ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE
SCHOOL OF LIFE SCIENCES

Master's project in Life Sciences and Technology

Functional Analysis and Knock-Down Phenotype of BCL9 and BCL9L in Cancer Cell Lines

Done by

Fanny Bringolf
Ch. Des Aubépines 22
1950 Sion

Supervised by
Prof. Michel Aguet, MD

In the Unit of Professor Michel Aguet
Swiss Institute for Experimental Cancer Research (ISREC)
EPFL

External Expert: Pof. Ivan Stamenkovic, MD
Lausanne (IUP)
Institute of Pathology, University of Lausanne (UNIL)

LAUSANNE, EPFL 2010

This master project was done in the laboratory of Prof. Michel Aguet in the Swiss Institute for Experimental Cancer Research (EPFL) from Semptember 2009 to the 20th August 2010, under the supervision of Frederique Baruthio, Norbert Wiedemann, Jennyfer Bultinck and Jürgen Deka. I would like to thank them and all the people who helped, supported and taught me during the realization of this work.

Abstract

Therapy-resistant tumor cells often display EMT and stem cell-like traits. The canonical Wnt pathway has been associated with these phenotypes in many cancer types. Our group has shown in mice that two proteins of this pathway, Bcl9 and Bcl9l, participate in controlling Wnt-mediated stem cell traits. We hypothesized that inhibition of BCL9/BCL9L could be a potential strategy to revert cells with stem-cell properties to a differentiated epithelial phenotype, and thus also re-establish their sensitivity to chemotherapy. Preliminary experiments in a human colon cancer cell line showed that BCL9L single knock-down (KD) surprisingly already completely abrogated expression of selected Wnt target genes, despite the presence of the second paralog BCL9, which was even up-regulated upon loss of BCL9L. This observation raised the question whether the remaining BCL9 protein might be dysfunctional, or whether BCL9 and BCL9L are, against our expectation, not functionally redundant. Thus, we characterized Bcl9 expression at transcriptional and genomic level by reverse transcription and DNA sequencing. By sequencing, we identified three different BCL9 mRNA species in SW480 cells (wild-type (WT) and two isoforms carrying short deletions), which we also confirmed to be present in other carcinoma cell lines. No abnormalities were found on the genomic level. We then assessed Bcl9 and Bcl9l expression at protein level using standard Western blotting. In the different cell lines the proteins showed distinct molecular weights, which may correspond to different post-translational modifications that could be associated with different functions. To generally explore BCL9 function, we silenced the gene using lentivirus vector-based expression of shRNAs. Further analysis of the observed Bcl9 KD on a broad panel of Wnt target genes and comparison with the Bcl9l KD phenotype will help elucidating Bcl9 and Bcl9L function.

Keywords: Wnt; BCL9/BCL9L; expression; knock-down

Table of Contents

Abbreviations

aa: amino-acid	miRNA : micoRNA
APC: Adenomatous Polyposis Coli	MOI: Multiplicity Of Infection
Apcdd1: Adenomatosis polyposis down-regulated 1	NHD: N-terminal Homology Domain
Arm: Armadillo	NLS: Nuclear Localization Signal
Bcl9: B-cell CLL/lymphoma 9	nt: nucleotide
bp: base pair	NTD: N-Terminal Domain
BSA: Bovine Serum Albumin	PBS: Phosphate Buffer Saline
cDNA: complementary DNA	PCR: Polymerase Chain Reaction
CHD: C-terminal Homolog Domain	PPT: Poly-Purine Track
CMV: CytoMegaloVirus	PyBD: Pygo Binding Domain
Ct: Cycle threshold	Pygo: Pygopus
CTD: C-Terminal	qRT-PCR: quantitative Real Time PCR
CTS: Central Termination Sequence	RISC: RNA-Induced Silencing Complex
ddNTP: DiDeoxyNucleoside Triphosphate	rpm: revolutions per minute
DMSO: DiMethyl SulfOxide	RRE: Rev Response Element
dNTPS: DeoxyNucleoside Triphosphates	RT: Reverse Transcriptase
dsDNA: double-stranded DNA	shRNA: short hairpin RNA
dsRNA: double-stranded RNA	shRNAmir:microRNA-adapted shRNA
EDTA: EthyleneDiamineTetraacetic Acid	siRNA: small interfering RNA
EMT: Epithelial-Mesenchymal Transition	ssDNA: single-stranded DNA
FBS: Foetal Bovine Serum	ssRNA: single-stranded RNA
GSK-3β: Glycogen Synthase Kinase 3β	TBS: Tris Buffer Saline
HMBS: HydroxyMethylBilane Synthase	Tm: melting Temperature
HRP: HorseRadish Peroxidase	tris: trishydroxymethylaminométhane
kb: kilo base	UTR: UnTranslated Region
KD: knock-down	VSV G: Vesicular Stomatitis Virus Glycoprotein
kDa: kilo Dalton	Wnt: Wingless integration site
Lef1/TCF: Lymphoid enhancer binding factor 1/Transcription Factor	WPRE: Woodchuck hepatitis virus Post-transcriptional Regulatory Element
Lgs: Legless	WT: wild-type
LRP: Leucine-responsive Regulatory Protein	α-: anti-
LTR: Long Terminal Repeat	β-catBD: β-catenin Binding Domain

List of Figures

List of Tables

I. Introduction

A. Cancer biology

Despite several decades of cancer research, cancer is still a major cause of death worldwide with about 7.9 million cases in 2007 (www.who.int/cancer, 25.06.2010). Cancer is a genetic disease caused by mutations in oncogenes or tumor suppressor genes that result in six physiological aberrant changes of the cancer cell: self-sufficiency in growth signals, insensitivity to growth-inhibitor signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan and Weinberg, 2000).

Therapies are designed through targeting these different mechanisms (Table 1). Despite all the investigative efforts, cancer remains a widespread, relentless and lethal disease. One of the major difficulties is that cancer cells are able to acquire resistance against treatments after which the cancer might obtain more malignant traits.

Therapeutic target or modality	Targeted process	Mechanism of action of therapeutics	Target example (drug)
Transformation	Apoptosis	Activation of apoptosis pathways	BCL2
	Signalling	Interference with signal transduction, response	ABL (Gleevec; Novartis)
	Invasion/metastasis	Inhibition of tumour spread	Cathepsin K
Immortalization	Senescence	Induction of senescence	Telomerase
Host	Angiogenesis	Interference with blood supply of tumour	VEGF (Avastin; Genentech/Roche)
	Tumour-associated membrane proteins	Antibody-directed cytotoxicity	CD20 (Rituxan; Biogen Idec/Genentech)
Traditional cytotoxics	Replication/cytokinesis	Interference with DNA synthesis, cell division	Microtubules (Taxol)
	Metabolism	Reduction of essential metabolite	Thymidylate synthase (5-FU)
Neocytotoxics	Protein turnover	Inhibition of acceleration of protein degradation	Proteasome (Velcade; Millennium Pharmaceuticals)
	Epigenetics	Remodelling chromatin, DNA methylation	HDAC interactions
	Stress response	Interference with cellular stress buffering	ATPase/chaperone superfamily

Table 1. Cellular processes targeted by actual therapies (Kamb et al., 2007).

A number of observations led to the concept of tumor cells with characteristics of stem cells, called tumor-initiating cells or cancer stem cells. They can be defined as cells having the ability to re-grow a tumor similar to the one they are originated from. They are not necessarily derived from normal stem cell (Zhou et al., 2009). The idea is that only a subset of the cells shares this capacity in a tumor and that tumors are organized as a stem cell hierarchy. However, not all the experimental models support this hypothesis (Zhou et al., 2009). These tumor-initiating cells are characterized by stem cell-like phenotypes, as drug resistance, extensive capability of self-renewal, sometimes quiescent nature (Pardal et al., 2003) and an epithelial-to-mesenchymal transition (EMT) state (Thiery, 2002). These characteristics might explain therapy resistance and dissemination of this tumor cells.

EMT is a process involved in a variety of biological events including embryogenesis, organ development, tissue regeneration, organ fibrosis, and cancer progression. EMT can be described as a shift from an epithelial phenotype, characterized by cell polarity, cell-cell adhesion and lack of motility, to a mesenchymal phenotype, characterized by motility, invasiveness and higher resistance to apoptosis (Figure 1). Typically, carcinoma cells which undergo an EMT express mesenchymal markers such as α-smooth muscle actin (α-SMA), fibroblast-specific protein 1 (FSP1), vimentin and desmin (Yang and Weinberg, 2008) and are localized at the invasive front of primary tumors (Thiery, 2002). β-catenin accumulates in the nucleus of EMT cells and their E-cadherin expression is lower, marking the disruption of cell-cell adhesion junction (Kalluri and Weinberg, 2009). It has also been shown that cells undergoing an EMT acquire stemness properties (Mani et al., 2008).

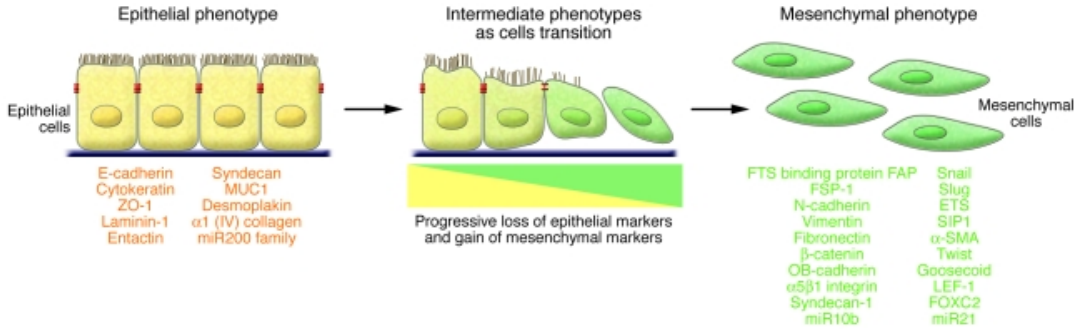


Figure 1. EMT mechanism and players (Kalluri and Weinberg, 2009).

Signals from the tumor-associated stroma, like hepatocyte growth factor (HGF), epidermal growth factor (EGF), platelet-derived growth factor (PDGF) and transforming growth factor-β (TGF-β), appear to induce expression of EMT-inducing transcription factors like Snail, Slug, zinc finger E-box binding homeobox1 (ZEB1), Twist, Goosecoid or FOXC2 (Kalluri and Weinberg, 2009). TGF-β is the main and best characterized of them. The general mechanism of TGF-β action in EMT is the phosphorylation of Smad proteins and so, activation of gene expression that lead to facilitated motility (Kalluri and Weinberg, 2009). Other non-canonical pathways activated by TGF-β are mitogen-activated protein kinases (MAP kinases) (e.g. ERK1/ERK2, p38 MAPK, JNK), growth and survival kinases (e.g. PI3K, AKT/PKB, mTOR), and small GTP-binding proteins (e.g. Ras, ThoA, Rac1, Cdc42) (Wendt et al., 2009). Signaling pathways, such as the Wnt (Vincan and Barker 2008), Shh and Notch pathways (Bailey, Singh et al. 2007) also contribute in triggering this transition.

B. The Wnt Pathway

The Wnt pathway is known as a key regulator of embryonic development and tissue renewal. Wnt drives at least three sub-signalings. The canonical Wnt pathway involves β-catenin protein and activates transcription of target genes. The planar cell polarity pathway involves Jun N-terminal kinase (JNK) and cytoskeletal modifications. The Wnt/Ca^{2+} pathway modulates intracellular Ca^{2+} level. An aberrant activation of the pathway, mainly the canonical, triggers cancer (Clevers, 2006). It is known that a majority of colorectal cancers are initiated by an activating mutation in the Wnt pathway, mainly in the adenomatous polyposis coli (APC) gene. Experimental evidences also show that dysregulation of the Wnt pathway plays a role in leukemias (Reya and Clevers, 2005) and in breast cancers (Ai et al., 2006; Howe and Brown, 2004; Milovanovic et al., 2004). Frequency of Wnt pathway abnormalities in different cancer types led to a substantial investigative effort on the role of Wnt signaling in tumor development worldwide.

The canonical Wnt pathway (Figure 2) is a critical signaling pathway driven by binding of Wnt ligand with the receptor complex. In addition to its essential role in cell-cell adhesion through its interaction with the transmembrane protein E-cadherin, β-catenin is a key player in this Wnt cascade. Its role depends on its cellular location and on differential interaction with various proteins.

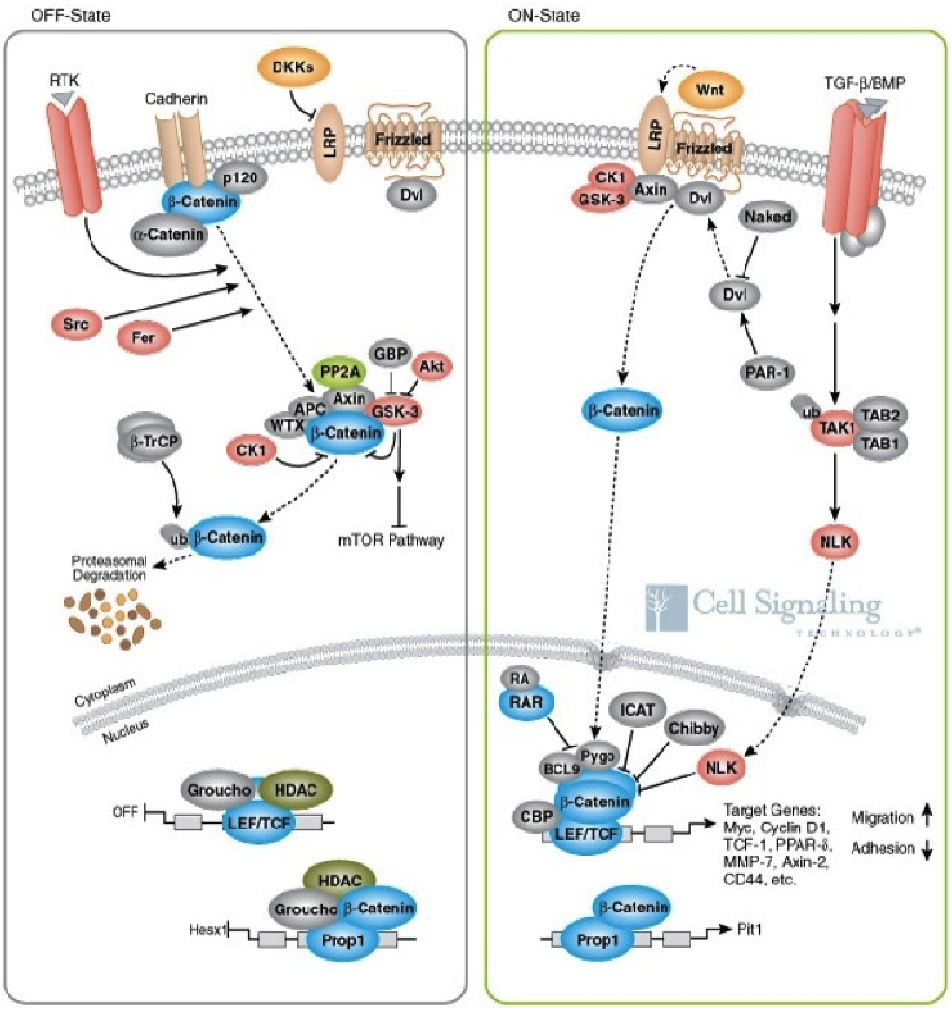


Figure 2. Canonical Wnt pathway. Adapted from http://www.cellsignal.com/.

In absence of a Wnt ligand, the so-called destruction complex composed of APC, Axin, glycogen synthase kinase 3 (GSK-3β), and casein kinase (CKI), maintains the level of free cytoplasmic β-catenin at low level. GSK-3β phosphorylates β-catenin leading to its ubiquitination by β-transducin repeat-containing protein (β-Trcp) and hence proteasomal degradation. In the absence of nuclear β-catenin, lymphoid enhancer binding factor 1/transcription factor (Lef1/TCF) complex represses the expression of Wnt target genes by recruiting a chromatin repressor, Groucho, and histone deacetylases (HDACs).

When a Wnt ligand binds to Frizzled and leucine-responsive regulatory protein (LRP) receptor complexes, LRP and Dishevelled (Dvl) bind Axin and GSK-3β from the destruction complex, which leads to inhibition of β-catenin phosphorylation by GSK-3β through a mechanism that is still unclear (Wu and Pan, 2010). This results in the stabilization and the nuclear accumulation of β-catenin. Once in the nucleus, β-catenin displaces with Groucho and binds to Lef1/Tcf. This leads to the transcription of Wnt target genes, such as c-Myc, Axin-2 and Cyclin D1 (Clevers, 2006).

β-catenin consists of 12 armadillo (Arm) repeats, R1-R12, and distinct N-terminal and Carbox-terminal domains (NTD and CTD, repectively). β-catenin C-terminal end, R10-R12 and CTD, binds different co-factors that modulate the transcriptional activation (Mosimann et al., 2009). These

co-factors act on chromatin modification, chromatin remodeling and the basal transcription machinery (Hecht et al., 2000; Daniels and Weis, 2002; Barker et al., 2001; Koh et al., 2002; Sierra et al., 2006; Yang et al., 2006). The N-terminal Arm repeat, R1, binds B-cell CLL/lymphoma 9 (Bcl9) (Kramps et al., 2002) and R1-R2 bind α-catenin (Pokutta and Weis, 2000). Phosphorylation of Tyrosine 142 in R1 disrupts α-catenin binding and promotes Bcl9 binding (Brembeck et al., 2004). E-cadherin and LEF1/TCF share the same binding site in the central β-catenin armadillo repeat (R3-R10) (Orsulic et al., 1999; Graham et al., 2000; Huber and Weis, 2001). E-cadherin binding to β-catenin is regulated by phosphorylation of a tyrosine in the last armadillo repeat (R12) (Brembeck et al., 2006). So, β-catenin action is finely tuned, switching from cell adhesion to transcriptional activation of Wnt target genes.

C. BCL9/BCL9L

The importance of Bcl9 proteins in the Wnt pathway was discovered through its Drosophila homolog, Legless (Lgs). Interaction of Lgs with its two binding partners, β-catenin and Pygopus (Pygo), is critical for normal development in flies (Hoffmans et al., 2005), but is not essential in mammalian cells (de la Roche et al., 2008). However, several observations imply that Bcl9 and/or Bcl9l do play a role in the transcription of Wnt targets in mammalian cells. Absence of Bcl9 resulted in attenuation of Wnt target gene transcription in Wnt-active cells like the colorectal cancer cell lines SW480 and COLO320, or Wnt-stimulated non-cancerous cell lines, HEK293T cells (de la Roche et al., 2008; Mani et al., 2009), as well as absence of Bcl9l in SW480 cells (Brembeck et al., 2004). In chemically induced colon adenocarcinomas in mice, induced depletion of Bcl9 and Bcl9l in intestinal cells shows down regulation of a subset of Wnt target genes, as Adenomatosis polyposis down-regulated 1 (Apcdd1) or Axin2 (Deka *et al.*, 2010). Furthermore, a reduced transcription of Wnt targets is observed in cells that express a β-catenin mutant unable to bind Bcl9 (Hoffmans et al., 2005).

Bcl9/Bcl9l depletion in the mouse intestine does not impair homeostasis, but impairs colon tissue-renewal in case of injury (Deka *et al.*, 2010). Depletion of Bcl9 in a mouse colon xenograft model induces a better survival of the host mice, suggesting that it attenuates tumor load and metastasis (Mani et al., 2009). However, depletion of both Bcl9 and Bcl9l in a chemical carcinogenesis model of murine colon adenorcarcinoma has no significant impact on the tumor frequency (Deka *et al.*, 2010), indicating that Bcl9/9l does not affect the number of tumor-initiating cells and their capacity to form tumors. However, we were able to show that, in Wnt active tumors, Bcl9 and Bcl9l are largely involved in the control of stem cell-like properties.

Recently, the role of Bcl9 and Bcl9l in human cancer has been investigated. Bcl9l has been found to be over-expressed in multiple myeloma tumors, colon cancer (Mani et al., 2009; Adachi et al., 2004; Sakamoto et al., 2007) and breast cancer (Toya et al., 2007). Bcl9 shows also a higher expression level in myeloma tumors and colon cancers, compared to normal plasma and colon cells, respectively (Mani et al., 2009). Still, that does not prove a direct role of Bcl9/Bcl9l in tumor formation, malignancy or spreading.

The direct Wnt target genes that are down-regulated in Bcl9/Bcl9l mutated murine colon tumors include mesenchymal (e.g. Vimentin), malignancy (e.g. Prox1) or stemness markers (e.g. Lgr5) (Deka *et al.*, 2010). Furthermore, subcellular distribution of epithelial and mesenchymal markers changes when either Bcl9, Bcl9l or both are absent from colon cancer cells (Deka *et al.*, 2010; Mani *et al.*, 2009; Brembeck *et al.*, 2004). Mesenchymal markers no longer co-localize with epithelial markers in Bcl9/Bcl9l-depleted colon tumor cells. Furthermore, the Bcl9/Bcl9l-negative tumors display a preserved basal lamina, indicating a less invasive phenotype (Deka *et al.*, 2010).

Therefore their inhibition would have a great therapeutic interest that will be set out in the next chapter.

Lgs, Bcl9 and Bcl9l share only limited sequence identity: Lgs and Bcl9 share less than 10% of their sequence (Kramps et al., 2002), and Bcl9 and Bcl9l about 35% (Brembeck et al., 2006). However, it has been demonstrated that both Bcl9 and Bcl9l can functionally replace Lgs (Hoffmans and Basler, 2007; Kramps et al., 2002). The reason for this functional redundancy is the fact that the three proteins contain seven common, highly conserved domains (Figure 3). Two of these homology domains, named β-catenin binding domain (β-catBD) and pygo binding domain (PyBD), are involved in binding β-catenin and Pygo, respectively. Pygo is thought to exchange or stabilize co-activators that bind the C-terminus of β-catenin and regulate chromatin remodeling (Mosimann et al., 2009; Stadeli et al., 2006). The three proteins (Lgs, Bcl9, Bcl9l) carry a nuclear localization signal (NLS), but it has been shown insufficient for the nuclear localization of Lgs (Kramps et al., 2002; Brembeck et al., 2006). The N-terminal homology domain (NHD) is only essential for the nuclear localization of Bcl9l (Brembeck et al., 2004), while Bcl9 and Lgs, who lack this NLS, require Pygo binding to be redistributed into the nucleus (Townsley et al., 2004).

