Ulm University Medical Center

Institute of Experimental Cancer Research

Director: Prof. Dr. med. Christian Buske

Role of the microRNA-196b in normal and malignant hematopoiesis

Dissertation submitted to obtain the doctoral degree of Human Biology (Dr. biol. hum.) of the Medical Faculty of Ulm University

Submitted by

Maria Götze (née Hollnagel)

Karl-Marx-Stadt (Chemnitz)

Germany

2022

Acting Dean: Prof. Dr. Thomas Wirth

1st Reviewer: Prof. Dr. med. Christian Buske

2nd Reviewer: Prof. Dr. Franz Oswald

Day of graduation: 25.11.2022

Results gained in my thesis have previously been published in the following publication:

Rawat VPS*, Götze M*, Rasalkar A, Vegi NM, Ihme S, Thoene S, Pastore A, Bararia D, Döhner H, Döhner K, Feuring-Buske M, Quintanilla-Fend L, Buske C: The microRNA miR-196b acts as tumor suppressor in Cdx2-driven acute myeloid leukemia. *Haematologica* 105 (6): 285-289 (2020).

*both authors contributed equally

This paper has been published under the Creative Commons Open Access license CC BY-NC 4.0 (https://creativecommons.org/licenses/by-nc/4.0/).

List of Contents

Abbreviations ..... **vi**

1. Introduction **1**

1.1 Normal Hematopoiesis 1

1.2 Malignant Hematopoiesis: Leukemia 5

1.3 MicroRNA biogenesis and function 11

1.4 MiRNAs and miR-196 in normal hematopoiesis 15

1.5 MiR-196 in solid cancer and leukemia 20

1.6 Aim of this study 29

2. Materials and Methods **30**

2.1 Materials 30

2.2 Methods 51

3. Results **73**

3.1 MiR-196 family expression in healthy and leukemic human hematopoietic cells 73

3.2 MiR-196b transcript determination in healthy human BM cells 74

3.3 MiR-196b promoter determination in healthy human BM cells 77

3.4 Effect of aberrant miR-196b expression in murine hematopoietic stem and progenitor cells on clonogenicity, cell growth, lineage differentiation and transformation ability 82

3.5 Effect of aberrant miR-196b expression on differential gene expression in murine hematopoietic stem and progenitor cells 89

3.6 Associations between mature miR-196b, its transcript variants and neighboring *HOX* gene targets in AML cells 90

3.7 Effect of aberrant miR-196b on clonogenicity, cell growth and leukemia associated transformation in murine HSPCs 95

3.8 Effect of aberrant miR-196b expression on differential gene expression in murine leukemic cells 102

3.9 Effect of aberrant miR-196b expression on clonogenicity, cell growth and candidate gene levels in human leukemic cells 108

4. Discussion **116**

4.1 One of the functions of miR-196b in healthy BM cells is to promote differentiation into myeloid cells *in vivo* 116

4.2 Transcriptional control of miR-196b in healthy human BM cells occurs via a specific promoter generating two primary miR-196b transcript variants 117

4.3 Association of miR-196b and *HOXA9* expression levels in normal and leukemic hematopoiesis 119

4.4 MiR-196b functions in a cell context dependent manner 127

4.5 Technical aspects for investigating miR-196b function 128

4.6 Identification of context dependent mRNA targets is essential for understanding miRNA function 130

4.7 Therapy 132

5. Summary 134

6. References 135

Acknowledgement 180

Curriculum vitae 181

Abbreviations

2-AP	2-Aminopurine
5-FU	5-Fluorouracil
7AAD	7 Aminoactinomycin
AF9	ALL1-Fused gene from chromosome 9
AGO	Argonaute
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
BM	Bone Marrow
bp	basepair
BRE	B Recognition Element
BSA	Bovine Serum Albumin
C. elegans	Caenorhabditis elegans
CBFß-MYH11	Core-Binding Factor subunit Beta – Myosin Heavy Chain 11
CCR4-NOT	Carbon Catabolite Repressor 4 (CCR4) - Negative on TATA (NOT) complex
CD	Cluster of Differentiation
cDNA	complementary single stranded Deoxyribonucleic Acid
Cdx	Caudal-type homeobox
CEBPA	CCAAT Enhancer Binding Protein Alpha
CEBPB	CCAAT Enhancer Binding Protein Beta
CFC	Colony Forming Cell
CFU-S	Colony Forming Unit-Spleen
CHD1	Chromodomain Helicase DNA Binding Protein 1
CI	Confidence Interval
CLL	Chronic Lymphocytic Leukemia
CLP	Common Lymphoid Progenitor
CML	Chronic Myeloid Leukemia
CMP	Common Myeloid Progenitor
CN	Cytogenetically normal
CpG-islands	Cytosine and Guanine dinucleotide enrichment in the genome
CSF	Colony Stimulating Factor
CTBP2	C-Terminal Binding Protein 2
DCE	Downstream Core Element
DEG	Differentially Expressed Genes
DEPC	Diethylpyrocarbonate treated water
DLX	Distal-less homeobox
DMEM	Dulbecco’s Modified Eagle’s Medium
DMSO	Dimethylsulfoxide
DN	Down
DNA	Deoxyribonucleic Acid
dNTP	deoxy Nucleosid Triphosphate
DPBS	Dulbecco’s Phosphate Buffered Saline
DPE	Downstream core promoter element

DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
dTTP	deoxy Thymine Triphosphate
E2F1	E2F Transcription Factor 1
EDTA	Ethylenediaminetetraacetic Acid
EGR1	Early Growth Response 1
ELN	European Leukemia Network
EST	Expressed Sequence Tag
ETO	Eight Twenty One
ETS	E Twenty Six
Evi-1	Ecotropic viral integration site-1
EZH2	Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit
FAB	French-American-British system
FACS	Fluorescence Activated Cell Sorting
FP	Forward Primer
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FLT3	FMS-like Tyrosine kinase 3
GC-boxes	Guanine and Cytosine rich regulatory element in transcription
GFP	Green Fluorescent Protein
GMP	Granulocyte-Macrophage Progenitor
GSEA	Gene Set Enrichment Analysis
GSX	GS homeobox
H3K27ac	acetylated Lysine 27 of Histone 3
H3K4me3	trimethylated Lysine 4 of Histone 3
HBS	Hepes Buffered Saline
HEPES	4-(2-HydroxyEthyl)-1-Piperazineethanesulfonic acid
HOX	clustered Homeobox gene
HP	Hairpin probe
Hprt	Hypoxanthine phosphoribosyltransferase 1
HSC	Hematopoietic Stem Cell
HSK	Housekeeping gene
i.p.	intraperitoneally
i.v.	intravenously
Inv (16)	Inversion within chromosom16
ITD	Internal Tandem Repeat
Kb	Kilo bases
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria Broth
LGMP I	Leukemic Granulo-Monocyte Progenitor
LIC	Leukemia Initiating Cell
Lin	Lineage
LN	Liquid Nitrogen
LLMPP	Leukemic Lymphoid-primed Multi-Potent Progenitor.
LMPP	Lymphoid-Primed Multipotent Progenitor
LSC	Leukemic Stem Cell

LT	Long-term
M^7G	7-Methylguanosine
MAPK	Mitogen-Activated Protein Kinase
MAZ	MYC Associated Zinc finger protein
MEP	Megakaryocyte-Erythrocyte Progenitor
$MgCl_2$	Magnesium Chloride
miRNA	micro Ribonucleic Acid
MLL	Mixed-Lineage Leukemia
MNC	Mononuclear Cell
MPP	Multipotent Progenitor
mRNA	messenger Ribonucleic Acid
MSX	Msh homeobox
MTE	Motif 10 Element
MXI1	MAX Interactor 1 dimerization protein
Myc	Myelocytomatosis proto-oncogene
n.a.	not available
NES	Normalized Enrichment Score
NK	Natural Killer
NPM1	Nucleophosmin 1
Nras	Neuroblastoma rat sarcoma viral oncogene homolog
NSG	Nonobese Diabetic (NOD)- Severe Combined Immunodeficiency (SCID)
nt	nucleotide
p	short arm of chromosome
P/S	Penicillin/Streptomycin
PAX	Paired box gene
PB	Peripheral Blood
PBX1	Pre-B-cell Homeobox 1 leukemia transcription factor
Pdx1	Pancreas and Duodenal Homeobox 1
Pep	Mouse strain C57Bl/6Ly-Pep3b x C3H/HeJ
PI	Propidium Iodide
PIC	Pre-initiating Complex
PML-RARa	Promyelocytic Leukemia (PML) - Retinoic Acid Receptor alpha (RARa) fusion gene
POLR2A	RNA Polymerase II subunit A
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA transcript
PS	Phosphatidylserine
PU.1	Purine box binding protein 1
q	long arm of chromosome
qRT-PCR	quantitative Real Time Polymerase Chain Reaction
RBC	Red Blood Cell
RP	Reverse Primer
RFP	Red Fluorescent Protein
RISC	RNA-induced Silencing Complex
DGCR8	DiGeorge syndrome Chromosomal Region gene 8

rmIL3	recombinant murine Interleukin 3
rmIL6	recombinant murine Interleukin 6
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute
RT	Room Temperature
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RUNX1	Runt-related Transcription factor (=AML1)
RUNX1T1	RUNX1 Translocation Partner 1 (=ETO)
Seq	Sequencing
shRNAs	short hairpin RNAs
ST	Short-term
SOC	Super Optimal broth with Catabolite repression
SUZ12	Suppressor of Zeste 12 protein homolog
t	translocation
TAE	Tris-acetate-EDTA
TAF1	Transcription initiation factor TFIID subunit 1
TALE	Three Amino acid Loop Extension
T-ALL	T-cell Acute Lymphocytic Leukemia
TBP	TATA box Binding Protein
TBS	Tris-buffered Saline
TCF7L2	Transcription Factor 7 Like 2
TCGA	The Cancer Genome Atlas
TF	Transcription Factor
TFAP2A (AP2)	Transcription Factor AP-2 Alpha
TFAP2C	Transcription Factor AP-2 Gamma
TFIIB	the general Transcription Factor IIB
Th	T-helper
TKD	Tyrosine Kinase Domain
TR	Tandem Repeat
TSS	Transcription Start Site
UCSC	University of California Santa Cruz
UTR	Untranslated Region
VCM	Virus Containing Medium
WBC	White Blood Cell
WHO	World Health Organization
YFP	Yellow Fluorescent Protein

1. Introduction

1.1 Normal Hematopoiesis

Hematopoiesis is classically described as a hierarchical organized system in which blood cells form and develop, resulting in about 10^{12} newly generated cells per day in vertebrates (**Figure 1**) (Doulatov et al. 2012, Fliedner et al. 2002).

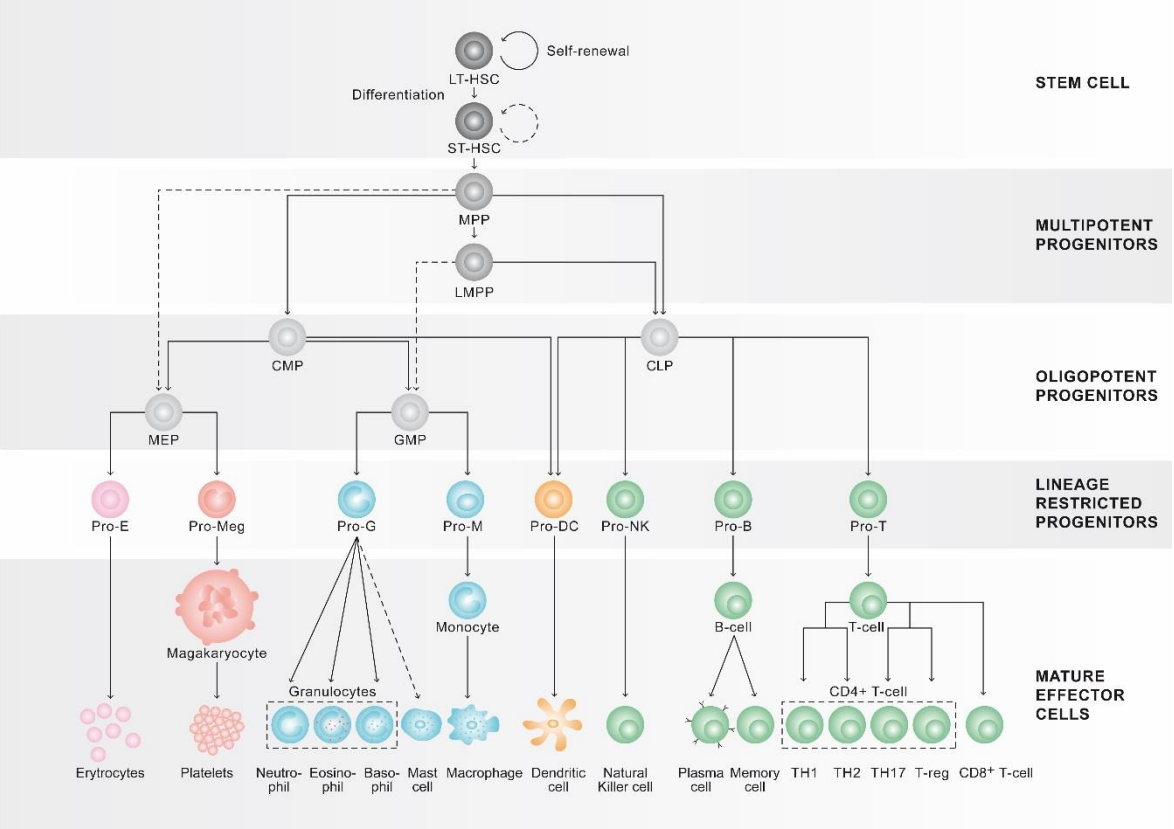


Figure 1. Hematopoietic hierarchy. The schematic illustrates differentiation stages from the most primitive stage of a long term (LT) hematopoietic stem cell (HSC) to its progenitor (Pro) and eventually its most functionally mature lymphoid and myeloid cell state. Dashed arrow lines indicate alternative pathways, except in the case of mast cells there, the dashed arrow line indicates a simplified pathway. ST-HSC: short-term HSC; MPP: multi-potent progenitor; LMPP: lymphoid-primed multi-potent progenitor; CLP, common lymphoid progenitor; CMP: common myeloid progenitor; MEP: megakaryocyte-erythroid progenitor; GMP: granulo-monocyte progenitor; E: Erythrocyte; Meg: Megakaryocyte; G: Granulocyte; M: Monocyte; DC: Dendritic cell; NK: Natural Killer cell; B: B-cell; T: T-cell; TH: T-Helper cell ($CD4^+$); T-reg: regulatory T-cell. The figure was created with Illustrator in the style of Dzierzak and Philipsen (2013) and according to information from a number of other publications (Zou and Restifo 2010, Petriv et al. 2010, Forsberg et al. 2006, Doulatov et al. 2012, Dörner et al. 2009, Dahlin and Hallgren 2015, Chotinantakul and Leeanansaksiri 2012, Adolfsson et al. 2005, Wang and Wagers 2011)

This process occurs in the bone marrow (BM) of upper arm and leg, hip and sternum after the embryonal development has been completed though asymmetric and symmetric cell division of hematopoietic stem cells (HSCs) at the top of the hierarchy (Prabhakar et al. 2009, Congdon and Reya 2008). Asymmetric cell division leads to two daughter cells: one being an exact copy of the mother cell and one a more differentiated cell. Symmetric cell division leads to either two more differentiated daughter cells or two exact copies of the mother cell. Which strategy is preferred depends on the current need of the organism. HSCs are classified into two groups. One group can reconstitute the blood long-term (LT, >six month) and the other group reconstitutes the blood short-term (ST, weeks to month) (Bartelmez et al. 1991). LT-HSCs have a self-renewal and multipotent differentiation phenotype. The self-renewal capacity allows the stem cell pool to persist lifelong. The multipotent differentiation capacity allows the stem cell pool to mature into functionally specialized blood cells. Both phenotypes enable LT-HSCs to sustainable replenish the complete hematopoietic system (Beckmann et al. 2007, Seita and Weissman 2010, Dick 2003, Morrison et al. 1995, Morrison et al. 1996, Reya 2003). ST-HSCs, which are the more differentiated daughter cells of LT-HSCs, have a limited self-renewal phenotype enabling this cell population to replenish the complete hematopoietic system only for a short period of time. The daughter cells of ST-HSCs are referred to as multipotent progenitor cells (MPPs), which do not self-renew anymore, but are highly proliferative and give rise to oligopotent progenitors (Morrison and Weissman 1994, Yang et al. 2005). These are classically classified as common lymphoid (CLP) and myeloid (CMP) progenitor cells (Kondo et al. 1997, Akashi et al. 2000). CMPs differentiate into granulocyte-macrophage progenitors (GMP) and megakaryocyte-erythrocyte progenitors (MEP). GMPs are restricted progenitors giving rise to monocytes, macrophages, granulocytes (neutrophils, eosinophils and basophiles) and mast cells. MEPs are restricted progenitors giving rise to megakaryocytes, platelets and erythrocytes. CLPs, which have developed from MPPs or lymphoid primed MPPs (LMPPs) differentiate into lineage restricted progenitors for B-, T- and natural killer (NK) cells (Adolfsson et al. 2005). Dendritic cells develop from both myeloid and lymphoid progenitors (Godin 2016, Wang and Wagers 2011). All cell populations express specific cell surface marker with which they can be isolated and identified. In particular, both hematopoietic stem cell populations lack expression markers characteristic for differentiated cells of hematopoietic lineages (Lin^-), but are double positive for the surface molecules stem cell antigen 1 (Sca-1/Ly6A) and the stem cell factor receptor (c-kit/Cd117). Thus, Sca-1 and c-kit are, in principle, primitive cell markers, which are lost during cellular differentiation. However, new markers are gained. Essential lymphocyte markers are CD19 and B220 for identifying all murine B-cell subsets independent of developmental stages as well as CD3 for identifying all murine T-cells, which can be specified for $CD4^+$ helper and $CD8^+$ killer cells (Team R&D Systems and Novus Biologicals 2015, Team Research and Development in Miltenyi Biotec 2019). An essential NK cell marker is Nk-p46 (Walzer et al. 2007). An essential monocyte marker is CD11b, which in combination with F4/80 and CD68 identifies macrophages (Taylor et al. 2004). Granulocytes also express CD11b, which in combination with $Gr1^{low}$ can be used to identify eosinophils and with $Gr1^{high}$ neutrophils and with FcεRI basophils (Min et al. 2012, Team Research and Development in Miltenyi Biotec 2019). Mast cells also express FcεRI, but they are additionally c-kit^{high} positive (Bochner and Schleimer 2001). Mast cells generated from murine BM cells *in vitro* express additionally to c-kit and FcεRI also Sca-1 (Weller et al. 2005, Drew et al. 2002). Erythrocytes are mainly $Ter119^+$ and megakaryocytes/platelets $CD41^{high}$ (Ikuta et al. 1990, Bakocevic et al. 2014, Zhang et al. 2007). Dendritic cells can be identified with the help of CD11c expression detection (Singh-Jasuja et al. 2013). However,

this classical hierarchical model is challenged these days through data generated using new single cell technologies, which suggest a rather continuous than a step-wise differentiation process in hematopoiesis (Cheng et al. 2020). That means, the HSC and MPP population is rather heterogeneous, with different lineage cell fates and thus, are not exclusively multipotent but some of the cells are lineage biased (Laurenti and Göttgens 2018).

The balance of this hierarchical or continuous hematopoietic system is maintained by a highly dynamic gene expression network with a central role assigned to chromatin remodelers and transcription factors (TFs) (Oh and Humphries 2012). Together they regulate growth factors such as granulocyte-, macrophage- and granulocyte-macrophage-colony stimulating factors as well as enzymes, adhesion molecules and other downstream transcription factors with the result of self-renewal, proliferation, differentiation or cell death (Alenzi et al. 2009, Aggarwal et al. 2012, Raghav and Gangenahalli 2018). Depending on the functional status of a hematopoietic cell TFs such as GATA, purine box binding protein 1 (PU.1), early growth response 1 (EGR1) and pre-B cell leukemia transcription factor 1 (PBX1) are involved in HSC differentiation or maintenance (including self-renewal and quiescence). Of the transcription factors, the homeobox family is particularly relevant for promoting HSC self-renewal.

GATA family transcription factors, are named like that, because they bind to the deoxyribonucleic acid (DNA) sequence: GATA found in promoters and enhancers of many target genes via zinc fingers (Evans et al. 1988, Omichinski et al. 1993, Whyatt et al. 1993, Zhang et al. 2009). There are six GATA family members amongst which GATA1-3 are classically considered in hematopoiesis (Orkin 1995). GATA1 is expressed in late progenitor cells and is essential for the maturation of erythrocytes, megakaryocytes, mast cells and eosinophils (Suzuki et al. 2003, Suzuki et al. 2009, Pevny et al. 1991, Shivdasani et al. 1997, Migliaccio et al. 2003, Yu et al. 2002). GATA2 is particularly high expressed in stem and early progenitor cells and is essential for the maintenance of HSCs throughout life (Minegishi et al. 2003, Suzuki et al. 2006, Menendez-Gonzalez et al. 2019, Tsai et al. 1994, de Pater et al. 2013). GATA3 contributes to T-lymphocyte development and is controversially discussed to regulate LT-HSC activations (Zaidan and Ottersbach 2018).

PU.1, which is encoded by the *SPI1* gene, harbors an e twenty six (ETS) domain, that recognizes the DNA sequence: GGAA of up to 3000 genes (Kodandapani et al. 1996). Additionally, PU.1 activity is substantially regulated by protein-protein interactions including GATA1 and GATA2 modifying PU.1's overall transcriptional activity. PU.1 controls maintenance and function of HSCs and regulates lymphoid differentiation (Rothenberg et al. 2019, Burda et al. 2010).

PBX1 and EGR1 are both essential for self-renewal in HSCs. EGR1, a member of the ETS family, binds DNA via its zinc finger domain and is highly expressed in HSCs and downregulated in proliferating (dividing) and migrating cells. It suppresses differentiation into granulocytes and erythrocytes and maintains the long term reconstitution capacity of HSCs (Krishnaraju et al. 2001, Min et al. 2008). PBX1 belongs to the three amino acids loop extension homeodomain containing TF family. It is preferentially expressed in LT-HSCs compared to expression levels in more mature cells (Forsberg et al. 2005). PBX1 is a target gene of the ecotropic viral integration site-1 (Evi-1), another TF for HSC maintenance and proliferation mediating its function (Shimabe et al. 2009, Yuasa et al. 2005, Goyama et al. 2008). But also PBX1 itself has been found to be essential for self-renewal of HSCs (Ficara

et al. 2008). It forms protein complexes with homeobox proteins by which their relatively nonspecific DNA-binding property is enhanced (Mann 1995, Piper et al. 1999).

Homeobox genes in normal hematopoiesis

Homeobox family genes are characterized by the possession of a particular 183 nucleotide long DNA sequence, the homeobox. It is a highly conserved region that encodes a 61 amino acid long protein domain. In transcription, the homeobox binds DNA on one side and co-factors on the other side enabling precise gene expression regulation. They have an important role in embryonic patterning and hematopoiesis (Bürglin 2005). Their specialized role in hematopoiesis has been in detail reviewed by Alharbi et al. (2013), Bhatlekar et al. (2018), Rawat et al. (2012). In short, there are two main classes of homeobox family genes: clustered (HOX) and non-clustered. The HOX gene class comprises 39 genes, which are organized in four gene clusters: A to D, localized in four different chromosomes. HOX A-C are essential during hematopoiesis by controlling the balance between proliferation and differentiation. HOXA genes are mostly expressed in myeloid, HOXB in erythroid and HOXC in lymphoid cells. On HSC and progenitor level, *HOXA9* is the most abundant HOX gene, whose expression level decreases with increased differentiation state (Laurenti et al. 2013, Skvarova Kramarzova et al. 2014, Argiropoulos and Humphries 2007). Functional studies with *Hoxa9* knock-out mouse cells revealed that less lymphoid and myeloid progenitor cells were available and the repopulating ability in competitive transplantation assays was diminished, but could be rescued by *Hoxa9* reintroduction. Thus, HOXA9 promotes self-renewal and proliferation, whereas it inhibits differentiation (Pineault et al. 2002, Sauvageau et al. 1994, So et al. 2004, Lawrence et al. 2005). There are a number of HOXA9 target genes such as *HOXA7*, which has been identified to be involved in myeloid/erythrocyte differentiation (So et al. 2004, Faber et al. 2009). Selectivity of Hoxa9 and specificity to target genes is controlled by the cofactors Pbx1 and Meis1, which are also homeobox genes but due to their low sequence identity to genes of the two classes (clustered and non-clustered) and the addition of a three amino acid loop extension (TALE) they cannot be assigned to any of the classes, but are classified as TALE homeobox genes (Dard et al. 2019, Dassé et al. 2012). The non-clustered homeobox gene class includes gene families, which harbor the homeobox, but appear rather divergent than clustered throughout the genome. These include Msh homeobox (MSX), Paired box (PAX) and Distal-less homeobox (DLX) genes as well as Engrailed and ParaHox genes. ParaHox genes show homology to different paralogue groups of HOX clusters apart from the homeobox. They are so to speak sisters of clustered HOX genes. ParaHox genes include GS homeobox 1 (Gsx1) and 2 (Gsx2), pancreas and duodenal homeobox 1 (Pdx1) and caudal-type homeobox 1 (Cdx1), 2 (Cdx2) and 4 (Cdx4). Cdx2 and Cdx4 are particularly relevant in embryonic and malignant adult hematopoiesis by inducing gene expressions of HOXA and HOXB cluster (Rawat et al. 2012). One prominent hematopoietic malignancy is blood cancer scientifically referred to as leukemia.

1.2 Malignant Hematopoiesis: Leukemia

Leukemia and its subclasses

Leukemia is one of the top 10 deadliest cancer types in the United States (U.S. Cancer Statistics Working Group). It is a disease of the BM and blood, in which abnormally differentiated white blood cells (WBC blasts) of a primitive malfunctioning phenotype clonally expand and inhibit healthy stem and progenitor cells to differentiate into well-functioning mature WBCs through a grossly perturbed gene expression network (Döhner et al. 2015, Wujcik 2003, German Cancer Research Center 2020). Depending on the time it takes for the disease to develop and the affected blood cell type, leukemia can be grouped into four categories: acute myeloid leukemia (AML), acute lymphocytic leukemia (ALL), chronic myeloid leukemia (CML) and chronic lymphocytic leukemia (CLL). CML and AML are the most common leukemia subtypes, but AML is the deadliest (Team for Medical and Editorial Content in The American Cancer Society 2020).

Acute Myeloid Leukemia (AML)

As AML accounts for 50% of leukemia related deaths, it is important to identify risk factors (Team for Medical and Editorial Content in The American Cancer Society 2020). Commonly recognized are age, male gender, certain blood disorders (e.g. thrombocythemia, myelodysplatic syndrome), certain chemotherapy drugs used in cancer treatment, cytogenetic and genetic abnormalities and environmental agents such as ionizing radiation (including therapeutic radiation), benzene (e.g. in dyes, glues, cleaning products) and smoking (Team for Medical and Editorial Content in The American Cancer Society 2018). In AML, primitive myeloid blast cells rapidly proliferate and differentiate abnormally to a lesser extent. Early symptoms caused by the reduction of normal erythroblasts and their progeny are: anemia, weakness, constant fatigue and paleness can occur. The reduction of normal granulocytes causes fever and infection. Decreased platelets level may result in sudden bleeding with following colored spots on the skin (bruises, petechiae, ecchymosis). Gingival enlargement has been reported to be another potential early sign of AML as Lim and Kim (2014) concluded form patients diagnosed with AML M4. Major cornerstones of current treatment consist of poly chemotherapy, partly followed by allogenic stem cell transplantation and in some cases targeted drug therapy such as FLT3 inhibitors (Heuser et al. 2020). However, about 80% of ≥65 years old patients die within five years from the time point of diagnosis (Klepin et al. 2019). Thus, in order to support treatment decisions diagnostics of this very heterogeneous disease are crucial. There are two classification systems known. The former French-American-British (FAB) system and the current World Health Organization (WHO) classification system. The FAB system is based on cytomorphology for grouping AML patients into eight subtypes (M0-M7) according to original cell type and maturation status of the leukemic cells. The minimum requirement for an AML diagnosis was a blast cell count of 30% in the BM for most of the subtypes (Catovsky et al. 1991). The WHO classification, additionally to the FAB criteria, includes genetic alterations and patient history based on which AML patients can be assigned to three groups: (1) AML with recurrent genetic abnormalities (2), AML with myelodysplasia-related changes and (3) AML not otherwise specified according to chromosomal aberrations, gene mutations, co-existing myelodysplasia and pre-existing therapies (**Table 1**) (Harris et al. 1999, Swerdlow et al. 2008). The minimum requirement for an AML diagnosis is a blast count of ≥20% in the

BM or peripheral blood (PB) for an AML diagnosis, unless an AML specific chromosomal aberration has been found.

Table 1. WHO classification of acute myeloid leukemia (AML) without considering myeloid neoplasms. Table based on information from Arber et al. (2016)

Classification	Specification
AML with recurrent genetic abnormalities	AML with t (8;211)(q22;q22.1);*RUNX1-RUNX1T1*AML with inv (16)(p13.1q22) or t (16;16)(p13.1;q22);*CBFB-MYH11*APL with *PML-RARA* AML with t (9;11)(p21.3;q23.3);*MLLT3-KMT2A* AMLwith t (6;9)(p23;q34.1);DEK-NUP214 AML with inv (3)(q21.3q26.2) or t (3;3)(q21.3;q26.2); *GATA2, MECOM* AML (medakaryoblastic) with t (1;22)(p13.3;q13.3);*RBM15-MKL1* *Provisional entity: AML with BCR-ABL1* AML with mutated *NPM1* AML with biallelic mutations of *CEBPA* *Provisional entitiy: AML with mutated RUNX1*
AML with myelodysplasia-related changes	None
AML not otherwise specified	AML with minimal differentiation AML without maturation Acute myelomonocytic leukemia Acute monoblastic/monocytic leukemia Pure erythroid leukemia Acute megakaryoblastic leukemia Acute basophilic leukemia Acute panmyelosis with myelofibrosis

t:translocation; p: short arm of chromosome; q: long arm of chromosome

Cytogenetics and mutations in AML

AML is a heterogeneous disease characterized by cytogenetics (chromosome-related) and mutations (gene-related). Chromosomal changes account for 50-60% of newly diagnosed AML patients. These patients have chromosomal aberrations such as t (11q23): mixed-lineage leukemia (*MLL*) abnormalities, t (15;17): promyelocytic leukemia - retinoic acid receptor alpha (*PML-RARa*), t (8;21): *AML1-ETO* and inversion (Inv) (16): core-binding factor subunit beta – myosin heavy chain 11 (*CBFß-MYH11*) or have a complex karyotype defined by at least three chromosomal aberrations. The other 40-50% of the AML patients do not have chromosomal changes. They suffer from so-called cytogenetically normal AML (CN-AML), which is characterized by multiple gene mutations (Walker and Marcucci 2012). The most common gene mutations are in nucleophosmin 1(*NPM1*) and FMS-like tyrosin kinase (*FLT3*), which depending on their type of co-occurrence, can classify CN-AML patients into the favorable or intermediate risk group defined by the old European Leukemia Network (ELN) classification system (Döhner et al. 2010). From 2017 onwards it was agreed that mutations, are the superior factor in risk stratification, irrespective of the background

cytogenetics (Döhner et al. 2017). This led to the recognition of new critical mutations and a simplification from four to a three risk groups, but the classification of CN-AML patients with *NPM1* and/or *FLT3* mutation remained the same, it only included now patients with abnormal karyotype (**Table 2**).

Table 2. Risk stratification according to ELN 2017 with the focus on *NPM1* and *FLT3* mutations. Table based on information from Döhner et al. (2017)

Risk category	Genetic abnormality
Favorable (55% OS)[#]	Mutated *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low*} *CEBPα* mutation *PML-RARa*, *AML1-ETO*, Inv 16
Intermediate (24% OS)[#]	Mutated *NPM1* and *FLT3*-ITD^{high*} Wild-type *NPM1* without *FLT3*-ITD or with *FLT3*-ITD^{low*}(without adverse risk genetic lesions)
Adverse (5% OS)[#]	CN-AML with wild-type *NPM1* and *FLT3*-ITD^{high*} *MLL* rearrangements Complex and monosomal karyotype *ASXL1*, *RUNX1* and *TP53* mutations

* low: low allelic ratio (<0.5), high: high allelic ratio (≥0.5), 'semi-quantitative assessment of [the] *FLT3*-ITD allelic ratio (using DNA fragment analysis) is determined as ratio of the area under the curve [for] *FLT3*-ITD divided by [the] area under the curve [for] *FLT3*-wild type' (Döhner et al., 2017, p. 432). OS: Overall survival. [#]Byrd et al. (2002)

Classically *NPM1* mutations are believed to be favorable and *FLT3*-ITD mutation unfavorable/adverse (Liu, He, et al. 2014, Small 2006, Stirewalt and Radich 2003, Kottaridis et al. 2003). However, *NPM1* mutations are only favorable, if no *FLT3* with internal tandem repeat (*FLT3*-ITD) or *FLT3*-ITD with low allele ratio is present. If mutated *NPM1* co-occurs with *FLT3*-ITD of high allele ratio, the patient is classified into the intermediate-risk group. *FLT3*-ITD is only adverse, if its allele ratio is high and *NPM1* is not mutated. But *FLT3*-ITD co-occurrence with mutated *NPM1* classifies the patient into the intermediate-risk group. If the *FLT3*-ITD allele ratio is low and no *NPM1* mutation is present, patients are classified into the intermediate risk-group, but it becomes a favorable situation if *NPM1* is mutated. This example shows that mutations of two genes must be complementary and occur together to be transforming. So it was believed for a long time that genetic aberrations of two different classes had to occur simultaneously in a single cell in order to malignantly transform that cell causing leukemia, a concept termed the 'two-hit hypothesis' (Gilliland and Griffin 2002, Naoe and Kiyoi 2013). Class I comprises mutant genes, which confer a proliferative and/or survival advantage, but do not affect differentiation such as *FLT3*-ITD or with mutation in the tyrosine kinase domain (*FLT3*-TKD), *RAS* mutations and receptor tyrosine kinase (*KIT*) mutations. Class II comprises mutant genes, which are involved in maintaining self-renewal and impair differentiation: chromosomal aberrations: t (8;21), t (15;17), Inv (16) and *MLL (*now termed *KMT2A*) rearrangements (*MLLr*); gene mutations: runt-related transcription factor 1 (*RUNX1*), *NPM1*, CCAAT/enhancer-binding protein alpha (*CEBPα*). This 'two-hit hypothesis' has being challenged in 2015 by Greaves, who postulated a new idea of a 'single big hit' hypothesis, in which one mutation is sufficient to cause certain leukemias (eg. *MLL*) by globally influencing molecular pathways resulting in

a cascade of different disease-related gene expressions (Greaves 2015). Additionally, new gene mutations, which could not be classically classified have been found and two new classes were defined, which co-occur with class I or class II mutant genes (**Figure 2**) (Naoe and Kiyoi 2013, Takahashi 2011, Kao et al. 2014). Class III comprises mutant genes, which are involved in DNA- and histone methylation: Tet Methylcytosine Dioxygenase 2 (TET2), Isocitrate Dehydrogenase (NADP(+)) 1/2 (IDH1/2), DNA (cytosine-5)-methyltransferase 3A (DNMT3A), ASXL Transcriptional Regulator 1 (ASXL) and Enhancer of Zeste Homolog 2 (EZH2). Class IV comprises mutant genes, which are involved in tumor suppression: Wilms Tumor 1 (WT1) and Tumor Protein P53 (TP53).

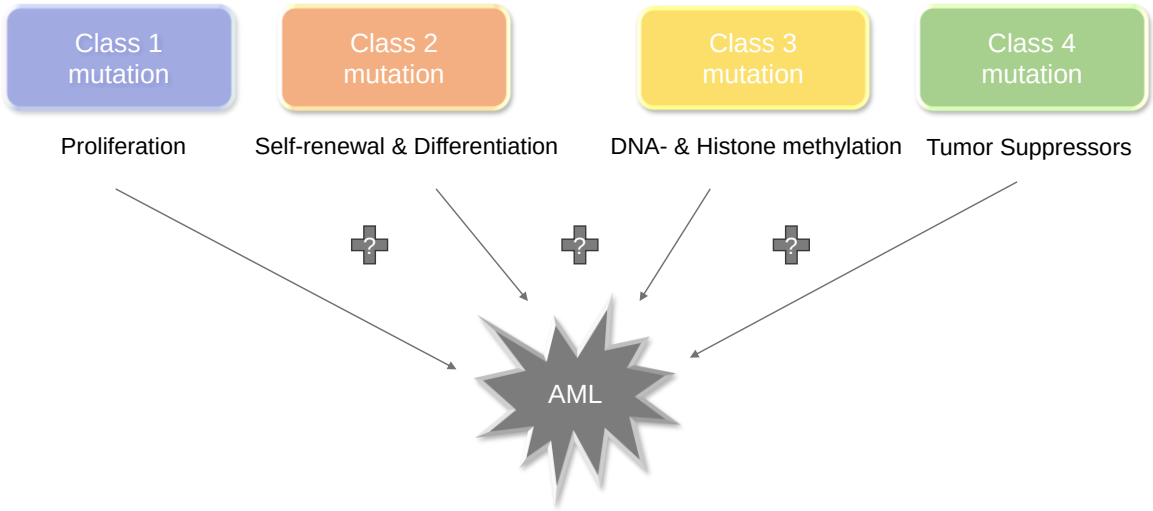


Figure 2. Gene class hypotheses in acute myeloid leukemogenesis. The schematic illustrates different classes of genes that alone or in combination are necessary for causing acute myeloid leukemia (AML).

Despite those very sophisticated diagnosis-relevant measures used to choose best treatment strategies, there is still one major problem: the hard to eradicate leukemia propagating leukemic stem cells (LSCs) (Hanekamp et al. 2017).

Leukemic stem cells (LSCs)

Leukemia-initiating cells (LICs) or leukemic-stem cells (LSCs) are a rare (1 in 1.000.000) cell population among leukemic cells with the property to initiate leukemia in a xenograft mouse model (Lapidot et al. 1994, Bonnet and Dick 1997, Villatoro et al. 2020). LSCs can be derived from stem as well as from progenitor cells, which were initially pre-leukemic (Valent et al. 2019). LSCs are able to self-renew and differentiate into more mature leukemic cells, thereby sustaining the continued growth of premature and mature leukemic bulk cells in BM and blood (**Figure 3**) (Hanekamp et al. 2017, Villatoro et al. 2020, Zagozdzon and Golab 2015, Chopra and Bohlander 2019).

