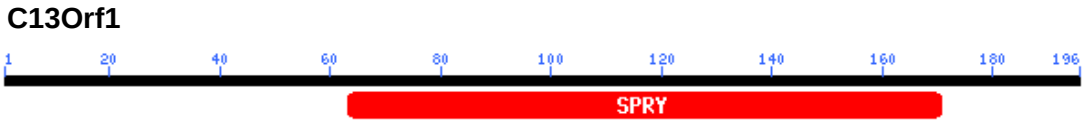


Figure 1.10 Schematic representation of the C13orf1 protein coding sequence. Depicted is the SPRY homology domain (red box). Numbers indicate amino acids. http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

The function of SPRY domains is not fully characterized, but in homologous proteins the SPRY domain is necessary for protein-protein interaction and homo-oligomerization (Blayney et al. 2004). The domain adopts a highly distorted, compact β-sandwich fold, with two additional short α-helices at the N-terminus (Woo et al. 2006). In an alignment of 20 different SPRY domain-containing proteins, the consensus residues were mostly hydrophobic and are likely important for maintaining the structural scaffold. Therefore, the tertiary structures of all SPRY domains are probably similar. The absence of any conserved residue at the protein-binding surface suggests that the SPRY domains have a varying individual specificity for

binding a specific partner protein (Woo et al. 2006). It seems that the majority of the SPRY domains, if not all, would be protein-interacting modules that bind a partner protein, with the target specificity determined by variation of the amino acid sequence on the ligand binding surface. In this work we aimed to identify the interaction partners of C13orf1 and whether they might be related to malignancy.

1.3.1.2. RFP2 and TRIM proteins

The *RFP2* gene (also known as *Leu5* or *TRIM13*) is a gene located within the minimally deleted region (Figure 1.7) and coding for a protein of 417 amino acids. It belongs to the family of TRIM proteins (Reymond et al. 2001, van Everdink et al. 2003) and contains a RING finger, B-box, coiled-coil domain and a transmembrane domain (Figure 1.11).

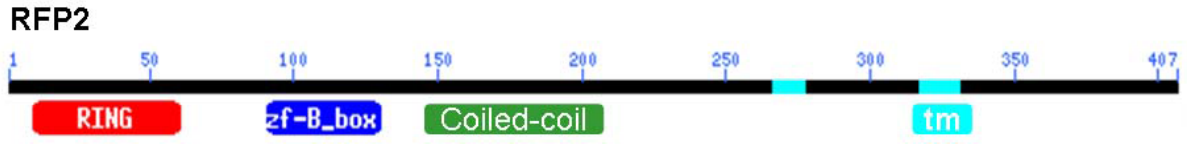


Figure 1.11 Schematic representation of the RFP2 protein coding sequence. Depicted are the homology domains: RING (red box), zinc finger-B-box (dark blue box), coiled-coil (green box) and transmembrane domain (light blue box). Numbers indicate the amino acids positions. http://www.ncbi.nlm.nih.gov/Structure/ cdd/wrpsb.cgi

To date, the TRIM family contains over 70 identified members. TRIM proteins are characterized by their TRIM domain (TRIpartite Motif) composed of a RING, a B-box and a coiled-coil domain. RING domains are cysteine and histidine-rich motifs that bind two zinc atoms and mediate protein-protein interactions. B-boxes are domains with yet unknown function. Coiled-coiled domains also mediate protein-protein interactions and are important for homo and heterodimerization of TRIM family members. Furthermore, the self-association properties of the coiled-coil region are responsible for the high molecular weight complexes formed by TRIM proteins *in vivo* (Reymond et al. 2001). A very interesting property of TRIM proteins is their characteristic localization to discrete cytoplasmic or nuclear structures. In this context it has been suggested that TRIM proteins define a set of different cellular compartments as a consequence of their propensity to form high-order molecular structures. Notably, the TRIM proteins have distinct patterns of expression and are not functionally interchangeable, suggesting that they are single functional entities. In a recent publication, Lerner et al. (Lerner et al. 2007) showed that RFP2 is an E3 ubiquitin ligase involved in endoplasmic reticulum-associated degradation (ERAD) and that is associated with the endoplasmic reticulum (ER) membrane through its

transmembrane domain. A remarkable property of RFP2 is its inherent instability due to self-degradation (Lerner et al. 2007) which makes it difficult to functionally characterize it biochemically. In addition to these functions, work by Tschuch et al. (Tschuch 2006) indicates a function of RFP2 in apoptosis and additional cellular processes. Therefore it would be interesting to further characterize other cellular pathways RFP2 might be involved in, and whether its ER-related activity could be linked to malignancy in CLL.

1.3.1.3. SETDB2 — a putative histone methyltransferase

The gene for *SETDB2* (also known as *CLLD8*) is the candidate gene farthest away from the minimally deleted region that was characterized in this work (Figure 1.7), and it codes for a protein of 719 amino acids. Its closest related human protein is SETDB1, to which it shows 45% identity and 61% similarity, with a very similar overall domain structure organization. SETDB2 as well as SETDB1 contain three conserved domains: a methyl-CpG binding domain, a preSET domain and a bifurcated SET domain (a SET domain with an insertion) (Figure 1.12).

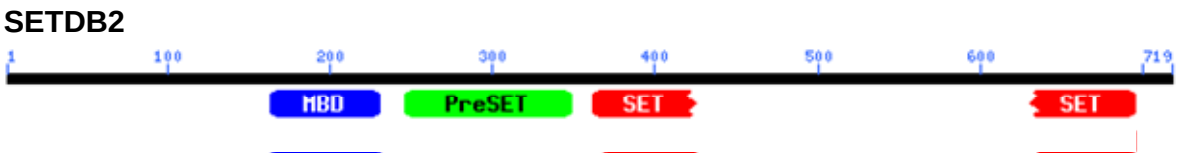


Figure 1.12 Schematic representation of the SETDB2 protein coding sequence. Depicted are the homology domains: methyl binding domain (blue box), preSET (green box) and bifurcated SET domain (red box). Numbers indicate the amino acids positions. http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

In mammals, the SET domain defines a family of highly modular proteins with histone methyltransferase (HMTase) activity. They catalyze methylation of the N-terminus of histone proteins, and this activity is based on the combination of the SET domain with adjacent cysteine-rich regions (known as preSET domain, (Jenuwein 2001)). Histones are evolutionarily highly conserved small globular proteins that form the central structure of DNA, the nucleosome (Kornberg & Thomas 1974, Olins & Olins 1974). Histones are involved in gene regulation through the modification of their N-terminus by addition of different chemical groups (like phosphate, methyl, acetyl, etc.), which are highly dynamic modifications and allow for both transient and long-term alterations in chromatin structure (Rea et al. 2000). These modifications enable the cell to adjust gene activity, but at the same time also maintain specific gene expression programs (Jenuwein 2001). Distinct patterns of methylation at specific

residues in the histone tails produce heritable marks that are used to recruit repressing or activating protein complexes (Breiling & Orlando 2002). For example, SETDB1 has been shown to play a key role in the organization of heterochromatin by specifically methylating lysine residues at position 9 in histone H3 (Schultz et al. 2002) and thereby inducing the recruitment of repressing complexes (Rea et al. 2000). The presence of the HMTase domains in SETDB2 points to a role as chromatin modifying protein and an involvement in transcriptional regulation. Therefore it would be crucial to further characterize its cellular function.

1.3.1.4. RCBTB1 and guanine exchange factors

The *RCBTB1* gene (also known as *CLLD7*, *GLP* or *E4.5*) is located in close proximity of the minimally deleted region in 13q14 (Figure 1.7) and codes for a protein of 531 amino acids. It contains two conserved domains: an N-terminal RCC1 domain and a C-terminal BTB/POZ domain (Figure 1.13).



Figure 1.13 Schematic representation of the RCBTB1 protein coding sequence. Depicted are the homology domains: RCC1 (blue box) and BTB/POZ (red box). Numbers indicate amino acids. http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

RCC1-containing proteins are guanine-nucleotide exchange factors and it seems that the RCC1 domain is essential for the catalytic activity (Renault et al. 1998). The BTB/POZ domain mediates homomeric or heteromeric interactions and is a protein-protein interaction domain. Thus, RCBTB1 seems to be a member of the guanine-nucleotide exchange factors (GEFs). GEFs target mainly members of the Ras superfamily of small monomeric GTP-binding (G) proteins. Small monomeric G–proteins act as molecular switches and control a wide variety of cellular functions (like growth and division, cell morphology, vesicular transport, etc.) (Day et al. 1998). The activity of all known small GTP-binding proteins is regulated by guanine nucleotide exchange and GTP hydrolysis, making GEFs crucial factors in the signalling cascades, where G-proteins are involved. In the work of Solomou et al. (Solomou et al. 2003), exon trapping was performed to identify novel genes in 13q14 and the *RCBTB1/E4.5* gene was isolated. Based on homology, it was speculated that RCBTB1/E4.5 could be involved in the regulation of cell cycle. In another study RCBTB1/GLP was associated with hormone responses (Guo et al. 2004). RCBTB1

was shown to localize in cytoplasmic bodies and to play a role in signaling cascades inducing hypertrophy of cells. An involvement of RCBTB1 in signaling pathways and cell cycle could be related to the development of malignancy and should be further characterized.

1.3.1.5. KPNA3 and nuclear import

The gene for Karyopherin α-3, *KPNA3*, is also located in the vicinity of the minimally deleted region in 13q14 (Figure 1.7) and codes for a protein of 521 amino acids. It contains a highly basic N-terminus and a strongly acidic C-terminal region (Takeda et al. 1997). KPNA3 belongs to the importin protein family involved in the transport of cargo from the cytoplasm into the nucleus. KPNA3 contains two homology domains: the importin-beta binding domain at the N-terminus that is involved in the interaction of KPNA3 with the importin-beta protein, and several armadillo repeats that are required for protein-protein interactions (Figure 1.14).



Figure 1.14 Schematic representation of the KPNA3 protein coding sequence. Depicted are the homology domains: improtin-beta binding domain (IBB) (red box) and armadillo repeats (ARM) (blue boxes). Numbers indicate amino acids. http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

The nucleo-cytoplasmic traffic of large molecules in the cell is mediated by specific nuclear import and export complexes. The importin-α and -β heterodimers are responsible for the import of proteins that contain classical nuclear-localization signals (NLSs). In heterodimerization, importin-α binds to the NLS (through the armadillo repeats) (Talcott & Moore 2000), and importin-β is responsible for association of the importin-cargo complex to the cytoplasmic side of the nuclear pore (Fagerlund et al. 2005). The process of loading and release of cargo is controlled by a gradient of RanGTP across the nuclear envelope: in the cytoplasm, low levels of RanGTP favor importin-cargo complex formation, while in the nucleus a high concentration of RanGTP favors cargo displacement from importins (Jäkel et al. 2002). KPNA3 has been shown to specifically be involved in the nuclear transport of RCC1 (the guanine exchange factor for Ran) (Talcott& Moore 2000). It is also probably involved in the import of NF-κB (a transcription factor), but due to difficulties in expressing KPNA3, experiments could only be performed with KPNA4 (Fagerlund et al. 2005). The correlation of KPNA3 with either cargoes would be very interesting

as they are involved in cellular responses like cell cycle control (RCC1), cell differentiation and apoptosis (NF-κB). Therefore, to get deeper insight into the possible role of KPNA3 in malignancy, it would be very interesting to further characterize its cellular functions.

1.3.1.6. Potential interactions among 13q14 candidate proteins

The function of domains of 13q14 proteins as deduced from homology suggests complex relationships among the proteins of the critical region. From interactions between proteins homologous to the 13q14 candidates it can be inferred that some 13q14 proteins might be part of the same molecular complexes or networks, or participate in similar cellular functions. The association of proteins in protein complexes enables the cell to perform highly complex processes, which take place in an ordered and regulated manner. These complexes can be of high molecular weight and involve as many as 88 proteins in yeast (Shi-Hua Zhang et al. 2007). Because of their homology to interacting proteins, gene products from 13q14 might have similar function and/or might interact with each other. Relations were found among 4 of the 13q14 proteins under study and are summarized in Figure 1.15 and described below.

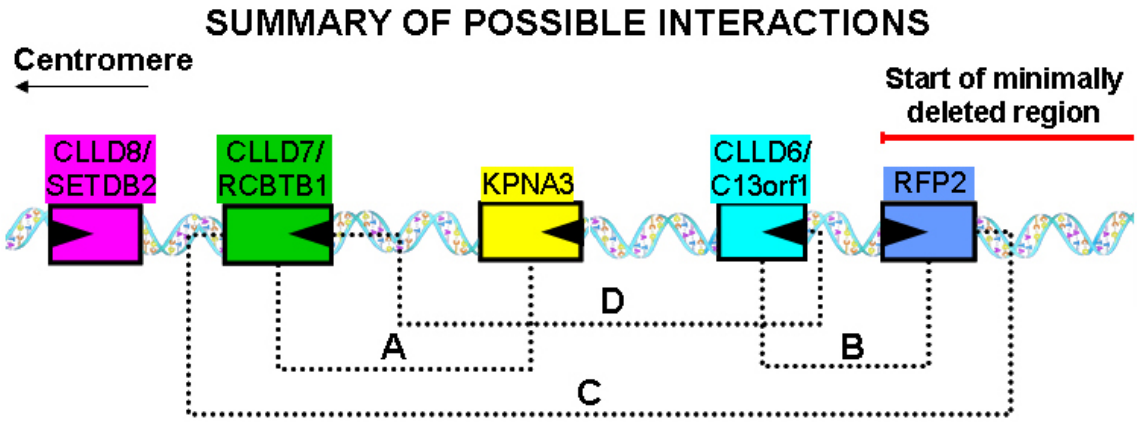


Figure 1.15 Possible interactions between 13q14 proteins. The interactions based on homology to non-13q14 proteins are indicated by dotted lines. Interactions are depicted in more detail below (A=Figure 1.16, B=Figure 1.17, C=Figure 1.18 and D=Figure 1.19).

One of the possible interactions is between KPNA3 and RCBTB1 (Figure 1.15 A and 1.16). This interaction is based on the finding that KPNA3 is an import protein for the guanine nucleotide exchange factor RCC1 (Talcott& Moore 2000), which in turn has high homology to the 13q14 protein RCBTB1, a putative guanine nucleotide exchange factor as well.

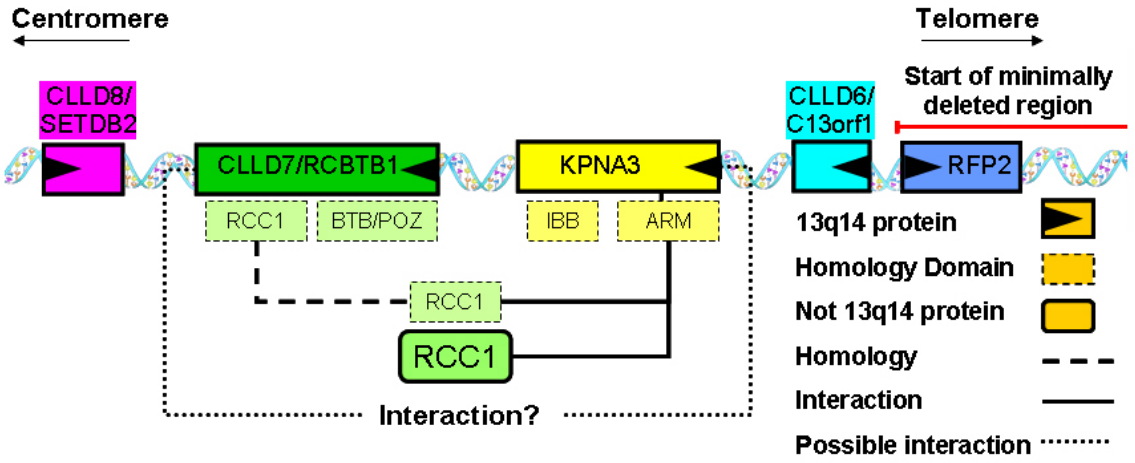


Figure 1.16 Possible interaction between RCBTB1 and KPNA3. The possible interaction is based on the biochemically found interaction between RCC1 and KPNA3. RCC1 and KPNA3 interact through the armadillo domain (ARM) and RCC1 domain respectively. The presence of the RCC1 homology domain in RCBTB1 suggests the possibility of interaction between RCBTB1 and KPNA3. Distances and sizes of genes are not drawn to scale.

Another 13q14 protein, RFP2, is a TRIM protein. The finding that some TRIM proteins like TRIM35 have been shown to be involved in apoptosis (Kimura et al. 2003), and that TRIM35 has a TRIM domain like RFP2 and a C-terminal SPRY domain like C13orf1 (another 13q14 protein), points to possible interactions of RFP2 and C13orf1 (Figure 1.17).

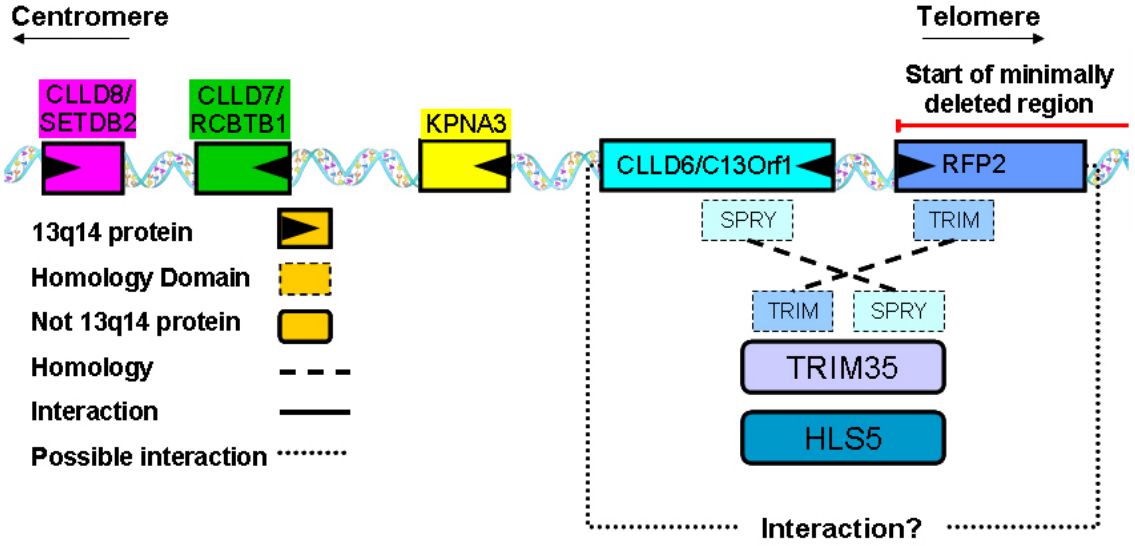


Figure 1.17 Possible interaction between RFP2 and C13orf1. The likely interaction is based on the presence of homologous domains from both 13q14 proteins in two single non-13q14 proteins. The two non-13q14 proteins TRIM35 and HLS5 both contain the two homology domains TRIM and SPRY, which are present in RFP2 and C13orf1 respectively. Distances and sizes of genes are not drawn to scale.

In addition, another TRIM protein, HLS5 (also containing a TRIM domain and a SPRY domain) is a candidate tumor suppressor gene in hematopoietic cells (Lalonde et al. 2004), further pointing to a possible functional association of RFP2 and C13orf1

and the tumor suppressor mechanism in CLL. Interestingly, another TRIM protein, TRIM5δ, has been shown to co-localize with two BTB/POZ containing proteins, BTBD1 and BTBD2. In their work, Xu et al. (Xu et al. 2003) could show that the interaction of TRIM5δ and BTBD1/2 is mediated through the TRIM domain and BTB/POZ domains. The presence of RCBTB1 (a BTB/POZ protein) and RFP2 (a TRIM protein) close to each other in the 13q14 region raises the possibility of an interaction of the two proteins that could also form a multi-subunit complex (Figure 1.18).

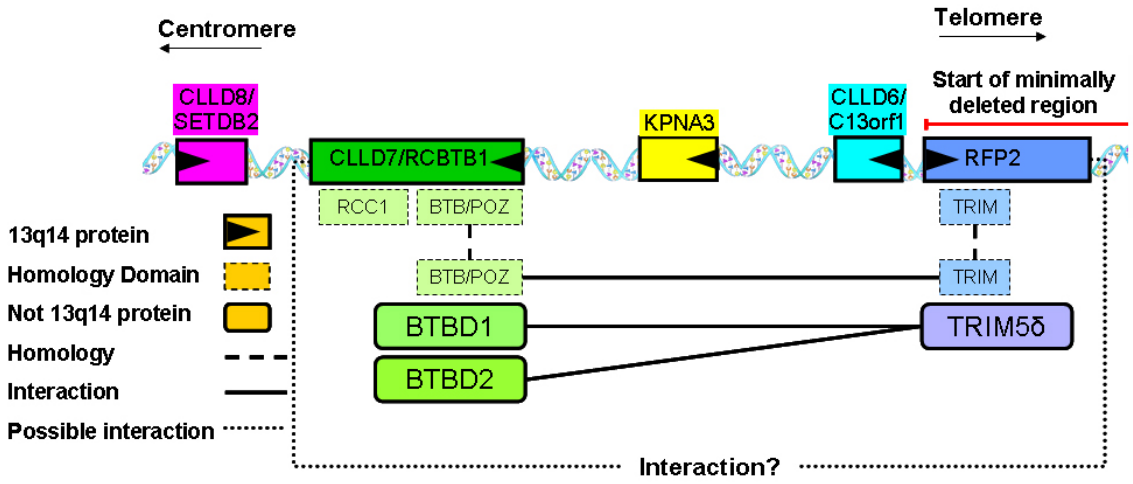


Figure 1.18 Possible interaction between RCBTB1 and RFP2. The possible interaction is based on the interaction of TRIM5δ with BTBD1 and BTBD2. Interaction requires the BTB/POZ domain in BTBD1 and BTBD2 and the TRIM domain in TRIM5δ. The presence of the interacting domains BTB/POZ and TRIM in RCBTB1 and RFP2 respectively suggest an interaction of RFP2 and RCBTB1 in 13q14. Distances and sizes of genes are not drawn to scale.

Further supporting the possibility that several 13q14 proteins interact is the existence of a giant protein (p619) that features two protein domains contained in two distinct 13q14 proteins. Among the different functional domains of p619, there are two RCC1-homologous domains (RLD-1 and RLD-2) with activity as guanine nucleotide exchange factors (Rosa et al. 1996). In addition, p619 has a polar stretch with homology to the SPRY domain. The fact that these domains are present in a single protein with multiple functions raises the possibility that RCBTB1 and C13orf1 are part of a multi-protein complex (Figure 1.19).

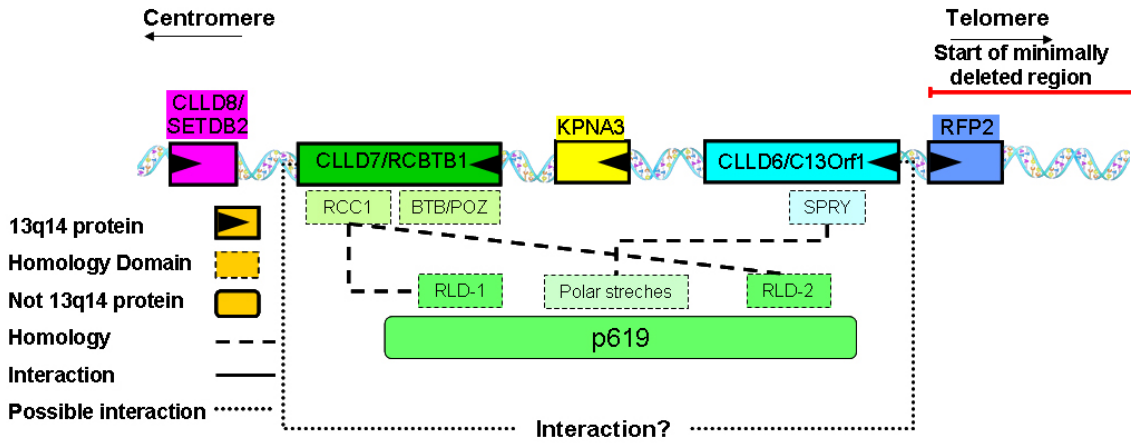


Figure 1.19 Possible interaction between RCBTB1 and C13orf1. The possible interaction is based on the presence of homology domains from two 13q14 proteins in one single protein. The protein p619 contains homology domains to RCC1 (RLD-1 and -2) and to SPRY (polar stretches), each one present in RCBTB1 and C13orf1 respectively. Distances and sizes of genes are not drawn to scale.

The presence of domains in distinct 13q14 proteins with homology to domains in other known proteins and protein complexes are intriguing, and raise the possibility that 13q14 proteins might be involved in similar cellular processes and could form a functional cluster. This is uncommon in the mammalian genome and occurs mostly in genes that are coordinately regulated like the globin- and Hox-gene clusters, histone genes and imprinted genes. The existence of a complex epigenetic regulatory mechanism in 13q14 that might regulate transcription of the candidate genes coordinately and the potential functional correlation of the gene products points to a similar organization in 13q14. The functional characterization of the 13q14 proteins will shed light onto their possible involvement in the same cellular processes.

1.4. Aims of this work

In this study, five protein-coding genes from the critical region in 13q14 were functionally analyzed. These genes were selected for analysis because of a) their localization with respects to the minimally deleted region in B-CLL patients, and b) their homology to other proteins involved in cellular pathways relevant for the development of malignant phenotypes. The hypothesis that some of these genes might be part of the same protein complexes and be involved in common cellular functions was also tested. To better characterize the proteins C13orf1, RCBTB1, SETDB2, RFP2 and KPNA3, two strategies were employed: a) identification of the intracellular localization of the proteins and, b) finding interaction partners for the proteins. To analyze the intracellular localization, GFP-fusion proteins were over-expressed and specific antibodies were raised to assess endogenous proteins. Cellular functions of the proteins were inferred from the pathways their binding partners are involved in, and binding partners were identified in a yeast-two-hybrid screen. Ultimately, elucidation of the cellular localization and functions of these proteins will contribute to understand the cellular mechanisms and pathways the genes from the critical region in 13q14 are involved in, and shed light on the molecular pathomechanism of B-CLL.

2. Material and Methods

2.1. Material

2.1.1. Primers for PCR

Primers for PCR are listed in Appendix 1.

2.1.2. Commercial Vectors

Vector Name	Supplier
pcDNA3.1 XL	Invitrogen, Karlsruhe
pCRT7-CT	Invitrogen, Karlsruhe
pTrcHis-NT	Invitrogen, Karlsruhe
pDEST-14	Invitrogen, Karlsruhe
pDEST-15 (NT)	Invitrogen, Karlsruhe
pDEST-17 (NT)	Invitrogen, Karlsruhe
pDEST53 (NT)	Invitrogen, Karlsruhe
pET43.1a (NT)	DB Bioscience Clontech, Heidelberg
pDS5	DB Bioscience Clontech, Heidelberg

2.1.3. Enzymes

Enzyme	Supplier
Ampli Taq DNA polymerase	Applied Biosystems, Weiterstadt
Benzonase Nuclease	Novagen/Merck, Darmstadt
High Fidelity polymerase	Roche Diagnostics, Mannheim
DNAse I	Roche Diagnostics, Mannheim
Restriction endonucleases:	Roche Diagnostics, Mannheim and
Acc65I=Asp718	New England Biolabs, Frankfurt am Mein
BamHI	
BspEI	
Eco RI	
Eco RV	
HindIII	
NcoI	
NotI	
NheI	
SalI	
XbaI	
XhoI	
XmaI	
Shrim alkaline phosphatase	Roche Diagnostics, Mannheim
T4 DNA ligase	New England Biolabs, Frankfurt am Mein

2.1.4. Reagents

Reagents	Supplier
Acetone	Merck, Darmstadt
Acrylamide/Bis-Acrylamide (30% w/v)	Bio-Rad Laboratories
Agarose	Sigma-Aldrich, Munich
Ampicillin	Roche Diagnostics, Mannheim
Bacto Agar	Difco Laboratories
Bacto Tryptone	Difco Laboratories
Bacto Yeast Extract	Difco Laboratories
Bicinchonic acid	Invitrogen, Karlsruhe
Bovine Serum Albumin (BSA)	New England Biolabs, Frankfurt am Mein
Canamycin	Roche Diagnostics, Mannheim
Complete EDTA-free Protease Inhibitor Cocktail Tablets	Roche Diagnostics, Mannheim
Complete Freud's adjuvant	Sigma-Aldrich, Munich
Coomassie brilliant blue R 250	Merck, Darmstadt
Copper-sulphate	Invitrogen, Karlsruhe
Deoxynucleotide set (dNTPs)	Amersham Bioscience, Freiburg
Dithiothreritol (DTT)	Sigma-Aldrich, Munich
Disodium Phosphate (Na_2HPO_4)	Merck, Darmstadt
DNA mass standard ladder	Gibco BRL/Invitrogen, Karlsruhe
EDTA	Merck, Darmstadt
Ethanol	Merck, Darmstadt
Ethidium Bromide	Sigma-Aldrich, Munich
Fluoromount-G	Southernbiotech
Formamide	Merck, Darmstadt
Glucose	Merck, Darmstadt
Glycerine	Roth
HCl	Merck, Darmstadt
Hepes	Sigma-Aldrich, Heidelberg
Herring Testes Carrier DNA	Clontech
Incomplete Freud's adjuvant	Sigma-Aldrich, Munich
Isoamylalcohol	Merck, Darmstadt
Isopropyl β-D Thiogalactosidase (IPTG)	Fluka/Sigma-Aldrich, Munich
Lithium Acetate (LiAc)	Sigma-Aldrich, Munich
Magnesium chloride ($MgCl_2$)	Merck, Darmstadt
Methanol	Merck, Darmstadt
NaOH	Merck, Darmstadt
Ni-NTA Agarose	Qiagen, Hilden
Nonidet P-40	Amersham Biosciences, Freiburg
Nuclease-free water	Ambion, Austin, USA
PEG 1500 (50% w/v)	Roche Diagnostics, Mannheim
Ponceau S	Sigma-Aldrich, Munich
Poly-dT primer	Roche Diagnostics, Mannheim
Phenylmethanesulfonyl fluoride (PMSF)	Fluka/Sigma-Aldrich, Munich
Rainbow Molecular Weight Marker	Amersham Biosciences, Freiburg
SDS (Sodium-Dodecyl-Sulfate)	Sigma-Aldrich, Munich
See Blue pre-stained standard	Invitrogen, Karlsruhe

Reagents	Supplier
See Blue Plus2 pre-stained standard	Invitrogen, Karlsruhe
Skim Milk Powder	Roth
SOC medium	Invitrogen, Karlsruhe
Sodium Acetate (NaAc)	Merck, Darmstadt
Sodium Citrate	Merck, Darmstadt
Sodium Chloride (NaCl)	Merck, Darmstadt
Sodium Phosphate (NaH_2PO_4)	Merck, Darmstadt
Sucrose	Merck, Darmstadt
TEMED (N,N,N,N-Tetramethylendiethylamin)	Bio-Rad Laboratories, Munich
T4 gene 32 protein	Roche Diagnostics, Mannheim
Tris Base	Gibco BRL/Invitrogen, Karlsruhe
Trypsin	Invitrogen, Karlsruhe
Tween-20 (Polyoxy-ethylene-sorbitan-monolaurate)	Sigma-Aldrich, Munich
Urea	Ambion, Texas
VECTASHIELD without DAPI	Linaris, Germany
2-Mercaptoethanol (β-ME)	Sigma-Aldrich, Munich
2-Propanol (Isopropanol)	Merck, Darmstadt

2.1.5. Commercial Antibodies

Antibody	Supplier
Anti-GFP	Roche Diagnostics, Mannheim
Anti-HA.11 monoclonal antibody	Covance, California
Anti-His_6	Roche Diagnostics, Mannheim
Anti-Protein A antibody	Covance, California
Anti-V5 Antibody	Invitrogen, Karlsruhe
Anti-Xpress Antibody	Invitrogen, Karlsruhe
Cy3-conjugated anti-guinea pig IgG	Dianova
Cy3-conjugated anti-mouse IgG	Dianova
Cy3-conjugated anti-rabbit IgG	Dianova
FITC-conjugated anti-mouse IgG	Dianova
FITC-conjugated anti-rabbit IgG	Dianova
Mouse monoclonal to beta Actin (ab3280)	Biozol/Abcam
Nus·Tag Monoclonal Antibody	Novagen, Darmstadt
Peroxidase-conjugated anti-guinea pig IgG	Dianova
Peroxidase-conjugated anti-mouse IgG	Dianova
Peroxidase-conjugated anti-mouse IgG	Amersham Biosciences, Freiburg
Peroxidase-conjugated anti-rabbit IgG	Amersham Biosciences, Freiburg
Peroxidase-conjugated anti-rabbit IgG	Dianova

2.1.6. Standard Buffers and solutions

- 5x TBE pH 8
  0,455M Tris-Borate
  10mM EDTA
- 5xLoading buffer for nucleotides electrophoresis
  100mM EDTA
  30% (v/v) Glycerin
  0,25% (w/v) Bromphenolblue
- 2x Laemmli loading buffer
  50mM Tris-HCl pH 8
  2,4% (w/v) SDS
  8% Glycerin
  0,2% (w/v) Bromphenolblue
- 5x SDS-PAGE run buffer
  25mM Tris Base
  0,2M Glycine
  10% SDS
- 10x Blotting Buffer
  25mM Tris base
  192mM Glycine

For Western blot analysis, the 1x blotting buffer was prepared fresh with 20% Methanol.

- TBS-T
  20mM Tris Base
  137mM NaCl
  3,8mM HCl
  0,1% Tween-20
- 1xPBS pH 7,4
  2,7mM KCl
  1,7mM KH_2PO_4
  137mM NaCl
  10mM Na_2HPO_4

2.1.7. Renaturation buffers for purified proteins

- Tris buffer pH 7,5
  50mM Tris
  2mM Mg-Oxalacetate
  5% Glycerol
- Phosphate buffer pH 7
  57,7mM Na_2HPO_4
  42,3mM NaH_2PO_4
  2mM Mg-Oxalacetate

5% Glycerol

- Phosphate buffer pH 8 — 93,2mM Na_2HPO_4
  8,80mM NaH_2PO_4
  2mM Mg-Oxalacetate
  5% Glycerol

For protein renaturation, 12 different buffers were prepared using the three buffers indicated above and adding to each one NaCl accordingly to obtain 20mM, 100mM, 200mM and 500mM NaCl concentrations in each buffer. Prior to dialysis of protein solutions, 0,5mM DTT was added to each buffer.

2.1.8. Buffers and solutions for cell extracts and lysates

- RIPA — 150,0 mM NaCl
  50,0 mM Tris/HCl, pH 8,0
  1,0 mM EGTA
  1,0 mM EDTA
  0,25 % Natrium-Desoxycholate
  1,0 % NP-40
  0,1 % SDS
  $H_2O_{Milli\text{-}Q.}$
- Hepes — 50mM Hepes
  150mM NaCl
  10% Glycerin
  0,2% NP-40
  20mM NaF
  1,5mM $MgCl_2$
  0,2mM EDTA
  $H_2O_{Milli\text{-}Q.}$
- Tris buffer — 20mM Tris (pH 7.5)
  150mM NaCl
  1mM EDTA
  1mM EGTA
  1% Triton X-100
  5mM K_2HPO_4

$H_2O_{Milli\text{-}Q.}$

To all cell lysis buffers the following components were freshly added:

- 1µl/ml Benzonase
- 1mM PMSF
- 1tablet/50 ml Complete EDTA-free Protease inhibitor cocktail (Roche)

Washing-buffer for pull-down assays was the same as the lysis buffer but without the components that were added fresh.

- Bacteria lysis buffer pH 7.5 — 100mM Na_3PO_4, 150mM NaCl, 2mM Mg-Oxalacetate, 5% Glycerol

2.1.9. Solutions for yeast transformation

- MIX1 — 1ml 1M LiAc, 0.5 ml 10xTE, 5ml 1M Sorbitol, 3,5ml ddH_2O
- MIX2 — 1.5ml 10x TE, 10ml PEG, 2ml ddH_2O

2.1.10. Growth media

- LB-Medium — 1 % (w/v) Bacto-Tryptone, 0,5 % (w/v) Bacto-Yeast-Extract, 1 % (w/v) NaCl, Autoclave
- YPD medium — 2% peptone, 1% yeast extract, 2% glucose, 20mg/l adenine, 20mg/l arginine, 20mg/l methionine, 20mg/l tyrosine, 20mg/l lysine

20mg/l isoleucine
20mg/l histidine
20mg/l uracil
20mg/l leucine
20mg/l tryptophan

- SDII medium: 2% peptone, 1% yeast extract, 2% glucose, 20mg/l adenine, 20mg/l arginine, 20mg/l methionine, 20mg/l tyrosine, 20mg/l lysine, 20mg/l isoleucine, 20mg/l histidine, 20mg/l uracil
- SDIV medium: 2% peptone, 1% yeast extract, 2% glucose, 20mg/l adenine, 20mg/l arginine, 20mg/l methionine, 20mg/l tyrosine, 20mg/l lysine, 20mg/l isoleucine

For preparing plates used for robotic spotting 1,5% agarose was added to the corresponding medium.

2.1.11. Cell lines

Cell lines are listed in Appendix 2.

2.1.12. Cell culture media and reagents

Medium/Reagent	Supplier
DMEM (Dulbecco's Modified Eagle Medium)	Gibco BRL/Invitrogen, Karlsruhe
RPMI	Gibco BRL/Invitrogen, Karlsruhe
IMEM (Iscove's, Modified Eagle Medium)	Gibco BRL/Invitrogen, Karlsruhe
HAT	Gibco BRL/Invitrogen, Karlsruhe
Hybridoma Express medium	PAA, Austria
FCS (fetal cow serum)	Biochrom
Penicillin (10000U/ml)	Gibco BRL/Invitrogen, Karlsruhe
Streptomycin (100µg/ml)	Gibco BRL/Invitrogen, Karlsruhe
L-Glutamine (200mM)	Gibco BRL/Invitrogen, Karlsruhe
Penstrep (10000U/ml)	Gibco BRL/Invitrogen, Karlsruhe
Na-Pyruvate (100mM)	Gibco BRL/Invitrogen, Karlsruhe
100x Non-essential Aminoacids	Gibco BRL/Invitrogen, Karlsruhe
Dimetyl sulphoxide (DMSO)	Sigma-Aldrich, Munich

- RPMI-complete
  - 90% RPMI
  - 10% FCS
  - 1mM L-Glutamine
  - 100u/ml Penstrep
  - 0,5mM Na-pyruvate
  - 1x Non-essential Aminoacids

2.1.13. Kits

Kit name	Supplier
BCA Protein Assay Kit	Pierce, München
ECL Western blot detection kit	Amersham Bioscience, Freiburg
Effectene transfection reagent	Qiagen, Hilden
High puritiy plasmid purification system	Marligen Biosciences, Inc, Ijamsville, MD
Lipofectamine 2000 transfection reagent	Invitrogen, Karlsruhe
PCR rapid purification kit	Marligen Biosciences, Inc
Plasmid Maxi Kit, Qiagen	Qiagen, Hilden
Plasmid Mini kit, Qiagen	Qiagen, Hilden
PRISM Big DyeDeoxy Terminator Kit	PE Applied Biosystems
Proteus G Midi kit	Prochem
QIAEX II gel extraction kit	Qiagen, Hilden
Rapid Translation System RTS 100, E.coli HY Kit	Roche Diagnostics, Mannheim
Rapid Translation System RTS E.coli Linear Template Generation Set, His tag	Roche Diagnostics, Mannheim
RNeasy-Mini Kit, Qiagen	Qiagen, Hilden
Transfection Kit (Nanofectin)	PAA, Austria

2.1.14. Other materials

Material	Supplier
Extra thick blot paper	Bio-Rad Laboratories, Munich
Hyperfilm ECL	Amersham, Freiburg
Immobilon^{TM}-P-PVDF Membran	Millipore, Schwalbach
MicroAmp Optical 96-Well Reaction Plates	PE Applied Biosystems, Weiterstadt
MicroAmp Optical Caps	PE Applied Biosystems, Weiterstadt
Microcentrifuge tubes, PCR clean, 0,5ml	Eppendorf, Hamburg
Microcentrifuge tubes, PCR clean, 1,5ml	Eppendorf, Hamburg
Microcentrifuge tubes, PCR clean, 2,0ml	Eppendorf, Hamburg
PCR tubes, 0,2ml	Molecular BioProducts, San Diego, USA
Protran Nitrocellulose Membrane	Whatman, Maidstone, UK
Sterile Filters	Millipore, Eschborn
Vacutainer Blood collection tubes	DB Bioscience Clontech, Heidelberg

2.1.15. Instruments

Instrument	Supplier
ABI PRISM 3100 Genetic Analyzer, 16 Capillary DNA Sequencer	PE Applied Biosystems, Langen
ABI PRISM 7900 Sequence Detection System	Applied Biosystems, Forster City, USA
ABI PRISM 7700 Sequence Detector	PE Applied Biosystems, Weiterstadt
Automatic developing machine	Amersham, Freiburg
Biofuge Fresco refrigerated tabletop centrifuge	Heraeus/Kendro, Hanau
Cary 50 Bio UV-photometer	Varian, Darmstadt
Centrifuge 5810 R	Eppendorf, Köln
Cytospin 3	Shandon Life Sciences International, England
EAS Gel Documentation System	Herolab, Wiesloch
EL800 Universal Microplate Reader	BIO-TEK, Neufahrn
Gel electrophoresis power supply	E-C Apparatus Corporation, USA
GeneAmp PCR-System 2400, 9600 and 9700	PE Applied Biosystems, Weiterstadt
Gene Pulser Xcell	Bio-Rad Laboratories, Munich
GS 6, GS 6KR, Centrifuge	Beckmann, Wiesloch
GSA- and SS34-Rotor	DuPont, Boston, USA
Heating block QBT2	Grant Instruments/CLF, Emersacker
HMT 702 C Microwave oven	Robert Bosch GmbH, Stuttgart
Mastercycler PCR-Maschine	Eppendorf, Köln
Microscope Wilovert S	Hund, Köln
Micro-centrifuge	NeoLab Laborbedarf, Heidelberg

Instrument	Supplier
Mini Trans-Blot Electrophoretic Transfer Cell	Bio-Rad Laboratories, Munich
Mini-PROTEAN 3-Dodeca cell	Bio-Rad Laboratories, Munich
Mini-Protean 3 gel and electrophoresis system	Bio-Rad Laboratories, Munich
Multifuge 3 SR	Heraeus / Kendro, Hanau
NanoDrop ND-1000 Spectrophotometer	NanoDrop Technologies, San Diego
Semi-dry Transfer Cell	Bio-Rad Laboratories, Munich
Servapor dialysis tubing	Serva Electrophoresis GmbH, Heidelberg
Sigma 4K10-Coldcentrifuge	Sigma-Aldrich, Heidelberg
Thermomixer compact	Eppendorf, Köln
Ultrospec 2000 Photometer	Amersham, Freiburg
Unimax 1010 Shaker	Heidolph Instruments, Schwabach
Varifuge 3.0/3.0R Laborcentrifuge	Heraeus/Kendro, Hanau
Waterbad SW22	Julabo Labortechnik, Seelbach
Waterfiltrationsmachine Millipore-Q UF Plus	Millipore, Eschborn
Weighing machine BL 610 und BL150S	Sartorius AG, Göttingen

2.2. Methods

2.2.1. Nucleic acids

Isolation of total RNA

Total RNA was isolated for preparing of cDNA in order to clone genes or for Real Time-PCR quantification of expression of genes. In both cases, RNA was isolated from cell pellets containing $5x10^5$ to 10^7 cells using the RNeasy-Mini kit from Qiagen following manufacturer instructions.

Isolation of plasmid DNA

Upon transformation of bacteria, colonies containing the plasmid of interest were picked and grown O/N at 37ºC in selective LB-medium in a shaking incubator at 220rpm. Isolation of plasmids was performed following the manufacturer protocol and using 400ml transfected bacteria culture for MAXI-PREPs and 5ml for MINI-PREPs (Qiagen Plasmid Maxi Kit, and Qiagen Plasmid Mini kit, Qiagen).

For transfection of human cell lines (section 2.2.9), plasmids with the least RNA and protein contaminations perform better. For this purpose, plasmids were purified with the "High purity plasmid purification system" kit from Marligene, following the manufacturer instructions.

Quantification of DNA and RNA

After purification of DNA and RNA molecules their concentration and purity was measured based on light absorbtion. Nucleic acids absorb light because they have aromatic ring structures which have a maximum absorption peak at 260nm wavelength. Their concentration can be estimated with a spectrophotometer by measuring the optical density (OD) at this wavelength and considering average absorptivity constants as follow:

- a solution containing pure double stranded DNA with OD=1 contains approximately 50µg/ml double stranded DNA
- a solution containing pure single stranded DNA with OD=1 contains approximately 33µg/ml single stranded DNA
- a solution containing pure RNA with OD=1 contains approximately 40µg/ml RNA

The purity of the sample can be assessed by measuring the absorbance at 280nm, as aromatic amino acids in proteins have a maximum absorbance at this wavelength. By calculating the ratio OD_{260}/OD_{280}, the purity of the DNA and RNA can be estimated. Pure DNA solutions have a ratio of 1.8, while pure RNA solutions have a ratio of around 2.0. Contamination with phenol or proteins cause a decrease of this ratio (pure protein solution have an OD=0.6). For plasmids that will subsequently be used for transfections of human cells, the purity has to be very high (1.75 to 1.85), while for sequencing or restriction digest less pure molecules can be used. DNA with an $OD_{260/280}$ ratio of below 1.6 was discarded and purification was repeated. Measurements of optical density were performed with an automatic spectrophotometer reader (NanoDrop)

Gel electrophoresis and analysis of DNA

Nucleic acid molecules are charged in neutral buffer solutions, therefore they can be separated according to their size by transfer through an agarose or polyacrylamide gel in an electrical field. At a pH of 7 DNA molecules are homogeneously negatively charged through their phosphate groups, so they will migrate from the anode to the cathode. Smaller molecules will be able to migrate faster, while bigger molecules will migrate slower through the matrix of the polymer. 1% agarose gels were prepared by melting the agarose in 1x TBE buffer. DNA solutions were mixed with 5x loading buffer and loaded in the wells. Electrophoresis was run at 140V for 20min for separating fragments of 0.5 to 5Kb in size. After electrophoresis the DNA in the gel was stained by soaking the gel for 10min into a solution of 1µg/ml ethidium-bromide at RT. Ethidium-bromide (3,8-Diamino-5-ethyl-6-phenylphenanthridinium bromide) is an organic dye with a planar structure that intercalates into DNA strands and emits fluorescence when excited with UV[5]. Images were taken with a digital camera.

Purification of nucleic acids and gel extraction

PCR products and plasmids were purified to remove enzymes, salts and other contaminants. Nucleic acid molecules were purified with a PCR purification kit following manufacturer protocol and eluted (PCR rapid purification kit, Marligen Biosciences, Inc.). For isolation of nucleic acids of distinct size, bands were cut from gels after electrophoresis. For gel extraction purification, nucleic acid were detected

[5] http://en.wikipedia.org/wiki/Ethidium_bromide

by low wavelength UV light illumination in order to avoid DNA-damage. Gel slices were processed with the "QIAEX II gel extraction" kit following manufacturer instructions (Qiagen, Hilden).

2.2.2. PCR

The polymerase chain reaction (PCR) is a technique to massively copy a specific DNA fragment (Kleppe et al. 1971, Mullis et al. 1986). It is based on the ability of DNA-polymerases to generate complementary strands of DNA out of a single stranded DNA template. The enzymes used can only elongate existing short oligonucleotide, called "primers" of 15-25 bp in length. The complementarity of the primers to a specific DNA sequence determines the fragment of DNA that will be amplified. In this work the PCR technique was used for amplifying genes of interest for cloning, for generation of templates for restriction cloning into the appropriate vectors (see section 2.2.3.2) and for testing of the cloned vectors for appropriate insert.

Primer design

Primers were designed for the amplification of the genes of interest in the region 13q14 (genes' IDs from NCBI are on table 3.1). There were various purposes for which primers were designed (Figure 2.1):

- for full-length gene amplification: primers were designed according to the gene sequences obtained from NCBI (table 3.1).
- for restriction cloning of the genes: primers were designed containing the restriction enzyme specific sequence.
- for sequencing of full length gene inserts: internal primers were designed according to the gene sequences obtained from NCBI (table 3.1)
- for nested PCR: primers were designed in the untranslated regions (UTR) of the genes

For a detailed list of primer sequences see Appendix 1.

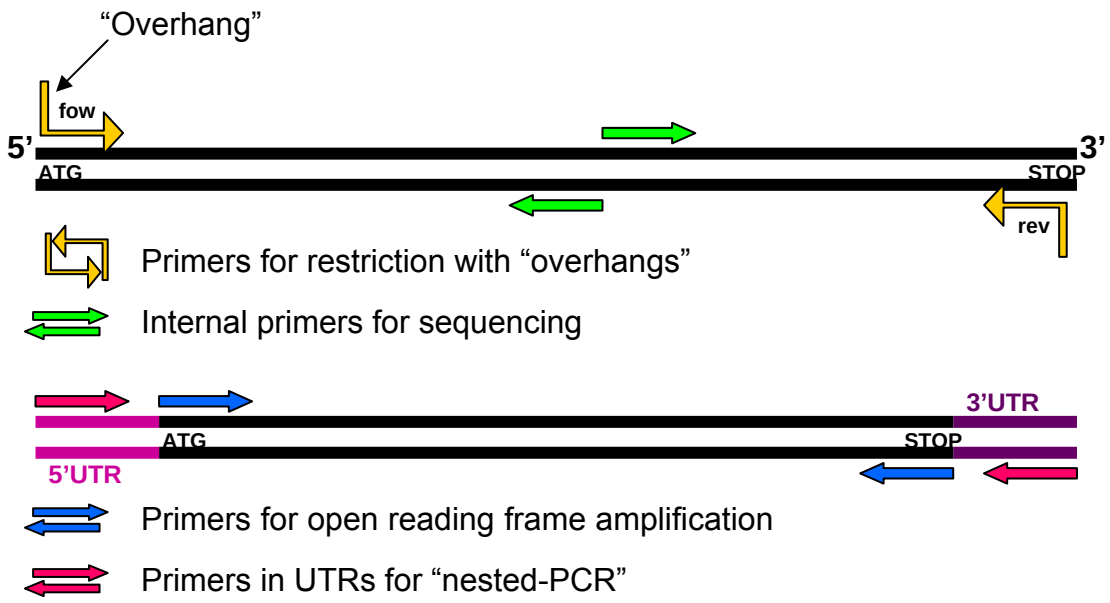


Figure 2.1 Schematic representation of primer design. Black lines represent coding DNA strands with 5’- and 3’-end indicated at the ends. Violet lines represent untranslated regions (UTR) in the 5’ and 3’ ends of the coding DNA. Arrows represent primers.