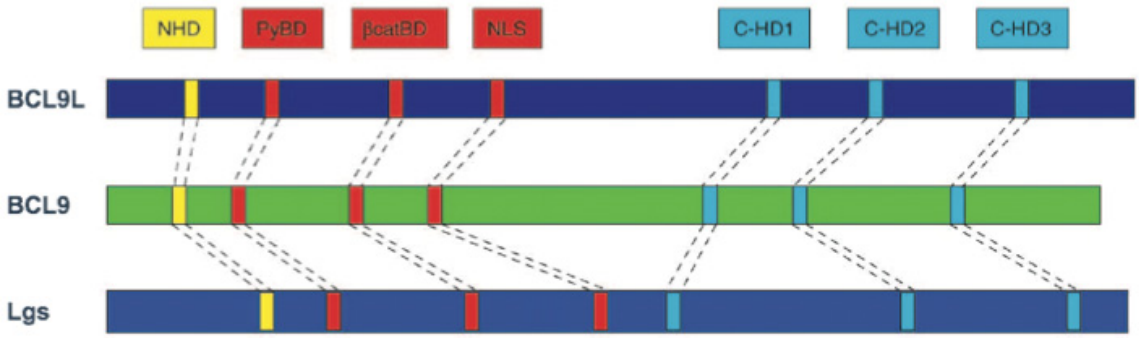


Figure 3. BCL9, BCL9L and Lgs homology domains. Adapted from (Brembeck et al., 2006)

Three additional homology domains are located in the C-terminal part of the Bcl9 proteins (C-terminal homology domain: CHD1 to 3). For these CHD no interaction partners have yet been found (Sustmann et al., 2008). However, a region between CHD2 and CHD3 was shown to act as a transactivation domain in Bcl9, and more modestly in Bcl9l.

D. RNA interference and lentiviral biology

1. RNA interference

The interfering RNA pathway is a natural mechanism of the cells for defense against some parasitic RNA, e.g. viral RNA, or for regulation of the transcriptional expression of messenger RNAs (mRNAs). It produces a small, about 20 nucleotides (nt), single strand RNA, micoRNA (miRNA), that can bind specifically or with some mismatches a targeted RNA and induces their degradation or represses their translation. Scientists investigate deeply this mechanism to copy miRNA features and create means to permanently and efficiently degrade targeted mRNA.

The mechanism of miRNA processing is more or less known. In the genome, miRNA sequences are located in introns as well as in non-coding exons. They are transcribed as mRNA by the RNA polymerase II and are regulated by polymerase II promoters. The direct miRNA transcript product, the pri-mRNA, forms a hairpin secondary structure with two ~40 nt single-stranded RNA (ssRNA) flanking sequences. The stem of this hairpin is about 3 helical turns (~30 base pairs (bp)). The loop of the hairpin is often large, more than 10 nucleotides, and seems to improve

the miRNA production efficiency (Zeng et al., 2005). After transcription, RNase III, Drosha, and the RNA binding protein, DGCR8 (alias Pasha) work in a complementary fashion to cleave the pri-miRNA from the mRNA with the release of the pre-miRNA form. DGCR8 recognizes the single stranded flanked sequences together with the double stranded stem. DGCR8 also binds Drosha and set the cleavage by Drosha 11th bp downstream the root of the miRNA hairpin (Zhao and Liu, 2009; Nowotny and Yang, 2009). As DGCR8 can also less efficiently recognize the loop, some abortive products can be produced. Introducing a small internal bulge 9-12 nucleotides after the stem loop in the 5' miRNA arm seems to avoid it (Han et al., 2006). Exportin 5 exports the pre-miRNA to the cytoplasm where Dicer removes the pre-miRNA terminal loop. The site where Drosha cleaves is very important as it determines the cleavage site for Dicer (Zhao and Liu, 2009; Nowotny and Yang, 2009). The crucial player in the last step to create the final miRNA product is the RNA-induced silencing complex, RISC. RISC, formed by Argonaute proteins associated with proteins non-essential for slicing (e.g. Fragile Xmental Retardation Protein or Dead-box RNA helicase gemin), incorporates the double strand-miRNA, keeps one strand and degrades the other. The choice of the single strand is not random. Usually the one with the less thermodynamically stable 5'end (Schwarz et al., 2003), mainly composed with uracil (U) nucleotides, is chosen. The bases composition of the strand seems also to play a role as an excess of purines favours the selection (Hu et al., 2009). The RISC coupled to the miRNA will then recognize through perfect or imperfect base pairing the miRNA complementary sequence of the mRNA, usually located in its 3'untransleted region (UTR). The complementary with the nt 2-8 of the 5'end of the miRNA are necessary for the recognition process. Other features such as AU-rich sequences, base pairing with nt 13-16 of the miRNA, targeting sequence in the 3'UTR at least 15 nucleotides from the stop codon and away from the centre of long UTRs, act as booster for the recognition (Zhao and Liu, 2009; Nowotny and Yang, 2009). Usually in mammalian cells, the matching between the target mRNA and the miRNA is not perfect and induces different inhibition mechanisms as translational initiation inhibition, post-initiation inhibition and mRNA decay pathway.

Artificial silencing RNAs inspired from the miRNA are called the small interfering RNA (siRNA), short hairpin RNA (shRNA) and the microRNA-adapted shRNA (shRNAmir). siRNA can be considered as miRNA, shRNA as pre-miRNA and the shRNAmir as pri-miRNA. The siRNA are transfected directly to the cell and its silencing is transient. shRNA are transcribed under the control of RNA polymerase III promoters. They are produced as single-stranded molecules of 50-70 nucleotides that form a stem-loop structure and exit the nucleus to be cleaved by Dicer and RISC. The shRNAmir contains same kind of sequence than shRNA plus flanked sequences based on mir-30a, a well known miRNA possibly targeting p53 of the apoptotic pathway (Li et al., 2010), and contains also the features necessary for miRNA processing, as the bulge at the 9-12 nt, the less thermodynamically stable 5'end and the loop (Figure 4). This shRNAmir can be integrated either in introns of protein-coding genes or in exons or introns of mRNA-like noncoding RNAs. Their transcription is driven by polymerase II promoter (Wiznerowicz and Trono, 2003) and they enter the natural process for miRNA maturation (Fewell and Schmitt, 2006). Both shRNA and shRNAmir induce stable silencing. siRNA, shRNA and shRNAmir sequence match perfectly with the target mRNA and so, RISC directly degrades the mRNA.

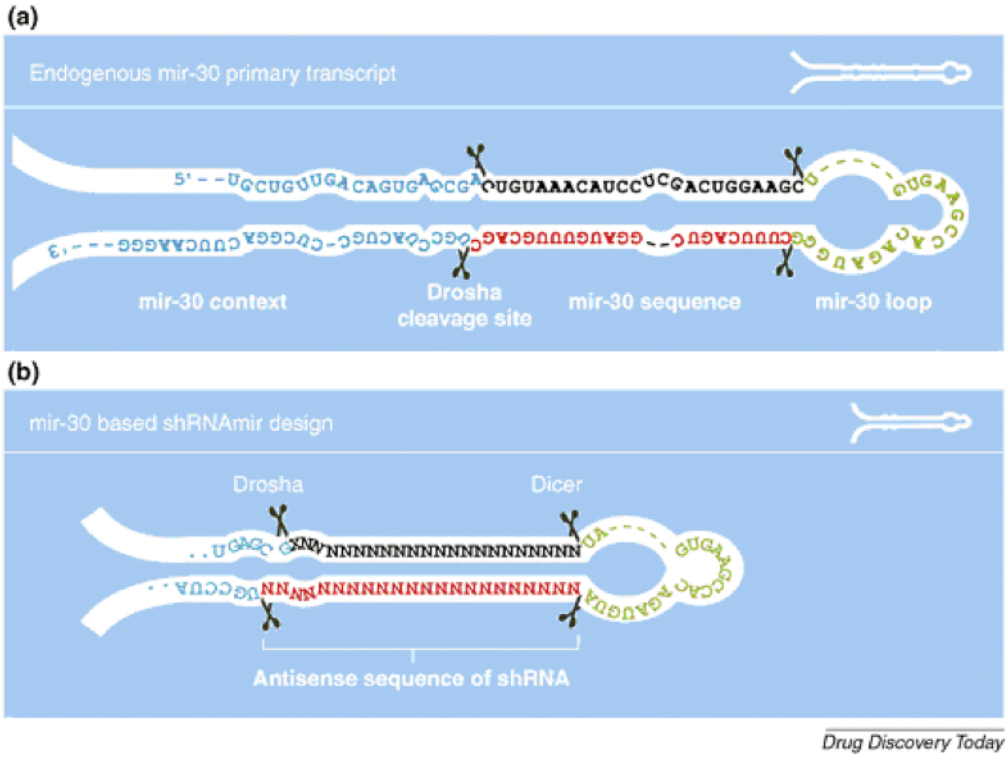


Figure 4. Mir-30 hairpin and mir-30 based shRNAmir design (Schwarz *et al.*, 2003).

2. Lentiviral biology

After the fusion with the cell, the viral RNA genome enters the host cell. Its own reverse transcriptase initiates its reverse transcrition and creates the viral double-stranded DNA (dsDNA). Then a complex is formed, called PreIntegrationComplex (PIC), with the integrase (IN), the reverse transcriptase (RT), the Gag matrix protein (MA), the Viral protein R (Vpr) protein and the nucleocapsid protein (NC) (Miller et al., 1997; Bukrinsky et al., 1993; Nermut and Fassati, 2003; Gallay et al., 1995). This complex allows the viral dsDNA to go into the nucleus and to be protected. IN form dimer with both ends of the viral DNA, specifically in the U5 and U3 region of the long terminal repeat (LTR) (Delelis et al., 2008). Finally, both dimers form a tetramer and interact with the host proteins, lens epithelium derived growth factor (LEDG)/p75, tethering factor for IN, and a transcription coactivator in the host cell (Llano et al., 2006; Hombrouck et al., 2007). Then IN cleaves and eliminates a dinucleotide adjacent to highly conserved CA dinucleotide from each 3' end of the DNA and then by transesterification introduces the viral dsDNA into the host genome (Delelis et al., 2008; Jegede et al., 2008). The choice of the integration site is not understood yet. The viral DNA is now able to use the host machinery to produce the different kind of proteins needed for its replication. Gag proteins are structural proteins involved in the formation of the virus capsid and the recruitment of the viral ssRNA in it. Pol proteins are involved in the reverse transcription and integration of the lentiviral ssRNA and envelope proteins in the binding and the fusion with the target cells. The lentiviral genome contains regulatory proteins involved in transcriptional transactivation (Tat) and translational regulator (Rev), and accessory proteins involved with the virion infectivity (Nef), T cell targeting (Vpn), G2 arrest of the host cells (Vpr) and replication in non permitive cells (Vif) (Hope and Trono, 2000). It contains also regulatory sequences promoting nuclear export and stabilizing lentiviral mRNA (Woodchuck hepatitis virus post-transcriptional regulatory element, WPRE), priming synthesis of the DNA positive strand after the reverse transcription and its entry into the nucleus (poly-purine track, cPPT and 3'PPT), terminating the DNA synthesis (central termination sequence, CTS), facilitating the transport from the nucleus to the cytoplasm of

unspliced transcripts (rev response element, RRE), triggering the packaging of the ssRNA into the lentiviral capsid (Psi). The LTRs are divided in three subregions, U3, R and U5, and are also key sequences for the regulation of the lentivirus genome. The U3 of the 5'LTR contain enhancer, promoter and transcription initiation sequences. The 3' LTR contain polyadenylation and transcription termination signals. U5 regions contain the Tat binding, packaging sequences and the site priming the reverse transcription (Trono, 2002). The TATA box is located roughly 25 bp from the beginning of the R sequence. LTR cis-acting sequences are required for the integration of the lentiviral DNA in the host genome.

3. Lentiviral system for gene delivery

The lentivector system use the capacity of the lentiviruses to integrate a gene of interest the host genome in a stable manner but is designed as to not induce viral replication. Its genome is split into three different plasmids and then transfected into producing eukaryotic cells. One plasmid encodes for the Pol and the Gag proteins, a second encodes the Env proteins, and a third contains the sequences for encapsidation and production of its genome and the gene of interest. Depending on the system used, the Env proteins are replaced by heterologous envelope glycoprotein from other virus (e.g. VSV G glycoprotein) (first generation system), the accessory proteins, Nef Vif, Vpr and Vpu, are removed (second generation system), LTRs are replaced by a strong heterologous promoter from cytomegalovirus (second or third generation system), the U3 region in the 3'LTR are deleted in the transfer vector (second or third generation system) and the protein Rev is isolated in a fourth plasmid (third generation) (Trono, 2002).

With these capacities shRNAmirs delivered with lentivectors induce less off-target effect than artificial siRNA delivery through liposomes, nanoparticles or bacteria. Moreover their integration in the host cell genome is stable overtime and allows them to be processed as miRNA.

II. Aims of the study

Despite the fact that efficient cancer treatments are currently available, major limitations as metastasis, drug resistance and relapse after the ending of the treatment remain. Moreover, these metastatic and resistant cancers are often correlated with poor prognosis. As EMT is implicated in tumor invasion and metastasis, and as therapy-resistant tumor cells often display EMT and stem cell-like traits, reverting these traits might significantly improve classical cancer therapies that target immortality and proliferative phenotypes.
Studies in loss-of-function mouse mutants revealed an essential role for Bcl9/Bcl9l in regulating a subset of Wnt target genes involved in controlling EMT and stem cell–related features (Deka *et al.*, 2010). Indeed, conditional ablation of Bcl9/Bcl9l in the intestinal epithelium resulted in decreased expression of intestinal stem cell markers and impaired regeneration of ulcerated colon epithelium. Adenocarcinomas with aberrant Wnt signaling arose with similar incidence in wild-type and mutant mice. However, transcriptional profiles were vastly different: whereas wild-type tumors displayed characteristics of EMT and stem cell–like properties, these properties were largely abrogated in mutant tumors. To validate that targeting Bcl9/Bcl9l function in colon and other Wnt-activated cancers may result in attenuated malignancy and restored responsiveness to therapy, a similar role for BCL9 and BCL9L needs to be shown in human cancer models. Therefore, BCL9 and BCL9L will be depleted in human Wnt active carcinoma cell lines that were selected based on the expression of mesenchymal and stem cell-associated genes (SW480 cells for colon and MDA-MB-231 cells for breast). The impact of BCL9/9L depletion will be assessed on several levels, including morphological changes, gene expression profiles, *in vitro* and *in vivo* response to therapy, and tumor-initiating capacity.

In *Drosophila*, either of the two Lgs homologues, BCL9 or BCL9L, could functionally reconstitute the Lgs loss-of-function phenotype (Kramps *et al.*, 2002; Brembeck *et al.*, 2004). Therefore, BCL9 and BCL9L are expected to have redundant functions. However, in colon carcinoma cell lines, single knock-down of BCL9 or BCL9L already results in phenotypical and gene expression changes (de la Roche *et al.*, 2008; Mani *et al.*, 2009; Brembeck *et al.*, 2004). Moreover, up-regulation of BCL9 in BCL9L-depleted SW480 cells could not rescue the effects on Wnt target gene expression. These observations might be explained either by different biological functions for BCL9 and BCL9L, or by a defect of one of the two proteins in human colon carcinoma cell lines.

To address these questions, we will compare BCL9 or BCL9L single and double knock-down, and try to rescue the BCL9L knock-down phenotype with BCL9 complementary DNA (cDNA), and vice versa. We already observed that BCL9L cDNA resistant to the BCL9L shRNAmir is not able to rescue the BCL9L knock-down phenotype, which is indicative of an off-target effect. However, it could also be that the BCL9L defect induces irreversible changes. This will be addressed by transfecting cells with the cDNA before adding the shRNAmir. If the presence of the cDNA can prevent the occurrence of the knock-down phenotype, off-target effects are excluded. Subsequently, the irreversibility of the knock-down phenotype will be assessed by looking at the chromatin accessibility of direct Wnt target genes in the parental *versus* the knock-down cell lines.

The goal of the present study is to clarify BCL9 expression and function in Wnt active human cancer cell lines. BCL9 genomic DNA, RNA and protein expression were characterized using standard methods, as Southern and Western blots. Different transcript populations of BCL9 were found in all tested cell lines and subsequently sequenced and identified. Finally, BCL9 expression was prevented using BCL9-specific shRNAmir-expressing lentiviruses and its effect on the expression of direct Wnt target genes was compared to BCL9L knock-down cell lines.

III. Materials and Methods

A. Human cell lines:

All cell lines (Table 2) were cultured on tissue culture treated dishes or flasks, at 37°C in humidified atmosphere with 5% CO_2. To passage the cells, cultures were washed with 1x phosphate buffer saline (PBS) and then incubated about 3 minutes, depending on the cell lines and their ease to detach, with 0.05% trypsin-ethylenediaminetetraacetic acid (EDTA) (Gibco, #25300) at 37°C. Addition of complete culture medium stops the trypsinization process due to the presence of foetal bovine serum (FBS). If needed, live cell number was determined with trypan blue, a Neubauer counting chamber and the following equation:

(number of cells counted in a large square of the Neubauer chamber) x 10^4 x (dilution factor of cells in trypan blue) = (the number of cells)/ (ml of culture medium).

For long time storage, cells were suspended in 65% culture medium, 25% FBS and 10% dimethyl sulfoxide (DMSO), stored in cryogenic tubes and frozen overnight into isopropanol containing box at -80°C. Isopropanol allows the cells to slowly cool down, by 1°C per minute, and avoids cells breaking. Finally, cells were transferred into liquid nitrogen at about -196°C.

To re-cultivate cells from liquid nitrogen storage, they were thawed at 37°C for a short time (~1minute) and transferred into culture medium, which was centrifugated to separate cells from the cytotoxic DMSO-containing medium, cells were then re-suspended in new culture medium.

Organ	Cell line	Type	Medium	Mutations*	Overexpression*
Colon	HCT116	Carcinoma	McCoy's 5a medium + 10% FBS	B-catenin, Ras, TGF-β receptor 2 (TGFBR2)	β-catenin, c-myc
	COLO320	Adenocarcinoma	RPMI 1640 medium + 10% FBS	APC (type I)	β-catenin, c-myc
	LS174T	Carcinoma	Dubelco's Modified Eagle's medium (DMEM) + 10% FBS	B-catenin, Ras, Tcf4, TGFBR2	APC, β-catenin, c-myc, p53
	SW480	Carcinoma	RPMI 1640-medium + 10% FBS	APC (type I), Ras, p53	APC, β-catenin, c-myc, p53
	HT29	Adenocarcinoma	McCoy's 5a medium + 10% FBS	APC (type II), p53	APC, β-catenin, c-myc, p53
Breast	MCF7	Adenocarcinoma	DMEM + 10% FBS		
	MDA-MB231	Adenocarcinoma	DMEM + 10% FBS		
	HS578T	Adenocarcinoma	DMEM + 10% FBS + 0.01mg/ml bovine insuline		
Kidney	HEK293T	Embryonic	DMEM + 10% FBS + 1% non-essential amino acid solution (NEAA)		

*(Gayet *et al.*, 2001; van Erk *et al.*, 2005)

Table 2. List of human cell lines used.

B. Genomic DNA extraction from human cells:

DNA was extracted from $5x10^6$ cells using DNeasy® Blood & Tissue Kit (spin column format) from Qiagen (#69504) according to the manufacturer's instructions. In brief, cells were harvested, centrifugated, resuspended in PBS and digested with proteinase K, a broad-spectrum serine protease, in a buffer with a high concentration of chaotropic salts creating appropriate binding conditions for nucleic acid to a silica-gel membrane. Adding ethanol ensures lysis and weakening of hydrophobic bonds. Then the silica-gel membrane selectively adsorbs nucleic acids while metabolites and cellular proteins are subsequently washed away. Finally, the DNA is eluted from the silica-gel membrane using a low-salt buffer and centrifugation.

If RNA-free genomic DNA is required, RNase A can be added before proteinase K. (Qiagen, 2002 #3).

C. RNA extraction from human cells:

Total RNA was extracted using NucleoSpin® RNA II from Macherey-Nagel (# 740 955.50). $5x10^6$ cells are lysed with a lysis solution containing large amounts of chaotropic ions inactivating RNases and creating appropriate binding conditions for the RNA absorption to silica membrane placed in a column. Contaminating DNA is removed by addition of rDNase solution when the lysate is already loaded onto the silica membrane. Salts, metabolites and macromolecular cellular components are removed through several washes steps. RNA was eluted by

centrifugation with RNase-free H_2O. The method yields up to 70µg of total RNA per NucleoSpin® RNA II column.

D. Protein extraction from human cells

Proteins were extracted from 10^6 -10^7 cells. Medium that contains secreted proteins was removed from cells followed by two cold PBS washes. Cells were lysed with CelLytic™ MT lysis buffer (Sigma, # C3228) mixed with a cocktail of proteases inhibitor, cOmplete Mini (Roche, # 04693124001). The lysis buffer consists of a mild detergent, bicine, and 150 mM NaCl mix that minimize interference with protein interaction. Proteases inhibitors target a broad spectrum of serine, cystein, metalloproteases and calcium-dependent, non-lysosomal cysteine proteases, calpains, to protect proteins from degradation. Cell lysate was mechanically harvested with a scraper and sonicated to shear DNA and reduce sample viscosity. Proteins were purified by centrifugation and the supernatant stored at -80°C.

E. Quantification of DNA, RNA and protein concentrations:

Concentrations of DNA, RNA and protein were quantified by UV/Vis spectroscopy using a NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Inc., USA). The NanoDrop® ND-1000 spectrophotometer is a full-spectrum (220-750 nm) spectrophotometer, using a pulsed xenon flash lamp as light source and a linear charge-coupled device (CCD) array for analysis of the light after passing through the sample.

Purines, pyrimidines and amino acids with aromatic rings absorb UV-light at certain wavelengths. For both RNA and DNA this maximal absorbance occurs at 260nm (A260). For proteins it occurs at 280nm (A280). The Beer-Lambert law predicts that the absorbance at a particular wavelength is proportional to the concentration of nucleic acid, the light path length from the source to the detector and the molar extinction coefficient or alternatively the specific absorption coefficient of the measured molecule.

Nucleic acid concentrations are measured through their absorbance at 260nm, in a range between 2ng/µl to 3700ng/µl (dsDNA), 3000ng/µl (RNA) or 2400ng/µl (ssDNA). Purity of DNA and RNA solution can be estimated looking at A_{260}:A_{280} ratio. A DNA preparation is considered as pure with a A_{260}:A_{280} ratio of ~1.8 and RNA with a A_{260}:A_{280} ratio of ~2.0. Generally, "pure" nucleic acid A_{260}:A_{230} ratio is higher than its A_{260}:A_{280} ratio and in a range of 1.8 to 2.0. A drastic lower A_{260}:A_{280} value may indicate presence of proteins, phenol or other agents and a drop of the A_{260}:A_{230} value may indicate co-purified contaminant.