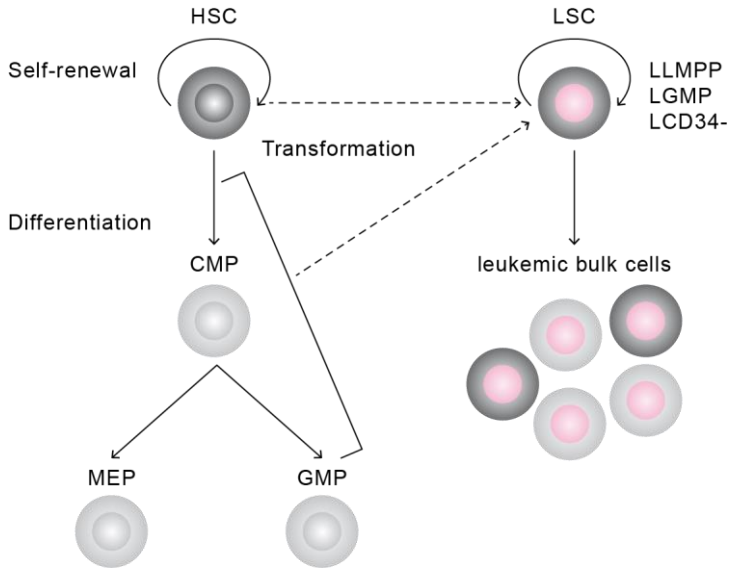


Figure 3. Leukemic transformation. The schematic illustrates different stages of primitive hematopoietic cells at which transformation to leukemia initiating/stem cells (LIC/LSC) can take place. HSC: hematopoietic stem cell; CMP: common myeloid progenitor, MEP: megakaryocyte-erythroid progenitor; (L)GMP: (leukemic) granulo-monocyte progenitor, LLMPP: leukemic lymphoid-primed multi-potent progenitor. The figure was created with Illustrator on the basis of Figure 1 and according to information from Pandolfi et al. (2013), Goardon et al. (2011), Dong et al. (2019), Horton and Huntly (2012), Heuser et al. (2011).

The LSC population suppresses the production of healthy blood cells by proliferating rapidly and creating a microenvironment, which is hospitable for them but not for healthy stem and progenitor cells. Moreover, due to the quiescence property LSCs are relatively chemotherapy-resistant. The eradication of LSCs is recognized as crucial aspect in leukemia therapy, which has resulted in ~20 years of intense research on the identification and molecular characterization of LSCs (Ishikawa et al. 2007, Laverdière et al. 2018, Saito et al. 2010, Zeijlemaker et al. 2016, Terwijn et al. 2014).

For LSC characterization a multi-parametric FACS analysis based on defined markers is used these days. Proof of LSC properties depends on validating their functional capacity to induce leukemic engraftment and leukemia-development in murine transplantation assays or in case of human LSCs in xenograft assays. John Dick's group paved the way for this cell surface marker based LSC characterization in the year 1990. They used two transmembrane proteins: CD34 and CD38 for detecting LSCs, which were later functionally validated by Eppert et al. (2011) through xenograft assays. Ng et al. (2016) and Terwijn et al. (2014) confirmed that the $CD34^+/CD38^-$ cell fraction harbors most LSC activity capable of initiating leukemia in primary transplanted mice and maintaining the disease in secondary and third transplantations. However, also healthy stem cells coexist with LSCs in the bone marrow, are $CD34^+/CD38^-$. Therefore a number of additional markers such as CD90, CD45RA, CD110 and CD123 to LSCs and its subpopulations LGMP and LLMPP next to the less potent $CD34^-$ cell population were established (**Figure 3**) (Taussig et al. 2010, Goardon et al. 2011). In 2016, Zeijlemaker et al. established a cocktail of 13 different cell surface markers including CD45, CD34, CD38, CD45RA, CD44, CD123, CD33, CLL-1, TIM-3, CD7, CD11b, CD22 and CD56 presenting a robust and reliable marker based method for LSC characterization and frequency estimation (Zeijlemaker et al. 2016). Besides these cell surface markers and cancer-causing mutations, there are LSC specific

gene expression signatures, which have been found to correlate with stemness and survival (Ng et al. 2016, Gentles et al. 2010, Eppert et al. 2011, Metzeler et al. 2013). The most recognized signature is the one from Ng et al. (2016). He and colleagues first identified 104 top differentially expressed genes (DEGs) in functionally defined LSCs. Those DEGs include 48 up- and 56 downregulated genes strongly correlating with global gene expression patterns of normal HSCs and MPPs suggesting a similar transcriptional program. A sparse regression analysis of these DEGs with survival in a large training cohort containing 495 AML patients revealed an association of 17 genes (LSC17 score) with inferior clinical outcome and with a prognostic impact in *de novo* (without previous bone marrow or cancer disease) AML (Ng et al. 2016). Recently, this LSC17 score became particularly valuable for a refined risk classification (low, intermediate, adverse) of younger AML patients <60 years of age beyond classical cytogenetic or molecular markers (Bill et al. 2020). Moreover, a LSC epigenetic signature by Jung et al. (2015) revealed differentially methylated DNA regions of HOXA cluster genes including hypomethylation of *HOXA5*, *HOXA6*, *HOXA7*, *HOXA9* and *HOXA10*.

HOXA9 in AML

HOX genes are established key transcription factors regulating normal hematopoiesis and their aberrant expression is strongly associated with leukemia (Rice and Licht 2007, Alharbi et al. 2013). Particularly in AML, *HOXA* and *HOXB* gene families were highly expressed, while *HOXC* and *HOXD* were not detected (Spencer et al. 2015, Andreeff et al. 2008). *HOXA* gene expression levels could be strongly correlated with poor survival rates compared to patients with non-*HOX* or only *HOXB* expression levels (Spencer et al. 2015, Chen, Qin, et al. 2019). Along, *HOXA* genes were found to be particularly high expressed in LSCs and highest in leukemic bulk cells of *MLLr* and CN-AML patients with *NPM1c*$^{+}$ mutation, which correlates with poor survival (Alcalay et al. 2005, Debernardi et al. 2003, Nagy et al. 2019, Wang et al. 2020, Ng et al. 2016, Jung et al. 2015). Especially *HOXA9* was found to be the best predictor for survival. (Andreeff et al. 2008, Golub et al. 1999). It is overexpressed in about 70% of AML cases (Aryal et al. 2021). Functional studies to investigate its oncogenic role revealed that *HOXA9* fusion with *NUP98*, *Hoxa9* wildtype overexpression and in collaboration with *Meis1* transform primary bone marrow cells and cause leukemia in mice (Borrow et al. 1996, Chung et al. 2006, Ziller 2021, Kroon et al. 1998, Bach et al. 2010). *HOXA9* attenuation via a histone methyltransferase interfering with *Hoxa9* gene expression or short hairpin RNA (shRNA) counteracts proliferation and self-renewal of AML cells leading to reduced disease progression (Lehnertz et al. 2014, Faber et al. 2009). HOXA9 mediated leukemogenesis occurs due to its binding to promoters (region close to the 5' untranslated region (UTR) of a gene) and enhancers (distal to the 5' UTR of a gene) regulating the transcription of key oncogenes (Zhong et al. 2018). HOXA9 also creates new enhancer regions involved in malignant gene transcription by altering histone methylation states (Sun et al. 2018). However, there is no ideal treatment strategy yet to target HOXA9 making the development of new strategies necessary by elucidating HOXA9 expression mechanisms and HOXA9-dependent transcriptional processes (Lambert et al. 2019). In *MLLr* leukemia, high *HOXA* levels including HOXA9 have been attributed to the function of epigenetic modulators (e.g. DOT1L) attracted by the fusion gene at the promoter site of *HOXA* genes (Yokoyama 2015, Aryal et al. 2021). In CN-AML patients with *NPM1c*$^{+}$, one model suggests that the loss of NPM1 function in the nucleolus due to its mutation induced transfer into the cytosol affects MLL driven *HOXA* gene expression (Kühn et al. 2016). On the other hand, *CDX2* is highly expressed in CN-

AML/*NPM1c*$^{+}$ patients, a gene that is completely absent in healthy hematopoietic cells of the adult, but acts as direct upstream regulator of especially *HOXA* genes in AML (Rawat et al. 2008, Rawat et al. 2012). Alternatively, *HOX* gene upregulation might be promoted via long non-coding ribonucleic acids (RNAs) on transcript level or via microRNAs (miRNAs) whose impact on post-transcriptional gene expression (e.g. *HOXA9*) level is reduced (Wang et al. 2011, Xu and Wen 2018, Li, Huang, et al. 2012).

1.3 MicroRNA biogenesis and function

MiRNAs

MiRNAs are small non-coding RNAs. The fact that they do not code for proteins and are only ~22 nucleotides (nts) long may lead to the assumption that miRNAs are not important for a cell. However, miRNAs are anything but insignificant for molecular cell biology. In 1993, the first miRNA was identified by joint efforts of the Ambros' and Ruvkun's laboratories, which revolutionized the field of molecular biology (Lee et al. 1993). They found, that the small non-coding RNA *lin-4* had the ability to downregulate *lin-14* gene expression post-transcriptionally by binding to its 3' untranslated region (UTR) in *Caenorhabditis elegans* (*C. elegans*). Seven years later Reinhart and colleagues from Ruvkun's lab discovered another small non-coding RNA called *let-7*, which controls the expression of genes that coordinate events during development of *C. elegans* (Reinhart et al. 2000). From that date on, these small endogenous RNA molecules were termed miRNAs, which are present also in several other organisms besides *C.elegans*, most of which being evolutionarily conserved (Meunier et al. 2013). According to the latest update in miRBase 2018 there are currently 1234 miRNAs in the murine and 1917 miRNAs in the human genome registered (Griffiths-Jones 2018). Its general function is to change the expression profile of cells rapidly and transiently in response to intracellular or extracellular changes (Wang, Zhang, Weber, et al. 2010, Kucherenko and Shcherbata 2018, Ghini et al. 2018, Nam et al. 2014, Thomson and Dinger 2016).

MiRNA biogenesis

MiRNA biogenesis starts in the nucleus with the transcription of miRNA genes by polymerase II or III acting on upstream regulatory regions, so-called promoters. In this way transcripts are generated undergoing a canonical (dominant) or non-canonical maturation process (Lee et al. 2004, Borchert et al. 2006, O'Brien et al. 2018) (**Figure 4**). The canonical pathway starts off with the transcription of a miRNA gene located between (intergenic) or within (intronic) protein-coding genes from a miRNA-specific or the host-gene promoter into a long non-coding primary transcript (pri-miRNA). The pri-miRNA can be up to 100 kilo bases (kb) long, harbor one or more hairpin structures/stem-loops (mono- or polycistronic), are capped and polyadenylated (Lee et al. 2002). It is further processed to a ~70-85 nucleotides (nts) long precursor miRNA (pre-miRNA) by a so-called microprocessor complex consisting of the ribonuclease enzyme Drosha and its binding partner DiGeorge syndrome chromosomal region gene 8 (DGCR8, in drosophila known as Pasha). Under energy consumption (carried out by Ran-GTP), Exportin 5 transports the pre-miRNA stem-loop out of the nucleus into the cytosol. Here, the endonuclease RNAse III named Dicer cleaves in collaboration with its cofactor the pre-miRNA 'loop' from the 'stem', leaving behind a double stranded miRNA containing 18-24 nts (Bartel 2004, Denli et al. 2004, Kim and Nam 2006, Grishok et al. 2001). For strand selection, the miRNA duplex is loaded onto

the Argonaute (AGO) protein, which prefers to keep the less stable 5p 'guiding' strand. The other 3p 'passenger' strand is left behind for degradation. AGO and the guiding strand form the so called RNA-induced silencing complex (RISC) to silence gene expression. However, not all miRNAs follow canonical processing, multiple non-canonical miRNA biogenesis pathways have been described, in which the transcription of miRNAs does not produce long pri-miRNAs and during maturation either Drosha/DGCR8 processing or Dicer cleavage are bypassed. Mirtrons and 7-methylguanosine (M^7G)-capped pre-miRNA are examples of a Drosha/DGCR8-independent process. Mirtrons are short hairpin introns, which were spliced out in the process of messenger RNA (mRNA) generation from pre-RNAs. M^7G pre-miRNAs are directly transcribed from DNA, consist of ~50 nts and are directly exported into the cytosol via Exportin 1 where the 3p strand is loaded onto the AGO to form the RISC (Xie et al. 2013). A microprocessor dependent, but Dicer independent process has been identified for endogenously transcribed short hairpin RNAs (shRNAs), whose maturation process is mediated via AGO2 slicing (Yang et al. 2010).

MiRNA function

At the end of the canonical and non-canonical miRNA biogenesis pathways, classically mature miRNA strands bind with their 2-8 nts long seeding region at the 5' end to partially complementary mRNA target sequence at its 3' UTR region and together with an AGO protein form the RISC complex (Jonas and Izaurralde 2015, Iwakawa and Tomari 2015, Ellwanger et al. 2011, Xu et al. 2014). Stability and specificity of such miRNA:target bindings can be supported by additional pairing of the miRNA's 3' end to mRNA preventing a fast dissociation of the RISC through which mRNA can escape degradation (Broughton et al. 2016, Wee et al. 2012). Then, there are three possibilities for gene silencing: (a) RISC-mediated prevention/repression of mRNA translation, (b) RISC-mediated mRNA degradation or (c) receptor-mediated activation of downstream signaling (Iwakawa and Tomari 2015). For translational prevention, the RISC complex inhibits the eukaryotic translation initiation factor 4 from binding to the mRNA cap, thus, ribosome recruitment and following translation is impaired (Humphreys et al. 2005). For translational repression, the RISC complex causes ribosome drop off or its proteolysis so that translation cannot be completed (Olsen and Ambros 1999, Petersen et al. 2006). Both translational inhibition mechanisms lead to decreased protein levels of target genes, but mRNA levels remain unaffected. For mRNA degradation a glycine-tryptophan (GW) repeats containing protein of 182 kDa binds to the RISC and recruits the deadenylases proteins poly(A)-nuclease 2 and 3 as well as the carbon catabolite repressor 4 (CCR4) - Negative on TATA (NOT) complex. CCR4-NOT recruits a decapping complex to the 5' mRNA site. For the activation of downstream signaling pathways miRNAs can bind to Toll-like receptors in cancer (Bayraktar et al. 2019). Beside these four established miRNA effector mechanisms in the cytosol, it has recently been shown, that miRNAs can traffic into the nucleus and cause mRNA degradation (Bottini et al. 2017, Pitchiaya et al. 2017).

Primarily, miRNA function begins with its paring to mRNA, which, however can occur in different ways determining the mode of action. Classically miRNA binds with its seeding region (2-8 nts) to the 3' UTR region of mRNA. Alternative miRNA binding sites are: the anchor (1 nt), central (9-12 nts), 3' supplementary (13-16 nts) and 3' tail (17-22 nts) binding domains. For example, an involvement of the central region in miRNA:mRNA pairing determines the catalytic conformational change of the RISC complex, which in case of too many mismatches does not take place anymore and silencing is prevented (Wee et al.

2012). Moreover, miRNAs can also bind to the coding region or the 5’ UTR of mRNAs (Lytle et al. 2007, Hafner et al. 2010, Chi et al. 2009). Furthermore, while most of the classical and alternative miRNA:mRNA pairings lead to gene silencing, miRNAs can also be gene activating, for example when miRNAs associate with proteins that resolve an inhibitory complex at a target mRNA or when miRNAs bind to promoters in the nucleus where they activate gene transcription (Xiao et al. 2017, Vasudevan 2012).

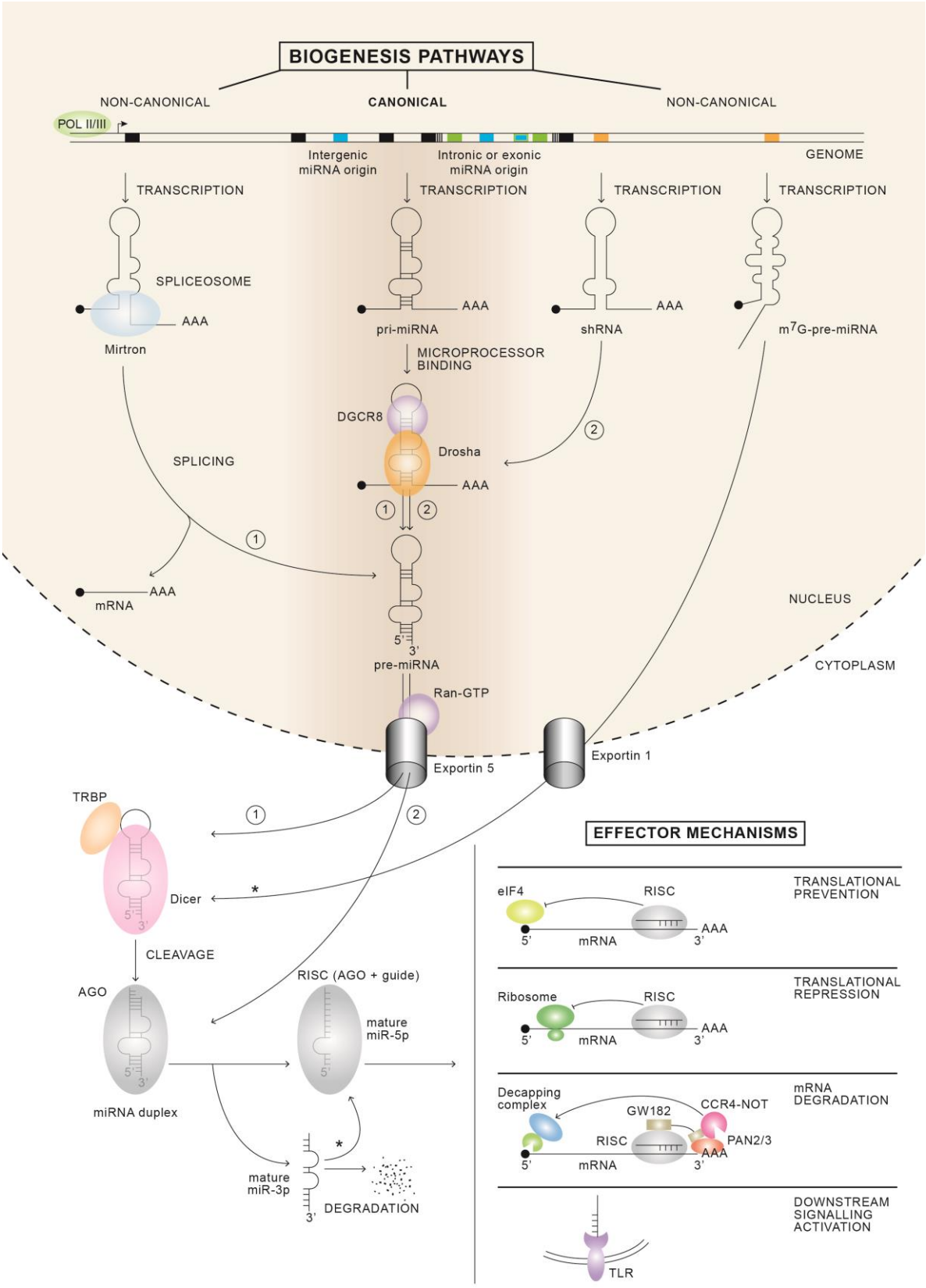


Figure legend on next page

Figure 4. MiRNA biogenesis and effector mechanisms. During miRNA biogenesis, Pol II or III initiate the transcription of transcripts, which are in case of the canonical pathway primary long-non coding miRNAs. These are further processed to pre-miRNAs by the microprocessor Drosha/DGCR8 in the nucleus. Via Exportin 5 shuttling, pre-miRNAs get into the cytosol, where they are cleaved by Dicer utilizing TRBP as cofactor for a time-efficient action. Then an AGO protein associates with the guiding strand and forms the so called RISC complex. This complex acts at complementary mRNA sites mediating mRNA silencing via different effector mechanisms. RISC binding can initiate inhibition of translation initiation by recruiting eIF4 to the mRNA cap or repression of the ribosome-mediated decoding mechanism. The recruitment of a decapping complex to the mRNA cap and a degradation complex to the tail results in mRNA degradation. Mature miRNA binding to TLR can lead to signal activation. During the non-canonical pathways mirtrons, transcribed shRNAs or m^7G-pre-miRNAs are processed slightly differently. Mirtrons occur as intronic splice products during mRNA processing and are further processed as pre-miRNAs (1). shRNAs are processed via Drosha/DGCR8 and without further cleavage associate with an AGO protein (2). M^7G-pre-miRNAs skip microprocessor binding, shuttle into the cytosol via Exportin 1 and then follow classical Dicer and RISC complex formation with the 3p strand. Blue boxes represent miRNA genes, black boxes protein coding genes. Black box with interrupted end or beginning represents a cut open gene revealing regions of exons represented as green boxes. Orange boxes are miRNA genes that are processed via a non-canonical pathway. m^7G: 7-methylguanosine cap; AAA: poly(A) tail; shRNA: short hairpin RNA; RNA: Ribonucleic acid; pri-miRNA: primary miRNA; miRNA: micro RNA; pre-miRNA: precursor miRNA; TRBP: The RNA-binding protein; Pol: RNA Polymerase; DGCR8: DiGeorge syndrome chromosomal region gene 8; AGO: Argonaute protein; RISC; RNA induced silencing complex; eIF4; eukaryotic translation initiation factor; CCR4-NOT: carbon catabolite repressor 4-Negative on TATA ; PAN2/3: poly(A)-nuclease 2/3; GW182:glycine-tryptophan (GW) repeats containing protein of 182 kDa; TLR: Toll-like receptor. The figure was created with Illustrator according to information from a number of publications: O'Brien et al. (2018), Ender and Meister (2010), Undi et al. (2013), Nakanishi (2016), Peng and Croce (2016), Elton and Yalowich (2015), Chen, Heikkinen, et al. (2019).

Overall, miRNA function is very diverse due to its variety of mRNA targets. One miRNA can target several mRNAs. Thus, miRNAs are implicated in almost every biological process including cell growth, tissue differentiation, self-renewal, cell proliferation, embryonic development and apoptosis by regulating signaling pathways and (de-) methylation (Sayed and Abdellatif 2011, Braconi et al. 2010, Duursma et al. 2008). They play an important role in a variety of different organisms and tissue types including the hematopoietic system of vertebrates.

1.4 MiRNAs and miR-196 in normal hematopoiesis

MiRNAs in normal hematopoiesis

MiRNAs are important regulators during hematopoiesis, which first became evident by detecting miRNA specific expressions in different hematopoietic tissues (Chen et al. 2004). Other studies reported that knockout of Ago2 involved in RISC formation or the Dicer protein involved in miRNA maturation, lead to sever reduction of hematopoietic stem cell population and a failure of lineage commitment (O'Carroll et al. 2007, Cobb et al. 2006, Koralov et al. 2008, Guo, Lu, et al. 2010). MiRNA expression profiles, determined by microarrays, next-generation sequencing, a bead based detection system, single molecule measurements or qRT-PCR, revealed miRNA enrichment at certain developmental stages that are involved in fine-tuning every step during hematopoiesis by maintaining stem cell properties and initiating or blocking differentiation (**Figure 5**) (Undi et al. 2013). This is accomplished by individually organized miRNAs, but also by groups of miRNAs organized in a single transcriptional unit termed miRNA clusters (e.g. one pri-miRNA serves for six different

mature miRNAs) and by miRNA families sharing significant sequence homology (Bartel 2004, Yu et al. 2006).

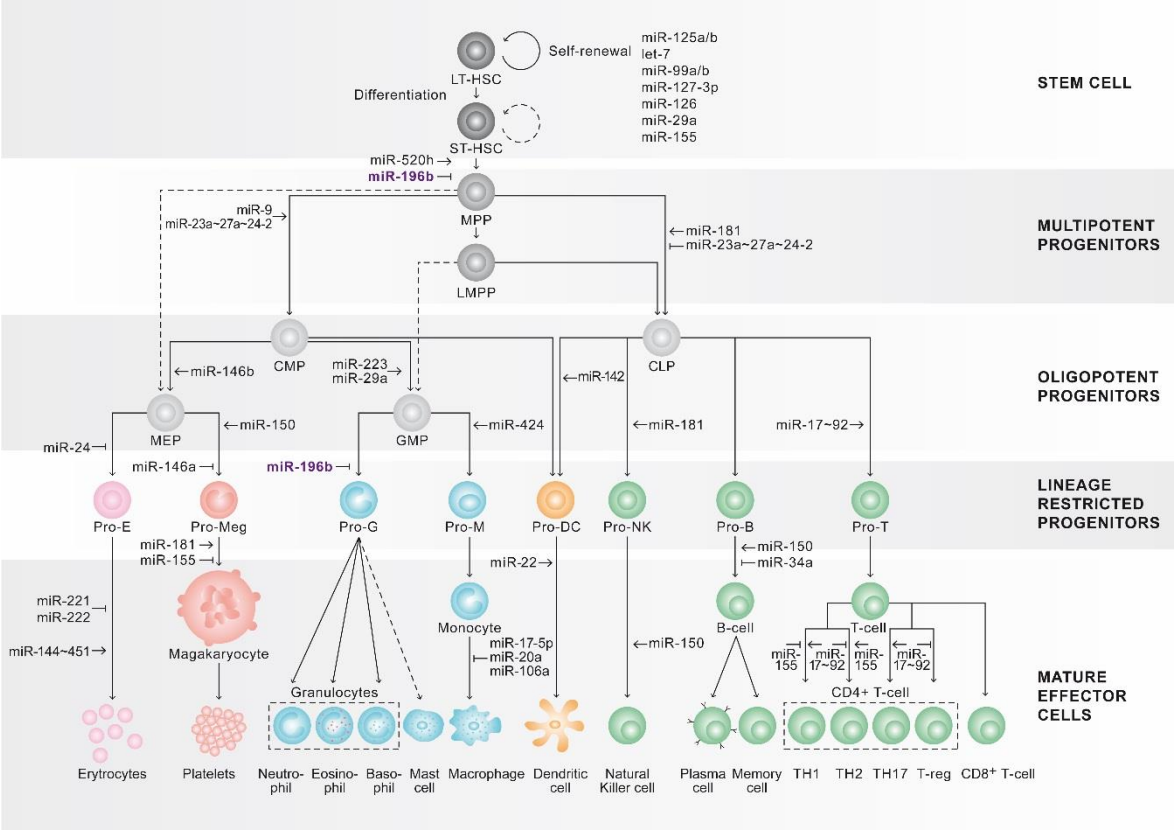


Figure 5. Regulation of hematopoiesis via microRNAs. Depicted are representative microRNAs along the hematopoietic hierarchy inspired by two reviews (Montagner et al. 2014, Bissels et al. 2012) and refined by information from functional studies (Guo, Lu, et al. 2010, Connell et al. 2010, Gerrits et al. 2012, Han et al. 2010, Li, Zhang, et al. 2012, Xiao et al. 2007, Georgantas et al. 2007, Senyuk et al. 2013, Bezman et al. 2011, Zhan et al. 2007, Labbaye et al. 2008, Zhai et al. 2014, Zardo et al. 2012, Gentner et al. 2015, Vian et al. 2014, Lechman et al. 2012, Laura et al. 2019, Hu et al. 2015, Hashimi et al. 2009, Li, Greeley, et al. 2012, Mildner et al. 2013, Cichocki et al. 2011, Velu et al. 2009, Romania et al. 2008, Ghani et al. 2011, Kurkewich et al. 2017, Popovic et al. 2009, Regelin et al. 2015, Banerjee et al. 2010, Jiang et al. 2011, Simpson et al. 2014, Liu, Jiang, et al. 2014). The figure was created with Illustrator on the basis of Figure 1. Inhibition or activation along a particular path are indicated by a blunt arrow or an arrow with a peak, respectively.

For example lymphoid lineage commitment at a late stage in hematopoiesis is controlled by miRNAs such as miR-181 and miR-150. Particularly in B-cell development miR-150 is engaged in the transition of pro-B to pre-B cells and miR-181 overexpression in lineage negative (Lin^-) cells cultured *in vitro* or transplanted into mice resulted in increased number of $CD19^+$ B-cells (Xiao et al. 2007, Chen et al. 2004). The B-cell function in immune response has been found to depend on miR-155 and the miR-17~92 cluster. Both miRNAs are additionally involved in T-cell development (Vigorito et al. 2007, Lai, Gonzalez-Martin, et al. 2016, Regelin et al. 2015). Lack of miR-155 contributes to the maturation of $CD4^+$ cells into T-helper (Th) 1 cells, whereas miR-155 overexpression promotes Th2 maturation (Banerjee et al. 2010). The miR-17∼92 cluster promotes Th1, Th2 and Th17 response and prevents regulatory T-cell differentiation (Kuo et al. 2019). The T-cell response is promoted by miR-150 (Chen et al. 2017). Additionally, miR-150 as well as miR-181 are involved in the

differentiation process of NK cells (Bezman et al. 2011, Cichocki et al. 2011). Myeloid lineage commitment at a late stage in hematopoiesis is controlled by miRNAs such as miR-223-424 and 146. MiR-223 promotes granulopoiesis in human myeloid $CD34^+$ cord blood (CB) progenitor cells investigated via knockdown and overexpression experiments using colony forming, cytokine-based differentiation culture and xenotransplantation assays (Gentner et al. 2015, Vian et al. 2014). MiR-424-20a-106a and 17-5p, on the other hand, are involved in monocytopoiesis, the activation of which involves miR-146a and miR-155 in an immune response, which are additionally involved in megakaryocytopoiesis (Rosa et al. 2007, Fontana et al. 2007, Boldin et al. 2011, O'Connell et al. 2007, Labbaye et al. 2008, Zhai et al. 2014, Romania et al. 2008) .

Particularly interesting are miRNAs regulating hematopoiesis at an early stage, because a dysregulation in primitive hematopoietic cells can enhance pathological stemness maintenance. MiRNAs involved in HSC maintenance are for example miR-125a and miR-125b, which were found to protect HSCs from apoptosis (Gerrits et al. 2012, Guo, Lu, et al. 2010, Ooi et al. 2010, Connell et al. 2010). Further down the hematopoietic hierarchy miR-520h is involved in regulating the differentiation to MPPs as shown by increased number of progenitor colonies upon miR-520h overexpression in human $CD34^+$ cells (Liao et al. 2008). Further differentiation of MPPs to either myeloid or lymphoid progenitors involves miRNAs such as the miR-23 cluster, while specifically early myeloid differentiation is attributed to miR-9 and early lymphoid lineage development to miR-181 (Kong et al. 2010, Kurkewich et al. 2017, Zhu et al. 2013). The specific role of miRNAs is accompanied by their up- and downregulation as well as their association with target genes. If a miRNA is highly expressed, its promoter is hypermethylated. If a miRNA is low expressed, its promoter is hypomethylated. A promoter is the deoxyribonucleic acid (DNA) region (~100-1000 bps) serving as platform for protein complexes to bind and initiate transcription of protein- coding or non-coding genes (Lopez-Serra and Esteller 2012, Lenhard et al. 2012, Butler and Kadonaga 2002, Haberle and Stark 2018, Juven-Gershon and Kadonaga 2010, Roy and Singer 2015, Valen and Sandelin 2011). Eukaryotic promoters are classified as core, proximal and distal promoters. The core promoter is the minimal portion (~100 bps) required for initiating transcription of the directly adjacent gene. It harbors the transcription start site (TSS) and functional sequence elements to which general or specific TFs bind, which recruit polymerase II to form the pre-initiating complex (PIC) for transcription. Canonical functional sequence elements are the TATA-box, the Initiator (Inr), which is functionally similar to TATA, the general transcription factor IIB (TFIIB), the B recognition element (BRE), the downstream core promoter element (DPE), the motif 10 element (MTE) and the downstream core elements (DCEs). TATA and Inr can co-exist and act in synergy with the DPE or BRE. Also the BRE-DPE synergy can be sufficient for core promoter function (Gershenzon and Ioshikhes 2005). Worth mentioning are the tandem repeats (TRs), which only recently have come into focus (Bessière et al. 2019, Quilez et al. 2016). TRs are unstable regions effect DNA methylation and neighboring gene expression. They are very sensitive to environmental changes by which they can contribute to evolutionary gene expression fine tuning (Vinces et al. 2009). Due to their instability TRs are prone to errors during replication, which are 100-10.000 times more frequent than point mutations (Lynch et al. 2008). Thus, TRs are an interesting hotspot for mutating events in disease development. Core promoters, which contain a TATA-box or an Inr near a well-defined TSS and when active acquire trimethylated lysine 4 of histone 3 (H3K4me3) and acetylated lysine 27 of histone 3 (H3K27ac) as well as specific TFs preferably binding close to TSSs are so-called focused (sharp) promoters. They most often regulate restricted gene

expression and thus are most common in terminally differentiated tissue specific cells. In contrast to these are so-called dispersed (broad) promoters, which mostly lack defined promoter motifs, but harbor distinct regions with elevated guanine cytosine (GC-boxes) and high density CpG dinucleotides (CpG-islands) typically common near TSSs and rare elsewhere in the genome. In these nucleotide arrangements cytosine bases are prone to methylation indicative for promoter activity and several TSSs within a broad nucleotide sequence of 50-100 bps prevail. Moreover, histone modifications H3K4me3 and H3K27ac flank transcriptionally active promoter regions for protein-coding and miRNA genes just like nucleosomes do (Barski et al. 2009). The region between the flanks is nucleosome free and co-maps with desoxyribonuclease I (Dnase I) hypersensitivity sites (The et al. 2007, Gross and Garrard 1988, Hebbes et al. 1994, Ozsolak et al. 2008). These are sites, which have lost their condense structure and are thus accessible for Dnase I but also other proteins such as TFs in both core and distal promoter regions (Keene et al. 1981, McGhee et al. 1981). Distal promoter regions can bend towards proximal promoters and thereby regulate core promoter activity. They contain the regulatory elements: enhancers, silencers and insulators (Levine and Tjian 2003). Enhancers are sequences usually found >100 kb upstream or downstream of the gene, which bind TFs, recruit orientation independently activating co-TFs and via DNA binding activate the PIC (Schoenfelder and Fraser 2019). Silencers have the opposite function to enhancers. They bind TFs, but recruit repressive co-TFs silencing the PIC. Insulators are sequences which insulate the PIC from the effect of enhancers or silencers.

The above mentioned genomic features help to identify potential miRNA promoters, which with their nature in transcription regulation point towards the physiological function of miRNAs. In hematopoiesis, particularly the miR-196 family gene deserves special attention due to its embedding in clustered *HOX* genes known for their role in self-renewal - a process that together with differentiation regulates hematopoiesis.

MiR-196 in normal hematopoiesis

The miR-196 family gene, as one of the three *HOX* gene associated miRNAs (miR-196, miR-10 and iab-4), is encoded within clustered HOX genes (Bhatlekar et al. 2018, Yekta et al. 2008, Braekeleer et al. 2014, Tanzer et al. 2005). MiR-196a-1 is located on chromosome 17 in the region between *HOXB9* and *HOXB13*, miR-196a-2 located on chromosome 12 between *HOXC9* and *HOXC10* and miR-196b on chromosome 7 in human between the highly evolutionary conserved region of *HOXA9* and *HOXA10* (Chen et al 2011). The mature microRNAs of miR-196a-1 and miR-196a-2 have the same nucleotide sequences, but compared to miR-196b one nucleotide is different (**Table 3**).

Table 3. MiR-196 family gene locations and mature sequences. Table based on information from Chen et al. (2011) and (Griffiths-Jones 2018)

miRNA	Gene location	Mature -5p sequence* (5'-3')
hsa-miR-196a-1	chr17: 48632490-48632559 [-]	UAGGUAGUUUCAUGUUGUUGGG
hsa-miR-196a-2	chr12: 53991738-53991847 [+]	UAGGUAGUUUCAUGUUGUUGGG
hsa-miR-196b	chr7: 27169480-27169563 [-]	UAGGUAGUUUCCUGUUGUUGGG

* Mature sequences were acquired from miRBase database (http://www.mirbase.org/). chr: chromosome. [+] and [-] indicate two single DNA strands, that are complementary to each other.

In addition to the remarkable genomic proximity of miR-196 genes and clustered *HOX* genes, miR-196 genes possess highly evolutionary conserved complementarity sites to the 3' UTR region of those *HOX* genes (e.g. *HOXA9*, *HOXA7*, *HOXB8, HOXC8, HOXD8*) in vertebrates, and indeed, preferentially regulate *HOX* genes in embryonic stem cells (Karagkouni et al. 2018, Agarwal et al. 2015, Tanzer et al. 2005, Yekta et al. 2008). MiR-196a was found to cleave *Hoxb8* in mouse embryos and to regulate limb development in mouse and chicken embryos via *Hoxb8 cleavage* (Mansfield et al. 2004, Yekta et al. 2004, Hornstein et al. 2005). The miR-196a homolog in drosophila *iab-4* inhibits the activity *Ubx*, a homolog to the *HOX6-8* paralogous gens in vertebrates (Ronshaugen et al. 2005). MiR-196b was found to be regulated by Hox proteins (Cdx2 and Hoxd13), thus Hox proteins acted upstream of miR-196b in murine embryonic stem (ES) cells (Fantini et al. 2015). In ES cells, HOX proteins are crucial TFs for body segmentation and eventually for the development of a complete organism (Mallo et al. 2010, Alexander et al. 2009). In hematopoietic stem cells, *HOX* genes are upregulated and known to play a key role in regulating self-renewal (Alharbi et al. 2013, Bhatlekar et al. 2018). Also miR-196a and miR-196b were found to be highly expressed in stem and progenitor cells by Petriv et al. (2010), who used microfluidic qPCR as miRNA expression quantification method. However, Connell et al. (2010) found only miR-196b and not miR-196a to be enriched in HSCs utilizing microarray assays. Thus, the genomic and transcriptional association between miR-196b and *HOX* genes in primitive hematopoietic cells was primarily the focus of further investigations. Popovic and his colleagues found a dramatic overexpression of this miR-196b in murine progenitor cells compared to more committed cells by qRT-PCR with highest expression in ST-HSCs indicating, that once miR-196b levels are down, differentiation can take place (Popovic et al. 2009). High miR-196b expressions in murine stem and progenitor cells with drastic reduction during transition from GMP to mature myeloid cells was further confirmed by Petriv et al. (2010) and a reduction in human maturing $CD34^+$ BM cells by Coskun et al. (2011). Moreover, a functional analysis by Popovic et al. (2009) revealed that upon miR-196b overexpression in murine bone marrow progenitor cells ($Lin^-/Ckit^+$), the serial colony forming potential was increased while differentiation was partially blocked *in vitro.* On the other hand, in the same year, Velu et al. (2009) published that an overexpression of miR-196b in Lin^- cells reduced the total number of colony forming units while differentiation into G-colonies was inhibited. These data are in line with the finding from Connell et al. (2010), who observed a disadvantage of miR-196b overexpressing stem and progenitor cells in BM engraftment. Taken together, expression patterns and functional *in vitro* data from Popovic et al. (2009) suggest that a failure of miR-196b reduction during myeloid differentiation may lead to a malignant cell transformation, whereas *in vitro* and *in*

vivo data from Velu et al. (2009) and Connell et al. (2010), respectively, suggest the opposite.