PCR reaction

A standard PCR reaction contained:

150ng template DNA
\+ 10pmol forward primer
\+ 10pmol reverse primer
\+ 4µl dNTP-mix (10mM each)
\+ 5µl polymerase buffer
\+ 1µl DNA polymerase
\+ H_2O to a total of 50µl

Cycling reactions consisted of a first denaturing step, a second annealing step, a third polymerization step with around 1min per 1000bps to be amplified, and an additional polymerization step of 10min at the end of all cycles. Finished reactions were held at 4°C. Temperatures and times varied depending on the gene and on the purpose of the PCR. For gene product amplification, a standard cycling would be:

5 min at 94ºC
30 sec at 94ºC
20 sec at 60ºC
1 min at 72ºC
X 20
10min at 72ºC

Reverse transcription

Reverse transcription of RNA into single stranded, complementary DNA (cDNA) was used to obtain gene coding sequences for cloning. In the reverse transcription oligo dT primers where used to amplify only the mRNA in order to enrich for coding sequences. For a total reaction volume of 20µl, 1µg of total RNA in 8µl was incubated together with 2µl of 5x 1st strand buffer and 1µl DNAse I for 20min at RT to allow digestion of DNA. 3µl of mastermix I, which contains 25mM EDTA, dNTP mix (10mM each) and 300ng/µl oligo dT primer, were added to the reaction and incubated in a Thermo cycler for 10min at 65ºC to melt secondary structures, followed by 10min at 25ºC to anneal oligo dT primers to the RNA. Afterwards, 5µl of mastermix II were added containing 5x 1st buffer, DTT 0.1M, and nuclease free H_2O at a ratio of 2:2:1. The reaction was then cooled to 42ºC for 2min, and 1µl of reverse transcriptase was added together with 0.2µl of T4 gene 32 protein. The mix was incubated at 42ºC for 50min and subsequently the enzyme inactivated at 95ºC for 10min. The reverse transcription reaction was stored at -20ºC until usage or at -80ºC for long-term storage.

Colony PCR

Colony PCR is a PCR reaction in which bacteria from a single colony are directly used as template to screen for positive colonies containing plasmids with inserts. With proper selection of primers, the orientation of inserts in the vector as well as the insert size can be determined.

PCR reaction mixture was 2.5µl 10x Taq-polymerase buffer, 1µl dNTP mix (10mM each), 0.5µl forward primer, 0.5µl reverse primer (each 10pmol), 0.25µl Taq-polymerase and 20.25µl H_2O per reaction. To 25µl PCR mix bacteria from a single colony was added with a white pipette tip. Thermo cycling conditions were 96ºC for 10min; 30 cycles of 94ºC for 30sec, 55ºC for 20sec and 72ºC for 1min; hold at 4ºC. The polymerization time was increased according to the expected product size.

Sequencing Reaction

Sequencing reactions were performed according to Sanger dideoxy sequencing chain termination method (Sanger et al. 1977). Sanger sequencing is based on the use of normal nucleotides (dNTP's) in addition to dideoxynucleotides (ddNTP's), the later terminating the DNA polymerization when integrated into a sequence by

preventing binding of further nucleotides. Sequencing was performed with the PRISM Big DyeDeoxy Terminator Kit. For each sequencing reaction (10.3µl), 3µl of Big Dye Terminator 3.1 mix (containing polymerase, buffer, dNTPs and fluorescently labeled ddNTP), 0.3µl of specific primer (10pmol) and 150 to 400ng template DNA were filled up with 7µl H_2O. The PCR reaction conditions for sequencing were 95ºC for 5min; 35 cycles of 95ºC for 7sec, 50ºC for 10sec and 60ºC for 4min and hold at 15ºC.
Sequencing reaction products were precipitated with 25µl of 100% Ethanol plus 2µl of 3M NaAc per reaction, mixed by briefly vortexing and centrifuged at 4000rpm for 30min at 18ºC. The precipitated products were washed once in 200µl 70% Ethanol, and centrifuged at 4000rpm for 30min at 18ºC. Precipitated products were dried at 70ºC for 5min, resuspended in 10µl formamide per reaction and used for sequencing in an automatic capillary sequencing machine.

Sequence Analysis

Sequences were analyzed using the Staden package software[6] (Staden 1996) and the gene's sequences given in Table 3.1 obtained from NCBI[7].

2.2.3. Cloning

Sequences of interest were cloned into appropriate vectors among others for restriction cloning and transformation in human cells.

2.2.3.1. TOPO TA cloning

TOPO TA cloning is a patented technology from Invitrogen (Invitrogen, Karlsruhe) for inserting PCR products directly into a TOPO vector. These vectors have topoisomerases covalently linked at their ends that twist the insert-DNA into the vector within minutes (for more information see www.invitrogen.com). TOPO TA cloning was used to directly clone the gene-specific PCR amplified products into vectors for restriction digestion, for sequencing and for expression. Typically 6µl of cloning reactions consisted of 2µl of purified PCR product, 1µl of 1.2 M NaCl; 0.06 M $MgCl_2$, 0.5µl of TOPO vector and 2.5µl H_2O. Cloning reactions were incubated at RT for 20min. Afterwards 3µl of the completed cloning reactions were used to transform one vial of TOP 10 Chemically competent cells (Invitrogen). For sequencing and

[6] http://www.mrc-lmb.cam.ac.uk/pubseq/

[7] http://www.ncbi.nlm.nih.gov/

restriction digestion inserts were cloned into TOPO TA pcDNA3.1 XL vector, while for protein expression inserts were cloned into TOPO expression vectors pCRT7-CT and pTrcHis-NT, as described above.

2.2.3.2. Restriction cloning

In restriction cloning, restriction enzymes with specific DNA recognition sequences are used to digest both the vector and the insert-DNA. Many restriction enzymes recognize palindromic sequences and produce staggered incisions, resulting in "sticky" ends that bind to complementary sequence in the opposite strand[8]. If different enzymes are used in 5' and 3' end of the amplified product, the vector is less prone to religation (as the ends are not complementary), and the DNA-fragment will be inserted directionally.

Digestion of expression vectors and PCR products with restriction enzymes

Vectors were usually digested in 50µl reactions (for a detailed list of vectors see section 2.1.2). Reactions contained 5µg vector, 5µl 10x enzyme specific buffer, 0.5µl 100x BSA (if needed), 1µl enzyme and H_2O to the total volume of 50µl. Incubation times varied from 2hrs to O/N at 37ºC. PCR products were usually digested in 20µl reactions containing 17µl of PCR purified product, 2µl 10x enzyme specific buffer, 0.2µl 100x BSA (if needed) and 1µl enzyme. Incubation times were usually O/N at 37ºC. For double digestions, if the enzymes had compatible buffers they were simultaneously added to the reaction, or else the DNA was purified in between digests. After digestions were completed, enzymes were heat inactivated 15min at 65ºC.

Dephosphorylation and ligation

Dephosphorylation of vector molecules is performed to minimize recircularization of linearized vector with compatible ends. Shrimp alkaline phosphatase (SAP) was used for dephosphorylation during 1hr at 37ºC, followed by a DNA purification step (see Purification of nucleic acids and gel extraction). During DNA ligation *in vitro* DNA ligase catalyzes the formation of a phosphodiester bond between one nucleotide containing a 5' phosphate group and another nucleotide containing a 3' hydroxyl group. Ligation was performed in 5µl reactions with 150ng vector DNA and insert DNA at a molecular ratio of 3:1 between insert and vector, 0.5µl of 10x ligation buffer, 0.5µl of dATP (10mM), 1µl T4 DNA-ligase, and H_2O to a total volume of 5µl. The

[8] http://en.wikipedia.org/wiki/Restriction_enzyme

reaction was incubated 1hr at RT. For transformation, 2.5µl of the ligation reaction was transformed into 25µl of chemically competent bacteria. A summary of vectors with inserts produced by restriction cloning in this work is listed in table 2.1.

Expression organism	Vector name	Gene cloned	Primers used (forward-reverse)
Bacteria	pET43.1a	C13orf1	C6M1_fow-m / C6F196_rev
		RCBTB1	C7-fow / C7-rev
		SETDB2	C8-fow / C8-ATV-Hind-Rev
		RFP2	RF-mam_fow / Rf-Rev
		KPNA3	Kp_fow / Kp_rev
	pPEP	C13orf1	C6-NcoI_fow / C6F196_rev
		RCBTB1	C7-fow / C7-rev
		SETDB2	C8-fow / C8-XmaI-stop-rev
		RFP2	Rf_fow / Rf_Rev
	pDS5	C13orf1	C6-NcoI_fow / c6-HindIII-rev
		RCBTB1	C7-fow / C7-rev
		SETDB2	C8-fow / C8-XmaI-stop-rev
		RFP2	Rf_fow / Rf_Rev
	$NHis_{10}$-zzTev-80N	C13orf1	C6-NcoI_fow / c6-BamHI-rev
		RCBTB1	C7-Acc65I-fow / C7-BamHI-rev
		SETDB2	C8-NcoI_fow / c8-BamHI-rev
		RFP2	Rf_fow / Rf-BamHI-rev
		KPNA3	Kp-NcoI_fow / Kp-HindIII-rev
	$NHis_{10}$-80N	C13orf1	C6-NcoI_fow / c6-HindIII-rev
		RCBTB1	C7-Acc65I-fow / C7-XmaI-stop-rev
		SETDB2	C8-NcoI_fow / C8-XmaI-stop-rev
		RFP2	Rf_fow / Rf-Xma-STOP-rev
		KPNA3	Kp-NcoI_fow / Kp-XmaI-rev
Mammalian cell lines	pEGFP-C1 and pEYFP-C1	C13orf1	C6M1_fow-m / C6F196_rev
		RCBTB1	C7-mam_fowI / C7-mam_rev-I
		SETDB2	C8-mam_fow / C8-XmaI-stop-rev
		RFP2	Rf_fow-Nhe-Bsp / Rf-Xma-STOP-rev
		KPNA3	Kp-mam_fow-I / Kp-XmaI-rev
	pBCHGS	C13orf1	C6-ATV-Bam-Fow / c6-HindIII-rev
		RCBTB1	C7-fow / C7-XmaI-stop-rev
		SETDB2	C8-ATV-EcoV-Fow / C8-ATV-Hind-Rev
		RFP2	RF-ATV-Bam-Fow / Rf-ATV-Hind-Rev
	pcDNA3-PA-1b	C13orf1	Inserts restricted out of pBTM117c vector with SalI-NotI to be cloned into the pcDNA3-PA-1b vector
		RCBTB1	
		SETDB2	
		RFP2	
		KPNA3	
	pTL1-HA1	C13orf1	Inserts restricted out of pBTM117c vector with SalI-NotI to be cloned into the pcDNA3-PA-1b vector
		RCBTB1	
		SETDB2	
		RFP2	
		KPNA3	
Yeast	pBTM117c	C13orf1	C6-SalI-fow-1 / C6-rev-178-Stop
		RCBTB1	C7-XhoI-fow-1 / C7-NotI-rev-531-Stop
		SETDB2	C8-SalI-fow-1 / C8-NotI-rev-719
		RFP2	RF-SalI-fow-1 / RF-NotI-rev-407
		KPNA3	Kp-SalI-fow-1 / KP-NotI-rev-521
	pACT4_1b	C13orf1	C6-SalI-fow-1 / C6-rev-178-Stop
		RCBTB1	C7-XhoI-fow-1 / C7-NotI-rev-531-Stop
		SETDB2	C8-SalI-fow-1 / C8-NotI-rev-719
		RFP2	RF-SalI-fow-1 / RF-NotI-rev-407
		KPNA3	Kp-SalI-fow-1 / KP-NotI-rev-521

Table 2.1 Summary of restriction cloning of 13q14 candidate genes. Restriction cloning was performed for cloning into vectors for expression in bacteria, mammalian cell lines and yeast (for yeast-two-hybrid). Indicated are the primers used for amplification and restriction (for more details see Appendix 1).

2.2.4. Transformation

Transformation is the process of genetically altering a host cell by the introduction of foreign genetic material. If the material is introduced into an animal cell, the process is generally termed transfection. Transformation of *E.coli* (bacteria) and *S.cerevisiae* (yeast), and transfection of human cell lines with plasmid DNA was performed to express the candidate 13q14 proteins and to functionally characterize them.

2.2.4.1. Chemical transformation

Chemical Transformation of *E.coli*

For chemical transformation, two different chemically competent strains of *E.coli* were used (TOP 10 cells from Invitrogen and XL1Blue from Stratagene). Chemically competent *E.coli* were mixed with plasmid DNA, incubated 20min on ice, heat shocked 30sec at 42ºC and promptly placed on ice for 1min. 250µl of SOC medium was added to each vial and incubated at 37ºC with vigorous shacking for at least 1hr to allow bacterial growth. 50-150µl of grown bacteria were spread accordingly to the vector's resistance onto selective agar plates containing Ampicillin or Kanamycin and grown O/N at 37ºC. The next day, resistant colonies were picked and tested for plasmid containing target gene-insert by colony PCR (see chapter 2.2.2).

Chemical Transfection of eukaryotic cells

Chemical transfection of eukaryotic cell lines was performed according to different protocols supplied with commercially available kits (Table 2.2). The cell lines used are listed in Appendix 2.

Company	Kit	Application
Invitrogen	Lipofectamine	PD, WB
Stratagen	Effectene	IF, PD, WB
PAA	Nanofectin	IF, WB, PD

Table 2.2 Transfection Kits for animal cell lines. The applications column indicates the preferred method for which the kits were used. IF=immunofluorescence, PD=pull-down, WB=Western blot analysis.

Chemical Transformation of Yeast

Chemical transformation with lithium acetate was used to transform yeast with bait and prey plasmids for Y2H screen. For each 96 well plate, a starter culture of yeast strains was prepared by inoculation of 3ml YPD medium and growth O/N at 30ºC.

This starter culture was mixed with 28ml YPD medium and grown 4hrs at 30ºC up to an OD_{600} of ca. 0.6. Yeast cells were centrifuged at 2000rpm during 5min at RT and cell pellets resuspended in 10ml TE buffer for washing. Cells were centrifuged again and pellets mixed with 1.1ml MIX1 and incubated 10min at RT. In each well of the plate, 3µg of a single plasmid was added together with 3µl of carrier DNA (herring testes carrier DNA, Clontech) to which 11µl of the cell pellet in MIX1 were also added. Then, 58µl of MIX2 were added to each well and incubated at 30ºC during 30min. After incubation, 8µl/well DMSO were added and the cells were heat-shocked 7min at 42ºC. Directly after heat-shock the cells were stamped onto SDII selective plates and let grow at 30ºC for 3 to 4 days. The yeast strains L40cca [MATa] and L40ccα [MATα] were used for baits and preys transformation respectively.

2.2.4.2. Electroporation

Bacteria

Electroporation was used for transformation of bacteria for protein expression. Cells were given one pulse of no more than 1mA for 2 sec. and immediately placed on ice. After no more than 1min, 1ml SOC medium was given and cells were grown at 37ºC with vigorous shacking for at least 1hr. 0.5-1ml from the grown bacteria were spread onto selective agar plates and grown O/N at 37ºC in an incubator. The next day, plates were washed with selective LB-medium and diluted in appropriate volume of growth media (for more details see section 2.2.6). For protein expression the strains Rosetta(DE3)pLysS and BL21(DE3)pLysS (both from Novagen) were used for electroporation.

2.2.5. In vitro protein expression

To test whether the 13q14 candidate proteins could be expressed by the *E.coli* transcription-translation machinery, an *in vitro* system was tested. Some proteins yield low amounts when expressed in bacteria because ribosomes can fall off the mRNA after few amino acids have been translated depending on the sequence in the first 25bp (B.Korn personal communication (DKFZ)). To test whether changes in the first 20bp of the published sequences for 13q14 candidate proteins improved protein expression in bacteria, conservative single nucleotide changes were introduced in the sequences. Primers containing conservative nucleotide exchanges (Appendix 3) were designed for all 13q14 candidate genes and used as templates for PCR

amplification. The amplified products were then used as templates for the *in vitro* translation. The *in vitro* system used was the Rapid Translation System RTS 100, *E.coli* HY Kit developed by Roche laboratories based on the method described by Zubay (Zubay 1973), Spirin et al. (Spirin et al. 1988) and others. Briefly, a DNA template is transcribed and translated in the same reaction. The reaction contains T7 RNA polymerase for transcription and the *E.coli* ribosomal machinery for translation. For transcription and translation linear templates like PCR products can be used as long as they contain specific regulatory elements like the T7 promoter and the ribosomal binding site. These regulatory elements were added to the target genes in a PCR reaction together with a His-tag for detection (Figure 2.2).

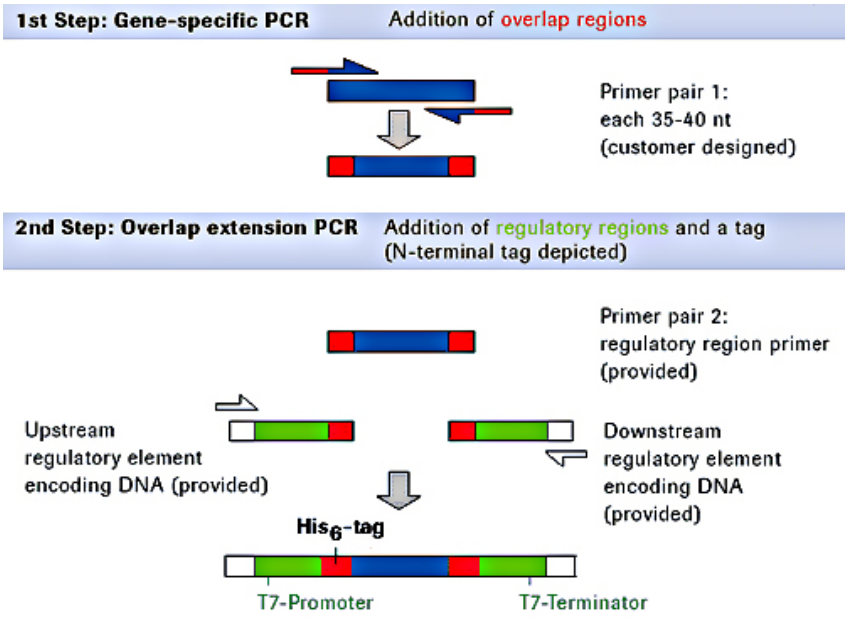


Figure 2.2 Schematic view of the steps for generation of the linear template for *in vitro* protein expression with the RTS system (Roche).

Amplification of the genes was performed with High Fidelity (HiFi) polymerase (Roche) and each 50µl reaction consisted of:

5µl 10x HiFi Buffer with $MgCl_2$
2µl dNTPs
1µl 10pmol/µl forward primer
1µl 10pmol/µl reverse primer
2µl 25ng/µl template DNA (clone with gene insert)
0,8µl HiFi-Polymerase
38,2µl H_2O

Addition of the required regulatory elements and the tag for detection was performed in a 50µl PCR-reaction for linear template generation consisting of:

5µl	Buffer 2
1,25µl	dNTPs(10mM)
2µl	T7 promotor
2µl	T7 terminator
1µl	C-terminal His
0,8µl	HiFi-polymerase
1µl	DNA template from first PCR
36,95µl	H_2O

After generation of the linear template, in vitro protein expression was performed with the Rapid Translation System RTS 100, *E.coli* HY Kit (Roche). Each 50µl reaction consisted of:

12 µl	*E.coli* lysates
10 µl	Reaction mix
12 µl	Amino acids
1 µl	Methionine
5 µl	Reconstitution buffer
10 µl	Linear template DNA

Reactions were incubated at 30ºC during 4hrs and then either directly analyzed or stored at -20ºC. For analysis, 5µl from each reaction were taken as the samples to load SDS-PAGE gels for Western blot analysis with anti-His antibody (for more details see sections 2.2.13 and 2.2.16.2 respectively).

2.2.6. Expression of recombinant protein in bacteria

E.coli bacteria were used to recombinantly over-express 13q14 candidate proteins C13orf1, RCBTB1, SETDB2 and RFP2 fused to a Tag that was used for purification and detection. Cells were electroporated (section 2.2.4.2) and grown in 1l LB-selective medium at 37ºC with vigorous shaking. Cell density was measured in a spectrophotometer at 600nm and when cells reached the logarithmic growth rate (OD_{600} = [0.6-0.8]) they were induced with 0.1mM IPTG. After induction, cells were grown for further 3-6hrs. Before harvesting, 5mM PMSF was added to the cultures and cells were collected by centrifugation at 4000rpm for 30min at 4ºC. Cell density was measured and cells were resuspended in 50 ml ice-cold lysis buffer pH 7.5 to an OD_{600} of 3.5 (lysis buffer volumes were calculated accordingly for each culture depending on the OD_{600}). Solubilized pellets were either used directly for preparing cell lysates or frozen in liquid Nitrogen and stored at -80°C until further usage.

Cell lysates from bacteria

Solubilized pellets were thawed at 4ºC and 1mM β-ME was added. Cells were lysated by sonication at 50% intensity, and 5 cycles of 2 to 5 min at 4ºC were applied. Proteins were collected by centrifugation at 45000rpm for 2hrs at 4°C. Pellets were either used directly for further processing or frozen at -80°C until purification.

2.2.7. Purification of recombinant proteins from bacteria

Nickel-NTA purification

Recombinant proteins containing a histidine tag can be purified via metal-affinity chromatography. Histidines are known to bind to Nickel ions (Ni) with high affinity, and this property was used for the purification of the recombinant proteins with Ni-NTA Agarose (Qiagen). The proteins C13orf1, RCBTB1, SETDB2 and the fragment I of RFP2 (S154-E264) had been expressed as recombinant proteins (section 2.2.6) containing a 10x His tag at the N-terminus. They were purified with Ni-NTA beads under denaturing conditions following the manufacturer instructions (The QIAexpressionist handbook, fifth edition, Qiagen).

2.2.8. Renaturation of purified recombinant proteins expressed in bacteria

Purification yielded denatured proteins in 8M urea buffer. Denatured proteins have lost their tertiary structure and are therefore enzymatically inactive. In order to optimize the renaturation conditions, the purified proteins were dialyzed using buffers of different compositions and pH: two phosphate-buffers and one tris-buffer were used with varying salt concentrations from 20mM NaCl to 500mM NaCl giving a total of 12 different buffers specified in section 2.1 and 3.1.2.1 (Table 3.10). 20µl purified protein in 8M Urea buffer were dialysed against 500ml of the corresponding end-buffer (Servapor dialysis tubing, Serva Electrophoresis) O/N at 4ºC. After dialysis, samples were collected and centrifuged at 45000rpm for 30min at 4ºC to remove precipitated protein. After centrifugation the supernatants were collected and prepared for SDS-PAGE (section 2.2.12).

2.2.9. Cell culture

Different cell lines were cultured at 37°C, 95% relative air humidity and 5% CO_2, as indicated by DSMZ (German resource center for biological material). For details of cell lines used, see appendix 2. Cells were split before reaching confluence. Adherent cells were washed once in 1x PBS and detached by 2-5min incubation at 37ºC in either 0.25% Trypsin diluted in 1x PBS, or 0.02%EDTA in 1x PBS. Once detached, 5ml of fresh, pre-warmed medium was added and cells were collected in a Falcon tube, followed by centrifugation at 1000rpm for 3-5min. Cell pellets were resuspended in fresh, pre-warmed media and diluted 1:5 to 1:10 in new medium-containing flasks.

2.2.10. Recombinant protein expression in cell lines

Recombinant proteins were transiently expressed in cell lines (for cell lines used see Appendix 2) for functional analyses using appropriate vector systems (see section 2.1 and Table 2.1 for vectors). After transfection, cells were cultured for 24 to 48hrs and then either fixed or harvested and lysed.

Cell fixation

Cultured cells were fixed for immunofluorescence and microscopy by washing once in 1x PBS with 1mM $MgCl_2$. Fixation was performed with the Methanol-Acetone method by embedding the cells in 100% Methanol at -20°C during 10min. Cover slips were dried from methanol and immediately submerged in 100% Acetone at -20°C during 5min. After acetone treatment the cover slips were let to dry and kept at RT until usage. Suspension cells were grown in culture media and centrifuged at 800rpm-5min at RT. Cells were resuspended in fresh media at $1x10^6$cells/ml. 100µl were spread onto a slide in a Cytospin (Cytospin 3, Shandon Life Sciences) and fixed with the Methanol-Acetone method as described above.

Cell lysates from cell lines

Cell lysates were prepared from cultured cells using one of either three lysis buffers: RIPA, Hepes or Tris buffer. All buffers were freshly supplemented with “Complete EDTA-free protease inhibitor cocktail” tablets (Roche), and 1µl/ml Benzonase (Merk) prior to lysis. All buffers were at 4ºC and 150µl buffer per $5x10^5$-10^6 cells were used for cell lysis. Cells were lysed in RIPA buffer for 30min on ice and lysates were used

for Western blot analysis detection of proteins, while lysis in Hepes and Tris buffers was performed for 1hr at 4°C in a mixing-rotor and lysates were also used for pull-down experiments and immunoprecipitation. After lysis, all samples were centrifuged at 13000rpm during 10min at 4°C. The supernatants were then either mixed with 3x loading buffer (Laemmlie) for acrylamid gel analysis, or further used for pull-down experiments or immunoprecipitation.

2.2.11. Measurement of protein concentration

BCA-Test

The BCA-Test is a Biuret reaction where peptide bonds react with copper ions in an alkaline solution. The reacted ions in turn react with bicinchonic acid (BCA) to give a violet colour which varies in intensity depending on the amount of protein present. Because peptide bonds occur with approximately the same frequency per gram of material in every protein, it can be used for estimation of protein concentration (Rehm 2002). Standard curves were made with BSA (bovine serum albumin) ranging from 2000µg/ml to 25µg/ml. Biuret reaction was performed with Copper-sulphate and Bicinchonic acid from Invitrogen (100µl Copper-sulphate in 5ml Bicinchonic acid). 10µl of protein solution was added to 200µl of reaction mixture, incubated at 37°C for 30min and then read in an ELISA reader at 550nm.

Estimation by UV spectroscopy at 280nm

Aromatic rings in amino acids have a maximum of absorbance at 280nm wave length that can be used to estimate protein concentrations. Absorbance measurements were performed by measuring 1µl of protein solution at 280nm with an automated spectrophotometer (NanoDrop).

2.2.12. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a method to separate proteins according to their size (Schagger& von Jagow 1987). Proteins are coated with the detergent sodium-dodecyl-sulfate (SDS) that confers a linear shape and a negative charge to the proteins. Because of similar shape of the denatured proteins and the fact that the net charge depends only on the size of the proteins, the polypeptides can be separated according to their molecular weight. In PAGE, the matrix used to separate the proteins is a polyacrylamide gel, a polymer that gives a highly reproducible pore size upon polymerization. In

subsequent electrophoresis, the charged molecules migrate in an electric field. A discontinuous system was used in this work, with a non-restrictive large pore gel stacked on top of the resolving gel. Samples were prepared by mixing with Laemmli solution, boiling at 96ºC for 5 to 10min and cooling them before loading together with See Blue or SeeBlue Plus2 pre-stained protein size standards (Invitrogen). Samples from pull-down experiments were always prepared without reducing agent while all other samples contained 5% β-ME.

Components	Stacking gel	Separation gel		
	5% Acrylamide	**10%** Acrylamide	**12%** Acrylamide	**15%** Acrylamide
1M Tris-HCL pH8,8	-	1,875 ml	1,875 ml	1,875 ml
1M Tris-HCL pH6,8	0,375 ml	-	-	-
PAA-30%	0,5 ml	1,6 ml	2 ml	2,5 ml
10% SDS	30 µl	50 µl	50 µl	50 µl
40% APS	15 µl	24 µl	24 µl	24 µl
TEMED	6 µl	10µl	10µl	10µl
H_2O	2,1 ml	1,4 ml	1 ml	0,6 ml
Total volume	3 ml	5ml	5ml	5ml

Table 2.3 Composition of polyacrylamide gels for SDS-PAGE. Indicated are the volumes of the different components used for stacking and separation gels.

The gels for SDS-PAGE were prepared as described in Table 2.3 and run in SDS-PAGE run buffer at 100V for 1 to 2 hrs depending on the size of the proteins and the acrylamide percentage of the gel (C13orf1 12% or 15%, RCBTB1, RFP2, KPNA3, SETDB2 10% or 12%). For casting and running SDS-PAGE gels the Mini-Protean 3 gel system (BIORAD) was used.

2.2.13. Protein staining in SDS-PAGE gels

After separation with SDS-PAGE, the gels were either further processed for Western blot analysis or stained with Coomassie brilliant blue R 250 (Merck). Gels were soaked in staining solution (40% (v/v) Methanol + 10% (v/v) glacial Acetic Acid + 50% (v/v) H2O + 0.05% (w/v) Coomassie brilliant blue R 250). Unbound Coomassie was removed by immersion in 250ml ddH_2O/gel and boiling in a microwave oven until staining of the gel was optimal for imaging.

2.2.14. Production of Antibodies

Polypeptides for antibody production were produced in bacteria as recombinant proteins containing a 10x His tag, and subsequently purified via metal ion chromatography (see sections 2.2.6 and 2.2.7). Because the His tag has a low antigenicity, it was not removed and animals were immunized with recombinant fusion proteins. Polyclonal antibodies were produced against the full length proteins C13orf1, RCBTB1 and SETDB2, and the fragment I (S154-E264) of RFP2. Monoclonal antibodies were produced against the full length proteins C13orf1, RCBTB1 and SETDB2, the fragment I (S154-E264) of RFP2, and two peptides of KPNA3 (Appendix 4). Polypeptides were produced as described (section 2.2.6), and peptides from KPNA3 were ordered at Peptide Specialty Laboratories (PSL, Heidelberg).

2.2.14.1. Polyclonal antibodies

Immunization of Guinea Pigs

Guinea pigs were immunized for the first time at 8 weeks of age, with 100µg of protein per animal diluted 1:1 with Complete Freud's adjuvant (Sigma). Subsequently animals were immunized periodically with 100µg of protein per animal diluted 1:1 with Incomplete Freud's adjuvant (Sigma).

Immunization schedule:

Immunizations	1st	Monthly	Every 6 months
Time	0	During 6 months	During 18 months
Antigen	100µg protein + complete Freud's adjuvant (1:1)	100µg protein + incomplete Freud's adjuvant (1:1)	100µg protein + incomplete Freud's adjuvant (1:1)

During this process, if necessary, serum was taken one week after immunization by heart-punction. After 24 months animals were bled. Blood was collected in Vacutainer Blood collection tubes (BD), left at RT during 1hr and blood cells sedimented by spinning at 2000rpm 1hr at RT. After sedimentation, serum was collected and aliquoted for storage at -80ºC. For immediate usage, serum was complemented with 0.02% Sodium Azide as preservative and kept at 4ºC.

Testing of polyclonal antibodies

To test whether the antibodies were performing in Western, serum from guinea pigs was diluted in blocking solution (5% skimmed milk powder, 1% BSA in 1x TBS) and used as primary antibody for Western. To test whether the antibodies were performing in immunofluorescence, serum was diluted in 1% BSA in PBS, and applied to the slides as primary antibody. Dilutions were performed as indicated in Table 3.14.

2.2.14.2. Monoclonal antibodies

Immunization of Mice

Mice were immunized with 50µg of protein per animal diluted 1:1 with Complete Freud's adjuvant (Sigma). Subsequent immunizations were performed with 50µg of protein in dilution buffer (4M Urea in 1xPBS).

Immunization schedule:

Immunization	1st	2nd	3rd	Fusion
Time	Day 0	Day 4	Day 6	Day 7
Antigen	50µg protein+complete Freud's adjuvant (1:1)	50µg protein	50µg protein	

After 7 days, dead animals were sunk in 70% ethanol for disinfection and dried. Under sterile conditions the lymph nodes in the popliteal space were extirpated and put in 10ml 37ºC warm pure RPMI medium. Lymph nodes were disrupted with a syringe by friction. For each fusion, the SP 2/0 cells from a big, confluent flask were taken and both cells (lymph-extracted and myeloma cell line) were centrifuged separately 10min at RT and 150xg. Both cell pellets were washed twice in 10ml pure RPMI medium. After second wash, cells were re-suspended in pure RPMI and mixed by pipetting up and down. Then, the mixture of cells was centrifuged 10min at RT and 150xg and supernatant was completely removed. For fusion of lymph cells and myeloma cells, 1,5ml polyethylene glycol (PEG) was added to the pellet in 1min by dropping it over the cells under continuous shaking. Cells were shaken 1min alone, and then 20ml pure RPMI medium was added. Cell suspension was centrifuged 10min at RT and 150xg, and incubated at RT for 5min. After incubation, the supernatant was completely removed and cells were re-suspended in 10ml of plating

medium (RPMI complete with 10%FCS, 1xHAT supplement (Hypoxanthine, Aminopterin and Thymidine), and 24µM β-Mercaptoethanol). 5ml were centrifuged as above, cells re-suspended in 1,5ml freezing medium (RPMI: FCS: DMSO, 7:2:1) and frozen as stock for hybridoma clones. The remaining 5ml were diluted in 100ml plating medium and plated in 5x 96well plates with 200µl hybridoma cell suspension/well. Plated cells were grown at 37°C and 5%CO_2 until hybridoma clones appeared 4 to 7 days after fusion.

Testing of monoclonal antibodies

The culture media of growing clones were used as monoclonal 1^{ry} antibodies to perform different tests for specificity. In ELISA and Western blot analysis the recombinant proteins that had been used for immunization were used as sample and 1:200 to 1:400 dilutions of blood of the immunized animals was used as positive control; 1xPBS and fresh medium were the negative controls. In immunofluorescence, cells recombinantly over-expressing the target proteins were used as sample, tag-antibodies were used as positive control, and 1xPBS and fresh medium as negative controls. Monoclonal antibodies producing specific signals were selected for purification.

Purification of monoclonal antibodies

Hybridoma clones were grown in 1l Hybridoma Express medium (PAA) 1 to 2 weeks and were centrifuged at 150xg during 10min at RT. After centrifugation, supernatants were collected and frozen at -20ºC until purification. Purification was performed with the Proteus G Midi kit (Prochem) following the manufacturer instructions. After purification, the concentration of the purified antibodies was measured (section 2.2.11) and they were tested again for ELISA, Western blot analysis and immunofluorescence to reassure their specificity. Antibodies were aliquoted for storage at -80ºC.

2.2.15. Immunological methods

2.2.15.1. ELISA

The Enzyme-Linked Immunosorbent Assay (ELISA) is a technique used to detect specific antibodies or substances with antigenic properties in a sample. There are different types of ELISA assays, but in this work the indirect ELISA was used. Briefly,

an antigen is used to coat the wells of a 96-well plate. A specific 1ry antibody that is tested is then added to the coated wells. For detection of specifically bound primary antibody, a 2ry antibody is added that is labeled with an enzyme that will transform a colorless substrate into a colored compound that can be detected in an automatic manner with a spectrophotometer at the appropriate wave length (Janeway et al. 2004). In this work the antigens used to immunize the animals (purified proteins) were used to coat the wells of 96-well plates. Assessed for specificity were sera from immunized animals, culture media from hybridomas or purified antibodies. The 2ry antibody was labeled with Horseradish Peroxidase (HRP).

2.2.15.2. Western blot analysis

The Western blot analysis is a method to detect proteins in a sample with specific antibodies, after being separated by SDS-PAGE (section 2.2.11). In Western blotting (Burnette 1981, Renart et al. 1979), the separated proteins will be transferred to a membrane where the proteins will be better accessible for antibodies and can be further analyzed. In this work, it had been experimentally determined that detection of proteins that were in low concentration was more efficient if blotting was performed wet. Thus, semi-dry blotting was used when the amount of protein to be detected was high in the sample, i.e. when using purified proteins as samples; when lower amounts of protein were present wet blotting was used, i.e. to assess precipitated proteins from pull-down experiments and when using whole cell extracts as protein sample. After separating the samples by SDS-PAGE, semi-dry blots were assembled directly on the catode and consisted on 1 sheet of extra thick blot paper (Bio-Rad), a nitrocellulose membrane (Whatman), the gel with the separated samples and another sheet of blotting paper. All components were wet in blotting buffer 2 to 5min prior to assembly and blot was run at 15mA during 1 to 2 hrs. For more sensitive detection or small amounts of protein, Western blot analysis was performed wet. The assembled blot was embedded in the blotting buffer and run at 250mA for 2 to 4hrs on ice. To verify transfer of proteins, membranes were stained in Ponceau S (Sigma) for 5min and de-stained by washing 5min in washing buffer. Membranes were blocked in blocking solution for 1hr or O/N at 4ºC. After blocking, the 1ry antibody diluted in 1% BSA solution was applied (at the concentrations described in Table 2.4) and incubated 1hr or O/N at 4ºC. After incubation of the 1ry antibody, membranes were washed 2 times for 5min and once for 10min at RT in TBS-T. 2ry antibody diluted in

blocking solution was applied (at the concentrations described in Table 2.4) and incubated 1hr at 4ºC. Membranes were washed as above and the blot was developed with ECL Western blot analysis detection kit (Amersham) as described by the manufacturer.

Antibody	Concentration	Antibody	Concentration
Ms-α-C13orf1	0,2µg/ml	Ms-α-GFP	0,4µg/ml
Ms-α-RCBTB1	0,5µg/ml	Ms-α-His	1:1000
Ms-α-RFP2	0,5µg/ml	Ms-α-Xpress	1:5000
Ms-α-KPNA3	0,2µg/ml	Ms-α-V5	1:5000
Gp-α-C13orf1	1:1000	Rb-α-PA	1:1000
Gp-α-RCBTB1	1:1000	Ms-α-HA	1:200
Gp-α-SETDB2	1:1000	Dn-α-Rb IgG-HRP	1:5000
Gp-α-RFP2	1:1000	Sh-α-Ms IgG-HRP	1:5000
Ms-α-Nus	1:10000	Sh-α-Gp IgG-HRP	1:5000

Table 2.4 Antibody concentrations used for Western blot analysis. The concentrations used for primary and secondary antibodies are indicated either in µg/ml or in dilution factor for stock solutions. Ms=mouse, Gp=guinea pig, Rb=rabbit, Dn=donkey, Sh=sheep, HRP=Horseradish peroxidase conjugated antibody.

2.2.15.3. Immunofluorescence

Immunofluorescence was used to detect target proteins of 13q14 and positive controls (GFP and poli-Histidine tags). In this work indirect immunofluorescence was used, and the antigen-specific 1^{ry} antibody was always detected by a 2^{ry} antibody that was fluorescently labeled. Fixed cells (section 2.2.10) were incubated with 100µl of primary antibody diluted in 1% BSA, 1x PBS at RT during 1hr in a humid chamber. Concentration of the primary antibody varied (Table 2.5). To detect the positive controls, α-GFP (mouse monoclonal anti-GFP, Roche) and α-His (mouse monoclonal anti-His6, Roche) antibodies were used as described by the manufacturer. Subsequently, slides were washed 3x in ~100ml of 1x PBS. After washing, slides were incubated with 100µl of Cy3- or FITC-labeled secondary antibody —goat-anti-mouse, or anti-guinea pig antibody (all from Dianova)—. Secondary antibodies were diluted 1:200 in 1% BSA, 1x PBS, 100µg/ml DAPI. 100µl of 2^{ry} antibody solution was added to each slide and incubated at RT during 1hr in a humid chamber. Slides were washed as above and mounted on microscope-slides with Fluoromount-G

(Southernbiotech), or VectaShield (Linaris). Slides were kept in the dark until analyzed under the microscope.

Mouse	Concentration	Guinea Pig	Dilution
α-C13orf1	0,5µg/ml	α-C13orf1	1:1000-1:2000
α-RCBTB1	0,5µg/ml	α-RCBTB1	1:1000-1:2000
α-SETDB2	2µg/ml	α-SETDB2	1:1000-1:5000
α-RFP2	0,5µg/ml	α-RFP2	1:1000-1:2000
α-KPNA3	2µg/ml		

Table 2.5 Antibody concentrations used for immunofluorescence. Concentrations of purified mouse monoclonal antibodies used are indicated in µg/ml. The concentrations used of polyclonal antibodies are indicated as the factor of dilution for the sera extracted from the animals.

2.2.16. Yeast-two-Hybrid (Y2H) screen

To identify binding partners for candidate proteins from 13q14, a Y2H screen was performed against a library of over 5000 clones containing human genes (see section 3.4. for details). Vectors containing target genes and corresponding reporter genes constitute the baits and prey constructs. The bait and prey constructs were transfected into mating-compatible yeast strains separately to create the corresponding library composed of the pool of yeast clones containing the candidate constructs. After baits and prey libraries were created, cells were mated and protein-protein interactions (PPIs) were screened for based on the reporter genes.

Automated Y2H screen

The genes C13orf1 (AF334405), RCBTB1 (AF334406), SETDB2 (NM_031915), RFP2 (NM_052811) and KPNA3 (NM_002267) from the chromosomal region 13q14 were selected as baits to mate with an existing prey library of human genes. Full length genes as well as fragments (Appendix 5) were cloned into bait vector pBTM117c and prey vector pACT4_1b (kindly provided by U.Stelzl (Max Planck Institute for Molecular Genetics, Berlin)). Yeast strains used for mating are described in Goehler et al. (Goehler et al. 2004). Baits were transformed into the L40cca [MATa] strain and the prey library was transformed into the L40ccα [MATα] strain (Stelzl et al. 2005). In order to test for interaction of 13q14 proteins with each other, 13q14 genes and fragments were also cloned in the prey vector and transformed in

the MATα strain. Baits and preys were arrayed in 96-well format as depicted in Figure 2.3.

Bait	1	2	3	4	5	6	7	8	9	10	11	12
A	C6-1	C7-4	C7-12	C8-3	C8-11	RF-4	RF-12	KP-7	BTM	C6-5	C7-1	RF-4
B	C6-2	C7-5	C7-13	C8-4	C8-12	RF-5	RF-13	KP-8	BTM	C7-17	C7-5	RF-6
C	C6-3	C7-6	C7-14	C8-5	C8-13	RF-6	KP-1	KP-9	BTM	C8-15	C7-9	RF-8
D	C6-4	C7-7	C7-15	C8-6	EMPTY	RF-7	KP-2	KP-10	BTM	RF-13	C7-16	RF-11
E	C6-5	C7-8	C7-16	C8-7	C8-15	RF-8	KP-3	KP-11	BTM	KP-11	C8-1	KP-1
F	C7-1	C7-9	C7-17	C8-8	RF-1	RF-9	KP-4	BTM	BTM	C6-1	C8-3	KP-4
G	C7-2	C7-10	C8-1	C8-9	RF-2	RF-10	KP-5	BTM	BTM	C6-3	C8-8	KP-8
H	C7-3	C7-11	C8-2	C8-10	RF-3	RF-11	KP-6	BTM	BTM	C6-4	C8-13	KP-9

Prey	1	2	3	4	5	6	7	8	9	10	11	12
A	C6-1	C7-4	C7-13	C8-4	C8-15	RF-10	KP-5	ACT	C6-5	C7-1	C8-6	RF-10
B	C6-2	C7-5	C7-14	C8-5	RF-1	RF-11	KP-6	ACT	C7-17	C7-3	C8-11	RF-11
C	C6-3	C7-6	C7-15	C8-7	RF-2	RF-12	KP-7	ACT	C8-15	C7-5	C8-12	KP-1
D	C6-4	C7-7	C7-16	C8-8	RF-3	RF-13	KP-8	ACT	RF-13	C7-7	C8-13	KP-4
E	C6-5	C7-8	C7-17	C8-9	RF-5	KP-1	KP-9	ACT	KP-11	C7-10	RF-2	KP-5
F	C7-1	C7-10	C8-1	C8-11	RF-6	KP-2	KP-10	ACT	C6-1	C7-16	RF-5	KP-6
G	C7-2	C7-11	C8-2	C8-12	RF-7	KP-3	KP-11	ACT	C6-3	C8-1	RF-6	KP-8
H	C7-3	C7-12	C8-3	C8-13	RF-9	KP-4	ACT	ACT	C6-4	C8-3	RF-7	KP-9

Figure 2.3 Representation of 13q14 Baits and Preys libraries in 96-well format. The bait library is represented by the 96-well plate containing the bait clones in the positions indicated. The prey library is represented by the 96-well plate containing the prey clones in the positions indicated. EMPTY=empty well, BTM=cells with the empty bait-vector pBTM117c, ACT=cells with the empty prey-vector pACT4_1b, C6=C13orf1, C7=RCBTB1, C8=SETDB2, RF=RFP2, KP=KPNA3. Numbers indicate the number of the fragment in the vector (as indicated in Appendix 5).

Screening for interactions was performed in an interaction matrix (Stelzl et al. 2005) with a 13q14 bait-plate as baits and the prey-library consisting of approx. 5000 human protein coding genes or the 13q14 prey-plate as preys. For the screening with the prey-library, the 13q14 baits were pooled for mating as depicted in Table 2.6. Positive pool interactions were selected by growing mated cells on SDIV selective medium plates (-Leu, -Trp, -Ura, -His) and back-mated for confirmation of single PPIs as described previously (Stelzl et al. 2005). The β-GAL assay was used to test the additional LacZ reporter gene (coding for the enzyme β-Galactosidase) and was performed as described by Goehler et al. (Goehler et al. 2004). In the assay, a chromogenic substrate was applied and only colonies where interaction occurred turned blue.

POOL	CLONES						
1	C6-F1	C6-F4	C7-F2	C7-F6	C7-F11	C7-F14	C7-F17
2	C6-F2	C8-F4	C8-F9	C8-F15	RF-F8	RF-F13	KP-F3
3	C6-F3	C7-F1	C7-F5	C7-F8	C7-F13	C7-F16	C8-F5
4	C6-F5	C7-F4	C8-F8	C7-F12	RF-F5	RF-F11	KP-F2
5	C7-F7	C7-F12	C7-F15	C8-F10	RF-F2	RF-F9	KP-F1
6	C6-F1	C7-F1	C7-F7	C7-F14	C8-F5	C8-F15	RF-F11
7	C6-F2	C7-F2	C7-F8	C7-F15	C8-F8	RF-F2	RF-F13
8	C6-F3	C7-F4	C7-F11	C7-F16	C8-F9	RF-F5	KP-F1
9	C6-F4	C7-F5	C7-F12	C7-F17	C8-F10	RF-F8	KP-F2
10	C6-F5	C7-F6	C7-F13	C8-F4	C8-F12	RF-F9	KP-F3
11	C6-F1	C7-F1	C7-F5	C8-F4	RF-F2	KP-F1	C8-F8
12	C6-F2	C7-F4	C7-F6	C8-F5	C8-F9	RF-F5	KP-F2
13	C6-F3	C7-F2	C7-F11	C7-F17	C8-F12	RF-F8	KP-F3
14	C6-F4	C7-F7	C8-F12	C7-F15	C8-F10	C8-F15	RF-F11
15	C6-F5	C7-F8	C7-F13	C7-F14	C7-F16	RF-F9	RF-F13

Table 2.6 Pools composition for 13q14 baits screening. Each pool (rows) consisted of 8 individual cultures of bait clones mixed together. C6=C13orf1, C7=RCBTB1, C8=SETDB2, RF=RFP2, KP=KPNA3, F=fragment (fragments are detailed in Appendix 5).

Validation of protein-protein interactions (PPIs)

For validation of PPIs, pull-down assays were performed. The technique is similar to immunoprecipitation, but precipitation is based on the property of Protein-A to bind to the heavy chain of immunoglobulins. The protein that is precipitated is fused to a Protein-A (PA) tag and pulled-down with an IgG-bound matrix. If binding partners are bound to the precipitated protein, they will be co-precipitated. For pull-down experiments, 13q14 proteins were cloned into the pcDNA3.1-PA vector and candidate binding partners into the pTL-HA vector and co-transfected into HEK-293 cells. The protein pairs were co-expressed as PA- and HA-tagged proteins, and 48hrs after transfection cell lysates were prepared (see section 2.2.8). As positive control two interacting proteins were used —SH3GL1-PA and SH3GL3-HA— kindly provided by U.Stelzl (Max Planck Institute for Molecular Genetics, Berlin). Non-interacting proteins were used as negative controls. Each interaction was assessed by co-transfection of HEK-293 cells in 2 separate experiments. In preparation, 30µl sepharose beads (IgG sepharose 6 fast flow, Amersham-GE healthcare) were washed 3x with 1:1 water and 2x with 1:1 lysis buffer and centrifuged at 2000rpm for 1min. For each interaction transfected cells of 2 wells were pooled. Cells were pelleted at 2000rpm for 5min; cell lysates were prepared with 250µl of lysis buffer and 1hr of incubation at 4ºC on the rotor; lysates were centrifuged at 13000rpm for

10min and the soluble fraction was used for the pull-down experiments. As input, 20µl of the supernatant were mixed with Laemmli loading buffer without reducing agent and kept refrigerated until gel loading. The remaining of the lysates were given to the 30µl of washed beads and bound for 4hrs with mixing. After binding, beads were washed 5x in 500µl Washing-buffer. After the last wash the supernatant was completely removed and 40µl of 1x Laemmli loading buffer without reducing agent was added and boiled for 5min. The whole sample was loaded for Western blot analysis (section 2.2.15.2). All pull-down experiments steps were performed at 4ºC if not indicated otherwise.

3. Results

From the critical region in chromosomal band 13q14 five candidate tumor suppressor proteins were chosen to be functionally characterized in detail: C13orf1, RCBTB1, SETDB2, RFP2 and KPNA3 (Table 3.1).