Protein concentrations are measured through their absorbance at 280nm (A280), in a limited range from 0.10mg/ml to 100mg/ml (NanoDrop Technologies, 2005).

F. Reverse transcription

Retroviruses contain a RNA-dependent DNA polymerase, the reverse transcriptase, that allow them to transcribe their genomic single-strand RNA (ssRNA) into double-strand DNA (dsDNA). This enzyme can be use *in vitro* to convert RNA into cDNA, for example from biological samples like cellular total RNA extract. An oligodeoxynucleotide primer that hybridizes to the mRNA is needed to start DNA extension (Figure 5). A primer that is specific to a unique region of the total mRNA will allow to generate the copy of a unique gene. Oligo(dT) primer binds the endogenous poly$(A)^+$ tails of mammalian mRNAs and will generate cDNA copies for the total mRNA population. Random hexamer primer will generate fragmentary copies of the entire population of RNA molecules.

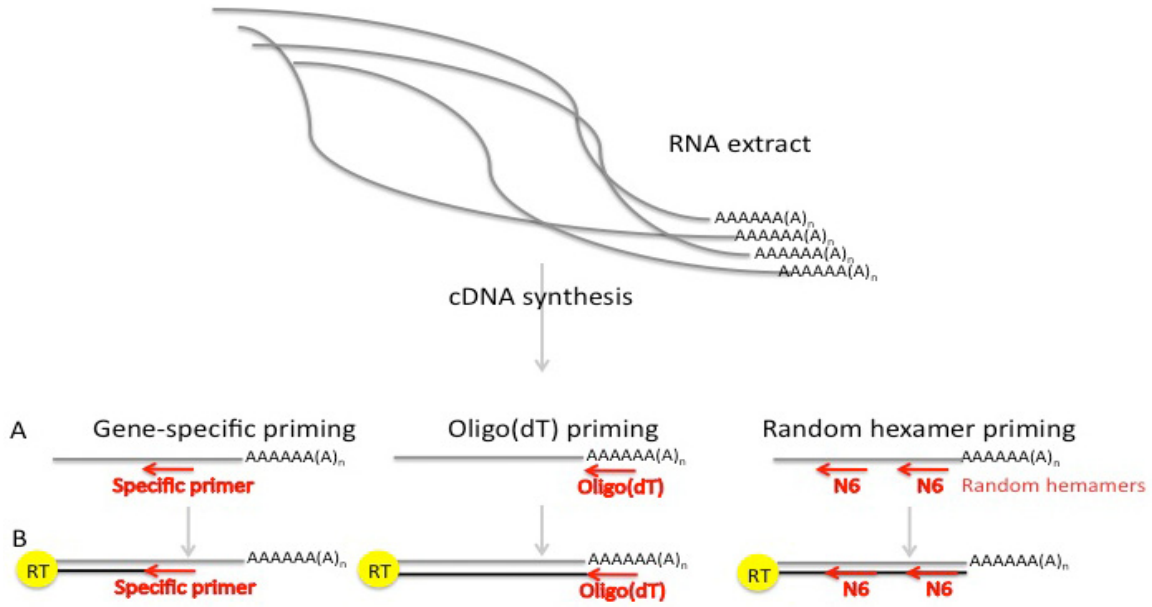


Figure 5. Alternative reverse transcription methods.

The reverse transcripts of BCL9 RNA were synthesized using the Moloney Murine Leukemia Virus Reverse Transcriptase, RNase H Minus (M-MLV RT [H-]) point mutant (Promega, # M368), which is efficient up to >5kb and lacks the RNase H activity to prevent degradation over long incubation times. Addition of a broad-spectrum RNase inhibitor, RNasin® Ribonuclease Inhibitors (Promega, # N211B), prevents degradation of the RNA template by RNases until the end of the process. 500ng of total mRNA was added to 2µl of the 1µM reverse primer, hBcl9_ORF_R2, heated at 70°C for 10 minutes and then incubated on ice for 10 minutes. To this mix was added 5µl of M-MLV RT buffer 5x, 1,25µl of deoxynucleoside triphosphates (dNTPs) 10mM, 1µl of RNasin®, 1µl of M-MLV RT (H-) and adjusted to 25µl with RNase free, sterile, H2O. It was incubated 10 minutes at 25°C, 2 hours at 42°C and 10 minutes at 95°C. The primer specific for the 3'UTR position in BCL9, hBcl9_ORF_R2: 5'-GGGAGAAGCCAGTCCAAAGGACG-3', was designed based on the NCBI mRNA BCL9 sequence in NCBI (NCBI reference sequence: NM_004326) and the consensus coding sequence (CDS) (ID CCDS30833.1) from NCBI CCDS database, using the Primer Express® Software v3.0 from Applied Biosystems. It was ordered at Microsynth.

For subsequent qRT-PCR, total RNA of cell lines was reverse transcribed using the PrimeScript™RT reagent Kit (TaKaRa, # RR037A), which consist of a mix of Oligo(dT) primers and random hexamer primers. Usually, 200ng of RNA extract was added to 2µl of 5x PrimeScript™ buffer, 0,5µl of PrimeScript™ RT enzyme mix I, 0,5µl of 50µM Oligo(dT) primer, 0,5µl of 100µM random hexamers and completed to 10µl with RNase free H2O. This was incubated 15 minutes at 37°C for the reverse transcription reaction and 5 seconds at 85°C to inactivate reverse transcriptase.

G. Quantitative real-time-PCR (qRT-PCR)

Quantitative real-time PCR is a method combining the cycling PCR and binding fluorescent dyes to DNA to quantify the relative amount of a specific mRNA. In this work, SYBR Green, who emits fluorescent light when bound dsDNA, was used. A reverse transcription usually precedes the qRT-PCR to convert the total cell RNA into cDNA. This cDNA is then amplified, which results in

increased fluorescence. At the start of the reaction, templates are at a low concentration to avoid that product renaturation competes with primer binding to the template. This should ensure amplification with a constant and exponential rate. After a certain number of cycles, which is variable among samples, the reaction rate ceases to be exponential due to product renaturation and becomes linear. After a certain time, the reaction reaches a saturation state. Only the exponential phase is extremely reproducible (Dharmaraj, 2010). To compare the different samples, an artificial fluorescent threshold, named cycle threshold (Ct), is fixed in the exponential phase. Then the number of cycles needed to reach it is determined.

As SYBR Green is a non-specific fluorophore binding any dsDNA, including primer dimer or other non-specific reaction product, detected samples should be analyzed individually to exclude data that are non-specific. Before data analysis, the dissociation curves of the samples were determined. The objective of dissociation curve analysis is to determine the melting temperature (Tm) of each qRT-PCR DNA product and identify presence of non-specific product as primer-dimer products. The Tm of specific product is higher than the Tm of primer-dimer products as the Tm depends on the length of the dsDNA. To determine the Tm of the samples, the qRT-PCR machine increases gradually the samples temperature and records changes of SYBR Green dye fluorescence. As the dsDNA dehybridizes around its Tm, SYBR Green dye fluorescence will decrease and the dissociation curve will display a peak at this temperature. Typically, a sample containing non-specific and specific qRT-PCR products will show two peaks.

The SYBR Green master mix used ((Applied Biosystems, # 4309159) contains an AmpliTaq® Fast DNA polymerase allowing instant hot start, minimizing non-specific product and reaction setup at room temperature. It contains also SYBR Green I dye, dNTPs and uracil DNA glycosylase (UDG) to reduce carryover contamination. 4,2µl of diluted cDNA (1:40) was mixed to 5µl of fast SYBR Green Master Mix 2x, 0,2µl of 22,5µM forward primer, 0,2µl of 22,5µM reverse primer and 0,4µl of RNase free sterile H_2O and loaded in one well of a 384-well plate in triplicate. To avoid contamination, plates were prepared under RNase and DNA decontaminated culture hood. After 1 minute centrifugation of the plate at 1000rpm, the qRT-PCR (Table 3) was run on the 7900HT Fast Real-Time PCR system from Applied Biosystems and analyzed with SDS2.3 program from Applied Biosystems. Then, the Ct values were normalized with the hydroxymethylbilane synthase (HMBS), a housekeeping gene involved in the metabolism of the porphyrin.

Normalized values, $\Delta Ct_{gene} = Ct_{sample} - Ct_{HMBS}$, were used for relative quantification of the genes between parental and modified cell lines. The relative expression was calculated as follow: $2^{-\Delta\Delta Ct}$ with $\Delta\Delta Ct = \Delta Ct_{gene,modified} - \Delta Ct_{gene,parental}$, , the percentage of gene expression compared to parental as follow: $100/2^{\Delta\Delta Ct}$ %, and the percentage of knock-down as follow: $100x(1-2^{-\Delta\Delta Ct})$%.

Stage	Temperature	Time	Cycles number
Start	50°C	2 minutes	1
Initial denaturation	95°C	10 minutes	1
Denaturation	95°C	15 seconds	40
Annealing-Polymerization	60°C	1 minute	
Dissociation curve	60°C-95°C	15 seconds	1
End	60°C	15 seconds	1

Table 3. qRT-PCR temperatures cycles

Primer pair sequences were as follow: Hmbs1&2_F4: 5'-CCAGCTCCCTGCGAAGAG-3', Hmbs1&2_R3: 5'-AAGCCGGGTGTTGAGGTTT-3', hApccd1_F2: 5'-ACTCAGACCCGGTGTGCAA-3', hApccd1_R1: 5'-GCTGTGGCCGCATCCA-3', hAxin2_F2: 5'-GCAGTGTGAAGGCCAATGG-3', hAxin2_R2: 5'-GCAGGCGGTGGGTTCTC-3', hFscn1_F1: 5'-GCGGCCAACGAGAGGAA-3', hFscn1_R2:

5'-TCCGAGTCCCCTGCTGTCT-3', hBCL9L_F: 5'-CCAATCTCAGCACCAGAATGTG-3', hBCL9L_R: 5'-TTGCTAGGCGAGATCTGGTTG-3', hBCL9_F: 5'-TCTCCAAAACCTCCCCTTCAG-3', hBCL9_R: 5'-GATATTCACAGAGGCAGGCTGG-3'.

H. Polymerase chain reaction (PCR)

The polymerase chain reaction is an *in vitro* technique that amplifies DNA templates. This reaction requires a thermostable DNA polymerase that synthesizes DNA accordingly to its DNA template, two oligonucleotide primers to prime reverse and forward synthesis, deoxynucleoside triphosphates (dNTPs) and template DNA or cDNA. PCR reaction is a succession of different temperature steps. The first one, named denaturation, around 95-98°C, unwinds and separates dsDNA. The second one, named annealing, is the temperature at which primers optimally binds to the complementary DNA sequence. The third one, named polymerization, is the optimal temperature for the extension of the oligonucleotide primers by the DNA polymerase. Repetition of these three steps together leads to exponential ($2^{\text{number of cycles}}$) amplification of the DNA template. A final polymerization step ensures that all single-stranded DNA (ssDNA) are fully extended. All primers were designed with Primer Express® Software v3.0 from Applied Biosystems and ordered at Microsynth.

1. Taq PCR

Amplification of specific regions in Bcl9 genomic DNA or cDNA was performed using a standard PCR with the Taq DNA polymerase (Biolabs, # M0267L). This thermostable DNA polymerase lacks a 3' to 5' exonuclease proofreading activity. 500ng of DNA, 1µl of cDNA or 500pg of BCL9 plasmid prep was mixed with 5µl of Thermopol Reaction buffer 10x, 1µl of 10mM dNTPs, 1µl of 10µM forward primer, 1µl of 10µM reverse primer, 0,25µl of Taq Polymerase and H_2O to a final volume of 50µl. The PCR reaction was as follow:

	Time	Temperature	Cycles
Initial denaturation	30 seconds	95°C	1x
Denaturation	30 seconds	95°C	35x
Annealing	60 seconds	52-60°C	
Polymerization	49 seconds	68°C	
Final polymerization	5 minutes	68°C	1x

Table 4. PCR temperatures cycles with Taq.

Primers were designed based on the NCBI DNA BCL9 sequence (NCBI reference sequence: NC_000001.10), mRNA BCL9 sequence (NCBI reference sequence: NM_004326) and the consensus coding sequence (CDS) (ID CCDS30833.1) from NCBI CCDS database. Primer sequences were as follows: reverse primers: Bcl9_intron1_R1: 5'-CCTACTCTCCCCACTTGTCTGATATT -3', Bcl9_intron6_R1: 5'-TTCCCCAGGACAGTGAAGGA-3', hBcl9_Exon3_R3: 5'-TGTGGGCTCTGTGGCAGTAGT -3', hBcl9_Intron6_R2: 5'-TGGCTGGGTCATTCCCATT -3', and forward primers: Bcl9_intron1_F1: 5'-CCTTGACCCCCAATAGAAACTG-3', Bcl9_intron6_F1: 5'-TTGTATGTCCCCAGATGTATTATGATC -3', hBcl9_Exon1_F3: 5'-CCCTACAGTGATGTCCCCATCT -3', hBcl9_Exon6_F1: 5'-CAAGTTTGCAATGCCCAGTTC-3'.

2. Phusion™ PCR of BCL9 cDNA

Amplification of the 4,824kb BCL9 cDNA was performed using a standard PCR with the high-fidelity DNA polymerase, Phusion™ (Finnzymes, # F-518). This polymerase was chosen because of its low error rate (4.4E-7), its proofreading activity and because of the large size of the

synthesized cDNA. 5µl of cDNA was mixed with 23.5µl H_2O, 10µl Phusion™ HE buffer 5x, 5µl of each 10µM primers, hBcl9_ORF_R1 (reverse) and hBcl9_ORF_F1 (forward), 1µl of 10mM dNTPs and 0,5µl of Phusion™ DNA polymerase. The thermo cycles were the following:

	Time	Temperature	Cycles
Initial denaturation	30 seconds	98°C	1x
Denaturation	7 seconds	98°C	30x
Annealing	30 seconds	56°C	
Polymerization	3 minutes	72°C	
Final polymerization	7 minutes	72°C	1x

Table 5. PCR temperatures cycles with Phusion™

Primers specific for the 5'UTR and 3'UTR position in BCL9 were designed based on the NCBI mRNA BCL9 sequence in NCBI (NCBI reference sequence: NM_004326) and the consensus coding sequence (CDS) (ID CCDS30833.1) from NCBI CCDS database. Primer sequences were as follows: forward primer, hBcl9_ORF_F1: 5'-CGGTGCAACACATGACCGAGG-3', and PCR reverse primer, hBcl9_ORF_R1: 5'-AGGAGGCCGCACCCAGAGA-3'.

a) A-overhang adding

In contrast to Taq DNA polymerase, Phusion™ polymerase does not add A-overhangs at the 3' end of the DNA product it synthesizes. Therefore, when needed for ligation of the PCR product with T-overhang bearing vector, A-overhangs at the 3'end were added to the RT-PCR cDNA product using the Taq DNA polymerase (Biolabs, # M0267L). 28µl of the extracted cDNA was incubated 20 minutes at 72°C with 0,64µl of 10mM dATP, 3,2µl of Buffer Thermo Pol 10x and 0,2µl of Taq DNA polymerase.

3. Phusion™ PCR of shRNAs

Amplification of the BCL9 shRNAs was performed with the Phusion™ DNA polymerase because of its low error rate. 10ng of the oligo template was added to 20µl fo 5x GC buffer, 5µl of 100% DMSO, 2µl of 10µM MIR30_XhoI_F, 2µl of 10µM MIR30_EcoRI_R, 2µl of Phusion™ (Finnzymes, # F-518) and 60,5µl of H_2O.

The reaction steps are listed in Table 6.

	Time	Temperature	Cycles
Initial denaturation	3 minutes	98°C	1x
Denaturation	10 seconds	98°C	10x
Annealing	10 seconds	54°C	
Polymerization	10 seconds	74°C	
Denaturation	10 seconds	98°C	20x
Annealing	10 seconds	64°C	
Polymerizaton	10 seconds	74°C	
Final polymerization	7 minutes	74°C	1x

Table 6. Lentiviruses PCR with Phusion™

Primers were designed specific for shRNAmir oligonucleotides and so to add restriction sites to the newly synthesized product. Primer sequences were as follow: MIR30_XhoI_F5'-CAGAAGGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG-3', and MIR30_EcoRI_R5'-CTAAAGTAGCCCCTTGAATTCCGAGGCAGTAGGCA-3'.

I. Purification of PCR product

PCR-amplified DNA was using QIAquick PCR purification kit (spin column format) from Qiagen (# 28104) according to the manufacturer's instruction. In brief, PCR product were mixed in a buffer at the appropriate pH for DNA binding in a silica membrane. DNA solution was loaded on the membrane. Salts, enzymes, unincorporated nucleotides (up to 40 nucleotides), agarose, dyes, ethidium bromide, oils and detergents do not bind the silica membrane and flow through the column. Salts are further washed away with ethanol containing buffer. Finally, the DNA was eluted from the silica membrane using 30-50µl of low-salt basic buffer and centrifugation. The kit allows purifying DNA product from 100 bp to 10kb.

J. Agarose Gel Electrophoresis

Agarose gel forms a mesh that slows down DNA according to their sizes. An electrical field is the driving force that allows DNA, a negatively charged molecule, to move towards the positive pole and migrate through the gel. According to the expected product size, the gel contains from 0.5% to 2% agarose dissolved in 1x trishydroxymethylaminomethane(Tris)-borate-EDTA (TBE) buffer with 0.005% of 0.10mg/ml ethidium bromide. Ethidium bromide contains a planar tricyclic phenanthridine ring that can intercalate between the base pairs of dsDNA and make van der Waals contacts with them. It also absorbs UV light of 302 to 366nm and re-emit at 590nm. Furthermore fluorescence of ethidium bromide is greatly enhanced by its contact to DNA.

In this work, DNA product was supplemented with a 5x DNA loading dye containing xylene cyanol and bromophenol blue. These two agents migrate at known sizes and help to approximate location of DNA molecules during the electrophoresis. In 0.5 to 2% agarose gels, xylene cyanol migrates around 3kb and bromophenol blue around 500 bp. In higher agarose concentration gels, xylene cyanol migrates approximately around 1kb and bromophenol blue around 100 bp. Gel elecrophoresis was performed at 100V in wide mini-sub® cell GT cuves (Bio-Rad). DNA ladders (50 bp, Fermentas, # SM0373, and 1kb, Fermentas, # SM0311) mixed with 6x DNA loading dye and composed of DNA fragments with known sizes was loaded parallel to the DNA samples. Separated DNA fragments were visualized and photographed with a UV transilluminator, Red™ Imaging System from Alpha Innotech. If needed, they were also visualized on a UV transilluminator table from UVP, then selectively cut with surgical blades for gel extraction.

K. DNA fragment Gel extraction

DNA was extracted from agarose gel using QIAquick Gel Extraction kit (spin column format) from Qiagen (# 28704) according to the manufacturer's instruction. In brief, cut bands were melt in a solubilization buffer at the appropriate pH for DNA binding on a silica membrane. Salts, enzymes, unincorporated nucleotides (up to 40 nucleotides), agarose, dyes, ethidium bromide, oils and detergents do not bind the silica membrane and flow through the column. Salts are further washed away with ethanol containing buffer. Finally, the DNA was eluted from the silica membrane using 30-50µl of low-salt basic buffer and centrifugation. The kit allows extracting DNA product from 100 bp to 10kb.

L. Restriction Digest

Restriction enzymes are enzymes found in bacteria that can cut double-stranded or single-stranded DNA at specific sites, called restriction sites. Three different restriction digest were performed, NotI (New England Biolabs, # R0189s) digest, NotI (New England Biolabs, # R0189s)/ AflIII (New England Biolabs, # R0541) digest and EcoRI (Invitrogen, # 15202-021)/XhoI

(Invitrogen, # 15231-012) digest. Single digest, NotI, contained 1µl of the restriction enzyme NotI, 10% of reaction buffer 3 (NEB, #B7003s), >1µg of DNA, 10% of bovine serum albumin (BSA) and the total volume of the reaction was adjusted with nuclease free water. Double digests contained 1µl of each restriction enzyme, NotI/AflIII or EcoRI/XhoI, 10% of TAS buffer, 1-5g of DNA and the total volume of the reaction was adjusted with nuclease free water.

TAS buffer is a reaction buffer composed of 330µl of 1M Tris-acetate at pH7,9, 130µl of 5M potassium-acetate, 100µl of 1M magnesium-acetate, 50µl of 100mM dithiothreitol, 40µl of 100mM spermidine and 350µl of sterile H_2O. Spermidine substitutes the BSA, so a reaction with TAS does not need adding of BSA.

Restriction digest mixtures were incubated 2 hours at 37°C.

1. Nucleotide removal

After restriction digest, nucleotides should be removed from DNA using the Nucleotide removal kit (spin column format) from Qiagen (# 28304) according to the manufacturer's instruction. A buffer at the appropriate pH for DNA binding in a silica membrane was added to the digested product. DNA solution was loaded on the membrane. Salts, enzymes, unincorporated nucleotides (up to 40 nucleotides), agarose, dyes, ethidium bromide, oils and detergents do not bind the silica membrane and flow through the column. Salts are further washed away with ethanol containing buffer. Finally, the DNA was eluted from the silica membrane using 50µl of low-salt basic buffer and centrifugation. The kit allows purifying DNA product from 100 bp to 10kb.

M. T4 Ligation

DNA ligases are enzymes that are able to repair single-standed discontinuities of dsDNA. It can be used to ligate a DNA insert into a plasmid DNA vector. The maps of the two different vectors used are shown in Figure **6** (pGEM®-T easy) and Figure 7 (MA-344). For ligation, the amount of insert and vector was determined with the following equation:

(amount of vector (ng)) x (size of the insert (kb)) x (insert:vector ratio) =
(amount of insert (ng)) x (size of the vector (kb))

The insert:vector ratio should be between 1 and 3 as recommended.

For each insert the reaction mix was as listed in Table 7 and Table 8.

DNA inserts:	BCL9 cDNA	Positive Control	Negative Controle (H_2O)
2x Rapid Ligation Buffer (Promega)	5µl	5µl	5µl
pGEM®-T Easy Vector (Promega, # A1360)	50ng	50ng	50ng
BCL9 PCR product	80-240ng	-	-
Control Insert DNA (Promega, # A363A)	-	2µl	-
T4 DNA ligase (Promega, # M180A)	1µl	1µl	1µl
Nuclease free water adjusted to a final volume of	10µl	10µl	10µl

Table 7. Ligation reactions for BCL9 cDNA.

The positive control is a DNA with a known size provided by the manufactured that allows to check afterwards the quality of the ligation reaction. The negative control reaction contains no DNA and allows to check self ligation of the vector and if the reagents does not contain any contaminant.