1.5 MiR-196 in solid cancer and leukemia

MiR-196 in solid cancer

Dysregulated miRNAs play critical roles in development and progression of various diseases, including heart failure, neuronal disease, solid and blood cancers (Roden and Lu 2016, Dwivedi et al. 2019). Additionally, miRNAs can serve as molecular biomarkers (packed into exosomes or microvesicles) not only for disease detection but also as basis for the implementation of therapeutic interventions (Hayes et al. 2014, Huang 2017). Particularly interesting are miRNAs embedded in *HOX* gene clusters, because they preferentially target neighboring *HOX*-mRNAs in embryogenesis, which, in cancers, are established transcription factors with a key role in oncogenic transformation (Hoss et al. 2014, Yekta et al. 2008). One of these *HOX* associated miRNAs is the miR-196 family gene, which is located in the intergenic region of *HOX* gene clusters and found to be either up- or downregulated in different cancers functioning as oncomir or tumor suppressor (**Table 4**). In most of the solid cancer types, miR-196a/b are overexpressed and have an oncogenic function by disturbing the normal posttranscriptional control of *HOX* and non-*HOX* targets. Cancer cells with an abnormal miR-196 family expression pattern have the capability to sustain proliferation, resist cell death and activate invasion. However, miR-196 can also have a tumor suppressive function considering different cancer types and less malignant stages, in which miR-196 family members primarily interfere with metastasis formation (summarized in **Table 4**).

Table 4. Oncogenic and tumor suppressor role of miR-196 in solid cancers.

		Expected oncogenic role		
Cancer type	MiRNA	Expression level	Assay(s)	Target gene(s)
Pancreas	miR-196a	Increased compared to normal (Zhang et al. 2008, Szafranska et al. 2008, Huang et al. 2014, Wang et al. 2009, Bloomston et al. 2007)	Migration, proliferation (Huang et al. 2014) Biomarker potential (Wang et al. 2009) Prognosis potential (Szafranska et al. 2008) Targeted target correlations, overall survival for patients (Bloomston et al. 2007)	NFKBIA(Huang et al. 2014)
	miR-196b	Increased compared to normal (Wang et al. 2017, Kanno et al. 2017)	Invasion (Wang et al. 2017, Kanno et al. 2017)	CADM1 (Wang et al. 2017)

Table continues on next page

Table 4. *Continued*

Cancer type	MiRNA	Expression level	Assay(s)	Target gene(s)
Pancreas	miR-196b	--	Biomarker potential, therapeutic target potential, proliferation, CpG methylation, overall survival for patients (Kanno et al. 2017)	--
			Apoptosis (Wang et al. 2017)	
Breast	miR-196a	Increased compared to normal (Hui et al. 2009, Hoffman et al. 2009, Jiang et al. 2018, Luthra et al. 2008)	Genome microarray, CpG methylation, cell cycle (Hoffman et al. 2009)	ANXA1 (Luthra et al. 2008)
			Proliferation, apoptosis (Luthra et al. 2008)	SPRED1 (Jiang et al. 2018)
			Metastasis, promoter ChIP-Seq, wound healing (Jiang et al. 2018)	
Esophagus	miR-196a	Increased compared to normal (Luthra et al. 2008, Luzna et al. 2011, Maru et al. 2009)	Progression (Maru et al. 2009)	KRT5, SPRR2C, S100A9 (Maru et al. 2009)
			Proliferation, apoptosis (Luthra et al. 2008)	ANXA1 (Luthra et al. 2008)
Stomach	miR-196a	Increased compared to normal (Sun et al. 2012, Pan et al. 2017)	Proliferation, cell cycle, clonogenecity *in vitro* and *in vivo* (Sun et al. 2012, Pan et al. 2017)	$p27^{kip1}$ (Sun et al. 2012)
				SMAD4 (Pan et al. 2017)
	miR-196b	Increased compared to normal (Lee et al. 2016, Tsai et al. 2010, Shen et al. 2018, Li, Wang, et al. 2016, Shao et al. 2018)	Promoter methylation (Tsai et al. 2010, Shao et al. 2018)	RAD23B (Shen et al. 2018)
			Radiosensitivity (Shen et al. 2018)	PI3K/AKT/ mTOR (Li, Wang, et al. 2016)
			Proliferation, invasion (Li, Wang, et al. 2016)	
	miR-196s	Increased compared to normal (Tsai et al. 2016)	Overall survival for patients, metastasis, biomarker (Tsai et al. 2016)	--
Colon/rectum	miR-196a	Increased compared to normal (Schimanski et al. 2009)	Expression, metastasis, migration, chemosensitivity, apoptosis, proliferation (Schimanski et al. 2009)	HOXA7, HOXB7, HOXC7, HOXD8, AKT (Schimanski et al. 2009)

Table continues on next page

Table 4. *Continued*

Cancer type	MiRNA	Expression level	Assay(s)	Target gene(s)
Colon/rectum	miR-196b	Increased compared to normal (Xu and Gu 2017, Mo et al. 2014)	Prognosis and biomarker potential (Xu and Gu 2017) Apoptosis (Mo et al. 2014)	FAS (Mo et al. 2014)
	miR-196s	Increased compared to normal (Shen et al. 2016, Ge et al. 2014)	Metastasis, prognosis, chemosensitivity (Shen et al. 2016) Overall survival of patients (Ge et al. 2014)	HOXB8 (Shen et al. 2016,
Lung	miR-196a	Increased compared to normal (Guerriero et al. 2016, Li, Yang, et al. 2016, Liu et al. 2012)	DNA methylation, proliferation, clonogenecity, migration, invasion (Liu et al. 2012) Cell growth, apoptosis, chemosensitivity (Li, Yang, et al. 2016)	FOXO1, $p27^{kip1}$, HOXA9 (Guerriero et al. 2016) HOXA5 (Liu et al. 2012)
Ovaries	miR-196a	Increased compared to normal (Yang et al. 2016)	Migration, invasion (Yang et al. 2016)	HOXA10 (Yang et al. 2016)
	miR-196b	Increased compared to an earlier stage (Chong et al. 2017)	Expression, invasion (Chong et al. 2017)	HOXA9 (Chong et al. 2017)
Osteosarcoma	miR-196a	Increased compared to normal (Shang et al. 2015)	Proliferation, apoptosis (Shang et al. 2015)	PTEN/AKT/FOXO1 pathway p21, p27, CYCLIN D1 (Shang et al. 2015)
	miR-196s	Increased compared to normal (Zhang et al. 2014)	Survival of patients (Zhang et al. 2014)	--
Glioblastoma	miR-196a	Increased compared to normal (Yang et al. 2014, Yang, Yang, et al. 2017)	Proliferation, invasion apoptosis, xenograft tumors (Yang et al. 2014, Yang, Yang, et al. 2017)	IκBα (Yang et al. 2014) ZMYND11 (Yang, Yang, et al. 2017)
	miR-196b	Increased compared to low grade glioblastoma (Ma et al. 2012)	Survival of patients, proliferation (Ma et al. 2012)	--

Table continues on next page

Table 4. *Continued*

Cancer type	MiRNA	Expression level	Assay(s)	Target gene(s)
Glioblastoma	miR-196s	Increased compared to anaplastic astrocytomas (Guan et al. 2010)	Survival of patients (Guan et al. 2010)	--
Liver	miR-196a	Increased compared to normal (Yang, Peng, et al. 2017)	Proliferation, apoptosis, clonogenecity (Yang, Peng, et al. 2017)	FOXO1 (Yang, Peng, et al. 2017)
	miR-196b	Increased compared to normal (Yu et al. 2018)	invasion, migration (Yu et al. 2018)	FOXP2 (Yu et al. 2018)
Expected tumor suppressor role				
Breast	miR-196s	--	Migration, metastasis (Li et al. 2010)	HOXC8 (Li et al. 2010)
Colon/rectum	miR-196b	Decreased in later tumor stages compared to earlier tumor stages (Stiegelbauer et al. 2017, Fantini et al. 2016) Increased compared to para-cancerous control tissue (Wang, Zhang, Zhang, et al. 2010, Fantini et al. 2016)	Prognosis potential (Stiegelbauer et al. 2017, Fantini et al. 2016) Migration, metastasis, chemosensitivity, proliferation (Stiegelbauer et al. 2017, Fantini et al. 2016)	HOXB7, GALNT5 (Stiegelbauer et al. 2017) GATA6 (Fantini et al. 2016)
Lung	miR-196b	Decreased compared to normal (Bai et al. 2017)	Metastasis, cell growth (Bai et al. 2017) Biomarker potential, tumor growth, promoter methylation, promoter histone marks (Tellez et al. 2016)	RUNX2 (Bai et al. 2017) FOS, AP-1 (Tellez et al. 2016)
Osteosarcoma	miR-196s	Decreased compared to normal (Cao et al. 2018)	Proliferation, migration (Cao et al. 2018)	HOXA9 (Cao et al. 2018)

Table continues on next page

Table 4. *Continued*

Cancer type	MiRNA	Expression level	Assay(s) and References	Target gene(s)
Melanoma	miR-196s	Decreased compared to normal (Braig et al. 2010, Mueller and Bosserhoff 2011)	Migration (Braig et al. 2010) Invasion, tumor growth (Mueller and Bosserhoff 2011) DNMT3B methylation (Micevic et al. 2016)	HOXB7 (Braig et al. 2010) HOXC8 (Mueller and Bosserhoff 2011)
Liver	miR-196b	--	Proliferation, apoptosis, chemoresistance (Rebucci et al. 2015)	IGF2BP1 (Rebucci et al. 2015)

miR-196s: microRNA-196a and microRNA-196b; '—' no information. NFKBIA: Nuclear Factor Kappa B Inhibitor Alpha, CADM1: Cell Adhesion Molecule 1, SPRED1: Sprouty Related EVH1 Domain Containing 1, ANXA1: Annexin A1, HOXC8: Homeobox C8, KRT5: Keratin 5, SPRR2C: Small Proline Rich Protein 2C, S100A9: S100 Calcium-binding protein A9, $p27^{kip1}$ (CDKN1B): Cyclin Dependent Kinase Inhibitor 1B, SMAD4: SMAD family member 4, RAD23B: Proteasome Aassociated Protein 23 homolog B, PI3K: Phosphoinositide 3-kinase, AKT: Protein kinase B, mTOR: mammalian Target Of Rapamycin, HOXA7: Homeobox A7, HOXB7: Homeobox B7, HOXC7: Homeobox C7, HOXD8: Homeobox D8, HOXA5: Homeobox A5, HOXB6: Homeobox B6, GLTP: Glycolipid Transfer Protein, GALNT5: Polypeptide N-Acetylgalactosaminyltransferase 5, GATA6: GATA Binding Protein 6, FOXO1: Forkhead Box Protein O1, HOXA9: Homeobox A9, RUNX2: Runt Related Transcription Factor 2, FOS: AP-1 Transcription Factor Subunit; AP-1: Activator Protein 1, HOXA10: Homeobox A10, p21 (CDKN1A): Cyclin Dependent Kinase Inhibitor 1A, IκBα (NFKBIA): NFKB Inhibitor Alpha, ZMYND11: Zinc Finger MYND-Type Containing 11, IGF2BP1: IGF2 mRNA Binding Protein 1, FOXP2: Forkhead Box P2; ChIP-Seq: Chromatin Immunoprecipitation combined with sequencing, CpG: Cytospine and Guanine joined by a phosphodiester bond.

MiR-196 in leukemia

Deregulation of miRNAs in leukemia is associated with cytogenetics and a small subset of miRNAs is correlated with survival (Garzon, Volinia, et al. 2008, Li et al. 2008, Marcucci et al. 2009, Wang, Li, et al. 2010, Díaz-Beyá et al. 2014, Jongen-Lavrencic et al. 2008, Schotte et al. 2008). Functionally, overexpression of miRNAs such as miR-125b, miR-22 and miR-29a can induce transformation of murine cells (Connell et al. 2010, Song et al. 2013, Han et al. 2010). Several LSC-associated miRNAs have been identified, among which for example miR-17~92 and miR-126 are able to maintain LSC in *MLL* models and in *in vivo* experiments with primary AML patient samples, respectively (Wong et al. 2010, Lechman et al. 2016). However, the role of miRNAs is not always that clear. For example, miR-196 family genes share the role of a tumor suppressor and an oncomir in leukemia as it does in solid cancers (**Table 4 and 5**). MiR-196a expression level was reported to be increased in AML and decreased in T-ALL compared to normal $CD34^+$ BM samples (Coskun et al. 2011). Among AML subtypes, miR-196a is low expressed in *MLLr* and *AML1/ETO*, but high in CN-AML and (CN-)AML with *NPM1* mutation compared to other AMLs (with *NPM1* wildtype) (Garzon, Volinia, et al. 2008, Cammarata et al. 2010, Coskun et al. 2011, Becker et al. 2010, Jongen-Lavrencic et al. 2008, Zhang et al. 2018). The role of miR-196a can be controversially discussed. On the one hand miR-196a was found to positively correlate with the proto-oncogenic *HOXA* and *HOXB* genes in CN-AML, but on the other hand enforced miR-196a levels inhibit *HOXB8* in APL (Debernardi et al. 2007, Kawasaki and Taira 2004). In addition, *ERG*, an adverse prognostic factor, was discovered as validated target gene of

miR-196a in AML and APL, but high miR-196a expression levels were associated with poor event-free survival in the *de novo* AML cohort by Zhang et al. 2018 (Coskun et al. 2011, Zhang et al. 2018). The role of the sister miRNA miR-196b can also be considered dichotomous. On the one hand miR-196b expression levels are high in the intermediate-risk AMLs: *MLLr*, CN-AML and (CN-)AML with *NPM1* mutation compared to healthy BM cells and other favorable AMLs such as *AML1/ETO* or (CN-)AML with *NPM1* wildtype (wt) (Jongen-Lavrencic et al. 2008, Cammarata et al. 2010, Becker et al. 2010, Popovic et al. 2009, Velu et al. 2014, Li et al. 2008). Additionally, miR-196b has been found to be co-expressed with *HOXA* genes and correlated with poor clinical outcome in AML (Braekeleer et al. 2014, Díaz-Beyá et al. 2014, Velu et al. 2014, Becker et al. 2010, Zhang et al. 2018). Along this line, functional studies revealed that enhanced miR-196b expression levels in the context of *MLL-AF9* in murine BM cells promote proliferation and disease progression by targeting *Cdkn1b* ($p27^{Kip1}$) (Popovic et al. 2009, Velu et al. 2014, Meyer et al. 2018). Interestingly, a combined antagomir treatment specific for miR-196b and miR-21 reduced the LIC frequency of transplanted *MLL-AF9* derived spleen cells by 30% (Velu et al. 2014). In these studies HOXA9 is considered as upstream regulator of miR-196b. This assumption was supported by Fantini's study in 2015, who found an active murine miR-196b promoter during embryogenesis with validated Hoxa9 binding sites (Fantini et al. 2015). On the other hand, miR-196b was found to be one of eight miRNAs downregulated in LSCs (de Leeuw et al. 2014). Additionally, results of primary and secondary transplantation assays from Li, Huang, et al. (2012) showed, that miR-196b initially delays leukemogenesis by targeting *Hoxa9/Meis1* and only later in the disease progression promotes leukemogenesis by targeting the tumor suppressor *Fas*. Moreover, as observed for solid tumors, miR-196b can act as discrete tumor suppressor in leukemia such as in T-ALL, CML and APL (Bhatia et al. 2011, Liu et al. 2013, Kawasaki and Taira 2004). Despite these functional studies of miR-196b in the *MLL-AF9* mouse model, its functional role in CN-AML has not been investigated till date. Similar to *MLLr* leukemias CN-AML samples have characteristically high *HOX* gene expression levels, but the molecular basis for this is unknown.

Table 5. Oncogenic and tumor suppressor role of miR-196 in leukemia subtypes.

		Expected oncogenic role		
Leukemia subtype	MiRNA	Expression level	Assay(s)	Target gene(s)
AML	miR-196b	Increased compared to ALL	qRT-PCR expression, survival of 53 AML and 32 ALL patients (Wang, Li, et al. 2010)	--
	miR-196s	Increased in AML-M5, $NPM1c^+$ and $FLT3\text{-}ITD^+$ AML compared to other AMLs	qRT-PCR expression, survival of 124 AML patients (Zhang et al. 2018)	--
CN-AML	miR-196a	--	qRT-PCR expression for gene target (HOXA and HOXB) correlation analysis (Debernardi et al. 2007)	--
	miR-196b	Increased compared to $CD34^+$ BM	Multiplex RT-PCR and survival (Díaz-Beyá et al. 2014)	--

Table continues on next page

Table 5. *Continued*

Leukemia subtype	MiRNA	Expression level	Assay(s)	Target gene(s)
MLLr-AML	miR-196b	Increased compared to non-MLL AMLs (t (15;17), inv16, t (8;21)) and normal MNC BM, but not to $CD34^+$ BM or CN-AML (Li et al. 2008, Popovic et al. 2009, Velu et al. 2014, Li, Huang, et al. 2012)	Bead-based microarray, Exiquon miRNA arrays and qRT-PCR expression detection (Li et al. 2008, Popovic et al. 2009, Velu et al. 2014, Li, Huang, et al. 2012) Co-expression analysis of miRNA with mRNA levels, overexpression, apoptosis (Li, Huang, et al. 2012) HOXA9 correlations, overexpression, CFC assay (Velu et al. 2014) Differentiation, proliferation in healthy murine BM (Popovic et al. 2009) Transplantation assays in the context of MLL-AF9 measuring survival (Velu et al. 2014, Li, Huang, et al. 2012) LIC frequency/activity and comparing the effect to drug treatment (Velu et al. 2014) Target characterization *in vitro* and *in vivo* (Meyer et al. 2018, Li, Huang, et al. 2012)	CDKN1B ($p27^{Kip1}$) (Meyer et al. 2018) FAS (Li, Huang, et al. 2012)
APL	miR-196s	--	Survival, qRT-PCR and proliferation after gain and loss of function in HL-60 (Zhang et al. 2018)	HOXA9, ANXA1, ERG
T-ALL	miR-196b	Increased compared to precursor B-ALL	qRT-PCR and microarray expression, CpG-island array, HOX-gene correlation, chemoresistance (Schotte et al. 2010)	--

Table continues on next page

Table 5. *Continued*

Expected tumor suppressor role

Leukemia subtype	MiRNA	Expression level	Assay(s)	Target gene(s)
AML	miR-196s	Increased compared to T-ALL and normal $CD34^+$ BM	qRT-PCR expression, miRNA enhancement, luciferase reporter assay in KG-1 cell line (Coskun et al. 2011)	ERG2 and ERG3
MLLr-AML	miR-196b	Increased compared to non-MLL AMLs and normal MNC BM, not $CD34^+$ BM	Bead-based assay, Exiquon miRNA arrays and qRT-PCR expression, luciferase reporter & mutagenesis assay, gain and loss of function assays combined with western blot assays on miRNA targets and primary transplantation assays (Li, Huang, et al. 2012)	HOXA9/ MEIS1
T-ALL	miR-196b	Decreased compared to normal T-cells	qPCR, target protein expression analysis after miRNA enhancement (Bhatia et al. 2011)	c-MYC
	miR-196s	Decreased compared to AML and similar to normal $CD34^+$ BM	qRT-PCR expression, miRNA enhancement, luciferase reporter assay in MOLT-4 cell line (Coskun et al. 2011)	ERG2 and ERG3
CML	miR-196b	Decreased compared to BM of healthy individuals	qRT-PCR, miRNA enhancement, target protein expression, luciferase reporter assay, promoter methylation, proliferation (Liu et al. 2013)	BCR-ABL1, HOXA9
APL	miR-196a	--	miRNA microarray, qRT-PCR, miRNA enhancement, differentiation, luciferase reporter assay (Kawasaki and Taira 2004)	HOXB8

Expression associations without functional implications

Leukemia subtype	MiRNA	Expression level	Assay(s)
AML with $NPM1^{mt}$	miR-196s	Increased compared to $NPM1^{wt}$ and other AMLs	qRT-PCR, low density array, multiplex and single RT-PCR (Coskun et al. 2011, Cammarata et al. 2010, Jongen-Lavrencic et al. 2008)
AML with $FLT3^{mt}$	miR-196s	Increased compared to other AMLs	low density array qRT-PCR (Cammarata et al. 2010)

Table continues on next page

Table 5. *Continued*

Leukemia subtype	MiRNA	Expression level	Assay(s)
AML1-ETO	miR-196a	Decreased compared to other AML subtypes	low density array qRT-PCR (Cammarata et al. 2010)
CN-AML	miR-196a	Decreased in NPM1c^{+} compared to NPM1^{wt}	miRNA microarray and single tube qRT-PCR expression (Garzon, Garofalo, et al. 2008)
	miR-196b	Decreased compared to AK-AML	miRNA microarray, single tube qRT-PCR expression for association with cytogenetics and survival (Garzon, Volinia, et al. 2008)
	miR-196s	Increased in NPM1c^{+} compared to NPM1^{wt} (Becker et al. 2010) Increased compared to AK-AML (Cammarata et al. 2010)	Microarray, expression for association with cytogenetics (Becker et al. 2010) Low-density array (qRT-PCR) (Cammarata et al. 2010)
MLLr-AML	miR-196a	Decreased compared to other AMLs	miRNA microarray (Garzon, Volinia, et al. 2008)
MLLr-ALL	miR-196b	Increased compared to precursor B-ALL, non-MLLr leukemias and normal BM	RT-PCR, bead-based microarray (Schotte et al. 2008, Popovic et al. 2009)
T-ALL	miR-196b	Increased compared to precursor B-ALL	qRT-PCR (Schotte et al. 2008)

miR-196s: microRNA-196a and microRNA-196b; '--' no information; *FAB classification; AML: Acute Myeloid Leukemia, AK: Abnormal Karyotype, CN: Cytogenetically Normal; MLLr: Mixed Lineage Leukemia rearranged, APL: Acute Promyelocytic Leukemia, T-ALL: Acute T-Lymphocytic Leukemia, CML: Chronic Myeloid Leukemia, B-ALL: Acute B-lymphoblastic Leukemia, AML1-ETO: Fusion of AML1 (RUNX1: Runt Related Transcription Factor1) and ETO (RUNX1T1: RUNX1 Translocation Partner 1), NPM1: Nucleophosmin 1, NPM1c^{+}: cytoplasmic NPM1, NPM1wt: wildtype NPM1, FLT3-ITD: Fms Related Tyrosine Kinase 3 – Internal Tandem Duplication; p27^{kip1} (CDKN1B): Cyclin Dependent Kinase Inhibitor 1B, HOX: Homeobox, HOXA9: Homeobox A9, ERG: Erythroblast Transformation Specific Related Gene, MEIS1: Meis Homeobox 1, FAS (APO1): Apoptosis Antigen 1, c-MYC: constitutively expressed MYC, BCR-ABL1: Fusion gene of BCR (Break Point Cluster Region) and ABL1 (Abelson Murine Leukemia Viral Oncogene Homolog 1A), HOXB8: Homeobox D8; qRT-PCR: quantitative Real Time Polymerase Chain Reaction; CFC: Colony Forming Cell assay, ChIP-Seq: Chromatin Immunoprecipitation combined with sequencing, BM: Bone Marrow; LIC: Leukemia Initiating Cell, CpG: Cytospine and Guanine joined by a phosphodiester bond.

1.6 Aim of this study

MiR-196b is long known to be co-expressed with *HOXA* genes, in *MLLr* and CN-AML patients (Braekeleer et al. 2014, Becker et al. 2010, Velu et al. 2014). In the *MLLr*-AML patient cohort, accounting for ~3% of all AML cases, expression levels of miR-196b particularly correlate with the expression of the proto-oncogene *HOXA9* and has been previously shown to promote leukemogenesis in a *MLLr*- driven murine mouse model by repressing the tumor suppressors *Fas* or *Cdkn1b* (Li, Huang, et al. 2012, Meyer et al. 2018, Velu et al. 2014, Popovic et al. 2009, Schoch et al. 2003). However, the contribution of miR-196b to the leukemogenesis in the CN-AML patient cohort, accounting for ~40-50% of all AML cases and typically characterized by overexpression of multiple homeobox genes including *HOXA9* and *CDX2*, is currently unknown (Walker and Marcucci 2012, Rawat et al. 2012). Therefore, the aim of this thesis was to define the function and provide mechanistic aspects of miR-196b by utilizing the *Cdx2*-driven murine AML progression model together with the analysis of primary human AML samples and AML cell lines.

The first goal was to understand the expression profile of miR-196b in normal human hematopoietic cells of different maturation stages and in AML subtypes using real time PCR and re-analyzing published datasets, respectively. The next aim was to study the transforming potential of miR-196b alone and its role in early events of leukemia initiation and progression via overexpression studies in healthy murine hematopoietic stem and progenitor cells (HSPCs) or in the context of *Cdx2*, respectively, *in vitro* and *in vivo*. To study the effect of miR-196b in particular on CN-AML associated mutations, AML cell lines were used in *in vitro* and *in vivo* assays. Assays with murine and human cells included semi-solid/liquid culture medium and short/long-term transplantation experiments. For an insight into the molecular mechanism, the miR-196b precursor transcript was identified via conventional PCR and a promoter region close by was analyzed via bioinformatical tools and luciferase assays. To deepen the molecular understanding, the relation between miR-196b transcript, mature miR-196b and its neighboring *HOXA9* target gene expression levels were investigated in healthy BM cells, CN-AML bulk and LSCs. Additionally, new miR-196b specific mRNA targets were identified on the basis of RNA-sequencing and real-time PCR data collected in *Cdx2*-induced murine HSPCs and a comparable human CN-AML cell line, respectively.

2. Materials and Methods

2.1 Materials

2.2.1 Chemicals and Reagents

Table 6. List of Chemicals and Reagents

Name	Company
5-Fluorouracil (5-FU, 50mg/ml)	Medac, Hamburg, Germany
Agarose	Sigma-Aldrich, Taufkirchen, Germany
Ammonium chloride (NH_4Cl)	Sigma-Aldrich, Taufkirchen, Germany
Ampicillin sodium salt	Sigma-Aldrich, Taufkirchen, Germany
Bovine Serum Albumin (BSA)	Serva, Heidelberg, Germany
Calcium chloride ($CaCl_2$)	AppliChem GmbH, Darmstadt, Germany
Cell Dissociation Buffer	Thermo Fisher Scientific, Waltham, Massachussetts, USA
Chloroform	Sigma-Aldrich, Taufkirchen, Germany
D-Glucose	Sigma-Aldrich, Taufkirchen, Germany
ddH2O	Ulm University, Ulm, Germany
DEPC treated water	Applied Biosystems/Ambion, Austin, TX, USA
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich, Taufkirchen, Germany
Ethylenediaminetetraacetic Acid (EDTA)	AppliChem GmbH, Darmstadt, Germany
Ethanol (99.5% purity)	Sigma-Aldrich, Taufkirchen, Germany
Ethidium Bromide	Sigma-Aldrich, Taufkirchen, Germany
Formaldehyde	Sigma-Aldrich, Taufkirchen, Germany
Formaldehyde solution (37.5%)	Carl Roth GmbH & Co. K, Karlsruhe, Germany
Gel loading dye, blue (6x)	New England BioLabs, Ipswich, MA, USA
GelRed®10000x	VWR Darmstadt, Germany

Table continues on next page

Table 6. *Continued*

Name	Company
Giemsa Azur-Eosin-Methylenblue Solution	Merck, NJ, USA
Glacial Acetic Acid	Sigma-Aldrich, Taufkirchen, Germany
Glycerol	Sigma-Aldrich, Taufkirchen, Germany
Glycine	Sigma-Aldrich, Taufkirchen, Germany
Heparin	Braun, Melsungen, Germany
HEPES	Sigma-Aldrich, Taufkirchen, Germany
Human serum	Sigma-Aldrich, Taufkirchen, Germany
Isopropanol (99.5% purity)	Sigma-Aldrich, Taufkirchen, Germany
Lipofectamine™ 2000	Invitrogen, Carlsbad, CA, USA
Luria Broth (LB) Agar	Carl Roth GmbH & Co. K, Karlsruhe, Germany
Luria Broth (LB) Medium	Carl Roth GmbH & Co. K, Karlsruhe, Germany
Magnesium Chloride ($MgCl_2$)	Sigma-Aldrich, Taufkirchen, Germany
May-Grünwald Solution	Merck, NJ, USA
Opti-MEM™ I Reduced Serum Medium	Invitrogen, Carlsbad, CA, USA
Polybrene (Hexadimethrine Bromide)	Sigma-Aldrich, Taufkirchen, Germany
Potassium chloride (KCl)	Merck, NJ, USA
Potassium hexacyanoferrate (II) trihydrate ($K_4Fe(CN)_6 \cdot 3H_2O$)	Sigma-Aldrich, Taufkirchen, Germany
Potassium hexacyanoferrate (III) $K_3Fe(CN)_6$	Sigma-Aldrich, Taufkirchen, Germany
Propidium iodide (1 mg/ml solution)	Sigma-Aldrich, Taufkirchen, Germany
Protamine sulfate	Sigma-Aldrich, Taufkirchen, Germany

Table continues on next page

Table 6. *Continued*

Name	Company
Puromycin (Puromycin Dihydrochloride)	AppliChem GmbH, Darmstadt, Germany
Quick-Load® DNA ladder (100 bp, 1 kb and 1 kb plus)	New England BioLabs, Ipswich, MA, USA
RBC lysis buffer (10x)	Biolegend, San Diego, USA
RetroNectin®	Takara/Clontech Inc., Saint Germaine-en-Laye, France
RNAse A (Kit P-2027)	Epigentek, Farmingdale, NY, USA
Sodium chloride (NaCl)	Sigma-Aldrich, Taufkirchen, Germany
Sodium hydrogen carbonate ($NaCO_3$)	AppliChem GmbH, Darmstadt, Germany
Sodium hydroxide (NaOH) pellets	Merck, NJ, USA
Sodium phosphate dibasic (Na_2HPO_4)	Sigma-Aldrich, Taufkirchen, Germany
Sodium phosphate monobasic (NaH_2PO_4)	Sigma-Aldrich, Taufkirchen, Germany
Super Optimal broth with Catabolite repression (SOC) medium for transformed cell growth	Invitrogen, Carlsbad, CA, USA
SYTOX™ Blue Dead Cell Stain (1 mM solution in dimethylsulfoxide)	Thermo Fisher Scientific Inc., Waltham, USA
TaqMan® Universal PCR Master Mix, No AmpErase	Applied Biosystems, Foster City, CA, USA
TransIT®-LT1 Transfection Reagent	Mirus Bio LLC, USA
Tris Base	Sigma-Aldrich, Taufkirchen, Germany
Trizma base (Tris)	Sigma-Aldrich, Taufkirchen, Germany
TRIzol® Reagent	Invitrogen, Carlsbad, CA, USA

Table continues on next page

Table 6. *Continued*

Name	Company
Trypan blue	Invitrogen, Carlsbad, CA, USA
Trypsin-EDTA (1X in HBS without calcium and magnesium)	PAN Biotechnologies, Aidenbach, Germany
Tween®-20	Sigma-Aldrich, Taufkirchen, Germany
UltraPure distilled DNase/ Rnase free	Thermo Fisher Scientific, Waltham, Massachussetts, USA
X-gal (5-bromo-4-chloro-3-indolyl-B-D-galactoside) in dimethylformamide	Sigma-Aldrich, Taufkirchen, Germany

2.2.2 Consumables

Table 7. List of Consumables

Name	Company
1 ml syringes (U-100 insulin syringe)	BD Biosciences, Palo Alto, CA, USA
10 cm adherent plates	Sarstedt, Nümbrecht, Germany
10 cm suspension plates	Sarstedt, Nümbrecht, Germany
14 ml Falcon round bottom tube with snap cap	BD Biosciences, Heidelberg, Germany
15 ml Falcon tube	BD Biosciences, Heidelberg Germany
3 ml syringes luer lock	BD Biosciences, Palo Alto, CA, USA
40 µm cell strainer Nylon	BD Biosciences, Palo Alto, CA, USA
50 ml tubes (polybrene conical)	BD Biosciences, Heidelberg, Germany
96-well plate (MicroAmp Fast optical 96-well reaction plate, 0.1 ml)	Applied Biosystems, Foster City, CA, USA
Cell culture pipettes (2, 5, 10 and 25 ml), sterile and disposable	Corning Inc., NY, USA

Table continues on next page

Table 7. *Continued*

Name	Company
CFC dishes (35 mm x 10 mm)	Greiner Bio-One, Kremsmünster, Austria
CryoTubes	VWR Darmstadt, Germany
Cytospin slides (Marienfeld pre-cleaned twin frosted slides)	Marienfeld, Lauda-Königshofen, Germany
Eppendorf tubes (1.5 and 2 ml)	Eppendorf, Hamburg, Germany
Fluorescence Activated Cell Sort (FACS) tubes with and without cell strainer cap	BD Biosciences, Heidelberg, Germany
Filter for 50 ml falcon (0.22 µm and 0.45 µm)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Filtercards for cytospin	Tharmac, Waldsolms, Germany
Inoculation loop	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
MicroAmo Optical Adhesive Film	Applied Biosystems, Foster City, CA, USA
Microvette CB 300 (Haparin coated plastic capillaries for blood collection)	Sarstedt, Numbrecht, Germany
Parafilm	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
PCR soft tubes (0.2 ml)	Biozym Scientific GmbH, Hess Oldendorf, Germany)
Petri dish for bacteriology	BD Biosciences, Palo Alto, CA, USA
Pipette tips – SafeSeal.Tips professional (10, 20, 200, 1000 µl)	Biozym Scientific, GmnH, Hessisch Oldendorf, Germany
Suspension cell culture flask (T25 and T75)	Sarstedt, Nümbrecht, Germany
Suspension cell culture plate (6- and 12-well, sterile, with lid)	Sarstedt, Nümbrecht, Germany
16 gauge needle	VWR Darmstadt, Germany

2.2.3 Kits

Table 8. List of Kits

Name	Use	Company
Annexin V APC/FITC Kit	Apoptosis detection method	BD Biosciences, San Jose, CA, USA
Dnase I Kit, Amplification Grade	DNA inactivating enzyme kit	Invitrogen, Carlsbad, CA, USA
Direct-zol™ Miniprep Kit	Total RNA isolation for cell lines and primary murine BM	ZymoResearch, Freiburg, Germany
Dual-Luciferase® Reporter Assay	For promoter activity detection via luciferase activity	Promega GmbH, Freiburg, Germany
Illustra GFX PCR DNA and Gel Band Purification Kit	For DNA elution from gels and clean-up of PCRs	GE Healthcare, Waukesha, WI, USA
PrimeScript RT-PCR Kit	cDNA synthesis for PCR of primary human samples	Takara, Saint Germaine-en-Laye, France
PureYield™ Plasmid Miniprep and Midiprep System Kit	For isolation of plasmid DNA from bacteria	Promega GmbH, Freiburg, Germany
RNeasy Mini Kit	Total RNA isolation for primary human samples	Qiagen, Hilden, Germany
Senescence-Detection Kit	Senescence detection in OCI-AML3 cells	Merckmillipore, Darmstadt, Germany
Taq Polymerase Kit	For conventional PCR	Invitrogen, Carlsbad, CA, USA
TaqMan MicroRNA Reverse Transcription Kit	For microRNA specific cDNA	Applied Biosystems, Foster City, CA, USA
ThermoScript™ RT-PCR System for First-Strand cDNA Synthesis	cDNA synthesis for PCR of all non-primary human samples	Invitrogen, Carlsbad, CA, USA
TruSeq RNA Library Prep Kit v2, Set A and B	Library preparation from total RNA for mRNA-Seq	Illumina, San Diego, CA, USA

2.1.4 Cells

Table 9. List of Prokaryotic Cells

Name	Company
Max Efficiency DH5 alpha (DH5α) competent *E. coli* cells	Invitrogen, Carlsbad, CA, USA
One shot™ Stbl3™ competent *E. coli* cells	Invitrogen, Carlsbad, CA, USA

Table 10. List of Mammalian Cell Lines

Name	Description
293T Phoenix Eco	Human embryonic kidney cell line, which was transformed with adenovirus E1a and carries temperature sensitive T antigen. Second-generation retrovirus producer cell line for the generation of helper-free ecotropic retrovirusus to deliver genes to dividing murine or rat cells
32D	Murine myeloid cell line. Established from long-term BM cultures of C3H/HeJ mice
Cdx2 induced leukemic cells	Established from *ex vivo* cell cultures of leukemic BM cells from mice (Rawat et al. 2004)
GP+E86	Ecotropic packaging cell line for retroviral gene transfer, derived from the mouse embryonic fibroblast cell line NIH-3T3
HEK293T	Human embryonic kidney cell line containing SV40 large T antigen
HL-60	Human AML cell line. Established from the peripheral blood of a 35-year old woman with AML (AML FAB M2) in1976. Carry amplified *MYC* gene
Kasumi-1	Human AML cell line. Established from the peripheral blood of a 7-year old Japanese boy with AML in 1989. T(8;21) translocation leading to *AML1-ETO* fusion gene and *KIT* mutation N822
Lenti X™ 293T	293T clone selected for high titer lentivirus production. Human packaging cell line, Clonetech, Mountain View, CA, USA

Table continues on next page

Table 10. *Continued*

Name	Description
OCI-AML3	Human AML cell line. Established from the peripheral blood of a 57-year-old man with AML (FAB M4) at diagnosis in 1987. *NPM1* gene mutation (type A) and the *DNMT3A* R882C mutation

Cell lines and related information were obtained from the German Collection of Microorganisms and Cell Culture (DSMZ, Braunschweig, Germany) or American Type Culture Collection (ATCC, Wesel, Germany) unless otherwise indicated. Suspension cells were maintained in culture flask of size T25 and T75 and cells were sub-cultured at regular intervals. Adherent cells were maintained in 10 cm dishes treated with polystyrene. Cells were passaged at ~80% confluency via trypsin-induced cell detachment. Incubation of all cells took place in incubators with 37°C, 5% carbon dioxide (CO_2) with >80% relative humidity.