Gene sequences from NCBI database for target genes on 13q14

Accession number	Gene Name, Alias	Base pairs	Protein Size (KD)
AF334405	CLLD6, C13orf1	588	21.6
AF334406	CLLD7, RCBTB1	1595	58.2
NM_031915	CLLD8, SETDB2	2159	81.8
NM_002267	KPNA3	1229	58.5
NM_052811	RFP2	1216	46.9

Table 3.1 List of target genes in 13q14. Depicted are: the genes, length in base pairs and molecular weights in KD for the target genes used in this work. The NCBI accession numbers are indicated under “Accession numbers” , the names given under these accession numbers and the alias used in this work if they differed are indicated under “Gene Name, Alias”, the size, in base pairs, of the coding sequences used for protein expression are indicated under “Base pairs” and the expected molecular weights for each protein is indicated in Kilo Daltons (KD) under “Protein size”.

To elucidate the function of 13q14 proteins, three strategies were taken: a) expression of recombinant fusion proteins in prokaryotic and eukaryotic cells; b) production of antibodies for the analysis of the endogenous proteins; c) finding binding partners with a yeast-two-hybrid (Y2H) screen. First, the gathering of information about the intracellular localization of the proteins should point to their function. Second, identification of binding partners will link the candidate genes to known processes and functions and ultimately lead to a better understanding of the cellular role of 13q14 proteins. To this end, we produced recombinant proteins, specific antibodies and performed a Y2H screen.

3.1. Expression of recombinant fusion proteins in bacteria

Proteins were expressed in bacteria to be used in the production of specific antibodies. For purification and detection purposes, the proteins were expressed as fusion proteins with a polypeptide tag. Expression of different fusion proteins was tested including Histidine tags, V5-tag, X-press-tag, Nus-tag and GST-tag (Table 3.2). The results for expression with the different tags varied but all inserts in vectors from 1 to 8 in Table 3.2 had very poor expression and detection of over-expressed protein was either undetectable with the methods used (Coomassie staining and Western blot analysis) or only detectable by Western blot analysis (Figure 3.3).

Vectors used for recombinant expression of proteins in bacteria

Nr	Vector Name	Tag	Remarks	Distributor
1	pTrcHis-NT	Xpress+His	Sp.Ab	Invitrogen
2	pCRT7-CT	V5+His	Sp.Ab/matrix	Invitrogen
3	pDEST-14	No tag	native protein	Invitrogen
4	pDEST-15 (NT)	GST	Sp.Ab	Invitrogen
5	pDEST-17 (NT)	His	Sp.Ab/matrix	Invitrogen
6	pET43.1a (NT)	Nus+His(+HSV)	Sp.Ab/matrix	DB
7	pPEP (NT)	6xHis	Sp.Ab	DKFZ
8	pDS5	No tag	native protein	DB
9	$NHis_{10}$-zzTev-80N (NT)	ZZ tag+10xHis	Sp.Ab/matrix	ZMBH
10	$NHis_{10}$-80N (NT)	10xHis	Sp.Ab/matrix	ZMBH

Table 3.2 List of the vectors that were used for recombinant protein expression in bacteria. The vectors names are indicated as well as the tags contained in those vectors. “Remarks” lists whether there were specific antibodies (Sp.Ab) and purification matrixes (matrix) available for these tags or whether the proteins would be expressed as native form without any tag. Distributors of vectors are also listed. The vector from DKFZ was kindly provided by Prof. H. Hermann and the vectors from ZMBH were kindly provided by Prof. D. Görlich. NT=N-terminal tag, CT=C-terminal tag.

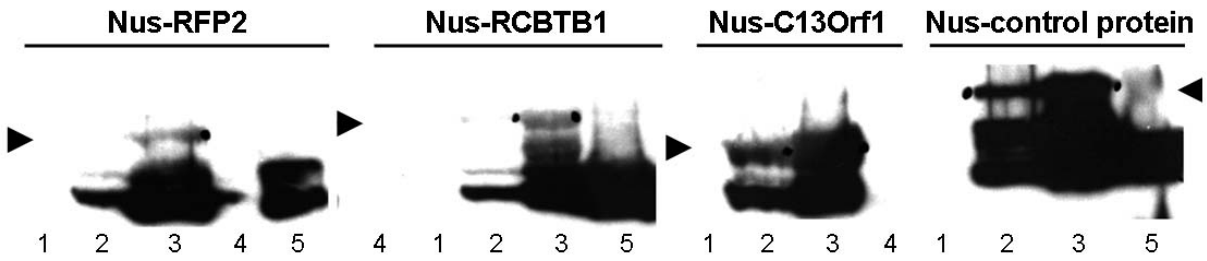


Figure 3.3 Western blot analysis of over-expressed 13q14 proteins. Proteins were expressed in bacteria in the vector pET43.1a (Nus and His tags) and detected with anti-His antibody. The arrow heads indicate the fusion proteins. 1=transformed cells before induction, 2=transformed cells 1hr after induction, 3=transformed cells 4hrs after induction, 4=untransformed cells 4hrs after induction, 5=cells transformed with empty vector.

To test the possibility of low expression levels of human genes in bacteria related to differences in codon usage in prokaryotic cells, the gene sequences were optimized to match the genetic code for bacterial translation by introduction of conservative point mutations (Appendix 3). Mutated sequences were expressed with an *in vitro* expression system (RTS, Roche Applied Science, Mannheim), but no remarkable changes in expression were detected with the improved sequences (Figure 3.4).

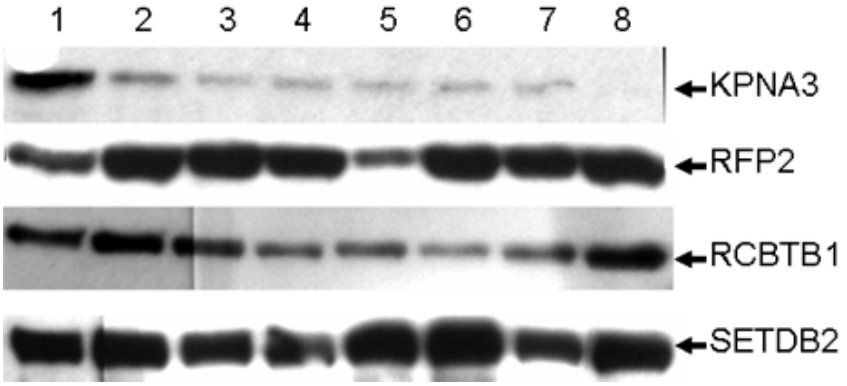


Figure 3.4 Western blot analysis of *in vitro* expressed 13q14 genes. All of the sequences contain a NT-His tag and were detected with anti-His antibody. 1 corresponds to the original gene sequence as in Table 3.1 and 2-8 correspond to conservative mutated sequences (Appendix 3). Arrows indicate in vitro expressed KPNA3, RFP2, RCBTB1 and SETDB2.

Therefore, it was concluded that the low expression levels were probably not due to incompatible codon usage, and the original sequences were taken for further vector testing. In contrast to the low levels of expression in commercially available vectors, expression in the vectors $NHis_{10}$-zzTev-80N and $NHis_{10}$-80N (courtesy of Prof. D. Görlich) was very high and detectable even with Coomassie staining for most of the candidate proteins (Figure 3.5).

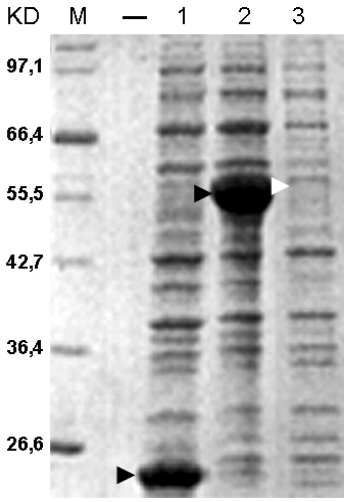


Figure 3.5 Coomassie staining of recombinant proteins expressed in *E.coli*. Proteins were separated by SDS-PAGE in 10% Acrylamide gel. The *E.coli* strain Star-Rosetta was used for transformation and cells were grown at 37ºC. 4hrs after induction, 10µl of total cell lysates were loaded in each well. C13orf1 (1) and RCBTB1 (2) show high over-expression (black arrow-heads) while KPNA3 (3) shows no over-expression (white arrow-head). KD=Kilo Daltons, M=marker, — =empty line.

The full length proteins C13orf1, RCBTB1 and SETDB2 as well as $RFP2_{S154\text{-}E264}$ (a polypeptide containing RFP2 amino acids S154 to E264) could be expressed recombinantly in bacteria. Best results were obtained with a strain deficient in several proteases (Star-Rosetta (Invitrogen, Karlsruhe)) and the vectors 9 and 10 (Table 3.2). The amount of over-expressed protein varied depending on (1) the protein and (2) the growth conditions (the temperature and concentration of IPTG used for induction were varied). Concerning the proteins, C13orf1 and $RFP2_{S154\text{-}E264}$ were expressed strongest followed by RCBTB1; SETDB2 was the weakest, and nothing was seen for KPNA3 and the full length of RFP2 with Coomassie staining. Only with Western blot analysis a weak band of the expected size was seen for RFP2 full length while no KPNA3 fusion protein was detectable (Figure 3.6.A). Concerning the conditions (Figure 3.6.B), changes in temperature and IPTG concentration did not have an effect on the not expressing proteins, while 37ºC and 1mM IPTG gave optimal and satisfactory results for the strong and weaker expressing proteins, respectively, and were therefore used as the standard conditions for expression.

From the results obtained in these tests, it can be concluded that the main factors responsible for low expression of the 13q14 candidate proteins are vector characteristics and degradation of the proteins. Therefore, for further experiments of purification, proteins were expressed in a protease deficient strain (Star Rosetta *E.coli*) using the vectors kindly provided by Prof. D. Görlich ($NHis_{10}$-80N and $NHis_{10}$-zzTev-80N).

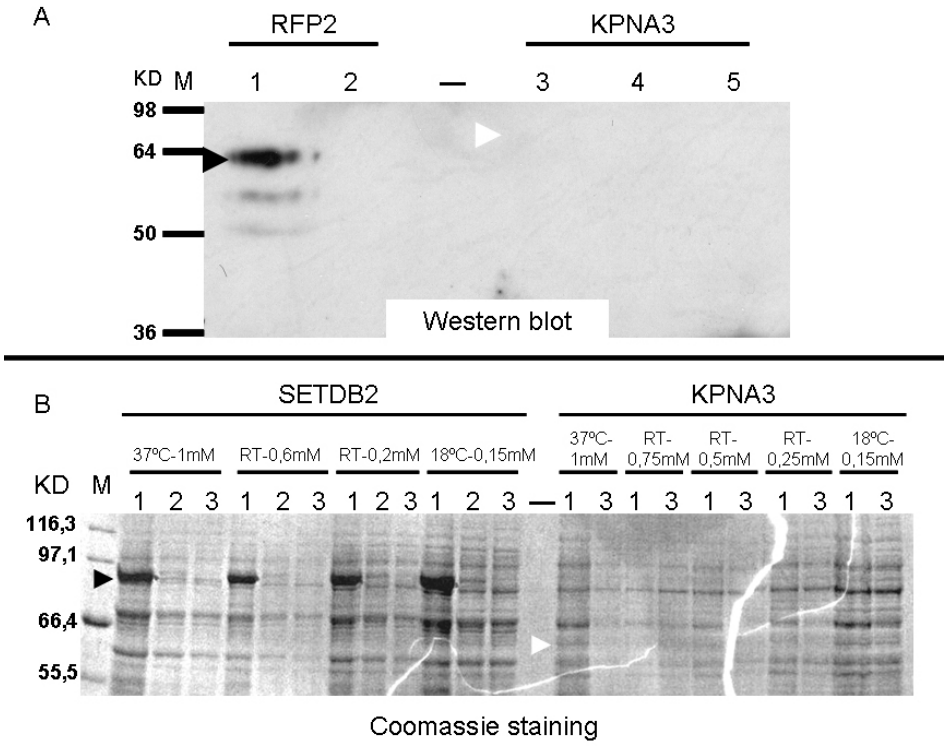


Figure 3.6 Expression tests for candidate proteins in *E.coli*. Proteins from 13q14 were tested for over-expression in Rosetta *E.coli*. **A.** Western blot analysis with α-His antibody to detect recombinant RFP2 and KPNA3. A specific band the expected size of RFP2 recombinant protein (marked with a black arrow-head) is seen (1) as compared with untransfected cells (2). For KPNA3 no band of the expected size can be detected (white arrow-head) at varying induction conditions (3= 0,4mM IPTG and 5= 1mM IPTG) as compared with the untransfected cells (4). **B.** Coomassie staining for recombinant SETDB2 and KPNA3. Growth conditions for expression are noted over the corresponding lines. Temperature indicates the temperature of growth and mM indicates the concentration of IPTG used for induction. SETDB2 expression varies little with conditions (prominent band of expected size for the recombinant protein marked with a black arrow-head) and is mainly insoluble (mainly present in the total lysates (1)). For KPNA3 no prominent band of the expected size can be detected (white arrow-head). 1=total lysate, 2=supernatant after centrifugation at 4000rpm during 15min., 3=supernatant after centrifugation at 45000rpm during 20min. KD=Kilo Daltons, M=marker, — = empty well.

3.1.1. Expression of recombinant RFP2 alters bacterial growth

In contrast to the other proteins, RFP2 full length was very weakly expressed from a single vector and only detectable in Western blot analysis (Figure 3.6). Interestingly, bacteria expressing the RFP2 full length protein grew slower compared to bacteria transformed with other 13q14 candidate proteins or the shorter $RFP2_{S154\text{-}E264}$ fragment (Figure 3.7 and 3.8). It is shown that expression of RFP2 full length protein or of its C-terminal domain —$RFP2_{S154\text{-}E409}$ — is compromised in bacteria and that under the conditions tested here, expression of the constructs causes decreased bacterial growth. Further experiments will be needed to clarify this effect, but it indicates that possibly the RFP2 protein could play a role in conserved pathways that are involved in cell proliferation.

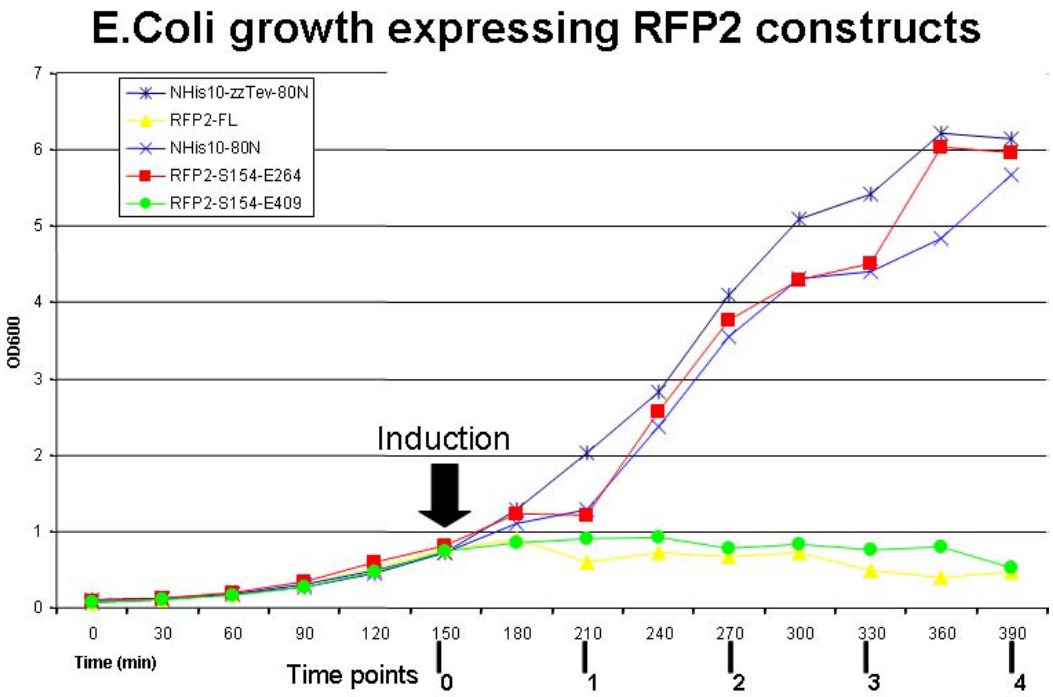


Figure 3.7 Graphic representation of growth in bacteria expressing RFP2. Indicated are growth measurements as optic density of the cultures at 600nm (OD600). At OD600~0.8 (time point 0 at 150min) cultures were induced with 1mM of IPTG, and samples were taken and measured at the indicated time points (dots). The time points indicated with black lines represent samples taken for SDS-PAGE and Coomassie staining shown in Figure 3.8. The cells with the empty vectors (dark and light blue lines) show growth, while after induction only the $RFP2_{S154-E264}$ construct (red line) grows like the empty vector transfected cells. Both, the full-length RFP2 expressing cells (yellow line) and the fragment containing the C-terminal part of the RPF2 protein ($RFP2_{S154-E409}$, green line) stopped growing after induction.

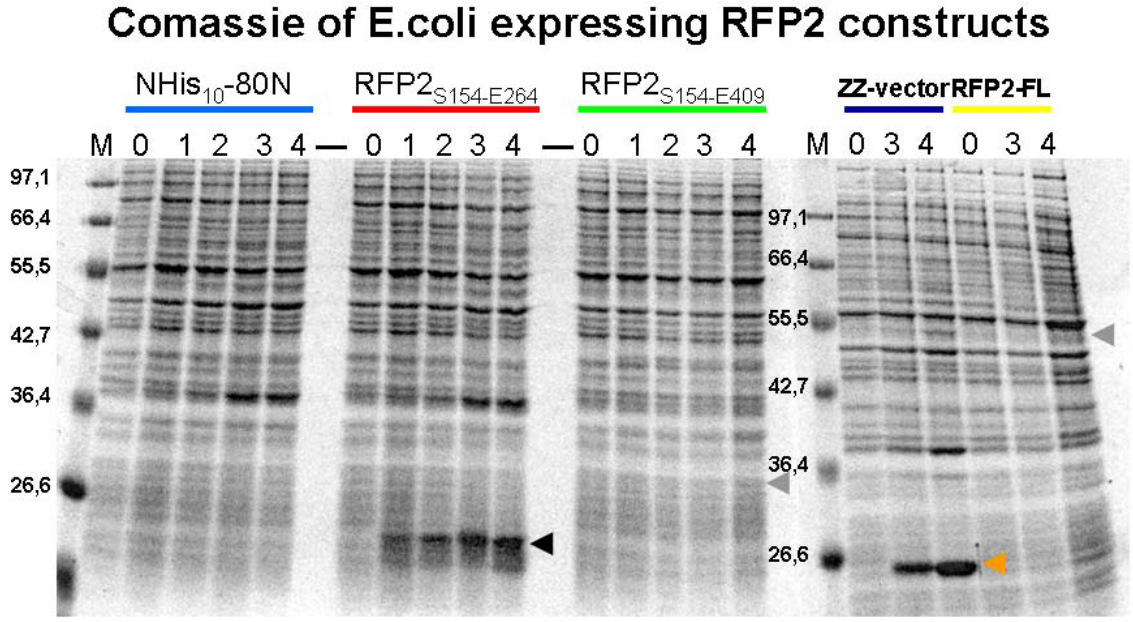


Figure 3.8 Coomassie staining of samples from the cultures depicted in Figure 3.7. The empty vectors (ZZ-vector and $NHis_{10}$-80N) show tag expression (orange arrow head for ZZ+His-tag and no detection for 10xHis-tag because it run out of the gel) together with the fragment expression of normal growing $RFP2_{S154-E264}$ (black arrow head), while the not growing cells show no protein expression (samples from $RFP2_{S154-E409}$ and RFP2-FL underlined green and yellow respectively (grey arrow-heads)).

3.1.2. Purification of recombinant 13q14 candidate tumor suppressor proteins

The candidate tumor suppressor proteins containing a Histidine (His) tag were purified by metal-affinity chromatography using Nickel (Ni) ions. Ni ions are immobilised on an inert matrix and bind His residues with high affinity. Two His residues suffice for binding to the Ni-matrix, but for the purification of 13q14 expressed proteins a 10x His-tag was used to further increase the binding of the recombinant proteins. Purification was performed under denaturing conditions as the 13q14 proteins were deposited mostly as insoluble form in inclusion bodies. The over expressed C13orf1, RCBTB1, SETDB2 and $RFP2_{S154\text{-}E264}$ could be highly purified as assessed by Coomassie staining (Figures 3.9 and 3.10). In the purification of C13orf1 there was a band between 47.7 and 55.5 KD, in addition to the expected size of ~23 KD, that was cut and analyzed by mass spectrometry. The results verified that it was the C13orf1 protein. Thus, it seems that the C13orf1 protein forms aggregates, possibly dimers, of very strong self-binding affinity as they are not disturbed by the reducing conditions of the loading buffer and boiling the sample (at 95ºC for 10min) before loading on the gel.

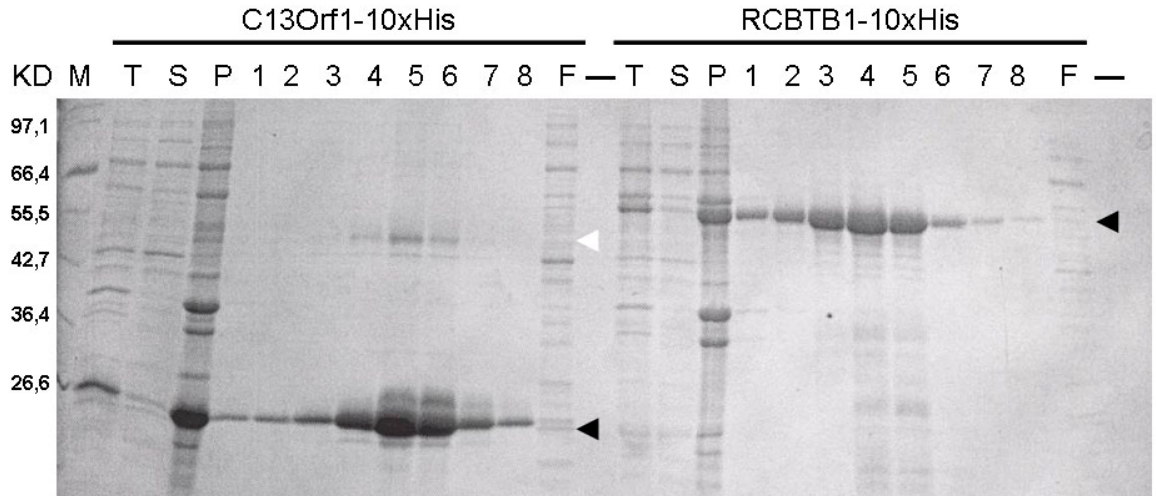


Figure 3.9 Coomassie staining of purification of recombinant proteins C13orf1 and RCBTB1. 13q14 proteins were expressed recombinantly with a 10xHis tag for purification. Total cell lysates (T) were centrifuged at 4000rpm to separate soluble proteins (S) and pellet inclusion bodies (P). After purification of the proteins (Ni-NTA, Quiagen) they were eluted by pH (1-8). Purified proteins are indicated with black arrow-heads The band between 47.7 and 55.5 KD (white arrow-head) corresponds to C13orf1 (analysis by mass-spectrometry). KD=Kilo Daltons, M=marker, F=flow through, — =empty line.

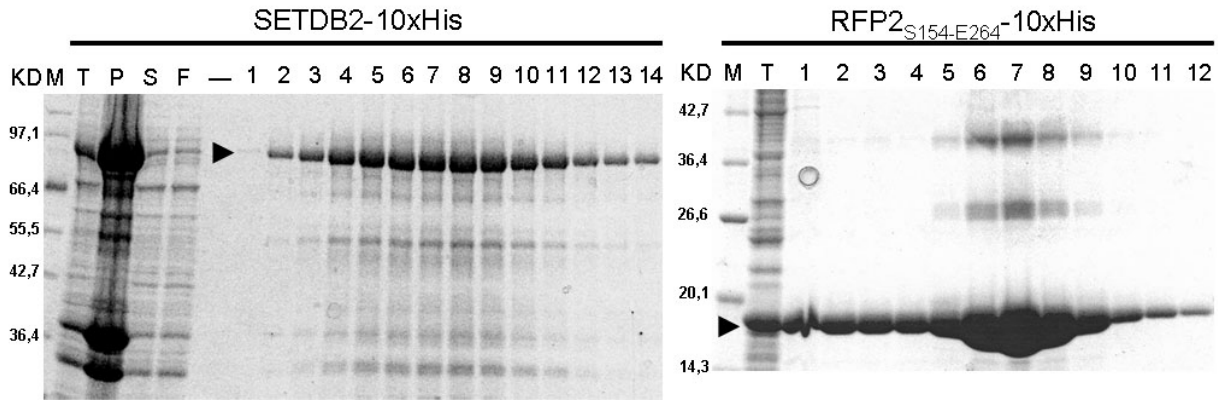


Figure 3.10 Coomassie staining of purification of recombinant proteins SETDB2 and RFP2$_{S154-E264}$. 13q14 proteins were expressed recombinantly with a 10xHis tag for purification. Total cell lysates (T) were centrifuged at 4000rpm to separate soluble proteins (S) and pellet inclusion bodies (P). After purification of the proteins (Ni-NTA, Quiagen) they were eluted by pH (1-14). Purified proteins are indicated with black arrow-heads. KD=Kilo Daltons, M=marker, F=flow through, — =empty line.

Smaller bands were detected in the expression and purification of some 13q14 proteins, especially for SETDB2 and RCBTB1. These fragments are likely degradation products or result from incomplete expression as they always showed the same pattern and were also present after expression in mammalian cells. For use in immunization, fragmentation is not a mayor problem since they constitute parts of the full length protein and would, if immunogenic, produce specific antibodies. After purification the eluted fractions with the highest concentrations were pooled and used as antigen for immunization and subsequent testing of the produced antibodies. The proteins with putative enzymatic potential were also assessed for renaturation.

3.1.2.1. 13q14 proteins renature independently of the salt concentration at high pH

For the antibody production the native conformation of the proteins is not always required. Nevertheless, it would be interesting to be able to renature the proteins because then they could be tested functionally. Among the proteins of interest on 13q14, two have predicted enzymatic activities: RCBTB1, with a predicted guanine nucleotide exchange factor activity, and SETDB2, predicted to be a histone methyl transferase (HMTase). In order to renature the recombinantly expressed proteins, 12 different buffers were tested for re-solubilizing the proteins (Table 3.10)

Buffer composition for renaturation test				
Buffer Nr	**Tris (mM)**	**Phosphate (mM)**	**pH**	**Salt (mM)**
1	50	—	7.5	20
2	50	—	7.5	100
3	50	—	7.5	200
4	50	—	7.5	500
5	—	100	7	20
6	—	100	7	100
7	—	100	7	200
8	—	100	7	500
9	—	100	8	20
10	—	100	8	100
11	—	100	8	200
12	—	100	8	500

Table 3.10 Buffer composition for renaturation test. Indicated are the buffer numbers, amount of Tris or Phosphate in mM, the pH and the amount of salt in mM.

The results are summarized in Figure 3.11 and indicate that for these two proteins, the factors determining solubility were a) the pH, with higher pH (8.0) resulting in more soluble protein, and b) the salt concentration, with more soluble protein at higher salt concentrations (>200mM NaCl). It remains to be tested whether the soluble proteins obtained will also be enzymatically active.

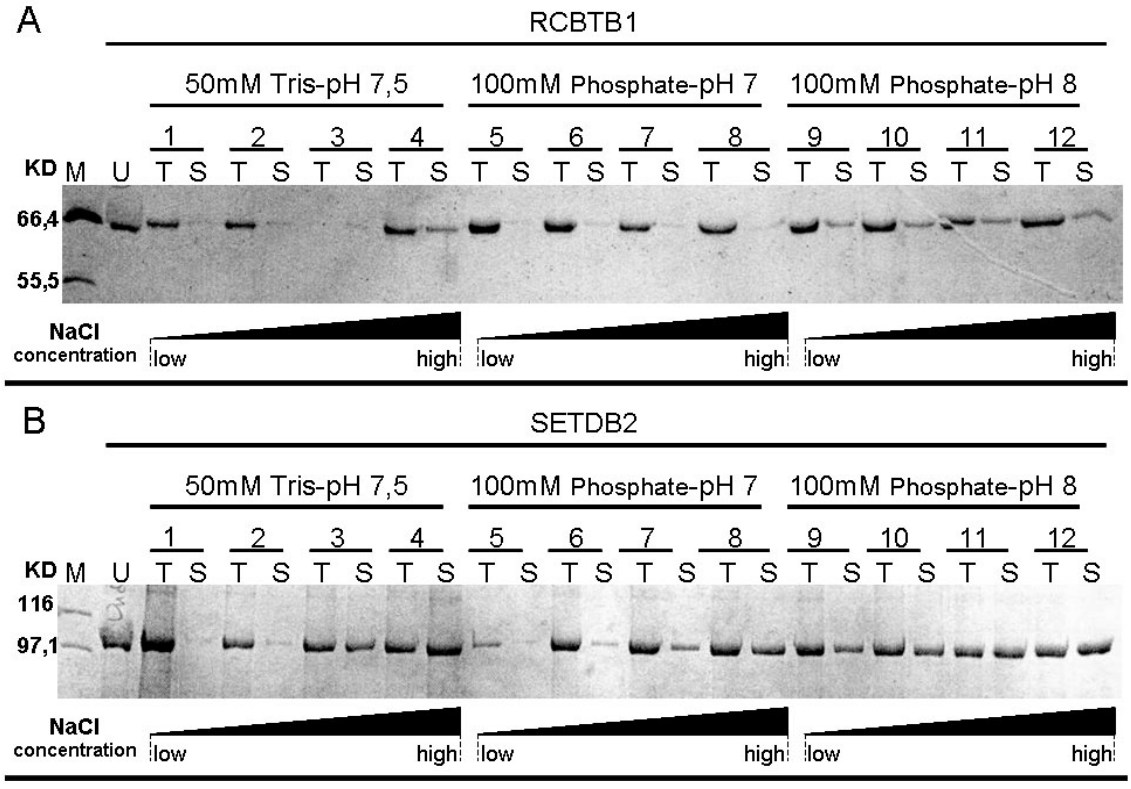


Figure 3.11 Solubility test for RCBTB1 and SETDB2. Recombinantly expressed proteins purified under denaturing conditions in 8M Urea were dialysed against the indicated buffers 12 to 18 hours. Dialysis was performed at 4ºC and in a dilution volume of 1:25000. Fractions were stained with Coomassie following SDS-PAGE. **A**. RCBTB1 re-solubilises at high pH (samples 9 to 12) and pH 7,5 and high salt (sample 4) but there is minimal impact of salt at low pH (samples 5 to 8). **B.** SETDB2 re-solubilises at high pH (samples 9 to 12) and at low pH if salt concentration is high (samples 3-4 and 7-8). 1=20mM NaCl, 2=100mM NaCl, 3=200mM NaCl, 4=500mM NaCl, U=sample before dialysis, T=total dialysed sample, S=supernatant after centrifugation at 50000rpm.

3.2. Production of specific antibodies for 13q14 candidate proteins

Once the target proteins on 13q14 were obtained in pure form and in high quantities in bacteria, they were used as antigens to produce specific antibodies. Specific antibodies are raised by immunizing animals repeatedly with the target antigen. After immunization either the blood-serum of the animals was used as antibody source (polyclonal antibodies) or the antibody-producing lymphocytes were isolated and fused to an immortal cell line to create hybridomas that in turn secrete specific antibodies (monoclonal antibodies). Polyclonal antibodies contain antibodies specific for different epitopes in the antigen and some show positive results when used in different techniques such as immunofluorescence, Western blotting or immunoprecipitation, which makes them very useful. In contrast to that, monoclonal antibodies are produced in a way that only one antibody-producing B-cell will be selected, and therefore the antibody will only recognize one of the possible epitopes of the antigen, giving highly specific antibodies with low background signal contamination. In this work, both polyclonal and monoclonal antibodies against 13q14 proteins were produced.

3.2.1. Polyclonal antibodies from Guinea pigs

For the production of polyclonal antibodies Guinea pigs were immunized with C13orf1, RCBTB1, SETDB2 and $RFP2_{S154\text{-}E264}$ 10xHis purified proteins. The polyclonal antibodies were tested a) against the purified proteins in Western blot analysis and b) against proteins over-expressed in mammalian cells as shown by Western blot analysis and immunofluorescence for light microscopy (Figure 3.12).

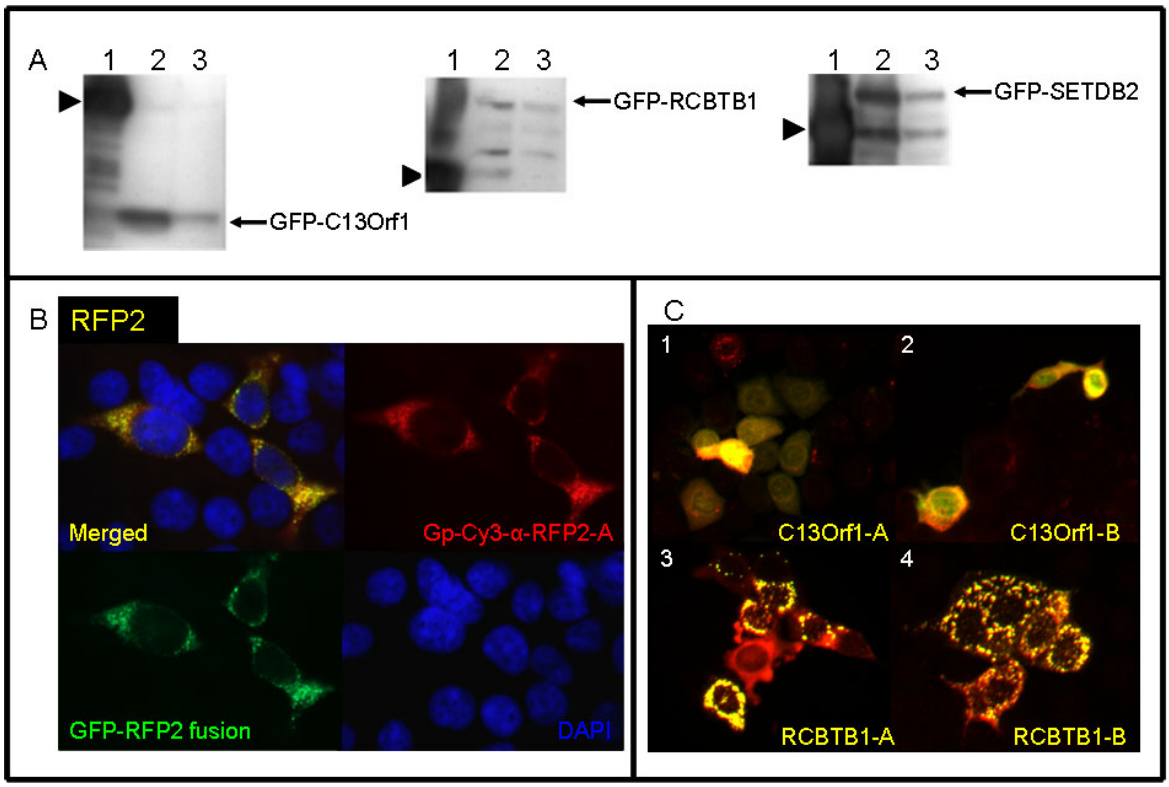


Figure 3.12 Guinea pig antibodies specific for 13q14 proteins. Specific antibodies were tested in Western blot analysis (A) and immunofluorescence (B and C). **A.** Analysis of α-C13orf1, α-RCBTB1 and α-SETDB2 specific antibodies. Positive controls (1) were purified proteins (Nus-C13orf1, 10xHis-RCBTB1 and 10xHis-SETDB2, all marked with a black arrow head). Arrows indicate the expected sizes for each protein expressed in HEK-293 (2) and MCF-7 (3). **B.** For GFP-RFP2 protein, the merged signal (yellow) indicates that the antibody (red) is specific for the target protein (green). DNA is stained with DAPI (blue). **C.** Merged images of staining with C13orf1-specific antibodies from animals A (1) and B (2) and RCBTB1 from animals A (3) and B (4). All antibodies were used at a dilution 1:1000. Immunofluorescence was performed in transfected HEK-293 and specific antibodies were detected with Cy3-labelled secondary antibody.

To elucidate the best working dilutions for the antibodies, different dilutions were tested by Western blot analysis (Figure 3.13) and by immunofluorescence for light microscopy. The signals of the specific antibodies were evaluated against endogenous fluorescence of the GFP moiety of the respective fusion proteins (examples shown in Figure 3.12). The antibodies were also tested over untransfected cell extracts but with negative results, and in untransfected fixed cells, which showed rather weak endogenous protein signals. The antibodies were considered to be specific based on the results with recombinant proteins from the Western blot analysis (positive detection of recombinant protein) and analysis of the immunofluorescence under the light microscope (co-localization of signals from GFP-fusion and antibody). The dilutions at which the antibodies worked best for over-expressed proteins in both techniques are summarized in Table 3.14.

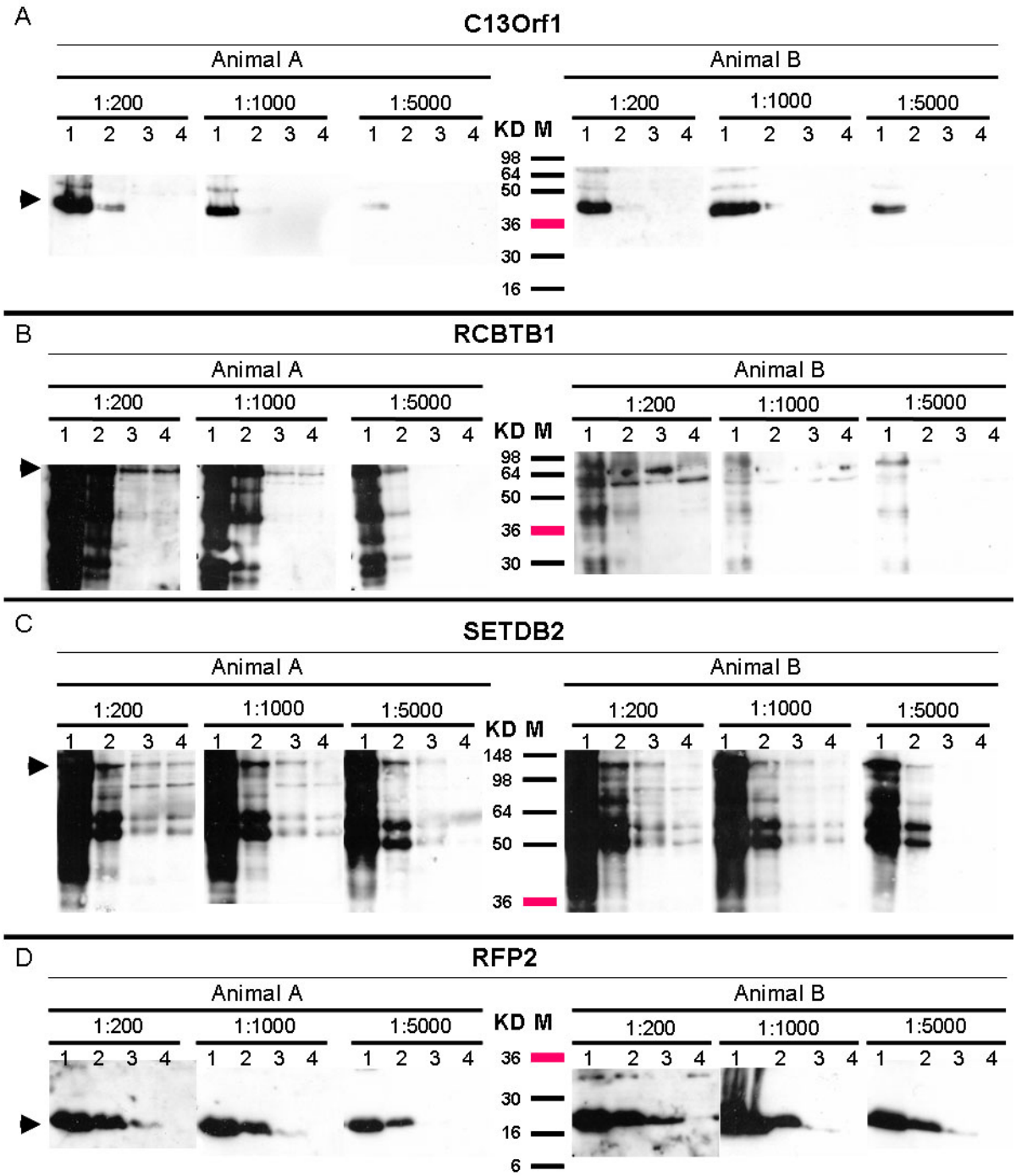


Figure 3.13 Western blot analysis for dilution series of polyclonal antibodies. Purified proteins were used as samples (1=1µg, 2=100ng, 3=10ng, 4=1ng). Polyclonal antibodies from 13q14 genes were at indicated dilutions and sera from two different animals (A and B) were used for comparison. Expected sizes for the purified proteins are indicated with black arrow-heads. **A.** Both C13orf1 sera present low background and recognize only high amounts of protein (bands in samples 1 or 2). **B.** Both RCBTB1 sera present some background. Antibody from animal B shows very weak signal at all concentrations and antibody from animal A recognizes mainly high amounts of protein (samples 1 and 2). **C.** Both SETDB2 sera present some background and recognize a wide range of protein amount (samples 1 to 4). **D.** Both RFP2 sera recognize mainly high amounts of protein (samples 1 and 2) and lower amounts at lower dilutions (samples 3 with dilutions 1:200 and 1:1000) with antibodies in serum A presenting slightly less background than serum B.

Polyclonal Antibodies		Western Blot		Immunofluorescence
Gene	Antibody	Dilution	Range	Dilution
C13Orf1	GpαC13Orf1-A	1:1000	µg	1:1000 to 1:2000
		1:200	>100ng	
	GpαC13Orf1-B	1:1000	µg	1:1000 to 1:2000
RCBTB1	GpαRCBTB1-A	1:1000	ng	1:1000 to 1:2000
	GpαRCBTB1-B	Bad signal→**Do not use**		1:1000 to 1:2000
SETDB2	GpαSETDB2-A	1:5000	>100ng	1:1000 to 1:5000
		1:200	ng	
	GpαSETDB2-B	1:5000	>100ng	1:1000 to 1:2000
		1:200	>10ng	
RFP2	GpαRFP2-A	1:1000	>100ng	1:1000 to 1:2000
		1:200	>10ng	
	GpαRFP2-B	1:1000	>100ng	1:1000 to 1:2000
		1:200	>10ng	

Table 3.14 Dilutions of polyclonal antibodies for over-expressed 13q14 proteins. Depicted are dilutions for Western blot analysis and Immunofluorescence. The "Antibody" column indicates the guinea pig (Gp) antibodies for each protein from animal A (-A) or B (-B). The range indicates the amount of proteins that was detected in Western blot analysis: µg=able to detect 1µg purified protein, ng=able to detect 1ng purified protein, >100ng=able to detect 100ng or more purified protein, >10ng=able to detect 10ng or more purified protein.

3.2.2. Monoclonal antibodies from Mice

For the production of monoclonal antibodies, mice were also immunized with C13orf1, RCBTB1, SETDB2 and $RFP2_{S154\text{-}E264}$ 10xHis recombinant proteins obtained from bacteria. KPNA3 could not be expressed recombinantly in bacteria —similar to Fagerlund et al. (Fagerlund et al. 2005)—, so two synthetic peptides conjugated to KLH (keyhole limpet hemocyanin protein) —one N-terminal and one C-terminal (Appendix 4)— were used for immunization. After immunization, B-cells extracted from the lymph nodes were fused to myeloma cells to generate hybridomas that produce antibodies and can grow indefinitely in culture. At least 100 clones that grew under selection were tested for each protein for their ability to bind to the antigen in ELISA (Figure 3.15), and the clones producing most sensitive antibodies were further tested in Western blot analysis and immunofluorescence. Ultimately, a single clone was selected and grown to produce sufficient amount of antibody-containing culture medium for purification of the monoclonal antibody. After purification, antibodies were re-tested for sensitivity and specificity by Western blot analysis and immunofluorescence in light microscopy (RCBTB1 is shown in Figure 3.16).

Elisa Reading

	1	2	3
A	2.201	1.844	1.621
B	2.142	0.099	1.576
C	2.110	0.169	0.078
D	1.990	0.347	0.063
E	0.119	0.804	0.090
F	0.187	1.825	0.139
G	0.164	0.149	0.074
H	0.276	0.191	0.087
	C13Orf1	RCBTB1	RFP2

Coated plate

	1		2		3	
A	C6 #13	+	C7 #33	+	RF #40	+
B	C6 #14	+	Medium	-	RF #125	+
C	C6 #35	+	PBS	-	Medium	-
D	C6 #50	+	C6 #14	-	PBS	-
E	Medium	-	C6 #35	-	C6 #13	-
F	PBS	-	C7 #33	+	C6 #35	-
G	C7 #33	-	RF #40	-	C7 #33	-
H	C8 #17	-	RF #125	-	C8 #17	-
	C13Orf1		RCBTB1		RFP2	

Figure 3.15 ELISA measurements for clones of C13orf1, RCBTB1 and RFP2. Plates (right) were coated with the purified antigen used for immunization (colour coded proteins indicated under the plate). Supernatant from the indicated cultured hybridomas was taken as 1^{ry} antibody for ELISA. Clones were considered to be positive for values over 1.0 in the ELISA reading (left) and are labeled in the same colour as their corresponding protein and marked in the right panel with a + symbol. Medium, PBS and medium from other hybridomas were used as negative controls and are marked with a – symbol in the right panel and are uncoloured in both panels. Hybridoma supernantants from C13orf1 (clones 13, 14, 35 and 50), RCBTB1 (clone 33) and RFP2 (clones 40 and 125) show high values for the corresponding antigens and low for the corresponding negative controls. C6=C13orf1, C7=RCBTB1, RF=RFP2, #=clone number.

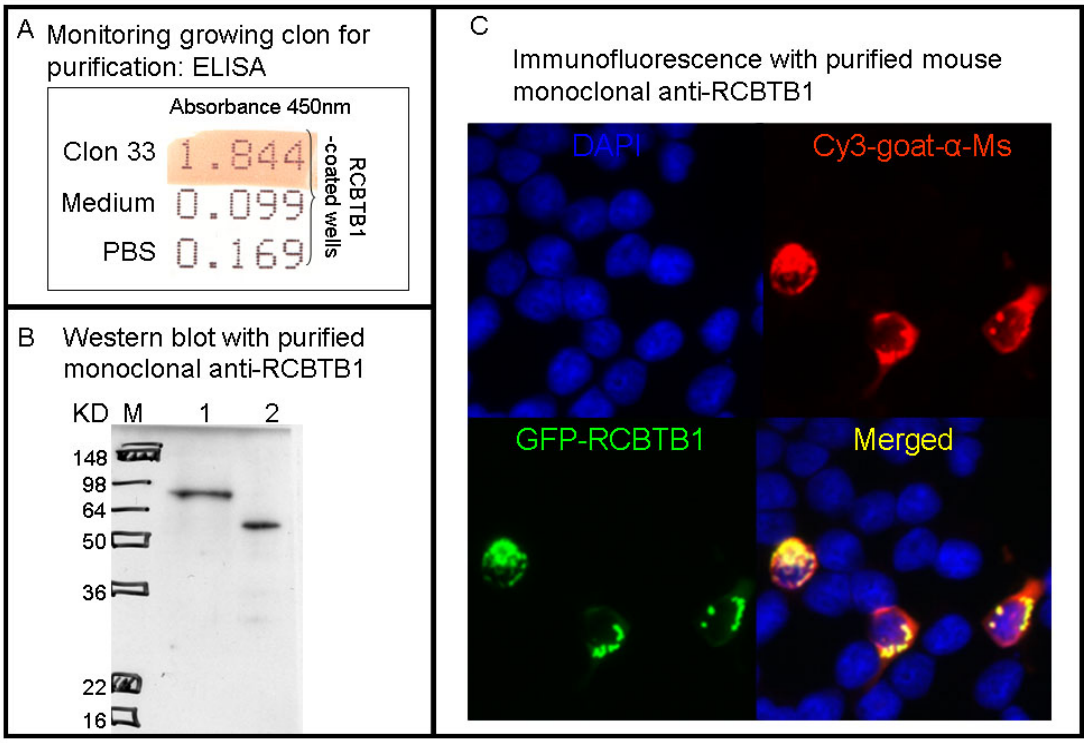


Figure 3.16 Example for selection of monoclonal antibodies for RCBTB1. Routinely, clones were tested for ELISA, Western blot analysis and immunofluorescence. Depicted are the tests performed with anti-RCBTB1 antibody (clone #33). **A.** ELISA values. **B.** Western blot analysis with 0.2µg/ml purified anti-RCBTB1. The antibody recognizes both, purified protein (sample 2) and over-expressed GFP-RCBTB1 fusion protein (sample 1). **C.** Immunofluorescence of transfected HEK-293 cells expressing GFP-RCBTB1 fusion protein (green). Cells were fixed and immuno-stained with 0.2µg/ml mouse-α-RCBTB1 and 1:400 Cy3-labelled Goat-anti-Mouse (Cy3-goat-α-Ms). Specific staining is seen as a red signal and overlay with the GFP-RCBTB1 signal is shown in yellow (merged). DNA is shown in blue with DAPI staining.

Different concentrations of antibody were used to determine the lowest concentration that could detect over-expressed protein in whole cell extract (example for C13orf1 is shown in Figure 3.17).

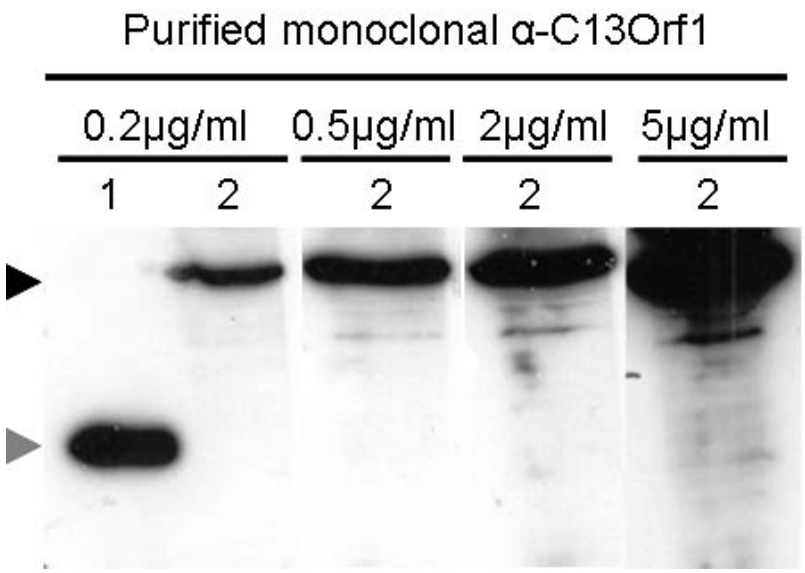


Figure 3.17 Sensitivity of monoclonal antibody against C13orf1. As control 200ng purified protein were loaded (sample 1). Samples (2) consisted of 20µg of total cell extract from cells over-expressing GFP-C13orf1. Purified antibody was at the concentrations indicated (0.2µg/ml to 5µg/ml). Black arrow-head indicates GFP-C13orf1 fusion protein and grey arrow-head indicates purified 10xHis-C13orf1.