	BCL9 shRNAmirs
10x ligation buffer (Roche)	3µl
MA-344	200ng
shRNAmirs	9,1ng
T4 DNA ligase (Roche, # 10799009001)	1µl
Nuclease free water adjusted to a final volume of	30µl

Table 8. Ligation reactions for BCL9 shRNAmirs.

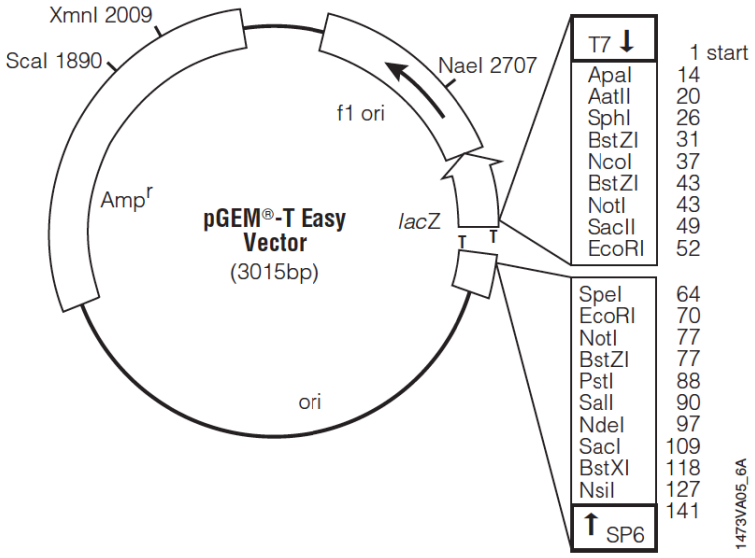


Figure 6. pGEM®-T Easy Vector Map and Sequence Reference Points from Promega Technical manual.

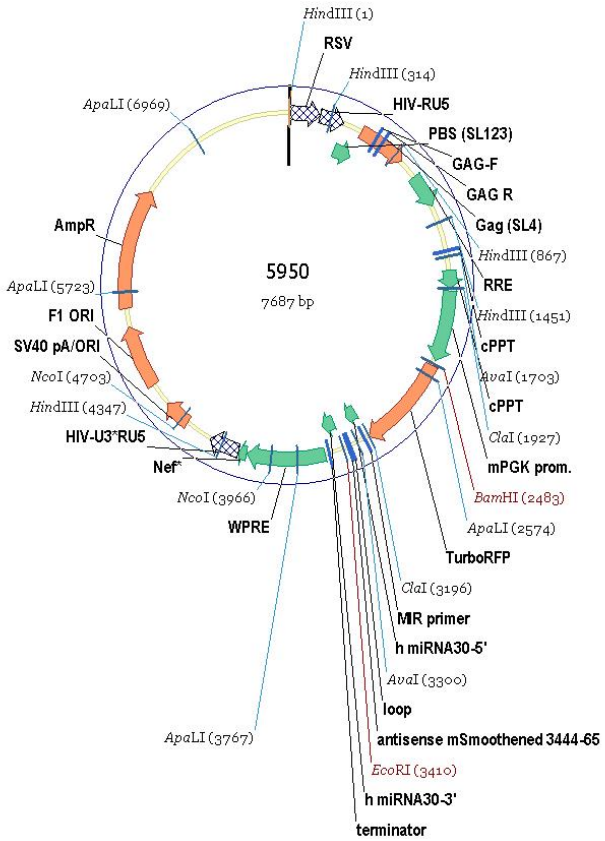


Figure 7. MA-344 vector also named pLVshRNAmir 5950.

N. Bacterial culture

Bacteria were culured in LB medium with ampicillin (100µg/ml) at 37°C and 220rpm agitation, overnight (maximum 16 hours). LB medium is a mixture of 10g of bacto-tryptone, 5g of bacto-yeast extract and 10g of NaCl in 900ml of H_2O. The pH is adjusted to 7,0 with 10M NaOH. The solution is then completed to 1L with H_2O and autoclaved. For long time storage, 800µl of bacteria culture was mixed to 200µl of glycerol and frozen at -80°C.

1. Bacterial Transformation

To enter a bacterial cell, plasmids require assistance to go through the outer and inner cell membranes. This process, called transformation, forms pores on the outer membrane, lowers inner membrane potential and therefore, facilitates DNA entry in the cell. Two different procedures were used: heat-shock transformation with chemically-competent cells and electroporation with electro-competent cells. Chemically-competent cells are DH5-α *E. Coli* treated with cold 50mM calcium chloride and electro-competent cells DH5-α *E. Coli* washed with cold water.

DH5α *E. Coli* cells transformed with pGEM®-T easy vector through standard heat shock or electroporation, were plated on LB agarose plate with ampicillin (100µg/ml), X-gal (20mg/ml) and IPTG (100mM). Plasmids used contain ampicillin resistance gene, which allow to select transformed bacteria. The insert integration interrupts the coding sequence of β-galactosidase, prevents its expression induced by IPTG and the formation of blue product from X-gal. The bacteria bearing pGEM®-T easy vector with an insert will stay white and bacteria bearing

pGEM®-T easy vector without insert will become blue. This allows to screen for colonies positive for the insert.

DH5α *E. Coli* cell transformed with MA-344 vector through electroporation were plated on LB agarose plate with ampicillin (100µg/ml). After overnight incubation at 37°C, clones were picked and put in liquid LB medium with ampicillin (100µg/ml) overnight. Plasmids were extracted by mini- or midi-prep.

a) Heat-shock transformation

10µl of ligation product, was added to 80µl of chemically competent bacterial cells and were incubated on ice for 30 minutes. Then they were briefly (1 minute) heat-shocked at 42°C followed by 5 minutes on ice. 900µl of LB medium was added to the cells, followed by a 2 hours incubation at 37°C in a shaker.

b) Electroporation

The ligation product was first precipitated with 1-Butanol, washed with 70% ethanol, air-dried and redissolved in 6µl of sterile, RNase free water. 3-4 µl of this was added to 60µl of electro-competent bacterial cells and put in electroporation chambers (Gene Pulse Cuvette, Bio-Rad, # 165-2089). The electric shock was provided by an electroporator (Bio-Rad Gene Pulser) set at a resistance of 200Ω, a capacitance of 250µF and a voltage pulse of 1.65. The time constant which determines the quality of the electric pulse should be between 2.5 and 4.5 to have an efficient transformation. 900µl of SOC medium was immediately added to the cells and the mix was incubated at least 45 minutes at 37°C in a shaker. SOC medium is a mix of 20g of bacto-tryptone, 5g of bacto-yeast extract, 0,5g of NaCl, 2,5ml of 1M KCl, 10ml of $MgCl_2$ and 5ml of 1M glucose in 900ml H_20. The pH is adjusted to 7,0 with 5M NaOH and the volume completed to 1L with H_2O. The medium is sterilized by autoclaving.

O. Plasmid extraction from transformed bacteria (*E. Coli*)

Plasmids were isolated from bacteria using a standard alkaline lysis, in combination with the detergent SDS and purified over resin column. The strongly anionic detergent at high pH breaks the bacterial walls, denatures proteins and chromosomal DNA by disrupting the base pairing, and releases plasmids. Plasmids remain intact thanks to their intertwined topology. They are then purified through an anion-exchange resin. Plasmids were isolated using Miniprep Jet quick 2.0 plasmid kit (Genomed, # 400 250), Qiagen Plasmid Midi kit (Qiagen, # 12143) and Maxiprep Jet quick 2.0 plasmid kit (Genomed, # 200 20).

1. Miniprep

4ml of overnight bacteria culture were centrifuged at 5000rpm for 5 minutes, resuspended in an RNase-containing solution and lysed in an alkaline SDS lysis buffer. The lysate is then neutralized by addition of potassium acetate which causes potassium dodecyl sulfate, genomic DNA, proteins and cell debris to precipitate. Then it was applied on a spin column with the anion-exchange resin that bind the DNA. Proteins, RNA and other impurities were washed away. Plasmids were eluted with 75µl of TE buffer by centrifugation.

2. Midiprep

100ml of overnight bacteria culture was centrifuged at 4300rpm, 4°C for 30 minutes, resuspended in an RNase-containing solution and lysed in an alkaline SDS lysis buffer. After

neutralization, the lysate was centrifuged at 4500rpm, 4°C for 1 hour. The supernatant was again centrifuged at 4500rpm, 4°C, for 30 minutes and applied on equilibrated column with the anion-exchange resin. Proteins, RNA and other impurities were washed away. Plasmids were eluted by gravity flow, precipitated with isopropanol, washed with 70% ethanol, air-dried and redissolved in 200µl of 5mM Tris-Cl, pH8.8 solution.

3. Maxiprep

500ml of overnight bacteria culture was centrifuged at 4300rpm, 4°C, for 30 minutes, resuspended in an RNase-containing solution and lysed in an alkaline SDS lysis buffer. After neutralization, the lysate was centrifuged at 4300rpm during 30 minutes and its supernatant was applied on an equilibrated column with the anion-exchange resin. Proteins, RNA and other impurities were washed away. Plasmids were eluted by gravity flow, precipitated with isopropanol, washed three times with 70% ethanol, air-dried and resuspended in 400µl of 10 mM, pH 8, Tris-chloride solution.

P. Sequencing

In 1977, Sanger et al. developed a DNA sequencing method using the capacity of DNA polymerase to integrate dideoxynucleoside triphosphate (ddNTP), a nucleoside analog instead of dNTPs at random positions during DNA strand copying. Furthermore, integration of theses ddNTPs induces a premature termination of DNA synthesis in a base specific manner. These ddNTPs lack 3'-hydroxyl group necessary for the phosphodiester linkage with next nucleoside and so, DNA polymerase cannot continue the DNA synthesis. As these terminations are random, DNA amplification with ddNTPs produces DNA fragments with various lengths that could be separated according to their size through capillary electrophoresis. Each of the four dideoxynucleotides terminators are labeled with a specific Big Dye. Big Dyes contain a fluorescein isomer as the donor dye and four dichlorhodamine dyes as the acceptors. Each of them have different wavelengths of fluorescence and emission allowing identification of the ddNTP linked to it. For DNA sequencing, BigDye® Terminator v1.1 Cycle Sequencing Kit from Applied Biosystems (# 4336774) was used. The big dye terminator mix contains AmpliTaq DNA polymerase FS, Big Dye labeled ddNTPs, dNTPs and an 5 x optimized buffer. Specific primers and DNA templates were added to the mix. The reaction setting was the following: 1µl of Big Dye Terminator, 1.5µl of 5x Big Dye buffer, 1µl of primer at 3µM, 500ng of DNA and H_2O to reach a 10µl reaction. The temperature cycles for the amplification were 95°C for 3 minutes, 30 cycles of (95°C for 30 seconds, primer temperature for 30 seconds and 60°C for 1 minute) and conservation at 4°C. If needed, samples were stored afterwards at -20°C. The DNA produced was precipitated by adding 25µl of 100% ethanol with 1µl of 3M sodium acetate at pH 5.2, incubating 10 minutes at room temperature and centrifugation at 13000rpm for 15 minutes. The pellet was then washed with 80% ethanol, air-dried and resuspended in 20µl highly deionized formamide (Hi-Di). Automatic capillary electrophoresis were performed by the sequencing service at CHUV, using the 3130xl Genetic Analyzers from Applied Biosystems. Results were sent back as chromatogram files.

Sequences of primers sequ_F1 to F5 and primers sequ_R2 to R11 were designed from Primer3 program (Skaletsky, 2000). Sequence of primer Bcl9_intron6_R2 was designed with Primer Express® Software v3.0 from Applied Biosystems. References for T7 and Sp6 promoter primers were given in the technical manual of pGEM®-T and pGEM®-T Easy Vector Systems (Promega) and their sequence on Promega website. Primer sequences were as follow: reverse sequencing primers: sp6: 5'-TATTTAGGTGACACTATAG-3', BCL9_sequR2: 5'CATGCCAGGTAATGCTAGTCC3', BCL9_sequR3: 5'TGGTGCATCGTGGGAGACT3', BCL9_sequR4: 5'-GGGTTATTCCCAGGCTCCAT-3', BCL9_sequR5: 5'-AGGGAACACTGCTGGACAAC-3', BCL9_sequR8: 5'-CTGGCTGGCTGTAGGAGGT-3'

BCL9_sequR11: 5'-CTTGTTGGAGAAAGGCTGGTGGT-3', and forward sequencing primers: t7: 5'-TAATACGACTCACTATAGGG-3', BCL9_sequF1: 5'-CTGGGTTTTCACACATGCAG-3', BCL9_sequF2: 5'-AGTATTTCCGCCGACTCCTT-3', BCL9_sequF4: 5'-ATCCCGCATTATTCCATCTG-3', BCL9_sequF5: 5'-GCCACCACTACCTCTCAACC-3', BCL9_sequR6: 5'-AGAGAATCCCGATGGCCTAT-3'.

Microsynth's primer walking sequencing service was also used. They began the sequencing from the two promoter primers, T7 and Sp6, that are respectively located before and after the insertional site in the pGEM®-T Easy vector.

shRNAmirs were sequenced by Microsynth's premium single strand sequencing service. We supplied the reverse sequencing primer: MA-344 rev: 5'-GGCCGCATTAGTCTTCCAATT-3'.

Chromatogram nucleotides sequences were aligned with the consensus coding sequence (CDS) (ID CCDS30833.1) from NCBI CCDS database using William Pearson's lalign program (http://www.ch.embnet.org/software/LALIGN_form.html) to identify point mutations and deletions.

Q. Western blot

The western blot is a technique used to detect specific proteins in a cell extract. Proteins are first separated by their size through SDS (sodium dodecyl sulfate)-polyacrylamide gel electrophoresis. Then proteins are transferred on nitrocellulose membrane. This membrane is then incubated with primary and secondary antibodies that recognize specific proteins. The secondary antibody is coupled to an enzyme that oxidizes a substrate and so induces chemoluminescence, which can be detected. With this method, a specific protein can be detected according to its known size, by the specificity of the primary antibody and the fluorescent band detected.

1. SDS-polyacrylamide gel electrophoresis

Tris-glycine SDS-polyacrylamide gel electrophoresis is a method to separate proteins according to their size by applying an electrical current. 150ng of protein extract was first treated with 4x Laemmli Buffer. This buffer contains 2,5ml of 1M Tris-HCl, 4ml of glycerol, 0,8g of SDS, 2ml of 0,1% bromophenol blue and 1,46ml of H_2O. 4% of β-mercaptoethanol was added just before use. SDS is a detergent that can dissolve hydrophobic molecules and so denature secondary and non-disulfide tertiary structures. As it bears negative charge, it ensures that all proteins are negatively charged in proportion to their mass and that they will only migrate according to their size and in the direction of the anode. β-mercaptoethanol reduces disulfides bonds due to the thiol group of the cysteins and so unfolds the proteins. The glycerol increases the density of the sample loaded in the gel and helps the sample to fall in the gel well. Bromophenol blue is a dye that allows checking the protein migration through the gel.

First, proteins are loaded in a stacking gel with a low density matrix and chloride ions. These leading ions migrate quicker than the proteins and so will trap them in a thin band. The second part of the gel, the resolving gel, has smaller pore size, a higher pH and higher salt concentration. This will induce the ionization of the glycine and allow proteins to be separated only based on their size. Due to the big size of our protein of interest, a gel with a low percentage was chosen allowing a better separation resolution. Composition of the stacking and the resolving gel is described in Table 9. Protein migration was performed in a running buffer composed of 3,02g of Tris (25mM), 18,8g of glycine (250mM), 5ml of 20% SDS and completed to a volume of 1L with H_2O. Protein migration was performed at 100V in the stacking gel and 120-150V in the resolving gel, in the Mini PROTEAN™ Tetra Cell system from Bio-Rad. A prestained

protein marker, PeqGOLD protein marker V (Peqlab, # 272110), with known peptide sizes was added to the gel to monitor time of migration.

Solution components	6% Resolving gel (2 gels)	Stacking gel (2 gels)
H_2O	10,6 ml	5,5ml
30% acrylamide mix (Bio-Rad, # 1610158)	4ml	1,3ml
1,5M Tris (pH8.8)	5ml	-
1,0M Tris (pH6,8)	-	1ml
10% SDS	200µl	80µl
10% ammonium persulfate	200µl	80µl
TEMED (Fluka, # 87689)	16µl	8µl

Table 9. Composition of the SDS-polyacrylamide gel.

2. Protein transfer

Proteins from the SDS-polyacrylamide gel are transferred on an Amersham Hybond ECL nitrocellulose membrane (GE Healthcare, # RPN2020D) using the Mini Trans-Blot cell chamber from Bio-Rad. The transfer buffer contains 2,9g of glycine (39mM), 5,8g of Tris (48mM), 3,7ml of 10% SDS and completed to 800ml by H_2O. 200ml of methanol was added just before use to minimize swelling of the gel during blotting and to increase the binding capacity of the membrane. Nitrocellulose membrane was pre-wet in the transfer buffer for at least 10 minutes. From cathode to anode, following elements, immersed in transfer buffer, were placed on top of each other in the holder cassette: foam pad, absorbent filter paper, SDS-page polyacrylamide gel, nitrocellulose membrane, absorbent filter paper and foam pad. Proteins were transferred overnight with a voltage of 30V at 4°C. As peqGold prestained marker proteins should also transfer, presence of their dye on the nitrocellulose membrane after transfer indicates that the transfer went well.

3. Immunodetection

The principle of this method is a 3-step detection. First, a primary antibody specifically binds the protein of interest. Then, a secondary antibody coupled to the horseradish peroxidase (HRP) specifically recognizes the primary antibody. Finally, the membrane is immersed in reagent, which contains amongst others luminol. HRP and hydrogen peroxide together induce excitation of the luminol by oxidation and so emission of a light spectrum with a maximum at 428nm until luminol reaches its ground state. Chemical enhancers can increase the light output approximately 1000 fold and extend the time of emission. Emitted light can be detected by a short exposure to blue-light sensitive autoradiography film.

First, nonspecific binding sites were blocked by incubation of the nitrocellulose membrane at room temperature, during 1 hour, under agitation, with tris buffer saline (TBS)-Tween plus 5% nonfat dry milk (blocking buffer). Then, the membrane was incubated 1-2 hours with the primary antibody (Table 10) diluted in blocking buffer. After three washes of 10 minutes with TBS-Tween, the membrane was incubated during maximum 1 hour with the secondary antibody (Table 10) diluted in blocking buffer. After three washes of 10 minutes with TBS-Tween and a short one with PBS, the membrane was incubated for 5 minutes with the Amersham ECL western blotting detection reagent (GE Healthcare, RPN2106). This reagent contains hydrogen peroxide, luminol and chemical enhancers as phenols in an alkaline buffer. Amersham Hyperfilm™ ECL (GE Healthcare, 2890637) were exposed to the emitted light during 1, 5, 20 or 60 minutes and then developed with the film processor, Konica SRX 101a. Marker protein sizes were then copied onto the film.

Primary antibody	Primary antibody dilution	Secondary antibody	Secondary antibody dilution
Anti-BCL9 (R&D, # AF3996)	1/1000	Anti-Goat/HRP (Dako, # P0160)	1/2000
Anti-BCL9L (R&D, # AF4967)	1/1000	Anti-Sheep/HRP (Dako, # P0163)	1/2000
Anti-V5 (GeneTech, # GTX30566)	1/250	Anti-Goat/HRP (Dako, # P0160)	1/2000

Table 10. List of primary and according secondary antibodies used and their dilution.

TBS-Tween buffer is composed of 3,025g of Tris (25mM, pH7,4), 87,6g of NaCl (0,15M), 0,05% of Tween®20 and completed to 1L by H_2O

R. shRNAs design

BCL9 shRNAmir sequences were designed by Open Biosystems (OB) or by Joerg Huelsken (JH) lab program, and ordered at Microsynth. JH program insert the anti-abortive product bulge between the 10th and the 11th nucleotide in the 3' arm shRNAmir sequence. Otherwise the two designs share the same features. Sequences used are the following: sh024_Bcl9: 5'-TGCTGTTGACAGTGAGCGCCGACCTCAGAGCAGAGTATTATAGTGAAGCCACAGATGTATAATACTCTGCTCTGAGGTCGTTGCCTACTGCCTCGGA-3' (OB), sh165_Bcl9: 5'-TGCTGTTGACAGTGAGCGCCACAGAGTAGCCCTAAGTCAATAGTGAAGCCACAGATGTATTGACTTAGGGCTACTCTGTGTTGCCTACTGCCTCGGA-3' (OB), sh161_Bcl9: 5'-TGCTGTTGACAGTGAGCGCAACGATGACTCTGACATTAAATAGTGAAGCCACAGATGTATTTAATGTCAGAGTCATCGTTTTGCCTACTGCCTCGGA-3' (OB), sh5090_JH_Bcl9: 5'-TGCTGTTGACAGTGAGCGCGTTCCAGGAGTCACTTACTATTTAGTGAAGCCACAGATGTAAATAGTAAGT-ACTCCTGGAACTTGCCTACTGCCTCGGA-3' (JH), sh5254_JH_Bcl9: 5'-TGCTGTTGACAGTGAGCGCTGTGGACTTGGCGTATCAATGATAGTGAAGCCACAGATGTATCATTGATAC-CCAAGTCCACAATGCCTACTGCCTCGGA-3' (JH), and sh5588_JH_Bcl9: 5'-TGCTGTTGACAGTGAGCGAACCTCAGAGCACGAGTATTAATTAGTGAAGCCACAGATGTAATTAATACTC-TGCTCTGAGGTCTGCCTACTGCCTCGGA-3' (JH). The loop sequence is in yellow, the mir-30 context in blue and the sense and antisense BCL9 targeting sequence in white.

S. Lentivirus production

The system used for lentivirus production is here the second generation implying three plasmids. The packaging plasmid, MA-201 (Figure 8) contains the Gag-Pol proteins, the envelope plasmid MA141 (Figure 9) contains the VSV-G protein gene, and the transfer vector MA-344 (Figure 7) contains the gene of interest and a fluorescent protein under the ubiquitous promoter of the phosphoglycerate kinase. Lentivirus-producing cells are the HEK293T as they are easily transfectable and grow easily. MA-344, MA-201 and MA-141 were lipotransfected with the FuGENE 6 transfection reagent from Roche (# 11814443). Plasmids are so integrated into

artificial cationic membrane vesicles (liposomes) that can fuse with negatively charged cell membrane.