Table 11. List of Primary Eukaryotic Cells

Name	Description	Origin
Murine 5-FU BM cells	For Gain of function studies	Murine BM harvested four days after 5-FU (150 mg/kg mouse body weight) intraveneous mouse tail injection
Healthy human Pooled Bone Marrow $CD34^+$ Cells	Bulk cells were used for miR-196b transcript identification and qRT-PCR analysis	CellSystems, St. Kathatinen, Germany
Healthy human Pooled Bone Marrow Mononuclear Cells	Bulk cells were used for miR-196b transcript identification and qRT-PCR analysis	CellSystems, St. Kathatinen, Germany
Healthy human Bone Marrow Mononuclear Cells	Sorted for $CD34^+$ subpopulations and $CD34^-$ in the scope of Susann Ihme's doctoral thesis	Lonza Cologne GmbH, Cologne, Germany
CN-AML patient cells	For miR-196b transcript identification and qRT-PCR	University Hospital, Ulm

Murine BM cells were enriched for murine stem- and progenitor cells by injecting 5-FU into the tail vein and harvesting the BM four days later. Cells were cultured at a concentration of $1x10^6$ /1 ml culture medium in incubators with 37°C, 5% carbon dioxide (CO_2) and a relative humidity of >80%. Human bone marrow cells were thawed using Lonza's protocol for 'thawing Poietics cells' (http://bio.lonza.com) or CellSystem's protocol.

Table 12. Characteristics of adult cytogenetically normal (CN)-AML patients. Patient samples were used in conventional PCR for miR-196b transcript identification and in qRT-PCRs for mRNA and miRNA level determination and correlation analysis. Table adapted from Rawat & Götze et al. 2020, Supplementary Materials, p. 20

Patient No.	Karyotype	Age	Sex	Mutations
1	46,XY	46	M	$FLT3^{wt}/NPM1c^{+}$
2	46,XY	44	M	$FLT3^{wt}/NPM1c^{+}$
3	46,XY	43	M	$FLT3^{wt}/NPM1c^{+}$
4	46,XX	55	F	$FLT3\text{-}TKD^{+}/NPM1c^{+}$
5	46,XY	48	M	$FLT3^{wt}/NPM1c^{+}$
6	46,XX	34	F	$FLT3^{wt}/NPM1c^{+}$
7	46,XX	58	F	$FLT3^{wt}/NPM1c^{+}$
8	46,XX	49	F	$FLT3^{wt}/NPM1c^{+}$
9	46,XX	51	F	$FLT3\text{-}TKD^{+}/NPM1c^{+}$
10	46,XX	34	F	$FLT3^{wt}/NPM1c^{+}$
11	46,XX	52	F	$FLT3\text{-}ITD^{+}/NPM1c^{+}$
12	46,XX	50	F	$FLT3^{wt}/NPM1c^{+}$
13	46,XX	44	F	$FLT3^{wt}/NPM1c^{+}$
14	46,XX	52	F	$FLT3\text{-}ITD^{+}/NPM1c^{+}$
15	46,XX	48	F	$FLT3\text{-}ITD^{+}/NPM1c^{+}$
16	46,XY	56	M	$FLT3\text{-}ITD^{+}/NPM1^{wt}$
17	46,XY	40	M	$FLT3\text{-}ITD^{+}/NPM1^{wt}$
18	46,XY	39	M	$FLT3\text{-}ITD^{+}/FLT3\text{-}TKD^{+}/\ NPM1^{wt}$
19	46,XX	51	F	$FLT3\text{-}ITD^{+}/NPM1^{wt}$
20	46,XY	31	M	$FLT3\text{-}ITD^{+}/NPM1^{wt}$
21	46,XX	61	F	$FLT3\text{-}ITD^{+}/NPM1^{wt}$

Table continues on next page

Table 12. *Continued*

Patient No.	Karyotype	Age	Sex	Mutations
22	46,XX	34	F	FLT3-ITD^{+}/NPM1^{wt}
23	46,XY	40	M	FLT3-ITD^{+}/NPM1^{wt}
24	46,XX	22	F	FLT3-ITD^{+}/NPM1^{wt}
25	46,XX	35	F	FLT3-ITD^{+}/NPM1^{wt}
26	46,XX	66	F	FLT3-ITD^{+}/NPM1^{wt}
27	46,XX	53	F	FLT3-ITD^{+}/NPM1^{wt}
28	46,XX	75	F	FLT3-ITD^{+}/NPM1^{wt}
29	46,XY	83	M	FLT3-ITD^{+}/NPM1^{wt}

Patient samples were diagnosed according to the FAB and the WHO classification (Harris et al. 1999, Swerdlow et al. 2008, Catovsky et al. 1991). The study was approved by the ethics committees of all participating institutions, and informed consent was obtained from all patients before they entered the study in accordance with the Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html). FLT3 mutated (mt)$^{+}$: FLT3-ITD/FLT3-TKD = fms-like tyrosine kinase receptor-3 with internal tandem duplications or mutations in the tyrosine kinase domain, NPM1c^{+}: cytoplasmic nucleophosmin 1, F = female, M = male.

Table 13. Characteristics of adult cytogenetically normal (CN)-AML patients. Patient was used for LSC identification and subsequent RNA- and miRNA-sequencing.

Patient No.	Karyotype	Age	Sex	Mutations
1	46,XY	23	M	FLT3-ITD^{+}/NPM1^{wt}
2	46,XX	52	W	FLT3-ITD^{+}/NPM1c^{+}
3	46,XX	43	W	FLT3-ITD^{+}/NPM1c^{+}
4	46,XY	70	M	FLT3^{wt}/NPM1^{wt}
5	46,XY	71	M	FLT3-ITD^{+}/NPM1^{wt}

Patient samples were diagnosed according to the FAB and the WHO classification (Harris et al. 1999, Swerdlow et al. 2008, Catovsky et al. 1991). The study was approved by the ethics committees of all participating institutions, and informed consent was obtained from all patients before they entered the study in accordance with the Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html). FLT3 mutated (mt)$^{+}$: FLT3-ITD/FLT3-TKD = fms-like tyrosine kinase receptor-3 with internal tandem duplications or mutations in the tyrosine kinase domain, NPM1c^{+}: cytoplasmic nucleophosmin 1, F = female, M = male.

2.1.5 Cell Culture Medium and Additives

Table 14. List of Cell Culture Medium and Additives

Name	Company
DMEM High Glucose (4.5 g/l)	Gibco, Thermo Fisher Scientific, Waltham, Massachusetts, USA
Dulbecco's PBS (without magnesium and calcium, sterile filtered)	Gibco, Thermo Fisher Scientific, Waltham, Massachusetts, USA
Fetal bovine serum (FBS, 0.22 µm filtered mycoplasma screened)	Gibco, Thermo Fisher Scientific, Waltham, Massachusetts, USA
Methylcellulose H4330	Stem Cell Technologies, Köln, Germany
Methylcellulose M3434	Stem Cell Technologies, Köln, Germany
Murine stem cell factor (mSCF)	Immunotools, Friesoythe, Germany
Penicillin/Streptomycin (P/S)	PAN Biotechnologies, Aidenbach, Germany
Recombinant murine interleukin 3 (rmIL3)	Immunotools, Friesoythe, Germany
Recombinant murine interleukin 6 (rmIL6)	Immunotools, Friesoythe, Germany
RPMI 1640	Gibco, Thermo Fisher Scientific, Waltham, Massachusetts, USA

Table 15. Cell Type Specific Culture Medium Constitutions

Cell Type	Culture Medium Composition
Murine HSPC (5-FU)	15% FBS in DMEM + cytokine cocktail diluted to 10 ng/ml mIL-6, 6 ng/ml mIL-3, 50 ng/ml mSCF + 1X P/S
OCI-AML3	20% FBS in RPMI + 1 X P/S
Kasumi-1	20% FBS in RPMI + 1 X P/S
HEK293T	10% FBS in DMEM + 1X P/S
Lenti X™ 293T	10% FBS in DMEM + 1X P/S
Phoenix Eco*	10% FBS in DMEM + 1X P/S

Table continues on next page

Table 15. *Continued*

Cell Type	Culture Medium Composition
32D	10% FBS in RPMI + 10 ng/ml mIL-3 + 1X P/S
Cytokine cocktail (100x)	50 mg/ml mSCF 10 mg/ml mIL6 6 mg/ml mIL3 Dissolved in DMEM

2.1.6 Enzymes

Table 16. List of Enzymes

Name	Company
DNA ligase T4	New England BioLabs (NEB), Ipswich, MA, USA
Restriction Enzymes	New England BioLabs (NEB), Ipswich, MA, USA

2.1.7 MiR-196b relevant sequences for cloning

Table 17. MiR-196b promoter, transcript variant and murine DNA fragment

Name	Sequence (5' to 3')
MiR-196b promoter	ACCTTCCACTTTATCCCGTTCACCACTTTTACAACAGGAGGACTAACCC GAGCCCCTGCAATTACTTTAGGCATCTATTTAAATATTACCTAGACGGTC GTAATTTGTCTGGGCCCTATAGCCCTGGTGCCGTAAGGTTTGTCGGCTT TTGTTCAGTTTTATGGCTTGCTAGTATATCTGGATTGTGGCTGTCTTGGA CCAGTGATTTCAGTTGAGAGGGGAGCTACATAGACAGAGGAAGCCAAA CAGGATTTCTTTCGGGAGCCCCAGGGAGGCTTGCAGAGGCCGGTTATT TTGCGGGTGTCTGGGTCACAGATGACCCTCAATCTGAATCGGGGAGCA GGCCTGGCCCACAGCGCATCTCTAGCCGCTGGAGGCAGCGGTTAGGT GGACCCGGGCTGGACTGCATGGGGGCCTCCCCCTCCCCCAGTACGGC CTCCCCCACCTGGCGGCACTCCTTAGGGCCACGCGTTTCCTGCTTGCC AGAGTGGGGGGGGCGGGGGCTAAGGGGAGGGGGCGCAGGAGCGCG CGCCCCGGCGGGCGGGCAGCTAGGAGGGAGAGGGGGAGAGCGAGAG GCGGGAGGCCGGGGCCAGACAGGGAGCCGGGATCTGTACCGTCCCT GGCGCTTTCGAGCCTTCCACGGCTACCGCCTCTTGGCGCCGGCTCGCT GGGCTGCAAGATTTGGGAAAGCGGCGCGCACACTGCTTCTCGGAGCT GCTCTCGTCTCGCCTGCTCCCCGCCTTTATTCCTCTCTCCCTGCCTTTC CTCCCTCTCTTTTCCTCCCTCTCCCTCCCAGGAATAC

Table continues on next page

Table 17. *Continued*

Name	Sequence (5’ to 3’)
MF139050	CTGGGGCACTCTGTTGCACTGGCGGGCGCAGGTTGCCTAGGGGCTGG GCTGGGCCGGGCCAGGCGCGATGGCAGGGTTCTCTCCTTGGCGGCGG CGGCAGCGGCGGAGGCGGCGGCGGCGGCGGGCGAGGCAGCCCTTC GCGGGCAGCACCAGAACTGGTCGGTGATTTAGGTAGTTTCCTGTTGTT GGGATCCACCTTTCTCTCGACAGCACGACACTGCCTTCATTACTTCAGT TGAAATCGTCTCCAGTTGATAGAGAAAAACAACCCAGCGAAGGCGCCTT CTCTGAAAACAATGCTGAGAATGAGAGCGGCGGAGACAAGCCCCCCAT CGATCCCAATAACCCAGCAGCCAACTGGCTTCATGCGCGCTCCACTCG GAAAAAGCGGTGCCCCTATACAAAACACCAGACCCTGGAACTGGAGAA AGAGTTTCTGTTCAACATGTACCTCACCAGGGACCGCAGGTACGAGGT GGCTCGACT
MF139051	CTGGGGCACTCTGTTGCACTGGCGGGCGCAGGTTGCCTAGGGGCTGG GCTGGGCCGGGCCAGGCGCGATGGCAGGGTTCTCTCCTTGGCGGCGG CGGCAGCGGCGGAGGCGGCGGCGGCGGCGGGCGAGGCAGCCCTTC GCGGGCAGCACCAGAACTGGTCGGTGATTTAGGTAGTTTCCTGTTGTT GGGATCCACCTTTCTCTCGACAGCACGACACTGCCTTCATTACTTCAGT TGAAATCGTCTCCAGATAACCCAGCAGCCAACTGGCTTCATGCGCGCT CCACTCGGAAAAAGCGGTGCCCCTATACAAAACACCAGACCCTGGAAC TGGAGAAAGAGTTTCTGTTCAACATGTACCTCACCAGGGACCGCAGGTA CGAGGTGGCTCGACT
mmu-miR-196b genomic fragment	GACTGGCTGAAAAGACATCGCGATTTGCTGTCCCCAACACTTAAATCTT GGGAGCCTGGAGCTGAGCGAAAGGAAGCCAGGAGAGCCAGGAACAGG CTCTACCTGGGCAGGGAAGCAGCCTCGGGACGCGTGGTCCCCGAGCA TCTGCCGCCTGAAGTCTGAGGACAGGACCCGGCGCGCCGGAGCTTAG TACCGCCCGGGGCCTGCCTGGCACCGTGAAACACCCAGGGCCTGGCC GAGAGCAGCTGGATCTAGGACGAGAAGAGAGCCTTGGGAAGCCCGGT GTAGCCCTGACTTCTGCCCTGGGGAGGGGGATAGGTAGGTGGCGAAA GTCTGCTT

2.1.8 Oligonucleotides

Table 18. List of Primers used for Conventional PCR and Sequencing

Name	Sequence (5’ to 3’) *	Usage
GAPDH gene	Fwd: CCGGGAAACTGTGGCGTGATGG Rev: AGGTGGAGGAGTGGGTGTCGCTGTT	Amplification
ß-ACTIN gene	Fwd: GGACTTCGAGCAAGAGATGG Rev: AGCACTGTGTTGGCGTACAG	Amplification
Human miR-196b transcript	Fwd: CTGGGGCACTCTGTTGCACT Rev: AGTCGAGCCACCTCGTACCTG	Amplification and Sequencing
Ph1 plasmid	Rev: GGATCCCAATATTTGCATGTCGC	Sequencing

Table continues on next page

Table 18. *Continued*

Name	Sequence (5' to 3') *	Usage
MIG & MIY plasmid	Fwd: CCCTTGAACCTCCTCGTTCG Rev: CCTAGGAATGCTCGTCAAG	Sequencing
LMP plasmid	Fwd: CCCTTGAACCTCCTCGTTCG	Sequencing
pGL3 basic plasmid	Rev: CTTTATGTTTTTGGCGTCTTCCA	Sequencing
pSMART hEF1a/TurboRFP plasmid	Rev: AGAGACAGCAACCAGGATTT	Sequencing

* Forward primers are defined as such, if they are amplifying towards the end of a gene. Reverse primers are defined as such, if they are amplifying towards the beginning of a gene.

2.1.9 Plasmids

Table 19. List of Plasmids

Name	Virus	Description	Company/ Reference
mmu-Cdx2-MIY	Retrovirus	The plasmid was kindly provided by Dr. V. P. S. Rawat (Ulm, Germany)	Rawat et al. (2004)
hsa-miR-196b-pSMART hEF1a/TurboRFP	Lentivirus	GSH11926-21362766 shMIMIC Human Lentiviral microRNA I-miR-196b-5p Hef1a-TurboRFP was used in comparison to the non-targeting Hef1a- TurboRFP SMART vector.	Dharmacon
LMP	Retrovirus	MSCV-LTRmiR-PIG (LMP). MSCV based retroviral vector, which expresses shRNAmir from the retroviral LTR promoter and has an EGFP sequence which serves as a marker for retroviral integration. This vector is maintained in the lab, but was originally kindly provided by Scott Lowe (New York, USA)	Dickins et al. (2005), He et al. (2005)
pGL3-basic		Luciferase Reporter Vector	Promega
pMD2.G	Lentivirus	Envelope plasmid for the production of lentiviral particles, contains VSV-G envelope protein.	Addgene

Table continues on next page

Table 19. *Continued*

Name	Virus	Description	Company/ Reference
pMIG/pMIY	Retrovirus	(MSCV-IRES-GFP/YFP) Murine stem cell virus (MSCV) derived vector. It is a bicistronic (co-expressing the gene of interest along with the reporter gene without the need of an additional promoter) retroviral vector with an internal ribosomal entry site (IRES) and an EGFP or EYFP sequence at the 3'end of the IRES. This vector is maintained in the lab, but was originally kindly provided by Prof. Dr. R. K. Humphries (Vancouver, Canada)	Pineault et al. (2002)
Ph1	Retrovirus	MDH1-PGK-GFP_2.0. Retroviral vector, which expresses miRNA from the H1 promoter and has a GFP sequence which serves as a marker for retroviral integration. This vector is maintained in the lab, but was originally kindly provided by Prof. David P. Bartel (Cambridge, Massachusetts)	Addgene Chen et al. (2004)
psPAX2	Lentivirus	2nd generation packaging vector for lentiviral particles, contains Gag, cPPT and HIV_Rev_NES.	Addgene

2.1.10 Antibodies for FACS

Table 20. List of Antibodies for FACS

Antigen	Tag	Target	Company
anti-human CD45	APC-H7	All nucleated hematopoietic cells	BD Pharmingen, Heidelberg, Germany
anti-mouse c-Kit/CD117	AlexaFluor700 or APC	Hematopoietic stem and progenitor cells	eBioscience, San Diego, USA or BD Pharmingen, Heidelberg, Germany
anti-mouse Sca-1 (Ly-6A/E)	PE-Cy7	Hematopoietic stem and progenitor cells	eBioscience, San Diego, USA
anti-mouse Gr-1 (Ly-6C/Ly-6G)	APC-Cy7 or PE	Granulocytes	Biolegend, San Diego, USA or BD Pharmingen, Heidelberg, Germany

Table continues on next page

Table 20. *Continued*

Antigen	Tag	Target	Company
anti-mouse Mac-1/CD11b	PE or APC	Myelomonocytic cells	BD Pharmingen, Heidelberg, Germany
anti-mouse B220	APC	B lymphocytes	BD Pharmingen, Heidelberg, Germany
anti-mouse CD19	PE	B-Lymphocytes	BD Pharmingen, Heidelberg, Germany
anti-mouse CD4	APC	T-lymphocytes	BD Pharmingen, Heidelberg, Germany
anti-mouse CD8a (Ly-2)	PE	T-lymphocytes	BD Pharmingen, Heidelberg, Germany
anti-mouse CD16/32	--	Purified anti-mouse AB for blocking of unspecific bindings	Biolegend, San Diego, USA

Sytox Blue from Invitrogen was used for dead cell staining

Human serum from was used for blocking of unspecific bindings

2.1.11 Mice

Parental strain mice were bred and maintained at the University Ulm in the animal facility of the Oberberghof.

Table 21. List of mouse strains and the respective irradiation intensity

Mouse strain	Abbreviation	Purpose	Irradiation in cGy
C57BL/6J × C3H/HeJ	C3HxB6	For murine cell transplantation experiments	1200
C57BL/6Ly-Pep3b x C3H/HeJ	PEP	For delta CFU-S	1200
NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ	NSG	For human cell transplantation experiments	325

2.1.12 Assays for qRT-PCR

All assays for qRT-PCR were purchased from Life Technologies and run on the TaqMan 7900HT system.

Table 22. List of Assays for qRT-PCR from Life Technologies

Name	Organism	Probe I.D.
Cdx2	Mouse	Mm01212280
HOXA7	Human	Hs00600844_m1
HOXA9	Human	Hs00365956_m1
Hprt	Mouse	Mm01545399_m1
MAPK1	Human	Hs01046830_m1
MAPK3	Human	Hs00385075_m1
MF139050*	Human	814305 A10 (E1-E2, T5, long transcript)
MF139051*	Human	814305 B1 (E1-E3, B4, short transcript)
miR-196a†	Human	Old: 495 and New: 241070_mat
miR-196b†	Human	002215
MSI2	Human	Hs01592569_m1
PBX1	Human	Hs00231228_m1
RNU19	Human	001003
RNU6b	Human	001093
RUNX1-RUNX1T1 fusion, short: AML1-ETO	Human	Hs03024752_ft
TBP	Human	Hs99999910_m1

* Primer/Probe assays were custom designed and thus had to be tested for equal efficiency. For that primer/prone assays were used on plasmid DNA harboring miR-196 gene at two different 1:10 dilution sets: set 1: 2ng to 20fg and set 2: 400pg to 400 atg. R^2 for both primer/robes was >0.996. R^2=1 means 100% efficiency.

†For gene expression analysis on miRNA level, miR-196a and miR-196b were assessed using the 'looped' real time PCR assay kit from Applied Biosystems, specifically designed to measure quantitatively the mature product (Chen et al. 2005).

2.1.13 Buffers and Solutions

Table 23. List of Buffers and Solutions

Name	Category	Composition
FACS & Sorting buffer	Buffer	For suspension cells: 2% FBS in PBS and live-dead cell stain, stored in dark at 4 °C For adherent cells: 1X DPBS + 1 mM EDTA and live dead cell stain
Hepes Buffered Saline (HBS)	Buffer	2X: 280 mM NaCl 10 mM KCl 1.5 mM Na_2HPO_4 $2H_2O$ 12 mM Dextrose (D^+-Glucose) 50 mM HEPES Adjust pH to 6.95-7.15, Filter sterilize buffer, Store 14 ml long term at -20°C. Once thawed keep at 4°C
Erythrocyte lysis buffer	Buffer	10X: 83.4 g NH_4Cl 10 g $NaHCO_3$ 2 ml EDTA (0.5M) Add ddH_2O to 1 Liter Adjust pH to 7.3, filter sterilize and store at 4°C for up to 6 month
Tris-acetate-EDTA (TAE) for running an agarose gel electrophoresis	Buffer	50X: 242 g Tris Base 57.1 ml glacial acetic acid 100 ml 0.5 M EDTA Add ddH_2O to 1 Liter Store at room temperature (RT)
Tris-buffered saline (TBS) for preparing the agarose gel	Buffer	10X: 80 g NaCl 24.2 g Tris Adjust pH to 7.4 Add ddH_2O to 1 Liter, autoclave
Freezing medium	Solution	FBS + 10% DMSO

Table continues on next page

Table 23. *Continued*

Name	Category	Composition
EDTA solution	Solution	0.5 M: 37.2 g disodium EDTA·$2H_2O$ Add ddH_2O to 200 ml Adjust to pH 8.0 Autoclave and store at RT
May-Grünwald-Giemsa	Solution	Dilute 4ml Giemsa Azur-Eosin-Methylenblue Solution in 200 ml ddH_2O
Formalin solution for fixing mice (4%)	Solution	100 ml ~ 40% formaldehyde 4 g $NaH_2PO_4 \cdot H_2O$ 6.5 g $Na_2HPO_4 \cdot 2H_2O$ Add ddH_2O to 1 Liter
BSA (2%) for blocking RetroNectin coated wells	Solution	50 ml ddH_2O in Falcon tube 1 g BSA on weighing paper Mix on shaker at 4°C Filter sterilize with 0.22 µm
ß-Gal staining	Solution	5 mM $K_4Fe(CN)_6 \cdot 3H_2O$ 5 mM $K_3Fe(CN)_6$ 2 nM Magnesium Chloride Heat mix to 50°C Add X-gal solution PBS for desired volume
Telleyesnickzky's	Solution	450 ml of 70% EtOH 25 ml glacial acetic acid 25 ml formaldehyde

2.1.14 Instruments and Labor Equipment

Table 24. List of Instruments and Labor Equipment

Name	Company
7900HT Fast real-time PCR System	Applied Biosystems, Foster City, CA, USA
BD FACS Aria III™ (cell sorter)	Becton Dickinson Biosciences, Heidelberg, Germany
BD FACS Fortessa™ LSR II	Becton Dickinson Biosciences, Heidelberg, Germany

Table continues on next page

Table 24. *Continued*

Name	Company
BD FACS Calibur Flow Cytometer	Becton Dickinson Biosciences, Heidelberg, Germany
Bioanalyzer 2100	Agilent Technologies Deutschland GmbH, Waldbronn, Germany
Centrifuges 5415R and 5810R	Eppendorf, Hamburg, Germany
Cytospin 2 Shandon Apparatus	Thermo Electron Corporation, USA
Cytospin* 4 Cytocentrifuge	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Freezers (-20 and -80°C)	Thermo Fisher Scientific Waltham, Massachusetts, USA
Hemocytometer/Neubauer improved	LO LaborOptik, Friedrichsdorf, Germany
HemaVet® 950	Drew Scientific, Waterbury, CT, USA
Illumina HiSeq 2000	Illumina, San Diego, CA, USA
Incubator for cell culture Heraeus	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Incubators for microbiology (Galaxy 170S and Innova 44 Incubator Shaker)	New Brunswick Scientific, Edison, NY, USA
Leitz Diavert Inverted Microscope	Ernst Leitz Wetzlar GmbH, Wetzlar, Germany
Liquid Nitrogen tanks	Jutta Ohst gemran-cryo, Jüchen, Germany
Magnetic Stirrer MR Hei-Standard	Heidolph Instruments GmnH & CO. KG, Schwabach, Germany
NanoDrop® 1000 Spectrophotometer	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Nikon Eclipse Ti Microscope	Nikon, Minato, Japan
Nikon Phase Contrast-2 Microscope ELWD 0-3	Nikon, Minato, Japan
peqSTAR 96 Universal Gradient Thermocycler	PEQLAB Biotechnology GmbH, Erlangen, Germany

Table continues on next page

Table 24. *Continued*

Name	Company
Pipette boy	Eppendorf, Hamburg, Germany
Refrigerators (4°C)	Liebherr, Bulle, Switzerland
4Special accuracy weighing machine CPA6235	Sartorius, Göttingen, Germany
Thermomixer	Eppendorf, Hamburg, Germany
UV gel imaging system	UVP, Cambridge, UK
Vac-Man®	Promega, Madison, WI, USA
Vacuum Eluator™	Promega, Madison, WI, USA
Vacuum Pump 2522Z-02	Welch, Fürstenfeldbruck, Germany
Vortex-Genie 2	Scientific Industries, Bohemia, NY, USA
Zeiss Axiovert 40C	Carl Zeiss MicroImaging GmbH, Göttingen, Germany

2.1.15 Software

Microsoft Office Professional Plus 2013

BD FACS Diva 8.0

Graphpad Prism 7.04

GSEA desktop application for Windows 4.1.0

AxioVision

Adobe Illustrator

2.2 Methods

2.2.1 Thawing and Freezing of Mammalian Cells

Mammalian cells were resuspended in freezing medium and stored in liquid nitrogen (LN) or -80°C freezers. Frozen cells were rapidly thawed in a 37°C water bath and transferred into a 15 ml flacon pre-filled with 9 ml DMEM or RPMI at RT. To wash off excess remnants of DMSO from the freezing medium, the falcon was inverted and centrifuged at 1500 rpm (453 rcf) for 5 min. The supernatant was discarded and the cell pellet was resuspended in the respective culture medium listed in Table 15 and placed into a culture flasks or cell culture plate in the incubator set to 37°C and 5% CO_2.
To freeze cells, up to $5x10^6$ human AML cells and up to $20x10^6$ primary murine BM cells were spun down at 1500 rpm (453 rcf) for 5 min at RT, resuspended in 1 ml freezing medium and aliquoted into CryoTubes. For short-term storage cells were stored at -80°C and for long-term storage were transferred later into LN tanks.

2.2.2 Cell Count

Cultured cells were counted manually using the hemocytometer (Neubauer chamber) by mixing 10 µl of the homogenized cell suspension with 10 µl trypan blue and placing 10 µl of the mixture into the hemocytometer. Dead cells took up the stain and appeared blue under the light microscope, whereas live cells remained bright due to an intact cell membrane. Cells from experimental sacrificed mice were counted by the HemaVet machine. BM, spleen and Peripheral blood (PB, diluted 1:5) were taken for the cell count in a total volume of 30 µl (*see 2.2.21 for details*). The HemaVet machine can distinguish between different cell types counting for example red (RBCs) and white blood cells (WBCs).

2.2.3 RNA Extraction & Purity and Concentration Determination

Extraction

Extraction of RNA was performed by using the RNeasy Mini Kit for primary human samples and the chloroform method or the Direct-zol Miniprep Kit for cell lines and primary murine BM samples. Using the chloroform method, cells in 1 ml TRIzol were incubated for 5 min at RT for cell lysis, whereas RNA was stabilized by this buffer. Then 200 µl chloroform was added, vortexed for 10 s and incubated for 2-3 min at RT. Then the mixture was centrifuged for 15 min at 13400 rcf at 4°C. Three phases were formed: (1) the organic phase - containing proteins, (2) the interphase - containing DNA and (3) the aqueous phase - containing RNA (Chomczynski and Sacchi 1987). The upper aqueous phase was transferred to a new tube and mixed with an equal volume of isopropanol by inverting the tube several times. The mixture was incubated for 10 min at RT and afterwards centrifuged for 10 min at 13400 rcf at 4°C to precipitate RNA. The RNA containing pellet was washed twice with 500 µl 75% ethanol (EtOH) by centrifugation at 7500 rcf for 5 min at 4°C and dried at RT for 15 min to remove residual EtOH. RNA was dissolved in DEPC treated or ddH_2O and heated 10 min at 55°C to completely dissolve the pellet. Using RNA extraction kits, manufacture's instructions were followed.

Purity and Concentration Determination

For measuring RNA purity and concentration, the spectrophotometer NanoDrop 1000 was used. RNA (and DNA) has its absorption maximum at 260 nm. Proteins have their absorbance near 280 nm. EDTA, carbohydrates, phenol, salts and other substances absorb light near 230 nm. The 260/280 nm ratio is used as primary and the 260/230 nm ratio as secondary measure for RNA purity. The expected value for both ratios is 2.0. If the ratio is appreciably lower than 2.0, it may indicate the presence of contaminants that absorb at 230 or 280 nm and can interfere with downstream applications. The RNA concentration was calculated by the NanoDrop software relating absorption values of the blank (H_2O) to absorption values of the sample. Then RNA was directly used for DNase I treatment and complementary single stranded deoxyribonucleic acid (cDNA) synthesis or stored at -80°C until further usage.

2.2.4 DNase I Treatment

The AML1-ETO specific qRT-PCR assay from Life Technologies required DNase treatment. For that the DNase I, amplification grade kit was used. Briefly, 1 µg RNA, 1 µl DNase I and 1 µl 10x DNase I buffer, ddH_2O to 10 µl were added into a RNase-free 1.5 ml Eppendorf tube on ice, mixed via flipping and shortly spun down, then incubated for 15 min at RT. DNase I was inactivated via 1 µl of 25 mM EDTA and an incubation time of 10 min at 65°C in a thermomixer.

2.2.5 cDNA Synthesis for RNA and miRNA

RNA was reverse transcribed into cDNA to create a more stable RNA construct being resistant to abundant RNases. For that, either the ThermoScript RT-PCR System for First-Strand cDNA Synthesis Kit was used for primary human samples or the PrimeScript RT-PCR Kit was used for all other samples according to manufacturer's instructions. For primary human samples both random hexamers and Oligo$(dT)_{20}$ were used in the cDNA reaction. For all other RNA samples only random hexamers were used in the cDNA reaction. MiRNA was reverse transcribed via the TaqMan MicroRNA Reverse Transcription Kit. Briefly, 2.5 or 5 µl of a 2 ng/µl RNA solution were combined with 5 or 10 µl of the reaction solution containing the reverse transcription buffer, dNTPs, the specific target primer, nuclease-free H_2O, RNase inhibitor and reverse transcriptase. The 0.2 ml PCR soft tube containing the reaction mixture was flicked, shortly spun down and incubated on ice for 5 min. The reverse transcription polymerase chain reaction (RT-PCR) was performed in a thermal cycler. Then, the cDNA was ready to use or alternatively stored short-term at 4°C or long-term at -20°C until further usage.

2.2.6 Quantitative Real Time (qRT)-PCR for mRNA and miRNA

For qRT-PCR of mRNA, specific Taqman assays from Life Technologies were used (*see Table 22*). The qRT-PCR reaction mix contained the following reagents: 1 µl TaqMan Gene Expression Assay (20x), 10 µl Taqman Universal PCR Master Mix (2x) and 1-2 µl cDNA template. H_2O was added to a total volume of 20 µl per reaction. The analysis was performed in duplicates or triplicates in a 96-well plate sealed with optical film using the TaqMan 7900HT machine. The settings in the SDS 2.4 software were as followed: standard curve and 40 or 45 cycles. Each program started with 10 min at 95°C and continued with

40 cycles á 15 s at 95°C and finished with 60°C for 60 s. The threshold value was usually 0.2 or, if necessary, was corrected according to the recommendations of the manufacturer's instructions. Always a H_2O and minus reverse transcriptase (RT) control sample was used to control for contaminations of the reagent and cDNA sample, respectively. An endogenous control (housekeeping gene, HSK) was used to improve reliability of the relative target expression changes. In human samples the TATA-box binding protein (*TBP*) was used as HSK and in murine samples hypoxanthine phosphoribosyltransferase 1 (*Hprt*) was used as HSK. For miRNA detection, the HSK genes were RNA U6 small nuclear 6 pseudogene (*RNU6-6P = RNU6b*) and RNA U19 small nuclear (*RNU19*). Samples in which the C_T-value of the HSK was >38 was excluded from further analysis.

Biochemistry

Gene specific Taqman probes in the qRT-PCR reaction mixture consist of a target specific nucleotide sequence with a non-fluorescent quencher on its 3' end a fluorescent reporter on its 5' end. During the standard qRT-PCR run, the quencher absorbs (quenches) the fluorescent signal from the reporter due to their close proximity (≤ 10nm) and dissipates the absorbed energy as heat. Once the Taqman probe anneals to the single strands of the target cDNA, polymerase will cleave the reporter gene off the probe, which leads to an increased distance to the quencher resulting in fluorescence signal. The fluorescence signal will accumulate over time proportional to target cDNA amplification. Thus, if the target cDNA template exists at high levels, the machine detects a fluorescence signal and transforms it into a value that states how many cycles (*C*) it took in order for the fluorescence signal to cross background signal (threshold = T), short: C_T-value.

Analysis

For qRT-PCR analysis, a fluorescence value correlating with gene expression level above threshold, the C_T-value, was compared between samples using a standard method (summarized in Table 25) in Microsoft Excel 2010. First, the delta (Δ) C_T-value was calculated by subtracting the C_T-value of the reference gene (HSK) from the C_T-value of the target gene. In this way it could be corrected for sample to sample variations such as differences in cDNA input due to different cDNA synthesis efficiencies or RNA input measurement discrepancies. Second, to compare test sample to a control sample, the ΔC_T-value of the control sample was subtracted from the ΔC_T-value of the test sample resulting in a $\Delta\,\Delta C_T$-value. In a third step, for an easier expression level comparison, the $\Delta\,\Delta C_T$-value or also the ΔC_T-value was transformed to Fold change values: $2^{-\Delta\Delta CT}$ or $2^{-\Delta CT}$, respectively (Livak and Schmittgen 2001, Schmittgen and Livak 2008).

Table 25. qRT-PCR analysis steps and their equations

Step	Calculation
0	C_T value provided by SDS 2.4
1	$\Delta C_T = C_T$ (a target gene) – C_T (a reference gene)
2	$\Delta\Delta C_T = \Delta C_T$ (a target sample) - ΔC_T (a reference sample)
3	Fold change (Control sample) = $2^{-\Delta\Delta CT}$ Fold change (HSK) = $2^{-\Delta CT}$

HSK: Housekeeping gene. qRT-PCR: quantitative real time polymerase chain reaction.

2.2.7 Conventional (Standard) PCR

Conventional PCR was used to identify the miR-196b transcript. For that, the DNA section of interest was specified and different primers were used in PCR reactions to test for an amplification.

DNA section specification

The DNA section, which was suspected to contain the base sequence for the miR-196b transcript, was determined on the basis of four criteria displayed by the USCS genome browser (**Figure 6**): (1) the DNA section was an Expressed Sequence Tag (EST) region, which (2) comprised the miR-196b stem-loop sequence, was (3) highly evolutionary conserved and (4) homologous to many other human ESTs. Therefore, the EST region 'CB127848' was found to be eligible for designing primers for the identification of miR-196b transcript(s) in the human system.

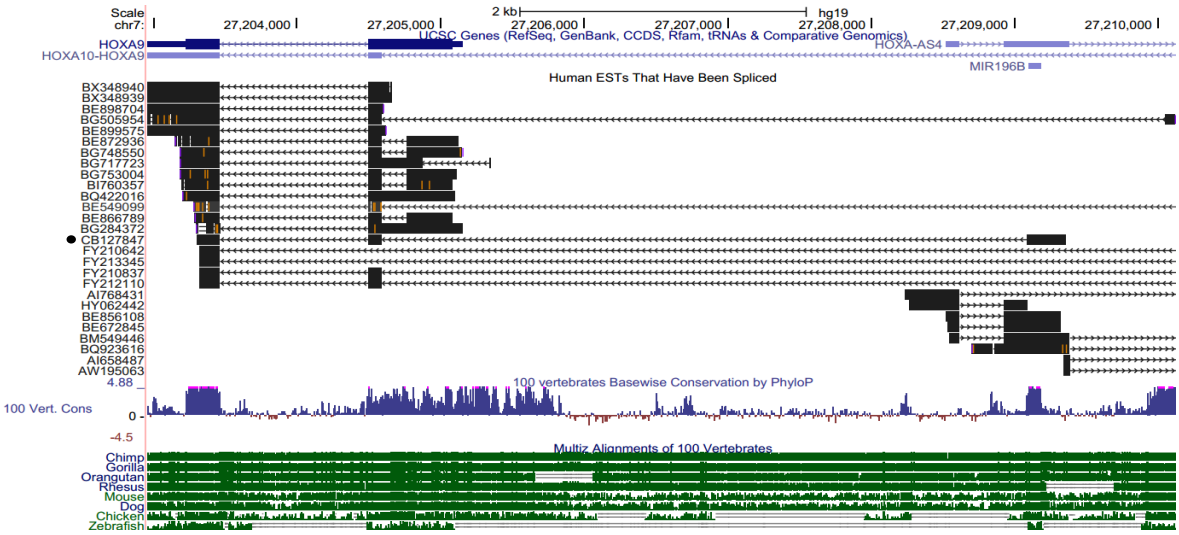


Figure 6. Visualization of criteria based on which the DNA section was specified for a miR-196b transcript identification. Depicted are homology and evolutionary conservation of 'human ESTs that have been spliced', which is a track provided by the UCSC Genome Browser (GRCh37/hg19) (Karolchik et al. 2014, Kent et al. 2002). *Legend continues on next page.*

Figure 6. *Legend continued*. The EST CB127847 encloses the stem-loop miR-196b coding region on its 5' end and an evolutionary highly conserved sequence on its 3' end, which also appears to be a homologous sequence to other ESTs (displayed by black blocks) and to Exon 2 of HOXA9. This EST is indicated by a dot to the left of its name. Red lines in the black blocks of the ESTs show a base difference between a transcript and the reference genome. Arrows indicate the direction of transcription. Blue bars indicate UCSC genes of which dark blue depicts coding transcript sequences and bright blue depicts non-coding transcript sequences. '100 Vert. Cons' track reveals sequence conservation across 100 vertebrates. Below, sequence conservation of selected individual organisms is displayed.