The same sensitivity test was performed for all 13q14 specific monoclonal antibodies and 4/5 antibodies showed specificity and sensitivity in Western blot and immunofluorescence assays (Table 3.18).

Protein	Clon	Subcl.	Western	Immunofluores.
C13Orf1	#13-1004	IgG1	0.2 µg/ml	0.5 µg/ml
RCBTB1	#33	IgG2	0.5 µg/ml	0.5 µg/ml
SETDB2	#40	IgG1	None	2 µg/ml
RFP2	#14	IgG1	0.5 µg/ml	0.5 µg/ml
KPNA3	NT#44	IgG1	0.2 µg/ml	2 µg/ml

Table 3.18 Monoclonal antibodies concentrations to detect 13q14 candidate proteins. The columns indicate: the protein used as antigen to raise the antibodies ("Protein"), the clon used for purification ("Clon"), the subclass of the antibody as tested in ELISA ("Subcl."), the lowest concentration of antibody used for Western blot analysis detection that gave signal ("Western") and the lowest concentration that gave strong signal in immunofluorescence ("Immunofloures.").

3.2.3. Expression of 13q14 genes in cell lines

Because of difficulties in detecting endogenous proteins with the specific antibodies, several cell lines were screened for expression of the candidate genes to determine, whether the genes were expressed in the cell lines being used for the experiments or in other cell lines available in the laboratory. Using RT-PCR, the expression of the candidate genes from 13q14 (C13orf1, RCBTB1, SETDB2 and RFP2) was compared to that of house-keeping genes (PGK and Lamin-B1) in 40 cell lines derived from various tissues. The data from both house-keeping genes was highly variable when compared to each other, as well as in relation to the different cell lines, indicating that the 13q14 genes present highly variable expression in different cell lines. Therefore it was decided to consider the data of each cell line separately, and compare it to the corresponding value of each of the house keeping genes. The data comparing to PGK1 (Figure 3.18.A) and Lamin-B1 (Figure 3.18.B) show that there are cell lines, in which the genes are expressed at levels around that of the house-keeping genes or even higher (values marked in orange in the Appendix 6). Interestingly, the genes were relatively low expressed in the cell lines used for detection experiments (Namalwa, HEK-293 and MCF-7). Although with high expression variability, the 13q14 candidate genes were detectable at transcription level, while the specific antibodies described in this work were not able to detect endogenous proteins strongly. This data highlights the already known low correlation between mRNA levels and protein levels for specific genes (Griffin et al. 2002, Gygi et al. 1999), and suggests the existence of post-transcriptional regulation of 13q14 candidate genes. Due to the lack of high correlation between expression data and endogenous protein detection with 13q14 specific antibodies, the approach of co-immunoprecipitation of interaction partners was not feasible and interaction partners were analyzed in a yeast-two-hybrid screen (section 3.4).

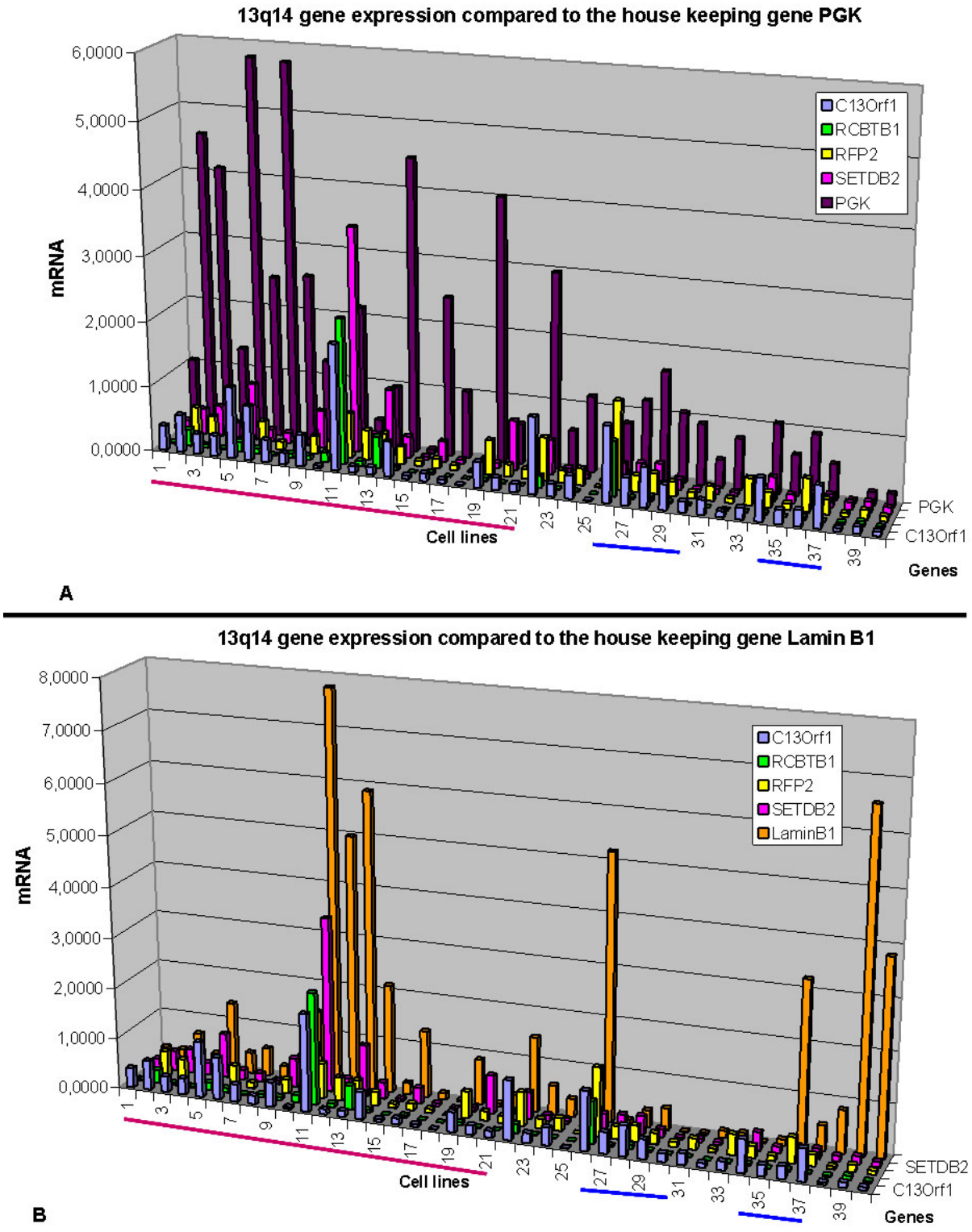


Figure 3.18 Real-time PCR analyses of 13q14 candidate genes compared to the house-keeping genes PGK1 and Lamin B1. The bars height (Y axe) indicate the absolute measurements in ng of mRNA present in the samples. 1 to 40 represents the different cell lines in the X axe (listed in Appendix 6). The different genes are represented in the Z axe colour-coded (blue=C13orf1, green=RCBTB1, yellow=RFP2, pink=SETDB2, violet=PGK and orange=Lamin B1). The comparison to both house-keeping genes (PGK1 (A) and Lamin-B1 (B)) show high variability in gene expression among different cell lines as well as lower expression of the candidate genes in most of the cell lines. Hematopoietic cell lines are marked with a violet line and central nervous system cells with a blue one.

From sections 3.1 and 3.2 it can be concluded that most proteins of interest from 13q14 could be expressed in bacteria and that the production of specific antibodies, both monoclonal and polyclonal, was successful. The polyclonal antibodies detected over-expressed proteins and gave weak signals for the endogenous proteins, while the monoclonal antibodies also detected over-expressed proteins but failed to clearly detect endogenous protein. Taking into consideration that some of the antibodies detected at least low levels of endogenous protein, the antibodies were considered satisfactory to be used in protein localization studies, and their results verified protein localization assessed with over-expressed proteins (chapter 3.3). To assess whether the low levels of endogenous protein detection correlated with low levels of gene expression, a real time PCR (RT-PCR) analysis was performed (section 3.2.3). Results showed that, for 13q14 genes, there is no high correlation between expression data and endogenous protein detection with specific antibodies. This low level of detection with the antibodies precluded a biochemical search for interaction partners of the 13q14 candidate proteins such as immunoprecipitation, which is why interaction partners were identified using a yeast-two-hybrid screen (chapter 3.4).

3.3. Intracellular localization of 13q14 proteins

The localization of 13q14 proteins within the cell was of major interest because it would help to functionally characterize the candidate proteins. It was approached by fluorescence microscopy of both over-expressed and endogenous proteins.

3.3.1. Recombinant protein detection of 13q14 genes

The intracellular localization of over-expressed 13q14 proteins was determined by fluorescence microscopy of HEK-293 cells transfected with vectors containing GFP-constructs with the candidate proteins. Images showed that the over-expressed proteins had intracellular localization according to what was expected from their homology domains (section 1). The proteins C13orf1, RCBTB1, RFP2 and KPNA3 all presented mainly cytoplasmic patterns. The degree of homogeneous cytoplasmic staining varied among the proteins, with C13orf1 (Figure 3.19) and KPNA3 (Figure 3.22) showing the most homogenous distribution with only few aggregates (Figure 3.19.A). The other two cytoplasmic proteins, RCBTB1 and RFP2, both presented strong aggregation patterns, but different from each other (Figures 3.20 and 3.21 respectively).

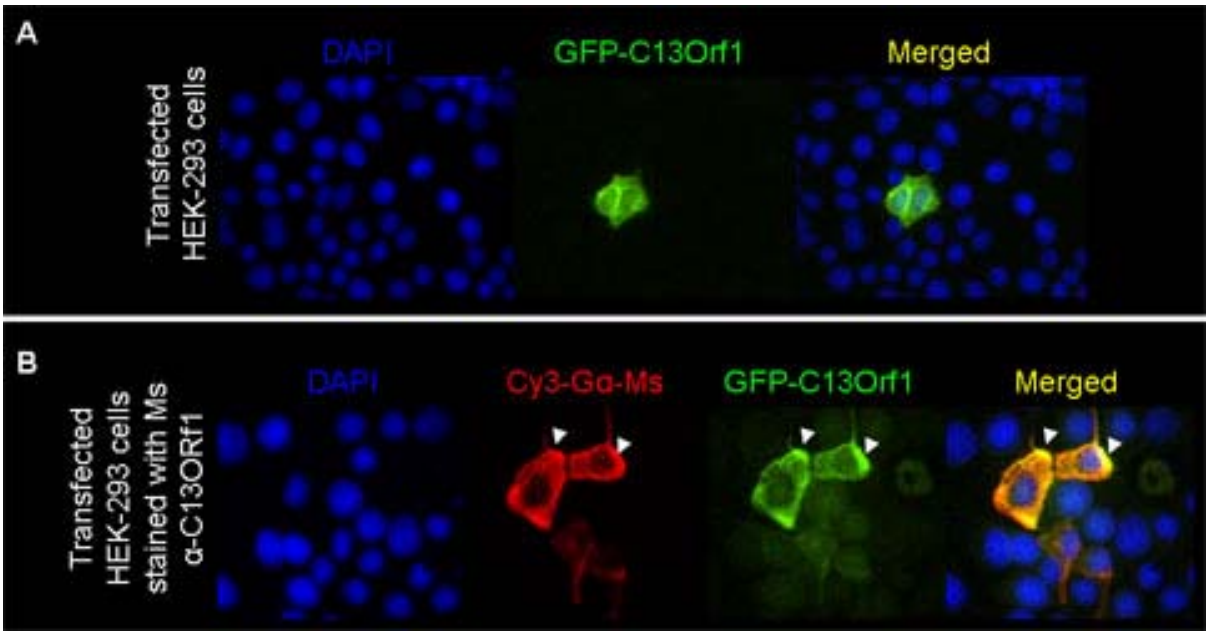


Figure 3.19 Localization of over-expressed GFP-C13orf1 in human culture cells. HEK-293 cells were transfected with a vector containing the fusion protein GFP-C13orf1. Transfected cells were fixed and prepared for microscopy (A) or immunofluorescence (B). **A.** Microscopy of GFP-C13orf1 (green) and DAPI (blue) localizes GFP-C13orf1 mainly in the cytoplasm. **B.** Immunofluorescence of GFP-C13orf1 (green) with mouse specific antibody against C13orf1 (red=Ms α-C13orf1) co-localizes both signals (yellow=merged) also mainly in the cytoplasm showing some aggregates (white arrow-heads). Cy3-Gα-Ms=Cy3-labelled Goat-anti-mouse.

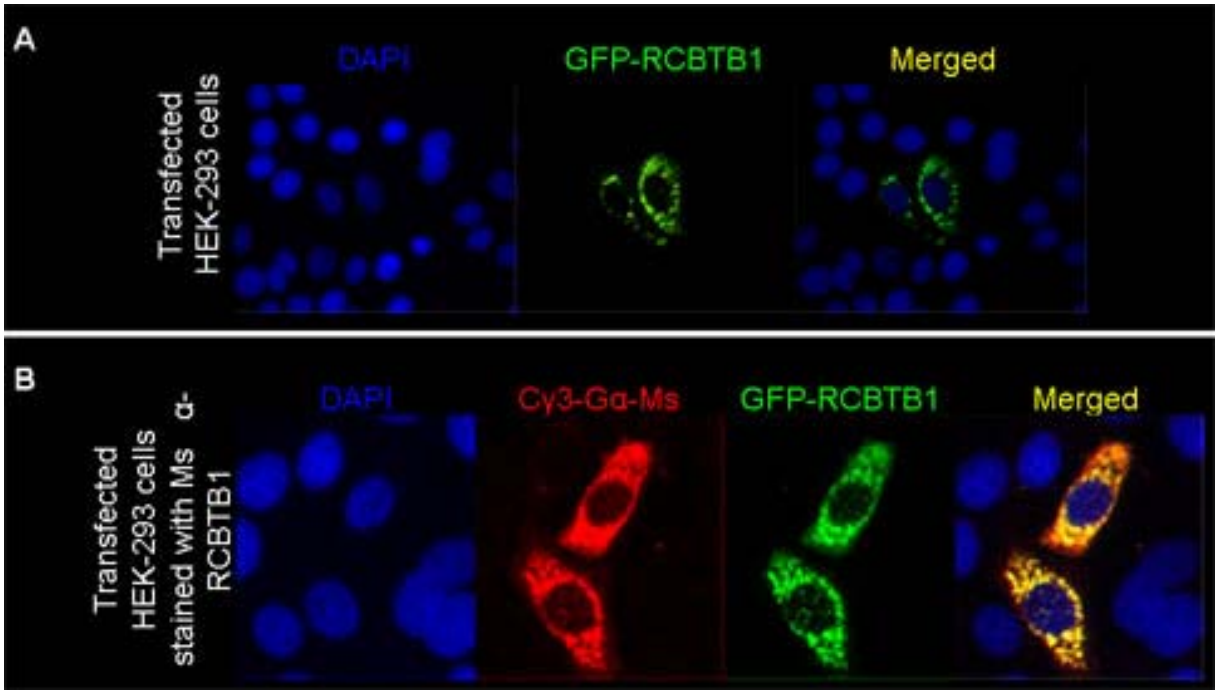


Figure 3.20 Localization of over-expressed GFP-RCBTB1 in human culture cells. HEK-293 cells were transfected with a vector containing the fusion protein GFP-RCBTB1. Transfected cells were fixed and prepared for microscopy (A) or immunofluorescence (B). **A.** Microscopy of the GFP-RCBTB1 (green) and DNA stained with DAPI (blue) localize GFP-RCBTB1 mainly in punctuate aggregates in the cytoplasm. **B.** Immunofluorescence of GFP-RCBTB1 (green) with mouse specific antibody against RCBTB1 (red=Ms α-RCBTB1) co-localize both signals (yellow=merged) also mainly in the cytoplasm, in punctuate aggregates. Cy3-Gα-Ms=Cy3-labelled Goat-anti-mouse.

Interestingly, a very similar aggregation pattern for RFP2 corresponding to the endoplasmic reticulum (ER) was described by others (Lerner et al. 2007), indicating that the similar pattern presented here most probably corresponds to the ER as well.

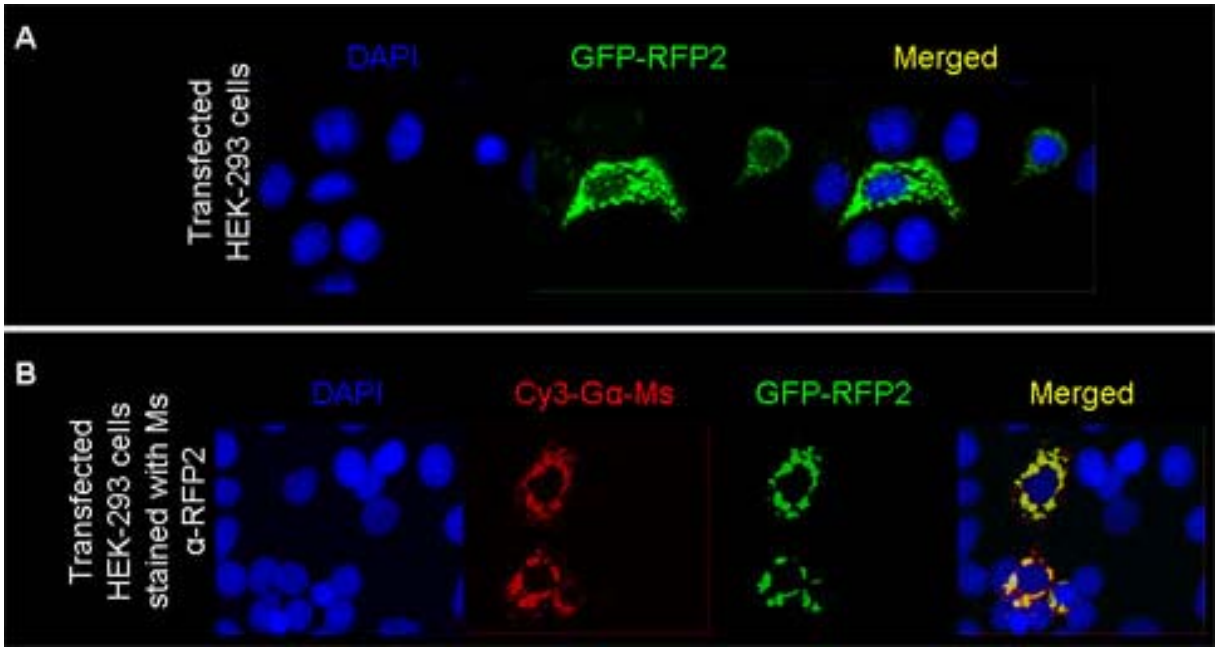


Figure 3.21 Localization of over-expressed GFP-RFP2 in human culture cells. HEK-293 cells were transfected with a vector containing the fusion protein GFP-RFP2. Transfected cells were fixed and prepared for microscopy (A) or immunofluorescence (B). **A.** GFP-RFP2 (green) transfected cells and DAPI-stained DNA (blue) localizes the protein in aggregates in the cytoplasm. This cytoplasmic pattern resembles that of the endoplasmic reticulum (ER). **B.** Immunofluorescence staining of GFP-RFP2 (green) with mouse specific antibody against RFP2 (red=Ms α-RFP2) also co-localize both signals (yellow=merged) mainly in the cytoplasm, in aggregates resembling the ER as well. Cy3-Gα-Ms=Cy3-labelled Goat-anti-mouse.

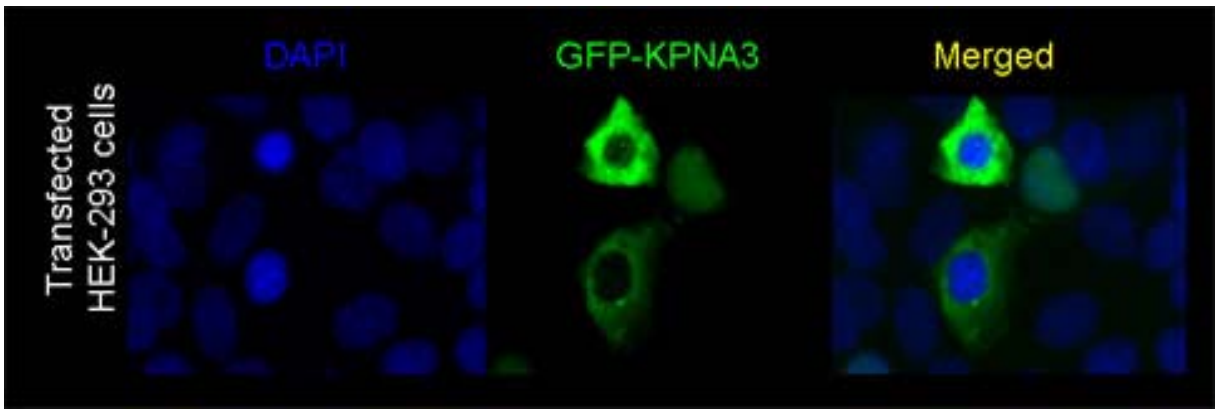


Figure 3.22 Localization of over-expressed GFP-KPNA3 in human culture cells. HEK-293 cells were transfected with a vector containing the fusion protein GFP-KPNA3. Transfected cells were fixed and prepared for microscopy. GFP-KPNA3 fusion protein (green) and DAPI-stained DNA (blue) localize KPNA3 mainly in the cytoplasm, in a homogeneous pattern.

In contrast to the other proteins, the homology domains of SETDB2 would predict a nuclear protein. Microscopy and immunohistochemistry of the over-expressed protein

revealed indeed that, in most of the cells analyzed, the SETDB2 signal is in the nucleus (Figure 3.23). The protein shows a more or less homogeneous distribution in the nucleus that seems to correlate with the DAPI stained DNA, being absent from the nucleoli (Figures 3.23.A and B). Interestingly, in some cells (less than 1%) an aggregation pattern in the cytoplasm was observed (Figure 3.23.C). In these cases, the protein possibly aggregates for unknown reasons, being then unable to re-localize from its abnormal distribution into the nucleus. Another change in the general distribution of SETDB2 was observed in dividing cells (Figure 3.23.D). It was seen that when nuclei dissolve for cell division, the protein shows a diffused distribution in the cytoplasm of the dividing cells and is excluded from the DNA. During cell division, SETDB2 probably dissociates from its original locations on the DNA of non-dividing cells, being in the cytoplasma until the daughter cells rebuild their cellular structure.

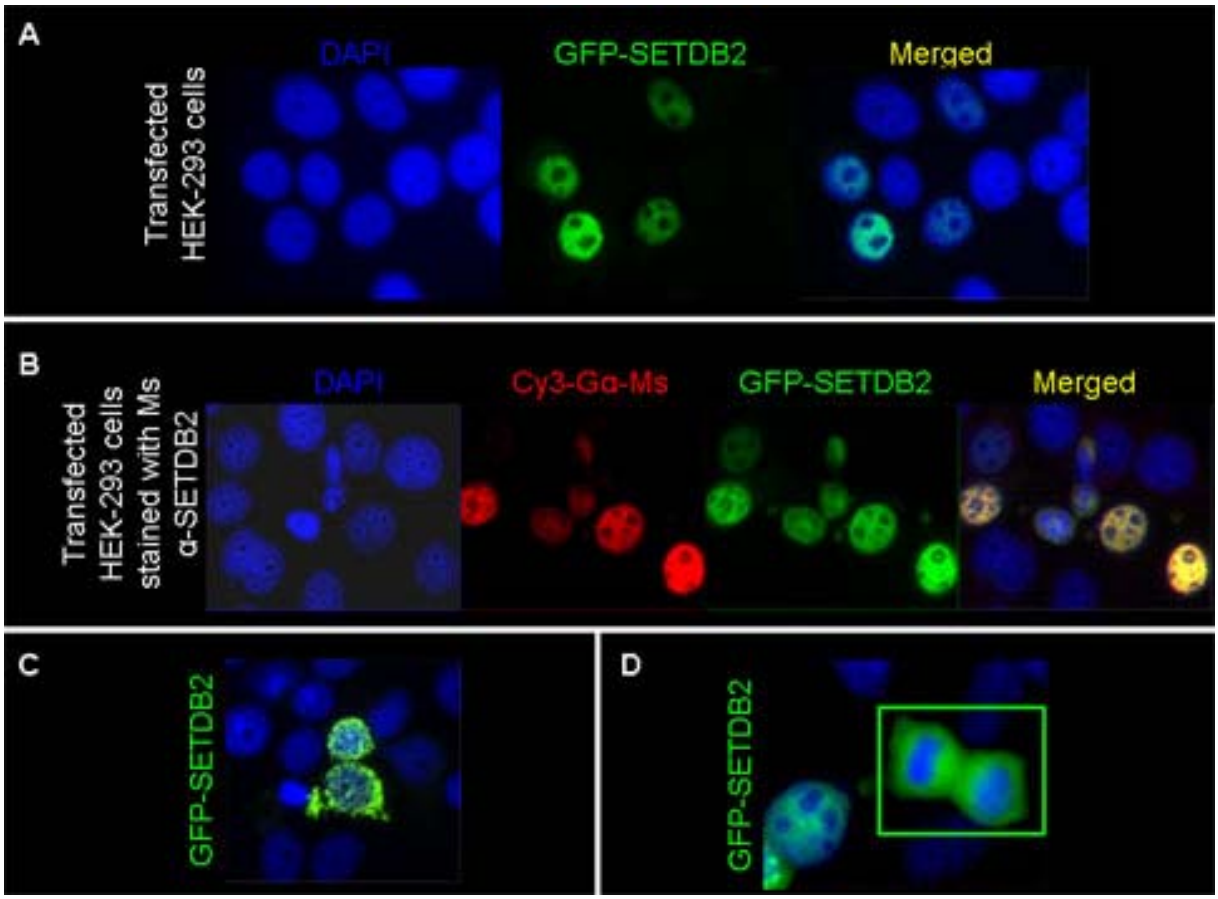


Figure 3.23 Localization of over-expressed GFP-SETDB2 in human culture cells. HEK-293 cells were transfected with a vector containing the fusion protein GFP-SETDB2. Transfected cells were fixed and prepared for microscopy (A, C and D) or immunofluorescence (B). **A.** GFP-SETDB2 (green) and DAPI stained DNA (blue) localize the protein homogenously in the nucleus but absent from the nucleoli. **B.** Immunofluorescence staining of GFP-SETDB2 (green) with mouse specific antibody against SETDB2 (red=Ms α-SETDB2) also co-localize both signals (yellow=merged) in the nucleus but absent from the nucleoli. **C.** Less than 1% of the cells show a different localization pattern for the protein in aggregates throughout the cell. **D.** In dividing cells (green square) SETDB2 is absent from the DNA and shows a diffused distribution throughout the cytoplasm. Cy3-Gα-Ms=Cy3-labelled Goat-anti-mouse.

3.3.2. Immuno-detection of endogenous proteins from 13q14

Interestingly, when immunofluorescence was performed with untreated fixed cells and endogenous protein was detected, the sub-cellular localization was in some cases different from that of the GFP-fusion proteins, indicating that, in some cases, the GFP-tag could alter the properties of the fusion protein, in this case its localization, as compared to that of the endogenous gene product. The localization of endogenous proteins was performed with the polyclonal antibodies and in two different cell lines: HEK-293 and Namalwa. The rationale was to compare endogenous expression of the candidate proteins in two cell lines of different histological origin, one of lymphocytic origin (Namalwa is derived from Burkitt's lymphoma) and one from a not blood-related origin (HEK-293 is an embryonic kidney cell line). The signals were rather weak for all of the proteins and apart from RFP2 (Figure 3.24) the localization was different between the two cell lines. For KPNA3 only data from Namalwa cells is available and it shows that endogenous protein localizes in the nucleoli of the cells (Figure 3.25), with some signal also being detected in other parts of the nucleus, mainly correlating to low-chromatin content regions as judged by DAPI staining. The endogenous C13orf1 (Figure 3.26) and RCBTB1 (Figure 3.27) both had a cytoplasmic homogeneous pattern in HEK-293, and more aggregates throughout the cell in Namalwa, indicating that the proteins might have specific functions in hematopoietic cells. There was also a variation detected in the amount of endogenous C13orf1 in HEK-293 cells (Figure 3.26.A). This could reflect differences in the protein expression depending on cellular status, e.g. cell cycle stage, and on the cell type, as this variation is not observed in Namalwa. Another protein showing a difference in the localization between the two cell lines was SETDB2 (Figure 3.28). In HEK-293 there seems to be very little endogenous SETDB2 protein (Figure 3.28.A) and it forms some aggregates in the cytoplasm surrounding the nucleus. In contrast, Namalwa cells present SETDB2 in the nucleus also in aggregates (Figure 3.28.B). Its localization in the nucleus of Namalwa cells suggests that SETDB2 (a predicted histone methyl transferase) could be active in these cells. The difference in the localization between the two cell lines for these genes could point to a blood-cell-specific function and could be related to an involvement in the development of B-CLL. Alternatively, the differences could also be derived from alterations suffered during the process of fixation. The later could also be the reason for the KPNA3 signal, although the highly specific nucleolar

pattern argues more in favor of a function-related localization. The specific localization might as well be related to KPNA3 functions in hematopoietic cells, and could also play a role in the development of malignancy.

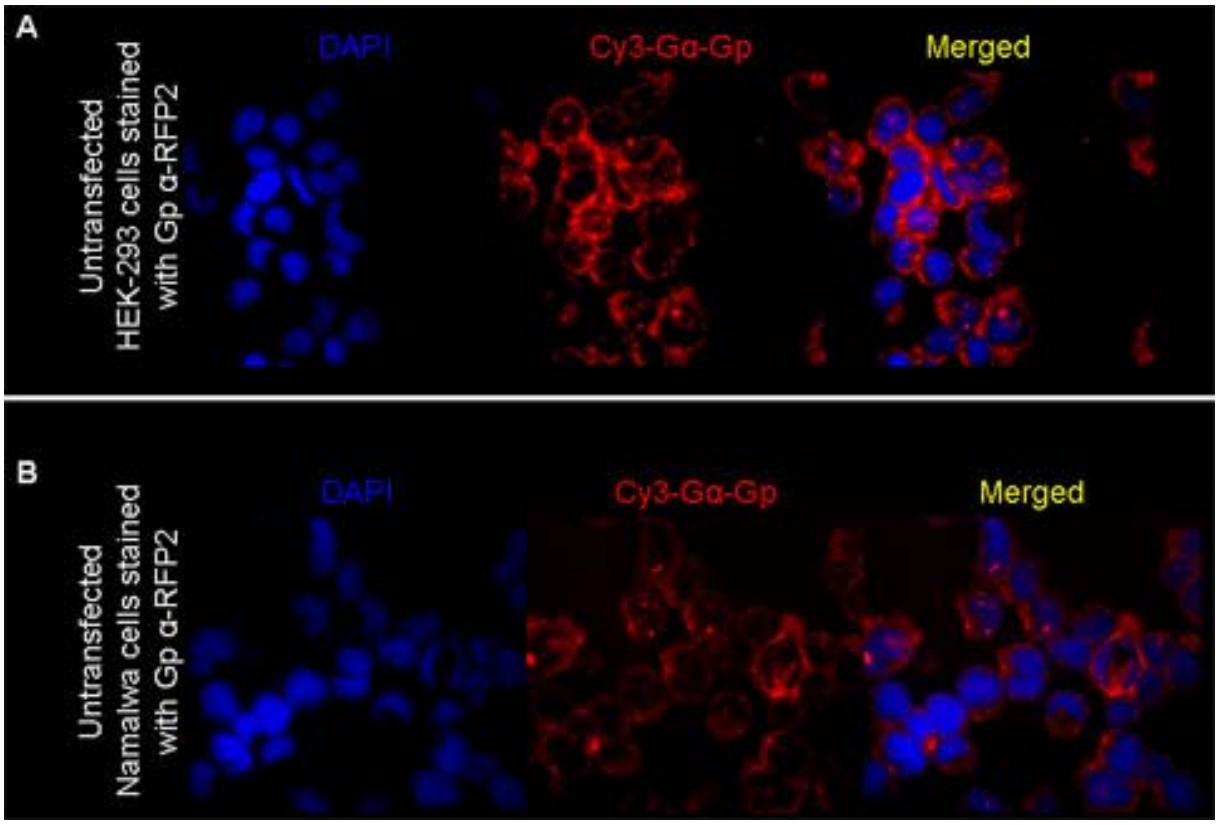


Figure 3.24 Localization of RFP2 in HEK-293 and Namalwa cells. Cells were fixed and stained for immunofluorescence with guinea pig anti-RFP2 (Gp α-RFP2) and detected with goat-anti-guinea pig (Cy3-Gα-Gp) (red). DNA was stained with DAPI (blue). **A.** HEK-293 cells show signal in the cytoplasm, specifically in a pattern resembling that of the endoplasmic reticulum (ER). **B.** Namalwa cells show signal mainly in aggregates in a very similar pattern than in A only weaker.

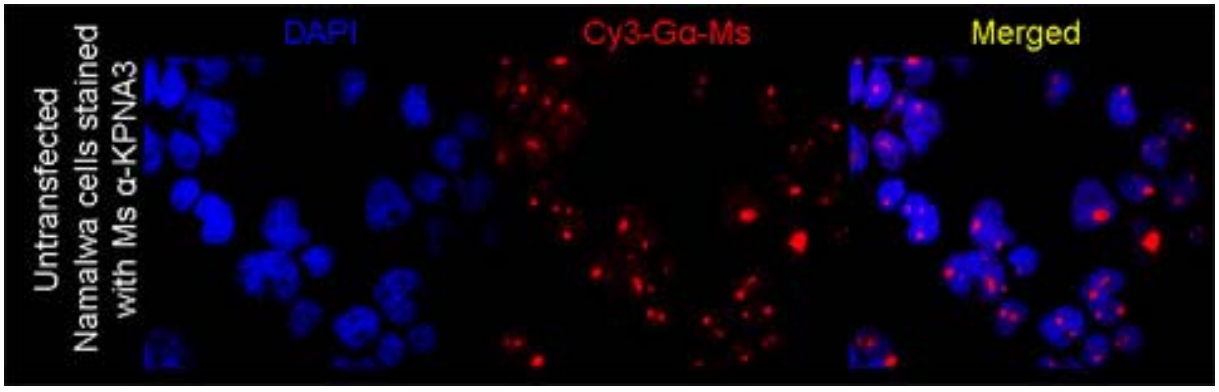


Figure 3.25 Localization of KPNA3 in Namalwa cells. Cells were fixed and stained for immunofluorescence with guinea pig anti-KPNA3 (Gp α-KPNA3) and detected with goat-anti-mouse (Cy3-Gα-Ms) (red). DNA was stained with DAPI (blue). In the merged picture the protein is mainly in aggregates in the nucleoli.

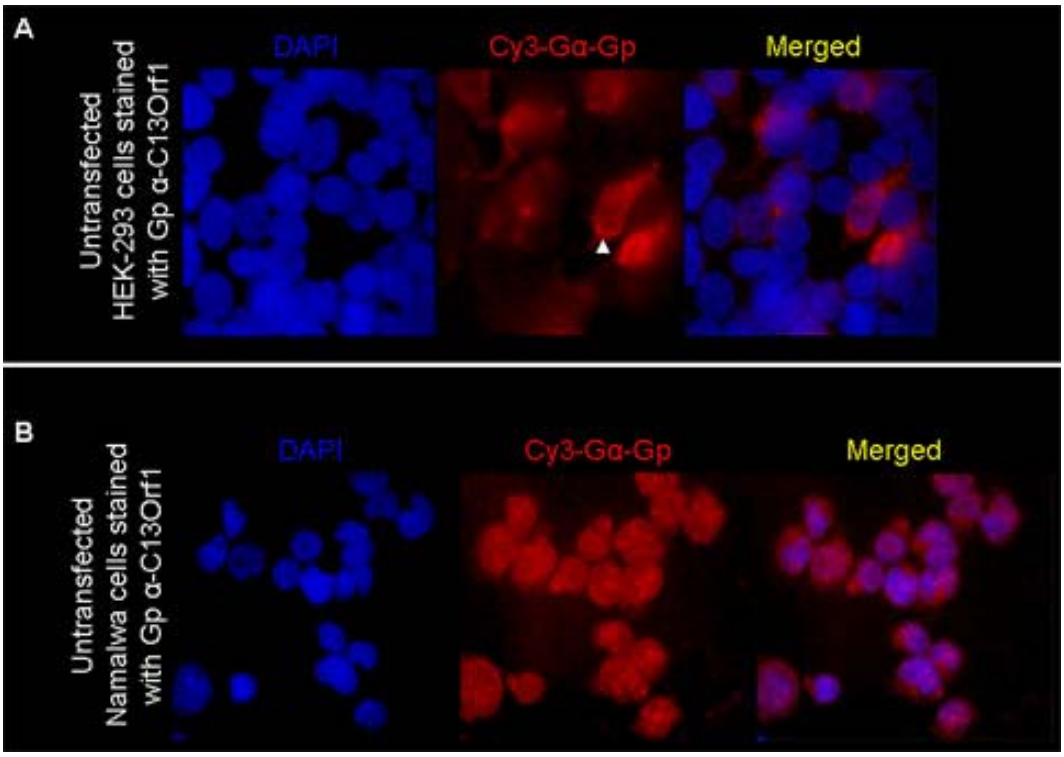


Figure 3.26 Localization of C13orf1 in HEK-293 and Namalwa cells. Cells were fixed and labeled for immunofluorescence with guinea pig anti-C13orf1 (Gp α-C13orf1) and detected with goat-anti-guinea pig (Cy3-Gα-Gp) (red). DNA was stained with DAPI (blue). **A.** HEK-293 cells show signal throughout the cell with some aggregates (white arrow-head). **B.** Namalwa cells show a homogeneous signal.

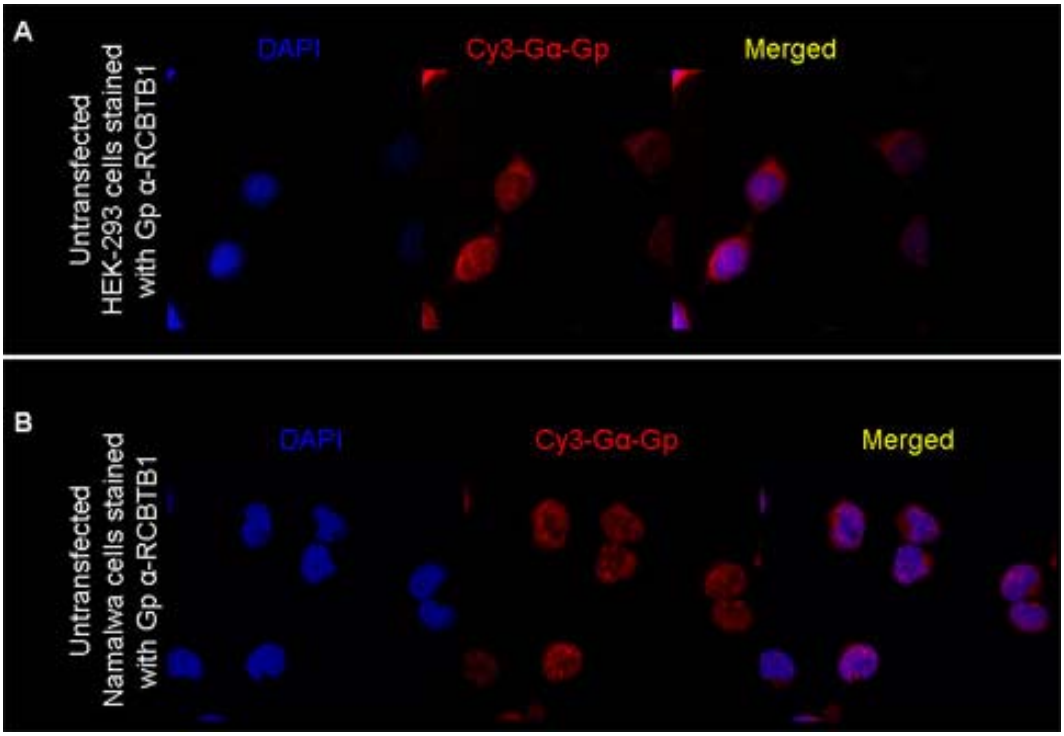


Figure 3.27 Localization of RCBTB1 in HEK-293 and Namalwa cells. Cells were fixed and stained for immunofluorescence with guinea pig anti-RCBTB1 (Gp α-RCBTB1) and detected with goat-anti-guinea pig (Cy3-Gα-Gp) (red). DNA was stained with DAPI (blue). **A.** HEK-293 cells shown signal mainly in the cytoplasm near the nucleus. **B.** Namalwa cells show signal throughout the cell in protein aggregates.

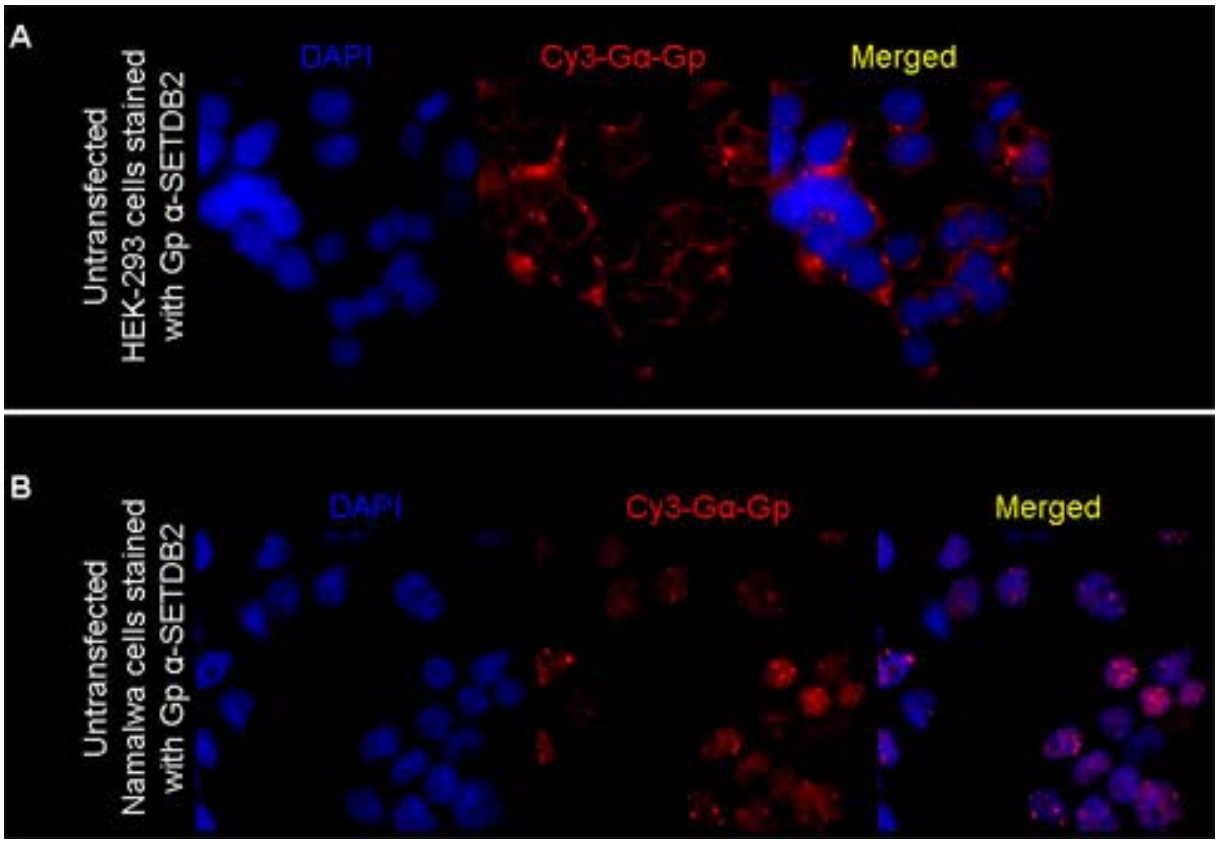


Figure 3.28 Localization of SETDB2 in HEK-293 and Namalwa cells. Cells were fixed and stained for immunofluorescence with guinea pig anti-SETDB2 (Gp α-SETDB2) and detected with goat-anti-guinea pig (Cy3-Gα-Gp) (red). DNA was stained with DAPI (blue). **A.** HEK-293 cells show signal mainly in the cytoplasm with some aggregates. **B.** Namalwa cells show signal mainly in the nucleus in aggregates.

To summarize, the 13q14 proteins showed different interesting intracellular localization, most of which could be related to their postulated functions, based on sequence homology. C13orf1, which has homology to the cytoplasmic interaction domain of receptors, is in the cytoplasm, and could function there via protein-protein interactions to bring interaction partners in close molecular proximity. Similarly, other proteins containing this domain, like the ryanodine receptors, are involved in dimerization of receptors (Lynda M. Blayney et al. 2004). RCBTB1 shows a patched-pattern in the cytoplasm and is proposed to be a guanine nucleotide exchange factor. The pattern shown here could indicate that it performs this function at determined cellular locations represented by the patches detected by fluorescence microscopy. The postulated function of SETDB2 as a histone methyl transferase fits well with its localization in the nucleus. For RFP2, recent publications postulate that it is a component of the ER-associated protein degradation machinery (Lerner et al. 2007),

and the localization data relates to the localization seen in that study. KPNA3 is known to be a nuclear import protein (Fagerlund et al. 2005), and the antibodies detected signal mainly in the nucleus, more precisely in the nucleoli and in low chromatin-containing regions, what could be indicative of specific regulation associated with lymphocytic cells (since it was observed in Namalwa). The differences in the localization of genes expressed as fusion proteins with GFP as compared to their endogenous-counterparts, could be due to the GFP moiety (Simpson et al. 2000) but suggested that if localization is affected by this non-endogenous tag, their functions might be as well. These, together with the weak detection results of the antibodies for endogenous proteins, made the yeast-two-hybrid (Y2H) screen the next approach to find interaction partners.

3.4. Identification of interaction partners of 13q14 proteins

To further characterize the 13q14 proteins functionally, a Y2H screen was performed to identify some of their interaction partners and link the candidate proteins with specific pathways and cellular functions. The Y2H screen was performed with the proteins of interest from 13q14 against more than 5000 human proteins from a brain library (Stelzl et al. 2005). The Y2H screen is based on the ability of a transcription factor (in this case GAL4) to activate transcription even when its DNA-binding domain (BD) and its activation domain (AD) are split, but in close proximity to each other. Each of these two domains is then expressed as a fusion protein: the BD fused to the so called “bait” protein and the AD fused to the so-called “prey” protein. The constructs contain reporter genes under the control of the gal-promoter, to which the GAL4 transcription factor (formed by proximity of domains BD and AD) bind. Upon interaction of the bait and prey proteins, the reconstituted transcription factor activates expression of the reporter genes. Positive protein-protein interactions (PPI) are selected based on expression of these reporters. For the screening with 13q14 proteins, the full lengths and different fragments of the proteins (selected to contain parts of the proteins homologous with described functional or structural domains) were used as baits (Appendix 5). Selected results of the Y2H screen were validated by pull-down experiments of over-expressed proteins in human cultured cells and are presented in section 3.4.2.

3.4.1. Interaction partners for 13q14 candidate proteins

For each of the proteins different constructs were designed with the help of the software Vector-NTI (Invitrogen): 5 constructs for C13orf1, 17 constructs for RCBTB1, 15 constructs for SETDB2, 13 constructs for RFP2 and 11 constructs for KPNA3 (Appendix 5). First they were tested with the empty prey-vector to eliminate auto-activating constructs for screening. At the same time they were tested with the same fragments in the prey-vector to see if interactions among 13q14 proteins could be detected (section 3.4.1.1). The library of more than 5000 prey proteins was screened for interaction partners with the not auto-activating 13q14 constructs, and results are presented in section 3.4.1.2. A schema of the screening process is depicted in figure 3.29.

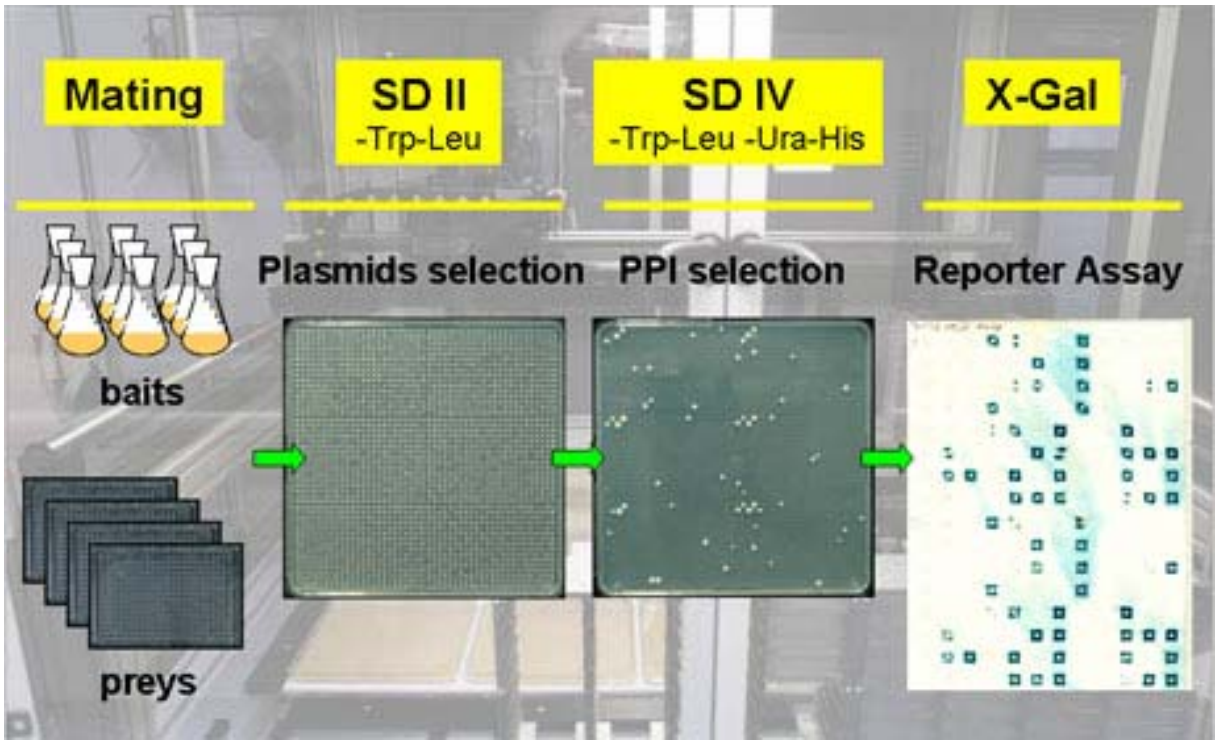


Figure 3.29 Schema of the yeast-two-hybrid (Y2H) automated screening process. Baits and preys are cloned into vectors containing reporter genes and transformed into mating-compatible yeast strains. After mating the cells are spotted onto plates containing selective medium lacking two essential amino acids (SDII) where only cells containing both plasmids —bait and prey— can grow. Grown cells are transferred to plates containing selective medium lacking four essential amino acids (SDIV) where bait-prey interactions are selected for. Each plasmid contains the reporter gene LacZ coding for the enzyme β-Gal for a last selection by a colorimetric reaction. If a chromogenic substrate is applied, colonies will turn blue only if interaction occurred. Trp=triptophane, Leu=leucine, Ura=uracil, His=histidine, X-Gal=5-bromo-4-chloro-3-indolyl-β-Dgalactopyranoside, PPI=protein-protein interaction.