$3x10^6$ low-passage HEK293T were seeded in 10cm dishes with 11ml of DMEM with 10% FBS. After 12 hours 5µg of MA-344, 3,75µg of MA-201, and 1,25µg of MA-141 were mixed. 18µl of Fugene 6 was pre-diluted into 582µl of serum-free DMEM and incubated for 5 minutes at room temperature. Then plasmid mix was added to it and further incubated for 25 minutes at room temperature. The latter was then added to the HEK293T cells in a drop-wise manner and the dish swirled to ensure distribution of the solution over the entire plate. Cells were incubated at 37°C during 48hours. Then, the supernatant was collected, filtered with a 0,45µm filter and centrifugated at 4°C, 19500rpm, for 1,5 hours under vacuum. Supernatant was quickly removed and the virus pellet immediately resuspended in 60µl of 10% FBS DME. Virus preparation were shaken 2 hours at 4°C and then aliquoted and stored at -80°C.

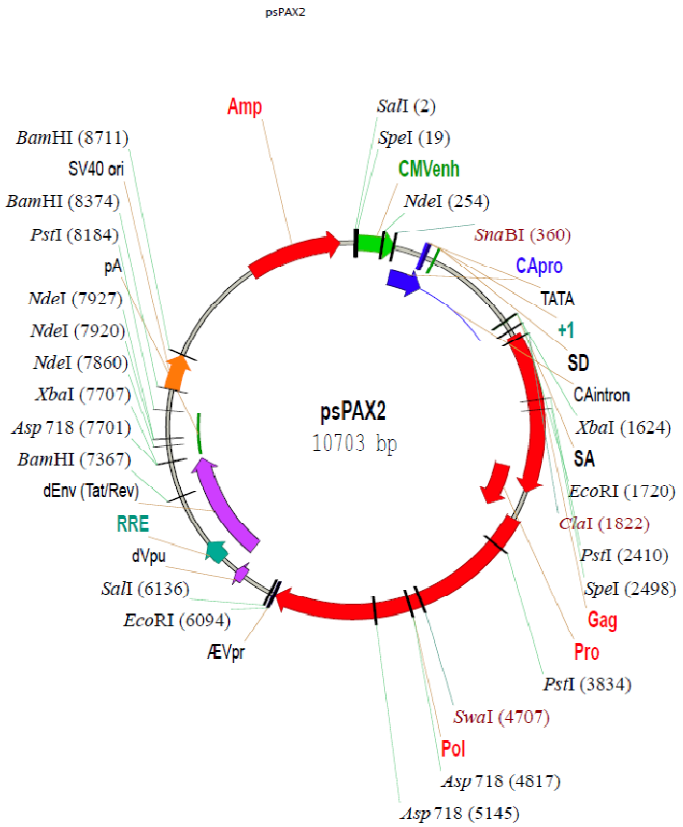


Figure 8. MA-201 packaging plasmid.

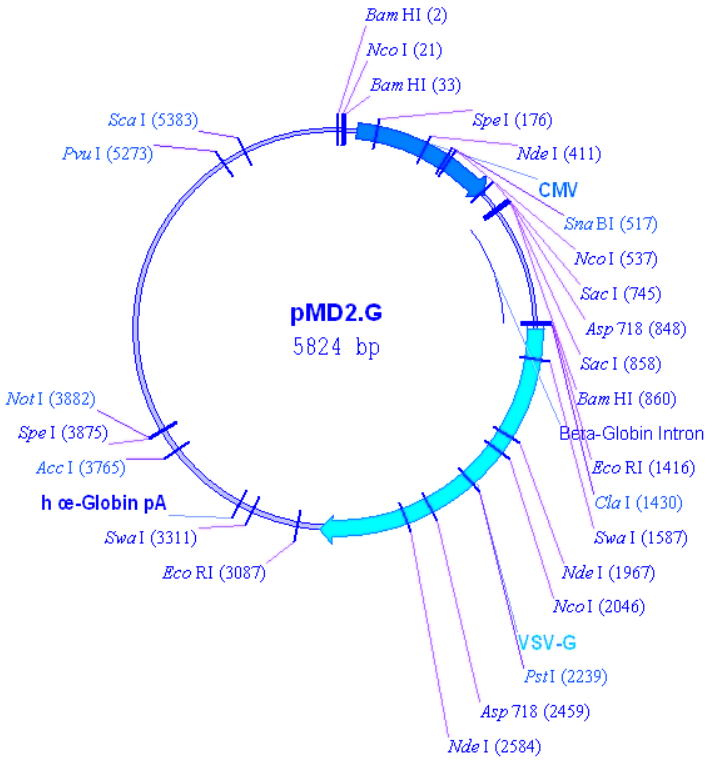


Figure 9. MA-141 envelope plasmid.

T. Virus infection

20000 cells per well were seeded in a 24-well plate. 12 hours later, cells were infected with the virus four times over 48 hours. Each round of infection was done with a MOI of 5 (Multiplicity Of Infection = number of infectious virus particles per cells). The Virus were diluted in culture medium containing 8µg/ml of polybrene and added to the cells. Polybrene is a cationic polymer that increases the efficiency of the infection by neutralizing the charge repulsion between the virus and the cell surface. After 12 hours of infection, medium was removed, cells were washed with PBS and cultured as usual.

U. Virus Titration

Virus titration determines the number of infectious particles in the preparation by infecting a known number of cells with different dilution of virus. Cells are considered infected when they are fluorescent because of the red fluorescent protein, turboRFP, expressed in the lentivector. TurboRFP is a protein excited by green light (maximum excitation at 553nm) and emits a red-orange light (maximum emission at 574nm).

50000 cells per well were seeded in a 24-well plate and incubated 24 hours. 24 hours is considered as the time needed for the cell population to double in number. At this time, the cells were infected with the viruses. The volumes of virus prep added to the cells were 0µl, 0,3µl (dilution 30), 1µl (dilution 10), 3µl (dilution 3) and 10µl (maximum volume, dilution 1). This was adjusted to a 600µl volume with culture medium containing 8µg/ml of polybrene. Number of cells expressing turboRFP was checked 48 hours later and the titer calculated. Assuming that the MOI is at the last dilution that gives 100% red cells, the calculation is the following: titer = number of viruses/µl of solution = (number of cells at day of infection (100000)) / (maximum volume of virus preparation x last dilution that showed 100% infected cells).

IV. Results

A. BCL9 mRNA in SW480 cell line sequencing

As a phenotype was observed in BCL9L stably knock-down SW480 cells despite of BCL9 expression, the functionality and integrity of BCL9 in SW480 cells was questioned. To see whether SW480 cells expresses wild-type, BCL9 mRNA was sequenced. BCL9 mRNA was reverse transcribed and amplified with BCL9-specific primers (Figure 10). A-overhangs were added to the obtained BCL9 cDNA and ligated into the T-overhang bearing pGEM-T easy vector.

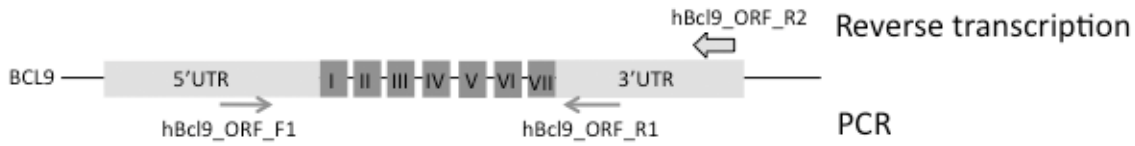


Figure 10. Primers used for reverse transcription of BCL9 mRNA and subsequent amplification.

Bacterial cells were transformed with the pGEM-T easy vector containing the BCL9 cDNA (BCL9-vector), an empty plasmid (negative control) or the pGEM-T easy vector containing a control DNA insert (positive control). 18 colonies (C1-18) of BCL9-vector transformed bacteria, 2 clones of the positive control and 2 of the negative control were picked. These plasmids were extracted by miniprep and >1µg of it was digested with NotI (Figure 11) to see whether they contain an insert of correct size (Figure 12). NotI digest of the negative control gives rise to a band of 2,981kb corresponding to the linearized pGEM-T easy vector. The positive control shows the 2,981kb of the pGEM-T easy vector and the 576 bp of the control DNA insert. 16 clones (C1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16 and 17) contain the 4,858kb insert of the BCL9 cDNA in addition to the 2,981kb pGEM-T easy vector and therefore should represent the BCL9 transcript population in the SW480 cells.

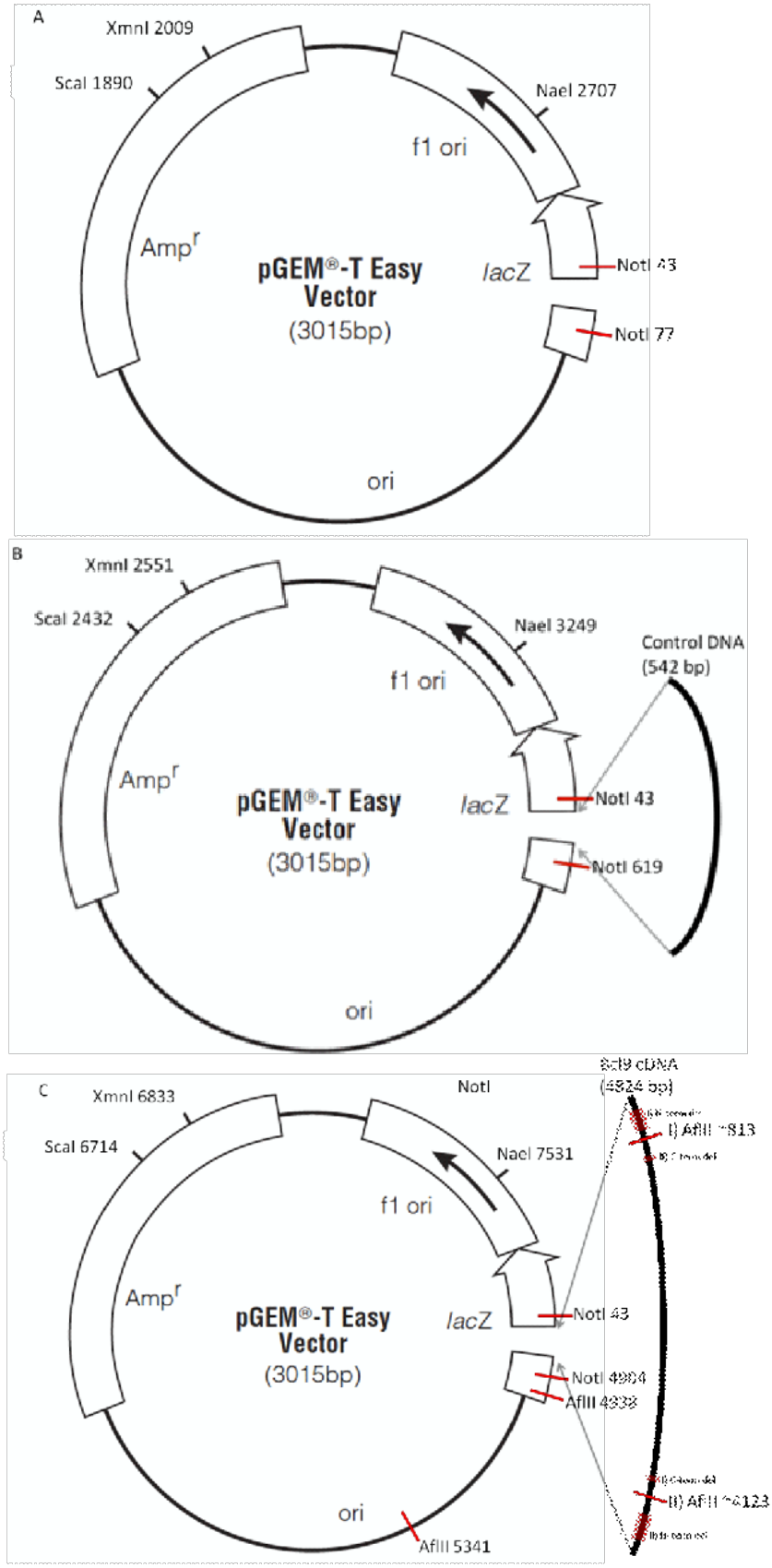


Figure 11. Restriction sites on: A: plasmid without insert. B: control DNA-containing plasmid. C: wild-type BCL9 cDNA-containing plasmids. The cDNA could be inserted in two different orientations, I or II. Positions of the deletions (I) or II) N-term del, and I) or II) C-term del) and the AflII restriction site (I) or II) AflIII) in BCL9 are depicted for both orientations.

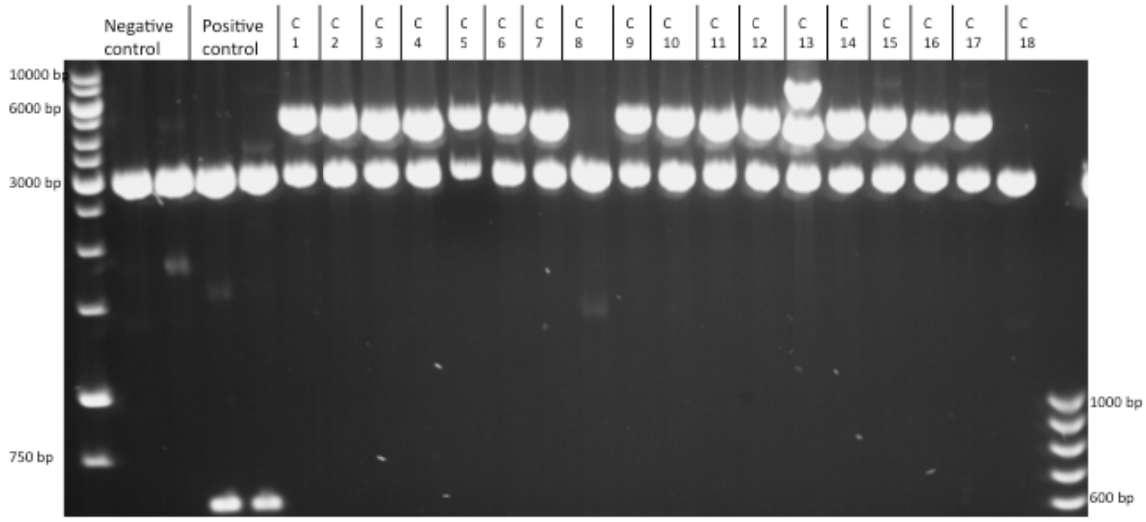


Figure 12. NotI digestion of BCL9 cDNA containing plasmid.

The BCL9 insert of clone 3 was sequenced by primer walking at Microsynth. The sequencing result shows two deletions and three points mutations compared to the reference sequence from NCBI. They are depicted below at nucleotide and amino-acid sequence level (Figure 13). Only one nucleotide mutation induces changes in the amino-acid sequence, a serine to proline transition, but it is not located in a known functional domain. A N-terminal deletion overlaps with the NHD of BCL9, which contains a NLS. However, this NLS is only functional in BCL9L (Townsley, Cliffe et al. 2004). With our current knowledge, we do not know whether this deletion has an impact on BCL9 function. A C-terminal deletion is located within the linker region between the CHD2 and CHD3 domains and could affect the transactivation function linked to this region (Sustmann, Flach et al. 2008). The two deletions do not change the reading frame of the sequence and do not insert any premature stop codon, and therefore they will not cause a truncated or largely different protein.

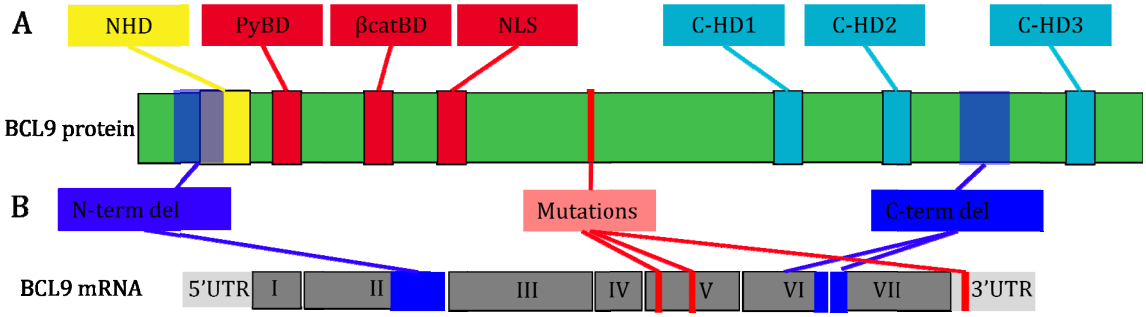


Figure 13. Conserved regions in BCL9 and deletions sequences. A BCL9 domains. B Deletions sequenced.

A NHD: NHD subdomain I: aa 94-123, NHD subdomain II: aa 124-139; PyBD: aa 177-205; βcatBD: aa 349-377; NLS: aa 469-483(Kramps et al., 2002); C-HD1: aa 839-982 (Adachi et al., 2004); C-HD2 and 3 with the linker region: aa 1002-1260. (Sustmann et al., 2008).

B Deletions and mutations sequenced: -N-terminal deletion: aa 50-123; -C-terminal deletion: aa 1055-1066; -First mutation: nt 1728 (A->G) in exon V, silent mutation; - Second mutation: nt 2011 (T->C) in exon V, aa 670 (S->P); -Third mutation: nt 4350, in the 3'UTR.

To test whether all BCL9 mRNA species contain these deletions, the other clones were tested for these two major deletions by restriction digest or specific PCR of the deleted regions.

The plasmid from clones 3, 4, 5, 11, 12 and 13 were extracted by maxiprep and 2µg was digested with AflII and NotI restriction enzymes. NotI only cuts the pGEM-T easy vector (Figure 11), while AflII cuts the pGEM-T easy vector twice and the BCL9 insert once giving rise to a short N-terminal fragment and a longer C-terminal fragment (Figure 11). The double digest should always produce a band of 403bp (AflII 4938-5341) and one of 2,455bp (AflII 5341-NotI 43). The size of the BCL9 N-terminal fragment should be 770bp if the clone is wild-type and 548bp if the clone has the N-terminal deletion. As shown in Figure 14, clones 4, 11 and 13 have a N-terminally deleted fragment, while clones 5 and 12 only have wild-type fragments. Clone 3 contains a mixed population of wild-type and N-terminally deleted fragment (Figure 14). The size of the BCL9 C-terminal fragment will be 4,089bp if the clone is wild-type and 4,053bp if the clone has the C-terminal deletion. However, no conclusions could be drawn for the C-terminal deletion, as this difference in size is very difficult to assess, as observed in Figure 14.

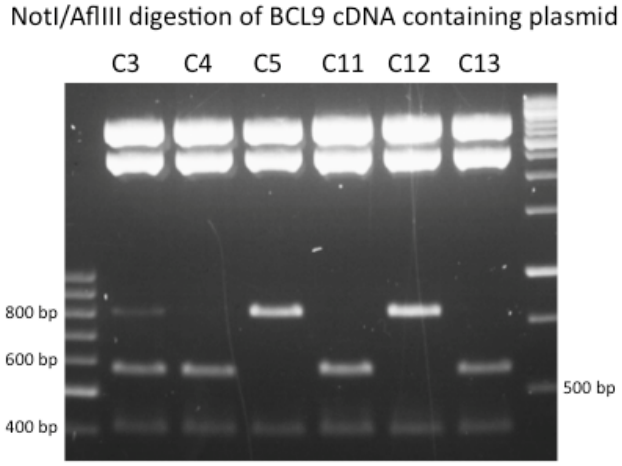


Figure 14. NotI/AflIII digestion of BCL9 cDNA-containing plasmid from clone C3, C4, C5, C11, C12 and C13.

The plasmids from clone 2, 4, 6, 7, 9 and 14 were extracted by maxiprep and the N-terminal region containing the N-terminal deletion was amplified (Figure 15, A). If the clone contains the N-terminal deletion, the PCR product should be 190bp, if the clone is wild-type, it should be 412bp. As shown in Figure 14 B, clones 4 and 7 have the BCL9 insert with the N-terminal deletion, while the BCL9 insert from clones 6, 9 and 14 is wild-type. The insert of clone 2 is a mixed population of wild-type and N-terminally deleted BCL9 (Figure 15, B).

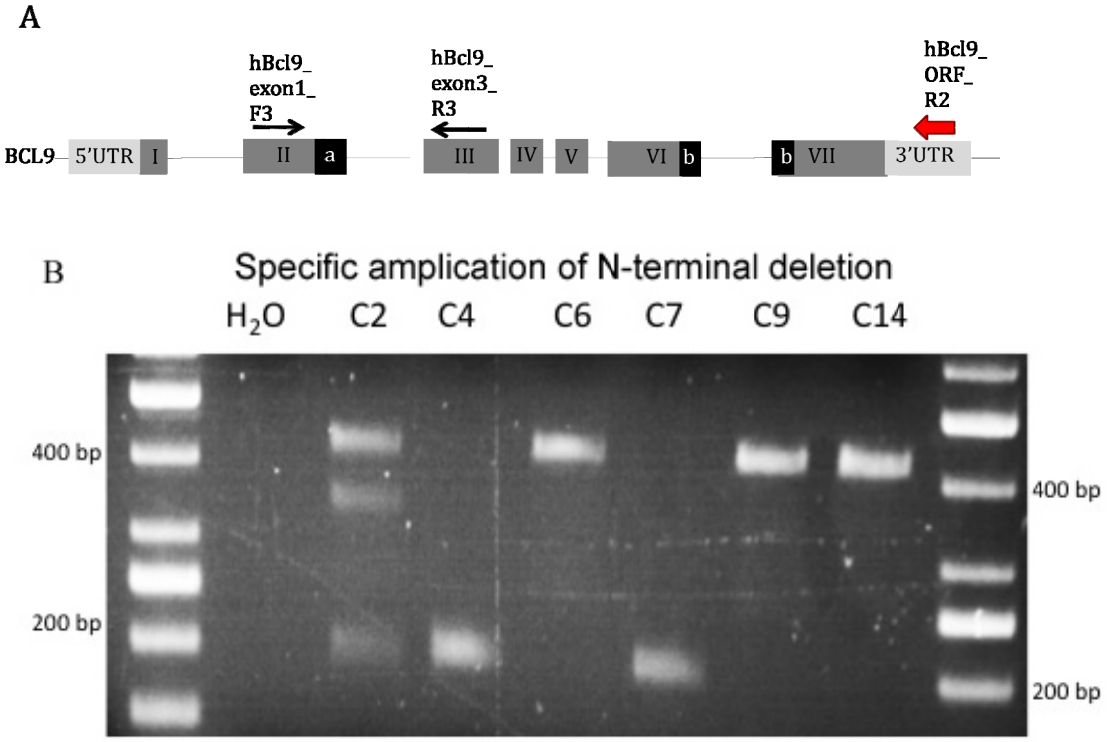


Figure 15. PCR of N-terminal fragment of BCL9 cDNA. A. Position of primers used and B. PCR results.