Primers positioning

Different primers were tested in miR-196b transcript amplification (**Figure 7**). Primers were designed using the bioinformatics tool Primer3 v 0.4. Different forward primers (FP) were designed at ~50 bp intervals upstream of the miR-196b stem-loop coding region with the initial FP being located in the stem and loop sequence region that gives rise to the hsa-miR-196b-5p, since this sequence will be in the primary transcript for sure. The reverse primer (RP) was designed to be located in exon 2 of *HOXA9*, which is a homologous sequence among known human ESTs including the CB127847 and additionally is highly evolutionary conserved. The final primer pair consist of the FP: 5'-CTGGGGCACTCTGTTGCACT-3' and RP: 5'-AGTCGAGCCACCTCGTACCTG-3' amplifying two miR-196b transcript variants.

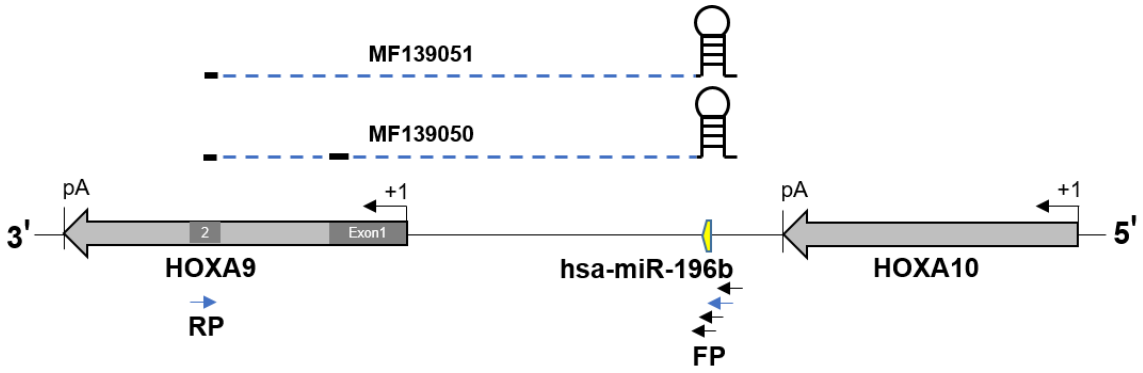


Figure 7. Schematic representation of primer positions for the identification of the miR-196b transcript. Grey block arrows represent *HOXA9* and *HOXA10* genes from their transcription start site TSS (+1) to their poly-A (pA) sequence. Dark grey regions within the gene highlight the exons of *HOXA9*, which was neglected for *HOXA10*. The miR-196b stem-loop sequence is indicated as yellow block arrow head. Thick black lines indicate the miR-196b transcripts MF139050 and MF139051. Spliced out nucleotides are indicated by the blue dotted lines. Primers used are presented by arrows. The final primer set used for the miR-196b transcript identification is resembled by blue arrows. FP: Forward Primer; RP: Reverse Primer. Figure is extracted from Rawat & Götze et al., 2020, Supplementary Materials, p. 37, CC BY-NC 4.0, https://creativecommons.org/licenses/by-nc/4.0/

PCR reaction

The TaqPolymerase Kit was used according to manufacturer's instructions. Briefly, a PCR reaction mix with a total volume of 25 µl was prepared containing PCR buffer, enhancer, $MgCl_2$, dNTP Mix (d*A*TP, d*T*TP, d*G*TP, d*C*TP), FP and RP, Taq DNA Polymerase, cDNA template and distilled DNase/RNase-free H_2O to make up 25 µl total volume. ß-Actin was used as endogeneous control. A water control instead of a cDNA template was used to ensure that reagents were not contaminated. PCR soft reaction tubes were placed in a thermal cycler an run with the following settings: (0) one cycle for the initial denaturation of

secondary structures at 95°C for 5 min, (1) polymerase activation and denaturation of double stranded cDNA at 94°C for 45 s, (2) annealing of the primer pair at 60°C for 30 s. (3) elongation at 72°C for 35 s in which Mg^{2+} ions bind to the active site of the polymerase. This enzyme has then a positive charge and can associate with the negatively charged primers and recruit dNTPs synthesizing the new strand (Pelt-verkuil et al. 2008). Step (1)-(3) were repeated 25 times. Then (4) one cycle at 72°C for 5 min terminated the PCR reaction. The PCR product was analyzed using agarose gel electrophoresis.

2.2.8 Agarose Gel Electrophoresis

Agarose gel electrophoresis was carried out for miR-196b transcript amplicons, helper plasmids or plasmids harboring an overexpression construct. The gel was prepared with 0.5-1.5% agarose in TBS buffer and 1X GelRed or 0.005% ethidium bromide, of which both substances intercalated with DNA and was made visible by UV light. The larger the DNA fragment was, the lower was the percentage of the agarose gel used for gel electrophoresis. The gel was placed into TAE buffer and samples, which were mixed with gel loading dye (bromephenol blue-based), were loaded into the gel pockets. As molecular weight size marker served 0.5 µg of a Quick-Load DNA ladder (100 bp, 1 kb or 1 kb plus), which was chosen according to the expected DNA size. For example a 100 bp ladder was used for miR-196b transcript identification and a 1 kb or 1 kb plus ladder was used for plasmids and their inserts. Electrophoretic separation was carried out by connecting the chamber containing the gel and TAE buffer to a voltage. The TAE buffer carried the current via its ions and at the same time was able to maintain the pH at a relatively constant level so that the charge of DNA, which depends on the pH, remained constant. The electrical field was applied for 60 min at 80-100 V. In this time, fragments will move from the negative pole to the positive pole of the chamber as the phosphate backbone of the DNA is negatively charged. Smaller DNA fragments will move further down the gel then larger fragments. For transcript identification, the amplicons were separately run from the housekeeping amplicons, because the size of both amplicons did not substantially differ from each other. Gel elution was performed using the Illustra GFX PCR DNA and Gel Band Purification Kit according to manufacturer's protocol and the eluate in the Elution buffer 6 was send for Sanger sequencing using the company Sequiserve. Particularly during cloning processes, as for miR-196b transcript variants into MSCV based vectors or its genomic fraction into the LMP or Ph1 vector, agarose gel electrophoresis was an important step. Also for cloning the miR-196b promoter into the promoter less pGL3-basic vector, agarose gels were used, in particular to verify cloning success, which at this point is depicted as a representative example for agarose gel pictures after electrophoresis (**Figure 8**).

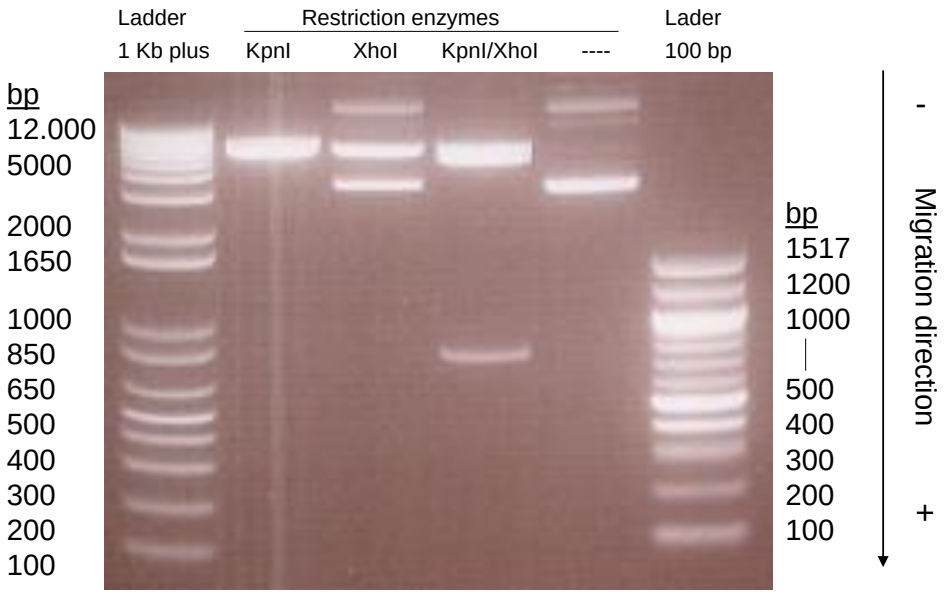


Figure 8. Agarose gel electrophoresis of the pGL3 expression vector containing the miR-196b promoter like sequence. The success of cloning the miR-196b promoter like DNA fragment into the pGL3 vector was tested by a double digestion using the restriction enzymes KpnI and XhoI, which release the promoter like gene containing 802 bps visualized in a 1% agarose gel lane 4. A single digestion with KpnI resulted in a single band in lane 2, depicting the linearized plasmid. A single digestion with XhoI resulted in three bands in lane 3, depicting the relaxed circular, linearized and superhelical DNA form in migration direction. A digestion reaction mix with water instead of enzymes lacks the linearized vector band and therewith proves activity of enzymes used during digestion. For molecular weight size estimation of the analyzed DNA fragments served two different ladders, a 1 kb plus ladder in lane 1 and a 100 bp ladder in lane 6.

2.2.9 Luciferase Assay

To prove promoter activity of the human miR-196b promoter like sequence determined by bioinformatics tools, the dual-luciferase reporter assay was performed. It is a system with which the activity of a promoter, which is essential for gene expression regulation, can be analyzed. For that, the 802 bps long promoter like fragment (see Table 17) was cloned into the pGL3-basic vector in front/upstream of the Photinus pyralis (firefly) luciferase reporter gene via the KpnI and XhoI restriction sites. The promoter orientation was the same as for the known HOXA9 expression direction. If the cloned DNA sequence had promoter activity, transcription of the luciferase gene would occur. Once the luciferase is transcribed and translated into a protein, this enzyme will show bioluminescence when it associates with the corresponding substrate. The produced light can then be measured. The amount of light measured corresponds to the activity of the analyzed promoter region.

For the transfection into eukaryotic cells, $5x10^4$ HEK293T cells were seeded into 24-well plates and DMEM supplemented with 10% FBS 16 h before transfection. Transient transfection was conducted out using Lipofectamine 2000 according to the manufacture's protocol. HEK293T cells were co-transfected with the Renilla reinformis (RL) (sea pansy) reporter vector and either the cloned promoter construct or the empty vector control. The RL construct was used as internal control for normalization (Behre et al. 1999). Firefly and Sea pansy luciferase activities were determined 24 h post transfection with the dual luciferase reporter assay system (Promega). Samples were lysed in 100 µl of passive lysis buffer (minimal coelenterazine, the natural substrate of RL, auto luminescence), and 20 µl

of lysates were used to measure activity upon luciferase specific substrate addition. Activity was assessed by determining the luciferase signal via a luminometer, whose readings were normalized against the internal control of a sea pansy luciferase and calculated as fold difference against the activity obtained with the pGL3-Basic promoter-free luciferase reporter vector.

2.2.10 Transformation of Heat-Competent *E. coli*

To amplify plasmid DNA, heat sensitive *E.coli* bacteria of type Stbl3 or DH5α were used depending on plasmid specifications according to manufacturer's recommendations (e.g. Stbl3 for MSCV and DH5α for Ph1 vector containing plasmids). The heat shock competent *E.coli* cells were produced in house and stored in aliquots at -80°C. They were thawed on ice. 50 µl of bacteria suspension were mixed gently via flipping with 1 µl of plasmid solution (1 ng-2 µg) and incubated on ice for 30 min. Then bacteria were heat-shocked at 42°C for 45 s in a thermomixer followed by 2 min on ice to enable plasmid DNA uptake. Afterwards 500 µl SOC media was added to the plasmid containing bacteria suspension and incubated for 1 h hat 37°C and 300 rpm in the thermomixer. To select for ampicillin resistant bacteria, 100-200 µl of the cell suspension were spread on an ampicillin-containing agar plate (100 µg/ml ampicillin) applying the 'spread plate method' with an inoculation loop. The plate was inverted and incubated overnight at 37°C in the dry microbiology incubator Galaxy 170S.

2.2.11 Plasmid DNA isolation

Single transformed heat competent *E.coli* colonies that have grown on an agar plate overnight were picked and put in 14 ml falcon round bottom tubes, each containing 5 ml LB medium + 100 µg/ml ampicillin. The tubes were incubated overnight at 35°C with 225 rpm in the Innova 44 shaker and a vented cap position. After successful pre-inoculation overnight, 1-4 ml of the bacteria solution were given into bacterial culture flasks containing 250 ml LB + 100 µg/ml ampicillin. Flasks were covered with an autoclaved tin foil cap and incubated overnight at 35°C with shaking at 225 rpm in the Innova 44 shaker. The next morning plasmid isolation was performed using the PureXield Plasmid MaxiPrep Kit according to manufacture's instructions utilizing the Vac-Man, Vacuum Eluator and Pump. After DNA extraction, the concentration and purity of plasmid DNA was determined using the NanoDrop spectrophotometer. Samples were sent for sequencing to LGC or Sequiserve when the construct was generated for the first time and also randomly later. In case of helper plasmids DNA was tested via characteristic restriction sites, which lead to dropouts of very specific sites visualized via agarose gels.

2.2.12 Glycerol Stock

Glycerol stocks were prepared from pre-cultures of bacteria containing the desired plasmid so that DNA could be efficiently amplified throughout the research period. 230 µl of 80% glycerol solution + 1 ml bacterial solution were added into one CryoTube, which was inverted 5x and stored at -80°C. When desired, the glycerol stock was taken out, few ice pieces were scratched off with a pipette tip, which were incubated in 250 ml LB medium + 100 µg/ml ampicillin without a time consuming pre-inoculation.

2.2.13 Transfection with Lenti or Retroviral Constructs

To produce a virus containing medium (VCM) from lentiviral plasmids, 293T LentiX cells were transfected via the calcium chloride ($CaCl_2$) method or the transfection reagent TransIT-LT1 in S2 lab conditions, because lentiviruses are able to infect humans. With the $CaCl_2$ method, LentiX cells at 50% confluency (achieved by splitting the cells at 80% confluency the day before transfection) in a 10 cm culture dish were transfected with a 1.4 ml mixture containing 15-20 µg plasmid DNA of interest, 10 µg psPAX2 helper plasmid, 5 µg pMD2.G helper plasmid, 88 µl 2M $CaCl_2$, 700 µl ddH_2O and 700 µl 2X HBS buffer with pH 7.15. The HBS buffer was bubbled with a 2 ml pipette while the DNA mixture was added dropwise. After a 30 min incubation at RT, the mixture was added dropwise to the LentiX culture plate containing 9 ml fresh culture medium, which was then placed into the incubator for 16-18 h. During the research period the $CaCl_2$ method was replaced by the transfection reagent TransIT-LT1. Briefly, 6 µg of each helper plasmid and of the plasmid containing the construct of interest were added to 1.5 ml Opti-MEM (pre-warmed to RT) and mixed completely via gentle up and down pipetting. Afterwards, 54 µl TransIT-LT1 reagent was added dropwise to the DNA mixture and mixed again via gentle pipetting using a 3 ml pipette. After 25 min incubation at RT, the mixture was added dropwise to the LentiX culture plate at ≥ 80% confluency. No fresh culture medium change was required before. To evenly spread the transfection reagent, the plate was moved in the shape of an eight along the work surface of the cell culture hood. After 16-18 h (overnight incubation) the medium was changed to 3 ml DMEM with 30% FBS and 1X P/S. After another 24 h and 48 h in the incubator, VCM was collected and filtered using a 3 ml syringe with a 0.45 µm filter to remove any non-viral particles. The VCM was either used directly for transduction or stored in 2 ml CryoTubes at -80°C.

To produce VCM from retroviral plasmids, GP+E86 or PhoenixEco cells were used (*for E86 method see the thesis of Edmaier (2013)*). PhoenixEco cells were transfected using the $CaCl_2$ method or the transfection reagent TransIT-LT1 as described for lentiviral VCM production. Different to the lentiviral VCM production was that the amount of the plasmid of interest was always 15 µg and 45 µl TransIT-LT1 were used, while helper plasmids as well as the high safety precautions could be omitted.

2.2.14 Transduction with Lenti- or Retrovirus

For a stable lentivirus-based integration of miR-196b into human AML cell lines, 1-2 $x10^6$ cells / 500-700 µl culture medium were mixed with 8 µg/ml polybrene and 500-700 µl VCM in a 15 ml Eppendorf tube. Cells were spin infected with 2500 rpm (1258 rcf) at 35°C for 60 min. The cell pellet was resuspended and cells cultured overnight with additional 1-2 ml fresh medium plus 8 µg/ml polybrene. The next day, cells were fed with additional fresh medium and placed in bigger culture dishes if required. 48 h later transduction was stopped by washing the cells with 1X DPBS. The cells were then cultured for another 48 h with puromycin for a chemically-based selection of successfully transduced cells or without puromycin for a subsequent FACS-based selection of successfully transduced cells. For murine *Cdx2* transformed *ex vivo* cells, 3 ml VCM / well of a 6-well suspension culture plate coated with RetroNectin and sealed with parafilm was spun down with 2500 rpm (1258 rcf) for 45 min at 4°C. The supernatant was taken off and $1x10^6$ cells / 2 ml culture medium were added. The next day, 1 ml fresh culture medium was added. Another day later, transduction was terminated by dissociating the cells from RetroNectin using the enzymatic-free dissociation buffer. Cells were used in an apoptosis assay 4 days after end of transduction.

For a stable retroviral overexpression in adherent cells, cells were incubated for 24 h with 2-3 ml VCM / well of a 6- well culture plate, while after 4-6 h 2-3 ml fresh culture medium was added. The next day, cells were plated in fresh culture medium for another 24 h and were FACS-purified afterwards. Suspension cells were spin infected, either 3x on subsequent days or 1x in combination with RetroNectin. Murine HSPCs were generated via a 4-day 5-FU incubation of the donor BM *in vivo*. Plates for transduction were prepared by coating the wells with RetroNectin according to manufacturer's instructions for 4 h at RT or overnight at 4°C with parafilm sealed around the plate. RetroNectin was taken off and stored at 4°C for second and third usage. RetroNectin coated wells were stored for 1-2 weeks at 4°C with 2 ml 1X DPBS on top. At the day of transduction, wells were blocked with 2% sterile BSA solution at RT for 30 min and washed 2x with 1X DPBS. 2-4 ml VCM were loaded onto the wells, then the plate was sealed with parafilm and centrifuged at 2500 rpm (1258 rcf) for 45 min at 4°C to concentrate the virus on the bottom of the plate. In parallel, for miR-196b/*Cdx2* co-transduction, 0.5-1x10^6 BM cells resuspended in 500 µl culture medium were spin infected with an equal amount of VCM + 5 µg/ml protamine sulfate with 1500 rpm (453 rcf) at 35°C for 60 min. Supernatant was decanted and BM cells were layered on top of the virus bound wells in 1-2 ml fresh culture medium and protamine sulfate (5 µg/ml). The next day, 1-2 ml fresh culture medium was added if needed. After 72 h, transduction was terminated by harvesting all BM cells using the enzyme-free dissociation buffer. One day later, transduced cells were selected via FACS and used for *in vitro* and *in vivo* assays.

Transduction efficiency was always reasonable high ranging from 4-7% for *Cdx2*/miR-196b transduction and from 20-60% for single gene transduction, but was maximally 60% to prevent too many integration off-targets, which potentially could lead to uncontrolled second hits relevant in leukemia initiation.

2.2.15 Fluorescence-Activated Cell Sort (FACS) and Immunophenotyping

Sorting

Successfully transduced cells were resuspended in sorting buffer and sorted out of the cell suspension via FACS Aria. The correct cell population was chosen according to size and granularity determined by forward and sideward scatter, respectively. Cells expressing the fluorescent reporter protein(s) were identified and sorted into collection buffer, which was the respective cell culture medium without cytokines. After the cells had recovered from sort stress in the incubator, cells were used for *in vitro* and/or *in vivo* assays. In case of murine BM cell transplantation assays, sorted cells were directly transplanted to avoid any additional artificial cell culture effect, which for established AML cell lines was negligibly small.

Immunophenotyping

Fluorescence based analyses were performed via the FACS Fortessa and its FACS Diva software. Antibodies and their conjugated fluorophores used are listed in Table 20. Single color stainings served as controls for the spectral overlap of fluorophores.

For a cell differentiation stage analysis of murine BM cells *in vitro*, cells from colony forming cell assays (CFCs) were stained for Gr1, Mac1, c-kit, Sca-1 fluorescence-conjugated antibodies. First, cells at the preferred concentration of 2x10^5 cells / 200 µl 1X DPBS were

incubated with the non-specific blocking agent CD16/32 (Fc block) in a FACS tube for 10 min at 4 °C. Cells were then stained with the antibodies for 20 min at 4°C in the dark. Then 3 ml 1X DPBS were added to the tubes and spun down with 1500 rpm (453 rcf) at 4°C to wash off excess antibodies. After decanting the supernatant, cells were resuspended in 1X DPBS containing the dead cell stain Sytox.

For investigating the lineage cell distribution *in vivo*, PB of mice transplanted with $1x10^6$ miR-196b transduced murine HSPCs and an equal amount of helper cells was gained from the tail vein 40 weeks post transplantation and stained for Gr1, Mac1, CD19, B220, CD4 and CD8 together with the live dead staining PI. This *in vivo* method for investigating the effect of an engineered gene overexpression on lineage cell distribution had also been used by David Bartle and Margaret A. Goodell in this research field (Chen et al. 2004, Challen et al. 2009).

For investigating *in vivo* growth of miR-196b transduced OCI-AML3 cells, genetically engineered human cells were injected into NSG mice, whose BM was later analyzed via fluorescent conjugated cell markers. After harvesting BM cells, the Fc-receptor was blocked by an incubation with 10% human serum (1:10 dilution in FACS buffer) for 20 min at 4°C to avoid false positives. Cells were then stained with the CD45 antibody for 10 min in the dark, washed and resuspended in 1X DPBS containing the dead cell stain Sytox.

2.2.16 RNA sequencing (RNA-Seq)

RNA was extracted from three biological replicates of murine HSPCs transduced with the empty LMP vector, miR-196b-LMP, *Cdx2*-MIY or *Cdx2*/miR-196b 48 h post sort using the Direct-zol Miniprep Kit following the manufacturer's instructions. Samples were tested for RNA quality using the Bioanalyzer. The library of samples with RIN value >9.5 was prepared using the TruSeq RNA Sample Preparation Kit version 2. Libraries were paired-end sequenced at the Genomics Core Facility of the University of Ulm using an Illumina HiSeq 2000 platform. RNA-Seq data were analyzed using the online analysis service provider Basepair (https://www.basepairtech.com/) with a pipeline including the following steps: (1) Reads were mapped to the murine transcriptome derived from the UCSC genome browser assembly mm10 using 'STAR' with default parameters, enabling an alignment of reads to only one location in the genome (Dobin et al. 2013). (2) Read counts were determined for each transcript using 'featureCounts' (Liao et al. 2014). (3) 'DeSeq2' was used as model to correct for differences in library size and to generate normalized count values (Love et al. 2014). Normalized count values were used to determine differentially expressed genes (DEGs) with cut off parameters of read count >10 and p-adjusted values (FDR-value, which is the p-value corrected for multiple hypothesis testing) <0.1 and <0.3. RNA-Seq data were visualized by heatmaps of hierarchical clustered genes generated by the online software GenePattern (https://www.genepattern.org/). To determine statistically significant enriched gene sets, the Enrichr (https://maayanlab.cloud/Enrichr/) platform or GSEA software provided from the Broad Institute (https://www.gsea-msigdb.org/gsea/index.jsp) was used (Chen et al. 2013, Kuleshov et al. 2016, Subramanian et al. 2005, Mootha et al. 2003)

2.2.17 Proliferation Assay

To study the proliferative potential of murine BM and human AML cells *in vitro*, proliferation assays were performed. Always an equal amount of successfully transduced control and test cells was plated in liquid culture medium. For the investigation of murine BM cells, 3-

$5x10^4$ / 0.5-1 ml cells were plated on day 0 and counted after 7 and 14 days. Within the 7 days cytokines were added after 2-3 days. For a day 14 cell count, cells were splitted 1:10 on day 7 and cultured further with fresh culture medium with the addition of fresh cytokines after 2-3 days. Cell counts on day 14 were adjusted as if all cells would have been cultured for 14 days. Since starting cell numbers varied due to different availabilities post sort, either percentage values were used for a concise analysis of biological replicates or cell numbers were recalculated. For the investigation of human AML cells, FACS purified and/or puromycin treated (1.5 µg for Kasumi-1 and -2.5 µg for OCI-AML3 for 48 h) cells were plated into liquid culture medium after 6-24 h recovery time from the selection procedure. Starting cell numbers were $2.5\text{-}8x10^4$ / 1.5-3 ml. for Kasumi-1 and $5x10^4$ / 3 ml for OCI-AML3 cells. Kasumi-1 cells were counted every 2 days for a total of 6 days. OCI-AML3 cells were counted on day 3 and day 6. Counted were only viable cells via trypan blue exclusion on the hematocytometer.

2.2.18 Colony-Forming Cell (CFC) Assay

To check the clonogenic and differentiation potential of murine BM and human AML cells *in vitro*, CFC assays were performed (Frisch and Calvi 2014). Through semi-solid medium supplemented with specific growth factors, committed progenitors were immobilized so that cell growth per cell became visible. An accumulation of ≥ 20 cells was considered a colony. For a concise analysis of biological replicates, results from all CFC experiments with genuine results in the control condition were combined.

Murine BM cells were plated directly after sorting into methylcellulose supplemented with cytokines (M3434) according to manufactures instructions using a 16 g needle on a 3 ml syringe. Always an equal amount of successfully transduced test and control cells was plated in duplicates into 6-well suspension culture plates. For investigating the effect of miR-196b overexpression alone, 300-500 cells per dish were initially plated. For investigating the effect of miR-196b in the context of *Cdx2* expression, 150-300 cells per dish were initially plated and 1000-2500 cells per dish were re-plated. Into the interspaces of a 6-well culture plate sterile ddH_2O was filled or when using a self-contained set up with single CFC-dishes, an extra CFC-dish without a lid was filled with ddH_2O. Colonies were scored with the help of a grid placed underneath the culture dish on day 7 to 9 after plating. Scored were CFU-GM (granulocyte/macrophage), CFU-M and CFU-G colonies for miR-196b only overexpressing cells, whereas lose, dense and burst-like colonies were scored for *Cdx2*/miR-196b overexpressing cells.

For the setup of CFC assays with human AML cells, FACS purified and/or puromycin treated (1.5 µg for Kasumi-1 and 2.5 µg for OCI-AML3 for 48 h) cells were plated into the methylcellulose H4330 after a 6-24 h recovery time from the selection procedure. To investigate the effect of miR-196b in the control cell line Kasumi-1, 2000 cells per dish were plated. To investigate the effect of miR-196b in OCI-AML3 cells, 500 cells per dish were plated. Colonies were counted 12 to 13 days after plating and scored as lose or dense (less or highly proliferative) colonies.

2.2.19 Apoptosis, Senescence and Cell Cycle Assay

To investigate possible reasons for a miR-196b induced proliferative and clonogenic disadvantage of AML cells, apoptosis, cell cycle and senescence assays were performed.

Apoptosis

Apoptosis was determined with the Annexin V Apopotsis Detection Kit I according to manufacturer's instructions. Briefly, cells were washed with cold 1X DPBS. Cells were then stained with Annexin V and 7-Aminoactinomycin (7AAD) in binding buffer for 15 min at RT in the dark. Stained cells were analyzed using the FACS Fortessa and its FACS Diva software. During this staining procedure Annexin V bound to phosphatidylserine (PS) present on the cell surface of early apoptotic cells. PS on the cell's surface is needed to get recognized by phagocytes. 7-AAD stained DNA in apoptotic cells, which had lost their plasma membrane integrity allowing 7-AAD to pass the cells membrane and stain DNA (Brauchle et al. 2014). Apoptosis in OCI-AML3 cells was determined directly after the end of the 6-day long proliferation assay (1.5 weeks after end of transduction). Cells from a murine *Cdx2* induced leukemic cell line, which was created in house via an *ex vivo* cytokine deprived long term culture, were directly examined 96 h after end of lentiviral miR-196b or control vector transduction. Test and control samples had been transduced with comparable efficiencies (Transduction efficiency of Ctrl-RFP: median=56.2%, 95% CI [45.8, 65.9] vs miR-196b-RFP: median=57%, 95% CI [54, 66.9]; n=3). Thus, a FACS-based purification could be avoided for those sort stress susceptible murine leukemic cells.

Cell Cycle

Cell cycle assays were performed by measuring the DNA content via PI staining. PI is a DNA binding dye that binds proportional to the DNA amount present in the cell varying according to the cell cycle phase: G0, G1, S, G2, M. Cells in G0 phase are non-cycling cells. In G1 phase, cells grow and have 2N nuclear DNA content. In S (synthesis) phase, DNA content doubles to 4N. In G2 phase, cells grow and maintain the DNA content of 4N. In M (mitosis) phase, cells divide resulting in two daughter cells with 2N nuclear DNA content. Thus, cells with more DNA content will fluoresce brighter than cells with less DNA content resulting in a detection peak further to the right on the frequency histogram generated by FACS Diva software of the FACS Fortessa machine. Whether miR-196b affects OCI-AML3 cell cycling was determined in those cells 1-2.5 weeks post end of transduction using the protocol 'Flow cytometric analysis of cell cycle with PI' from the company Abcam. An equal number of test and control cells was synchronized by culturing $3x10^5$ cells in 3 ml culture medium with FBS reduced to 1% for 16 h. Afterwards, cells were cultured for another 48 h in normal culture medium with 20% FBS. Then, $3\text{-}5x10^5$ cells were fixed with 1 ml ice cold 70% EtOH for 45 min on ice, washed 2x with 1X DPBS, treated with 50 µl Ribonuclease A of a 100 µg/ml stock to stain only DNA and no RNA with 200 µl PI from a 50 µg/ml stock. Fixation destroyed the RFP signal so that there was no signal interference with PI and cells could be regularly analyzed for their allocation to the cell cycle phases: G0/G1, S and G2/M phases on the FACS Fortessa.

Senescence Assay

To determine senescence, which is an arrested cell state in which cells do not divide upon a stimulus (e.g. serum) anymore, β-galactosidase, an enzyme associating with this cell state, was determined by adding the artificial substrate X-gal, which turns blue, when cleaved by ß-galactosidase (Itahana et al. 2007).

OCI-AML3 cells, which were cultured for two weeks after end of transduction, were seeded onto already prepared RetroNectin coated wells of a 12-well suspension plate with always

an equal amount of test and control cells (0.3-1x10^6 / 2 ml). Within the next 48 h, cells had settled down and were washed 1x with 1 ml 1X DPBS and fixed with 500 µl fixative solution of the Senescence-Detection Kit from Merckmillipore for 15 min at RT. Then, cells were washed 2x with 1 ml 1X DPBS and stained with a ß-galactosidase staining solution made in house (*see Table 22*) over night at 37°C in an incubator. Cells were observed for the development of blue color using the Nikon Phase Contrast-2 Microscope. Images were captured by the digital camera attached to the microscope. As positive control for the ß-galactosidase staining method served irradiated cells from an AML cell line. For that, 1x10^6 cells in 10 ml culture medium in a 15 ml flacon were irradiated with 3 Gy. Afterwards 5x10^5 cells / 3ml per well of a 12-well plate were cultured for 72 h.

2.2.20 Delta Colony Forming Unit on the Spleen (ΔCFU-S) assay

To investigate the role of miR-196b in BM cell transformation, the ΔCFU-S assay was performed. 5-FU mobilized primary murine BM cells from PEP donor mice were transduced with the empty LMP vector, miR-196b-LMP, *Cdx2*-MIY or *Cdx2*/miR-196b using E86 cells and cultured for7 days in murine BM culture medium. Then the day 0 equivalent of 1x10^4 cells without helper cells was injected per mouse into the tail vein of lethally irradiated C3HxB6 recipient mice. 3-6 mice were injected per experimental condition. The recovery of CFU-S cells was quantified by determining the number of macroscopic colonies on the spleen 12 days post injection after fixation in Telleyesnickzky's solution (*see Table 22*).

2.2.21 BM transplantation assays and mouse assessments

To study the effect of miR-196b overexpression *in vivo*, transduced murine or human hematopoietic cells were transplanted into mice and analyzed for leukemic features since this is a commonly accepted quick and efficient technique (Nguyen et al. 2016). Experimental mice were bred, housed and maintained at the University of Ulm, Oberberghof animal facility. Mice in individually vented cage systems were fed with autoclaved chow and supplied with drinking water, which contained the antibiotic ciproflaxin on the day of the experiment. The animal studies were approved by the Ethics Committee of the University of Ulm and abided by the tenets of the revised World Medical Association Declaration of Helsinki (http://www.wma.net/e/policy/b3.htm).

Murine BM transplantation

For the investigation of murine HSPCs in long term transplantation experiments, BM cells of C3HxB6 donor mice were enriched for HSPCs via 5-FU injection. Injections were performed into the lateral mouse tail vein using 1 ml syringes, which are gentle to the mouse and no sharing of cells is expected. Four days after 5-FU injection, BM cells were harvested via leg and hip crush. After a mild RBC lysis, cells were cultured at a concentration of 1x10^6 cells / ml in the respective culture medium (*see Table 15*). BM cells were retrovirally transduced with miR-196b, *Cdx2* or both. Reporter gene positive cells were highly FACS purified and injected into 8-16 weeks old lethally (1200 cGy) irradiated C3HxB6 mice. Injected were 1x10^6 miR-196b or LMP-empty vector positive cells together with 1x10^6 helper cells per mouse. Also injected were 1 or 1.5x10^5 *Cdx2* or *Cdx2*/miR-196b positive cells together with 3 or 4.5x10^5 helper cells per mouse, respectively. On one hand helper cells were used to carry test cells and on the other hand to provide short-term radioprotection. Mice were monitored for signs of sickness in particular for leukemic symptoms including

crouching, frizzled body hair, paleness in the feet and heavy breathing. Moribund mice were sacrificed and analyzed.

Human BM transplantation

For the investigation of human AML cell growth *in vivo* after miR-196b transduction, lentivirally transduced OCI-AML3 cells were FACS purified and injected into 12-15 weeks old sublethally (325 cGy) irradiated and intraperitonally immunoglobulin (IvIG) treated NSG mice. Injected were $0.5x10^6$ cells per mouse.

Mouse sacrifice

Mice transplanted with OCI-AML3 cells were sacrificed 4 weeks post transplantation, since this was the time point when control mice showed signs of sickness. Mice transplanted with murine BM cells were followed up to 400 days post transplantation and were sacrificed once diseased within this time. Mice were sacrificed by rendering them into unconsciousness with CO_2 followed by a cervical dislocation. BM, spleen and PB were taken for analysis. BM was harvested by crushing all hip and/or leg bones. BM cells were spun down at 1500 rpm (453 rcf) for 5 min at 4°C and resupsended in 10 ml 1X DPBS. The spleen was removed and weighed on a special accuracy weighing machine. A single cell suspension was prepared by pressing the organ with the back of a 3 ml syringe plunger through a 40 µm cell strainer under addition of 1X DPBS. Spleen cells were spun down at 1500 rpm (453 rcf) for 5 min at 4 °C and resuspended in 1.5-3 ml 1X DPBS. PB was drawn from the heart directly after sacrifice with a 1 ml syringe prepared with EDTA to prevent coagulation. For mice which were planned to be sent for histopathology, the whole mouse was preserved in 4% formalin as well as half of the spleen and one femur in a separate tube. Red blood cell lysis was performed for PB, spleen and BM when analyzed on the FACS by incubating the cells in ammonium chloride buffer or RBC lysis buffer from Biolegend.

Mouse analysis

BM of mice transplanted with genetically engineered OCI-AML3 cells was analyzed to determine the effect of miR-196b on *in vivo* cell growth. For that the fluorescent signal of the red fluorescent reporter protein (RFP) and the CD45 human cell marker were measured on the FACS Fortessa.

BM, spleen and PB of mice transplanted with genetically engineered murine BM cells were analyzed to determine, whether a mouse diseased from leukemia. For that RBC and WBC counts were taken by the HemaVet, BM and spleen cells were stained with May-Grünwald-Giemsa solution and evaluated on microscope slides, the fluorescent signal of the reporter protein yellow (YFP) or/and green fluorescent protein (GFP) of the integrated vector in BM samples was determined to assess *in vivo* growth and a histopathological analysis of selected organs preserved by 4% formalin was performed at the core facility for mouse pathology, University of Tübingen, Tübingen, Germany under the supervision of Prof. Leticia Quintanilla-Martinez de Fend.

2.2.22 Cyto-Morphology

Cyto-morphological analysis of murine BM and spleen cells or human AML cells was done by transferring $1\text{-}1.5x10^5$ murine or $8x10^4$ human cells on a glass slide. Cells in 250-300 µl

PBS were centrifuged at 350 rpm for 5 min using the cytospin apparatus (funnel) and the cytospin centrifuge. Air dried slides were stained with the May-Grünwald-Giemsa method. Briefly, slides in a slide holder were placed in a bucket filled with May-Grünwald reagent for 3 min. Then slides were placed in a bucket filled with tap water for 5 min for washing. Afterwards, slides were stained for 1 h in a 1:50 diluted Giemsa solution. Then, slides were washed two times á 5 min in tap water and were air dried afterwards. Blast counts were performed using the Nikon Eclipse Ti Microscope with a 600x original magnification. Images were captured by the digital camera attached to the microscope.

2.2.23 Exploitation of data generated outside this study

TCGA data

For miR-196b and *HOXA* gene expression level evaluations in AML patients, public TCGA data were downloaded, processed by the lab's bioinformatician Prof. Dr. M. Mulaw providing RPMI and FPKM values, respectively (Cancer Genome Atlas Research et al. 2013).