3.4.1.1. Autoactivation test for 13q14 constructs hints at interaction among proteins of 13q14

For the five 13q14 proteins of interest, 60 baits and 55 preys were transformed and mated with the empty vectors as well as with each other. Interactions between candidate proteins of 13q14 would reveal that 13q14 proteins might form part of macromolecular complexes, linking DNA organization of 13q14 genes to their protein function (Mertens et al. 2002). To find such interactions would be very interesting as this phenomenon, known as functional clustering, is not very commonly found in the mammalian genome (Blumenthal 1998, Lercher et al. 2002, Sproul et al. 2005). Interestingly, interactions between proteins from 13q14 with each other and with themselves could be detected (Table 3.30). These interactions were only found when construct A was used as the bait and construct B as the prey, but not in the reverse experiment. This could be due to conformational limitations that would for example allow interaction with one combination of tag to protein but not with the other. To date validation experiments were unsuccessful due to difficulties with co-expression of constructs. Therefore, to proof the interactions the data still needs to be validated with other methods.

Possible interactions of 13q14 genes with each other

	Results showing dimerisation	
Baits	**Interaction**	**Fragments**
C13orf1	C13orf1-C13orf1	F3-F3
RCBTB1	RCBTB1-RCBTB1	F8-F8
	Results showing heterologous interactions	
Baits	**Interaction**	**Fragments**
C13orf1	C13orf1-RCBTB1	F3-F5, F3-F7
C13orf1	C13orf1-SETDB2	F3-F8, F3-F15
C13orf1	C13orf1-RFP2	F3-F2
RCBTB1	RCBTB1-SETDB2	F6-F8, F5-F9
SETDB2	SETDB2-RFP2	F10-F2

Table 3.30 Summary of the results from the Y2H screen for 13q14 genes with each other. The "baits" column indicates the gene fragments used as bait for the mating. The "interaction" column indicates the genes involved in the interaction. The "fragments" column indicates the fragments found in the interaction, being the first fragment the one corresponding to the bait, and the second to the prey. If more than one interaction was found, the second interaction is also shown separated by commas from the first one. Interactions of protein fragments with themselves are depicted under "results showing dimerisation" and interactions between different proteins are indicated under "results showing heterologous interactions".

From the result of the interaction between the baits and the empty prey-vector, 25 out of the 60 baits (42%) were removed for being auto-activating (Figure 3.31) and the rest was used for screening with the prey library.

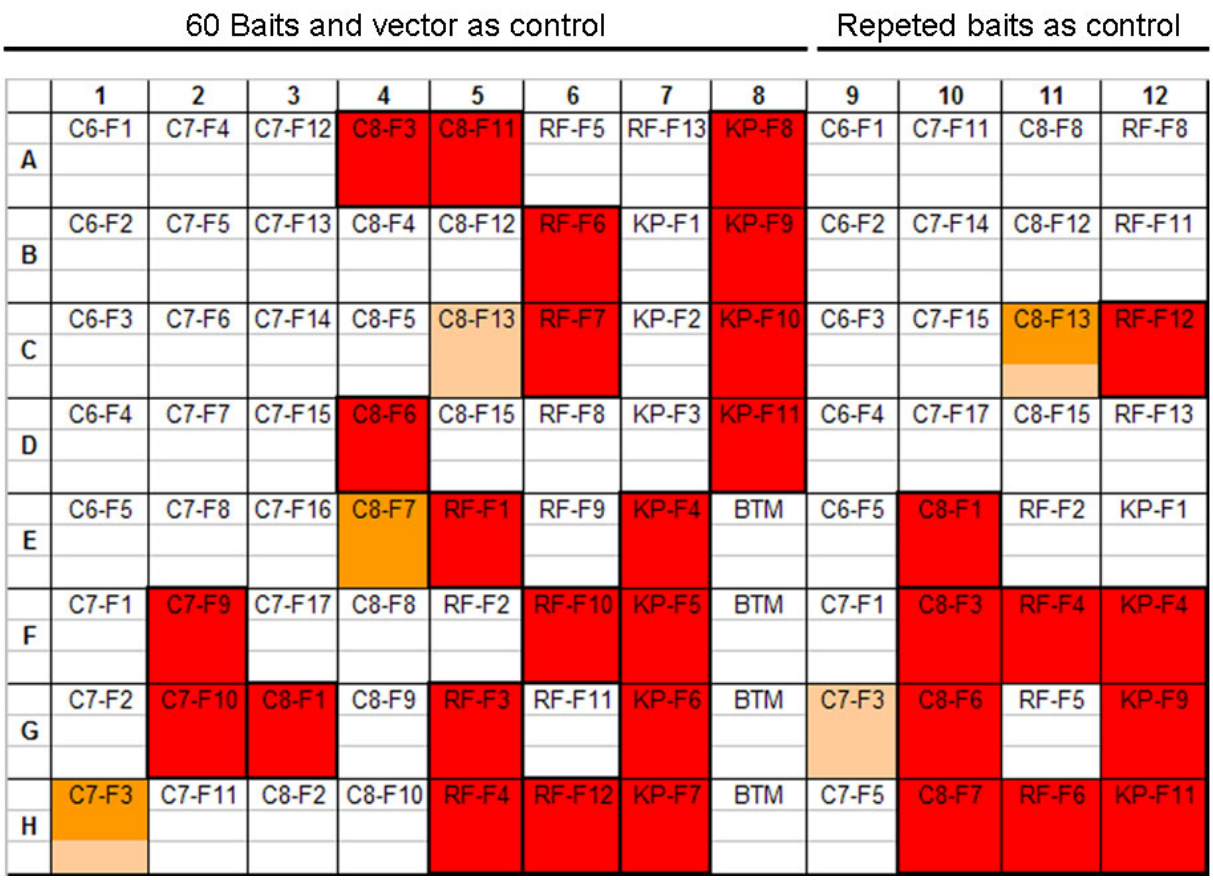


Figure 3.31 Plate showing the positions of the baits for mating indicating auto-active baits. Auto-active baits (coloured) gave positive growth in selective media when mated with the prey-vector alone, and were therefore dismissed for further analysis. The mating was performed three times and the colour indicates: Red=auto-active in triplicates, orange=auto-active in duplicates, salmon=auto-active in single, white=no auto-active. C6=C13orf1, C7=RCBTB1, C8=SETDB2, RF=RFP2, KP=KPNA3, F=fragment number (fragment information is given in Appendix 5).

3.4.1.2. Yeast-two-hybrid screen with 13q14 constructs reveal 53 interaction partners

The not auto-activating 35 baits previously identified were used to screen a library containing 5632 human sequences cloned in the prey vector (Stelzl et al. 2005). In total 196 high confidence interactions between 58 proteins were detected. Each interaction was performed in triplicate, so interactions showing positive growth for all three replicas were considered of higher confidence, but very rare in our screening (0.56%). More commonly interactions were detected in duplicates (8.38%) or 1 time only (91.06%). All of the interactions were positive for the two growth reporter genes (His and Ura), and most of them also for the LacZ colorimetric reporter. These interactions are summarized in Appendix 7, and in order to confirm these results interactions were further validated with biochemical approaches.

3.4.2. Validated interaction partners from the yeast-two-hybrid screen

For the validation of the constructs, the full length baits and some of the fragments were cloned in vectors containing the protein-A (PA) tag. The preys were cloned in vectors containing a different tag: the nonapeptide YPYDVPDYA from the human influenza haemagglutinin protein ("HA"), or obtained from commercial providers with the FLAG peptide patented by Sigma-Aldrich (used as 3xFLAG tag). The advantage of the PA-tag was that pull-down experiments could be performed, facilitating the assay of many interactions (PA-tagged proteins were directly precipitated with a matrix coated with IgG). For the pull-down experiments, human cell lines were transfected with both plasmids and after binding and washing, bound proteins were directly eluted in loading buffer and analyzed by Western blot analysis. Either protein was detected with respective antibodies against the tags: the antibody against PA to check precipitation, and the antibodies against HA or FLAG to detect the co-precipitated protein. Despite extensive washing of the pulled-down proteins, a weak band was almost always detected in the negative controls, which was in those cases considered as the background level. Interactions were considered as positive, if more precipitated protein as the background could be detected.

		13q14 genes expressed as PA-fusion proteins									
		C13orf1		**RCBTB1**		**SETDB2**		**KPNA3**		**RFP2**	
Method →		Y2H	P-D	Y2H	P-D	Y2H	P-D	Y2H	P-D	Y2H	P-D
Binding partners expressed as HA-fusion proteins	**eEF1A**	+	W	+	+	-	+	-	-	-	NT
	eEF1G	-	+	-	+	+	+	+	W	-	?
	EED	-	-	+	+	-	+	+	+	-	+
	CRMP1	+	+	+	+	+	+	-	-	-	-
	RXRA	+	§	+	§	+	§	+	§	-	§
	GNB2	-	?	+	+	+	+	+	-	-	-
	C11Orf2	-	-	-	NT	+	+	-	NT	-	NT
	RIPAIN	+	+	+	+	+	+	-	-	-	-

Table 3.32 Summary of validated protein interactions from the Y2H screen. The columns indicate the 13q14 candidate genes used for the interaction, as well as the methods that detected the interaction (Y2H=yeast-two-hybrid, P-D=pull-down), and the rows indicate the 8 binding partners validated. W=weak, +=positive interaction, -=negative interaction, NT=not tested, ?=interaction cannot be determined due to cross-reaction, §=interaction cannot be determined due to unspecific bands.

From the 53 interaction partners of the 13q14 candidate proteins, 19 target genes with a HA-tag, and 7 clones with the 3xFLAG tag could be expressed (49%). Of

these, 14 genes could be co-expressed in HEK-293 cells with the 13q14 constructs and conclusive results were obtained for 8 binding partners (Table 3.32) and are described in the next sections.

3.4.2.1. eEF1A and/or eEF1G are found as interaction partners for four of the 13q14 candidate genes tested

The translation elongation factors eEF1A and eEF1G (Thornton et al. 2003) were found to interact with C13orf1 and RCBTB1, and SETDB2 and KPNA3 respectively. Pull-down and Western blot analysis for eEF1A (Figure 3.33) and eEF1G (Figure 3.34) revealed more interactions than detected in the Y2H screen alone. It was found that C13orf1 and RCBTB1 precipitate not only eEF1A but also eEF1G. As expected, from the two fragments tested for RCBTB1-eEF1A, the fragment positive in the Y2H screen show the strongest binding ($RCBTB1_{301-467}$). The interactions found in the pull-down experiments but not in the Y2H screen suggest that some interactions might be jeopardized in the Y2H screen by the poor expression of full length proteins, as indicated by the few interactions found with full length proteins (Appendix 7).

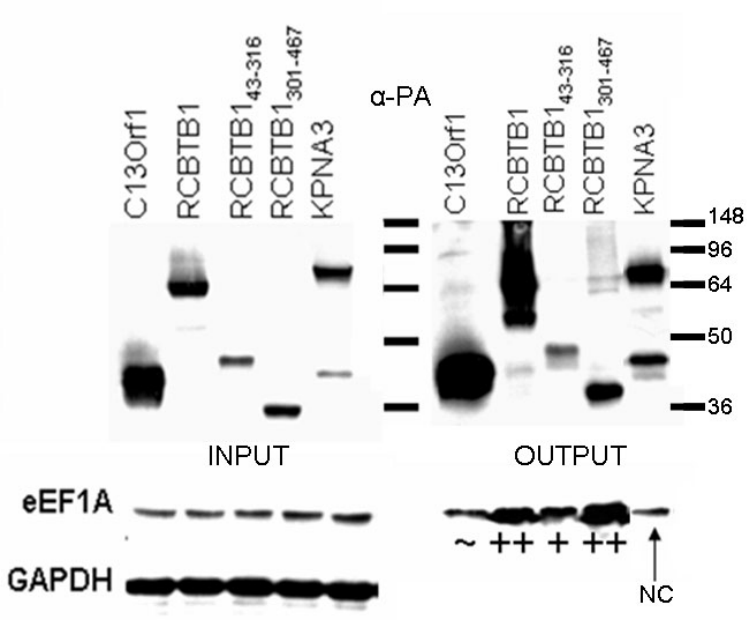


Figure 3.33 Western blot analysis of the interaction between 13q14 candidate proteins and eEF1A. Pull-down was performed for full-length candidate proteins and two RCBTB1 fragments as indicated. Baits for the pull-down experiments were detected with anti-Protein A antibody, and eEF1A was detected with anti-HA antibody. As loading control GAPDH antibodies were used. eEF1A was co-precipitated with RCBTB1 and its fragments. For C13orf1, eEF1A was rather weakly precipitated and KPNA3 set the threshold for background as negative control (NC). $RCBTB1_{43-316}$=RCBTB1 fragment containing amino acids 43 to 316, $RCBTB1_{301-467}$=RCBTB1 fragment containing amino acids 301 to 467, ~=weak interaction, +=positive interaction, ++=strong interaction.

The positive result from the Y2H screen for SETDB2-eEF1G was validated with more precipitated eEF1G protein than the other interactions, while KPNA3-eEF1G shows the weakest interaction although the interaction was also detected in the Y2H screen. Some results like that of RFP2-eEF1G cannot be analyzed because both recombinant proteins are of very similar size and the PA-tag is known to cross-react with secondary antibodies as it binds to their heavy chain. Because of this cross-reaction, results for PA-fusion proteins with HA-fusion proteins of similar size could not be analyzed with certainty and were left out.

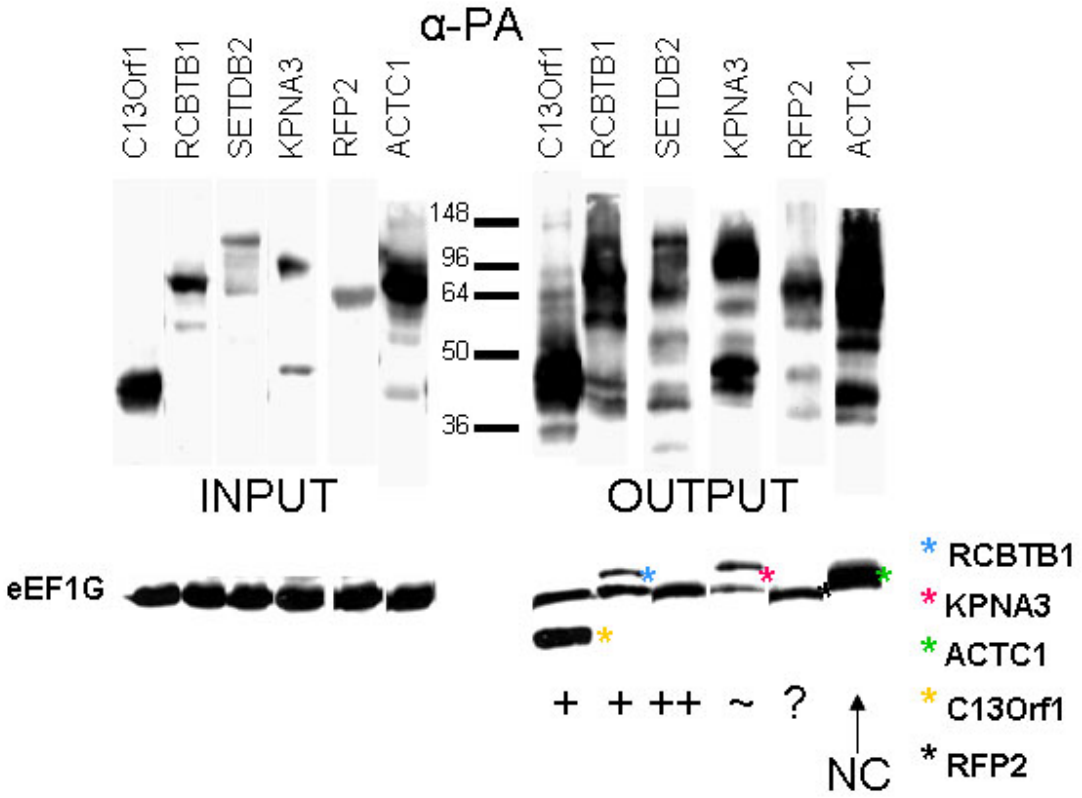


Figure 3.34 Western blot analysis of the interaction between 13q14 candidate proteins and eEF1G. Pull-down was performed for full-length candidate proteins and ACTC1 as negative control (NC). Baits for the pull-down experiments were detected with anti-Protein A antibody, and eEF1G was detected with anti-HA antibody. eEF1G was co-precipitated with C13orf1, RCBTB and SETDB2. KPNA3 was rather weak but they were all positive in comparison to the negative control ACTC1 were no band of the eEF1G-size was detected. RFP2 could not be analyzed as the sizes of eEF1G and RFP2 are very similar and the band seen cannot be assigned with certainty to any of the two proteins. Asterisks indicate the Protein A-fusion protein bands. ~=weak interaction, +=positive interaction, ++=strong interaction, ?=not analyzed.

3.4.2.2. EED is found as interaction partner of RCBTB1, SETDB2, KPNA3 and RFP2

The EED protein is an essential component of a polycomb group (PcG) complex involved in transcriptional repression during development (Richie et al. 2002). In the

Y2H screen it was found as an interaction partner of three 13q14 proteins and two of them, RCBTB1 and KPNA3 plus the fragment RCBTB1$_{301-467}$ showed positive results in the validation experiments (Figure 3.35). The RCBTB1$_{301-467}$ fragment showed weaker interaction than that of the full length protein indicating that the tertiary structure of the protein might be involved in the binding, and implying that the molecular context of the fragments tested in the screening can have an impact on the interaction with the target proteins. In contrast, the interaction of EED with C13orf1 found in Y2H screen could not be validated with pull-down experiments. Surprisingly, it was also found that SETDB2 and RFP2 interact with EED in the pull-down experiments, although they were negative in the Y2H screen. These results also indicate that some interactions might be concealed in the Y2H screen by the poor expression of full length proteins. It is interesting to note that for the negative controls (ACTC1 and the empty vector containing the PA-tag) some weak interactions were seen, with the strongest being the one with the empty vector. As washing was performed with stringent conditions, these weak bands detected with the negative controls were considered the background for this pull-down experiments set-up.

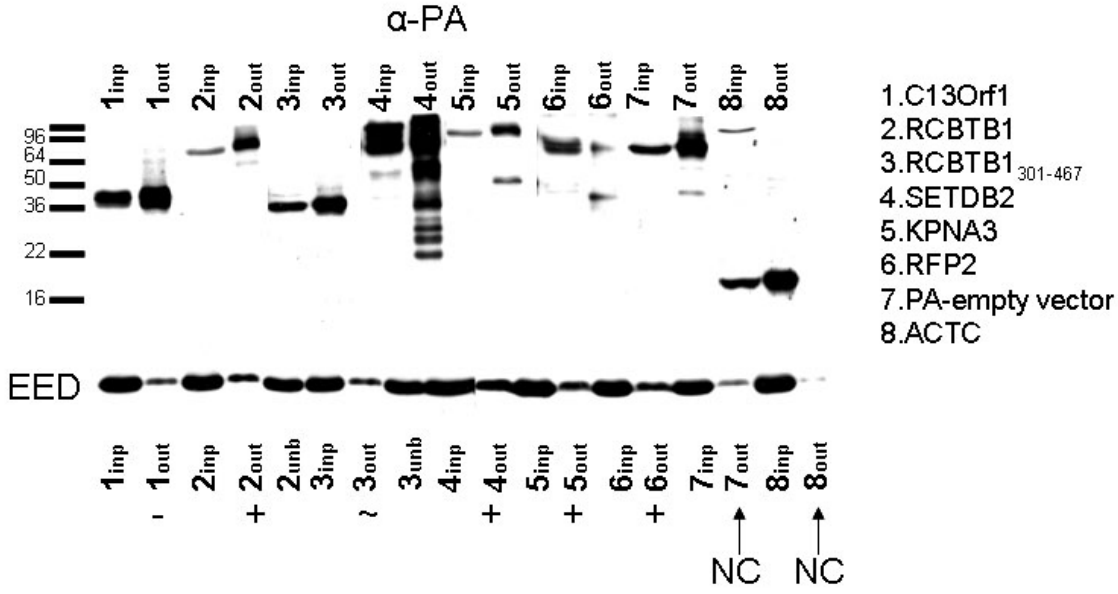


Figure 3.35 Western blot analysis of the interaction between 13q14 candidate proteins and EED. Pull-down was performed for full-length 13q14 candidate proteins (1-6), one fragment of RCBTB1 (3) and the negative controls ACTC1 (8) and the empty vector (7). Baits for the pull-down experiments were detected with anti-Protein A antibody, and EED was detected with anti-HA antibody. NC=negative control, ~=weak interaction, +=positive interaction, -=negative interaction.

3.4.2.3. CRMP-1 is found to bind to C13orf1, RCBTB1 and SETDB2

CRMP-1 is a member of a family of proteins (Collapsin Response Mediator response Protein) involved in intracellular signaling (Shih et al. 2001). In the Y2H screen

CRMP-1 was found as an interaction partner of C13orf1, RCBTB1 and SETDB2, all of which, together with the RCBTB1 C-terminal fragment $RCBTB1_{468-531}$, were validated by pull-down assays (Figure 3.36). In contrast, CRMP-1 was neither precipitated with the negative control (KPNA3) nor with RFP2 (data not shown). It should also be noted that for the positive pull-down experiments substantial amounts of protein can be seen in the unbound fractions, which could reflect a non-equivalent protein expression for the two fusion proteins being over-expressed, and/or sub-optimal buffer conditions.

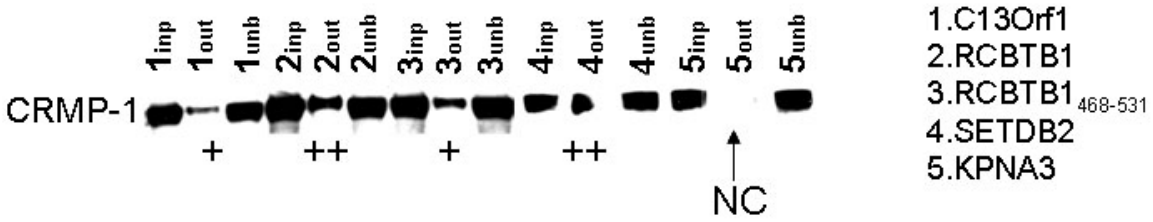


Figure 3.36 Western blot analysis of the interaction between 13q14 candidate proteins and CRMP-1. Pull-down was performed for full-length candidate proteins, one RCBTB1 fragment and KPNA3 as negative control (NC). CRMP-1 was detected with anti-HA antibody. CRMP-1 was co-precipitated with C13orf1 (1), RCBTB (2,3), and SETDB2 (4). KPNA3 (5) was completely negative. +=positive interaction, ++=strong interaction.

3.4.2.4. RXRA presents difficulties for validation

RXRA is a member of one of the families of retinoid acid (RA) receptors (Jones et al. 1993). It was found as interaction partner of C13orf1, RCBTB1, SETDB2 and KPNA3 in the Y2H screen. In Figure 3.37 it is shown that over-expression of RXRA resulted in production of two FLAG-positive bands in the cell extracts. The two bands could be related to posttranslational modifications, but to the best of our knowledge this phenomenon is not observed in other studies with RXRA. Therefore, although a highly interesting interaction partner, the validation could not be performed.

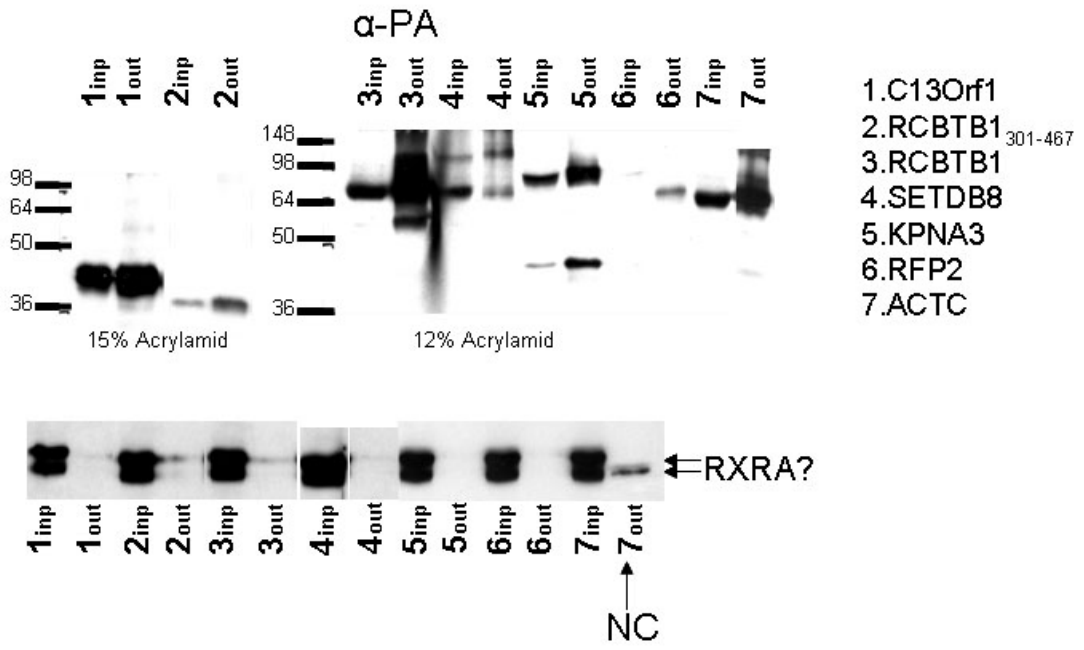


Figure 3.37 Western blot analysis of the interaction between 13q14 candidate proteins and RXRA. Pull-down was performed for full-length candidate proteins (1-6), one RCBTB1 fragment (2) and ACTC (7) as negative control (NC). RXRA was detected with anti-FLAG antibody, and two positive bands of similar size were detected, questioning the quality of the RXRA clone.

3.4.2.5. GNB2 interacts with RCBTB1 and SETDB2

GNB2 is a member of the large family of Guanine Nucleotide Binding proteins that are involved in signal transduction (Downes& Gautam 1999). In the Y2H screen, GNB2 was found as interaction partner of RCBTB1, SETDB2 and KPNA3, and they could be validated by pull-down experiments (Figure 3.38). The RCBTB1-full length protein showed stronger interaction than the fragment containing the C-terminal amino acids 301 to 467, both indicating that proper folding of the protein might be required for proper interaction and that the poor expression of full length protein in the Y2H screen might be masking positive results in the screening. For KPNA3 a weak interaction with GNB2 is seen compared to the negative controls ACTC and RFP2 (Figure 3.38.A). However, in order to show interaction more robustly, experimental conditions would need to be optimized. For C13orf1 no conclusive results can be presented, because it has a very similar size to GNB2, and as mentioned above, the proteins of similar sizes cannot be evaluated because of the unspecific binding of the PA-tag to antibodies (Figure 3.38.B).

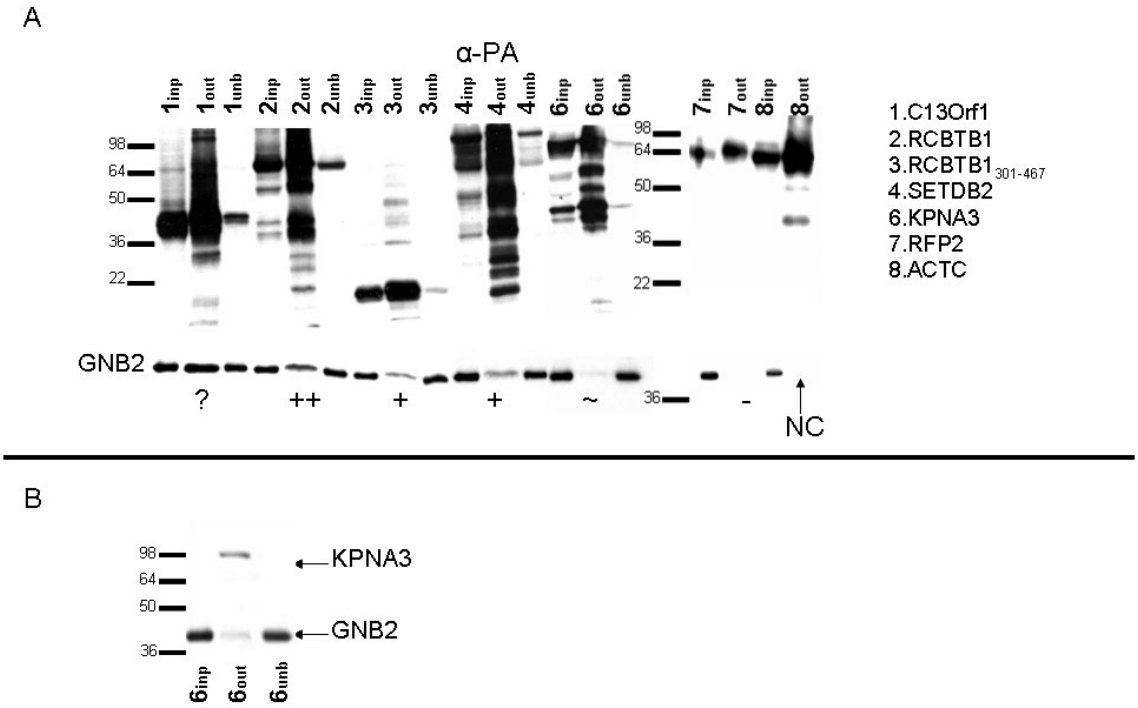


Figure 3.38 Western blot analysis of the interaction between 13q14 candidate proteins and GNB2. Pull-down was performed for full-length candidate proteins (1-7), one fragment of RCBTB1 (3) and ACTC1 (8) as negative control (NC). Baits for the pull-down experiments were detected with anti-Protein A antibody and GNB2 was detected with anti-HA antibody. **A.** Western blot analysis showing the results for GNB2 interaction with 13q14 proteins: GNB2 was co-precipitated with RCBTB and SETDB2, very weakly with KPNA3 and was completely negative both with RFP2 and the negative control. **B.** Western blot analysis with anti-HA antibody of KPNA3-GNB2 precipitation showing the band of GNB2 and the PA-KPNA3 band produced by unspecific binding of antibodies to the PA-tag. ~=weak interaction, +=positive interaction, ++=strong interaction, -=negative interaction, ?=not analyzable.

3.4.2.6. C11orf2 interacts with SETDB2

C11orf2 is a gene of unknown function described to be a nuclear protein (O'Brien et al. 2000). In the Y2H screen, C11orf2 was found only as interaction partner of SETDB2, and the data were validated with pull-down experiments. The Western blot analysis shows that C11orf2 could be precipitated with SETDB2 but not with the negative control C13orf1 (Figure 3.39). To date, no functional conclusions can be drawn from this data since C11orf2 is a protein of unknown function. Nevertheless it is interesting to note that SETDB2 is a predicted histone methyl transferase and that C11orf2 is predicted to be a nuclear protein as well.

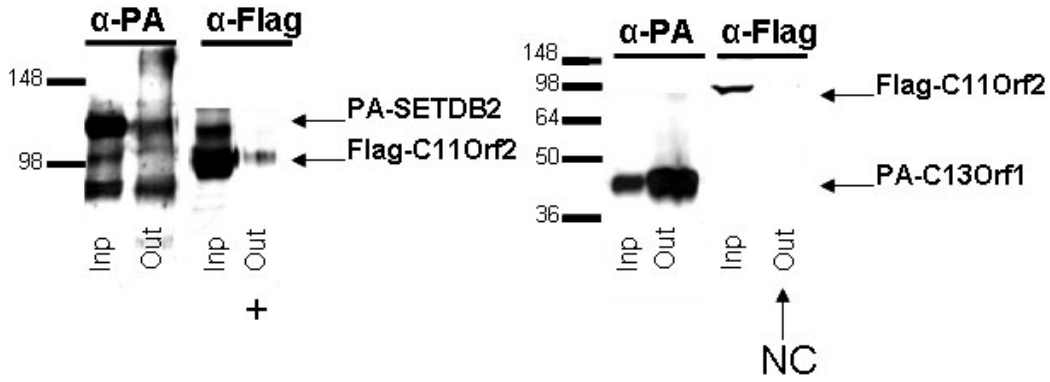


Figure 3.39 Western blot analysis of the interaction between SETDB2 and C11orf2. Pull-down was performed for full-length SETDB2 and C13orf1 was used as the negative control (NC). Bait proteins were detected with anti-PA antibody and C11orf2 was detected with anti-FLAG antibody. C11orf2 was co-precipitated with SETDB2 but not with the negative control C13orf1. +=positive interaction.

3.4.2.7. RPAIN interacts with C13orf1, RCBTB1 and SETDB2

The Replication Protein A Interacting protein (RPAIN) is a transport protein for RPA (Chen et al. 2005). It was found in the Y2H screen as interaction partner of C13orf1, RCBTB1 and SETDB2 and this data were validated with pull-down experiments (Figure 3.40). In the pull-down experiments there was an extra negative control, SETDB2-PA with TIFF-FLAG (the later an unrelated protein kindly provided by H. Bierhoff (DKFZ, Heidelberg)) to prove that PA proteins don't pull-down FLAG-tagged proteins unspecifically.

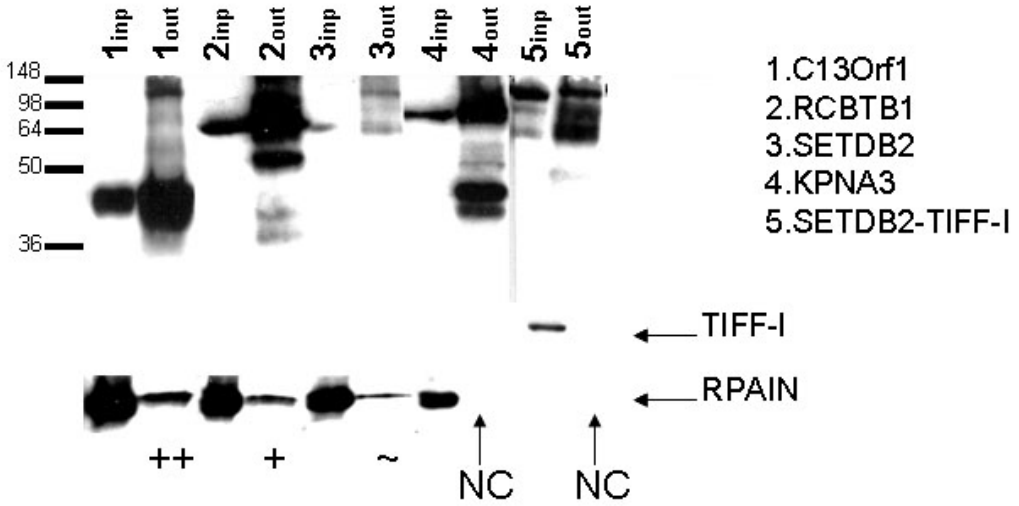


Figure 3.40 Western blot analysis of the interaction between 13q14 candidate proteins and RPAIN. Pull-down was performed for full-length candidate proteins (1-4) and KPNA3 was one of the negative controls (NC). A further negative control (5) was co-transfection of SETDB2 with an unrelated protein, TIFF-I, with the same tag as RPAIN. Bait proteins were detected with anti-PA antibody and RPAIN was detected with anti-FLAG antibody. RPAIN was co-precipitated with C13orf1, RCBTB1 and SETDB2 but not with the negative control KPNA3. SETDB2 did not precipitate the other negative control, TIFF-I, either. ~=weak interaction, +=positive interaction, ++=strong interaction.

To summarize the results of the Y2H screen, a total of 196 interactions between 58 proteins were found by screening with 13q14 candidate proteins. From the interaction partners, 8 could be co-expressed with the 13q14 target proteins and 17 interactions were validated with pull-down experiments. The fact that precipitated proteins were not enriched in comparison to the inputs indicates that the conditions for each single interaction were not optimal. For the data that could not be validated due to PA-cross-reactivity, further studies with other tags are needed. Interestingly, interactions among 13q14 proteins and interaction partners not detected in the Y2H screen were found in the pull-down experiments, indicating that the poor expression of full length proteins in the Y2H screen screening is a major problem for a comprehensive finding of interaction partners with this method. This result highlights both the importance of testing different methods to find interaction partners and the intrinsic difficulties for expression of 13q14 proteins in cells from different origins like bacteria and yeast. To conclude, the novel interactions found here for the 13q14 candidate proteins constitute a milestone for the further characterization of their function and their possible involvement in the development of cancer. The relevance of these interactions as related to cancer and malignancy will be discussed at the end of the next chapter.

4. Discussion

B-CLL is a leukemia that is characterized by a defined set of well described genomic abnormalities. The most common aberration is loss of a critical region in chromosome band 13q14 that is found in over 50% of B-CLL patients (Döhner et al. 2000). The high frequency of deletion of 13q14 in CLL cells suggests that this region harbors a central pathomechanism involved in malignant transformation. Analyses of the sequence of the critical region in CLL cells failed to identify point mutations in any of the genes (Baranova et al. 2003, Hammarsund et al. 2004, Mertens et al. 2000, Migliazza et al. 2001, van Everdink et al. 2003). Therefore, the tumor suppressor function is probably not related to point mutations. Recent functional studies performed in our group (Tschuch 2006) showed that modulation of the mRNA levels of 13q14 candidate genes results in changes in the transcriptome signature: over-expression or down-regulation of the non-coding gene *DLEU2* and the three protein-coding genes *C13orf1*, *RFP2* and *KPNA3* resulted in de-regulation of genes that are functionally involved in the repression of transcription, inhibition of proliferation, hematopoiesis, signal transduction and the NF-κB pathway. These results indicate that several genes in the critical region are involved in cellular functions that could functionally contribute to the development of CLL. Since the underlying pathomechanism of the disease is still unclear, it is important to analyze the functions of 13q14 candidate genes in detail and to establish molecular tools for validation of these gene functions in primary cells. The focus of this work is a) analyzing the intracellular localization of the five 13q14 protein-coding genes *C13orf1*, *RCBTB1*, *SETDB2*, *KPNA3* and *RFP2* and b) identifying their interaction partners.

4.1. Subcellular localization of 13q14 proteins

Localization of proteins to specific subcellular compartments points to the cellular function of the protein (Huh et al. 2003). Two strategies were employed in order to localize 13q14 candidate proteins intracellularly: first, localization of the proteins was tested by expressing them as fusion proteins with green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* (Niedenthal et al. 1996). Fluorescence emitted by GFP when the fusion protein is excited with light at 395nm was used to detect recombinantly over-expressed 13q14-GFP-fusion proteins. In a second approach,

specific antibodies produced in this work were used to identify the localization of endogenous proteins in two different cell lines, one of embryonic kidney origin (HEK-293) and the other of hematopoietic origin (Namalwa).

4.1.1. Cytoplasmic 13q14 proteins

Three of the 13q14 proteins analyzed —C13orf1, RCBTB1 and RFP2— were found to be localized in the cytoplasm. This is in line with the fact that none of them contains homology to DNA-binding domains or to other conserved domains that would suggest localization in the nucleus. The different proteins localized to distinct cellular compartments and varied somewhat between the recombinantly expressed GFP-fusion proteins and endogenous proteins, as well as among different cell lines. In general, signals from endogenous proteins were much weaker and less defined. This indicates that the gene products from 13q14 are not present in high amounts in the cells. It has been described that GFP-fusion proteins can localize to different cellular sub-compartments as compared to the endogenous proteins (Simpson et al. 2000). Therefore, it cannot be ruled out that the patterns from the recombinant signals might be influenced by the GFP-moiety of the recombinant fusion proteins. This might especially be the case for RCBTB1, where differences in the signal pattern between endogenous and recombinant proteins were observed. In general, very weak staining of endogenous proteins was also detected in the nucleus, most accentuated in Namalwa cells and especially for C13orf1. The subtle differences in localization between the various cell lines indicate that the proteins might have tissue-specific functions or be differentially regulated in these cells. However, it is also possible that this differential localization can be simply explained by the small amount of cytoplasm of the hematopoietic cell lines or by the process of fixation. Further experiments including confocal-microscopy and different fixation methods will help to clarify this issue, as well as analysis of the endogenous 13q14 candidate proteins in other cell lines will help to reveal whether the different localization between hematopoietic and non-hematopoietic cells can be generalized. For RFP2, localization to the cytoplasm has recently been also shown by others (Lerner et al. 2007). Specifically, Lerner et al. could show RFP2 to be a membrane protein of the endoplasmic reticulum (ER). In the here presented work, the pattern of localization of RFP2 was very similar to that of Lerner et al. Thus, RFP2 is located at the ER. For RCBTB1 and C13orf1, immunofluorescence signals were not attributed to a sub-

cellular organelle. The lack of transmembrane domains in RCBTB1 and C13orf1 as assigned by the NCBI Conserved Domains database also indicates that RCBTB1 and C13orf1 are probably not directly associated to membranous cytoplasmic organelles, which would be revealed by distinct localization. Localization of the candidate proteins was also expected to shed light on whether some of the proteins could be components of the same macromolecular complexes (Section 1.2.3.6); if this would be the case, the proteins should have similar localizations. While in the data presented here, no obvious similar localization could be seen for any of the three proteins located in the cytoplasm, some overlap of localization could exist within the cytoplasmic compartment as the patterns were diffuse to some extent. It could also be possible that these proteins belong to the same complexes, but only at distinct time points, only interacting to perform certain cellular processes at those specific time points. Therefore, the localization data presented here do not rule out the possibility that some of the genes products might still be part of the same complexes or more probably, of the same processes. Indeed, results discussed later derived from interaction studies point to an involvement of several proteins in the same cellular processes.

4.1.2. 13q14 proteins localized to the nucleus

Two of the five 13q14 proteins that were analyzed in this work were partly localized in the nucleus and likely have nuclear functions: SETDB2 is probably a histone-methyl-transferase (HMTase) and KPNA3 likely acts as a nuclear transport protein. The signal of endogenous SETDB2 in Namalwa cells as well as the over-expressed GFP fusion-protein in HEK-293 confirmed localization to the nucleus. The localization of endogenous SETDB2 in HEK-293 cells was surprising; however it was detected in a diffuse pattern with aggregations surrounding the nucleus in the cytoplasm (Figure 3.28.A). The pattern of recombinant GFP-SETDB2 fusion protein in HEK-293 and of endogenous SETDB2 in Namalwa was also different, although both proteins were localized in the nucleus. The reason for the unexpected difference in localization is unclear, but the diffuse pattern in the nucleus of HEK-293 that over-express SETDB2 indicate that the GFP-tag affects proper localization of the protein within the nucleus (Simpson et al. 2000). Different localization of endogenous 13q14 candidate proteins in the two cell lines could be attributed to different regulation of the proteins or different function in different cell types. In cells like HEK-293, SETDB2 could be

regulated by excluding the protein from the nucleus, resulting in a cytoplasmic pattern that was observed in this work. It would be interesting to study, whether the cytoplasmic localization is relevant for function of the protein and whether the GFP moiety affects binding of the fusion protein to its target sites and/or binding partners, thereby changing its localization. In contrast to SETDB2, KPNA3 should be transiently localized to the nucleus. KPNA3 has been shown to be a transport protein for NF-κB proteins (Fagerlund et al. 2005). Transport proteins cycle in and out of the nucleus with their cargoes, thus localizing only temporarily in the nucleus (Jäkel et al. 2002). In the experiments presented in this work, localization of KPNA3 was discrepant between recombinantly over-expressed fusion protein and localization of the endogenous protein. Surprisingly, the endogenous protein localized exclusively in the nucleoli in Namalwa cells, what could indicate that the protein detected represent storage in the nucleoli to regulate its functions (Olson et al. 2000). Unfortunately, endogenous KPNA3 could not be detected in HEK-293 cells. The GFP KPNA3-fusion protein localized mainly to the cytoplasm, which could be due to the GFP-tag (Simpson et al. 2000) or to the lack of appropriate signals that trigger import of KPNA3 to the nucleus. The unusual localization for endogenous KPNA3 to the nucleoli of transformed Namalwa cells is interesting: studies from our group (Tschuch 2006) and others (Fagerlund et al. 2005) have implicated KPNA3 in the transport of the apoptosis-related factor NF-κB. Since malignant cells are characterized by a failure to undergo apoptosis, it is conceivable that confinement of endogenous KPNA3 to the nucleoli of Namalwa could impede regulation of apoptosis by NF-κB via restriction of its transport protein to the nucleoli. To this end, it would be interesting to study whether specific transport-inducing signals, like apoptotic signals triggering NF-κB import, are able to change localization of KPNA3 in Namalwa cells, and also whether localization of KPNA3 to the nucleoli alters transport of NF-κB. It is interesting that the different localization is seen in a hematopoietic cell line of malignant origin, and although Namalwa is derived from a B-cell lymphoma other than B-CLL, it would be of major interest to study, whether the lack of transport of NF-κB by KPNA3 in Namalwa cells could lead to a failure of apoptosis.

In summary, the localization of the 13q14 candidate proteins indicates that (a) the proteins RFP2, C13orf1 and RCBTB1 probably have a cytoplasmic function; (b) the SETDB2 protein seems to have a nuclear role although some cytoplasmic-related regulation or function might also occur under certain conditions; and (c) the KPNA3 protein could be involved in defects of apoptosis in hematopoietic cells via transport of NF-κB into the nucleus. In the localization of 13q14 proteins with specific antibodies presented in this work, only little amounts of endogenous proteins could be detected. For the analysis of interaction partners of the candidate proteins, it was required to analyze whether the small amounts of endogenous proteins detected with the specific antibodies correlated with low expression levels, and whether it would be feasible to use the antibodies for co-immunoprecipitation experiments to find interaction partners.

4.2. Antibody detection and expression levels of endogenous 13q14 proteins

Antibodies specific for 13q14 candidate proteins had been produced as a tool to understand the function of 13q14 proteins and identify cellular processes these proteins are involved in. Both, specific monoclonal and polyclonal antibodies were produced, as their different characteristics would be of advantage for future experiments. The most remarkable difference is that polyclonal antibodies are a mixture of antibodies that recognize different epitopes, while monoclonal antibodies are highly specific against one single epitope of the antigen, and either property can be of importance in different approaches. In addition, polyclonal antibodies were produced in guinea pigs and monoclonal antibodies in mice, which open the possibility of performing simultaneous detection of different proteins like in co-localization or co-immunoprecipitation studies. The antibodies performed well in immunofluorescence and Western blot analysis using over-expressed 13q14 proteins (section 3.2). The specificity of antibodies was high as shown by immunofluorescence in co-localization experiments using over-expressed GFP-fusion proteins and in Western blot analysis using proteins purified from bacteria, the later being the original antigens used for immunization. In spite of the antibodies being able to detect their target proteins, immunoprecipitation studies were to date unsuccessful for unknown reasons. The most likely reason for lack of

immunoprecipitation is that conditions were inappropriate for proper binding of antibody to the target protein. Also, it cannot be ruled out that the antibodies might be unsuitable for immunoprecipitation. In addition, detection of over-expressed proteins compared to weak detection of endogenous proteins suggests that the proteins are probably present at low levels endogenously. To test this possibility, RT-PCR analysis of gene expression of 13q14 candidate genes was performed with cDNA from 40 different cell lines (section 3.3). Similar to protein levels, only low expression could be detected with the exception of a few cell lines (Appendix 6), indicating that probably additional factors like post-translational processing of the endogenous proteins also play a role in 13q14 protein production.

In conclusion, detection of endogenous proteins proved to be difficult in the cell lines tested. Based on RT-PCR data it seems that the 13q14 candidate genes were expressed in the cell lines used for analysis, but that the proteins were found at concentrations too low for the detection experiments and immunoprecipitation to be feasible. Therefore, in order to analyze the function of 13q14 proteins, binding partners were isolated with a yeast-two-hybrid (Y2H) screen.