The plasmids from clone 3, 4, 5, 11, 12 and 13 were extracted by maxiprep and the C-terminal region of the BCL9 insert containing the C-terminal deletion was amplified with specific primers, which are shown in **Figure 16** A. If the clone contains a BCL9 insert with the C-terminal deletion, the PCR product should be 178bp, while a wild-type PCR product should be 214bp. From **Figure 16** B, it becomes clear that clone 4 contains a BCL9 insert with the C-terminal deletion, clones 5, 11, 12 and 13 a wild-type BCL9 insert and clone 3 a mixed population of C-terminally deleted and wild-type BCL9 insert (**Figure 16**, B).

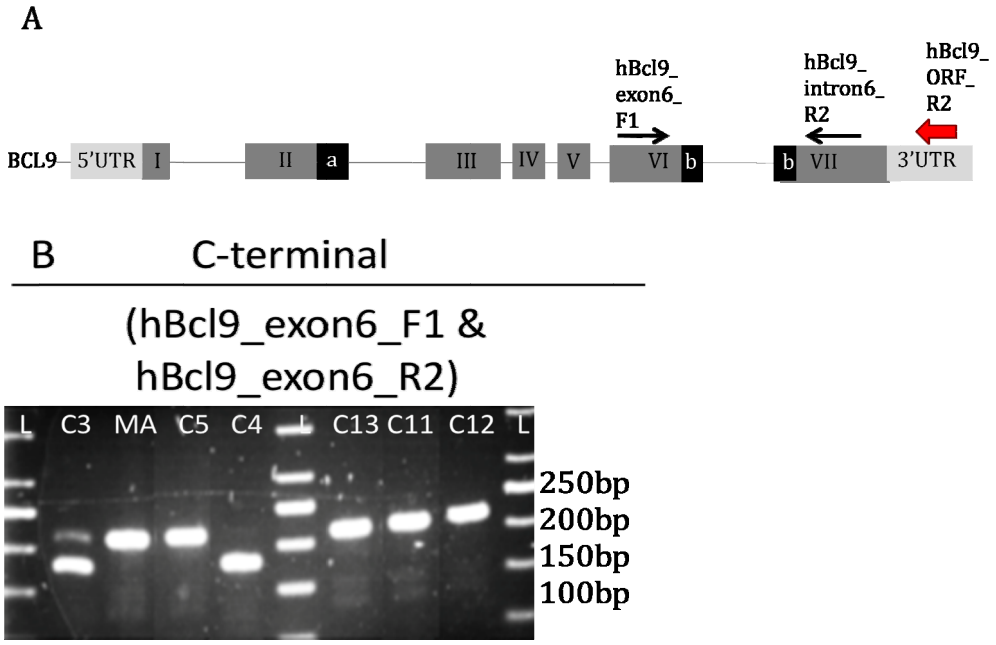


Figure 16. PCR of C-terminal fragment of BCL9 cDNA. A. Position of primers used and B. PCR results.

Clones	Digest N-terminus	PCR	
		N-terminus	C-terminus
C2	ND	Deletion/WT	ND
C3	Deletion/WT	ND	Deletion/WT
C4	Deletion	Deletion	Deletion
C5	WT	ND	WT
C6	ND	WT	ND
C7	ND	Deletion	ND
C9	ND	Deletion	ND
C11	Deletion	Deletion	WT
C12	WT	WT	WT
C13	Deletion	Deletion	WT
C14	ND	WT	ND
MA-365, a WT BCL9 cDNA	ND	WT	WT
ND = not done; WT = wild-type			

Table 11. Plasmid phenotypes for N-terminal digest, C-and N-terminal amplifications of BCL9 insert.

So 3 different populations could be observed: wt BCL9 insert, both N- and C-terminally deleted BCL9 insert and BCL9 insert with N-terminal deletion only. The mixed populations of clones 2 and 3 could be due to a contamination of one clone by another or integration of two different plasmids in the same bacteria. The different plasmids tested (C2, C3, C4, C5, C6, C7, C9, C11, C12, C13, C14 and MA-365) are listed in Table 11 with the analysis used (NotI/AflII digested, N- and C-terminal amplified) and the corresponding phenotype observed.

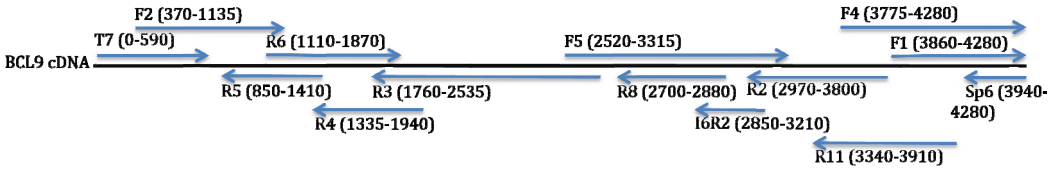


Figure 17. Positions of C12 sequenced fragments for each primer.

To exclude further anomalies in regions outside the C- and N-terminally deleted regions, we sequenced the entire BCL9 insert from a clone that did not contain the C- and N-terminal deletion (C12) (Figure 17). This sequence did not reveal any differences with the wild-type BCL9 sequence from NCBI CCDS database (ID CCDS30833.1). We can conclude that SW480 cells contain wild-type BCL9 mRNA in addition to at least two shorter forms.

B. BCL9 mRNA populations in other cell lines

To assess whether the heterogenous BCL9 mRNA population was intrinsic to SW480 cells, BCL9 mRNA from other colon (LS174T and COLO320 cells) and breast (MDA-MB231 and MCF7 cells) carcinoma cell lines was analyzed as for the SW480 cell line through BCL9-specific reverse transcription and specific amplification of the two deleted regions (Figure 18, A). The plasmids characterized previously (C3, C4, C5, C11, C12, C13 and MA-365) were used as controls for the different PCR products.

All cell lines show a wild-type (412bp) and a deleted band (190bp) in the N-terminal region (Figure 18, B). Intensity of the bands suggests that the proportion of the two populations is different between all the cell lines. COLO320 (Colo) and LS174T (LS) cells show a brighter wild-type band, while similar intensity for both wild-type and deleted band could be observed in SW480 (SW) and MDA-MB231 (MB) cells.

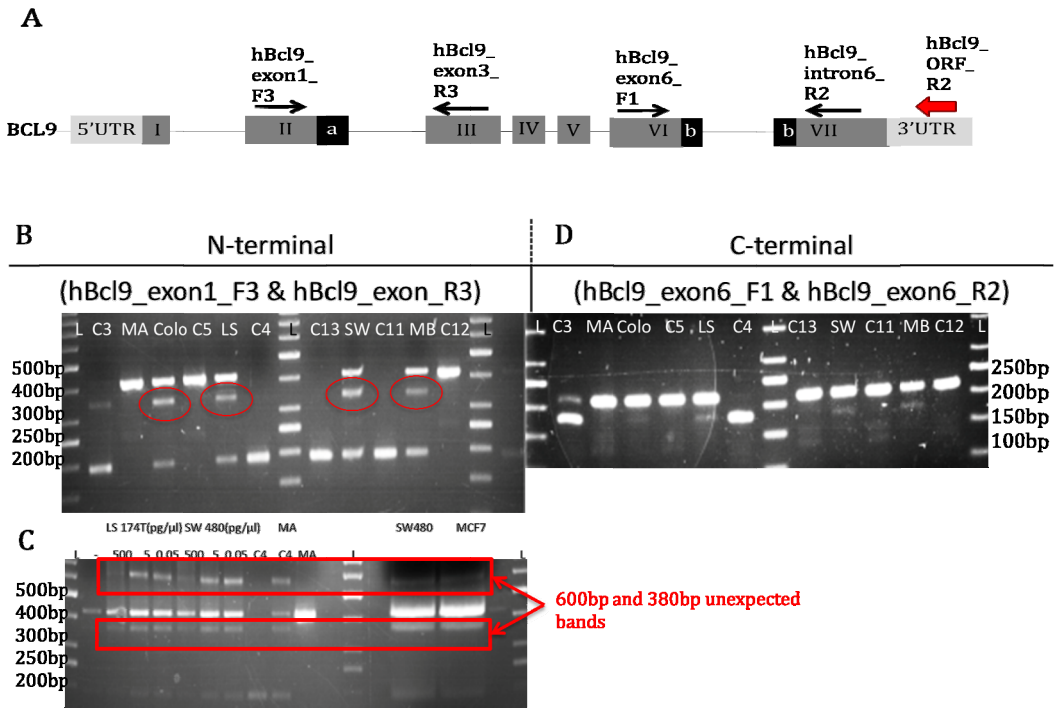


Figure 18. PCR primer position and PCR results of BCL9 transcripts. A. Primers used for the BCL9-specific reverse transcription (red arrow), for the N-terminal region (a) and the C-terminal region

(b) PCR (black arrows). B. N-terminal PCR results. C. Confirmation of the artifact. D. C-terminal PCR results.

Unexpected additional bands appear at ~380bp and ~600bp. If similar cDNAs are amplified together until saturation of the PCR reaction, two single strands that differ only by the deletion can hybridize together during the last annealing phase of the PCR. These asymmetric dsDNAs will migrate at unknown sizes as they do not have same negative charges repartition as symmetric dsDNA and have a physical structure different from symmetric dsDNA. To verify that the additional bands are indeed artificial, the same N-terminal region PCR was performed with a starting mixture containing equal amounts of a wild-type plasmid (MA-365) and a deleted plasmid (C4). If the unexpected band is caused by the mixture of N-terminally deleted DNA strand and wild-type DNA strand, this PCR will give rise to the four bands observed in the amplification of the total BCL9 cDNA population. If the unexpected band is another deleted mRNA transcript, only the N-terminally deleted C4 band (190bp) and the wild-type MA-365 band (412bp) will be observed. This PCR was also performed using DNA at different input (0.05, 5, 500 pg/µl) from one of the unexpected bands (380bp). If the extracted DNA from the unexpected band (380bp) contains the wild-type strand and the N-terminally deleted strand, its amplification will give the four bands observed in the amplification of the total BCL9 cDNA population. If the extracted DNA contains another deletion than the expected, its amplification will just give the same 380bp band. Both reactions confirm that the unexpected bands are artifacts caused by the presence of the deleted and the wild-type BCL9 cDNA (Figure 18, C).

All cell lines show both wild-type and deleted band for the C-terminal region of BCL9. There the difference of band intensity is more significant with the wild-type population Figure 18, D).

These results confirm presence of a heterogeneous population of BCL9 RNA transcript in all carcinoma cell lines tested, but they do not allow to distinguish whether there are two (e.g wild-type BCL9 and both N-and C-terminally deleted BCL9), three (e.g wild-type BCL9, N-terminally deleted BCL9 and C-terminally deleted BCL9) or four different BCL9 mRNA populations (e.g. wild-type BCL9, N-terminally deleted BCL9, C-terminally deleted BCL9, N-terminal and C-terminally deleted BCL9) in these cell lines.

C. Absence of C- and N-terminal deletions in BCL9 genomic DNA

To better understand the heterogeneity of the BCL9 mRNA population presence of these deletions was tested on the genomic level. Specific amplification of the regions that should contain the deletions (Figure 19, A) was performed on genomic DNA isolated from SW480 and HT29 cells, another colon carcinoma cell line. The expected PCR product for the C-terminal region is 638bp for a wild-type gene and 605bp for a C-terminally deleted gene. For the N-terminal region PCR, the wild-type product should be 520bp and the deleted one 298bp. As can be observed in Figure 19 B, the genomic BCL9 DNA in SW480 and HT29 cells is wild-type, and does not contain the N- and C-terminal deletions observed in a subpopulation of the BCL9 transcripts.

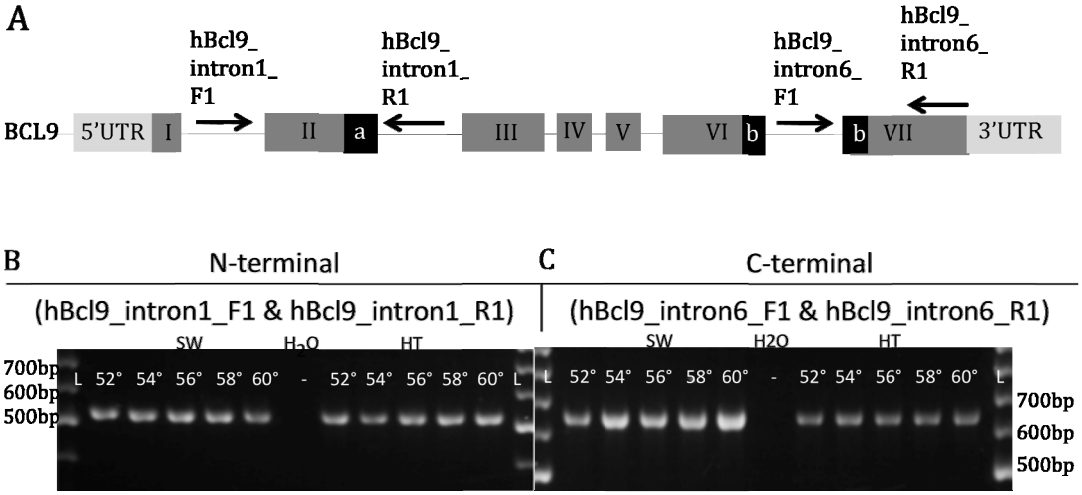


Figure 19. PCR primer position and PCR results of BCL9 genomic DNA. A. Primers used for the N-terminal region (a) and the C-terminal region (b) PCR (black arrows). B. N-terminal PCR results. C. C-terminal PCR results.

One explanation for the generation of three different mRNAs from a single gene is alternative splicing.

At this point, two colon carcinoma cells lines and four different colon and breast carcinoma cell lines show a mixed population of BCL9 transcript, including the wild-type one. Still, the presence of this wild-type transcript does not necessarily mean the protein is functional.

D. BCL9/BCL9L proteins

To further investigate BCL9 function, analysis of the BCL9 protein was performed in different carcinoma cell lines. BCL9L expression was also tested to compare with BCL9 expression profiles. Table 12 shows the relative mRNA levels of BCL9 and BCL9L of the cell lines analyzed on the protein level.

Cell line	% BCL9 mRNA expression level compared to SW480	% BCL9L mRNA expression level compared to SW480	Features
COLO320*	925%	5%	
HCT116*	330%	26%	
HT29*	226%	48%	
LS174T	No data	No data	
SW480	100%	100%	
BCL9L KD SW480*	426.4%	13.1%	BCL9L KD cells
BCL9 cDNA x SW480*	SW480 transfected with a plasmid (pLV) containing a BCL9 cDNA resistant to 3'UTR BCL9 shRNAmir		
BCL9L cDNA x SW480*	SW480 transfected with a plasmid (pLV) containing a BCL9L cDNA resistant to 3'UTR BCL9L shRNAmir		
pLV x SW480*	SW480 transfected with only the empty vector (pLV)		
HS578T*	148%	97%	
MCF7*	258%	94%	
MDA-MB231*	129%	78%	
*Results communicated by Frederique Baruthio and Jennyfer Bultinck.			

Table 12. Cell lines analyzed and their BCL9/BCL9L mRNA expression level. KD = knock-down.

120ng of protein extract from HS578T cells and 150ng of protein extract from the other tested cell lines are run on a 6% tris-glycine SDS-polyacrylamide gel, transferred onto a nitrocellulose membrane and immunodetected with BCL9- and BCL9L- specific primary antibodies. Manufacturer's predicted size for BCL9 and BCL9L was 200kDa and 200-220kDa, respectively. However, molecular weights calculated with the ExPASy bioinformatics tool was 149kDa for wild-type BCL9, 148kDa for putative C-terminally deleted BCL9, 142kDa for putative N-terminally deleted BCL9, 141kDa for N-and C-terminally deleted BCL9 and 157kDa for BCL9L. The antigen used for BCL9 antibody production is located outside any of the putative deleted regions of BCL9. As COLO320, LS174T, SW480 and MDA-MB231 cells express the 3 different populations of BCL9 mRNA (chapter IV.2.), three BCL9 protein bands can be expected. As control we used BCL9L knock-down SW480 cells, which show, in addition to lower BCL9L level, an extensive up-regulation of the BCL9 mRNA compared to parental SW480 cells

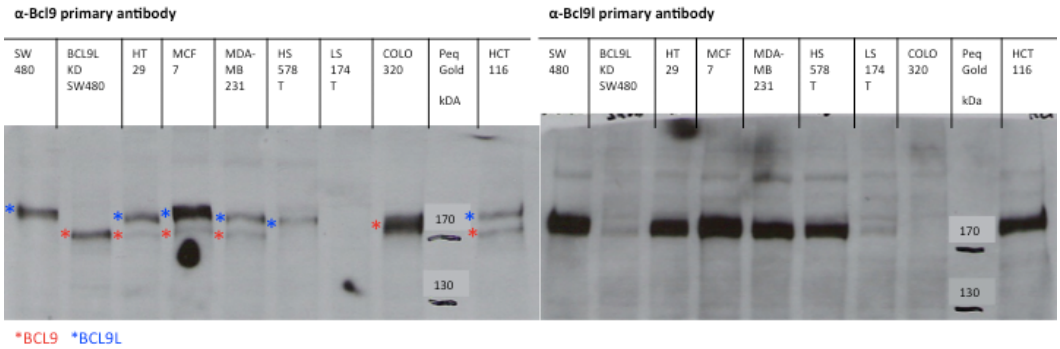
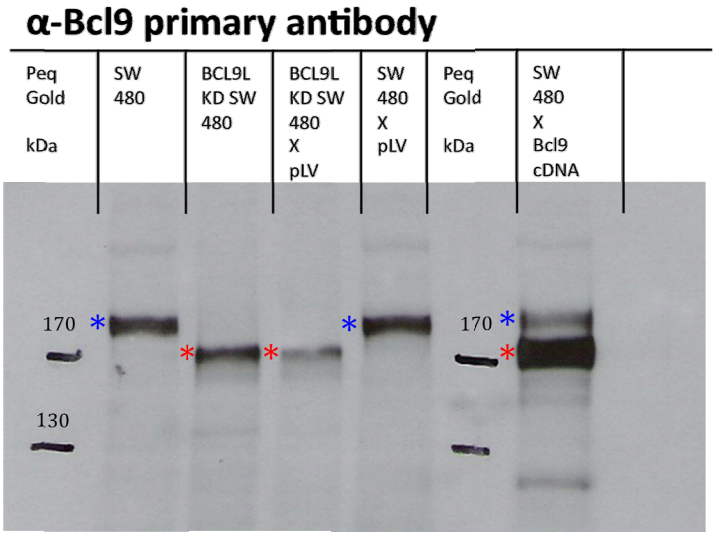


Figure 20. BCL9- and BCL9L-specific Western blot. The different cell lines analyzed are COLO320, HCT116, HT29, LS174T, SW480, BCL9L KD SW480, HS578T, MCF7 and MDA-MB231. KD = knock-down.

The specificity of the BCL9L primary antibody was confirmed as BCL9L-specific Western blot on the protein extract of BCL9L knock-down SW480 cells clearly shows the disappearance of the BCL9L band present in the parental SW480 cells protein extract (Figure 20, right panel). However, the band size (~180kDa) was lower than expected.

The BCL9 primary antibody seemed to recognize both BCL9 and BCL9L proteins (Figure 20, left pannel). Indeed, the up-regulation of BCL9 mRNA in the BCL9L knock-down SW480 cells compared to parental SW480 cells is reflected on protein level for the band slightly over 170kDa [*]. Furthermore, another band [*] appears at the same size as with the BCL9L-specific antibody in case of the parental SW480, while this band disappears for the BCL9L knock-down SW480 cells. Therefore, this band is likely to be BCL9L. These results are also confirmed with protein extract from SW480 cells expressing BCL9 cDNA as shown in Figure 21. Parental SW480 cells and SW480 cells expressing the control cDNA both show the BCL9L band [*]. This band is also present in SW480 cells expressing exogenous BCL9 cDNA, but in these cells an additional band appears with the same size as in BCL9L knock-down SW480 cells [*]



*BCL9 *BCL9L

Figure 21. Western blot with α-Bcl9 primary antibody. The different cell lines analyzed are SW480, BCL9L KD SW480, SW480xpLVxBCL9L shRNAmir, SW480xpLV and SW480xBCL9L cDNA. KD = knock-down.

As can be observed in Figure 20, HCT116, HT29, SW480, HS578T, MCF7 and MDA-MB231 cells all show an intense BCL9L-specific band, while a weaker BCL9L band is present in LS174T, just slightly more intense than in the BCL9L knock-down SW480 cells, and no BCL9L band can be observed in COLO320 cells. Noteworthy is that the level of BCL9L protein in COLO320 cells is even lower than the residual protein in BCL9L knock-down SW480 cells. SW480 and MCF7 cells seem to have the same high level of BCL9L protein, which correspond to BCL9 mRNA expression. HCT116, HT29 and MDA-MB231 cells express BCL9L at a slightly lower level. No conclusion can be drawn for HS578T cells because of the lower protein input. The relative BCL9L protein expression levels correlate well with the mRNA level of both the colon and breast cell lines. In the breast cell lines, in MCF7 cells BCL9L protein expression level is higher and in MDA-MB231 cells lower than what could be expected from their BCL9L mRNA expression level (Table 12 and Figure 20). The size of the BCL9L protein seems identical in all cell lines.

LS174T, SW480 and HS578T cells have neglectable levels of BCL9, and HCT116, HT29, MCF7 and MDA-MB231 cells a relatively low BCL9 protein expression. COLO320 cells have an intense BCL9-specific protein band, even more intense than in BCL9L knock-down SW480 cells. As for the

BCL9L, these expression levels correspond quite well with the BCL9 mRNA expression levels observed in Table 12, except for MDA-MB231 cells that show a more marked BCL9 protein band than expected from qRT-PCR results. All tested cell lines show the same BCL9 protein size except the COLO320 cells. Its band locates between the BCL9 band and the BCL9L band of all other cell lines.

BCL9L protein is higher expressed than BCL9 protein in SW480, HS578T and MCF7 cells; they are expressed to a similar extend in HCT116, LS174T and MDA-MB231 cells; and in COLO320 and HT29 cells BCL9 is the most abundant protein of the two.

All the cell lines tested seem to have only one band for BCL9, while we expected three bands coming from the three different transcripts. Either the resolution of the Western blot is too low to distinguish the three different bands or only one of the BCL9 mRNAs is translated into protein.

To further address this question, we increased the migration time of 150ng of protein extract from the same cell lines in a 6% tris-glycine SDS-polyacrylamide gel until all protein markers lower than 170kDa were out of the gel. This should also results in a better separation of BCL9 and BCL9L. This is important as the only working BCL9-specific antibody available recognizes both proteins, which have similar sizes. Results are shown in Figure 22, Figure 23 and Figure 24.

BCL9 [*] and BCL9L [*] cDNA-expressing SW480 cells and control SW480 cells were tested and revealed with anti-BCL9, anti-BCL9L and anti-V5 primary antibodies (Figure 22).