Microarray data from John Dick's group

For *HOXA9* expression level evaluations in functionally validated LSCs of AML patients and healthy CB cells, publically available datasets: GSE76009 and GSE42414, respectively, were used (Ng et al. 2016, Laurenti et al. 2013). In the GSE76009 dataset human LSCs were defined by phenotype (CD34/CD38) and *in vivo* engraftment. In both datasets samples were scanned on a BeadArray™ Reader using BeadScan software (v3.2). For this study, data were analyzed using the GEO2R software to get microarray expression values (https://www.ncbi.nlm.nih.gov/geo/geo2r/). The expression value in healthy CB cells of one person was generated by calculating the mean of three different progenitor cell types (MLP, GMP, MPP). Since *HOXA9* expression values of two different datasets were compared with each other, *HOXA9* values were presented relative to *TBP* expression values.

MiRNA- and mRNA-Seq data from Dr. Susann Ihme

In this study, miR-196b and its potential target gene expression levels in LSCs of CN-AML patients were determined by reanalyzing miRNA- and mRNA-Seq data originating from Dr. S. Ihme (Ihme 2014). In short, LSCs were identified by immunophenotyping and engraftment check. Immunophenotyping of LSCs and $CD34^-$ cells from CN-AML samples as well as corresponding healthy hematopoietic BM subpopulations was done according to Goardon et al. (2011). Lymphoid-primed multi-potential progenitor (LMPP): $CD34^+CD38^-$ $CD45RA^+CD90^-$ and granulocyte-macrophage progenitor (GMP): $CD34^+CD38^+$ $CD45RA^+$ $CD123^+CD110^-$. To prove the leukemic potential of LSCs, stained cells were sorted and transplanted into sublethally irradiated NSG mice. Engraftment (positive, if reporter gene signal >1%) and a leukemic phenotype of the mice were analyzed. Stained and sorted LSCs as well as their corresponding healthy counterpart were also used for sequencing. For miRNA-Seq of healthy BM subpopulations samples were combined into 2 pools, each pool harboring cells from 3 individuals, in order to achieve sufficient input material. For all other sequencing experiments such combination of material was not necessary.

For all miRNA-Seq experiments, RNA was isolated using the miRNease Mini Kit and sequencing libraries were prepared using Illumina TruSeq (small) RNA preparation kit. Technical replicates were sequenced as single end reads to increase sensitivity for low

expressed miRNAs on the Illumina HiSeq 2000. Sequence length was 50 bps with a depth of 40 million reads. Expression analysis was done using the basepair platform (https://www.basepairtech.com/). For miRNA expression analysis, the adapter sequence, which is a chemically synthesized oligonucleotide needed to be detected and sequenced by the sequencing machine, was removed from raw single-end reads using 'cutadapt'. Reads missing the adapter sequence were excluded from further analysis. For reads that were at least 16 bps long after the removal of the adapter sequence were retained for further analysis including the following steps: (1) Reads were mapped to the human transcriptome derived from the UCSC genome assembly hg19 using 'STAR' (Dobin et al. 2013). (2) Read counts were determined for each transcript using 'featureCounts' (Liao et al. 2014). (3) Instead of determining DEGs via DeSeq2, normalized count values of pre-selected topic relevant genes were log-transformed to transcript per million (TPM) values and their expression levels were analyzed as done by Wan et al. (2016) in their miRNA analysis study.

For all mRNA-Seq experiments, the same principle as for miRNA data was applied. Only differences were: (1) that samples were sequenced as paired end reads to detect mRNA isoforms and (2) that no adapter trimming was necessary during the mRNA expression value analysis.

2.2.24 Bioinformatics tools used in miR-196b transcript, promoter and RNA-Seq analysis

In order to extract additional information beyond the empirical data of this study, bioinformatics tools were used. Each bioinformatics tool uses other operational models with different strengths and weaknesses, thus, a combination of different bioinformatics tools increases the confidence of predictions.

Evolutionary conservation

To investigate evolutionary conservation of miR-196b transcripts and promoter, the UCSC genome browser data base (https://genome.ucsc.edu/) and the ECR browser (https://ecrbrowser.dcode.org/) were used (Karolchik et al. 2014, Ovcharenko et al. 2004). The USCS genome browser was used to visualize alignment data with 5 different species against the human genome hg19. The ECR browser was used to quantify evolutionary conservation. ECR default parameters were adjusted in order to get conservation values for a maximal region of the sequence in query. ECR length was set to a minimum of 2 bps, ECR similarity to 1% and the coordinate system was set to 'relative position'. For testing homology of the miR-196b promoter, the chromosomal location chr7: 27209384-27210185 was used. For testing homology of the miR-196b transcript, three different chromosomal locations of the long miR-196b transcript variant MF139050 were used: chromosomal locations 1: chr7: 27.203.316-27.203.460, 2: chr7: 27.204.497-27.204.586, 3: chr7: 27.209.086-27.209.336 consisting of 145 bps, 90 bps and 251 bps, respectively. The program showed the number of bases it had used for alignment and its homology to other species in percentages. Finally the %-values for all three locations were combined. If the number of bases used for alignment by the ECR browser was not the full length sequence, percentages were back calculated to absolute numbers of bases and given in percent of the total sequence length.

Promoter prediction tools

Several promoter prediction tools were applied on a pre-selected DNA region, which could be the miR-196b promoter, to reduce the risk of false-positive results (Liu and States 2002). Pre-selected was a DNA region 201 bps upstream of the miR-196b gene consisting of ~1 kbs. It is located between the *HOXA9* and *HOXA10* gene on the minus strand: chr7:27209384-27210185. Initially the Promoter 2.0 was used, but due to its poor accuracy (Gangal and Sharma 2005) it was then surpassed by more advanced bioinformatics tools such as GPMiner (http://gpminer.mbc.nctu.edu.tw/index.php), Softberry tools (http://www.softberry.com/) (including TSSG, TSSW, FPROM), ZLAB tools (https://zlab.bu.edu/zlab/gene.shtml) (including Comet, Possum, Cister), YAPP (http://www.bioinformatics.org/yapp/cgi-bin/yapp.cgi) and EMBOSS CpGPlot (https://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/). These tools together with data provided by the USCS genome browser (e.g. CpG island, histone methylation marks, DNA accessibility) (https://genome.ucsc.edu/) revealed promoter relevant features. Additionally, the 'updated GenHancer' track in the UCSC genome browser revealed an overlap between the pre-selected DNA promoter-like region in this study with a dedicated larger promoter region and additionally indicated possible enhancers in its vicinity.

Table 26. Promoter prediction tools

Name	Specifications
GPMiner (Lee et al. 2012)	Predicts for example TATA-, CAAT-, GC- box, tandem repeats, DNA stability, CpG islands and transcription factor binding sites (TFBSs). Considered were those regulatory promoter elements, which fulfilled published conventions in respect of the TSSs (Narang et al. 2005, Kanhere and Bansal 2005, Ponger and Mouchiroud 2002, Srikant et al. 1997).Timestamp: 03/2020
Softberry (Solovyev et al. 2010, Solovyev et al. 2006, Bajic et al. 2006)	Collection of bioinformatics software of which TSSG, TSSW and FPROM were used. Timestamp: 03/2020 TSSG had a threshold value LDF of >4 TSSW had a threshold value LDF of >3.7 FPROM had a threshold value LDF of >0.0467 Maximum sensitivity and specificity are reached with an LDF of ≥12.14 in the FPROM program. TSSG is the most reliable Softberry tool with 1 false positive result in 5000 bps compared to 1 false positive result in about 4000 bps by TSSW or FPROM*.
ZLAB (Frith et al. 2002, Fu et al. 2004, Frith et al. 2001)	Collection of bioinformatics software of which Comet, Possum and Cister were used. However, they did not predict promoter features for the pre-selected miR-196b promoter sequence. Timestamp: 03/2020

Table continues on next page

Table 26. *Continued*

Name	Specifications
YAPP (Gershenzon and Ioshikhes 2005)	Eukaryotic core promoter predictor, suitable for predicting core promoter motifs, but not for CpG features. Strength of predictions is indicated by weight matrix scores for promoter elements (e.g. DPE, BRE). Performance score can be maximally 1. TSSs are estimated relative to promoter elements. Timestamp: 03/2020
EMBOSS CpG Plot (Madeira et al. 2019)	CpG feature detection. Timestamp: 03/2020
UCSC genome browser (Kent et al. 2002)	annotation files: ENCODE regulation: H3K27Ac/H3K4me3 in 7 cell lines CD34 DNaseI: Nuclease accessible sites in CD34+/- human BM ENCODE open chromatin by DNaseI and FAIRE: Open chromatin accessible for regulatory proteins Updated GeneHancer: enhancer and promoter SwitchGear TSS: TSSs Transcription Factor ChIP-Seq clusters CpG islands

All programs were used with default settings. Except for YAPP the threshold for TSS detection was set to 0.6. * Reference: http://www.softberry.com/berry.phtml?topic=tssg&group=help&subgroup=promoter

Transcription factor (TF) prediction tools

TFs are the key factors in reconstructing regulatory gene networks. They regulate transcription by binding to promoters or enhancers (Boeva 2016). For TF determination, the miR-196b promoter region was used in several TF prediction tools such as the GPMiner (http://gpminer.mbc.nctu.edu.tw/index.php), Mapper (http://bio.chip.org/mapper), Jaspar (http://biogrid-lasagna.engr.uconn.edu/lasagna_search/), TFBind vertebrate (http://tfbind.hgc.jp/), P-Match (http://gene-regulation.com/cgi-bin/pub/programs/pmatch/bin/p-match.cgi), Alibaba 2.1 (http://gene-regulation.com/pub/programs/alibaba2/index.html) or the USCS genome browser providing Transcription Factor ChIP-Seq clusters. Considered were predicted transcription factor binding sites (TFBSs) of the minus and the plus strand, since till date it is not clear whether binding of TFs to the minus or the plus strand is decisive for the direction of transcription (Andersson et al. 2015, Almada et al. 2013, Ntini et al. 2013, Duttke et al. 2015). Presented were TFs predicted by at least 3 tools or 2 tools if 1 of them was the USCS genome browser track for ChIP-Seq validated data. For narrowing down relevant TFs on the miR-196b promoter region, gene lists of the mentioned prediction tools were overlapped using the automated online Venn diagram tool (https://bioinformatics.psb.ugent.be/webtools/Venn/). In order to not miss any important TFs with this automated approach, non-overlapping TFs were manually checked for aliases names using the GeneCards platform (https://www.genecards.org/) and compared to the rest of TFs in this non-overlapping gene list. This was done particularly thoroughly for hematologically relevant genes. By using both

the automatic and manual algorithms sources of errors in the TF overlapping process were minimized. In order to determine a local concentration (enrichment) of TF binding along the miR-196b promoter region, binding sites of the most commonly predicted TFs were summed up and attributed to 3 specific promoter locations: (1) proximal to the miR-196b gene: (chr7: 27209384-27209651), (2) in the middle (chr7: 27209651-27209918) or (3) distal (chr7: 27209918-27210185). TFBSs of one TF was considered to be enriched at one of the 3 promoter locations, if >40% of the TFBS were concentrated at this promoter location.

Table 27. Transcription factor (TF) prediction tools

Name	Settings
GP miner (Lee et al. 2012)	Default settings. Timestamp: 03/2020
Alibaba 2.1 (Grabe 2002)	Default settings. Timestamp: 03/2020
TFBind vertebrate (Tsunoda and Takagi 1999)	Default settings, Timestamp: 03/2020
P-Match (Chekmenev et al. 2005)	Settings: Group of matrices: vertebrates; Cut-off selection: to minimize sum of both error rates; profile: immune specific; untick high quality matrices only and tick to 'include matrices by the user', which was defined to select for matrices 'to minimize the sum of both error rates' which are false positives and negatives. This was done to increase the output. High quality results were included and extended to few more TFs TRANSFAC® positional weight matrices and site alignments from 2005 are used in this program to predict TFBSs. This program can be accessed on the link provided in the text only if one logs in.
Mapper (Marinescu et al. 2005)	Mulit-genome Analysis of Positions and Patterns of Elements of Regulation. Default settings.Timestamp: 07/2017
Jaspar (Lee and Huang 2013)	Was accessed via the LASAGNA-Search 2.0 platform by selecting Jaspar Core Matrices for all vertebrates TFs. Notably, a lot of error messages appear before the list of TFs, which can be ignored according to provider. Timestamp: 03/2020
UCSC genome browser (Kent et al. 2002)	Transcription Factor ChIP-Seq clusters. Annotation file: 'Txn Factor ChIP' as part of the 'Integrated Regulation Encode' super-track. 'TFBS conserved' and 'ChIP data' switched on

Heatmap generation

To visualize normalized count values of miR-196b specific candidate genes from RNA-Seq data, a heatmap was generated using GenePattern (http://www.genepattern.org/) (Reich et al. 2006). For that, 3 modules were utilized. First, the module 'Select Feature Rows' was used to slice the file harboring all normalized count values for the genes of interest. Second, the module 'Hierarchical Clustering' was used to generate the .atr and .cdt file required in the third module. Default parameters with activated 'Pearson correlation' for measuring column (sample) distance and a Pairwise average linkage as clustering method were used. Third, the module 'Hierarchical Clustering Image' was used to visualize normalized count values of the gene selection. Column size was set to 100, row size to 10 and description of row was activated. Apart from these adjustments default parameters were used throughout all modules.

MiR-196b specific target gene list determination

One miRNA can target several mRNAs. For the identification of such targets over 1000 bioinformatics programs are available (Chen, Heikkinen, et al. 2019). Each uses other operational models with different strengths and weaknesses making it important to use a combination of different bioinformatics tools in order to increase the confidence of predictions. To determine miR-196b specific DEGs and candidate genes in this study, a list of miR-196b specific target genes was created on the basis of four different miRNA-target platforms. MiRTarBase 7.0 (updated link: https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/php/index.php) and TarBase 6.0-8.0 (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8%2Findex) provided experimentally validated miR-196b specific target genes in mice (Vergoulis et al. 2012, Huang et al. 2020). TargetScan 7.1 (http://www.targetscan.org/mmu_71/) and miRDB (http://www.mirdb.org/) provided predictions of miRNA targets in humans (Agarwal et al. 2015, Wang 2016). Notably, TargetScan is the most user friendly and well revised bioinformatics approach. It uses seed sequence matching and a free energy variable, which gives likelihood for miRNA/RNA interaction, based on RNA folding ability.

Enrichment Analysis

Gene set enrichment analysis allows for biological insights of sequencing data. While single gene expression levels might vary between studies, their involvement in specific pathways underlying a certain phenotype might be the same (Subramanian et al. 2005). Therefore, not only the largest difference of single gene expression levels is relevant for grasping the in this study empirically investigated phenotype of miR-196b in a (Para)Hox gene BM cell context, but also, if not even more important are overrepresented classes of DEGs in a large set of single genes.

To investigate gene set enrichment in murine HSPCs transduced with *Cdx2* or *Cdx2*/miR-196b, RNA-Seq data were analyzed using the Enrichr platform (https://maayanlab.cloud/Enrichr/) and the GSEA desktop application (Chen et al. 2013, Kuleshov et al. 2016). The Enrichr platform provided gene set libraries and was particularly good for getting a quick overview in which pathways pre-ranked lists of genes, based on DeSeq2 and FDR-values, accumulate. The GSEA desktop application was used to analyze gene expression values that have not been ranked beforehand. Thus, normalized count

values of genes unfiltered or filtered for potential miR-196b targets were internally ranked within a particular gene matrix accessed on the MsigDB platform. Gene sets not available on the MsigDB platform could be created manually, which was done for the LSC signature used in this study. Gene names of upregulated LSC genes published by Ng et al. (2016) and Gentles et al. (2010) available on MsigDB were combined to one LSC signature list (.gmt file). After doublet deletion the manually created LSC signature list consisted of 76 genes.

The GSEA settings were:

Table 28. GSEA software settings

Parameter	Setting	Remarks
Number of permutation	1000	default
Collapse:	Collapse	default
Permutation type:	Gene_set	recommended when <7 samples in each phenotype
Chip platform	local .chip	provided by BasePair

General criteria for setting the focus on certain pathways were: (a) statistical significance (p-value), (b) the consideration of false positive results resulting from multiple testing (FDR-value), (c) the number of gene scores when using Enrichr and (d) normalized enrichment scores (NES) when using the GSEA software. In addition to the statistical parameters, the analysis was always paired with biological expertise to give the data a hypothetical direction.

2.2.25 Statistical Analysis

Data were presented as individual data points and in case of a sample size >3, also the median with 95% Confidence interval (CI) was displayed. Mann-Whitney *U* test was used for the comparison of two independent conditions. One sample Wilcoxon signed rank test was used for comparing the experimental condition to the control condition. Kruskal-Wallis with Dunn's post hoc test was used if more than two independent conditions were compared. P-values were given by the exact test. Kaplan-Meier plots were generated for a survival analysis using the log-rank test for the comparison of two independent conditions. For the association studies expression level changes were analyzed as well as statistical correlations were examined via scatterplots and the determination of the spearman correlation coefficient (r) together with the 95% CI.

Data acquisition and evaluation was done with PRISM Graph pad software 7.04 (La Jolla, California, USA) and Microsoft Excel 2010. All evaluations must be taken as explorative data analysis.

3. Results

3.1 MiR-196 family expression in healthy and leukemic human hematopoietic cells

To evaluate the expression pattern of miR-196a and miR-196b in healthy human hematopoietic cells, qRT-PCR was performed on primitive BM cells (BM CD34$^+$) and more differentiated mononuclear BM cells (BM MNCs). The expression levels of miR-196a and miR-196b were higher in BM CD34$^+$ compared to BM MNCs, showing ~11 and ~21 times higher expression values, respectively (**Figure 9A**). Notably, miR-196b was higher expressed than miR-196a in both BM CD34$^+$ and MNCs, showing ~9 and ~5 times higher expression values, respectively. Similarly, in AML cells, miR-196b was 55 times higher expressed than miR-196a regardless of the AML karyotype: Inversion 16 (Inv 16), AML1-ETO (t(8;21)), PML-RARA (t(15;17)) and CN-AML subtypes, assessed in a published miRNA-Seq dataset of a large AML patient cohort (Cancer Genome Atlas Research et al. 2013) (**Figure 9B-C**). MiR-196b expression levels were particularly high in CN-AML patients, of which 70% had 16 to 261 times higher expression values compared to the lowest and highest miR-196b median expression level of the three AML subtypes with a low median miR-196b expression level: Inv 16, t(8;21) and t(15;17) (**Figure 9C**).

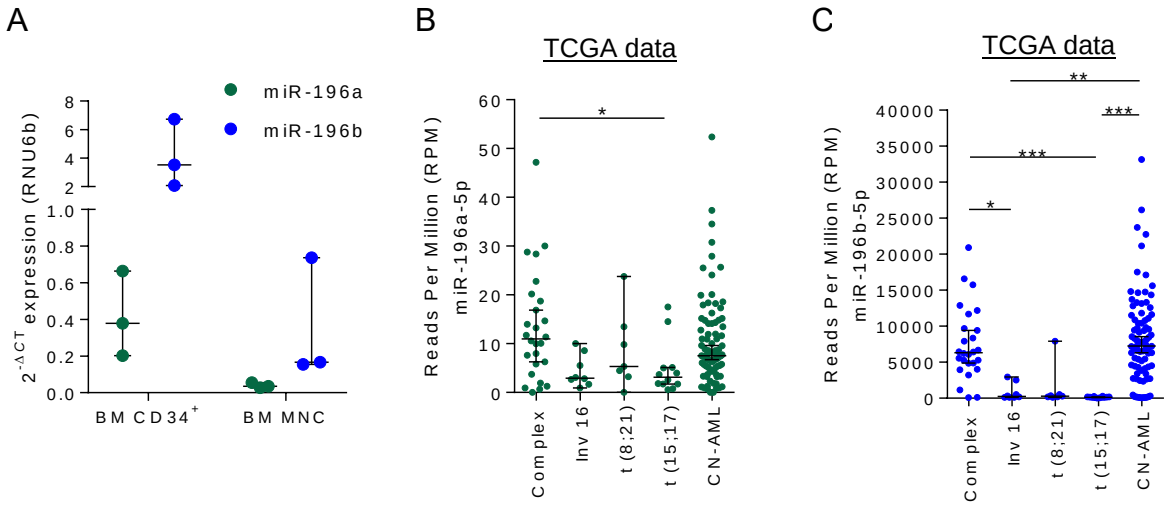


Figure 9. Expression of miR-196 family members in healthy BM cells and AML patient samples. (A) Expression levels of miR-196a and miR-196b in bone marrow fractions were determined via qRT-PCR, normalized on the housekeeping gene *RNU6b* and transformed into $2^{-\Delta CT}$-values, n=3. Presented are individual values displaying the median, minimum and maximum value. No inference statistical test was applied. BM: Bone Marrow; MNC: Mononuclear Cell; C_T: Cycle Threshold; RNU6b (RNU6-6P): RNA, U6 small nuclear 6, pseudogene. **(B-C)** Expression levels of miR-196a (B) and miR-196b (C) in AML subtypes were determined via the analysis of published microRNA-Sequencing data from The Cancer Genome Atlas and are presented as Reads Per Million (RPM) values (Cancer Genome Atlas Research et al. 2013). According to the karyotype patients were categorized into complex karyotype (Complex, n=26), inversion 16 (Inv16, n=8), AML1-ETO (t(8;21), n=7), PML-RARA (t(15;17), n=12) and cytogenetically normal acute myeloid leukemia (CN-AML: n=86 (miR-196b) and n=85 (miR-196a). One miR-196a expression value was 248 times higher than the average value and therefore declared as outlier and excluded from the analysis. Presented are individual values with the median and 95% confidence interval. Statistical significance was determined using the Kruskal-Wallis and Dunn's post hoc tests.

Based on the expression levels, these data suggested a more prominent role for miR-196b over miR-196a in healthy BM and AML cells. Henceforth, the focus of this study lied on miR-196b.

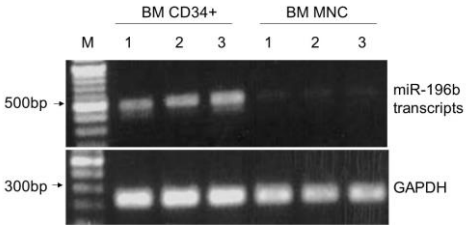
3.2 MiR-196b transcript determination in healthy human BM cells

Based on the different expression levels of miR-196b between primitive and more mature human BM cells, the next aim was to identify key molecular elements involved in its expression regulation. Therefore, the miR-196b transcript was determined (identified, characterized and validated) in BM CD34$^+$ and BM MNCs and its correlation with mature miR-196b evaluated.

Identification

In healthy human BM CD34$^+$ and BM MNCs, a conventional PCR reaction using a primer pair located in the *HOXA9* and miR-196b stem-loop sequence locus on chromosome 7 identified two PCR products containing ~500 bps and ~400 bps (**Figure 10A**). Sanger sequencing revealed that the shorter sequence lacked a 90 bp long fragment, which was present in the longer transcript and harbored two mismatches in the 5' vicinity of the missing fragment (**Figure 10B**). The rest of the shorter sequence was 99.5% homologous to the longer sequence. Both, the long (486 bps) and the shorter (396 bps) miR-196b transcript sequence have been deposited as MF139050 and MF139051, respectively, at the NCBI gene bank data base.

A



B

```
MF139050      CTGGGGCACTCTGTTGCACTGGCGGGCGCAGGTTGCCTAGGGGCTGGGCTGGGCCGGGCC  60
MF139051      CTGGGGCACTCTGTTGCACTGGCGGGCGCAGGTTGCCTAGGGGCTGGGCTGGGCCGGGCC  60
              ************************************************************

MF139050      AGGCGCGATGGCAGGGTTCTCTCCTTGGCGGCGGCGGCAGCGGCGGAGGCGGCGGCGGCG  120
MF139051      AGGCGCGATGGCAGGGTTCTCTCCTTGGCGGCGGCGGCAGCGGCGGAGGCGGCGGCGGCG  120
              ************************************************************

MF139050      GCGGGCGAGGCAGCCCTTCGCGGGCAGCACCAGAACTGGTCGGTGATTTAGGTAGTTTCC  180
MF139051      GCGGGCGAGGCAGCCCTTCGCGGGCAGCACCAGAACTGGTCGGTGATTTAGGTAGTTTCC  180
              ************************************************************

MF139050      TGTTGTTGGGATCCACCTTTCTCTCGACAGCACGACACTGCCTTCATTACTTCAGTTGAA  240
MF139051      TGTTGTTGGGATCCACCTTTCTCTCGACAGCACGACACTGCCTTCATTACTTCAGTTGAA  240
              ************************************************************

MF139050      ATCGTCTCCAGTTGATAGAGAAAAACAACCCAGCGAAGGCGCCTTCTCTGAAAACAATGC  300
MF139051      ATCGTCTCCAGATA----------------------------------------------  254
              ********** *

MF139050      TGAGAATGAGAGCGGCGGAGACAAGCCCCCCATCGATCCCAATAACCCAGCAGCCAACTG  360
MF139051      ----------------------------------------------ACCCAGCAGCCAACTG  270
                                                             ****************

MF139050      GCTTCATGCGCGCTCCACTCGGAAAAAGCGGTGCCCCTATACAAAACACCAGACCCTGGA  420
MF139051      GCTTCATGCGCGCTCCACTCGGAAAAAGCGGTGCCCCTATACAAAACACCAGACCCTGGA  330
              ************************************************************

MF139050      ACTGGAGAAAGAGTTTCTGTTCAACATGTACCTCACCAGGGACCGCAGGTACGAGGTGGC  480
MF139051      ACTGGAGAAAGAGTTTCTGTTCAACATGTACCTCACCAGGGACCGCAGGTACGAGGTGGC  390
              ************************************************************

MF139050      TCGACT   486
MF139051      TCGACT   396
              ******
```

Figure 10. Identification of miR-196b transcripts in healthy human BM cells. (A) Amplification products of miR-196b transcripts after conventional PCR visualized in a 1.2% agarose gel. *GAPDH* was used as housekeeping gene. n=3. **(B)** Sequence alignment of MF139050 (long transcript 5'-3') with MF139051 (short transcript 5'-3') using the Clustal Omega 0 (1.2.4) tool (Madeira et al. 2019). Stars (*) indicate identical bases. Dashed line indicates missing bases. No sign indicates a mismatch. BM: Bone Marrow; MNC: Mononuclear Cell; M: Molecular weight size marker; 1-3: Lane numbering indicating different biological replicates; bp: basepair; GAPDH: Glyceraldehyde 3 Phosphate Dehydrogenase. A: Adenine; C: Cytosine; G: Guanine; T: Thymine. Figure extracted from Rawat & Götze et al., 2020, Supplementary Materials, p. 37.

Characterization

An alignment of the miR-196b transcripts with annotated RNAs revealed that they encode the miR-196b stem-loop sequence and harbor parts of the *HOXA9* but not the *HOXA10* coding mRNA sequence. Additionally, miR-196b transcripts aligned with sequence parts of the non-coding RNAs: HOXA10-HOXA9 and HOXA-AS4 (*bars in* **Figure 11**). An accumulated alignment of the miR-196b transcript across 100 vertebrates revealed high homology across species (*histogram in* **Figure 11**) indicating functional importance. Highest conservation existed in mammals with 94.0%, 92.4% and 88.7% sequence identity in rhesus macaque, dog and mouse, respectively. In chicken 76.8% and in frog 63.6% sequence identity with MF139050 was detected using the Evolutionary Conserved Region (ECR) browser.

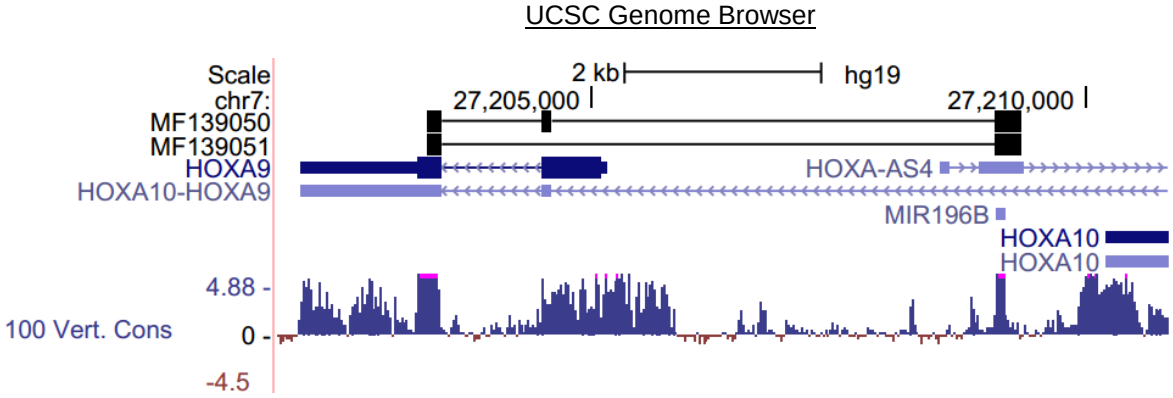


Figure 11. **Characterization of miR-196b transcripts.** Sequence alignment of MF139050 and MF139051 with each other and with sequence data from various organisms provided by the University of California Santa Cruz (UCSC) Genome Browser assembly hg19 (Kent et al. 2002, Kent 2002). Blue bars indicate UCSC genes of which dark blue depicts coding transcript sequences and bright blue depicts non-coding transcript sequences. '100 Vert. Cons' track reveals sequence conservation across 100 vertebrates. Figure extracted from Rawat & Götze et al., 2020, p. 286

Validation

The two miR-196b transcript sequences, MF139050 and MF139051, were validated by cloning their PCR products into the MIY vector, constitutively expressing them in different cell types and measuring the production of mature miRNA via qRT-PCR. In the human fibroblast cell line HEK293T, miR-196b expression increased on average 19- and 58- times upon MF139050 and MF139051 overexpression, respectively (*row 1 and 2 in* **Table 29**). This could also be shown in the murine myeloid cell line 32D, in which miR-196b expression levels increased 34065- (normalization to *RNU19*) to 34489- (normalization to *RNU6b*) times upon MF139050 overexpression (*row 3 in* **Table 29**). Also in primary murine HSPCs, MF139050 overexpression led to an averaged 108- (normalization to *RNU19*) to 209- (normalization to *RNU6b*) fold increase of miR-196b (*row 4 in* **Table 29**). Thus, both transcripts are able to generate considerably high mature miR-196b expression levels.

Table 29. Validation of miR-196b transcripts in cell lines and primary BM cells. Expression levels of mature miR-196b after constitutive transcript overexpression in (a) HEK293T (n=3), (b) 32D (n=1) and (c) murine HSPCs (n=2) were determined via qRT-PCR, normalized on the housekeeping gene *RNU6b* or *RNU19*, calculated relative to the control condition and transformed into $2^{-\Delta\Delta CT}$-values.

Cell type	Overexpression construct	$2^{-\Delta\Delta CT}$-values of miR-196b (RNU6b/RNU19)		
		Exp. 1	Exp. 2	Exp. 3
HEK293T	MF139050	19.74	20.18	17.12
HEK293T	MF139051	30.83	111.67	31.41
32D	MF139050	34489.33/ 34064.79	--	--
5-FU BM	MF139050	410.0363/ 209.4134	8.7544/ 7.6211	--

No inference statistical test was applied. Dashed line means no data available. HSK: Housekeeping gene; C_T: Cycle Threshold; RNU6b (RNU6-6P): RNA, U6 small nuclear 6, pseudogene; RNU19: RNA, U19 small nuclear; HEK293T: Human Embryonic Kidney cell transformed with large T-antigen; 32D: Murine myeloid cell line; Ctrl: Control; HSPCs: Hematopoietic Stem and Progenitor Cells.

Association of miR-196b transcript with mature miR-196b

Since miRNA expression and processing is regulated in a precise and collaborative way, the association of the newly identified miR-196b transcripts and its mature form was tested by determining qRT-PCR values utilizing custom primer probes of comparable efficiency tested a priory (*see footnote of Table 22).* As for mature miR-196b expression (Figure 9), both the shorter (MF139051) and longer transcript (MF139050) were higher expressed in healthy human BM $CD34^+$ cells compared to BM MNCs, showing 8 and approximately 5 times higher expression values, respectively (**Figure 12A**). Interestingly, in human BM $CD34^+$ cells, MF139051 was 2x higher expressed than MF139050, whereas in human BM MNCs only MF139051 was detectable. An association analysis of MF139051 with mature miR-196b expression values revealed that, if ΔC_T-expression values of MF139051 increased, ΔC_T-expression values of miR-196b increased accordingly (**Figure 12B**). As a side note, for MF139050 the amount of only three qRT-PCR based data points made an association analysis inapplicable.

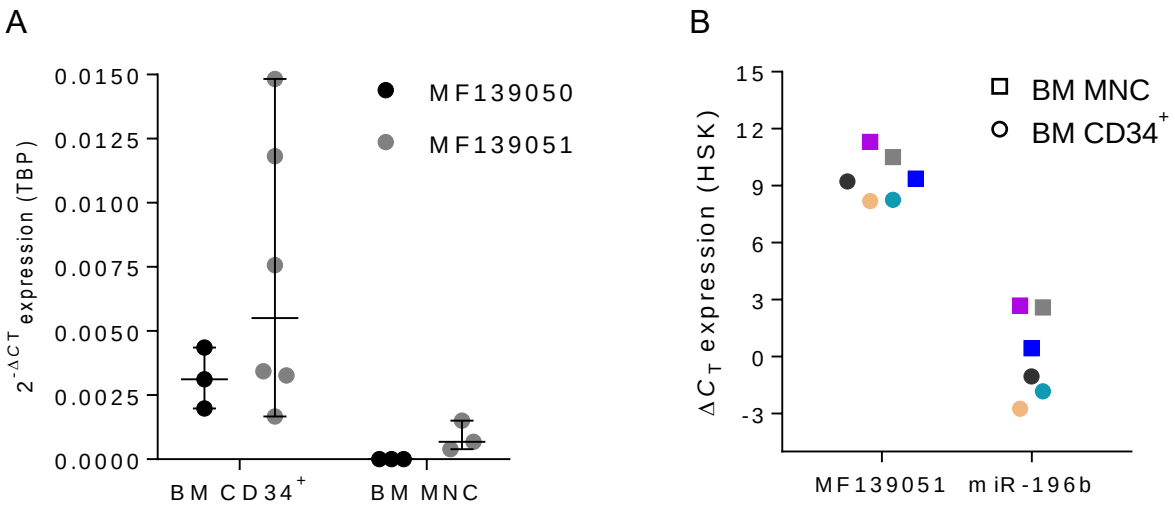


Figure 12. Expression of miR-196b transcripts and its relation to mature miR-196b in healthy human BM cells. (A) Expression levels of MF139050 (BM CD34⁺: n=3; BM MNC: n=3) and MF139051 (BM CD34⁺: n=6; BM MNC: n=3) in bone marrow fractions. Presented are individual values with the median and 95% confidence interval. **(B)** Association of MF139051 with miR-196b in BM CD34⁺ (n=3) and MNCs (n=3) was determined by comparing naturally occurring MF139051 expression level variations between individuals with corresponding miR-196b levels. Colors indicate data affiliations. **(A-B)** Individual expression levels were determined via qRT-PCR, normalized on the housekeeping gene *TBP* (for transcripts) or *RNU6b* (for miR-196b) and are presented as $2^{-\Delta CT}$ (A) or ΔC_T-(B) values. No inference statistical test was applied. BM: Bone Marrow; MNC: Mononuclear Cell; C_T: Cycle Threshold; TBP: TATA-box binding protein. RNU6b (RNU6-6P): RNA, U6 small nuclear 6, pseudogene.

3.3 MiR-196b promoter determination in healthy human BM cells

Next, as another key molecular element involved in miR-196b expression regulation, the promoter was determined (identified and validated) and characterized by using especially publicly available data on the human genome provided by the UCSC genome browser and promoter prediction tools (*see Table 26 and 27*).

Identification

A ~1 kb long intergenic DNA sequence nearby the miR-196b gene locus was identified as potential miR-196b promoter region on the basis of publicly available data on the human genome provided by the UCSC genome browser (**Figure 13A)**. This region was found to be highly conserved using an interspecies comparison among mammals with 95.5%, 83.4% and 77.9% sequence identity in rhesus macaque, dog and mouse, respectively, whereas in chicken 57.1% and in frog 22.8% sequence was detected using the Evolutionary Conserved Region (ECR) browser (**Figure 13B**). Notably, this promoter region is 3.5 kb away from the classical *HOXA9* promoter suggesting a *HOXA9* independent miR-196b promoter. Additionally, bioinformatics tools (*see chapter 2.2.24*) predicted TSSs and specific core promoter elements, which will be addressed later in detail under the heading ‘characterization’.