4.3. The yeast-two-hybrid screen approach

Interaction partners of 13q14 candidate proteins were identified in a Y2H screen. The results of the screen help to understand the functional roles of 13q14 proteins by relating them to cellular processes through their interaction partners. This screening method was chosen, because it allows circumventing the problem of low level expression/presence of 13q14 candidate proteins. In order to screen for interaction partners of 13q14 proteins, a total of 60 plasmid expression constructs were made for five candidate proteins (Appendix 5). To sort out auto-active “bait” proteins prior to the screening with a “prey” library of over 5000 human proteins (Stelzl et al. 2005), a pre-screening was performed. To test the possibility of interactions among 13q14 proteins, in the pre-screening not only the “prey” empty-vector was used to test the 13q14 “baits” but also the candidate constructs in the “prey” vector as well. In this pre-screening interaction among 13q14 proteins were also found (Table 3.30). These molecular interactions of 13q14 proteins with each other were of high interest as they support the notion of 13q14 proteins being components of the same macromolecular complexes. These interactions could only be detected in specific combinations of

bait- and prey-proteins, which is likely due to interaction constrains related to the position of the tag. So far, validation of these interactions has not been possible due to problems co-expressing the 13q14 proteins in mammalian cell lines. Nevertheless, the results from the Y2H screen suggest interactions between 4/5 of the 13q14 candidate proteins. It seems that C13orf1 could be physically linking all four interacting 13q14 proteins as it gave positive Y2H screen results with all other three proteins. However, additional experiments are needed to characterize them in more detail. The 36 fragments and full-length tagged 13q14 proteins that were not auto-activating in the pre-screen were then tested against the Y2H-library of more than 5000 human proteins (Stelzl et al. 2005). Strong and robust positive interactions would have been expected to appear in triplicates, but most of the interactions were only found once. From these, the most confident 53 interactions (the ones positive for at least the two metabolic reporter genes) were selected for further validation with pull-down assays. In these experiments, the two proteins were recombinantly co-expressed in cell lines. The genes were expressed as fusion proteins either with 1) a HA-tag, which is a small peptide consisting of 11 amino acids from an immunogenic epitope of the hemmaglutinin protein of the influenza virus, or with 2) a FLAG-tag, which is a synthetic hydrophilic octapeptide, or with 3) a PA-tag, which is a modified domain of staphylococcus protein-A that can be directly precipitated with a matrix coupled to immunoglobulin (Ig), which interacts with PA. In validation with pull-down assay, positive interaction would be reflected in an enrichment of both interaction partners in the precipitate. While PA-proteins that were directly precipitated could always be enriched, candidate interaction partners were co-precipitated but not enriched to the same extent. The reasons for the lack of enrichment of the co-precipitated proteins are unknown, but are likely sub-optimal conditions for interaction and consistent binding throughout the experimental procedure. Due to the large number of candidate interactions observed in the Y2H screen, interactions were tested by pull-down experiments using only one general precipitation buffer. However, a number of factors including ionic strength, pH, detergents and reducing agents among others are critical in protein-protein interaction. Thus, further optimization of conditions for the pull-down assays are required to increase the amount of co-precipitated interacting protein that can be recovered. Still, the candidate interactions described here will help to relate 13q14 proteins to certain

pathways (described in the next sections) and thus constitute a primary cellular context for 13q14 proteins in the network of the proteome.

4.3.1. Novel interactions identified in pull-down experiments

Interactions found in the Y2H screen were validated by testing 13q14 proteins with any of the candidate binding partners in order to have a high number of negative controls. Interestingly, in these experiments, not only Y2H screen interactions were validated but also interactions were detected among candidate interaction partners from the YH2 screen and 13q14 proteins that were not positive in the original Y2H screen. This can be due to different reasons: first, the validation experiments could have produced false positive results. However, the pull-down assay is generally a more robust method for detecting single protein interactions than the original Y2H screen and it is therefore more probable that the pull-down experiments data reflects a real interaction. Second, it cannot be ruled out that the interactions found in the Y2H screen are reflecting only a part of the interactions of the proteins analyzed, e.g. due to vector constraints. One of the major drawbacks in the performance of the Y2H screen was the low number of interactions detected with the full-length 13q14 proteins. The expression of full-length proteins for genes encoded in 13q14 had also been difficult in bacteria and animal cell lines, possibly reflecting an inherent property of the 13q14 proteins, e.g. self-ubiquitination of RFP2 (Lerner et al. 2007). Expression of 13q14 proteins in bacteria proved that several vectors were not efficient and produced only minute amount of recombinant protein (see section 2.2.1). Unfortunately, vectors used in the Y2H screen were pre-determined and could not be chosen to optimize expression levels of the recombinant proteins. This could explain why interactions could be detected in pull-down experiments with specific vectors but not in the original Y2H screen. A third reason could have been that the full length protein is required for the interaction to take place. For the three-dimensional structure of a protein, the context of each part of the protein can be crucial. Without its context, a protein fragment may lack its proper structure and thereby lose its interacting capabilities or the proper context for an interaction to take place. In the selection of fragments used for the Y2H screen, conserved domains of the candidate proteins were used as template to design fragments. However, how the sequence of a protein relates to its folding state is still unknown, so it is possible that the chosen fragments would not properly reproduce the interaction interfaces

present in the full length protein. Thus, some of the interactions of that protein might not be reproduced with fragments used in the Y2H screen.

It has also to be mentioned here that some of the interactions detected in the Y2H screen could not be validated in subsequent pull-down experiments. As mentioned above, the context of the different domains in a protein can have an important impact on its structure and function. Along the lines of this concept, some of the interactions detected in the Y2H screen might be falsely positive because conserved domains present in some fragments but out of their full length protein context might interact unspecifically. Therefore, only the candidate binding partners validated in pull-down experiments will be discussed here in the following sections.

4.3.2. Common interaction partners of 13q14 proteins

From the 53 interaction partners found in the original Y2H screen using 5 candidate proteins from 13q14 as bait, 36% interacted with three or more 13q14 proteins, indicating a broad overlap of common interaction partners for the candidate proteins. The fact that different 13q14 candidate proteins show common interaction partners favors the view that these proteins might be related functionally, e.g. by constituting different entities in common macromolecular complexes or, alternatively, by interacting with the same protein to form different complexes in a condition-dependent manner. Further experiments are needed to elucidate how the common interaction partners relate to the different 13q14 candidate proteins. Nevertheless, the finding of common binding partners for different 13q14 proteins relates the 13q14 proteins to some extent to the same cellular processes and might be an indicator of how functionally intertwined the proteome network actually is. The extent and nature of the functional relationships among the proteins presented here, and its relevance for CLL remain to be further investigated.

4.4. Cellular processes involving 13q14 proteins

The finding of interaction partners was of major interest for the elucidation of the function of the 13q14 candidate proteins. In particular, it has been shown that physical association between a novel protein and a well-characterized one readily indicates that the novel protein has a function related to that of the well-characterized protein (Takashi Ito et al. 2001). In the next sections, the possible involvement of the 13q14 candidate proteins in protein translation, apoptosis, gene expression and

signal transduction will be discussed, as well as the connection of these functions with the development of cancer.

4.4.1. Involvement of 13q14 genes in the regulation of translation

The regulation of translation is critical for the proper function of cells and for determination of the cellular phenotype. Biological and phenotypic complexities are ultimately derived from changes in protein concentration and localization, posttranslational modification and the combinatorial diversity of protein-protein interactions. The crucial role of regulation of translation is highlighted by the fact that protein abundance in the cell has a poor correlation with mRNA abundance. This has been shown in a study in yeast (Gygi et al. 1999), where variability between mRNA and protein concentration among several hundred genes ranged from 10- to 100-fold, indicating that gene transcription is only one of the regulatory layers determining the cell phenotype and that protein translation plays a major role in the determination of the proteins that will be present in the cell. The process of protein synthesis or translation is one of the most complex biochemical processes occurring in the cell and requires several hundreds of different proteins and diverse RNAs (Lamberti et al. 2004). Translation is known to be regulated at different levels, and the role of regulation of translation initiation is well studied in cancer (Holland et al. 2004). In contrast, the process of translation elongation is less well studied although there are also studies linking the expression of translation elongation factors (EFs) to cancer. In the present work it was found that EFs directly bind to 13q14 proteins, with eEF1A1 co-precipitating with RCBTB1 and C13orf1 (Figure 3.33) and eEF1G co-precipitating with SETDB2 and to a lesser extent with C13orf1 and RCBTB1 (Figure 3.34). Interestingly, eEF1A is known to play a general role in malignancy. In one of the relevant studies, Talapatra et al. (Talapatra et al. 2002) describe a protective role for eEF1A against ER stress-induced apoptosis. However, although the authors argue that this function is accomplished through inhibition of the ER-specific unfolded protein response (UPR), no direct link between the ER and eEF1A is proposed. The interaction between eEF1A and eEF1G and C13orf1 described here points to a 13q14-protein link between the ER and elongation factors since C13orf1 interacted in the Y2H screen with RFP2 which in turn is an ER membrane-associated protein involved in ER-mediated protein degradation (Lerner et al. 2007). The interaction of RFP2 through C13orf1 with eEF1A and eEF1G could be the mediator between

signals in the cell inducing ER-related apoptosis and the effector processes in the ER that trigger apoptosis. This does not exclude that RFP2 might be involved in other pathways related to apoptosis as well (like NF-kB, (Matsuda et al. 2003)). In another study Thornton et al. (Thornton et al. 2003) relate the oncogenic potential of eEF1A to overall increase in protein production. They propose that selectivity in the proteins being produced may underlie the mechanism for neoplastic transformation but argue that it is hard to understand how the elongation factors could attain such specificity. The interaction of eEF1A with RCBTB1 could point to a control mechanism in protein translation specificity through an alternative guanine-nucleotide exchange factor (GEF) for eEF1A, which would confer eEF1A with selectivity. Indeed, RCBTB1 was also shown to interact with eEF1G (the protein in the elongation complex bringing the GEF to activate the catalytic protein eEF1A), but it remains to be studied whether the three proteins, RCBTB1/eEF1G/eEF1A, can form a complex *in vivo* and whether RCBTB1 is catalytically active as a GEF for eEF1A *in vitro*. A certain condition, e.g. an apoptotic signal, would be transduced to the translation machinery of the cell by eEF1G recruiting RCBTB1 to activate eEF1A, thereby activating the production of certain proteins, e.g. apoptotic proteins. Alternatively, or in addition to, interaction with unknown regulatory proteins might be the key for favoring translation of certain kind of proteins. This role might be played by C13orf1, another candidate 13q14 protein shown to interact with both eEF1A and eEF1G and supposed to be involved in protein-protein interactions (van Everdink et al. 2003). C13orf1 could be involved in the regulation of elongation by recruiting alternative GEFs to eEF1G and to eEF1A. In support of this hypothesis, Y2H screen data in this work also showed that C13orf1 interacts with RCBTB1 (Table 3.30), although this data still has to be validated with other methods. The last candidate protein described to interact with the elongation factors is SETDB2, a potential histone-methyl transferase (Mabuchi et al. 2001). In the context of its interaction with EFs, it would be interesting to study whether SETDB2 exerts a catalytic activity as a methyltransferase, and whether this activity is restricted to histones or can be extended to other substrates. Indeed, endogenous SETDB2 has been found to localize in the cytoplasm surrounding the nucleus in certain cell lines (Figure 3.28), what could be a hint to a specific function under specific conditions or in specific cell types as methyltransferase of cytoplasmic proteins like the EFs.

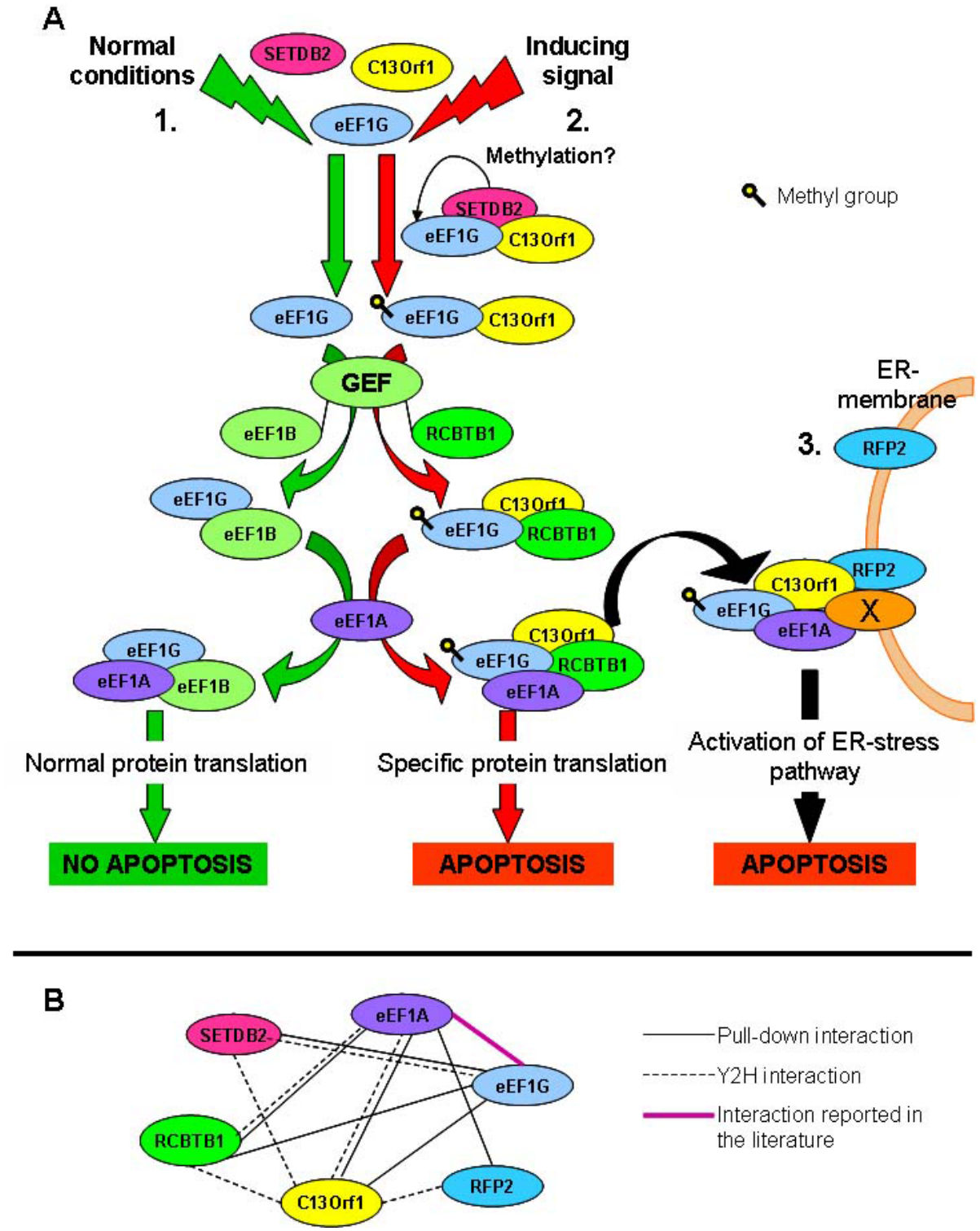


Figure 4.1 Model of the interactions of 13q14 proteins with elongation factors. A. C13orf1, RCBTB1 and SETDB2 interact with elongation factors eEF1A and eEF1G. A model is suggested where: 1. under normal conditions eEF1G recruits the guanine exchange factor (GEF) eEF1B to eEF1A and the elongation complex translates normally (green pathway). 2. Under specific inducing signals C13orf1 recruits SETDB2 and eEF1G, thereby modifying eEF1G to specifically bind other proteins probably via methylation. This complex recruits RCBTB1 as alternative GEF for eEF1A thereby inducing translation of specific proteins, e.g. apoptosis-related proteins (red pathway). 3. Under specific conditions, C13orf1 might recruit the EFs to the ER-membrane protein RFP2 to induce activation of the ER-stress related apoptotic pathway, probably also involving still unknown proteins (X) (black pathway). **B.** Network of interactions between 13q14 proteins and elongation factors described in this work. Numbers do not necessarily imply sequential processes.

This would be of interest as methylation of at least eEF1A seems to be related to increased activity (Lamberti et al. 2004), and increased activity of eEF1A has in turn been related to different types of cancer (Thornton et al. 2003). The interactions described above suggest that activity of elongation factors could be modulated by 13q14 proteins (Figure 4.1). Under certain conditions, C13orf1 would a) recruit SETDB2 to selectively activate eEF1G and b) help to recruit RCBTB1 as GEF to the elongation complex that in turn could preferentially translate certain proteins, e.g. apoptosis or growth related proteins. Since expression of C13orf1, RCBTB1 and RFP2 is significantly down-regulated in B-CLL cells (Mertens et al. 2002, Tschuch 2006), detrimental loss of C13orf1 and RCBTB1 activity could lead to decreased translation of specific genes like apoptosis-promoting genes. In addition, loss of RFP2 could contribute to the lack of induction of apoptosis via elongation factors and the ER, altogether leading to a pre-malignant apoptosis-resistant state.

In summary, the candidate 13q14 proteins were directly related to the cellular process of protein translation through EFs, and indirectly related to processes associated with the EFs such as apoptosis. Indeed, they could constitute the missing link between elongation factors and their role in apoptosis and malignancy. The interactions presented here open new intriguing avenues for further investigation. It will be interesting to determine, whether the 13q14 proteins confer indeed selectivity to the translation mechanism via their interaction with elongation factors, whether this interaction induces the preferential production of pro-apoptosis proteins or other specific factors related to malignancy and whether certain conditions might be responsible for triggering that selectivity.

4.4.2. Involvement of 13q14 proteins in regulation of gene expression

Regulation of gene expression is a critical process involved in most cellular functions and performed by specific regulatory proteins. Two binding partners of 13q14 proteins have been associated with regulation of gene expression: EED and RXRA. EED (Embryonic Ectoderm Development) is a member of the Poylcomb-group (PcG) family of regulatory proteins involved in gene regulation. The PcG-proteins regulate expression of genes involved in development, but in the last years additional functions have been attributed to different PcG genes, like X-chromosome inactivation (Chadwick et al. 2003), or globally modulating histone modifications in

brain (Kim et al. 2007). Especially interesting in the context of the B-CLL pathomechanism is the regulation of hematopoiesis by EED and its role as a tumor suppressor in hematopoietic precursors (Lessard et al. 1999). In this work, four of the candidate genes localized in 13q14 were shown to interact with EED by pull-down experiments (Figure 3.35). One of these genes, KPNA3, is known to be a transport protein and could be involved in the transport of EED into the nucleus. Interestingly, loss of EED function results in the increase in the frequency of pre-B lymphoid progenitors. Thus, loss of KPNA3 as seen in CLL could result in miss-localization and loss of function of EED, which could eventually increase the frequency of pre-B-lymphoid progenitors and thus increase the probability of one of these B-cells to become malignant. Indeed, in mutant EED mice monoclonal B-cell lymphoma was developed in some older mice and in 100% of the animals, if they were treated with some genotoxic agent (Lessard et al. 1999). Therefore, it would be interesting to test the hypothetical link between loss of function of KPNA3/EED and malignant transformation in further experiments. Also interesting in relation with the function of EED is the interaction with SETDB2, as the later is supposedly a HMTase and therefore a chromatin remodeling protein, and the former exerts a function as a transcription regulator. However, as SETDB2 is not significantly regulated in B-CLL patients, it remains to be determined, whether this novel interaction between SETDB2 and EED plays a role in the development of CLL. The other two proteins localized in 13q14 that are here shown to interact with EED are localized in the cytoplasm and therefore are unlikely to be directly involved in its function as transcriptional regulator in the nucleus. Nevertheless, they could have regulatory functions and/or be involved in as yet unknown functions of EED.

An additional DNA regulatory protein found as interaction partner of C13Orf1, RCBTB1, SETDB2 and KPNA3 is the transcription factor Retinoid X Receptor Alpha (RXRA) (Appendix 7). The experiments to validate these results produced an uncharacterized two-band pattern for the RXRA protein (Figure 3.37), that could be due to a compromised identity of the RXRA-clone. Therefore, the possibility of some real interaction taking place between 13q14 genes and RXRA (as found in the Y2H screen) remains to be further characterized until further experiments clarify the identity of the clone provided by the RZPD (German resource center for genomic research). Nevertheless, the possible implications of these interactions can be analyzed: RXRA is a member of the nuclear receptor (NR) superfamily of

transcription factors and constitutes one of the isotypes of NRs binding retinoic acids (RA). NRs play a crucial role in many cellular processes such as regulation of cell growth, proliferation, immune response and apoptosis, all of which are determined by ligand binding. Unlike the well known RARA (Retinoid Acid Receptor Alpha), RXRA only binds to 9-cis-RA (9cRA) and is able to homodimerise or heterodimerise in order to bind to DNA (Chambon 1996). Interestingly, RXRA has recently been shown to be involved in mitochondrium-mediated and ER stress-related apoptosis (Tiwari et al. 2006). In one study, the effects on apoptosis were correlated with the re-localization of RXRA from the nucleus to the cytoplasm (Zeng et al. 2006), and in the present work the proteins C13Orf1 and RCBTB1 interacting with RXRA were shown to localize in the cytoplasm and were related to ER-mediated apoptosis (sections 3.4 and 4.4.1 respectively). These 13q14 proteins might therefore be part of cytoplasmic function of RXRA modulating apoptosis. SETDB2 and KPNA3 are the other 13q14 proteins that interact with RXRA. While SETDB2 has been localized to the nucleus in the present study, and could therefore be involved as HMTase in the silencing of genes by RXRA, KPNA3 could be involved in the transport into the nucleus of RXRA since it has been characterized as an import protein for other transcription factors (Fagerlund et al. 2005). Further studies are needed to clarify the significance of 13q14 proteins in the diverse nuclear and cytoplasmic functions of RXRA, and in its transport into the nucleus. It will be highly interesting to clarify, whether the 13q14 proteins are related to the functions of RXRA associated with apoptosis and thus play a role in the development of the malignant phenotype in B-CLL.

4.4.3. Additional interaction partners of 13q14 proteins

Several additional interaction partners were found for 13q14 proteins (table 3.32). However, for these candidates it is more difficult to associate a function to the pathogenesis of B-CLL. Nevertheless, since they constitute new links in the cellular protein-interaction network these interaction partners will be briefly described here:

CRMP-1

CRMP-1 was identified as interaction partner of C13Orf1, RCBTB1 and SETDB2 (Figure 3.36). CRMP-1 is a member of the family Collapsin Response Mediator Protein (CRMP) (Deo et al. 2004) that participate in signaling pathways and signal transduction. Among other pathways, CRMPs have been related to the immune system signaling (Shih et al. 2003) and it is proposed that CRMPs might be involved

in lymphoid cell-cell interaction and B-cell receptor signaling (Bismuth et al. 2002). Interestingly, CRMP-1 has been directly related to tumor invasion and metastasis in malignancies like lung cancer through as yet unknown pathways (Shih et al. 2003).

GNB2

The signal transduction protein GNB2 was found as interaction partner of RCBTB1 and SETDB2 (Figure 3.38). GNB2 is a G protein coupled to membrane receptors whose function is to transduce the signal from receptors to effector molecules (Downes& Gautam 1999). GNB2 has been related to different effector functions such as inactivation of low-voltage Ca^{2+} channels and modulation of gene expression via the glucocorticoid receptor (GR). In the presence of specific signals, GNB2 associates with GR in the cytoplasm and co-migrates to the nucleus, thereby inhibiting GR-activation of target genes (Kino et al. 2005). Interestingly, the GNB2/GR complex has also been shown to co-migrate to the cell membrane after stimulation with other input signal, indicating that the diverse responses elicited by glucocorticoids might depend on the intracellular location of the GR (Kino et al. 2005). The GNB2 binding partners presented here could be involved in both the genomic and the non-genomic functions of GNB2/GR. Interaction of GNB2 with SETDB2 might provide the link to genomic inhibitory effects: once triggered, translocation of the GNB2/GR complex into the nucleus would induce binding of the complex to GR-target genes but GNB2 recruitment of the HMTase SETDB2 to those sites could result in inhibition of gene expression. In contrast, the 13q14 protein RCBTB1 is localized in the cytoplasm and could thus be involved in the cytoplasmic function of GNB2/GR. Whether these interactions are related to pathways that could lead to malignancy are issues that require further investigation.

C11Orf2

The protein C11Orf2 was found as interaction partner of SETDB2. To our knowledge there is no function assigned to date to the C11Orf2 protein. However, it should be noted that C11Orf2 is predicted to be a nuclear protein (O'Brien et al. 2000), and that SETDB2 is also predicted to be a nuclear protein with putative HMTase activity and has been shown here to be localized to the nucleus (Figure 3.23).

RPAIN

The last protein validated in this work as an interaction partner of 13q14 proteins is RPAIN (RPA Interacting Protein, also known as “RIP”). Several isoforms of the RPAIN protein exist (Chen et al. 2005), and all isoforms interact with and are

involved in the transport of the nuclear protein RPA and in multiple processes related to DNA metabolism. Some of these isoforms are very interesting as they have functions related to the DNA-damage response, like isoform β that is involved in the transport within the nucleus of RPA to PML bodies, from where RPA is released upon DNA damage (Park et al. 2005). The interaction between RPAIN and the 13q14 candidate proteins could relate 13q14 proteins to the DNA-damage response and also hint to a role of these proteins in malignancy through this interaction partner.
The interactions presented here open new possibilities of inter-connection in the protein network involved in the different cellular processes and might be helpful for our understanding of how those processes are related to cancer.

4.4.4. Lessons learned from binding partners of 13q14 proteins

The identification of binding partners of a candidate protein of unknown function frequently allows attribution of functional characteristics to the candidate protein. Several interaction partners for 13q14 candidate proteins were found that are involved in different cellular processes like protein translation, DNA-regulation, signal transduction and apoptosis. The fact that several interaction partners bind to more than one 13q14 protein implies that the 13q14 proteins are possibly involved in similar cellular functions. The data presented in this work provides first hints towards possible functions of the 13q14 proteins, but further experiments are needed to characterize their specific roles in each cellular pathway in greater detail. The most prominent example is C13orf1, which is shown to interact with many different proteins. It seems that the function of C13orf1 is primarily based on protein-protein interaction. Indeed, binding of C13orf1 to different proteins favors the notion of a function as a “bridging” protein or recruitment factor involved in physically interconnecting other proteins. The data presented here suggests that RCBTB1 protein is also involved in several pathways, most remarkably hematopoietic and apoptosis-associated pathways. The interaction of RCBTB1 with different proteins, such as eEF1A and RXRA, link RCBTB1 to apoptosis and make it a strong candidate for the tumor suppressor gene (TSG) located in 13q14. Another strong candidate for the postulated TSG and the most interesting protein in the 13q14 region in relation to B-CLL seems to be the RFP2 protein. In previous studies, it had been established that the RFP2 gene is included in the smallest region of deletion in chromosomal band 13q14 and it is the only gene significantly down-regulated in all B-CLL patients

analyzed (Mertens et al. 2002). In a recent study, the RFP2 protein has been characterized as a membrane component in the endoplasmic reticulum (ER) and is suggested to be involved in ER-mediated protein degradation (Lerner et al. 2007), with no obvious link to the development of B-CLL. The interaction partners of RFP2 described in this work (elongation factors and EED) link RFP2 as ER-membrane protein to ER-mediated apoptosis and hematopoiesis, highlighting its possible role in B-CLL. Also, the finding that SETDB2 is an interaction partner of transcription factors like EED and RXRA strongly suggest that it may indeed function as HMTase playing a role in the regulation of gene expression. Its binding to cytoplasmic proteins like eEF1G points to a second function, and it will be highly interesting to clarify this putative dual function. The last protein to be analyzed was KPNA3, an importin that shuttles proteins as cargo into the nucleus. Several cargoes have been described that relate KPNA3 to apoptosis —notably NF-κB (Fagerlund et al. 2005)—, and the localization of KPNA3 in the nucleoli of Namalwa cells shown in this work could point to a role of KPNA3 in a pathogenic alteration of the NF-κB pathway. Here, EED is described as an additional novel binding partner for KPNA3, and its role as a transcription factor involved in myeloid and B-cell precursors' development opens the possibility for regulatory interactions (Lessard et al. 1999, Richie et al. 2002).

It should be noted that the novel interactions presented in this work not only relate 13q14 proteins to specific cellular pathways but also add novel cross-links (Figure 4.2) in the general protein network. The data presented here shows that the links of 13q14 proteins to the B-CLL phenotype are intricate and that the relation to apoptosis and other potentially pathogenic processes might be indirect, explaining why the pathomechanism residing in the 13q14 region has not been uncovered to date even though it is under intense investigation.

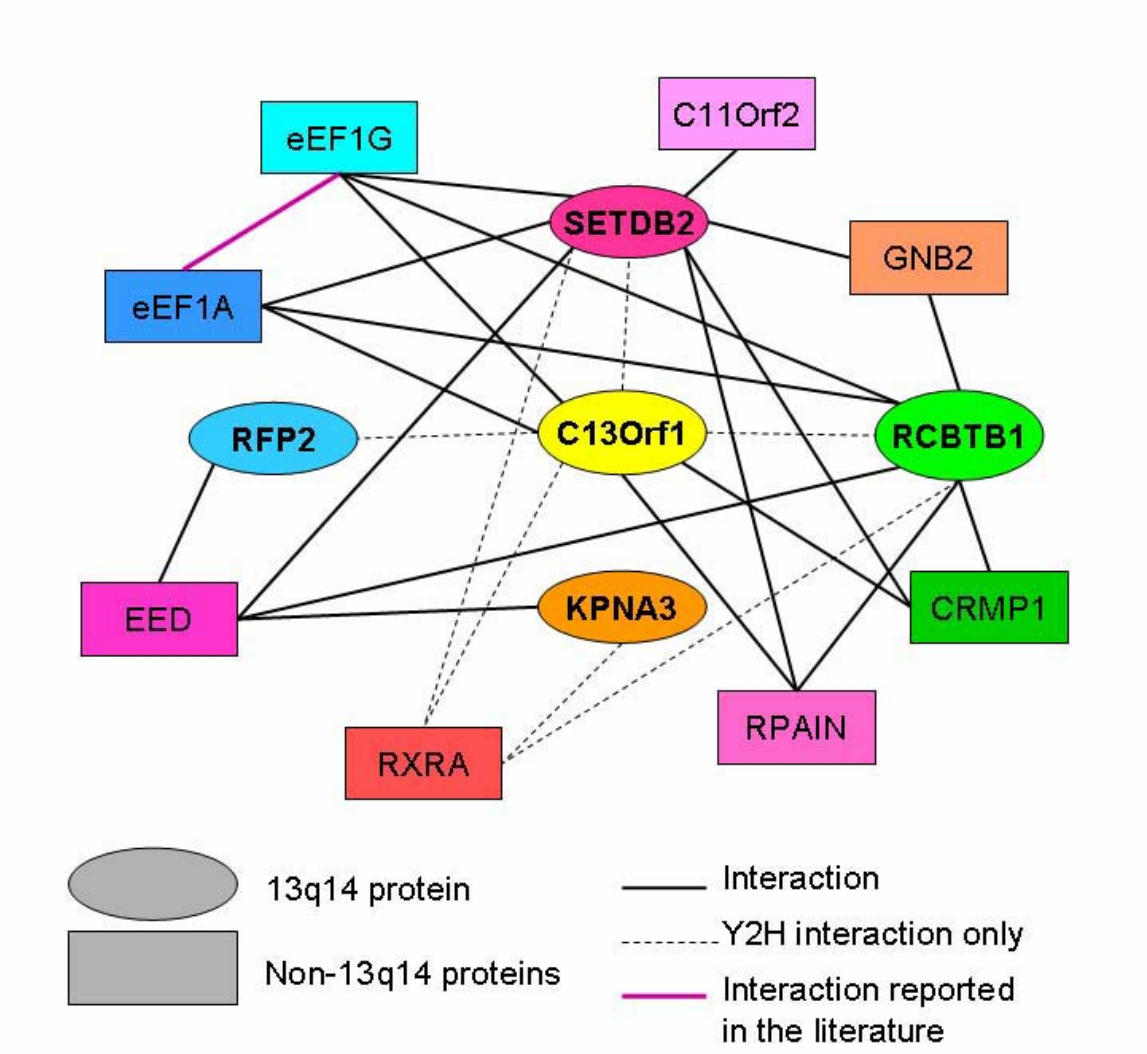


Figure 4.2 Schematic representation of the most relevant protein interactions discussed in this work. Interconnections between proteins are depicted here and indicate a complex network between the different proteins. eEF1A and eEF1G are protein-translation factors involved in apoptosis and cell proliferation; EED is a transcription factor involved in hematopoiesis; RXRA is a transcription factor involved in apoptosis; RPAIN is a transport protein for RPA involved in DNA-damage response; CRMP1 is a signal transduction factor involved in signaling pathways; GNB2 is G protein involved in glucocorticoid functions; C11Orf2 is a nuclear protein of unknown function.

4.5. Concluding remarks

The goal of this work was the functional characterization of five of the proteins (C13orf1, RCBTB1, SETDB2, RFP2 and KPNA3 (Figure 1.7)) encoded in the genomic region 13q14 that is deleted in over 50% of B-CLL patients and in other types of cancer. To date it is unclear, how deletion of a critical region in 13q14 results in the development of cancer. Results from previous studies that include modulation of the expression of 13q14 genes (Tschuch 2006), the role of epigenetics (Mertens et al. 2006) and the non-coding genes of the region (Calin et al. 2002, Mertens et al. 2006, Wolf et al. 2001) prompted us to study the protein-coding genes of 13q14 as possible candidates for the phenotype of B-CLL. Defining the cellular localization of 13q14 proteins allowed confinement of a function of these proteins either in the cytoplasmic and/or the nuclear compartment. The additional data obtained from a Y2H screen allowed to associate 13q14 proteins with specific cellular processes and pathways like transcription regulation, protein translation, signaling pathways and apoptosis. Furthermore, 13q14 proteins could be directly correlated with processes that could explain characteristics of the phenotype of B-CLL cells, especially their capability to escape apoptosis. Using the annotated function of these interaction partners, 13q14 proteins could be related to the process of programmed cell death through a) factors involved in elongation of translation, b) the endoplasmic reticulum and c) the control of gene expression. The data presented in this work constitute a starting point towards understanding the role of the proteins analyzed here in the pathogenesis of B-CLL. The next steps will be to characterize the role of the protein-protein interactions in the molecular pathogenesis of B-CLL. In addition, the protein-protein interactions presented here add to our knowledge of the complex protein network underlying cellular processes and the development of complex diseases such as cancer. The elucidation of these networks constitutes a step towards the development of better diagnostic and therapeutic tools for the treatment of cancer.

5. Appendix

Appendix 1

Primers

List of primers used for cloning and sequencing of 13q14 target genes. Gene-specific primers for full length gene amplification from cDNA were used for four genes: C13orf1, RCBTB1, SETDB2 and RFP2. Three genes required "nested-PCR" for amplification from cDNA: C13orf1, RCBTB1 and SETDB2. Red sequences indicate the start codon (ATG) and the stop codon, if they are present in the sequences. For sequencing, internal primers were designed for RCBTB1 and SETDB2, were not required for C13orf1, and were kindly provided by C.Tschuch and D.Mertens (DKFZ) for KPNA3 and RFP2. Gene-specific primers for restriction cloning were designed to include the specific sequences for the restriction enzymes, and are indicated together with the corresponding name of the enzymes colour-coded. The Y2H screen primers present no colour-coding since all the forward primers have SalI (gtcgac) and all the reverse primers have NotI (gcggccgc) as restriction sites. RCBTB1 has XhoI (ctcgag) restriction sites in forward primers because there is a SalI restriction site in the gene's sequence.

Gene-specific primers for full length gene amplification

Gene	Primer Name	Sequence from 5' to 3'	Features
C13Orf1	C6-f2	atggccacctcggtgttgtgctg	Start
C13Orf1	C6-rev	TCAGAAGATTTGCTGTTCAAATAATAT	Stop
RCBTB1	C7-fow	atggtggatgtcggaaagtgg	Start
RCBTB1	C7-rev	ttcagttcttaaaggctccacatttac	Stop
SETDB2	C8-fow	atgggagaaaaaaatggcgatgcaaaaac	Start
SETDB2	C8-rev	ttttctacatttattaaccccacattg	12bp before Stop
RFP2	RF-fow	atggagctg cttgaagaag atctc	Start
RFP2	RF-rev	gtggta ctgaacaatg tggcagaa	24bp before Stop

Gene-specific primers for "nested-PCR" in 5'- and 3'-UTR

Gene	Primer Name	Sequence from 5' to 3'
C13Orf1	C6-f1	gcgatggccacctcggtgttgtg
C13Orf1	C6-r1	gtatactgaatacattggttccttagacg
C13Orf1	C6-r2	tggtatactgaatacattggttccttag
RCBTB1	C7-f1	tcggcttatttataggaattgcttgaag
RCBTB1	C7-f2	ataggaattgcttgaagccagag
RCBTB1	C7-r1	gttcttaaaggctccacatttact
RCBTB1	C7-r2	tccacatttactggctttagcaatg
SETDB2	C8-f1	tataagcgacatcaaaagatgggag
SETDB2	C8-r1	cataaaccccatgacaaaggaaaacg
SETDB2	C8-r2	cacagcagtgaacaagataccctatg

Internal primers

Gene	Primer Name	Sequence from 5' to 3'
RCBTB1	C7-If1	tgaactatagtaactgtctaggaactggag
RCBTB1	C7-If2	cacctcacccacttctcctgcacc
RCBTB1	C7-Ir1	acacctctccatcagctgccagagc
RCBTB1	C7-Ir2	ctggaacatggatcgaaaatgctcacac
SETDB2	C8-If1	acctgcttgagacagagtgtaac
SETDB2	C8-If2	gagtctgtcactccagaagataatg
SETDB2	C8-Ir1	agtcacaggaatcagtaaacatgctg
SETDB2	C8-Ir2	tcgcctggttacatctgctccttg

Appendix 1 — Continued

Gene-specific primers for restriction cloning

Gene	Primer Name	Sequence from 5' to 3'	Features
C13Orf1	C6-ATV-Bam-Fow	gagagaggatccatggccacctcggtgttgtgc	BamHI
C13Orf1	C6M1_fow-m	gaAAGCTTtcatggccacctcggtgttgtgc	HindIII
C13Orf1	C6-NcoI_fow	gagagaccatggccacctcggtgttgtgc	NcoI
C13Orf1	C6F196_rev	tctctcgagacaaatacattcagaagatttgctgttc	XhoI
C13Orf1	c6-BamHI-rev	gagagaGGATCCgaagatttgctgttcaaataatattt	BamHI
C13Orf1	c6-HindIII-rev	gagagaAAGCTTtcagaagatttgctgttcaaataatattt	HindIII
RCBTB1	C7-fow	gagaattcggatccatggtggatgtcggaaagtgg	Eco RI-Bam HI
RCBTB1	C7-Acc65I-fow	agaagaggtaccatggtggatgtcggaaag	Acc65I=Asp718
RCBTB1	C7-mam_fowI	taGCTAGCaTCCGGAatggtggatgtcggaaagtgg	NheI-BspEI
RCBTB1	C7-mam_rev-I	taGGATCCccTCTAGAgttcttaaaggctccacatttac	BamHI-XbaI
RCBTB1	C7-rev	tctctcgaggttcttaaaggctccacatttac	XhoI
RCBTB1	C7-BamHI-rev	agaagaggatccgttcttaaaggctccacat	BamHI
RCBTB1	C7-XmaI-stop-rev	agaagacccgggttagttcttaaaggctccacat	XmaI
SETDB2	C8-NcoI_fow	gagagaccatgggagaaaaaaatggcgatgc	NcoI
SETDB2	C8-fow	gaggatccatgggagaaaaaaatggcgatgc	BamHI
SETDB2	C8-mam_fow	gaAAGCTTtcatgggagaaaaaaatggcgatgc	HindIII
SETDB2	C8-ATV-EcoV-Fow	gagagaGATATCatgggagaaaaaaatggcgatgc	EcoRV
SETDB2	C8-ATV-Hind-Rev	gagagaAAGCTTtaatatttttttctacatttattaaccccac	HindIII
SETDB2	C8-XmaI-stop-rev	gagagacccgggttatctacatttattaaccccacattgg	XmaI
SETDB2	c8-BamHI-rev	gagagaggatcctctacatttattaaccccacattgg	BamHI
KPNA3	Kp_fow	gacccgggattcatggccgagaaccccagcttg	XmaI/SmaI-EcoRI
KPNA3	Kp-NcoI_fow	gagagaccatggccgagaaccccagcttg	NcoI
KPNA3	Kp-mam_fow-I	taGCTAGCaTCCGGAatggccgagaaccccagcttg	NheI-BspEI
KPNA3	Kp-XmaI-rev	gagagacccgggttaaaattaaattcttttgtttgaaggttggctg	XmaI
KPNA3	Kp_rev	tctctcgagtgtttgaaggttggctgttggatc	XhoI
KPNA3	Kp-HindIII-rev	gagagaAAGCTTaaaattaaattcttttgtttgaaggttggctg	HindIII
RFP2	RF-ATV-Bam-Fow	gagagaGGATCCatggagctgcttgaagaagatctc	BamHI-NcoI
RFP2	Rf_fow-Nhe-Bsp	gaGCTAGCaTCCGGAatggagctgcttgaagaagatctc	NheI-BspE I
RFP2	Rf_fow	gaGGATCCatggagctgcttgaagaagatctc	BamHI-NcoI
RFP2	Rf_fow-Nhe-Bsp	gaGCTAGCaTCCGGAatggagctgcttgaagaagatctc	NheI-BspE I
RFP2	Rf-BamHI-rev	gaa gaaGGATCCttctgccacattgttcagtaccac	BamHI
RFP2	Rf-Xma-STOP-rev	gagagaCCCGGGttataatagtttatatttgcacacaaattctgc cac	XmaI
RFP2	Rf_Rev	cacCTCGAGAAGCTTttctgccacattgttcagtaccac	XhoI-Hind III
RFP2	Rf-ATV-Hind-Rev	gagagaAAGCTTtaatagtttatatttgcacacaaattctgccac	HindIII
$RFP2_{S154-E254}$	RFP2-S154_fow	caCCATGGGATCCtccctcttccagagctttgagacc	NcoI-BamHI
$RFP2_{S154-E254}$	RFP2-E254-b_rev	gagCTCGAGttactcctgcatctgttgcagaaatac	XhoI

Yeast-two-hybrid primers:

C13Orf1

Primer Name	Sequence from 5' to 3'
C6-SalI-fow-1	AGAGAGTCGACGATGGCCACCTCGGTGTTGTGCTGCCTGCGG
C6-fow-62	AGAGAGTCGACGCAAAACAAAAGCTATTTTGAATTCAAAATCCAG
C6-NotI-rev-196	AGAGAGCGGCCGCTCAGAAGATTTGCTGTTCAAATAATATTTT
C6-rev-61-Stop	AGAGAGCGGCCGCTTAATGTAAAGGTGCGCTGGCTAAACAACTTCC
C6-rev-178-Stop	AGAGAGCGGCCGCTTAAAACTCACTGAACTGGCAATCCAAAATTGC

SETDB2

Primer Name	Sequence from 5' to 3'
C8-fow-165	AGAGAGTCGACGCATGCAAAGACAAACTCTCATTCTTCAGCA
C8-fow-217	AGAGAGTCGACGCAGTTGGCTCGGAATTACCCAAAGCAAAAAGA
C8-fow-290	AGAGAGTCGACGATAACAAAATGCGCATGTCTTCAACTGACAGC

Appendix 1 — Continued

Primer Name	Sequence from 5' to 3'
C8-fow-354	AGAGAGTCGACGCAAGTGAGGTTACAGGTGTTCAAAACTGAG
C8-fow-429	AGAGAGTCGACGGAAGTTCTCCCATTAGGATTGGAAACACAT
C8-NotI-rev-719	AGAGAGCGGCCGCTTATAATATTTTTTTTCTACATTTATTAA
C8-rev-164-Stop	AGAGAGCGGCCGCTTATCGTCTTTGGAAGTGACATTTGATTGGCAG
C8-rev-216-Stop	AGAGAGCGGCCGCTTAAACATAGGTATTGAAAGAAAAGTTATCTGT
C8-rev-289-Stop	AGAGAGCGGCCGCTTAGTCTATGCAGCCCTCAGAGCAGTCACAGGA
C8-rev-353-Stop	AGAGAGCGGCCGCTTAAGGACCATGTTGGACAACTCGGTTTTGACA
C8-rev-639-Stop	AGAGAGCGGCCGCTTAAAGGAAGCGGCCGACATTTCCTTCTTTGTT
C8-rev-699-Stop	AGAGAGCGGCCGCTTAAACCCCACATTGGCAGAAGATTTCCTTCTC
C8-SalI-fow-1	AGAGAGTCGACGATGGGAGAAAAAAATGGCGATGCAAAAACTTTCTGGAT

KPNA3

Primer Name	Sequence from 5' to 3'
Kp-fow-106	AGAGAGTCGACGAGACCTGCTCTTGTGGAGGTACATTTCCTAAGC
Kp-fow-159	AGAGAGTCGACGCCTCAGATGACCTTACTTCCACTGCCTCCA
Kp-NotI-rev-229	AGAGAGCGGCCGCTTATAGAAACATACCTGTAATCCCAGCTAC
Kp-rev-102-Stop	AGAGAGCGGCCGCTTAGTCACTGGATAACAGTTTTCTTGCTGCCTG
Kp-rev-158	AGAGAGCGGCCGCAGGGAAAAAGTTACCATTCCAGGTCGGCGCTG
Kp-SalI-fow-1	AGAGAGTCGACGATGAGTAGTCAGGAACTGGTCACTTTGAATG
Kp-fow-103	AGAGAGTCGACGAGAAATCCACCGATTGATGACTTAATAAAA
Kp-fow-272	AGAGAGTCGACGGGAGGTAATGAACAGATACAGATGGTTATT
Kp-fow-398	AGAGAGTCGACGAGTGGCAGAAAAGATCAGGTTGAGTACCTT
Kp-NotI-rev-521	AGAGAGCGGCCGCTTAAAAATTAAATTCTTTTGTTTGAAGGTT
Kp-rev-271-Stop	AGAGAGCGGCCGCTTAATCTGTCAAGTATGACAGAGCCCAAACAGT
Kp-rev-397-Stop	AGAGAGCGGCCGCTTATATTGTTAAGTTGCTGATTGCCCAAGCAGC
Kp-rev-442-Stop	AGAGAGCGGCCGCTTAGGCCATTATCAGAATGTTTTTTAGACCATC
Kp-SalI-fow-1	AGAGAGTCGACGATGGCCGAGAACCCCAGCTTGGAGAACCAC

RFP2

Primer Name	Sequence from 5' to 3'
RF-fow-132	AGAGAGTCGACGGAAGATGCCTATGCTCAGGAAAGGGATGCCTT
RF-fow-250	AGAGAGTCGACGCAGATGCAGGAGTTTAGAGAGAAAATCAAA
RF-fow-319	AGAGAGTCGACGTTGCTAATCCTTCTGCTTGGCCTTGTCATT
RF-fow-89	AGAGAGTCGACGCCCAAAATGCCAGTATGCAAAGGACACTTG
RF-NotI-rev-407	AGAGAGCGGCCGCTTATAATAGTTTATATTTGCACACAAATTC
RF-rev-131-Stop	AGAGAGCGGCCGCTTAAATAGAACAGAAGACATGTTTGGTGTGCTC
RF-rev-254-Stop	AGAGAGCGGCCGCTTACTCCTGCATCTGTTGCAGAAATACAATGGG
RF-rev-318-Stop	AGAGAGCGGCCGCTTATAACTTATAAAAGCTCCAGGGAATCTTGCT
RF-rev-341-Stop	AGAGAGCGGCCGCTTATAATGACCATTCTAGGAACATGGTAGGACC
RF-rev-57-Stop	AGAGAGCGGCCGCTTAGCATGTAGGACACTTGAATGGAGCTGGTCT
RF-rev-88-Stop	AGAGAGCGGCCGCTTAAGAGATCTTGATCTTGTTATACTTTTCCAC
RF-SalI-fow-1	AGAGAGTCGACGATGGAGCTGCTTGAAGAAGATCTCACATGC

RCBTB1

Primer Name	Sequence from 5' to 3'
C7-NotI-rev-531-Stop	AGAGAGCGGCCGCTTAGTTCTTAAAGGCTCCACATTTACTGGCTTT
C7-rev-300-Stop	AGAGAGCGGCCGCTTAGGCTGCAGACGTGTGGGCAGAGTGACAGGC
C7-rev-316-Stop	AGAGAGCGGCCGCTTACTGACCCCGGCACTGGCCCCACATGTACAC
C7-rev-369-Stop	AGAGAGCGGCCGCTTAAGTTTCTGGACTATCAAATTCTTTCTTCAG
C7-rev-467-Stop	AGAGAGCGGCCGCTTAAATTCCTCTCTTGATAATGTGCTGACAAAG
C7-rev-92-Stop	AGAGAGCGGCCGCTTATTCGGTGCTGAGAAGAACATGTGGTCCACT
C7-fow-301	AGAGACTCGAGGAAGACGCAGGGTGGGCACGTGTACATGTGG
C7-fow-370	AGAGACTCGAGGGCTGATCTGAAGTTTCGAATTGATGGAAAA
C7-fow-43	AGAGACTCGAGGGTCTTTGTATTTGGACTGAACTATAGCAAC
C7-fow-468	AGAGACTCGAGGACTGTGGAGAATGCCTTTTCGCTATTCTCT
C7-fow-93	AGAGACTCGAGGGATGGAGTGGTTTATGCCTGGGGCCACATGGG
C7-XhoI-fow-1	AGAGACTCGAGGATGGTGGATGTCGGAAAGTGGCCCATCTTC

Appendix 2

Cell lines used in this work

Indicated are the ordering numbers (DSMZ Nr.) in the German resource center for biological material.

HEK-293 DSMZ Nr. ACC 305
Also called 293, this is a human embryonic kidney cell line established by transformation of a human primary embryonic kidney with adenovirus type 5 (Graham et al. 1977). Culture medium: 90% DMEM + 10% FBS.

HeLa DSMZ Nr. ACC 57
Cell line derived from a human cervix carcinoma and established from the epitheloid cervix carcinoma of a patient (Scherer et al. 1953). Culture medium: 90% DMEM + 10% FBS.

Namalwa DSMZ Nr. ACC 24
Cell line derived from a human Burkitt lymphoma established from the tumor mass of a patient with Burkitt lymphoma (Nadkarni et al. 1969). Culture medium: 90% RPMI + 10% FBS.

Ramos DSMZ Nr. ACC 603
Cell line derived from a human Burkitt lymphoma established from the ascetic fluid of a patient with American-type Burkitt lymphoma (Klein et al. 1975). Culture medium: 80% RPMI + 20% FBS.

L1236 DSMZ Nr. ACC 530
Cell line derived from a human Hodgkin lymphoma established from the peripheral blood of a patient with Hodgkin lymphoma (Joos et al. 2003, Kanzler et al. 1996, Wolf et al. 1996). Culture medium: 90% RPMI + 10% FBS.

MedB-1 Kindly provided by Dr. Tews.
Cell line derived from a human primary mediastinal B-cell lymphoma (PMBL) established from the tumor mass of a patient with PMLB (Moller et al. 2001). Culture medium: 90% RPMI + 10% FBS.