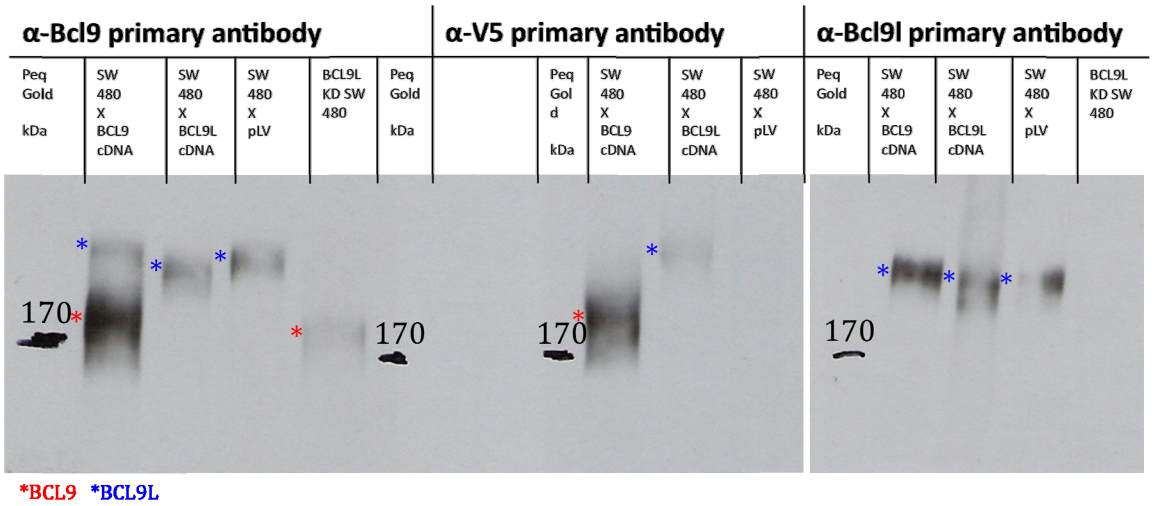


Figure 22. Western blot with α-Bcl9, α-V5 or α-Bcl9l primary antibody. The cell lines analyzed are SW480xBCL9 cDNA, SW480xBCL9L cDNA, control SW480, BCL9L KD SW480 and SW480. KD = knock-down.

The BCL9 and BCL9L cDNA contain a V5 tag, which means that exogenous BCL9 and BCL9L proteins produced from these cDNAs can be detected using an anti-V5 primary antibody. As already mentioned before, protein extract from SW480 cells expressing exogenous BCL9 show a clear BCL9 band at the same size than in BCL9L knock-down SW480 cells. This band also appears in the V5 Western blot confirming that the upcoming band originates from the expression of the BCL9 cDNA.

The results for the protein levels of BCL9 and BCL9L were reproduced for COLO320, HCT116, HT29, SW480, BCL9L knock-down SW480 and HS578T cells, (Figure 23 A/B). However, MCF7 and MDA-MB231 cells show both very different BCL9/BCL9L levels among the different experiments

(Figure 20, Figure 23 A/B). BCL9L band of MCF7 cells is less intense than BCL9L band of SW480 cells in Figure 23 A compared to Figure 20 and Figure 23 B. Relative BCL9 protein level correlates well with the qRT-PCR results in all experiments and relative BCL9L protein level in none of them. For MDA-MB231 cells, BCL9L band is slightly the same in all the experiment, but BCL9 band is too intense in Figure 23 B compared to the other (Figure 20 and Figure 23 A) and does not correlate with the qRT-PCR result.

BCL9 migrates at a specific size in COLO320 cells compared to the other cells. The likelihood that it is due to different post-translational modification is high. A previous paper (Mani et al., 2008) described BCL9 knock-down in this cell line and showed its effect on Wnt target genes (e.g. cMyc, CyclinD1, CD44, VEGF), proliferative, adhesive, invasive and tumor metastasis potential. So, the question whether these post-translational modifications are necessary for BCL9 function arises and the doubt about BCL9 functionality in the other cell lines persists. Still, as BCL9 of COLO320 cells is different from all the others, it could be an abnormally modified BCL9 protein.
the BCL9 protein single band is confirmed, indicating that with this better separation, we could only observe one BCL9 band. Therefore, we conclude that only one BCL9 protein is translated out of the three different BCL9 mRNAs.

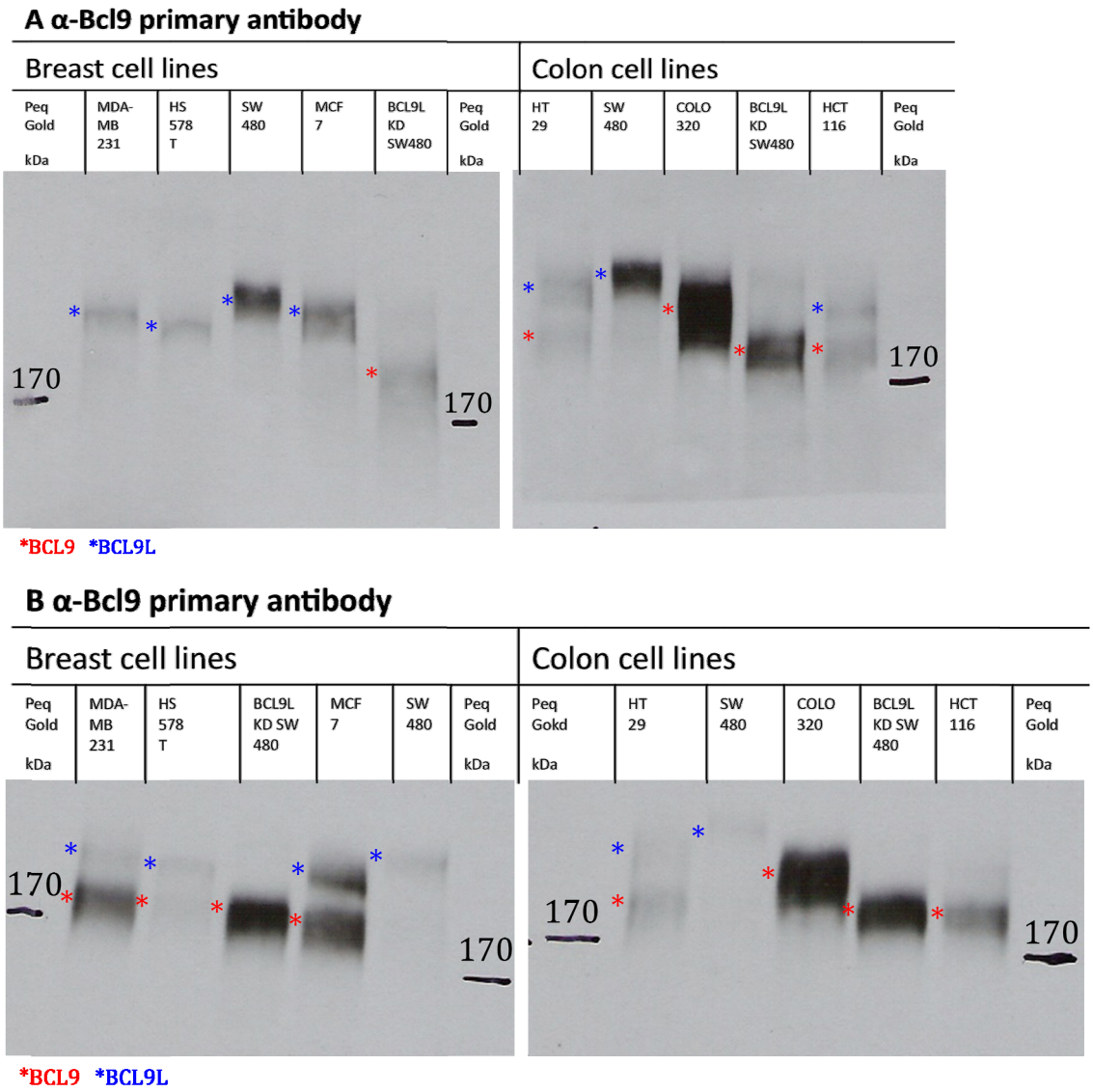


Figure 23. Western blot with α-Bcl9 primary antibody. The cell lines analyzed are MDA-MB231, HS578T, BCL9L KD SW480, MCF7, SW480, HT29, COLO320 and HCT116. KD = knock-down.

The longer migration time now revealed differences in BCL9L protein size between the cell lines, which was not evident on the previous blot. BCL9L of SW480 cells has the biggest size, BCL9L from HT29, MCF7 and MDA-MB231 cells is slightly smaller, and HS578T and HCT116 cells have a size in the middle of BCL9 blurred band of COLO320 cells (Figure 23). BCL9L Western blot confirms this result for SW480, HS578T, MCF7 and MDA-MB231 cells (Figure 24). The size variations should not be more than 20kDa as they were difficult to observe on the previous gel.

SW480 cells that express exogenous BCL9L have a slightly smaller BCL9L protein compared to control SW480 cells (Figure 22). The size differences for BCL9L in the different cell lines could indicate post-translational modifications as glycosylation, phosphorylation or ubiquitination. This could mean that the exogenous BCL9L protein is different post-translationnally modified and its functionality can no longer be guaranteed.

Based on all these previous results it becomes clear that it will be necessary to deplete both BCL9 and BCL9L in the human cell lines to investigate their biological functions.

Figure 24. Western blot with α-Bcl9l antibody. The cell lines analyzed are MDA-MB231, HS578T, MCF7 and SW480.

E. BCL9 knock-down in SW480, COLO320 and HEK293T cells*

To assess the function of BCL9 in different tumor cell lines, the efficiency of several BCL9 shRNAmirs was tested in SW480, COLO320 and HEK293T cells. COLO320 cells were chosen to test the efficacy of the shRNAmirs, as BCL9 depletion in this cell line was already described before (Mani et al., 2009). Gene signature analysis upon BCL9 knock-down in SW480 cells will learn us whether BCL9 is functional in these cells. Moreover, it will also allow to compare the effect of BCL9 and BCL9L knock-down on Wnt, EMT and stem-cell markers.

The BCL9 shRNAmir oligonucleotide sequences were designed either by Open Biosystem or the Joerg Huelsken lab programme. sh024, sh5090 and sh5588 should target the 3'UTR region of BCL9 mRNA, sh165 the coding sequence, sh161 the N-terminal deletion region and MA-362 (non-targeting shRNAmir) does not target any human genes. The lentivector, MA-344, contains an XhoI restriction site after the h_mirRNA30-5' region and a EcoRI restriction site before the h_mirRNA30-3' region. ShRNAmir oligonucleotides were designed to add XhoI and EcoRI restriction sites at the N- and the C-terminus, respectively. This allows to insert the shRNAmir in right orientation in the plasmid. Indeed, the shRNAmir needs to be transcribed in a specific direction, as its sequence is asymmetric. The DNA polymerase chosen was the Phusion™ as one nucleotide difference in the shRNAmir will already affect its mRNA degradation capacity. PCR products were separately run in a 2% agarose gel and bands at 138-139bp were gel extracted. This allows verifying that the PCR worked well and purifying the product from the other components of the PCR reaction. Then the plasmid vector MA-344 and the amplified shRNAmirs were digested with XhoI and EcoRI. Restriction digest products were purified and covalently linked to MA-344 by T4 ligation. Resulting plasmids were transformed in electro-competent DH5α *E. Coli* by electroporation. Clones were tested by XhoI and EcoRI digest and run it in a 2% agarose gel. Clones showing 115bp and 7577bp fragments were considered as containing BCL9 shRNAmir inserts and were sequenced. The plasmids containing one of the six BCL9-specific shRNAmir or the non-targeting shRNAmir were transfected together with the envelope and the packaging plasmids in HEK293T cells, who will produce viruses. Finally, the viruses were titrated on HEK293T, COLO320 and SW480 cells. Resulting titers are listed in Table 13.

Cell lines	shRNAmir	Titer (particles/ml)
HEK293T	non-targeting	$3x10^7$
	5090	undetermined
	5588	$3x10^7$
	024	$3x10^7$
	161*	$3x10^7$
	165*	$6,6x10^6$
	5254*	$5\ x10^6$
COLO320	non-targeting	10^7
	5090	undetermined
	5588	10^7
	024	10^7
	161*	$6,6x10^6$
	165*	$6,6x10^6$
	5254*	$2,5x10^6$
SW480	non-targeting	10^7
	5090	undetermined
	5588	10^7
	024	10^7
	161*	$6,6x10^6$
	165*	$6,6x10^6$
	5254*	$2,5x10^6$
*Results communicated by Norbert Wiedemann.		

Table 13. Titration for the different BCL9 shRNAmirs viruses in HEK293T, COLO320 and SW480.

The titer of the 5090 virus prep was too low and therefore, excluded of the tested shRNAmirs. HEK293T cells were infected four times with viruses non-targeting, 5588 or 024 shRNAmirs, SW480 and COLO320 cells with viruses containing the non-targeting, 5588, 024, 161 or 165 shRNAmirs, with a total MOI of 20. For the viruses containing the 5254 shRNAmir the MOI applied was 44. Unfortunately SW480 cells with 5588 shRNAmir and COLO320 with 5588 shRNAmir were contaminated by bacteria and were discarded from the analysis. Total RNA was extracted and reverse transcribed with the TaKaRa kit. Relative RNA level of HMBS, APCDD1, AXIN2, BCL9 and BCL9L were determined by qRT-PCR. HMBS is a housekeeping gene and is necessary for the first normalization step of the qRT-PCR data. APCDD1 and AXIN2 are two genes that were down-regulated in the BCL9L knock-down SW480 cells, and are direct targets of the Wnt pathway. BCL9L was also tested to ensure that an effect on the two preceding genes is solely due to BCL9 knock-down. The relative expression levels of the genes were determined versus parental cells and not versus non-targeting cells.

The percentage of knock induced for the different shRNAmirs is listed in Table 14. BCL9 knock-down levels are limited and range from 40% to 69%. These levels should be improved by either increasing the MOI or picking clones with good knock-down level and re-infecting them. BCL9 knock-down does not induce a prominent AXIN2 and APCDD1 down-regulation even if BCL9L expression was lower. In SW480 cells, BCL9 knock-down does not have an effect on BCL9L expression level, but still, induces a slight AXIN2 and APCDD1 down-regulation. In COLO320 cells, BCL9 knock-down induces more varying changes on BCL9L expression level, but this could be attributed to its low expression and therefore less accurate qRT-PCR results. AXIN2 and APCDD1 are less affected by the BCL9 knock-down than in SW480 cells. ShRNAmir-161, targeting only the wild-type BCL9 mRNA show an up-regulation of BCL9L, AXIN2 and APCDD1 expression level in COLO320 cells. BCL9 knock-down levels should be improved to at least a level of 75%.

Only afterwards, analysis of its effect would be extended to a larger subset of Wnt target genes and compared with BCL9L knock-down phenotypes.

	Gene:	BCL9	BCL9L	AXIN2	APCDD1
Cell line:	shRNAmir:				
HEK293T	5588	46.0%	45.4%	3.5%	9.3%
	24	56.4%	23.0%	9.6%	9.6%
	MA362	-17.9%	9.4%	-20.4%	-13.4%
Colo320	5254	69.0%	39.2%	20.0%	14.2%
	24	66.1%	29.3%	11.9%	10.0%
	161	53.2%	-52.8%	-54.3%	-102.3%
	165	40.0%	28.8%	15.2%	18.0%
	MA362	13.1%	36.7%	2.0%	2.6%
SW480	5254	41.1%	10.3%	17.3%	35.4%
	24	51.5%	11.5%	22.2%	34.1%
	161	48.3%	3.3%	12.5%	40.0%
	165	59.7%	3.0%	25.8%	46.4%
	MA362	4.2%	12.4%	3.2%	-19.6%
	BCL9L*	86.9%	-325.7%	97.8%	99.6%

*Results communicated by Norbert Wiedemann and calculated relative to SW480 expressing non-targeting shRNAmir and not to parental SW480.

Table 14. Percentage of knock-down for Bcl9 shRNAmir in HEK293T, COLO320 and SW480 cells.

*This experiment was performed together with Norbert Wiedemann.

F. BCL9/BCL9L proteins in BCL9 knock-down cell lines

To further investigate the impact of BCL9 knock-down on BCL9 and BCL9L protein levels, analysis of the BCL9/BCL9L protein was performed in SW480 and COLO320 cells expressing BCL9. Proteins were extracted from the COLO320 cells expressing BCL9 shRNAmir 024, 5254 and 161, the SW480 cells expressing shRNAmir 024, 5254 and 161, the parental COLO320 cells, the parental SW480 cells and the BCL9L knock-down SW480 cells**.**

150ng of protein extract were run in a 6% tris-glycine SDS-polyacrylamide gel, transferred onto a nitrocellulose membrane and immunodetected with BCL9 and BCL9L specific primary antibodies. The gel was run until all protein markers lower than 170kDa were out of the gel. SW480 cells and BCL9 knock-down SW480 cells were also used as reference for BCL9L and BCL9 size respectively.

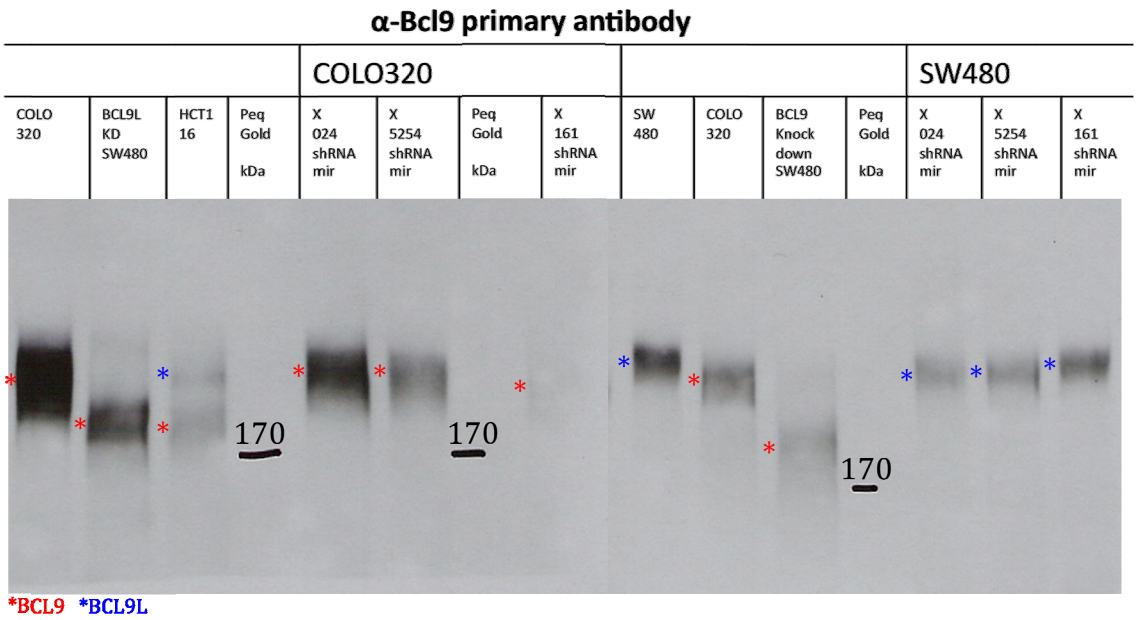


Figure 25. Western blot with α-Bcl9 primary antibody. Cell lines analyzed are HCT116, COLO320 expressing BCL9 shRNAmir 024, 5254 and 161, the SW480 expressing shRNAmir 024, 5254 and 161, the parental COLO320, the parental SW480 and the BCL9L knock-down SW480. KD = knock-down.

For SW480 cells, it is difficult to observe a down regulation of BCL9, as its expression is already borderline detectable. SW480 cells expressing BCL9 shRNAmir seem to express less BCL9L protein compared to parental SW480 cells (Figure 25), which is not in correspondance with mRNA (Table 14).

In COLO320 cells, on protein level, shRNAmir 161 is the most potent shRNAmir to knock-down BCL9 with almost no residual BCL9 protein visible, followed by shRNAmir 5254 and finally shRNAmir 024 (Figure 25). This does not correlate with knock-down of BCL9 mRNA level (Table 14).

As described previously in chapter IV.E, the 161 shRNAmir targets the N-terminally deleted region of BCL9 mRNA and so should only induce the degradation of the wild-type BCL9 mRNA. The absence of BCL9 protein in the 161 shRNAmir expressing COLO320 cells could indicate that only the wild-type BCL9 mRNA is translated into protein and thus, that the BCL9 mRNA percentage come from the non-degraded N-terminally deleted BCL9 mRNA population, which is not translated in protein. So, the qRT-PCR could overestimate the quantity of further translated mRNA. Moreover, this observation plus the fact that only one BCL9 protein band was observed would confirm that the BCL9 protein observed in the cell lines is a product of solely the wild-type BCL9 mRNA and not a deleted protein. In none of the BCL9 shRNAmir expressing COLO320 cells, BCL9L band appears and this correlates with the mRNA level observed in the qRT-PCR experiment.

V. Summary and Discussion

We showed in a mouse model that *Bcl9/Bcl9l*–mutant colon tumors lack EMT and stem cell–like features. Both features favor malignancy and resistance to therapy, and therefore, targeting Bcl9/Bcl9l function might have great therapeutic potential for colon cancer treatment (Deka *et al.*, 2010). However, these results should first be validated in human cancer. Therefore, Wnt active cell lines that contain EMT and stem-cell like traits were depleted of BCL9 and BCL9L using shRNAmir-expressing lentiviruses, followed by the assessment of phenotypical and gene expression changes. Preliminary results show that up-regulation of BCL9 in BCL9L knock-down SW480 cell does not rescue the induced phenotype. Others also reported effects upon single knock-down of BCL9 or BCL9L in colon cancer cell lines. Together, this questions the presumed redundancy between these paralogs.

Within our lab's effort to address the functional redundancy between BCL9 and BCL9L, the current study focused in a first part on assessing BCL9 (and BCL9L) expression pattern in different colon and breast carcinoma cell lines. On transcript level, three different populations of BCL9 mRNA could be found in all cell lines tested. More strikingly, the three transcripts were expressed from wild-type genomic BCL9 DNA at least in the two cell lines that were tested. Despite these different mRNA populations, all cell lines seem to produce BCL9 protein with one specific size, which varies between the lines, and which is very particular for COLO320 cells. The same is valid for BCL9L protein, although the variations are smaller than for BCL9. These size differences hint at the presence of possible post-translational modifications. In contrast to breast cancer cell lines, the relative expression levels between the colon cancer cell lines are consistent between mRNA and protein results for both BCL9 and BCL9L.
The second part of the study aimed at assessing BCL9 *versus* BCL9L function. Therefore, different BCL9 shRNAmir were designed and tested. Once a good knock-down is obtained, its effect on gene expression will be investigated and compared to the effect of BCL9L knock-down.