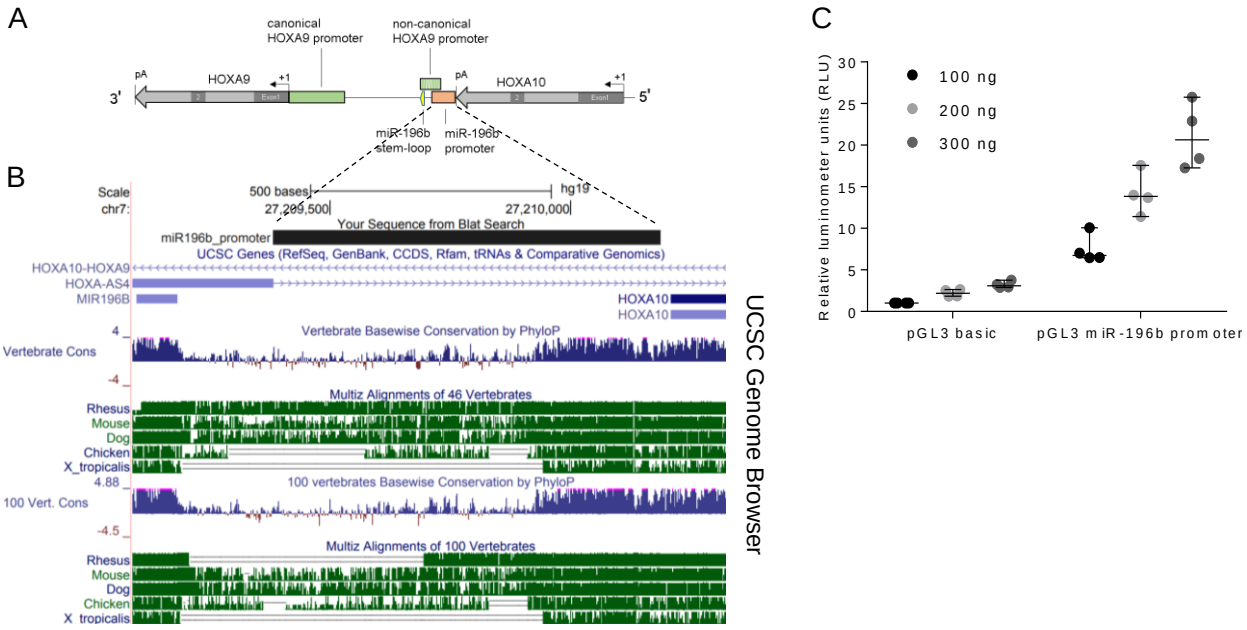


Figure 13. MiR-196b promoter region identification and validation. (A) Schematic representation of the human HOXA9-10 genomic region to scale. Grey block arrows represent the *HOXA9* and *HOXA10* gene from their transcription start site (+1) to their poly-A (pA) sequence. The known miR-196b stem-loop sequence is indicated by a yellow block arrow. The canonical and non-canonical HOXA9 promoter sequence (Sequence from SwitchGear Genomics Prod. ID: S720614 and S717602, respectively) are indicated as green squares not patterned and patterned, respectively. The predicted miR-196b promoter region is indicated by an orange square. **(B)** Conservation scores of the miR-196b promoter region across 100 vertebrates in total, provided by the University of California Santa Cruz (UCSC) Genome Browser and is displayed in blue histograms (Kent et al. 2002). Below a pairwise alignment of individual animals to the human genome (GRCh37/hg19) is displayed in green histograms. Double line indicates one or more unalignable bases. 'Multiz Alignments of 46 Vertebrates' conservation data is based on genome data from the version 2009. 'Multiz Alignments of 100 Vertebrates' conservation data is based on genome data from the version 2015. UCSC genes are displayed in blue bars: dark blue depicts coding and bright blue non-coding transcript sequences. The miR-196b promoter region is depicted as black bar. X.tropicalis: Western clawed frog Xenopus tropicalis. **(C)** Luciferase activity of the miR-196b promoter like sequence determined by bioinformatics tools was assayed in HEK293T using three different plasmid concentrations presented by three different colors: black, light and dark grey. Luminescence is given as relative values to the control. Presented are individual values with the median and 95% confidence interval. No inference statistical test was applied. n=4. pGL3: plasmid harboring the luciferase gene; HEK293T: Human Embryonic Kidney cell transformed with large T-antigen.

Validation

The prospective miR-196b promoter region was validated by the luciferase reporter assay, in which the promoter region (chr7: 27.209.384-27.210.185) was cloned into the pGL3-luciferase basic vector lacking a promoter and HEK293T cells were transfected with different concentrations of this plasmid. The promoter region showed activity across all three plasmid concentrations used (**Figure 13C**). Upon 100 ng test plasmid the activity signal (luminescence) was 7.5-fold higher compared to the control plasmid. Upon 200 ng and 300 ng test plasmid the activity signal doubled and tripled proportionally.

Characterization

Following the observation that the promoter region nearby the miR-196b gene locus was indeed active, promoter-related regulatory elements that potentially influence transcript expression were evaluated.

First, six TSSs were predicted by several well-established bioinformatics tools (*see Table 26*), of which TSS 1-2 were located at the beginning (5'-end) and TSS 3-6 at the end (3'-end) of the promoter region (**Figure 14A**). TSSs 1-2 were close to a TATA-box or to its alternative: an initiator (INR). TSS 3-6 on the other hand were embedded in CpG islands. Interestingly, TSS 6 was also allocated to tandem repeats, an activity feature, but more importantly to a TF IIB recognition element (BRE) and a downstream core promoter element (DPE). The cooperation of BRE and DPE has been proposed in transcription regulation of particularly TATA-less promoters (Gershenzon and Ioshikhes 2005). The results suggest that the identified promoter harbors both a TATA-driven and a TATA-less promoter region proximal and distal to the miR-196b gene, respectively. In addition to these six predicted TSSs, another experimentally validated TSS outside the promoter region nearest to the miR-196b gene and located within a CpG island was revealed by the UCSC genome browser (**Figure 14B**). It is probably a TSS of the non-canonical *HOXA9* promoter.

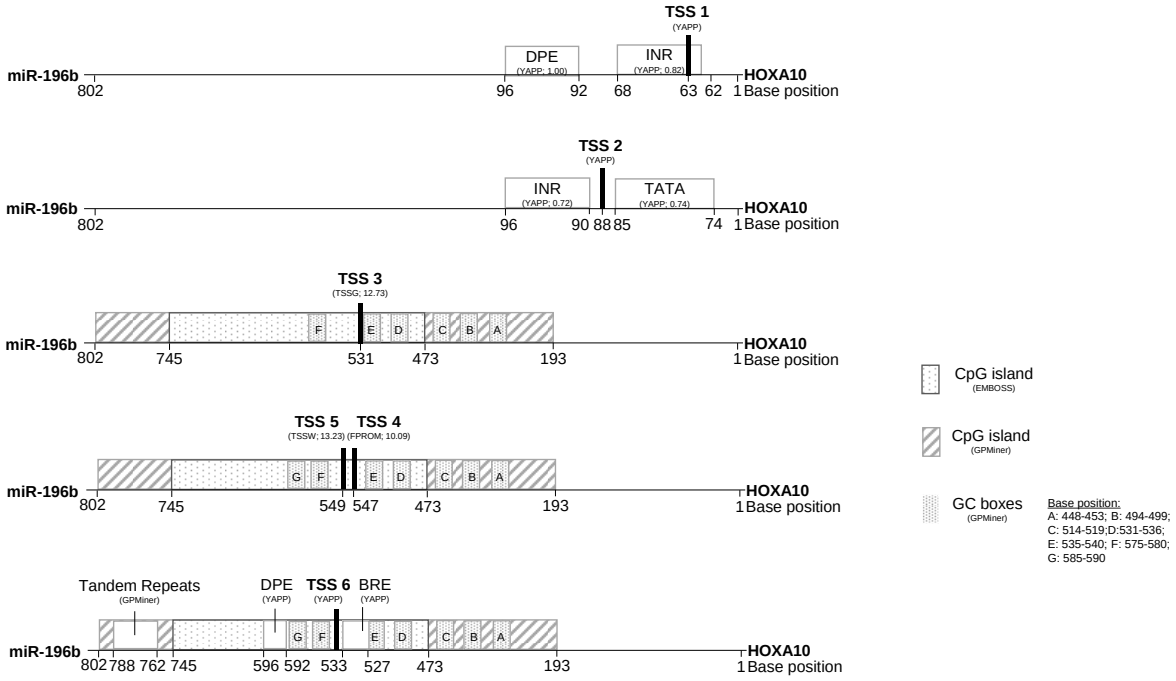
Second, epigenetic data revealed two distinct histone mark peaks. H3K27Ac (acetylated lysine residue 27 of histone 3) and H3K4Me3 (trimethylated lysine residue 4 of histone 3), which are generally associated with active transcription were located at the proximal and distal promoter location (**Figure 14B**). Consistent with this, chromatin accessibility data revealed DNase hypersensitivity and nuclease sensitivity sites across the entire promoter region with highest scores proximal and distal to the miR-196b gene in blood system relevant cells (B-cells, $CD34^+$ cells $CD14^+$, acute promyelocytic HL60 cell line) (**Figure 14B**). Thus, the transcriptional activity might be concentrated at these two promoter locations.

Third, possibly relevant regulatory proteins controlling promoter activity were identified by analyzing publicly available ChIP-seq data across the entire promoter region. The classical transcriptional activator 'RNA polymerase II' and the transcriptional repressor 'enhancer of zeste homolog 2' (EZH2), were detected in the majority of cell types with strong sequencing signals (**Figure 14B**). In addition to the experimentally validated, but very cell type specific TFs, 20 other TFBSs were identified by the usage of six online prediction tools (*see Table 27*) (**Figure 14C** left). Most commonly predicted TFs were SP1, MZF1, CEBPA, AML1, NF1 and NFKAPPAB. Their TFBSs were found to be locally enriched (>40% of all TFBSs) across the promoter region for each TF, but overall there was no particular concentration of TFBSs along the promoter (**Figure 14C** right).

Beyond the proximal core promoter elements, also three distal promoter elements termed 'enhancers' could be identified: by visualizing published data on the UCSC genome browser platform (**Figure 14B**). One enhancer was located ~40 kbs upstream and two enhancers were located ~125 kbs and ~140 kbs downstream.

Together all these regulatory promoter elements provide information on the activity and the dynamic of the promoter region consequentially affecting transcription of miR-196b.

A



B

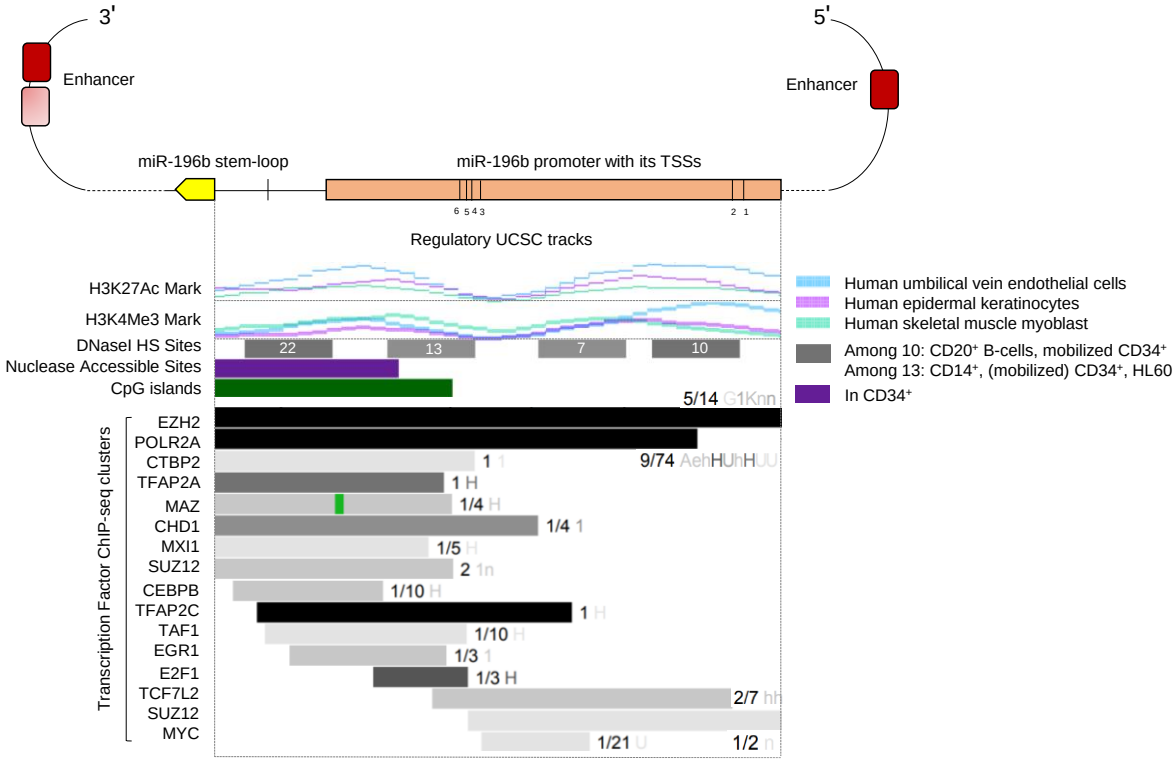


Figure 14 continues on next page

Figure 14 continued

C

Mapper	Jaspar	TFBind	P-Match	Alibaba 2.1	UCSC	GPMiner
AML1	AML1	AML1	AML1			
ARNT	ARNT	ARNT				
CEBPA	CEBPA	CEBPA	CEPBA	CEBPA		CEBP
		CEBPB	CEBPB		CEBPB	
COMP1		COMP1	COMP1			
CREB1	CREB1	CREB1				
DDIT3		DDIT3	DDIT3			
EGR1		EGR1		EGR1	EGR1	
ELK1		ELK1	ELK1			
E2F		E2F				E2F
E2F1					E2F1	
HNF1A		HNF1A		HNF1A		
IRF1		IRF1		IRF1		
LYF1		LYF1	LYF1			
MAZ					MAZ	MAZ
MZF1	MZF1	MZF1	MZF1			MZF1
NFKAPPAB	NFKAPPAB	NFKAPPAB	NFKAPPAB			
NF1		NF1		NF1		NF1
NHLH1	NHLH1	NHLH1				
PAX2		PAX2	PAX2			
	PBX1	PBX1	PBX1			
PU.1	PU.1			PU.1		
SP1	SP1	SP1	SP1	SP1		
	TFAP2A			TFAP2A	TFAP2A	TFAP2A
					TFAP2C	TFAP2C
		YY1	YY1	YY1		

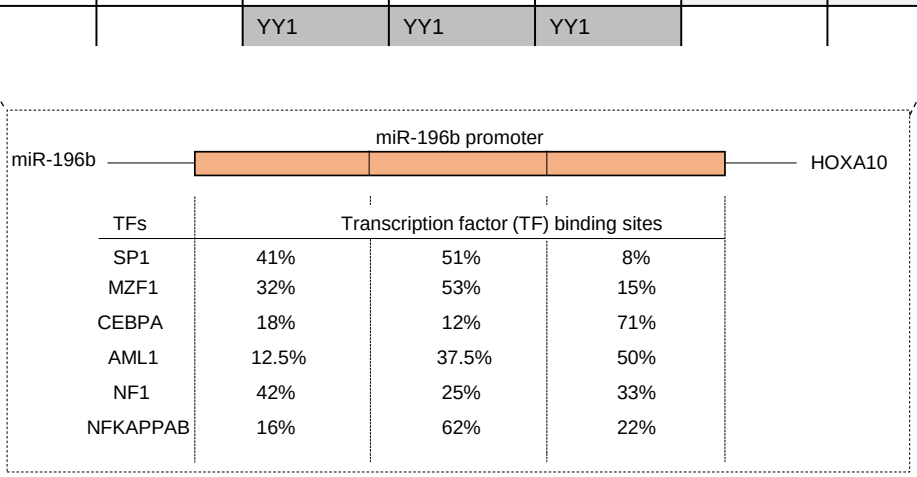


Figure legend on next page

Figure 14. Bioinformatically determined regulatory elements of the miR-196b promoter-like region. (A) Schematic diagram showing predicted sequence-based proximal core promoter elements. Transcription start sites (TSSs) and other regulatory elements were predicted by the online tools: YAPP, the Softberry tools: TSSG, TSSW and FPROM as wells by GPMiner and EMBOSS CpG. They are indicated in brackets together with a performance score. Threshold value of TSSG was 4, of TSSW was 3.7 and of FPROM was 0.0467. YAPP's performance score could be maximal 1. INR: initiator; DPE: downstream core promoter element. **(B)** Schematic diagram showing experimentally validated proximal and distal core promoter elements to the miR-196b gene, provided by the UCSC genome browser (Kent et al. 2002). Depicted are the miR-196b stem-loop sequence (yellow arrow block), the proximal promoter (orange block) with its predicted TSSs and an annotated TSS, CHR7_M0176_R2 outside the promoter sequence. Two enhancers (red rounded blocks) located ~40 kbs (dashed line) upstream and ~125-140 kbs (dashed line) downstream surround the proximal promoter. Underneath six annotation files, whose data is associated with promoter regions and active gene regulation are displayed to scale (*see Table 26 and 27 for name of annotation files*). The histone marks: H3K27Ac and H3K4Me3 identified in three different cell type are represented as lines. DNaseI hypersensitivity sites (DHSs) investigated among 125 diverse cell and tissue types are represented as grey blocks with numbers inside indicating the amount of cell types with a positive result for that DHS site. One nuclease accessible site identified in $CD34^+$ cells is displayed as dark purple and the CpG island 53 as dark green block. Transcription factor (TF) clusters are depicted as grey boxes. The darkness of the box is 'proportional to the maximum signal strength observed in any cell type contributing to the cluster' (*cited from display conventions for this UCSC track*). Green highlight indicates highest scoring of the canonical TF sequence fragment. To the right numbers indicate the amount of assays with positive hits out of the total amount of assays performed. Letters abbreviate cell line names with positive hits and are color coded alike the TF cluster conventions. Dark and light shadings indicate high and low scores, respectively. 1: H1-hESC; G: GM12878; K: K562; n: NH-A (EZH2), NHLF (EZH2), NT2-D1 (SUZ12); A: A549; e: ECC-1; h: HCT-116 (POLR2A, TCF7L2), HEK293 (TCF7L2); H: HeLa-S3; U: HUVEC. In case of multiple allocations to one letter the corresponding gene name is indicated in brackets. **(C)** The table shows a selection of TFs predicted by six online tools (*see Table 27*) additionally to ChIP-seq validated data provided by the UCSC genome browser (Kent et al. 2002). To the right the distribution of transcription factor binding sites (TFBSs) within the promoter are depicted for the six most commonly predicted and blood system relevant TFs. (**B-C)** For the full form of gene name abbreviations see section 'Abbreviations'.

3.4 Effect of aberrant miR-196b expression in murine hematopoietic stem and progenitor cells on clonogenicity, cell growth, lineage differentiation and transformation ability

The expression pattern of miR-196b in healthy human BM cells (Figure 9) suggests, that miR-196b plays a specialized role in regulating hematopoiesis on stem- and progenitor level. To ascertain miR-196b's functional impact, the miR-196b stem-loop sequence was constitutively expressed in enriched murine HSPCs and subjected to various *in vitro* and *in vivo* assays (**Figure 15**).

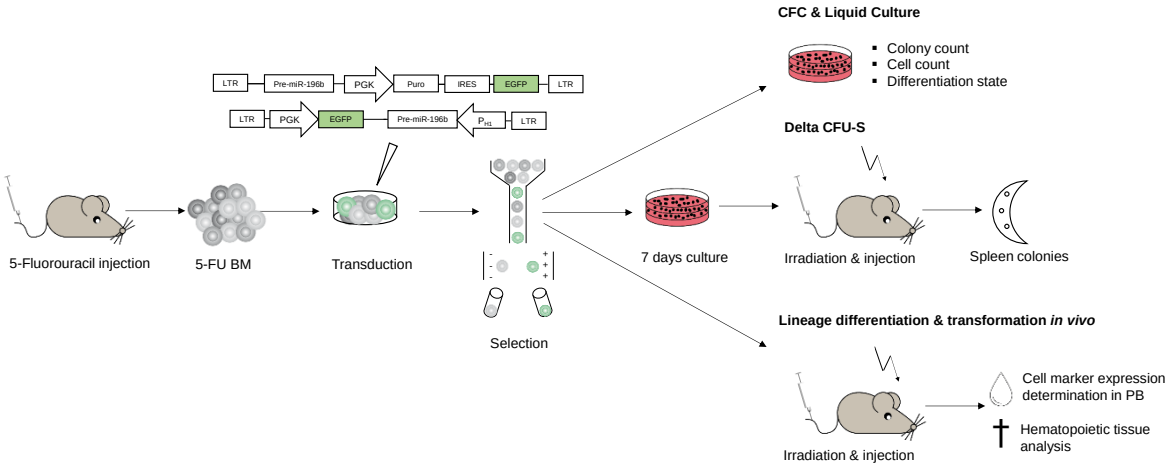


Figure 15. Experimental design for investigating the functional impact of miR-196b in healthy murine hematopoietic stem and progenitor cells. Bone marrow (BM) enriched for murine HSPCs via 5-Fluorouracil treatment was extracted and retrovirally transduced. After a FACS-based selection for reporter gene positive cells, *in vitro* and *in vivo* assays were performed.

In vitro, after miR-196b overexpression with the retroviral MSCV/LTR-miRNA-PIG (LMP) or Ph1 vector in HSPCs, the clonogenic potential of progenitor cells was investigated via colony forming cell (CFC) assays in semi-solid medium. In parallel, the ability of the cells to cease differentiation and grow over several days was tested by performing cell counts in liquid cultures. *In vivo*, the stem- and progenitor function was investigated via an evaluation of colony forming units on the spleen (CFU-S) indicative for the short-term repopulating ability. Then, the lineage differentiation potential was tested by detecting the GFP along with myeloid and lymphoid cell markers in PB of with miR-196b positive cells transplanted mice. Moreover, control and miR-196b mice were followed up till disease or for 400 days and in case of disease, animals were sacrificed and their hematopoietic tissue analyzed.

In vitro clonogenic and cell growth potential as well as the cells' differentiation state of miR-196b overexpressing murine HSPCs

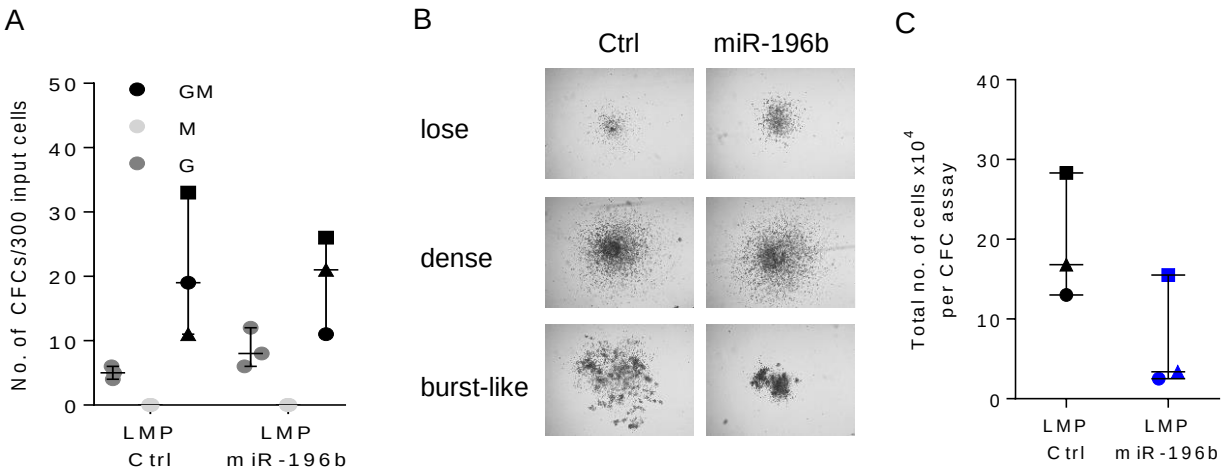
In CFC assays, a stable overexpression of miR-196b with the LMP vector in murine HSPCs (2970-fold (*RNU19*) or 4614-fold (*RNU6b*) relative to the control) did not affect (difference <20%) the median total amount of colonies compared to control cells across all three experimental replicates (Ctrl: median=24, 95% CI [15, 39] vs miR-196b: median=29, 95% CI [23, 32]; n=3) (**Figure 16A**). The type of cells forming the colonies were granulocytes and macrophages, which prominently appeared in a tight (dense) formation. Less frequent were dispersed (loose) or burst-like colonies (**Figure 16B**). The median number of colonies with cells originating from progenitors that gave rise to granulocytes (G) was 1.6- fold (37.5%) increased upon miR-196b overexpression, but remained unaffected (difference <20%) in colonies consisting of a granulocyte-macrophage (GM) mix. Harvested cells were over 98% GFP positive and the total amount of miR-196b positive cells was decreased by 79.9% [45.3, 80.7] on day 7 (**Figure 16C**). Total colony and cell numbers demonstrated that colonies from cells with miR-196b overexpression were smaller with approximately 80% less cells per colony than colonies from control cells (Ctrl: median=5.42x10^3, 95% CI [4.31x10^3, 18.87x10^3] vs miR-196b: median=1.09x10^3, 95% CI [1.06x10^3, 5.35x10^3]; n=3). Additionally, May-Grünwald-Giemsa stainings of cells from the CFC assays revealed terminally differentiated macrophages and granulocytes under experimental and control condition (**Figure 16D**). This was specified by a cell surface marker analysis, which

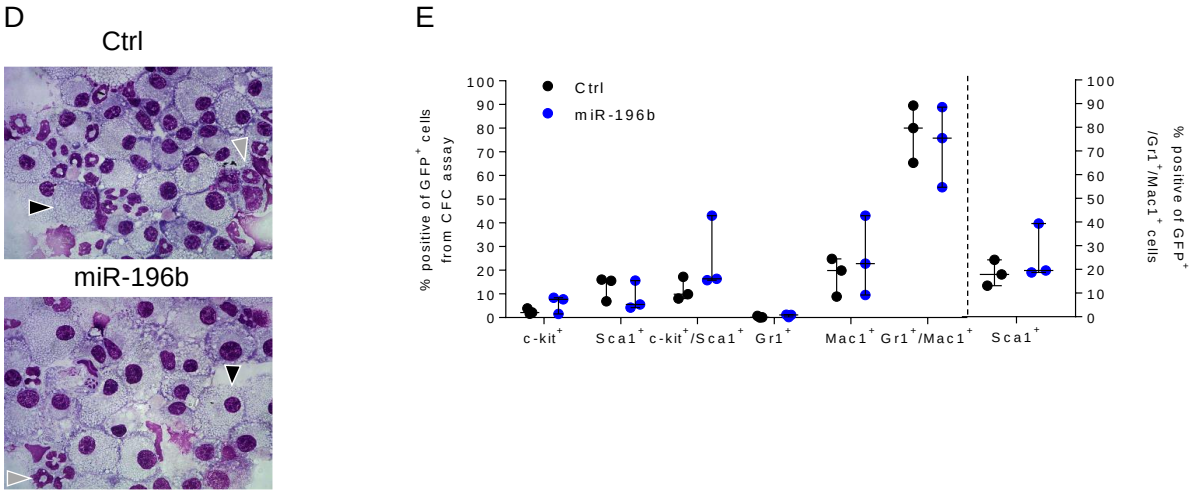
confirmed a myeloid immunophenotype – being very similar between miR-196b overexpressing and control cells (**Figure 16E**). Cells were predominantly (>75%) Gr1/Mac1 double positive from which over 1/4th were positive for the primitive cell marker Sca1 in both conditions. Notably, cells were not replatable.

When using the Ph1 vector as alternative to the LMP vector for miR-196b overexpression, murine HSPCs behaved as seen with LMP-miR-196b. The median total amount of colonies was not affected (difference <20%; Ctrl: median=114, 95% CI [56, 125] vs miR-196b: median=120, 95% CI [58, 141]; n=5) (**Figure 16F**), G colonies increased 1.2- fold without meeting statistical significance (p=0.0635) and the amount of GM colonies remained unaffected. The median total amount of cells per CFC assay was decreased by 33.33% [-5.26, 63.64], a difference which was statistically not significant (p=0.5476) (**Figure 16G**). Total colony and cell numbers demonstrated that colonies from cells with miR-196b overexpression were smaller with approximately 38% less cells per colony than colonies from control cells, which however did not meet statistical significance (Ctrl: median=14.4x10^3, 95% CI [5.36x10^3, 20.33x10^3] vs miR-196b: median=8.89x10^3, 95% CI [2.59x10^3, 16.67x10^3]; p=0.246; n=5).

In addition to the evaluation of cell numbers from CFC assays, cells transduced with LMP-miR-196b and the empty vector were plated into liquid culture medium and counted 7 and 14 days later. At both time points at least 93% of the cells were still GFP positive. Cell counts after 7 days revealed a reduction of LMP-miR-196b transduced cells by 74.5% [64, 87] compared to control cells without meeting statistical significance (p=0.125) (**Figure 16H**). 75% of miR-196b transduced and control cells co-expressed Gr1/Mac1, out of which about 1/3rd were Sca1 positive (**Figure 16I**). After 14 days, nearly all cells (96%) in both conditions co-expressed c-kit/Sca1 resembling mast cells (data not shown). Thus, miR-196b was not able to cease differentiation or transform murine HSPCs into a leukemic state after this long term culture.

Taken together, under *in vitro* conditions, miR-196b did not affect the frequency of clonogenic cells and their self-renewal/differentiation potential.





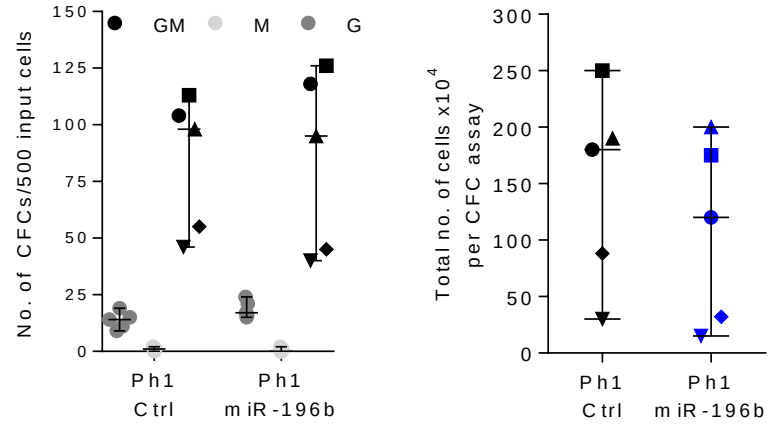


Figure continues on next page

Figure 16 continued

H

Liquid culture after 7 days

Number of miR-196b transduced cells in % to control

Exp. 1	Exp. 2	Exp. 3	Exp. 4	Mean	SD	Median
27%	13%	36%	26%	25.5%	9.47%	26.5%

I

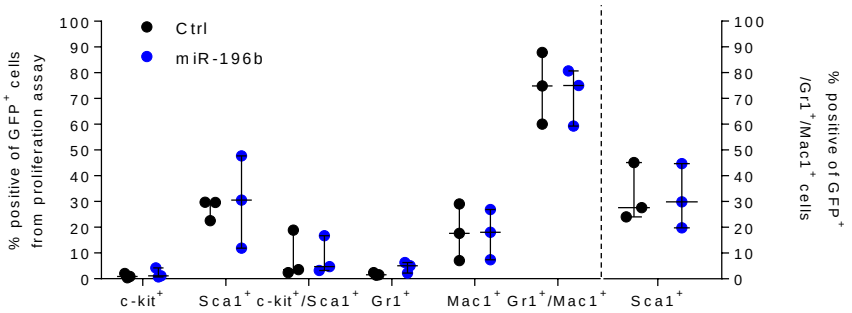


Figure 16. Effect of miR-196b overexpression in murine stem- and progenitor cells on clonogenicity, cell growth and differentiation state *in vitro*. (A-E) Analysis of LMP-miR-196b transduced murine hematopoietic stem- and progenitor cells (HSPCs) in colony forming cell assays (CFCs). **(A)** Number of colonies generated by 300 initially plated cells in methylcellulose, a semi-solid medium. Three different colony types were counted: G: granulocyte, M: megakaryocyte and GM: granulocyte/macrophage colony forming units. n=3. **(B)** Representative microscope photos of different colony morphologies with a 40x original magnification. **(C)** Total number of cells harvested from the CFC assay. n=3. **(D)** Representative microscope photos from cells on cytospins after May-Grünwald-Giemsa staining with a 600x original magnification. Black arrow indicates macrophage. Grey arrow indicates granulocyte. Figure extracted from Rawat & Götze et al. 2020, p. 287 **(E)** Immunophenotypic analysis for myeloid ($Gr1^+$ and $Mac1^+$) and primitive (c-kit^+ and $Sca1^+$) cell marker expression. **(F-G)** Analysis of Ph1-miR-196b transduced murine HSPCs in CFCs: **(F)** Number of colonies generated by 500 initially plated cells in methylcellulose. Three different colony types were counted: G: granulocyte, M: megakaryocyte and GM: granulocyte/macrophage colony forming units. n=5. **(G)** Total number of cells harvested from the CFC assay. n=5. **(F-G)** Statistical significance was determined using the Mann-Whitney *U* test. (**H**) Number of miR-196b transduced cells after 7 days in liquid culture medium in % to control. n=4. Statistical significance was determined using the one sample Wilcoxon Signed Rank Test. **(I)** Immunophenotypic analysis of LMP-miR-196b transduced murine HSPCs after 7 days in liquid culture medium for myeloid ($Gr1^+$ and $Mac1^+$) and primitive (c-kit^+ and $Sca1^+$) cell marker expression. n=3. In all scatter plots (**A, C, E, I**) individual values are presented with the median and 95% confidence interval. Symbol shape of black data points in A, C, G and F indicates data affiliation of each BM sample. No inference statistical test was applied unless otherwise indicated.

In vivo short-term repopulating capacity, lineage differentiation and transforming potential of murine HSPCs overexpressing miR-196b

The CFU-S assay, which was performed by injecting 10.000 LMP-miR-196b transduced cells per lethally irradiated mouse after *in vitro* culture for one week (ΔCFU-S), showed a significant (p=0.0175) 3-fold increase of colonies on the spleen compared to control spleens (**Figure 17**). Thus, miR-196b augments the short-term repopulating ability, however to a small extent, when compared to proto-oncogenes in leukemia models (Rawat et al. 2004, Cusan 2012).

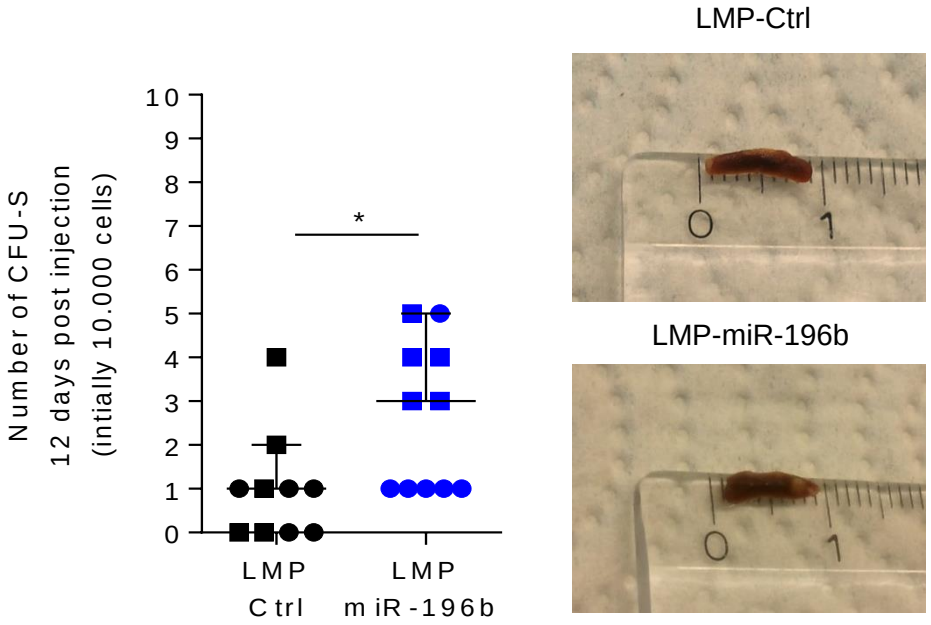


Figure 17. Effect of miR-196b on short-term repopulating ability *in vivo*. Total number of colonies on the spleen (CFU-S) 12 days post transplantation. 10.000 LMP-miR-196b or empty vector transduced hematopoietic stem and progenitor cells were kept in culture for 7 days and transplanted per lethally irradiated mouse. The experiment was conducted twice with always 5-6 mice per experimental condition. Presented are individual values with the median and 95% confidence interval. Shape of the data points indicates data affiliation to the independent experimental replicates. Statistical significance was determined using the Mann-Whitney *U* test. To the right representative spleen photos.

The PB analysis of mice transplanted with $1x10^6$ miR-196b transduced HSPCs together with $1x10^6$ untreated total BM cells, showed disturbed lineage specific cell distribution 40 weeks post transplantation (**Figure 18**). In evaluated mice, GFP^+ and GFP^- cells contributed to the differentiation into all lineages (GFP^+: median=59%, 95% CI [25, 82] vs median=41%, 95% CI [18, 75]; n=4). However, GFP^+ cells contributed mainly to myeloid cell differentiation ($Gr1^+$ and $Mac1^+$) by expressing 2.6 times more Gr1 and significantly (p=0.0286) 15- times more Mac1 compared to GFP^- cells. On the other hand, GFP^- cells contributed mainly to lymphoid cells with roughly two times more B- ($B220^+$ and $CD19^+$) and T-cell ($CD4^+$ and $CD8^+$) marker expressions of which the difference for $CD8^+$ cells was statistical significant (p=0.0286).

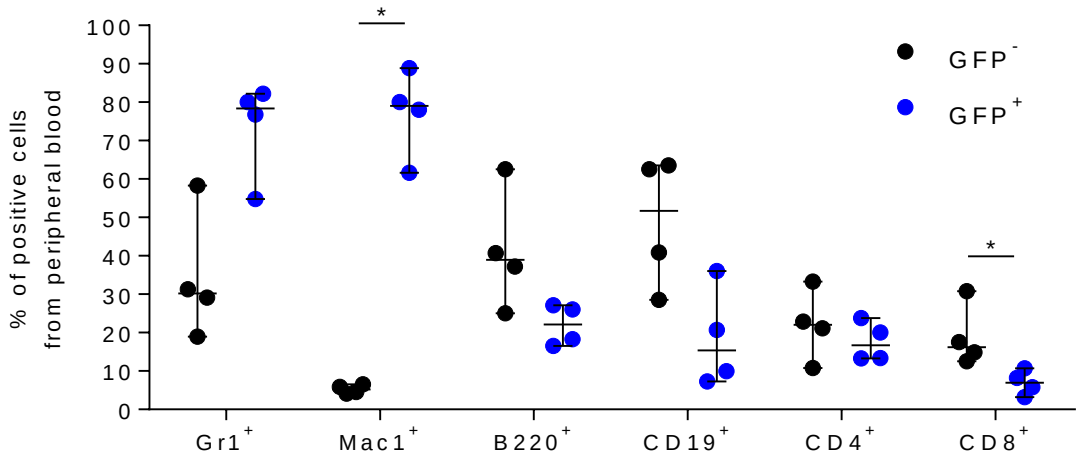


Figure 18. Effect of miR-196b on lineage differentiation potential *in vivo*. Immunophenotypic analysis of peripheral blood from four miR-196b mice in terms of myeloid (Gr1 and Mac1) and lymphoid (B-cell: B220 and CD19; T-cell: CD4 and CD8) cell marker, 40 weeks post transplantation, quantified with the FACS Calibur system. Presented are individual values with the median and 95% confidence interval. Statistical significance was determined using the Mann-Whitney *U* test

The long-term follow up of 15 transplanted mice revealed that the majority of mice (10/15) lived longer than 400 days. Only 5/15 mice became moribund after a long median latency of 314 days post-transplantation (**Table 30**). Out of these 5 mice, 3 mice were diagnosed with AML by an external histopathological analysis (#1, #2, #3), suggesting that miR-196b might have a mild transforming potential *in vivo*. One mouse (#4) was not leukemic but showed a shift towards more myeloid cells in PB with a lymphoid/myeloid ratio <2:1 (normal lymphoid/myeloid ratio: 4:1) and for another mouse (#5) no diagnosis relevant findings could be documented. The long latency for mouse #4 and #5 and the lack of AML indicates, that those mice became moribund for other reasons. Of note, also 1/4 control mice transplanted with $1x10^6$ empty vector transduced cells died after long latency (344 days).