KM-H2 Kindly provided by Prof. Zentgraf.
Cell line derived from a human Hodgkin lymphoma established from the pleural effusion of a patient with Hodgkin lymphoma (Kamesaki et al. 1986). Culture medium: 90% RPMI + 10% FBS.

SP 2/0 DSMZ Nr. ACC 146
Mouse myeloma cell line described to not synthesize or secrete immunoglobulin chains and to not grow in HAT medium (Shulman et al. 1978). Culture medium: 92,5% I-MEM + 7,5% FBS.

Appendix 3

Primers for RTS

KPNA3

Primer Name	Sequence
Kp-RTS-fow1	CTTTAAGAAGGAGATATACCATGGCCGAGAACCCAGC
Kp-RTS-fow2	CTTTAAGAAGGAGATATACCATGGCAGAAAATCCAAGTTTAGAGAACCACCGCATC
Kp-RTS-fow3	CTTTAAGAAGGAGATATACCATGGCAGAAAACCCATCATTAGAGAACCACCGCATC
Kp-RTS-fow4	CTTTAAGAAGGAGATATACCATGGCAGAAAACCCAAGTTTAGAGAACCACCGCATC
Kp-RTS-fow5	CTTTAAGAAGGAGATATACCATGGCAGAAAATCCATCATTAGAGAACCACCGCATC
Kp-RTS-fow6	CTTTAAGAAGGAGATATACCATGGCCGAAAATCCAAGTTTAGAGAACCACCGCATC
Kp-RTS-fow7	CTTTAAGAAGGAGATATACCATGGCAGAGAATCCAAGTTTAGAGAACCACCGCATC
Kp-RTS-fow8	CTTTAAGAAGGAGATATACCATGGCAGAAAATCCAAGTCTTGAGAACCACCGCATC
Kp-RTS-rev1	TGATGATGAGAACCCCCCCCCAAAATTAAATTCTTTTGTTT

RFP2

Primer Name	Sequence
Rf-RTS-fow1	CTTTAAGAAGGAGATATACCATGGAGCTGCTTGAAGAA
Rf-RTS-fow2	CTTTAAGAAGGAGATATACCATGGAATTACTTGAAGAAGATCTCAC
Rf-RTS-fow3	CTTTAAGAAGGAGATATACCATGGAATTACTGGAAGAAGATCTCACATG
Rf-RTS-fow4	CTTTAAGAAGGAGATATACCATGGAATTATTGGAAGAGGATCTCACATGCCCT
Rf-RTS-fow5	CTTTAAGAAGGAGATATACCATGGAATTATTGGAAGAAGATCTCACATG
Rf-RTS-fow6	CTTTAAGAAGGAGATATACCATGGAATTATTAGAAGAGGATCTCACATGCCCT
Rf-RTS-fow7	CTTTAAGAAGGAGATATACCATGGAATTACTGGAAGAGGATCTCACATGCCCT
Rf-RTS-fow8	CTTTAAGAAGGAGATATACCATGGAATTATTGGAGGAAGATCTCACATGCC
Rf-RTS-rev1	TGATGATGAGAACCCCCCCCTAATAGTTATATTTGCAC

C13orf1

Primer Name	Sequence
C6-RTS-fow1	CTTTAAGAAGGAGATATACCATGGCCACCTCGGTGTTG
C6-RTS-fow2	CTTTAAGAAGGAGATATACCATGGCTACTAGTGTTTTATGTTGCCTGCGGTGCTGC
C6-RTS-fow3	CTTTAAGAAGGAGATATACCATGGCTACTAGTGTATTATGTTGCCTGCGGTGCTGC
C6-RTS-fow4	CTTTAAGAAGGAGATATACCATGGCCACTAGTGTTTTATGTTGCCTGCGGTGCTGC
C6-RTS-fow5	CTTTAAGAAGGAGATATACCATGGCTACTTCAGTATTATGTTGCCTGCGGTGCTGC
C6-RTS-fow6	CTTTAAGAAGGAGATATACCATGGCCACTAGTGTATTATGTTGCCTGCGGTGCTGC
C6-RTS-fow7	CTTTAAGAAGGAGATATACCATGGCTACCTCAGTATTATGTTGCCTGCGGTGCTGC
C6-RTS-fow8	CTTTAAGAAGGAGATATACCATGGCTACAAGTGTATTATGTTGCCTGCGGTGCTGC
C6-RTS-rev1	TGATGATGAGAACCCCCCCCGAAGATTTGCTGTTCA

RCBTB1

Primer Name	Sequence
C7-RTS-fow1	CTTTAAGAAGGAGATATACCATGGTGGATGTCGGAAAG
C7-RTS-fow2	CTTTAAGAAGGAGATATACCATGGTTGATGTAGGAAAATGGCCCATCTTCACT
C7-RTS-fow3	CTTTAAGAAGGAGATATACCATGGTCGATGTAGGAAAATGGCCCATCTTCACT
C7-RTS-fow4	CTTTAAGAAGGAGATATACCATGGTTGATGTTGGAAAATGGCCCATCTTCACT
C7-RTS-fow5	CTTTAAGAAGGAGATATACCATGGTCGATGTTGGAAAATGGCCCATCTTCACT
C7-RTS-fow6	CTTTAAGAAGGAGATATACCATGGTTGATGTTGGTAAATGGCCCATCTTCACT
C7-RTS-fow7	CTTTAAGAAGGAGATATACCATGGTCGACGTAGGAAAATGGCCCATCTTCACT
C7-RTS-fow8	CTTTAAGAAGGAGATATACCATGGTTGATGTAGGTAAATGGCCCATCTTCACT
C7-RTS-rev1	TGATGATGAGAACCCCCCCCGTTCTTAAAGGCTCCA

SETDB2

Primer Name	Sequence
C8-RTS-fow1	CTTTAAGAAGGAGATATACCATGGGAGAAAAAAATGGCG
C8-RTS-fow2	CTTTAAGAAGGAGATATACCATGGGAGAAAAAAACGGTGATGCAAAAACTTTCTG
C8-RTS-fow3	CTTTAAGAAGGAGATATACCATGGGAGAAAAAAACGGCGATGCAAAAACT
C8-RTS-fow4	CTTTAAGAAGGAGATATACCATGGGAGAAAAAAATGGTGATGCAAAAACTTTCTG
C8-RTS-fow5	CTTTAAGAAGGAGATATACCATGGGAGAAAAAAACGGTGACGCAAAAACTTTCTGGA
C8-RTS-fow6	CTTTAAGAAGGAGATATACCATGGGAGAAAAGAACGGTGATGCAAAAACTTTCTG
C8-RTS-fow7	CTTTAAGAAGGAGATATACCATGGGAGAAAAAAATGGCGACGCAAAAACTTTCTGGA
C8-RTS-fow8	CTTTAAGAAGGAGATATACCATGGGAGAAAAGAATGGCGATGCAAAA
C8-RTS-rev1	TGATGATGAGAACCCCCCCCTAATATTTTTTTTCTACATTT

Appendix 4

Peptide sequences for KPNA3

Amino acid sequences of KPNA3 peptides synthesized for production of monoclonal antibodies. There is also a commercially available antibody from Abcam (goat polyclonal anti-KPNA3, ordering number ab6038). Sequence comparison showed that the peptide used for production of the commercially available antibody (underlined sequence) is a part of the Kp-CT peptide chosen in this work.

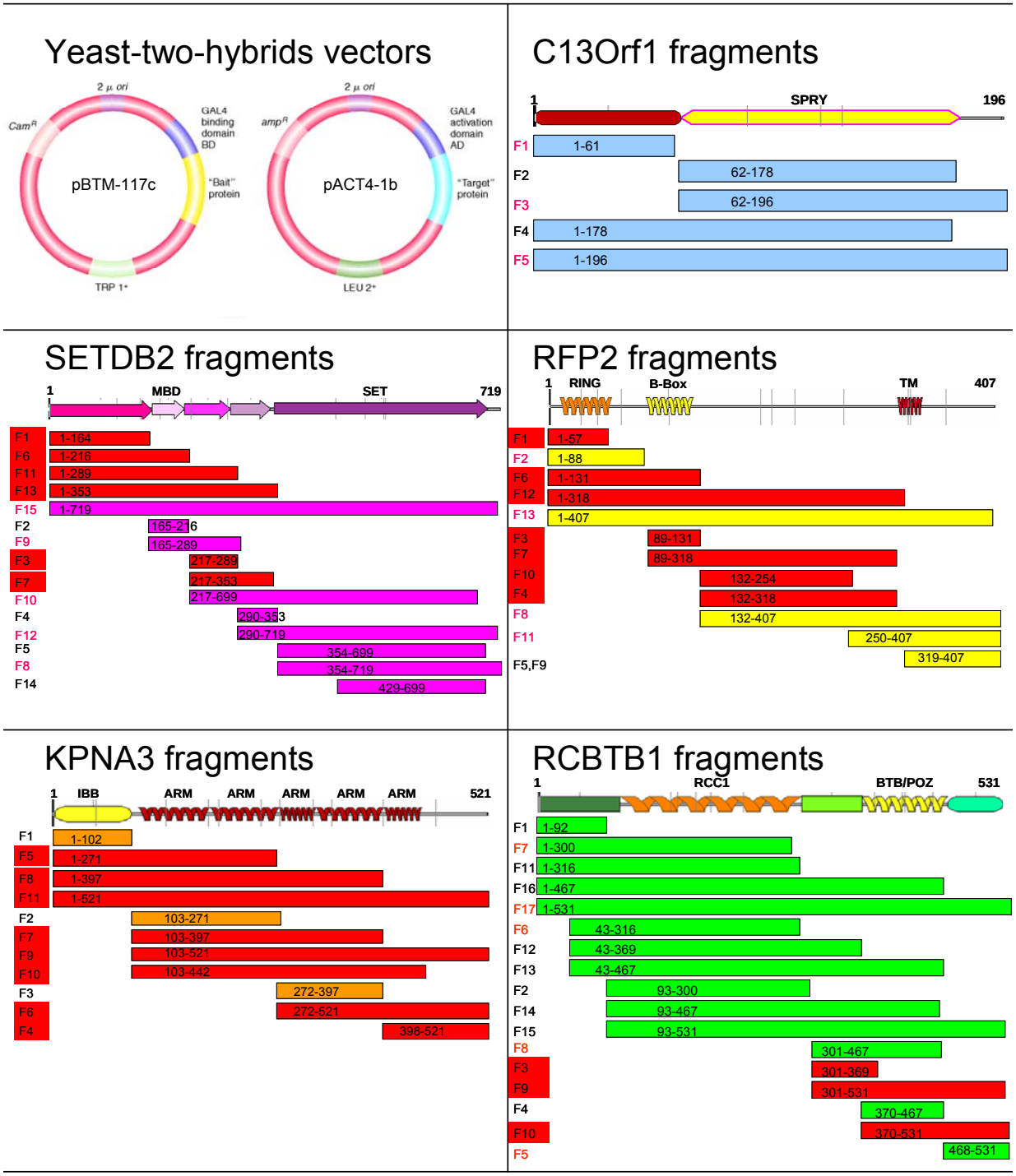
KPNA3 N-terminal peptide:
Kp-NT = MAENPSLENHRIKSFKNKGRDC

KPNA3 C-terminal peptide:
Kp-CT = CQGGTYNFDPTANLQTKEFNF

Appendix 5

Fragments of 13q14 candidate proteins used for Y2H screen

Schematic view of the vectors used for Y2H screen cloning and of the fragments of each 13q14 gene. Red numbers indicate fragments cloned into validation vectors, numbers in red boxes indicate auto-active fragments, numbers separated by dashes indicate amino-acids within the fragment. SPRY=SPRY domain; MBD= methyl binding domain; SET= SET domain; RING= RING finger domain; B-Box=B-Box domain; TM=transmembrane domain; IBB=importin Beta binding domain; ARM=armadillo repeat; RCC1=regulator of chromosome condensation domain; BTB/POZ=BTB/POZ domain.



Appendix 5 — Continued

Fragments of the different 13q14 genes used in the Y2H screen. Indicated are fragment number ("Fragment"), the insert size in base pairs ("Insert size") and the amino acids included in the fragment ("Amino Acids"). Fragment number followed by "FL" indicates fragment containing full length gene sequence.

RCBTB1		
Fragment	**Insert size**	**Amino Acids**
1	276	1-92
2	621	93-300
3	204	301-369
4	291	370-467
5	189	468-531
6	819	43-316
7	900	1-300
8	498	301-467
9	690	301-531
10	483	370-531
11	948	1-316
12	981	43-370
13	1272	43-467
14	1122	93-467
15	438	93-531
16	1401	1-467
17-**FL**	**1593**	**1-531**

SETDB2		
Fragment	**Insert size**	**Amino Acids**
1	492	1-164
2	153	165-216
3	216	217-289
4	189	290-353
5	1035	354-699
6	648	1-216
7	408	217-353
8	1095	354-719
9	375	165-289
10	1446	217-699
11	867	1-289
12	1287	290-719
13	1059	1-353
14	630	429-639
15-**FL**	**2157**	**1-719**

C13Orf1		
Fragment	**Insert size**	**Amino Acids**
1	183	1-61
2	348	62-178
3	402	62-196
4	534	1-178
5-**FL**	**588**	**1-196**

RFP2		
Fragment	**Insert size**	**Amino Acids**
1	171	1-57
2	264	1-88
3	123	89-131
4	561	131-318
5	264	319-407
6	390	1-131
7	687	89-318
8	823	131-407
9	66	319-341
10	369	131-254
11	471	250-407
12	954	1-318
13-**FL**	**1221**	**1-407**

KPNA3		
Fragment	**Insert size**	**Amino Acids**
1	306	1-102
2	504	103-271
3	375	272-397
4	369	398-521
5	813	1-271
6	747	272-521
7	882	103-397
8	1191	1-397
9	1254	103-521
10	1017	103-442
11-**FL**	**1563**	**1-521**

Appendix 6

Candidate gene expression in relation to Lamin B1

Real-time PCR analysis of 13q14 candidate genes compared to the house-keeping gene Lamin B1. The left part of the table contains the absolute measurements of mRNA (ng). The right part of the table contains the ratio for each gene to the corresponding value of the house-keeping gene (Lamin B1). Values were considered to indicate that the gene was expressed at more or less the same rate (around 1), much less (around 0) or at higher rate (>1) than the house-keeping gene. Orange marked values (*) indicate that in those cell lines expression is around that of the house-keeping gene or higher. Cell lines used for immunofluorescence are marked with a green arrow-head and those used to make lysates for Western blot analysis and immunoprecipitation with a red arrow-head. Hematopoietic cell lines are marked with a violet line and central nervous system cells with a blue line.

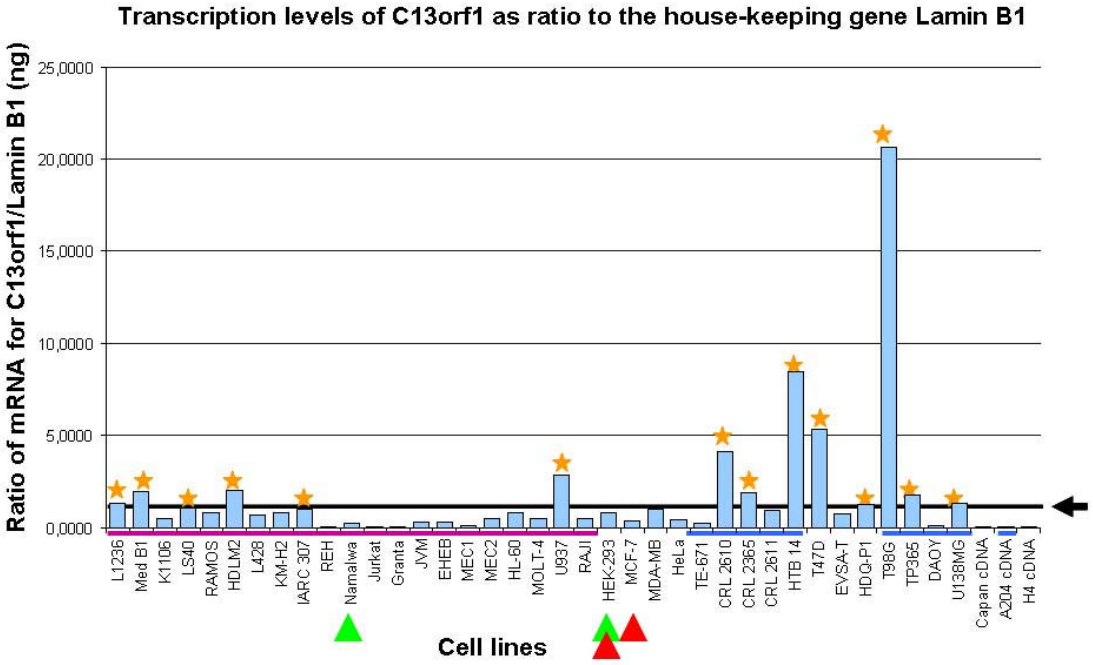
Cell lines	C13Orf1	RCBTB1	RFP2	SETDB2	LaminB1
L1236	0,3829	0,0307	0,2095	0,1779	0,2946
Med B1	0,5786	0,2583	0,5099	0,3903	0,2978
K1106	0,3124	0,0486	0,3967	0,4654	0,6662
LS40	0,3188	0,0255	0,2706	0,2385	0,3009
RAMOS	1,1012	0,1529	0,4710	0,8761	1,3776
HDLM2	0,8387	0,0871	0,4112	0,1938	0,4122
L428	0,3488	0,0169	0,1224	0,1713	0,5450
KM-H2	0,1745	0,0712	0,1382	0,0834	0,2308
IARC 307	0,4884	0,0636	0,2743	0,5695	0,4884
REH	0,0425	0,1442	0,2265	0,2339	1,4325
Namalwa	1,9382	2,2309	0,6939	3,4562	7,7819
Jurkat	0,1002	0,0810	0,4592	0,3105	4,9821
Granta	0,1188	0,4860	0,4456	1,0210	5,8882
JVM	0,5451	0,0470	0,2690	0,3300	2,1408
EHEB	0,0587	0,0014	0,0815	0,0498	0,2300
MEC1	0,1216	0,0441	0,1444	0,3144	1,3172
MEC2	0,0590	0,0029	0,0783	0,0581	0,1227
HL-60	0,0138	0,0005	0,0110	0,0104	0,0168
MOLT-4	0,3957	0,0559	0,5279	0,0812	0,8907
U937	0,1896	0,0081	0,1857	0,7596	0,0672
RAJI	0,1304	0,0045	0,1435	0,1194	0,2797
HEK-293	1,1710	0,1820	0,6657	0,5276	1,4771
MCF-7	0,1951	0,0220	0,1920	0,1323	0,5725
MDA-MB	0,3550	0,0148	0,2562	0,0631	0,3577
HeLa	0,0097	0,0003	0,0035	0,0029	0,0256
TE-671	1,1675	0,8371	1,3461	0,3036	5,2075
CRL 2610	0,4237	0,0205	0,2553	0,3393	0,1018
CRL 2365	0,6072	0,0761	0,3150	0,3624	0,3309
CRL 2611	0,3873	0,0273	0,2191	0,1588	0,4208
HTB 14	0,1701	0,0102	0,1025	0,0434	0,0203
T47D	0,2255	0,0108	0,2210	0,0797	0,0424
EVSA-T	0,0901	0,0007	0,0759	0,0555	0,1334
HDQ-P1	0,1794	0,0166	0,4087	0,1648	0,1461
T98G	0,6554	0,0351	0,2342	0,3571	0,0317
TP365	0,2175	0,0077	0,0986	0,1273	0,1246
DAOY	0,2488	0,0374	0,5085	0,1898	3,2360
U138MG	0,6427	0,0195	0,2286	0,1979	0,5036
Capan cD	0,0497	0,0289	0,0555	0,1124	0,8453
A204 cDNA	0,1064	0,0762	0,1362	0,1584	6,5533
H4 cDNA	0,0675	0,0321	0,0793	0,1085	3,8307

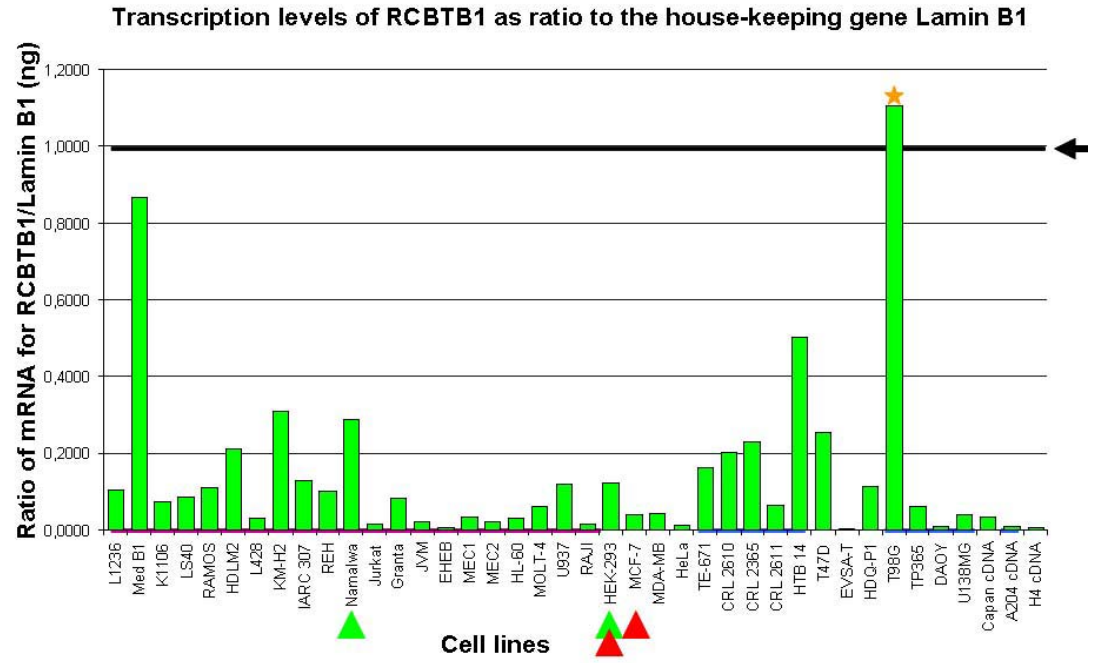
		Ratio zu LaminB1			
Cell line nr.	Cell lines	C13ORF1	RCTBT1	RFP2	SETDB2
1	L1236	1,2998	0,1042	0,7111	0,60403526
2	Med B1	1,9431	0,8676	1,7124	1,31068596
3	K1106	0,4689	0,0729	0,5955	0,69854847
4	LS40	1,0595	0,0848	0,8993	0,79282879
5	RAMOS	0,7993	0,1110	0,3419	0,63596645
6	HDLM2	2,0349	0,2112	0,9977	0,47018867
7	L428	0,6400	0,0310	0,2245	0,31423483
8	KM-H2	0,7561	0,3087	0,5989	0,36138355
9	IARC 307	1,0000	0,1302	0,5616	1,16610632
10	REH	0,0297	0,1006	0,1581	0,16326708
11	Namalwa	0,2491	0,2867	0,0892	0,44413339
12	Jurkat	0,0201	0,0163	0,0922	0,06232856
13	Granta	0,0202	0,0825	0,0757	0,17340387
14	JVM	0,2546	0,0219	0,1257	0,15416366
15	EHEB	0,2555	0,0062	0,3546	0,21674411
16	MEC1	0,0923	0,0335	0,1096	0,23867186
17	MEC2	0,4808	0,0234	0,6377	0,47371213
18	HL-60	0,8182	0,0311	0,6554	0,61844972
19	MOLT-4	0,4443	0,0627	0,5927	0,09113535
20	U937	2,8224	0,1212	2,7653	11,3093271
21	RAJI	0,4663	0,0159	0,5131	0,42699493
22	HEK-293	0,7928	0,1232	0,4507	0,35719956
23	MCF-7	0,3407	0,0385	0,3354	0,23105272
24	MDA-MB	0,9925	0,0413	0,7164	0,17641791
25	HeLa	0,3786	0,0127	0,1370	0,11221085
26	TE-671	0,2242	0,1607	0,2585	0,05829404
27	CRL 2610	4,1606	0,2015	2,5069	3,33229553
28	CRL 2365	1,8347	0,2299	0,9519	1,09522663
29	CRL 2611	0,9205	0,0649	0,5207	0,37742595
30	HTB 14	8,3958	0,5012	0,5627	2,14366008
31	T47D	5,3213	0,2540	5,2141	1,88111286
32	EVSA-T	0,6752	0,0052	0,5692	0,41595264
33	HDQ-P1	1,2276	0,1138	2,7963	1,12769902
34	T98G	20,6657	1,1053	7,3848	11,261559
35	TP365	1,7451	0,0620	0,7908	1,02133594
36	DAOY	0,0769	0,0116	0,1571	0,05864848
37	U138MG	1,2762	0,0387	0,4540	0,39295495
38	Capan cDNA	0,0588	0,0341	0,0656	0,13294508
39	A204 cDNA	0,0162	0,0116	0,0208	0,02417071
40	H4 cDNA	0,0176	0,0084	0,0207	0,02831793

Appendix 6

Candidate gene expression in relation to Lamin B1

Real-time PCR analysis of 13q14 candidate genes compared to the house-keeping gene Lamin B1. The values in the graphic are obtained dividing the absolute measurements of mRNA (ng) for each gene in each cell line by the absolute measurements of mRNA (ng) for the house-keeping gene in that same cell line. Values are considered to indicate that the 13q14 gene is expressed at more or less the same rate (around 1 (black bolded line and arrow)), much less (around 0) or at higher rate (>1) than the house-keeping gene. Orange marked values (orange star) indicate that in those cell lines expression is around that of the house-keeping gene or higher. Cell lines used for immunofluorescence are marked with a green arrow-head and those used to make lysates for Western blot analysis and immunoprecipitation with a red arrow-head. Hematopoietic cell lines are underlined with a violet line and central nervous system cells with a blue line.

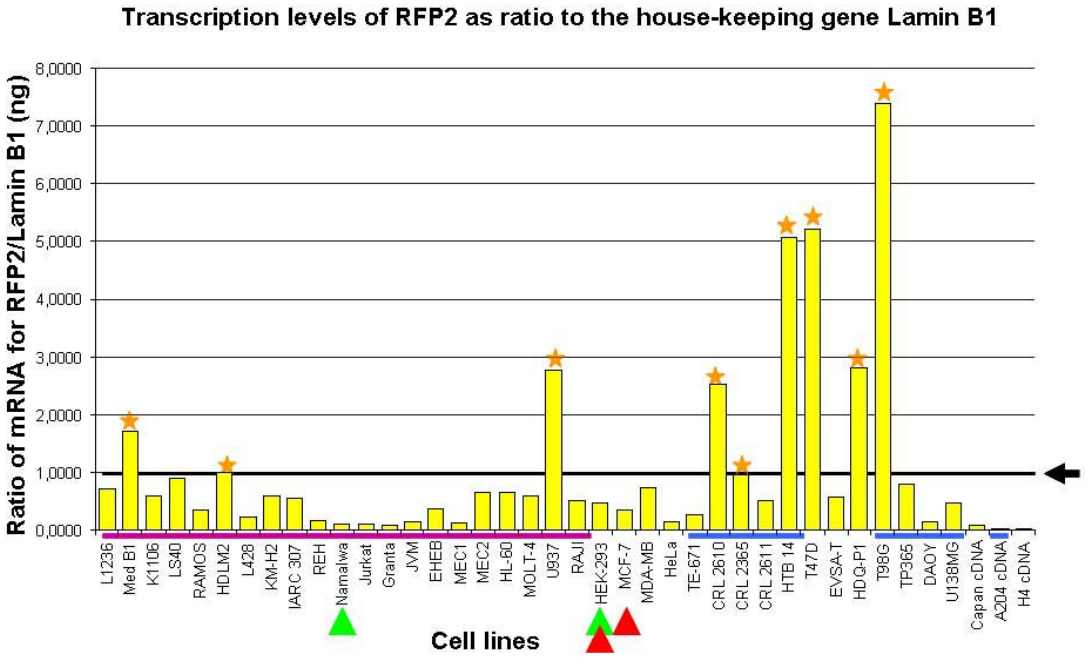


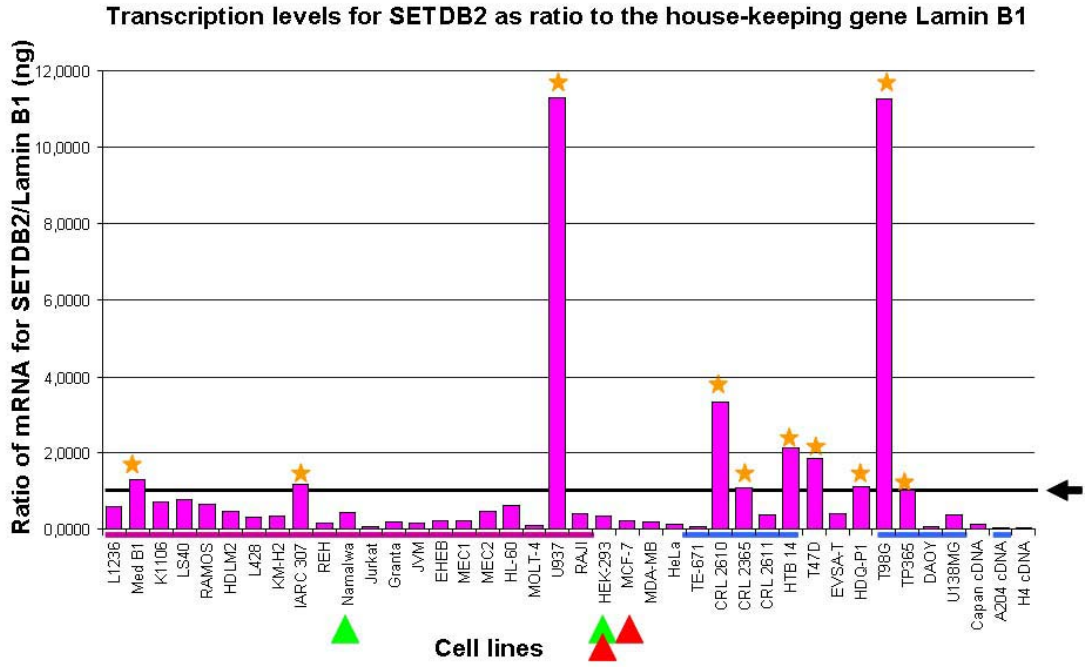


Appendix 6

Candidate gene expression in relation to Lamin B1

Real-time PCR analysis of 13q14 candidate genes compared to the house-keeping gene Lamin B1. The values in the graphic are obtained dividing the absolute measurements of mRNA (ng) for each gene in each cell line by the absolute measurements of mRNA (ng) for the house-keeping gene in that same cell line. Values are considered to indicate that the 13q14 gene is expressed at more or less the same rate (around 1 (black bolded line and arrow)), much less (around 0) or at higher rate (>1) than the house-keeping gene. Orange marked values (orange star) indicate that in those cell lines expression is around that of the house-keeping gene or higher. Cell lines used for immunofluorescence are marked with a green arrow-head and those used to make lysates for Western blot analysis and immunoprecipitation with a red arrow-head. Hematopoietic cell lines are underlined with a violet line and central nervous system cells with a blue line.





Appendix 6 — Continued

Candidate gene expression in relation to PGK1

RT-PCR analysis of 13q14 candidate genes compared to the house-keeping gene PGK1. The left part of the table contains the absolute measurements of mRNA (ng). The right part of the table contains the ratio for each gene to the corresponding value of the house-keeping gene (PGK=PGK1). Values were considered to indicate that the gene was expressed at more or less the same rate (around 1), much less (around 0) or at higher rate (>1) than the house-keeping gene. Orange marked values (*) indicate that in those cell lines expression is around that of the house-keeping gene or higher. Cell lines used for immunofluorescence are marked with a green arrow-head and those used to make lysates for Western blot analysis and immunoprecipitation with a red arrow-head. Hematopoietic cell lines are marked with a violet line and central nervous system cells with a blue line.

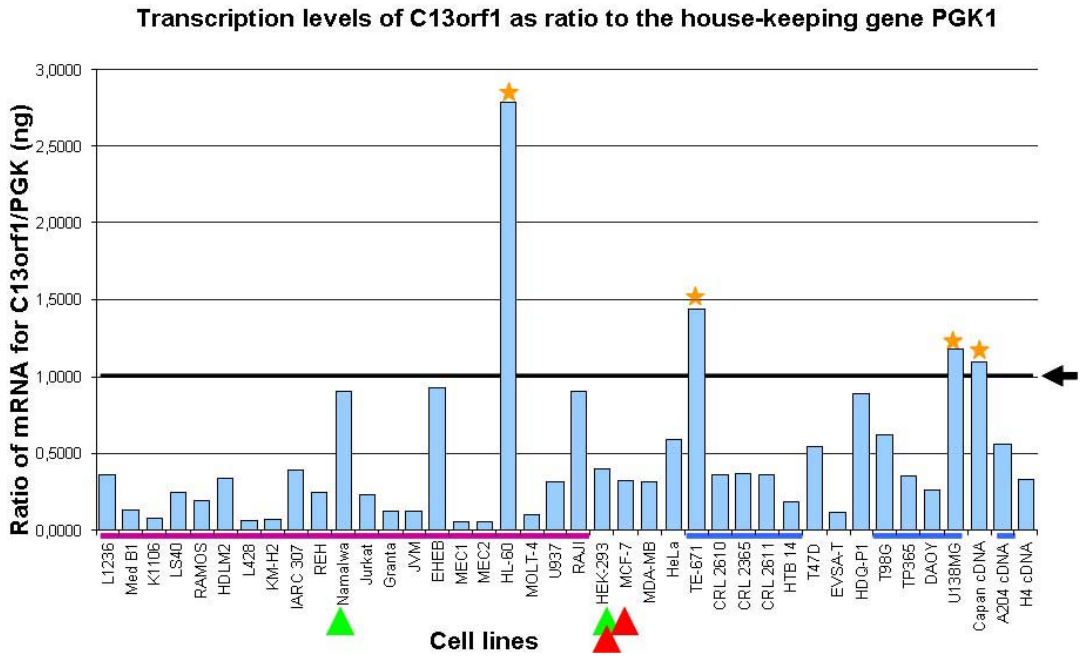
Cell lines	C13Orf1	RCBTB1	RFP2	SETDB2	PGK
L1236	0,3829	0,0307	0,2095	0,1779	1,0525
Med B1	0,5786	0,2583	0,5099	0,3903	4,5856
K1106	0,3124	0,0486	0,3967	0,4654	4,0894
LS40	0,3188	0,0255	0,2706	0,2385	1,3129
RAMOS	1,1012	0,1529	0,4710	0,8761	5,7829
HDLM2	0,8387	0,0871	0,4112	0,1938	2,4800
L428	0,3488	0,0169	0,1224	0,1713	5,7482
KM-H2	0,1745	0,0712	0,1382	0,0834	2,5504
IARC 307	0,4884	0,0636	0,2743	0,5695	1,2569
REH	0,0425	0,1442	0,2265	0,2339	0,1754
Namalwa	1,9382	2,2309	0,6939	3,4562	2,1485
Jurkat	0,1002	0,0810	0,4592	0,3105	0,4318
Granta	0,1188	0,4860	0,4456	1,0210	0,9724
JVM	0,5451	0,0470	0,2690	0,3300	4,4781
EHEB	0,0587	0,0014	0,0815	0,0498	0,0637
MEC1	0,1216	0,0441	0,1444	0,3144	2,4335
MEC2	0,0590	0,0029	0,0783	0,0581	1,0246
HL-60	0,0138	0,0005	0,0110	0,0104	0,0050
MOLT-4	0,3957	0,0559	0,5279	0,0812	4,0111
U937	0,1896	0,0081	0,1857	0,7596	0,6125
RAJI	0,1304	0,0045	0,1435	0,1194	0,1440
HEK-293	1,1710	0,1820	0,6657	0,5276	2,9667
MCF-7	0,1951	0,0220	0,1920	0,1323	0,6051
MDA-MB	0,3550	0,0148	0,2562	0,0631	1,1551
HeLa	0,0097	0,0003	0,0035	0,0029	0,0163
TE-671	1,1675	0,8371	1,3461	0,3036	0,8106
CRL 2610	0,4237	0,0205	0,2553	0,3393	1,1854
CRL 2365	0,6072	0,0761	0,3150	0,3624	1,6554
CRL 2611	0,3873	0,0273	0,2191	0,1588	1,0654
HTB 14	0,1701	0,0102	0,1025	0,0434	0,9286
T47D	0,2255	0,0108	0,2210	0,0797	0,4163
EVSA-T	0,0901	0,0007	0,0759	0,0555	0,7689
HDQ-P1	0,1794	0,0166	0,4087	0,1648	0,2017
T98G	0,6554	0,0351	0,2342	0,3571	1,0556
TP365	0,2175	0,0077	0,0986	0,1273	0,6196
DAOY	0,2488	0,0374	0,5085	0,1898	0,9606
U138MG	0,6427	0,0195	0,2286	0,1979	0,5459
Capan cD	0,0497	0,0289	0,0555	0,1124	0,0455
A204 cDNA	0,1064	0,0762	0,1362	0,1584	0,1930
H4 cDNA	0,0675	0,0321	0,0793	0,1085	0,2029

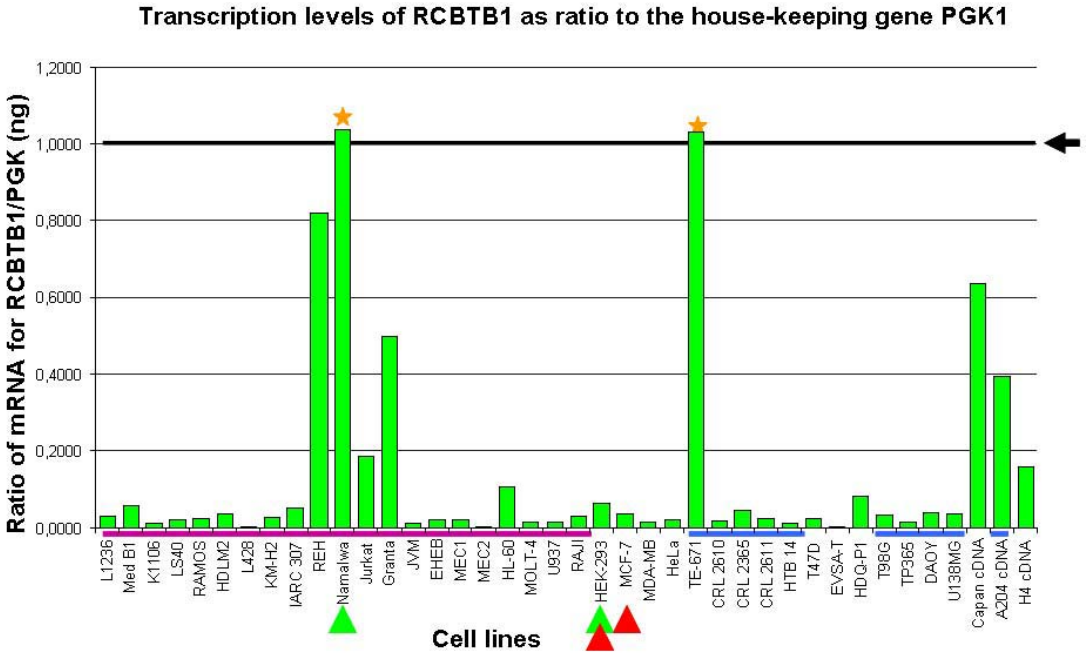
		Ratio zu PGK			
Cell line nr.	Cell lines	C13ORF1	RCTBT1	RFP2	SETDB2
1	L1236	0,3638	0,0292	0,1990	0,1691
2	Med B1	0,1262	0,0563	0,1112	0,0851
3	K1106	0,0764	0,0119	0,0970	0,1138
4	LS40	0,2428	0,0194	0,2061	0,1817
5	RAMOS	0,1904	0,0264	0,0815	0,1515
6	HDLM2	0,3382	0,0351	0,1658	0,0781
7	L428	0,0607	0,0029	0,0213	0,0298
8	KM-H2	0,0684	0,0279	0,0542	0,0327
9	IARC 307	0,3886	0,0506	0,2182	0,4531
10	REH	0,2422	0,8217	1,2913	1,3332
11	Namalwa	0,9021	1,0383	0,3230	1,6086
12	Jurkat	0,2321	0,1876	1,0635	0,7191
13	Granta	0,1222	0,4998	0,4583	1,0500
14	JVM	0,1217	0,0105	0,0601	0,0737
15	EHEB	0,9228	0,0225	1,2809	0,7830
16	MEC1	0,0500	0,0181	0,0593	0,1292
17	MEC2	0,0576	0,0028	0,0764	0,0567
18	HL-60	2,7782	0,1058	2,2255	2,1001
19	MOLT-4	0,0987	0,0139	0,1316	0,0202
20	U937	0,3095	0,0133	0,3032	1,2402
21	RAJI	0,9058	0,0309	0,9967	0,8295
22	HEK-293	0,3947	0,0613	0,2244	0,1778
23	MCF-7	0,3224	0,0364	0,3173	0,2186
24	MDA-MB	0,3073	0,0128	0,2218	0,0546
25	HeLa	0,5927	0,0198	0,2144	0,1756
26	TE-671	1,4402	1,0326	1,6606	0,3745
27	CRL 2610	0,3574	0,0173	0,2154	0,2863
28	CRL 2365	0,3668	0,0460	0,1903	0,2189
29	CRL 2611	0,3636	0,0256	0,2057	0,1491
30	HTB 14	0,1831	0,0109	0,1104	0,0468
31	T47D	0,5417	0,0259	0,5308	0,1915
32	EVSA-T	0,1171	0,0009	0,0988	0,0722
33	HDQ-P1	0,8895	0,0825	2,0262	0,8171
34	T98G	0,6209	0,0332	0,2219	0,3383
35	TP365	0,3510	0,0125	0,1591	0,2054
36	DAOY	0,2591	0,0389	0,5294	0,1976
37	U138MG	1,1774	0,0357	0,4188	0,3625
38	Capan cDNA	1,0924	0,6340	1,2186	2,4688
39	A204 cDNA	0,5513	0,3951	0,7058	0,8208
40	H4 cDNA	0,3329	0,1584	0,3911	0,5347

Appendix 6 — Continued

Candidate gene expression in relation to PGK1

RT-PCR analysis of 13q14 candidate genes compared to the house-keeping gene PGK1. The values in the graphic are obtained dividing the absolute measurements of mRNA (ng) for each gene in each cell line by the absolute measurements of mRNA (ng) for the house-keeping gene in that same cell line. Values are considered to indicate that the 13q14 gene is expressed at more or less the same rate (around 1 (black bolded line and arrow)), much less (around 0) or at higher rate (>1) than the house-keeping gene. Orange marked values (orange star) indicate that in those cell lines expression is around that of the house-keeping gene or higher. Cell lines used for immunofluorescence are marked with a green arrow-head and those used to make lysates for Western blot analysis and immunoprecipitation with a red arrow-head. Hematopoietic cell lines are underlined with a violet line and central nervous system cells with a blue line.

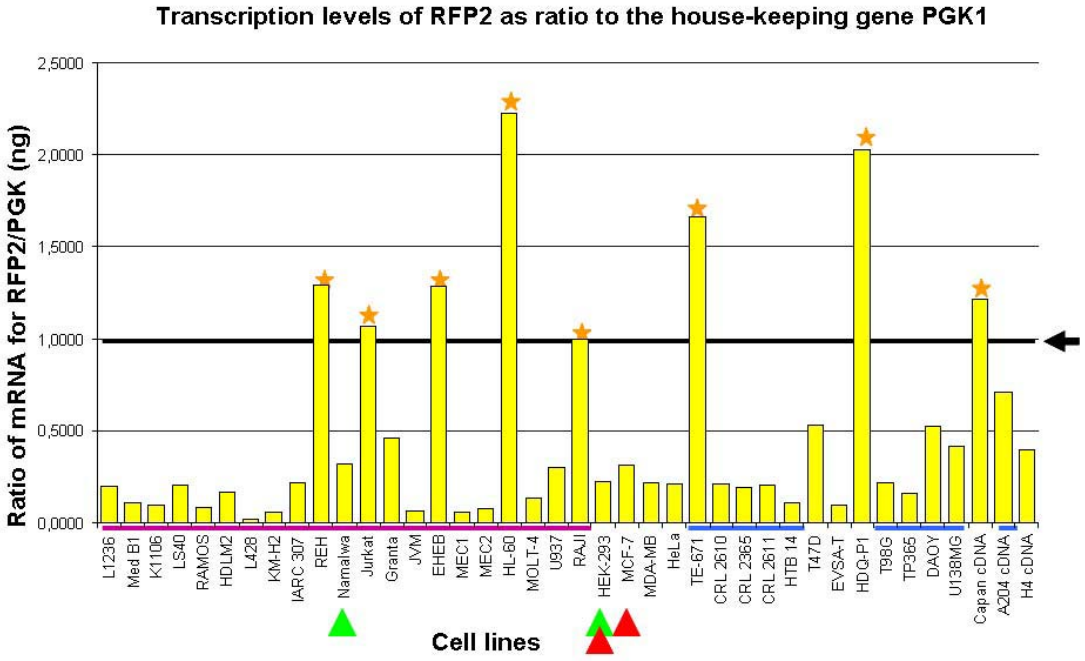


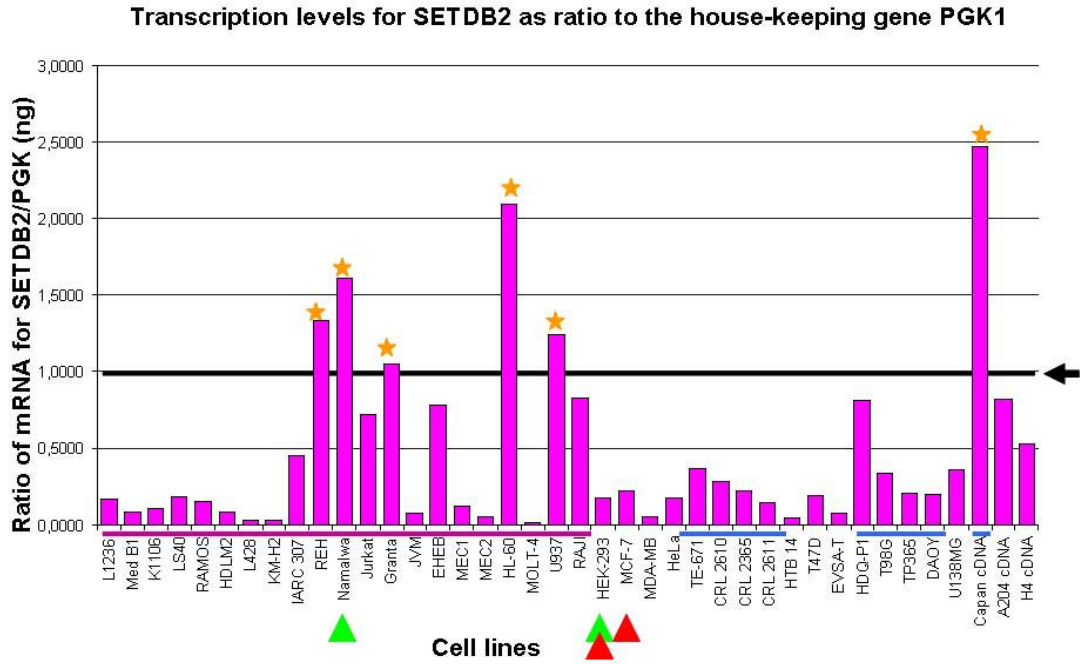


Appendix 6 — Continued

Candidate gene expression in relation to PGK1

RT-PCR analysis of 13q14 candidate genes compared to the house-keeping gene PGK1. The values in the graphic are obtained dividing the absolute measurements of mRNA (ng) for each gene in each cell line by the absolute measurements of mRNA (ng) for the house-keeping gene in that same cell line. Values are considered to indicate that the 13q14 gene is expressed at more or less the same rate (around 1 (black bolded line and arrow)), much less (around 0) or at higher rate (>1) than the house-keeping gene. Orange marked values (orange star) indicate that in those cell lines expression is around that of the house-keeping gene or higher. Cell lines used for immunofluorescence are marked with a green arrow-head and those used to make lysates for Western blot analysis and immunoprecipitation with a red arrow-head. Hematopoietic cell lines are underlined with a violet line and central nervous system cells with a blue line.





Appendix 7

Interaction partners of 13q14 candidate genes from the Y2H screen

Interaction partners of the 13q14 gene fragment involved in the interaction are given in the rows and columns, respectively. Numbers in the fields indicate how many times the interaction was reproduced in the screen in the triplicate experiments. Empty cells depict no interaction detected. Interaction partners listed in white boxes indicate HA-tag constructs were used for validation. Interaction partners listed in orange boxes indicate FLAG-tag constructs were used for validation. Interaction partners listed in red or light-blue boxes indicate that interaction was not analyzed. C6=C13orf1, C7=RCBTB1, C8=SETDB2, RF=RFP2, KP=KPNA3.