Three BCL9 mRNA populations - two carrying short deletions (4,059kb and 4,023kb) and the wild-type (4,281kb)- have been identified in SW480, LS174T, COLO320, MCF7 and MDA-MB231 cell lines using Southern blot, RT-PCR and sequencing methods. Willis *et al.* screened different human tissues (spleen, thymus, prostate, testis, ovary, small intestine, colon and blood) by Northern blot with a probe derived from the 1q21 break point region of BCL9, and found three different BCL9 mRNA transcripts (6,3kb, 4,2kb and 1,6kb) (Willis *et al.*, 1998). The two biggest transcripts were present in all tissues tested with the 6,3kb transcript being higher expressed than the one of 4,2kb. The small transcript was only present in the spleen, the thymus and the small intestine. So, only the 6,3kb and not the 4,2kb transcript present in human colon samples was found by us in the colon carcinoma cell lines.

Despite the presence of shorter mRNAs for BCL9, we could not find these deletions at the genomic level in SW480 and HT29 cells. Furthermore, wild-type BCL9 was sequenced on mRNA level. These deleted mRNAs could be the result of alternative splicing, which is a complex process determined not only by intron sequences present at the exon boundaries, but also by other auxiliary sequences present in adjacent intron and exon. These sequences are recognized by splicing factors, which help or prevent the function of the spliceosome (Tejedor and Valcarcel, 2010). A standard predicting bioinformatics tool (Reese and Harris, 1997) confirms the likelihood of such alternative splicing sites at the N-terminal deleted region. However, no putative acceptor site could be found for the C-terminal deleted region.

The repartition of the different BCL9 mRNA populations seems to differ between the cell lines. This could be explained by the tissue specificity of the alternative splicing process (Barash *et al.*, 2010). As our method is not quantitative, these differences should, however, be validated by qRT-PCR with primer pairs specific for the different transcripts.

As endogenous BCL9 expression is not able to rescue the BCL9L knock-down phenotype in SW480 cells, the presence of a functional BCL9 in SW480 cells was questioned. We found that SW480 cells express at least a wild-type BCL9 mRNA and thus, that the putative non-functionality of BCL9 cannot be attributed to the absence of a wild-type BCL9 mRNA and protein.

To get a clearer view on BCL9 and BCL9L function in different carcinoma cell lines, their expression was further analyzed on protein level by standard Western blotting. BCL9L size, close to 200kDa, and the relative expression levels of BCL9 and BCL9L in SW480 cells correlate with previous studies (Brembeck *et al.*, 2004; Adachi *et al.*, 2004). Same anti-BCL9L antibody was used by Mani *et al*. to detect BCL9L in normal colon mucosa, HCT115, HCT116, HT29, SW620, COLO201, COLO320, SW403 and SW678 cell lines (Mani *et al.,* 2009, Supplementary Figure 8). BCL9L protein expression correlates with our results in HCT116, HT29 and COLO320 cells. We also observed small size differences of the BCL9 and BCL9L protein between the carcinoma cell lines, which were not reported before.
The size difference for both BCL9 and BCL9L between the different cell lines could be explained by post-translational modifications as phosphorylation, which slows down protein migration in a SDS-polyacrylamide gel, and glycosylation, which increases the size of a protein. To address this, new methods should be introduced, such as mass spectrometry, which determines the amino acid composition and identifies post-translational modifications, or 2D western blot migration, which reveals phosphorylation differences and allows to use antibodies recognizing glycosylated sites. Post-translational modifications might change BCL9 and BCL9L function. BCL9 size in COLO320 cells differs from all the other cell lines tested. Moreover, BCL9 seem to be functional in COLO320 cells as Mani's *et al.* (Mani *et al.*, 2009) observed an effect on Wnt target gene expression and phenotypical changes upon transient BCL9 RNA silencing. The question then arises whether BCL9 in the other cell lines, which has a different size, is functional. We assessed the functionality of both BCL9 and BCL9L proteins in the different carcinoma cell lines by means of shRNAmir-mediated gene knock-down.

The knock-down efficiency of different BCL9 shRNAmir sequences was tested in COLO320 and SW480 cell lines, and their effect on the expression of BCL9L, AXIN2 and APCDD1 evaluated by qRT-PCR. The depletion of BCL9 mRNA is only partial (40% to 69%), and induces changes in BCL9L ranging from 3% to 11,5% in SW480 cell line and from 28,8 to 39,2% in COLO320 cell line. This effect on BCL9L mRNA was, however, also observed with the non-targeting shRNAmir, so is most probably an off-target effect. BCL9 knock-down induces a slight down-regulation of AXIN2 (12,5-25,8%) and APCDD1 (34,1-46,4%) in SW480 and even less in COLO320 (11,9-20% and 10-18%, respectively). The only exception is COLO320 cells that expresses shRNAmir161 where BCL9L, AXIN2 and APCDD1 are up-regulated.
De la Roche *et al.* tested BCL9 and BCL9L single gene, transient silencing in SW480 cell line. Partial BCL9 knock-down (~50%) induced down-regulation of both BCL9L (~50%) and AXIN2 (~35%) (de la Roche *et al.*, 2008). We, however, did not observe an effect on BCL9L and AXIN2 with this BCL9 knock-down level. Moreover, as BCL9L knock-down (~70%) induced down-regulation of AXIN2 (~50%) with neglectable down-regulation of BCL9 (10%) (de la Roche *et al.*, 2008), it is possible that the effect of BCL9 knock-down on AXIN2 observed by de la Roche is due to the down-regulation of BCL9L. However, it is difficult to draw a straightforward conclusion with this limited parental knock-down. Therefore, more efforts should be undertaken to

increase the knock-down level of BCL9 by testing different shRNAmir sequences, MOI and infection methods. Once a decent knock-down is obtained (at least 75%), additional direct Wnt targets should be quantified such as VEGF and CD44, which were analyzed in the Mani paper, vimentin, Lgr5 and other EMT and stem-cell marker. Finally, these gene signatures should be compared between BCL9- and BCL9L-depleted cells to clarify whether these two paralogs are functionally redundant.

The effect of BCL9-specific shRNAmir024, 5254 and 161 on BCL9 and BCL9L protein levels in COLO320 cells and SW480 cells was further analyzed by Western blot. SW480 parental cells seem to express BCL9 protein at a very low level, and therefore, it was difficult to observe further down-regulation. Still, it should be noted that BCL9L protein levels in SW480 cells are lower than expected based on the qRT-PCR results. In COLO320, the 53% knock-down on mRNA level induced by shRNAmir161 seemed to be completely underestimated, as no BCL9-specific protein band could be observed in these cells. As shRNAmir161 targets only the wild-type BCL9 mRNA, these results could imply that BCL9 protein is translated only from the wild-type transcript and that the residual BCL9 mRNA observed by qRT-PCR is coming from the deleted forms. The shRNAmir161-induced up-regulation of BCL9L on mRNA level (52,8%) seems to be too low to be observed in the Western blot.
In COLO320 cells, shRNAmir024 and shRNAmir5254 lower BCL9 protein level but not to the same extent as shRNAmir161, despite a higher BCL9 knock-down on mRNA level. This could be explained by the fact that the degradation of the wild-type BCL9 mRNA will be more efficient in case of shRNAmir161 than of shRNAmir024 or shRNAmir5254, as the latter target both the deleted and the wild-type BCL9 mRNA. When BCL9 is only translated from the wild-type mRNA, BCL9 protein level should be lower in COLO320 cells expressing shRNAmir161 compared to shRNAmir024 or shRNAmir5254, which is exactly what we observed.
What remains puzzling is that complete depletion of BCL9 protein (shRNAmir161) induces up-regulation of BCL9L, AXIN2 and APCDD1, while a partial depletion (shRNAmir024 or shRNAmir5254) results in down-regulation of those genes. Moreover, despite similar BCL9 mRNA knock-down (50%), shRNAmir161 induces up-regulation of BCL9L, AXIN2 and APCDD1 in COLO320 cells but not in SW480 cells, where AXIN2 and APCDD1 are actually rather down-regulated. It is clear that these results need to be repeated and confirmed more accurately, e.g. by including loading controls in the Western blots.

In conclusion, we have clearly shown that human colon and breast carcinoma cell lines carry wild-type BCL9 genomic DNA and mRNA but that there are differences at the protein level (both in size and expression level) between the different cell lines.
These results cannot yet answer the question whether BCL9 and BCL9L are functional in the different cell lines, neither can it assess the redundancy between BCL9 and BCL9L. Therefore, we need to optimize the BCL9 knock-down level, assess its specificity and extend the analysis on EMT and stem cell-like related genes. Only afterwards, the redundancy question can be addressed by comparing phenotypes and gene expression changes between BCL9 and BCL9L knock-down cell lines and by reconstituting the cells with BCL9L and BCL9 cDNA, respectively. These single and double knock-down cell lines will be extremely useful in characterizing BCL9 and BCL9L function in human cancer and, as tested cell lines are Wnt active and contain EMT and stem cell-like traits, we should be able to see whether BCL9 and/or BCL9L inhibition in these human cells has the same effect as we observed in the mouse. Double knock-down cell lines should be more differentiated and less mesenchymal and therefore, less metastatic and malignant when injected into immuno-compromised mice. To see whether BCL9 and/or BCL9L depletion renders chemo-resistant cancer cells more susceptible to therapy, we should compare the IC50 value of currently used cancer drugs between wild-type, single and double BCL9/9L knock-down cell lines.

Small molecules that inhibit the interaction between BCL9/BCL9L and β-catenin are already available in our lab and should be tested on the same panel of cell lines. Comparison of their effect with shRNAmir-mediated knock-down of BCL9 and/or BCL9L should further clarify whether they are potent inhibitors of BCL9 and BCL9L function, and thus whether they can abrogate EMT and stem cell-like properties of tumor cells.

VI. References

Adachi, S., Jigami, T., Yasui, T., Nakano, T., Ohwada, S., Omori, Y., Sugano, S., Ohkawara, B., Shibuya, H., Nakamura, T. &Akiyama, T. (2004). Role of a BCL9-related beta-catenin-binding protein, B9L, in tumorigenesis induced by aberrant activation of Wnt signaling. *Cancer Res* 64(23): 8496-8501.

Ai, L., Tao, Q., Zhong, S., Fields, C. R., Kim, W. J., Lee, M. W., Cui, Y., Brown, K. D. &Robertson, K. D. (2006). Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. *Carcinogenesis* 27(7): 1341-1348.

Barash, Y., Calarco, J. A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B. J. &Frey, B. J. (2010). Deciphering the splicing code. *Nature* 465(7294): 53-59.

Barker, N., Hurlstone, A., Musisi, H., Miles, A., Bienz, M. &Clevers, H. (2001). The chromatin remodelling factor Brg-1 interacts with beta-catenin to promote target gene activation. *EMBO J* 20(17): 4935-4943.

Brembeck, F. H., Rosario, M. &Birchmeier, W. (2006). Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. *Curr Opin Genet Dev* 16(1): 51-59.

Brembeck, F. H., Schwarz-Romond, T., Bakkers, J., Wilhelm, S., Hammerschmidt, M. &Birchmeier, W. (2004). Essential role of BCL9-2 in the switch between beta-catenin's adhesive and transcriptional functions. *Genes Dev* 18(18): 2225-2230.

Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G. &Stevenson, M. (1993). Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A* 90(13): 6125-6129.

Clevers, H. (2006). Wnt/beta-catenin signaling in development and disease. *Cell* 127(3): 469-480.

Daniels, D. L. &Weis, W. I. (2002). ICAT inhibits beta-catenin binding to Tcf/Lef-family transcription factors and the general coactivator p300 using independent structural modules. *Mol Cell* 10(3): 573-584.

de la Roche, M., Worm, J. &Bienz, M. (2008). The function of BCL9 in Wnt/beta-catenin signaling and colorectal cancer cells. *BMC Cancer* 8: 199.

Deka, J., Wiedemann, N., Anderle, P., Murphy-Seiler, F., Bultinck, J., Eyckerman, S., Stehle, J. C., Andre, S., Vilain, N., Zilian, O., Robine, S., Delorenzi, M., Basler, K. &Aguet, M. (2010). Bcl9/Bcl9l Are Critical for Wnt-Mediated Regulation of Stem Cell Traits in Colon Epithelium and Adenocarcinomas. *Cancer Res*.

Delelis, O., Carayon, K., Saib, A., Deprez, E. &Mouscadet, J. F. (2008). Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology* 5: 114.

Dharmaraj, S. (2010).RT-PCR: The basics.

Fewell, G. D. &Schmitt, K. (2006). Vector-based RNAi approaches for stable, inducible and genome-wide screens. *Drug Discov Today* 11(21-22): 975-982.

Gallay, P., Swingler, S., Song, J., Bushman, F. &Trono, D. (1995). HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. *Cell* 83(4): 569-576.

Gayet, J., Zhou, X. P., Duval, A., Rolland, S., Hoang, J. M., Cottu, P. &Hamelin, R. (2001). Extensive characterization of genetic alterations in a series of human colorectal cancer cell lines. *Oncogene* 20(36): 5025-5032.

Graham, T. A., Weaver, C., Mao, F., Kimelman, D. &Xu, W. (2000). Crystal structure of a beta-catenin/Tcf complex. *Cell* 103(6): 885-896.

Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., Sohn, S. Y., Cho, Y., Zhang, B. T. &Kim, V. N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125(5): 887-901.

Hanahan, D. &Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* 100(1): 57-70.

Hecht, A., Vleminckx, K., Stemmler, M. P., van Roy, F. &Kemler, R. (2000). The p300/CBP acetyltransferases function as transcriptional coactivators of beta-catenin in vertebrates. *EMBO J* 19(8): 1839-1850.

Hoffmans, R. &Basler, K. (2007). BCL9-2 binds Arm/beta-catenin in a Tyr142-independent manner and requires Pygopus for its function in Wg/Wnt signaling. *Mech Dev* 124(1): 59-67.

Hoffmans, R., Stadeli, R. &Basler, K. (2005). Pygopus and legless provide essential transcriptional coactivator functions to armadillo/beta-catenin. *Curr Biol* 15(13): 1207-1211.

Hombrouck, A., De Rijck, J., Hendrix, J., Vandekerckhove, L., Voet, A., De Maeyer, M., Witvrouw, M., Engelborghs, Y., Christ, F., Gijsbers, R. &Debyser, Z. (2007). Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog* 3(3): e47.

Hope, T. J. &Trono, D. (2000).Structure, Expression and Regulation of the HIV Genome.

Howe, L. R. &Brown, A. M. (2004). Wnt signaling and breast cancer. *Cancer Biol Ther* 3(1): 36-41.

Hu, H. Y., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y. H., Chen, W. &Khaitovich, P. (2009). Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* 10: 413.

Huber, A. H. &Weis, W. I. (2001). The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* 105(3): 391-402.

Jegede, O., Babu, J., Di Santo, R., McColl, D. J., Weber, J. &Quinones-Mateu, M. (2008). HIV type 1 integrase inhibitors: from basic research to clinical implications. *AIDS Rev* 10(3): 172-189.

Kalluri, R. &Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *J Clin Invest* 119(6): 1420-1428.

Kamb, A., Wee, S. &Lengauer, C. (2007). Why is cancer drug discovery so difficult? *Nat Rev Drug Discov* 6(2): 115-120.

Koh, S. S., Li, H., Lee, Y. H., Widelitz, R. B., Chuong, C. M. &Stallcup, M. R. (2002). Synergistic coactivator function by coactivator-associated arginine methyltransferase (CARM) 1 and beta-catenin with two different classes of DNA-binding transcriptional activators. *J Biol Chem* 277(29): 26031-26035.

Kramps, T., Peter, O., Brunner, E., Nellen, D., Froesch, B., Chatterjee, S., Murone, M., Zullig, S. &Basler, K. (2002). Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. *Cell* 109(1): 47-60.

Li, J., Donath, S., Li, Y., Qin, D., Prabhakar, B. S. &Li, P. (2010). miR-30 regulates mitochondrial fission through targeting p53 and the dynamin-related protein-1 pathway. *PLoS Genet* 6(1): e1000795.

Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W. &Poeschla, E. M. (2006). An essential role for LEDGF/p75 in HIV integration. *Science* 314(5798): 461-464.

Mani, M., Carrasco, D. E., Zhang, Y., Takada, K., Gatt, M. E., Dutta-Simmons, J., Ikeda, H., Diaz-Griffero, F., Pena-Cruz, V., Bertagnolli, M., Myeroff, L. L., Markowitz, S. D., Anderson, K. C. &Carrasco, D. R. (2009). BCL9 promotes tumor progression by conferring enhanced proliferative, metastatic, and angiogenic properties to cancer cells. *Cancer Res* 69(19): 7577-7586.

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., Brooks, M., Reinhard, F., Zhang, C. C., Shipitsin, M., Campbell, L. L., Polyak, K., Brisken, C., Yang, J. &Weinberg, R. A. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133(4): 704-715.

Miller, M. D., Farnet, C. M. &Bushman, F. D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J Virol* 71(7): 5382-5390.

Milovanovic, T., Planutis, K., Nguyen, A., Marsh, J. L., Lin, F., Hope, C. &Holcombe, R. F. (2004). Expression of Wnt genes and frizzled 1 and 2 receptors in normal breast epithelium and infiltrating breast carcinoma. *Int J Oncol* 25(5): 1337-1342.

Mosimann, C., Hausmann, G. &Basler, K. (2009). Beta-catenin hits chromatin: regulation of Wnt target gene activation. *Nat Rev Mol Cell Biol* 10(4): 276-286.

Nermut, M. V. &Fassati, A. (2003). Structural analyses of purified human immunodeficiency virus type 1 intracellular reverse transcription complexes. *J Virol* 77(15): 8196-8206.

Nowotny, M. &Yang, W. (2009). Structural and functional modules in RNA interference. *Curr Opin Struct Biol* 19(3): 286-293.

Orsulic, S., Huber, O., Aberle, H., Arnold, S. &Kemler, R. (1999). E-cadherin binding prevents beta-catenin nuclear localization and beta-catenin/LEF-1-mediated transactivation. *J Cell Sci* 112 (Pt 8): 1237-1245.

Pardal, R., Clarke, M. F. &Morrison, S. J. (2003). Applying the principles of stem-cell biology to cancer. *Nat Rev Cancer* 3(12): 895-902.

Pokutta, S. &Weis, W. I. (2000). Structure of the dimerization and beta-catenin-binding region of alpha-catenin. *Mol Cell* 5(3): 533-543.

Reese, M. &Harris, N. (1997).Splice Site Prediction by Neural Network.

Reya, T. &Clevers, H. (2005). Wnt signalling in stem cells and cancer. *Nature* 434(7035): 843-850.

Sakamoto, I., Ohwada, S., Toya, H., Togo, N., Kashiwabara, K., Oyama, T., Nakajima, T., Ito, H., Adachi, S., Jigami, T. &Akiyama, T. (2007). Up-regulation of a BCL9-related beta-catenin-binding protein, B9L, in different stages of sporadic colorectal adenoma. *Cancer Sci* 98(1): 83-87.

Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. &Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115(2): 199-208.

Sierra, J., Yoshida, T., Joazeiro, C. A. &Jones, K. A. (2006). The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. *Genes Dev* 20(5): 586-600.

Skaletsky, S. R. a. H. J. (2000).Primer3 on the WWW for general users and for biologist programmers. . In *Bioinformatics Methods and Protocols: Methods in Molecular Biology.*, 365-386 (Ed T. Humana Press, NJ).

Stadeli, R., Hoffmans, R. &Basler, K. (2006). Transcription under the control of nuclear Arm/beta-catenin. *Curr Biol* 16(10): R378-385.

Sustmann, C., Flach, H., Ebert, H., Eastman, Q. &Grosschedl, R. (2008). Cell-type-specific function of BCL9 involves a transcriptional activation domain that synergizes with beta-catenin. *Mol Cell Biol* 28(10): 3526-3537.

Tejedor, J. R. &Valcarcel, J. (2010). Gene regulation: Breaking the second genetic code. *Nature* 465(7294): 45-46.

Thiery, J. P. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* 2(6): 442-454.

Townsley, F. M., Cliffe, A. &Bienz, M. (2004). Pygopus and Legless target Armadillo/beta-catenin to the nucleus to enable its transcriptional co-activator function. *Nat Cell Biol* 6(7): 626-633.

Toya, H., Oyama, T., Ohwada, S., Togo, N., Sakamoto, I., Horiguchi, J., Koibuchi, Y., Adachi, S., Jigami, T., Nakajima, T. &Akiyama, T. (2007). Immunohistochemical expression of the beta-catenin-interacting protein B9L is associated with histological high nuclear grade and immunohistochemical ErbB2/HER-2 expression in breast cancers. *Cancer Sci* 98(4): 484-490.

Trono, D. (2002). *Lentiviral vectors.*

van Erk, M. J., Krul, C. A., Caldenhoven, E., Stierum, R. H., Peters, W. H., Woutersen, R. A. &van Ommen, B. (2005). Expression profiling of colon cancer cell lines and colon biopsies: towards a screening system for potential cancer-preventive compounds. *Eur J Cancer Prev* 14(5): 439-457.

Wendt, M. K., Allington, T. M. &Schiemann, W. P. (2009). Mechanisms of the epithelial-mesenchymal transition by TGF-beta. *Future Oncol* 5(8): 1145-1168.

Willis, T. G., Zalcberg, I. R., Coignet, L. J., Wlodarska, I., Stul, M., Jadayel, D. M., Bastard, C., Treleaven, J. G., Catovsky, D., Silva, M. L. &Dyer, M. J. (1998). Molecular cloning of translocation t(1;14)(q21;q32) defines a novel gene (BCL9) at chromosome 1q21. *Blood* 91(6): 1873-1881.

Wiznerowicz, M. &Trono, D. (2003). Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference. *J Virol* 77(16): 8957-8961.

Wu, D. &Pan, W. (2010). GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem Sci* 35(3): 161-168.

Yang, C. K., Kim, J. H., Li, H. &Stallcup, M. R. (2006). Differential use of functional domains by coiled-coil coactivator in its synergistic coactivator function with beta-catenin or GRIP1. *J Biol Chem* 281(6): 3389-3397.

Yang, J. &Weinberg, R. A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* 14(6): 818-829.

Zeng, Y., Yi, R. &Cullen, B. R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J* 24(1): 138-148.

Zhao, S. &Liu, M. F. (2009). Mechanisms of microRNA-mediated gene regulation. *Sci China C Life Sci* 52(12): 1111-1116.

Zhou, B. B., Zhang, H., Damelin, M., Geles, K. G., Grindley, J. C. &Dirks, P. B. (2009). Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. *Nat Rev Drug Discov* 8(10): 806-823.

Lausanne, August 20th 2010 Fanny Bringolf