RCBTB1

Symbol	C7-F1	C7-F11	C7-F13	C7-F15	C7-F2	C7-F4	C7-F5	C7-F6	C7-F8
ABHD3									
ACP5							1		
ACTC									
ACVR1									
ASB13		1							
C11orf2									
C14orf1		2	1				1		2
C14orf32		1							
C17orf25							1		
C17orf45									
COL4A2	1	3					1		1
CRIP2		1	1	1					1
CRMP1			1				1		3
DBI									
DBN1									
EED		2	2						2
EEF1A1							1		2
EEF1G									
EXT2							1		
FLJ22471									
FLJ31413		1				1	1		
FLNA		1							
GNB2		1	1				1		1
HKR3		1			1				
HNRPA2B1							2		
IDH2								1	
KIAA1377	1	1					2		
KPNA1									
MGC33584		1							
MICAL1									
NR2F1		1							
NR3C1		1					3		
PHF17									
PICK1									
PITPNA									
PLXNA2									
PTDSS2									
PTN		1					3		
RAI1			1						
RASSF1							3		
RIP							1		
RPL13									
RXRA		1					1		2
SERPINA4	1		1						
SETDB1		2	1						1
SPRED1									
STARD9									
SUMO2		1				1			

Appendix 7 — Continued

C13Orf1

Symbol	C6-F1	C6-F2	C6-F3	C6-F4	C6-F5
ABHD3					
ACP5	1				
ACTC					
ACVR1					
ASB13					
C11orf2					
C14orf1					
C14orf32					
C17orf25		1			
C17orf45					
COL4A2			1		
CRIP2					
CRMP1		1			
DBI					
DBN1					
EED					
EEF1A1		1			
EEF1G					
EXT2					
FLJ22471					
FLJ31413		1			
FLNA					
GNB2					
HKR3			1	1	
HNRPA2B1					
IDH2					
KIAA1377					
KPNA1	1				
MGC33584		1			
MICAL1					
NR2F1					
NR3C1		1			
PHF17		1			
PICK1					
PITPNA					
PLXNA2					
PTDSS2					
PTN		1			
RAI1				1	
RASSF1					
RIP		1		1	
RPL13					
RXRA		1	1		
SERPINA4					1
SETDB1	1				
SPRED1					
STARD9					
SUMO2					
TLR9					
TMOD3					
TRA2A	1			1	
TRIM28					
ZNF266				1	

Appendix 7 — Continued

	SETDB2					KPNA3		RFP2	
Symbol	C8-F10	C8-F15	C8-F4	C8-F5	C8-F9	KP-F1	KP-F3	RF-F13	RF-F8
ABHD3						2			
ACP5						2			
ACTC					2	1			
ACVR1	1								
ASB13					1	3			
C11orf2		1			2				
C14orf1					1				
C14orf32									
C17orf25	1				1				
C17orf45					1				
COL4A2					2				
CRIP2	1				2				
CRMP1				1					
DBI	1					3			
DBN1						2			
EED						2			
EEF1A1									
EEF1G	1					1			
EXT2					1				
FLJ22471						3			
FLJ31413	1				2				
FLNA						2			
GNB2					1	1			
HKR3	2				1	1			
HNRPA2B1					1				
IDH2					1				
KIAA1377	1				1	1			
KPNA1									1
MGC33584					2				
MICAL1						1			
NR2F1					1				
NR3C1					2		1		
PHF17					2				
PICK1	2								
PITPNA					2				
PLXNA2						1			
PTDSS2	1								
PTN					2				
RAI1	3								
RASSF1	3	1							
RIP	2		1		1				
RPL13					1	2			
RXRA					2	1			
SERPINA4	1					1		1	
SETDB1		1			1	1			
SPRED1	1		2						
STARD9					1				
SUMO2					3	1			
TLR9						1			
TMOD3						3			
TRA2A	1					2			
TRIM28						2			
ZNF266	2				1				

6. Bibliography

Auer, and, SR, Cotter, FE 2007: The 13q and 11q B-cell chronic lymphocytic leukaemia-associated regions derive from a common ancestral region in the zebrafish. British journal of haematology 137, 443-453.

Baranova, A et al. 2003: Distinct organization of the candidate tumor suppressor gene RFP2 in human and mouse: multiple mRNA isoforms in both species- and human-specific antisense transcript RFP2OS. Gene 321, 103-12.

Baranova, A, Ivanov, D, Tyazhelova, T, Yankovsky, N 2004: Structural and functional characterization of human chromosome 13q14 and its potential tumor suppressor genes. Molekuliarnaia biologiia 38**(2)**, 165-173.

Bismuth, G, Boumsell, L 2002: Controlling the Immune System Through Semaphorins. . Science's STKE 128, 1-4.

Blayney, LM, Zissimopoulos, S, Ralph, E, Abbot, E, Matthews, L, Lai, FA 2004: Ryanodine receptor oligomeric interaction: identification of a putative binding region. The Journal of biological chemistry 279**(15)**, 14639-48.

Blumenthal, T 1998: Gene clusters and polycistronic transcription in eukaryotes. BioEssays 20, 480–487.

Breiling, A, Orlando, V 2002: SET domain proteins reSET gene expression. Nature structural biology 9**(12)**, 894-6.

Bullrich, F, Veronese, ML, Kitada, S, Jurlander, J, Caligiuri, MA, Reed, JC, Croce, CM 1996: Minimal region of loss at 13q14 in b-cell chronic lymphocytic leukemia. Blood 88**(8)**, 3109-3115.

Burnette, WN 1981: "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Analytical biochemistry 112**(2)**, 195-203.

Calin, CDD, Masayoshi Shimizu*, Roberta Bichi*, Simona Zupo†, Evan Noch*,, Hansjuerg Aldler*, SR, Michael Keating‡, Kanti Rai§, Laura Rassenti¶, Thomas Kipps¶,, Massimo Negrini*, FB, and Carlo M. Croce 2002: Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proceedings of the National Academy of Sciences of the United States of America 99**(24)**, 15524-15529.

Carter, A, K., L, P.D., S, A.R., P 2004: Detection of p53 dysfunction by flow cytometry in chronic lymphocytic leukemia. British journal of haematology 127, 425-428.

Chadwick, BP, Willard, HF 2003: SETting the stage. Eed-Enx1 leaves an epigenetic signature on the inactive X chromosome. Developmental cell 4**(4)**, 445-7.

Chambon, P 1996: A decade of molecular biology of retinoic acid receptors. Faseb J 10**(9)**, 940-54.

Chen, JZ, Huang, SD, Ji, CN, Pang, RY, Xie, Y, Xue, JL 2005: Identification, expression pattern, and subcellular location of human RIP isoforms. DNA and cell biology 24**(7)**, 464-9.

Chiorazzi, N, Allen, SL, Ferrarini, M 2005: Clinical and laboratory parameters that define clinically relevant B-CLL subgroups. Current topics in microbiology and immunology 294, 109-33.

Cimmino, Calin, Fabbri, Iorio, Ferracin, Shimizu, Wojcik, Aqeilan, Zupo, Dono, Rassenti, Alder, Volinia, Liu, Kipps, Negrini, Croce, CM 2005: miR-15 and miR-16 induce apoptosis by targeting BCL2. Proceedings of the National Academy of Sciences of the United States of America 102**(39)**, 13944-13949.

Corcoran, M 2004: DLEU2 encodes an antisense RNA for the putative bicistronic RFP2/LEU5 gene in humans and mouse. Genes Chromosomes Cancer 40, 285-297.

Crespo, P, León, J 2000: Ras proteins in the control of the cell cycle and cell differentiation. CMLS, Cell. Mol. Life Sci. 57, 1613-1636.

Damle, TarunWasil, FF, Fabio Ghiotto, Angelo Valetto, Steven L. Allen, Aby Buchbinder,, Daniel Budman, KD, Jonathan Kolitz, Stuart M. Lichtman, Philip Schulman, Vincent P. Vinciguerra,, Kanti R. Rai, MF, and Nicholas Chiorazzi 1999: Ig V Gene Mutation Status and CD38 Expression As Novel Prognostic Indicators in Chronic Lymphocytic Leukemia. Blood Vol 94**(6)**, 1840-1847.

Day, GJ, Mosteller, RD, Broek, D 1998: Distinct subclasses of small GTPases interact with guanine nucleotide exchange factors in a similar manner. Molecular and cellular biology 18**(12)**, 7444-54.

Deo, RC, Schmidt, EF, Ehabazi, A, Togashi, H, Burley, SK, Strittmatter, SM 2004: Structural bases for CRMP function in plexin-dependent semaphorin3A signaling. The EMBO Journal 23**(1)**, 9-22.

DePuy, SD, Yao, J, Hu, C, McIntire, W, Bidaud, I, Lory, P, Rastinejad, F, Gonzalez, C, Garrison, JC, Barrett, PQ 2006: The molecular basis for T-type Ca2+ channel inhibition by G protein beta2gamma2 subunits. Proceedings of the National Academy of Sciences of the United States of America 103**(39)**, 14590-5.

Döhner, H, Fischer, K, Bentz, M, Hansen, K, Benner, A, Cabot, G, Diehl, D, Schlenk, R, Coy, J, Stilgenbauer, S, et al. 1995: p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. Blood 85**(6)**, 1580-9.

Döhner, H, Stilgenbauer, S, Benner, A, Leupolt, E, Krober, A, Bullinger, L, Döhner, K, Bentz, M, Lichter, P 2000: Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 343**(26)**, 1910-6.

Döhner, H, Stilgenbauer, S, Döhner, K, Bentz, M, Lichter, P 1999: Chromosome aberrations in B-cell chronic lymphocytic leukemia: reassessment based on molecular cytogenetic analysis. Journal of molecular medicine (Berlin, Germany) 77**(2)**, 266-81.

Downes, GB, Gautam, N 1999: The G protein subunit gene families. Genomics 62**(3)**, 544-52.

Fagerlund, R, Kinnunen, L, Kohler, M, Julkunen, I, Melen, K 2005: NF-{kappa}B is transported into the nucleus by importin {alpha}3 and importin {alpha}4. The Journal of biological chemistry 280**(16)**, 15942-51.

Falco, GD, Giordano, aA 2006: pRb2/p130: a new candidate for retinoblastoma tumor formation. Blood 25, 5333-5340.

Gaidano, G, Balleruni, P 1991: p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocitic leukemia. Proceedings of the National Academy of Sciences of the United States of America 88**(12)**, 5413-5417.

Goehler, H et al. 2004: A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. Molecular cell 15**(6)**, 853-65.

Graham, FL, Smiley, J, Russell, WC, Nairn, R 1977: Characteristics of a human cell line transformed by DNA from human adenovirus type 5. The Journal of general virology 36**(1)**, 59-74.

Griffin, TJ, Steven P. Gygi, Trey Ideker, Beate Rist, Jimmy Eng, Leroy Hood, Aebersold, aR 2002: Complementary Profiling of Gene Expression at the Transcriptome and Proteome Levels in Saccharomyces cerevisiae. Molecular & Cellular Proteomics 1, 323–333.

Guo, DF, Tardif, V, Ghelima, K, Chan, JS, Ingelfinger, JR, Chen, X, Chenier, I 2004: A novel angiotensin II type 1 receptor-associated protein induces cellular hypertrophy in rat vascular smooth muscle and renal proximal tubular cells. The Journal of biological chemistry 279**(20)**, 21109-20.

Gygi, SP, YVAN ROCHON, B. ROBERT FRANZA, AEBERSOLD, AR 1999: Correlation between Protein and mRNA Abundance in Yeast. MOLECULAR AND CELLULAR BIOLOGY 19**(3)**, 1720–1730.

Hamblin, TJ, Zadie Davis, Anne Gardiner, David G. Oscier, Stevenson, aFK 1999: Unmutated Ig VH Genes Are Associated With a More Aggressive Form of Chronic Lymphocytic Leukemia. Blood Vol 94 **(6)**, 1848-1854.

Hammarsund, M, Corcoran, MM, Wilson, W, Zhu, C, Einhorn, S, Sangfelt, O, Grander, D 2004: Characterization of a novel B-CLL candidate gene--DLEU7--located in the 13q14 tumor suppressor locus. FEBS letters 556**(1-3)**, 75-80.

Hanahan, D, Weinberg, aRA 2000: The Hallmarks of Cancer. Cell 100, 57–70.

Holland, EC, Sonenberg, N, Pandolfi, PP, Thomas, G 2004: Signaling control of mRNA translation in cancer pathogenesis. Oncogene 23, 3138-3144.

Huh, WK, Falvo, JV, Gerke, LC, Carroll, AS, Howson, RW, Weissman, JS, O'Shea, EK 2003: Global analysis of protein localization in budding yeast. Nature 425**(6959)**, 686-91.

Ivanov, D, Tyazhelova, T, Lemonnier, L, Kononenko, N, Pestova, A, Nikitin, E, Prevarskaya, N, Skyma, R, Panchin, YV, Yankovsky, N, Baranova, A 2003: A new human gene KCNRG encoding a potassium channel regulating protein is a cancer suppressor gene candidate located in 13q14.3. FEBS letters 539, 156-160.

Jaffe, HNL, Stein 2001: World Health Organization Classification of Tumours. Vardiman (Eds), Lyon.

Jäkel, S, Mingot, JM, Schwarzmaier, P, Hartmann, E, Görlich, D 2002: Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. EMBO 21**(3)**, 377-386.

Janeway, C, Travers, P, Walport, M, Shlomchik, M 2004: Immunobiology: The Immune System in Health and Disease - Sixth Edition. Garlan Science**(ISBN: 0-8153-4101-6)**.

Jenuwein, T 2001: Re-SET-ting heterochromatin by histone methyltransferases. Trends in cell biology 11**(6)**, 266-73.

Jones, KA, Fitzgibbon, J, Woodward, KJ, Goudie, D, Ferguson-Smith, MA, Povey, S, Wolfe, J, Solomon, E 1993: Localization of the retinoid X receptor alpha gene (RXRA) to chromosome 9q34. Annals of human genetics 57**(Pt 3)**, 195-201.

Joos, S, Granzow, M, Holtgreve-Grez, H, Siebert, R, Harder, L, Martin-Subero, JI, Wolf, J, Adamowicz, M, Barth, TF, Lichter, P, Jauch, A 2003: Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including REL and JAK2. International journal of cancer 103**(4)**, 489-95.

Juliusson, Oscier DG, Fitchett M 1990: Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. N Engl J Med 323, 720-724.

Kalachikov, S et al. 1997: Cloning and gene mapping of the chromosome 13q14 region deleted in chronic lymphocytic leukemia. Genomics 42**(3)**, 369-377.

Kamesaki, H, Fukuhara, S, Tatsumi, E, Uchino, H, Yamabe, H, Miwa, H, Shirakawa, S, Hatanaka, M, Honjo, T 1986: Cytochemical, immunologic, chromosomal, and molecular genetic analysis of a novel cell line derived from Hodgkin's disease. Blood 68**(1)**, 285-92.

Kanzler, H, Hansmann, ML, Kapp, U, Wolf, J, Diehl, V, Rajewsky, K, Kuppers, R 1996: Molecular single cell analysis demonstrates the derivation of a peripheral blood-derived cell line (L1236) from the Hodgkin/Reed-Sternberg cells of a Hodgkin's lymphoma patient. Blood 87**(8)**, 3429-36.

Kapanazade, B, Makeeva, N, Corcoran, M, Jareborg, Hammarsund, M, Baranova, A, Zabarovsky, E, Vorontsova, Merup, Gahrton, G, Jansson, Yankovsky, N, Einhorn, S, DG, O, Grander, D, Sangfelt, O 2000: Comparative sequence analysis of a region on human chromosome 13q14, frequently deleted in B-cell chronic lymphocytic leukemia, and its homologous region on mouse chromosome 14. Genomics 70, 327-334.

Kim, SY, Levenson, JM, Korsmeyer, S, Sweatt, JD, Schumacher, A 2007: Developmental regulation of Eed complex composition governs a switch in global histone modification in brain. The Journal of biological chemistry 282**(13)**, 9962-72.

Kimura, F, Suzu, S, Nakamura, Y, Nakata, Y, Yamada, M, Kuwada, N, Matsumura, T, Yamashita, T, Ikeda, T, Sato, K, Motoyoshi, K 2003: Cloning and characterization of a novel RING-B-box-Coiled-coil protein with apoptotic function. The journal of biological chemistry 278**(27)**, 25046-25054.

Kino, T, Tiulpakov, A, Ichijo, T, Chheng, L, Kozasa, T, Chrousos, GP 2005: G protein beta interacts with the glucocorticoid receptor and suppresses its transcriptional activity in the nucleus. The Journal of cell biology 169**(6)**, 885-96.

Klein, G, Giovanella, B, Westman, A, Stehlin, JS, Mumford, D 1975: An EBV-genome-negative cell line established from an American Burkitt lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by in vitro infection. Intervirology 5**(6)**, 319-34.

Kleppe, K, Ohtsuka, E, Kleppe, R, Molineux, I, Khorana, HG 1971: Studies on polynucleotides. XCVI. Repair replications of short synthetic DNA's as catalyzed by DNA polymerases. Journal of molecular biology 56**(2)**, 341-61.

Knudson, AG 2001: Two genetic hits (more or less) to cancer. Nature Reviews Cancer 1, 157-162.

Kornberg, RD, Thomas, J 1974: Chromatin Structure: Oligomers of the Histones. Science, 865-868.

Kraus, M, Alimzhanov, MB, Rajewsky, N, Rajewsky, K 2004: Survival of resting mature B lymphocytes depends on BCR signaling via the Igalpha/beta heterodimer. Cell 117**(6)**, 787-800.

Küppers, R, KLEIN, U, HANSMANN, ML, AND, RAJEWSKY, K 1999: Cellular origins of B-cell lymphomas. The New England Journal of Medicine, 1520-1529.

Küppers, R 2005: Mechanisms of B-cell lymphoma patogenesis. Nature Reviews Cancer 5, 251-262.

Lamberti, A, Caraglia, M, Longo, O, Marra, M, Abbruzzese, A, Aracari, P 2004: The translation elongation factor 1A in tumorigenesis, signal transduction and apoptosis: Review Article. Amino Acids 26, 443-448.

Lane, DP 1992: Cancer. p53, guardian of the genome. Nature 358**(6381)**, 15-6.

Lalonde, JP, Lim, R, Ingley, E, Tilbrook, PA, Thompson, MJ, McCulloch, R, Beaumont, JG, Wicking, C, Eyre, HJ, Sutherland, GR, Howe, K, Solomon, E, Williams, JH, Klinken, SP 2004: HLS5, a novel RBCC (Ring finger, B Box, Coiled-coil) family member isolated from a hemopoietic lineage switch, is a candidate tumor suppressor. The journal of biological chemistry 279**(9)**, 8181-8189.

Lam, KP, Kuhn, R, Rajewsky, K 1997: In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. Cell 90**(6)**, 1073-83.

Lercher, MJ, Urrutia, AO, Hurst, LD 2002: Clustering of housekeeping genes provides a unified model of gene order in the human genome. Nature genetics 31, 180-183.

Lerner, M, Corcoran, M, Cepeda, D, Nielsen, ML, Zubarev, R, Ponten, F, Uhlen, M, Hober, S, Grander, D, Sangfelt, O 2007: The RBCC Gene RFP2 (Leu5) Encodes a Novel Transmembrane E3 Ubiquitin Ligase Involved in ERAD. Mol Biol Cell 18**(5)**, 1670-1682.

Lessard, J, Armin Schumacher, Unnur Thorsteinsdottir, Maarten van Lohuizen, Magnuson, T, Sauvageau, aG 1999: Functional antagonism of the Polycomb-Group genes eed and Bmi1 in hemopoietic cell proliferation. Genes & Dev. 13, 2691-2703.

Liu, Y, Szekely, L, Grander, D, Soderhall, S, Julliusson, G, Gahrton, G, Linder, S, Inhorn, S 1993: Chronic lymphocytic leukemia cells with allelic deletions at 13q14 commonly have one intact RB1 gene: evidence for a role of an adjacent locus. Proceedings of the National Academy of Sciences of the United States of America 90, 8697-8701.

Lynda M. Blayney, Spyros Zissimopoulos, Emma Ralph, Eleanor Abbot, and, LM, Lai., FA 2004: Ryanodine receptor oligomeric interaction: identification of a putative binding region. JBC Papers in Press.

Mabuchi, H, Fujii, H, Calin, G, Alder, H, Negrini, M, Rassenti, L, Kipps, TJ, Bullrich, F, Croce, CM 2001: Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. Cancer Res 61**(7)**, 2870-7.

Matsuda, A, Suzuki, Y, Honda, G, Muramatsu, S, Matsuzaki, O, Nagano, Y, Doi, T, Shimotohno, K, Harada, T, Nishida, E, Hayashi, H, Sugano, S 2003: Large-scale identification and characterization of human genes that activate NF-kappaB and MAPK signaling pathways. Oncogene 22**(21)**, 3307-18.

Mertens, D, Wolf, S, Bullinger, L, Ohl, S, Schaffner, C, Döhner, H, Stilgenbauer, S, Lichter, P 2000: BCMSUN, a candidate gene for B-cell chronic lymphocytic leukemia and mantle-cell lymphoma, has an independently expressed homolog on 1p22-p31, BCMSUN-like. International journal of cancer 88**(5)**, 692-7.

Mertens, D, Wolf, S, Schroeter, P, Schaffner, C, Döhner, H, Stilgenbauer, S, Lichter, P 2002: Down-regulation of candidate tumor suppressor genes within chromosome band 13q14.3 is independent of the DNA methylation pattern in B-cell chronic lymphocytic leukemia. Blood 99**(11)**, 4116-21.

Mertens, D, Wolf, S, Tschuch, C, Mund, C, Kienle, D, Ohl, S, Schroeter, P, Lyko, F, Döhner, H, Stilgenbauer, S, Lichter, P 2006: Allelic silencing at the tumor-suppressor locus 13q14.3 suggests an epigenetic tumor-suppressor mechanism. Proceedings of the National Academy of Sciences of the United States of America 103**(20)**, 7741-6.

Migliazza, A, Francesc Bosch, Hirokazu Komatsu, Eftihia Cayanis, Stefano Martinotti, Elena Toniato, Ernesto Guccione, Xiaoyan Qu, Minchen Chien, V. V. V. Murty, Gianluca Gaidano, Giorgio Inghirami, Peisen Zhang, Stuart Fischer, Sergey M. Kalachikov, James Russo, Isidore Edelman, Argiris Efstratiadis, and Riccardo Dalla-Favera 2001: Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. Blood 97**(7)**, 2098-2104.

Moller, P, Bruderlein, S, Strater, J, Leithauser, F, Hasel, C, Bataille, F, Moldenhauer, G, Pawlita, M, Barth, TF 2001: MedB-1, a human tumor cell line derived from a primary mediastinal large B-cell lymphoma. International journal of cancer 92**(3)**, 348-53.

Mullis, K, Faloona, F, Scharf, S, Saiki, R, Horn, G, Erlich, H 1986: Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor symposia on quantitative biology 51 Pt 1, 263-73.

Nadkarni, JS, Nadkarni, JJ, Clifford, P, Manolov, G, Fenyo, EM, Klein, E 1969: Characteristics of new cell lines derived from Burkitt lymphomas. Cancer 23**(1)**, 64-79.

Niedenthal, RK, Riles, L, Johnston, M, Hegemann, JH 1996: Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. Yeast (Chichester, England) 12**(8)**, 773-86.

O'Brien, KP, Tapia-Paez, I, Stahle-Backdahl, M, Kedra, D, Dumanski, JP 2000: Characterization of five novel human genes in the 11q13-q22 region. Biochemical and biophysical research communications 273**(1)**, 90-4.

Olins, AL., Olins DE. 1974. Spheroid chromatin units (v bodies). Science **183**(122): 330-2.

Olson, MOJ, Dundr, M, Szebeni, A 2000: The nucleolus: an old factory with unexpected capabilites. Trends Cell Biol. 10, 189-196.

Overbeck, AF, Brtva, TR, Cox, AD, Graham, SM, Huff, SY, Khosravi-Far, R, Quilliam, LA, Solski, PA, Der, CJ 1995: Guanine nucleotide exchange factors: activators of Ras superfamily proteins. Molecular reproduction and development 42**(4)**, 468-76.

Park, J, Seo, T, Kim, H, Choe, J 2005: Sumoylation of the novel protein hRIP{beta} is involved in replication protein A deposition in PML nuclear bodies. Molecular and cellular biology 25**(18)**, 8202-14.

Popescu, NC 2000: Comprehensive genetic analysis of cancer cells. J.Cell.Mol.Med. 4**(3)**, 151-163.

Quilliam, LA, Khosravi-Far, R, Huff, SY, Der, CJ 1995: Guanine nucleotide exchange factors: activators of the Ras superfamily of proteins. Bioessays 17**(5)**, 395-404.

Rajewsky, K 1996: Clonal selection and learning in the antibody system. Nature 381**(6585)**, 751-8.

Rea, S, Eisenhaber, F, O'Carroll, D, Strahl, B, Sun, Z, Schmid, M, Opravil, S, Mechtler, K, Ponting, C, Allis, D, Jenuwein, T 2000: Regulation of chormatin structure by site-specific histone H3 methyltransferases. Nature 406, 593-599.

Rehm, H 2002: Der Experimentator: Proteinbiochemie / Proteomics. 4. Auflage. Spektrum Heidelberg: Akad. Verlag 2002.

Renart, J, Reiser, J, Stark, GR 1979: Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. Proceedings of the National Academy of Sciences of the United States of America 76**(7)**, 3116-20.

Renault, L, Nassar, N, Vetter, I, Becker, J, Klebe, C, Roth, M, Wittinghofer, A 1998: The 1.7 A crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. Nature 392**(6671)**, 97-101.

Reymond, A, Maroni, G, Fantozzi, A, Ballabio, A 2001: The tripartite motif family identifies cell compartments. EMBO 20**(9)**, 2140-2151.

Richie, ER, Schumacher, A, Angel, JM, Holloway, M, Rinchik, EM, Magnuson, T 2002: The Polycomb-group gene eed regulates thymocyte differentiation and suppresses the development of carcinogen-induced T-cell lymphomas. Oncogene 21**(2)**, 299-306.

Ripolles, L, Margarita Ortega, FO, Ana Gonzalez, JL, Jesus Ojangurene, Joan Alfons Solerf,, Juan Berguag, MDC, Marıa Rosa Caballın 2006: Genetic abnormalities and clinical outcome in chronic lymphocytic leukemia. Cancer genetics and cytogenetics 171, 57-64.

Rosa, JL, Casaroli-Marano, RP, Buckler, AJ, Vilaró, S, Barbacid, M 1996: p619, a giant protein related to the chromosome condensation regulator RCC1,

stimulates guanine nucleotide exchange on ARF1 and Rab proteins. EMBO 15**(16)**, 4262-4273.

Rowntree, C, Duke, V, Panayiotidis, P, Kotsi, P, Palmisano, G, Hoffbrand, A, Foroni, L 2002: Deletion analysis of chromosome 13q14.3 and chatacterization of an alternative splice from of LEU1 in B cell chronic lymphocytic leukemia. Leukemia 16, 1267-1275.

Sanger, F, Nicklen, S, Coulson, AR 1977: DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74**(12)**, 5463-7.

Schaffner, C, Stilgenbauer, S, Lichter, P 1999: Somatic ATM mutations indicate a pathogenic role of the ATM in B-cell chronic lymphocytic leukemia. Blood 94**(2)**, 748-753.

Schagger, H, von Jagow, G 1987: Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Analytical biochemistry 166**(2)**, 368-79.

Scherer, WF, Syverton, JT, Gey, GO 1953: Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. The Journal of experimental medicine 97**(5)**, 695-710.

Schotta, G, Lachner, M, Sarma, K, Ebert, A, Sengupta, R, Reuter, G, Reinberg, D, Jenuwein, aT 2004: A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. GENES & DEVELOPMENT 18, 1251–1262.

Shih, JY, Lee, YC, Yang, SC, Hong, TM, Huang, CY, Yang, PC 2003: Collapsin response mediator protein-1: a novel invasion-suppressor gene. Clinical & experimental metastasis 20**(1)**, 69-76.

Shi-Hua Zhang, Xue-Mei Ning, Hong-Wei Liu, Zhang, aX-S 2007: Prediction of Protein Complexes Based on Protein Interaction Data and Functional Annotation Data Using Kernel Methods. Submitted.

Shulman, M, Wilde, CD, Kohler, G 1978: A better cell line for making hybridomas secreting specific antibodies. Nature 276**(5685)**, 269-70.

Schultz, DC, Ayyanathan, K, Negorev, D, Maul, GG, Rauscher, FJ, 3rd 2002: SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes & development 16**(8)**, 919-32.

Simpson, J, Wellenreuther, R, Poustka, A, Pepperkok, R, Wiemann, S 2000: Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing. EMBO 1**(3)**, 287-292.

Solomou, EE, Sfikakis, PP, Kotsi, P, Papaioannou, M, Karali, V, Vervessou, E, Hoffbrand, AV, Panayiotidis, P 2003: 13q deletion in chronic lymphocytic leukemia: characterization of E4.5, a novel chromosome condensation regulator-like guanine nucleotide exchange factor. Leukemia & lymphoma 44**(9)**, 1579-85.

Spirin, AS, Baranov, VI, Ryabova, LA, Ovodov, SY, Alakhov, YB 1988: A continuous cell-free translation system capable of producing polypeptides in high yield. Science (New York, N.Y 242**(4882)**, 1162-4.

Sproul, D, Gilbert, N, Bickmore, aWA 2005: THE ROLE OF CHROMATIN STRUCTURE IN REGULATING THE EXPRESSION OF CLUSTERED GENES. Nature Reviews genetics 6, 775-781.

Staden, R 1996: The Staden sequence analysis package. Molecular biotechnology 5**(3)**, 233-41.

Stelzl, U et al. 2005: A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome. Cell 122, 957–968.

Stilgenbauer, L Bullinger, Lichter, P, Döhner, aH, (GCLLSG), atGCSG 2002: Genetics of chronic lymphocytic leukemia: genomic aberrations and V_H gene mutation status in pathogenesis and clinical course. Leukemia 16, 993-1007.

Stilgenbauer, S, Döhner, K, Bentz, M, Lichter, P, Döhner, H 1998: Molecular cytogenetic analysis of B-cell chronic lymphocytic leukemia. Annals of hematology 76**(3-4)**, 101-10.

Stilgenbauer, S, Leupolt, E, Ohl, S, Weiss, G, Schroder, M, Fischer, K, Bentz, M, Lichter, P, Dohner, H 1995: Heterogeneity of deletions involving RB-1 and the D13S25 locus in B-cell chronic lymphocytic leukemia revealed by fluorescence in situ hybridization. Cancer Res 55**(16)**, 3475-7.

Stilgenbauer, S 2006: Clonal evolution in chronic lymphocytic leukemia: aquisition of high-risk genomic aberrations associated with unmutated vh, resistance to therapy, and short survival. Submitted.

Takashi Ito, Tomoko Chiba, Ritsuko Ozawa, Mikio Yoshida, Masahira Hattori, Sakaki, aY 2001: A comprehensive two-hybrid analysis to explore the yeast protein interactome. PNAS 98**(8)**, 4569–4574.

Takeda, S, Fujiwara, T, Shimizu, F, Kawai, A, Shinomiya, K, Okuno, S, Ozaki, K, Katagiri, T, Shimada, Y, Nagata, M, Watanabe, T, Takaichi, A, Kuga, Y, Suzuki, M, Hishigaki, H, Takahashi, E, Shin, S, Nakamura, Y, Hirai, Y 1997: Isolation and mapping of karyopherin alpha 3 (KPNA3), a human gene that is highly homologous to genes encoding Xenopus importin, yeast SRP1 and human RCH1. Cytogenetics and cell genetics 76**(1-2)**, 87-93.

Talapatra, S, Wagner, J, Thompson, aC 2002: Elongation factor-1 alpha is a selective regulator of growth factor withdrawal and ER stress-induced apoptosis. Cell Death and Differentiation 9, 856-861.

Talcott, B, Moore, MS 2000: The nuclear import of RCC1 requires a specific nuclear localization sequence receptor, karyopherin alpha3/Qip. The Journal of biological chemistry 275**(14)**, 10099-104.

Thornton, S, ·, NA, Purcell, D, Lee, J 2003: Not just for housekeeping: protein initiation and elongation factors in cell growth and tumorigenesis. Journal of molecular medicine 81, 536-548.

Tiwari, M, Kumar, A, Sinha, RA, Shrivastava, A, Balapure, AK, Sharma, R, Bajpai, VK, Mitra, K, Babu, S, Godbole, MM 2006: Mechanism of 4-HPR-induced apoptosis in glioma cells: evidences suggesting role of mitochondrial-mediated pathway and endoplasmic reticulum stress. Carcinogenesis 27**(10)**, 2047-58.

Tschuch, C 2006: Functional characterization of candidate genes in the pathogenesis of B-cell chronic lymphocytic leukemia (B-CLL). Inaugural-Dissertation.

van Everdink, WJ, Baranova, A, Lummen, C, Tyazhelova, T, Looman, MW, Ivanov, D, Verlind, E, Pestova, A, Faber, H, van der Veen, AY, Yankovsky, N, Vellenga, E, Buys, CH 2003: RFP2, C13orf1, and FAM10A4 are the most likely tumor suppressor gene candidates for B-cell chronic lymphocytic leukemia. Cancer genetics and cytogenetics 146**(1)**, 48-57.

Vidalain, P-O, Boxem, M, Hui Ge, Siming Li, Vidal, aM 2004: Increasing specificity in high-throughput yeast two-hybrid experiments. Methods 32 363–370.

Weidemann, W, Stelzl, U, Lisewski, U, Bork, K, Wanker, EE, Hinderlich, S, Horstkorte, R 2006: The collapsin response mediator protein 1 (CRMP-1) and the promyelocytic leukemia zinc finger protein (PLZF) bind to UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase (GNE), the key enzyme of sialic acid biosynthesis. . FEBS Letters 580, 6649-6654.

Weinberg, RA 2007: The biology of Cancer. Garland Science.

Wolf, J, Kapp, U, Bohlen, H, Kornacker, M, Schoch, C, Stahl, B, Mucke, S, von Kalle, C, Fonatsch, C, Schaefer, HE, Hansmann, ML, Diehl, V 1996: Peripheral blood mononuclear cells of a patient with advanced Hodgkin's lymphoma give rise to permanently growing Hodgkin-Reed Sternberg cells. Blood 87**(8)**, 3418-28.

Wolf, S, Mertens, D, Schaffner, C, Korz, C, Döhner, H, Stilgenbauer, S, Lichter, P 2001: B-cell neoplasia associated gene with multiple splicing (BCMS): the candidate B-CLL gene on 13q14 comprises more than 560 kb covering all critical regions. Human molecular genetics 10**(12)**, 1275-85.

Woo, J, Imm, J-H, Min, C-K, Kim, K-J, Cha, S-S, Oh, B-H 2006: Structural and functional insights into the B30.2/SPRY domain. EMBO 25, 1353-1363.

Xu, L, Wei, Y, Reboul, J, Vaglio, P, Shin, TH, Vidal, M, Elledge, SJ, Harper, JW 2003: BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. Nature 425**(6955)**, 316-21.

Zanotti, R, Ambrosetti, A, Lestani, M, Ghia, P, Pattaro, C, Remo, A, Zanetti, F, Stella, S, Perbellini, O, Prato, G, Guida, G, Caligaris-Cappio, F, Menestrina, F, Pizzolo, G, Chilosi, M 2007: ZAP-70 expression, as detected by immunohistochemistry on bone marrow biopsies from early-phase CLL patients, is a strong adverse prognostic factor. Leukemia 21**(1)**, 102-9.

Zeng, J-Z, Sun, D-F, Wang, L, Cao, X, Qi, J-B, Yang, T, Hu, C-Q, Liu, W, Zhang, X-K 2006: *Hypericum sampsonii* induces apoptosis and nuclear export of retinoid X receptor-alpha. Carcinogenesis 27**(10)**, 1991-2000.

Zubay, G 1973: In vitro synthesis of protein in microbial systems. Annual review of genetics 7, 267-87.

Summary

Among all types of leukemia, B-cell chronic lymphocytic leukemia (B-CLL) is the most common leukemia in adults in the Western countries. The most frequent genomic aberration in B-CLL is deletion of the critical chromosome region 13q14, which is deleted in over 50% of patients. The high frequency of the deletion in B-CLL and its presence in other types of cancer points to the involvement of at least one of the genes from this region in a tumor suppressor mechanism. To elucidate the possible role of the protein-coding genes of the region in the pathomechanism associated with 13q14, the possible functions of the five candidate proteins C13orf1, RCBTB1, SETDB2, RFP2 and KPNA3 were analyzed in this work. To this end, the intracellular localization of the proteins as well as their interaction partners was studied. The intracellular localization of the proteins was analyzed in human cell lines, via proteins over-expressed with a GFP fusion tag and by immunohistochemistry using specific antibodies produced for the study of the endogenous proteins. The localization proved to be in accordance with their assigned putative functions, based on homology to well characterized proteins. Interestingly, there were some differences in the localization of some of the 13q14 candidate proteins between a lymphoma cell line and an embryonic kidney cell line of non-malignant origin, suggesting that changes in the intracellular localization that affect protein function could play a role in malignancy. Towards elucidation of the candidate proteins' functions, the identification of interaction partners for 13q14 candidate proteins was performed by a yeast-two-hybrid (Y2H) screen. A pre-screen revealed interaction among 4/5 of the 13q14 candidate proteins and the screen with a library of over 5000 human proteins resulted in 53 candidate interaction partners. Overlap of interactions was of special interest, since the homology analysis of the proteins suggested coexistence in macromolecular complexes. The analysis of the Y2H screen data revealed that 36% of the identified partners interacted with three or more 13q14 proteins and this, together with the interactions of 13q14 proteins with each other, further supports the hypothesis of functional relations among the analyzed 13q14 proteins. Validation was performed by pull-down experiments for 8 selected candidate interaction partners: eEF1A, eEF1G, EED, CRMP-1, RXRA, GNB2, C11Orf2 and RIPAIN. The interaction partners related the 13q14 proteins to cellular processes like protein translation,

regulation of gene expression, signal transduction and apoptosis, thereby linking them to possible pathways that could be involved in malignancy. Notably, interactions of 13q14 proteins with translation elongation factors could be involved in the apoptotic effects described for the factor eEF1A, e.g. by conferring specificity to the elongation complex to produce apoptosis-related proteins. In addition, the interactions with signaling mediator proteins like CRMP-1 and transcription factors like EED and RXRA indicate that 13q14 proteins are involved in signal transduction pathways and gene expression regulation. The links established in this work between 13q14 proteins and a broad range of cellular functions altered in malignancy, support their pathogenic role in the 13q14 deletion. Furthermore, the intertwined interactions of 13q14 proteins described here point to a possible functional gen-cluster in this region.

In conclusion, the work presented here shows that the 13q14 proteins a) localize corresponding to their predicted functions based on homology to known proteins, b) fit to a model of a functional cluster and involvement in the same cellular pathways, since they interact with each other and to a large extent with the same proteins, and c) affect pathways which are involved in the development of malignant phenotypes like signal transduction, protein translation, gene expression regulation and apoptosis. These findings constitute a basis for further studies towards the characterization of the functions of 13q14 candidate proteins and their possible roles in the development of cancer.

Zusammenfassung

Chronisch lymphatische Leukämie vom B-Zell Typ (B-CLL) ist die am meisten vorkommende Leukämie bei Erwachsenen in der westlichen Welt. Die häufigste genomische Abberation in B-CLL ist die Deletion der kritischen Region 13q14, die in über 50% der Patienten zu beobachten ist. Das häufige Auftreten dieser Deletion bei B-CLL und anderen Arten von Krebs deutet auf die Beteiligung von mindestens einem Gen dieser Region bei einem Tumorsuppressormechanismus hin. Um die Rolle der proteinkodierenden Gene der kritischen Region 13q14 im Pathomechanismus herauszuarbeiten, wurden die fünf Kandidatenproteine C13orf1, RCBTB1, SETDB2, RFP2 and KPNA3 in der vorliegenden Arbeit analysiert. Zu diesem Zweck wurden die intrazelluläre Lokalisierung sowie die Interaktionspartner der Proteine untersucht. Die intrazelluläre Lokalisierung der Proteine wurde in humanen Zelllinien anhand von überexpremierten Proteinen mit einem GFP-Fusions-Marker untersucht. Darüber hinaus wurden die endogenen Proteine über Immunohistochemie unter Verwendung spezifischer Antikörper, die für diese Studie hergestellt wurden, dargestellt. Die Lokalisierung dieser Proteine stimmt mit ihren mutmaßlichen Funktionen überein, die auf der Homologie mit gut charaktersierten Proteinen basiert. Interessanterweise wurden Unterschiede in der Lokalisierung einiger 13q14 Kandidatenproteine zwischen einer embryonalen Nieren-Zelllinie nicht-malignem Ursprungs und einer Lymphom-Zelllinie festgestellt. Dies deutet darauf hin, dass eine Relokalisation eine Änderung der Proteinfunktion vorliegen und mit dem malignen Status zusammenhängen könnte. Um die Funktionen der Kandidatenproteine aufzuklären, wurden Interaktionspartner von 13q14 Kandidatenproteinen mittels eines umfangreichen Yeast-Two-Hybrid (Y2H) Screen indentifiziert. In einer Voruntersuchung wurden Wechselwirkungen zwischen vier von fünf der 13q14 Kandidatenproteine gezeigt. Der Y2H-Screen wurde mit einer aus über 5000 Humanproteinen bestehenden Humanproteinbibliothek durchgeführt und lieferte 53 Kandidatenproteine. Die Überlappung von Interaktionen ist von besonderem Interesse, da die Homologiebetrachtung der Proteine auf eine Ko-Existenz in makromolekularen Komplexen hindeutete. Die Analyse der Y2H-Screen-Ergebnisse hat gezeigt, dass 36% der identifizierten Partner mit drei oder mehr 13q14 Proteinen interagieren. Dieser Befund, sowie die Interaktion der 13q14

Proteine untereinander, unterstützten die Hypothese eines funktionalen Zusammenhangs zwischen den untersuchten 13q14 Proteinen. Mittels Pull-Down-Experimenten wurden für 8 ausgewählte Kandidaten (eEF1A, eEF1G, EED, CRMP-1, RXRA, GNB2, C11Orf2 und RIPAIN) die Interaktionen validiert. Anhand der gefundenen Interaktionen lässt sich ein Zusammenhang zwischen 13q14 Proteinen und zellulären Prozessen wie Proteintranslation, Regulierung der Genexpression, Signalübertragung und Apoptose herstellen. Sie verbinden somit 13q14 mit möglichen Signalwegen, die bei der Entwicklung von bösartigen Tumoren eine Rolle spielen. Insbesondere könnten die Interaktionen von 13q14 Proteinen mit Translationselongationsfaktoren an den apoptotischen Effekten, die z.B. für den Faktor eEF1A bereits beschrieben wurden, beteiligt sein. Die Rolle der 13q14 Proteine könnte beispielsweise die Spezifität des Elongationskomplexes für die Produktion Apoptose-bezogenes Proteine beeinflussen. Die Interaktion einiger 13q14 Proteine mit intrazellulären Signalvermittlern wie CRMP-1 und Transkriptionsfaktoren wie EED oder RXRA weist zusätzlich auf eine Bedeutung bei der Signalübertragung und bei der Regulierung der Genexpression hin. Die Zusammenhänge zwischen 13q14 Proteinen und ein breites Spektrum an zellulären Prozessen, die an der Bildung von bösartigen Phänotypen beteiligt sind, unterstützen die pathogene Rolle der 13q14 Proteine bei 13q14 Deletionen. Darüber hinaus weisen die hier beschriebenen ineinander greifenden Interaktionen der 13q14 Proteine auf einen möglichen funktionellen Gen-Cluster in dieser Region hin.

Zusammenfassend kann festgestellt werden, dass die 13q14 Proteine a) dort lokalisieren, wo es durch Homologievergleiche mit bekannten Proteinen vorausgesagt war, b) im Einklang mit einem funktionalen Cluster-Modell und einer Beteiligung an den gleichen zellulären Signalwegen stehen, da die 13q14 Proteine sowohl miteinander als auch weitgehend mit den selben Proteinen interagieren, c) einen Einfluß auf Signalübertragung, Proteintranslation, Regulierung der Genexpression und Apoptose nehmen. Die Ergebnisse dieser Arbeit stellen damit eine Basis für weitere Untersuchungen zur Charakterisierung der Funktionen der 13q14 Kandidatenproteine und deren Rolle bei der Entwicklung von Krebs dar.

Publications

Original articles

Cordula Tschuch, Angela Schulz, Armin Pscherer, Wiebke Werft, Axel Benner, Agnes Hotz-Wagenblatt, **Leticia Serra Barrionuevo**, Peter Lichter and Daniel Mertens (2008). „Off-target effects of siRNA specific for GFP". BMC Molecular Biology 2008, 9:60

"Candidate tumor suppressor genes localized in 13q14.3 form a functional cluster that is involved in NFκB signalling". In preparation.

Posters

Leticia Serra Barrionuevo, Daniel Mertens, Melanie Ruppel, Cordula Tschuch, Stephan Wolf, Peter Lichter. Monoclonal Antibodies against proteins in 13q14. PhD student poster presentation, German Cancer Research Center, Heidelberg, December 2004.

Leticia Serra, Daniel Mertens, Melanie Ruppel, Cordula Tschuch, Stephan Wolf, Peter Lichter. Proteomics of 13q14. PhD student poster presentation, German Cancer Research Center, Heidelberg, February 2005.

Leticia Serra Barrionuevo, Daniel Mertens, Sphefan Wolf, Peter Lichter. Insight into proteins from 13q14. 14th International AEK Cancer Congress. Frankfurt am Mein, February 2007

Curriculum vitae

Name: Leticia Serra Barrionuevo

Title: Dipl. Biologin (at Granada's Universtiy, Spain)

Date of birth: 26.01.1978

Place of birth: Málaga, Spain

WORKING EXPERIENCE

Jun 2005 – Present	PhD student at the University of Ulm. Department of Physiological Chemistry. Head: Prof. Dr. Thomas Wirth
Jan 2003 – Jul 2007	PhD student and practical work at the German Cancer Research Center, Heidelberg. Department of Molecular Genetics. Head: Prof. Dr. Peter Lichter
May 2002 – Jan2003	Research assistant at the German Cancer Research Center, Heidelberg. Department of Molecular genetics. Head: Prof. Dr. Peter Lichter
Sep 2001 – May 2002	Research assistant at the Zoology institute, Freiburg. Department of Developmental Biology Head: Prof. Dr. Wolfgang Driever
Sep 1998 – Jun 1999	Laboratory Trainee at Granada's University, Spain Department of Cell Biology Head: Prof. Dr. Julio Navascués Martínez

EDUCATION

Aug 1999 – Aug 2000	European Grant of the Erasmus program to study at Uppsala University, Sweden
Oct 1996 – Jul 2001	Biology studies at Granada's University, Spain
Jul 1995 – Aug 1996	Spanish Government Grant (High School program) to study the English language two summers in England.
Sep 1992 – Jul 1996	High School at the IES Emilio Prados in Málaga, Spain

Acknowledgements

My sincere thanks go to Prof. Dr. Thomas Wirth and Prof. Dr. Peter Lichter for the opportunity to do my PhD thesis, for inspiring me to highest quality research and for the outstanding collaboration that have made this work possible.

My sincere thanks go also to my supervisor Dr. Daniel Mertens for his support and advice. I deeply appreciate his great patience and valuable comments, which have been of invaluable help during proofreading of this dissertation.

I would also like to thank Dr. Stephan Wolf and Dr. Armin Pscherer for their great contribution to the progress of this work. The friendly smiles and sharing of their extensive knowledge and experience were a precious help.

I'm sincerely thankful to Prof. Dr. Dirk Görlich, to Prof. Dr. Hans-Walter Zentgraf, and to Prof. Dr. Erich Wanker for the opportunity to learn new techniques in their groups and the excellent work together. Especially I thank Thomas Güttler, Susanne Franz, Carolin Böhm, Theresia Knöbel-de-Ledezma, Dr. Ulrich Stelzl, Dr. Maciej Lalowski, Martina Zenkner, Raphaele Foulle, Christian Hänig and the rest of the people in their groups for the outstanding work and for being so friendly and patient while I worked in their labs.

For great support and advice in the lab I would like to sincerely thank Magdalena Schlotter, Sybille Ohl and Petra Schröter. I would also like to thank Stefanie Hofmann, Andrea Wittmann, Laura Pucio and Frauke Devens for the excellent lab atmosphere and prompt, friendly reactions whenever help was required.

For great technical assistance and hard work during his internship, I thank Raphael Rüppert.

For great help in the lab for the studies of localization of the proteins I would like to thank Angela Philippen.

For helping with the optimization of the usage of the monoclonal antibodies produced in this work, I would like to thank Sybille Ohl.

For the patient listening, wise advices, and the great conversations, not only about science but about everything else, I thank Melanie Ruppel, Dr. Cordula Tschuch, Alexandra Farfsing, Suzin Choi, Angela Schulz and Frauke Devens.

For great hold along the years through good and bad times, I'm profoundly thankful to Dr. Daniel Göttel.

My sincere thanks go to the whole Lichter-group for the continuous, qualified support and the great working atmosphere. The friendly dynamic of the group was very stimulating and a highly enriching experience. Specially, I would like to thank Prof. Dr. Peter Lichter for giving me the opportunity to work in his lab, for the inspiring ideas and his considerate support.

I'm most thankful to all my friends and family for their precious support, without which this work wouldn't have been possible.

I thank my whole family for their love and attentive support. Their valuable perspectives of problems not only enrich me, but contribute to my understanding of the relative importance of issues. I thank my mum, Francisca Barrionuevo, for her love and also her husband, Ildefonso Álvarez, for their thoughtful and caring support; my uncle, Rafael Barrionuevo, for his intelligent comments and for always believing in me, and also his wife, Elisabeth Trujillo for their loving concern; the rest of my family Maria Barrionuevo, Dr. Francisco J. García, Maria José Barrionuevo, Eduardo Garrido, Ana Barrionuevo, Carmelo de Médicis, Maria Serra and Alejandro Romero, for everything they share with me. I deeply thank my family for their inspiring lives and the enriching relationship we share beyond family boundaries.

For always being there for me and showing me her love and care, I thank my greek sister Ioanna Tzinieri.

For 25 years of friendship and showing me how to be a better person and a more sensitive human being, I thank my friend Maria del Carmen Trujillo.

For his friendship and sharing with me his inspiring personality, I would like to thank Fernando Moreno.

For being not just a friend but like a brother to me, I would like to thank my cousin, Alejandro Romero.

I thank my friends in Spain, Maria del Carmen Trujillo, Sara Vargas, Maria Isabel Sabastro, Elizabeth Santana, Raquel Cordón, Rocío García, Almudena Setién, Eduardo Vargas, Carlos Sabroso and Kike Vargas, for being there for me since we were children and showing me that friendship is independent from distance and time. To their partners I'm very thankful for the time we spend together and for being also my friends no matter how little we sea each other.

I would like to thank my friends in Germany Mara Amorós, Lucía Cespón, Julia Wittwer, Rafael González, Heike Gerdes and Rico Grösse-Schuttle for their caring support through the ups and downs of my work.

For the great conversations and long-distances friendships I would like to thank Ana Camargo, Kepa Santisteban, Antony Chen, Patric Jern, Mitja Ucakar, Philipp Heimbach and Klaus Kirchmayr.

At last, I would like to thank Dr. Martin Volland for his endless love and support, and for sharing with me the most exciting and inspiring points of view about work and life that I could have ever imagined.