Taken together, under *in vivo* conditions, miR-196b promotes myeloid skewing without being a major driver of myeloid leukemogenesis.

Table 30. Characteristics of diseased mice transplanted with miR-196b transduced murine HSPCs. Successfully transduced cells (GFP^+) were transplanted into mice, which were followed up for 400 days. Normal range for (a) white blood cell (WBC) counts: 1.8-$10.7x10^6$ and (b) red blood cell (RBC) counts: 6.36-$9.42x10^9$. PB: Peripheral Blood. NA: Not analyzed.

Mouse no.	Days of death post transplant	GFP signal in BM (%)	Diagnosis relevant findings
1	245	86.5	Histopathology diagnosed AML.
2	217	7.67	Histopathology diagnosed very primitive stage of tumor cells AML M0.
3	314	97.08	Histopathology diagnosed AML. Spleen weight: 465 mg

Table continues on next page

Table 30. *Continued*

Mouse no.	Days of death post transplant	GFP signal in BM (%)	Diagnosis relevant findings
4	391	97.80	WBC in PB: $80x10^6$/ml and RBC in PB: $2x10^9$/ml Spleen weight: 500 mg; Blast count in BM<20% Percentages of myeloid cells in PB: $Gr1^+$: 5.4%, $Mac1^+$: 23.2%, $Gr1^+/Mac1^+$: 30.2%
5	363	97.00	NA

In conclusion of *in vitro* and *in vivo* data, miR-196b mainly attenuates proliferation and acts pro myeloid differentiation in stem- and progenitor cells. Moreover, miR-196b does not transform murine HSPCs into a leukemic state.

3.5 Effect of aberrant miR-196b expression on differential gene expression in murine hematopoietic stem and progenitor cells

To investigate the genetic program underlying the miR-196b induced pro-myeloid differentiation phenotype on stem- and progenitor level, RNA from LMP-miR-196b transduced cells was paired-end sequenced and compared to results obtained in empty vector transduced murine HSPCs. MiR-196b overexpression led to 155 DEGs out of which 110 (71%) were up- and 45 (29%) genes downregulated (DEGs: p<0.00172, FDR<0.1, fold>1.15) (**Figure 19A**). Downregulated DEGs showed significant (p<0.001, FDR<0.001) miR-196b target gene enrichment, which was not the case for upregulated DEGs determined by the 'TargetScan microRNA 2017' database on the Enrichr platform. Out of 45 downregulated DEGs, 26 (58%) genes were identified as miR-196b targets (**Figure 19B**). Among the downregulated miR-196b specific target genes the RNA-binding protein musashi homolog 2 (*Msi2*), *Hoxa9* and lysine-specific demethylase 5A (*Kdm5a* or *Rbp2*) were significantly (p<0.001, FDR=0.0521) associated with an abnormal hematopoietic stem cell signature identified by the MGI Mammalian Phenotype level 4 database on the Enrichr platform and are involved in self-renewal as well as in differentiation according to literature (Sauvageau et al. 1994, Kharas et al. 2010, Klose et al. 2007).

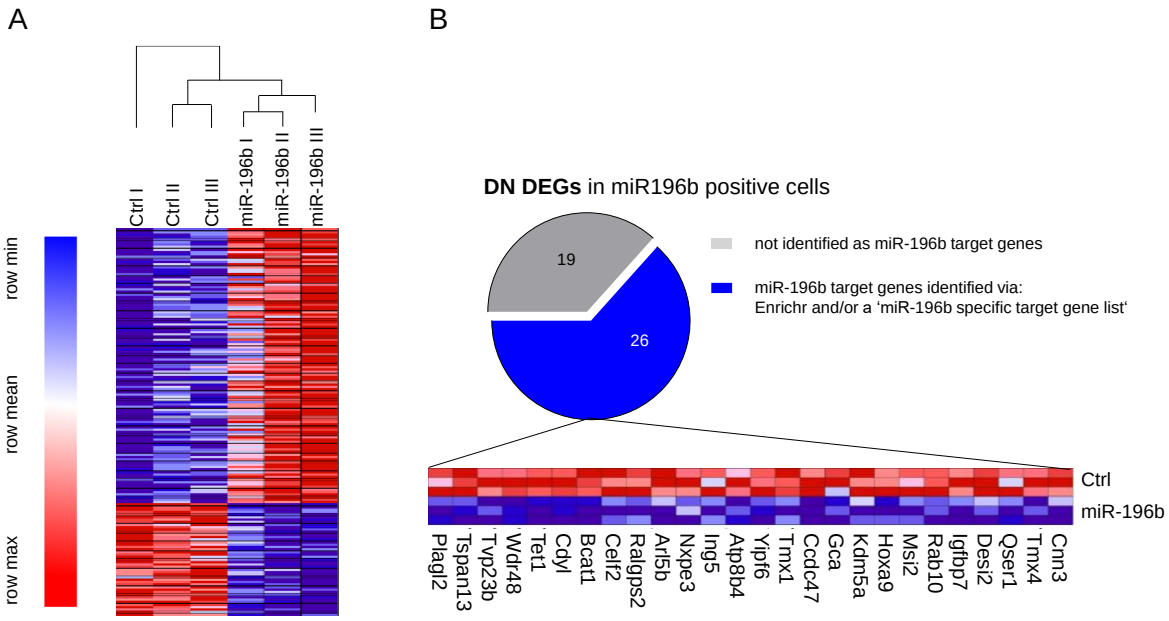


Figure 19. Effect of miR-196b overexpression in murine HSPCs on differential gene expression. RNA was isolated 96 h post transduction (48 h post sort) of murine bone marrow cells and paired-end sequenced. Differentially expressed genes (DEGs) were identified via DESeq2 provided by the Basepair platform. **(A)** Normalized count values of DEGs were hierarchically clustered and imaged as heatmap. FDR<0.1; n=3. **(B)** Pie chart depicts miR-196b target genes identified by the 'TargetScan microRNA 2017' database on the Enrichr platform and a list generated on the basis of four different miR-196b target databases (miRTarBase 7.0, TarBase 6.0-8.0, TargetScan 7.1 and miRDB). Genes and their normalized transcript variant 1 RNA-Sequencing expression count values are presented by a heatmap. **(A-B)** Heatmaps were generated using the GenePattern software. Low and high expression levels are indicated by blue and red color shades, respectively.

3.6 Associations between mature miR-196b, its transcript variants and neighboring *HOX* gene targets in AML cells

Although miR-196b did not affect the frequency of clonogenic cells and did not induce leukemic transformation in the mouse experiments presented above, data from the literature indicate co-expression of *HOXA* genes with miR-196b in human AML samples suggesting a functional collaboration between the two (Braekeleer et al. 2014). Thus, expression levels of miR-196b and specifically neighboring *HOXA* genes were analyzed and related to each other in AML bulk as well as in LSCs.

As expected, miR-196b and *HOXA* gene expression patterns were very similar in the different AML subtypes, analyzed from published miRNA- and mRNA-Seq datasets (The Cancer Genome Atlas) (**Figure 20 A-B**). Most highly expressed were miR-196b and *HOXA9* in CN-AML patients. In an additional qRT-PCR expression analysis, miR-196b and *HOXA9* were determined in CN-AML bulk cells (molecular subgroups: NPM1c^{+} and FLT3^{mt}) compared to healthy BM cells comprising BM MNCs and the matching healthy counterpart: BM CD34^{+} cells. *HOXA9* was strongly elevated in CN-AML bulk cells with significant 35- (p=0.0091) and 12- (p=0.0433) times higher expression values compared to BM MNCs or CD34^{+} cells, respectively, and irrespectively of the molecular subgroups (**Figure 20C**). Also miR-196b levels were elevated in CN-AML bulk cells with significant 17.6-times (p=0.0416) higher expression values when compared to BM MNCs, however, not when compared to

BM CD34^{+} cells (**Figure 20D**). Notably, miR-196b was significantly ~2-fold higher expressed in NPM1c^{+} vs FLT3^{mt} (NPM1c^{+}: median=3.647, 95% CI [1.69, 6.925] vs FLT3^{mt}: median=2.012, 95% CI [0.509, 3.634]; p=0.0257). Interestingly, the miR-196b transcript MF139051 was 4-fold (p=0.378) higher expressed in CN-AML bulk than in BM MNCs, however, without meeting statistical significance (**Figure 20E**).

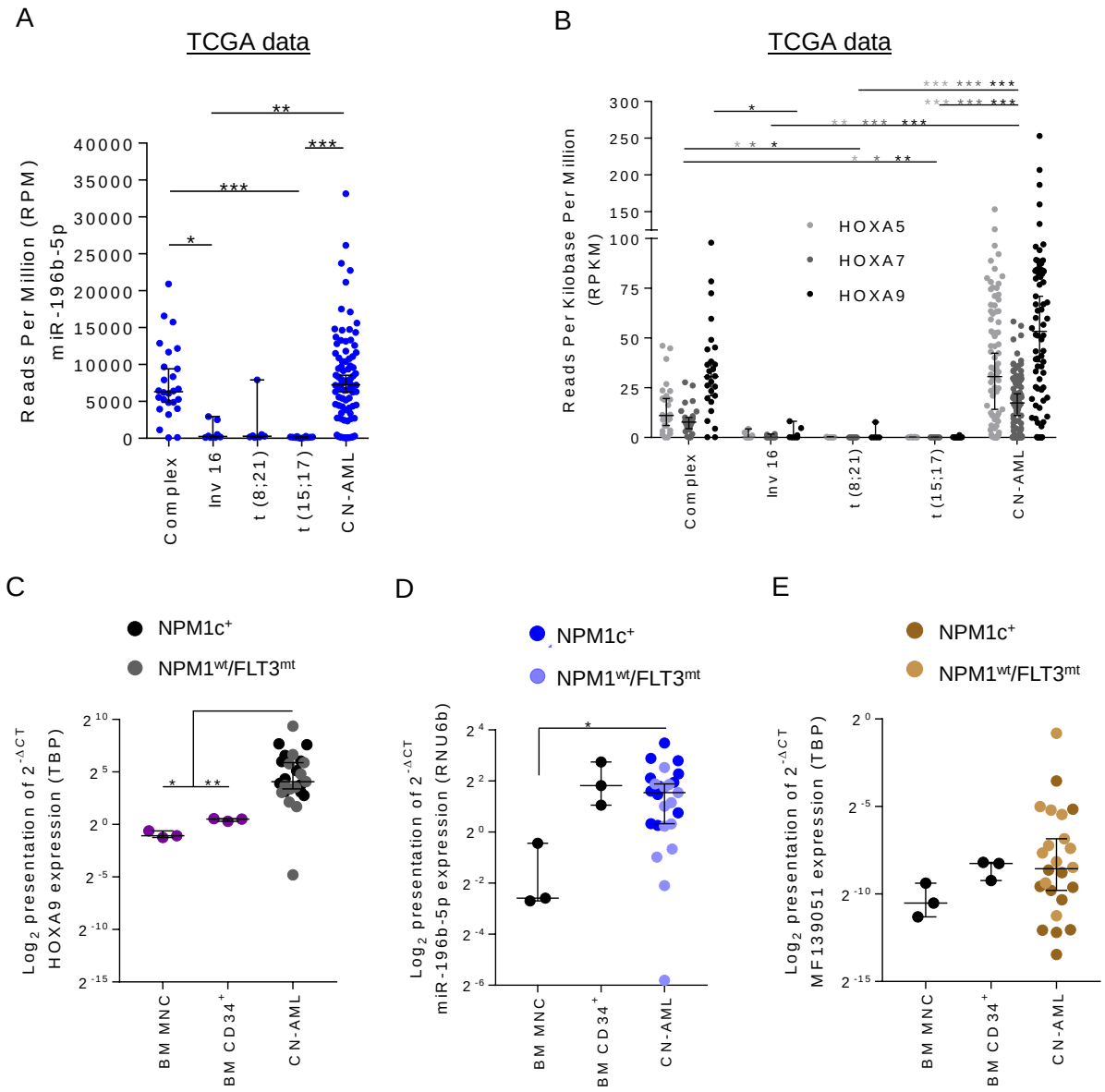


Figure 20. Expression of miR-196b transcript, its mature form and *HOXA* genes in AML bulk samples. (A-B) Expression of miR-196b-5p and *HOXA* genes in AML subtypes determined via miRNA- and mRNA-Sequencing, respectively (published data set from Cancer Genome Atlas Research et al. (2013)). According to the karyotype, patients were categorized into complex karyotype (Complex with ≥ 3 abnormalities, n=26), inversion 16 (Inv16, n=8), AML1-ETO (t(8;21), n=7), PML-RARA (t(15;17) without trisomy 8, n=12) and cytogenetically normal acute myeloid leukemia (CN-AML: n=86). Presented are individual values with the median and 95% confidence interval. Statistical significance was determined using the Kruskal-Wallis and Dunn's post hoc test. **(C-E)** Expression levels of (C) *HOXA9*, (D) miR-196b and (E) MF139051 in CN-AML patients compared to healthy bone marrow (BM) mononuclear (MNC) and CD34^{+} cells, determined via qRT-PCR, normalized on the housekeeping gene *TBP* for *HOXA9* and MF139051 or *RNU6b* for miR-196 and transformed into $2^{-\Delta CT}$-values. *Legend continues on next page.*

Figure 20. *Legend continued.* … NPM1c^{+} (n=12, patient no: 1-12) and FLT3^{mt}/NPM1^{wt} (n=13 for *HOXA9* and miR-196b, patient no: 16-24, 26-29; n=12 for MF139051, patient no: 16, 18-24; 26-29). BM CD34^{+} (n=3) and MNCs (n=3). Presented are individual values with the median and 95% confidence interval. Statistical significance was determined using the Kruskal-Wallis and Dunn's post hoc tests. Significant results between the control groups: 'BM MNC' and 'BM CD34^{+}' are not displayed. **(C-E)** For patient characteristics see Table 12 in the method section 2.1.4. C_T: Cycle Threshold; RNU6b (RNU6-6P): RNA, U6 small nuclear 6, pseudogene; TBP: TATA-box binding protein. NPM1c^{+}: Cytoplasmic nucleophosmin 11, FLT3 mutated (mt): FLT3-ITD/FLT3-TKD = fms-like tyrosine kinase receptor-3 /internal tandem duplication /tandem kinase domain.

The linear relation of *HOXA9* with MF139051 and mature miR-196b, which emerged in healthy BM cells, was lost in CN-AML bulk cells (**Figure 21 A-B**).

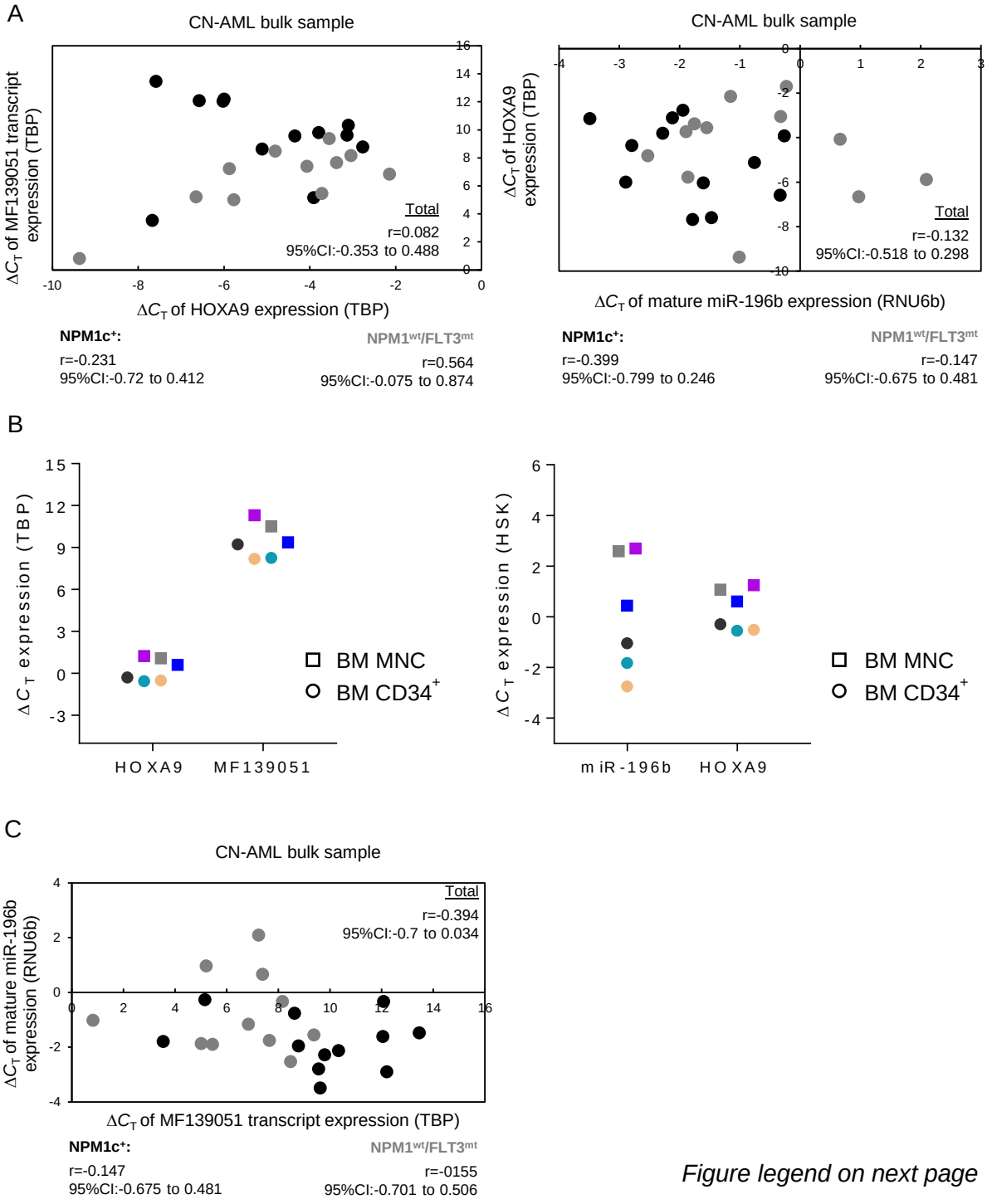


Figure legend on next page

Figure 21. Relation between *HOXA9* and miR-196b in AML bulk versus healthy BM samples. **(A-B)** ΔC_T-expression values of *HOXA9* with miR-196b transcript MF139051 (left) and mature miR-196b with *HOXA9* (right) in (A) CN-AML bulk and (B) healthy human bone marrow (BM) samples. Colors indicate data affiliations. **(C)** ΔC_T-expression values of MF139051 with miR-196b in CN-AML samples. $NPM1c^+$ (n=12, patient no: 1-12), $FLT3^{mt}/NPM1^{wt}$ (n=11 for MF139051:miR-196b, patient no:16, 18-24, 26, 28, 29; n=12 for *HOXA9*:miR-196b, patient no: 16-24, 26, 28, 29), healthy BM (n=6). **(A-C)** Data are depicted in scatterplots. For patient characteristics see Table 12. C_T: Cycle Threshold; RNU6b (RNU6-6P): RNA, U6 small nuclear 6, pseudogene; TBP: TATA-box binding protein. $NPM1c^+$: Cytoplasmic nucleophosmin 11, FLT3 mutated (mt): FLT3-ITD/FLT3-TKD = fms-like tyrosine kinase receptor-3 with internal tandem duplication or mutations in the tyrosine kinase domain.

In addition, the linear relation between transcript and mature miRNA observed in healthy BM cells, was also lost in CN-AML bulk samples (Figure 12B and **21C**).

In LSCs, the relation between *HOXA9* and miR-196b was similar to what was observed in CN-AML bulk samples. *HOXA9* was significantly 1.7 fold ($p<0.001$) higher expressed in LSCs ($CD34^+/CD38^+$) compared to healthy primitive cells, based on the analysis of published microarray data (**Figure 22A**) (Laurenti et al. 2013, Ng et al. 2016). However, miR-196b expression levels in LSCs, which were extracted from CN-AML samples, were not elevated. In fact, miR-196b was 1.5-3.6 times lower expressed in LSCs than in the matching healthy cells (**Figure 22B**). This result is based on the analysis of miRNA-Seq data generated in this lab by Dr. Susann Ihme, (Ihme 2014). In addition, while in healthy BM subpopulations of different differentiation stages the expression pattern of *HOXA9* and miR-196b was similar to each other, this expressional relation was lost in LSC subpopulations of different leukemia initiation potential (**Figure 22 C-D**). Thus, also in LSCs the correlation of *HOXA9* and miR-196b expression levels are likely to be disrupted as observed in AML bulk cells (Figure 21 A and C).

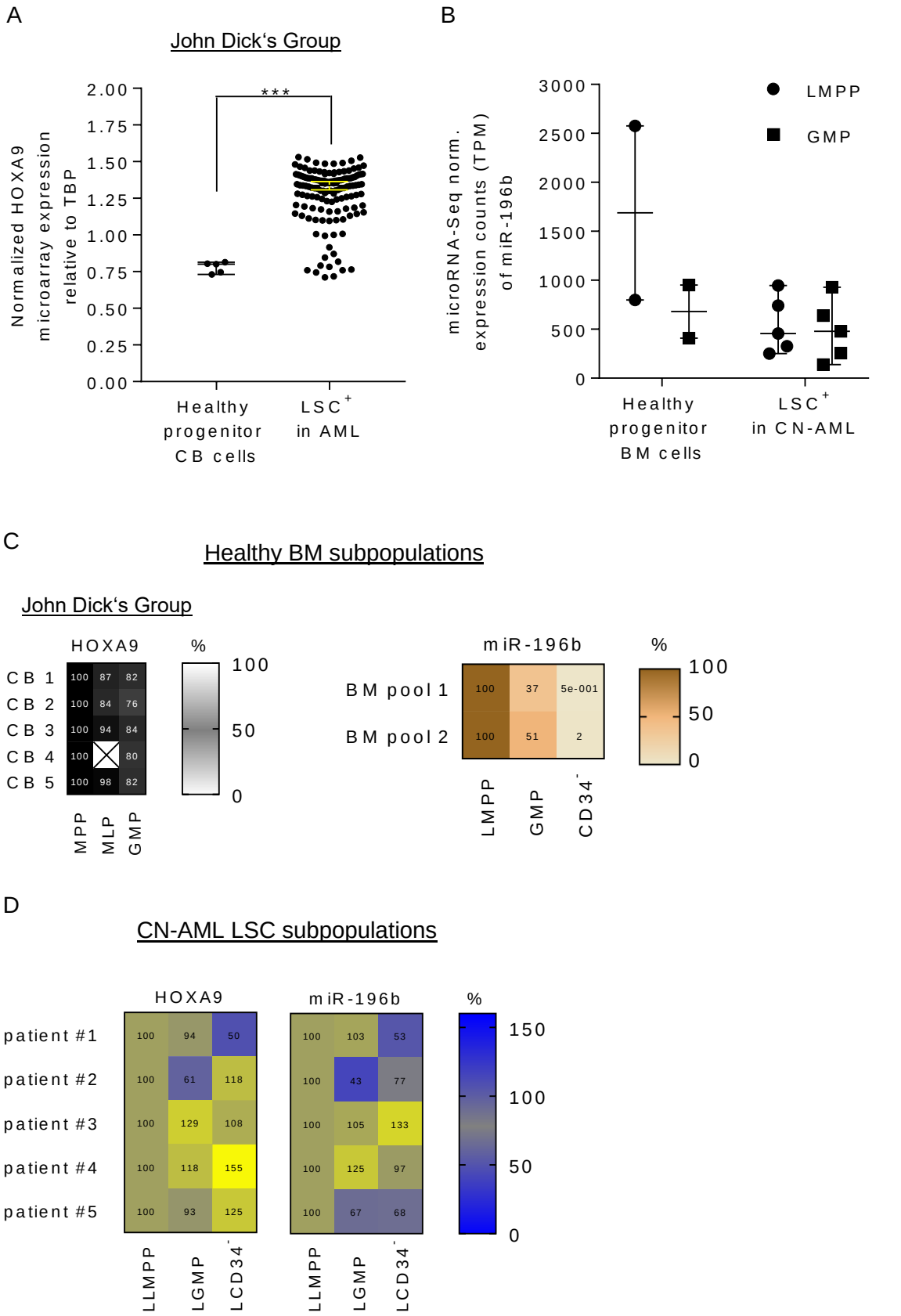


Figure 22. Expression levels of miR-196b and *HOXA9* in LSCs of AML patients. (A) Microarray expression values of *HOXA9* (ILMN_1702479) in functionally validated LSC⁺ (n=138) compared to healthy cord blood (CB) progenitor cells: (MPP, MLP and GMP, n=4-5) from the published microarray datasets GSE76009 (LSC) (Ng et al. 2016) and GSE4241 (CB) (Laurenti et al. 2013).

Legend continues on next page.

Figure 22. *Legend continued*. Presented are individual quantile normalized log2 transformed values relative to *TBP* with the median and 95% confidence interval. Statistical significance was determined using the Mann-Whitney *U* test. **(B)** MicroRNA-Seq expression values of miR-196b in functionally validated LSCs (leukemic LMPPs and GMPs) of cytogenetically normal acute myeloid leukemia (CN-AML) patients (n=5) compared to healthy directly matching counterparts (n=2 pools). Sequencing data in LSCs of AML samples was generated in the scope of Dr. Susann Ihme's thesis (Ihme 2014). For patient characteristics see Table 13. Presented are individual TPM values with the median and 95% confidence interval. No inference statistical test was applied. **(C) Left:** Heatmap visualizes a microarray based expression pattern of *HOXA9* (ILMN_1702479) in healthy progenitor cord blood (CB) subpopulations of decreasing differentiation stages from left to right retracted from the published microarray datasets GSE4241 (Laurenti et al. 2013). Presented are individual quantile normalized log2 transformed values relative to MPP in %. n=4-5. **Right:** Heatmap visualizes a miRNA-Seq based expression pattern of miR-196b in primitive bone marrow (BM) subpopulation pools. Presented are TPM values relative to LMPP in %. n=2 pools. Figure modified from Rawat & Götze et al. 2020, p. 286, CC BY-NC 4.0, https://creativecommons.org/licenses/by-nc/4.0 **(D)** MRNA-Seq and miRNA-Seq expression patterns of *HOXA9* and miR-196b, respectively, in leukemic LMPPs and GMPs as well as $CD34^-$ cells of CN-AML patients. For patient characteristics see Table 13. Presented are individual TPM values relative to LMPP in %. n=5.Fgiure modified from Rawat & Götze et al. 2020, p. 286, CC BY-NC 4.0, https://creativecommons.org/licenses/by-nc/4.0 (**A-D**) TPM: log-transformed transcript per kilobase million normalized count values; TBP: TATA-box binding protein. LMPP: lymphoid-primed multi-potential progenitor cells, GMP: granulocyte-monocyte progenitor, LSC: leukemic stem cell. MPP: multi-potent progenitors, MLP: multi-lymphoid progenitors.

3.7 Effect of aberrant miR-196b on clonogenicity, cell growth and leukemia associated transformation in murine HSPCs

Since miR-196b was not higher expressed in CN-AML, but was rather at the same level as in BM $CD34^+$ cells, it was hypothesized, that increased miR-196b levels might be able to attenuate the leukemic phenotype by targeting *HOX* genes, particularly *HOXA9* mRNA. To investigate this hypothesis, miR-196b stem-loop sequence was constitutively expressed in murine cells overexpressing *Cdx2*, known to induce *Hox* gene driven AML in mice as a single factor and to transform murine HSPCs *in vitro* (Rawat et al. 2004). Thus, miR-196b was co-overexpressed with *Cdx2* and compared to *Cdx2* alone in various *in vitro* and *in vivo* assays (**Figure 23**).

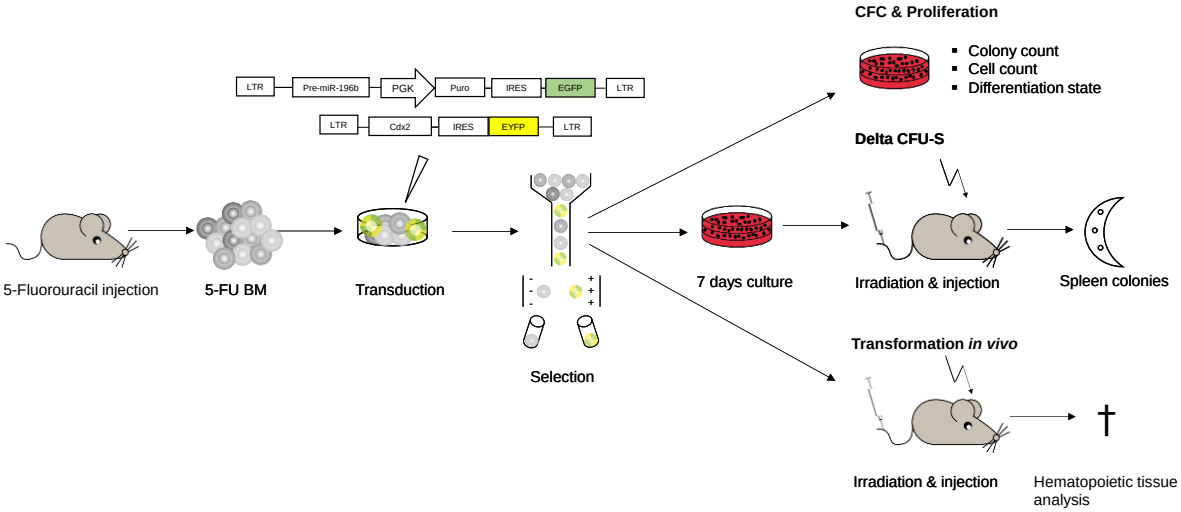


Figure 23. Experimental design for investigating the functional impact of miR-196b in a *Cdx2* induced leukemic mouse model. Bone marrow (BM) enriched for murine HSPCs via 5-Fluorouracil treatment was extracted and retrovirally transduced. After a FACS-based selection for reporter gene positive cells, *in vitro* and *in vivo* assays were performed.

In vitro, the effect of miR-196b on *Cdx2* induced enhanced clonogenic and cell growth potential was investigated via CFC assays in semi-solid medium and proliferation assays in liquid cultures, respectively. *In vivo*, first, the stem- and progenitor function was investigated via an evaluation of the colony forming units on the spleen (CFU-S) indicative for the short-term repopulating ability. Then, the ability of miR-196b to perturb *Cdx2*-induced leukemic cell transformation was investigated by following up *Cdx2* and *Cdx2*/miR-196b transplanted mice until disease or for 400 days and in case of disease animals were sacrificed and hematopoietic tissue analyzed.

In vitro clonogenic and cell growth potential as well as the cells' differentiation state of *Cdx2*/miR-196b co-overexpressing murine HSPCs

In primary CFC assays, a stable overexpression of miR-196b in the context of *Cdx2* (miR-196b relative to *Cdx2*: 3707-fold (*RNU19*) or 7218-fold (*RNU6b*); Cdx2 relative to empty vector control: 255471-fold (*mHprt*)) reduced the number of colonies in each experiment (**Figure 24A)** compared to *Cdx2* alone. Colonies were reduced by 9% to 40% with a median reduction by 27%. Replating cells from the primary and secondary CFC assay revealed a median reduction of 80% and 91%, respectively, demonstrating that miR-196b overexpression led to progressive colony decrease (**Figure 24 B-C**). Harvested cells were over 92% *Cdx2*-YFP or *Cdx2*-YFP/miR-196b-GFP positive and the median total amount of *Cdx2*/miR-196b positive cells from the primary, secondary and tertiary CFC assays decreased by 48% [12.07, 62.05], 81% [79.52, 96.55], and 98.71% [92.15, 99.94], respectively (**Figure 24 D-F**). Total colony and cell numbers implicated that colonies from *Cdx2*/miR-196b co-overexpressing cells were smaller than from only *Cdx2* expressing cells. Indeed, *Cdx2*/miR-196b co-expressing cells contained approximately 42%, 10% and 64% less cells per colony in the first, second and third plating, respectively, based on absolute numbers without extrapolations (Cells per colony for *Cdx2* vs *Cdx2*/miR-196b in 1st CFC assay: median=27.08x10^3, 95% CI [18.05, 33.9] vs median=15.61x10^3, 95% CI [14.05, 32.79]; 2nd CFC assay median=8.81x10^3, 95% CI [8.62, 9.89] vs median=7.96x10^3, 95% CI [3.15, 8.89] and 3rd median=10.23x10^3, 95% CI [6.36, 11.36] vs median= 3.66, 95% CI [1.18, 5.75]; n=3).

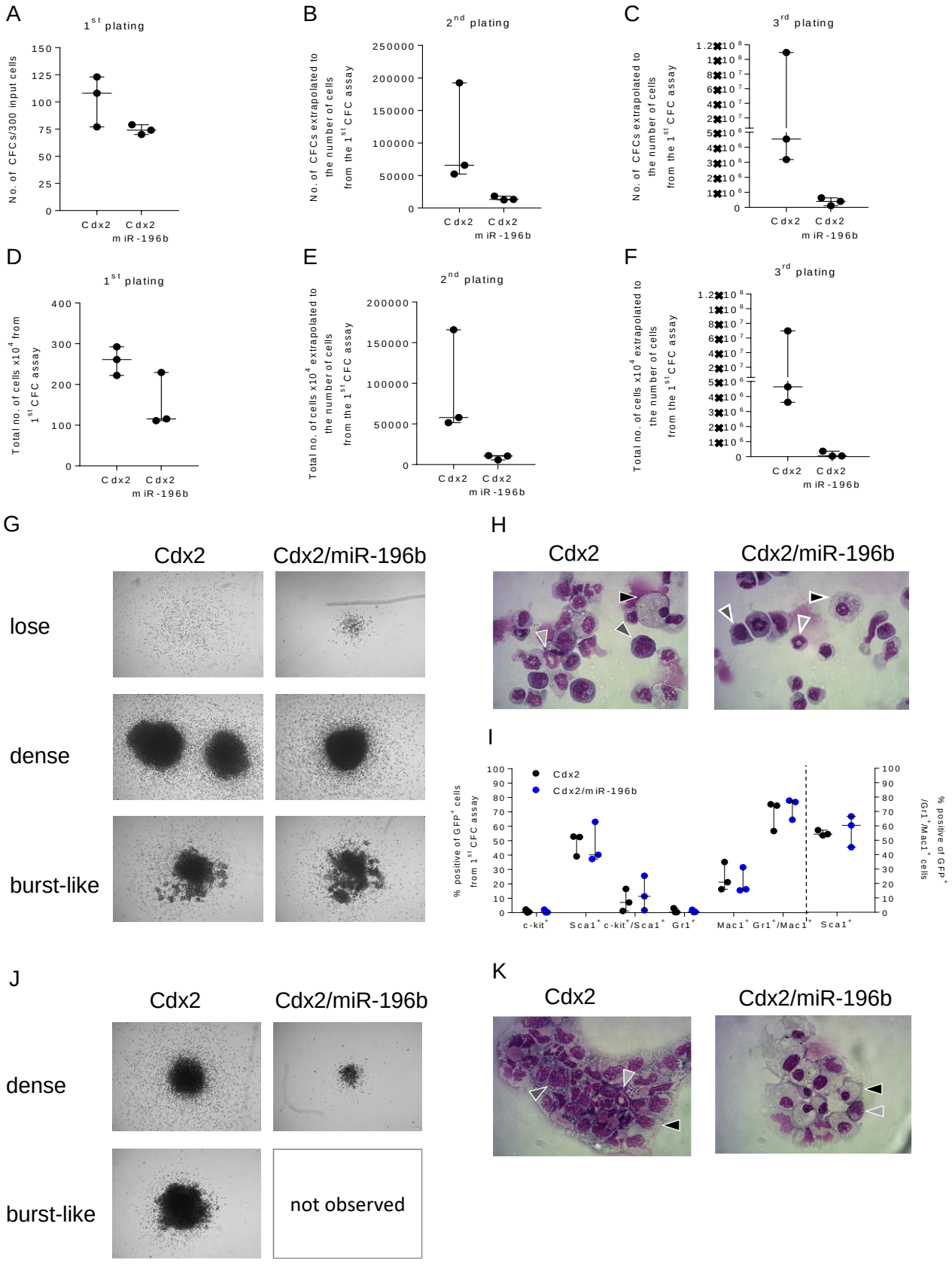


Figure legend on next page

Figure 24. Effect of *Cdx2*/miR-196b overexpression murine stem- and progenitor cells on clonogenicity and differentiation state of *in vitro*. (A-C) Total number of colonies and **(D-F)** total number of cells harvested from colony forming cell (CFC) assays generated by 300 initially plated murine hematopoietic stem and progenitor cells (HSPCs) constitutively expressing *Cdx2* or *Cdx2*/miR-196b. n=3. **(G-I)** Colony morphology and cell type analysis in primary CFC assays. **(G)** Representative microscope photos of different colony morphologies with a 40x original magnification. **(H)** Representative microscope photos from cells on cytospins after May-Grünwald-Giemsa staining with a 600x original magnification. Black arrow indicates macrophage. Dark grey arrow indicates primitive myeloid cell. Light grey arrow indicates granulocyte. **(I)** Immunophenotypic analysis for myeloid ($Gr1^+$ and $Mac1^+$) and primitive (c-kit^+ and $Sca1^+$) cell marker expression. n=3. **(J-K)** Colony morphology and cell type analysis in tertiary CFC assays. (J) Representative microscope photos of different colony morphologies with a 40x original magnification. Dense colony pictures extracted from Rawat & Götze et al. 2020, p. 287 (K) Representative microscope photos from cells on cytospins after May-Grünwald-Giemsa staining with a 600x original magnification. Black arrow indicates macrophage. Dark grey arrow indicates primitive myeloid cell. Light grey arrow indicates granulocyte. In all scatter plots (**A-F, I**) individual values are presented with the median and 95% confidence interval. No inference statistical tests were applied.

Colony morphology in the primary CFC assays was not affected by *Cdx2*/miR-196b co-expression compared to *Cdx2* alone (**Figure 24G**). In both conditions, GM colonies formed and appeared prominently as tight (dense) and less frequent as dispersed ('loose') or burst-like colonies. Additionally, May-Grünwald Giemsa stainings of cells from the CFCs revealed terminally differentiated macrophages and granulocytes, but also premature myeloid cells in both conditions (**Figure 24H**). This was specified by a cell surface marker analysis, which confirmed a very similar myeloid immunophenotype in *Cdx2* alone and miR-196b co-expressing cells (**Figure 24I**). Cells were predominantly (>74%) Gr1/Mac1 double positive from which more than $3/4^{th}$ were positive for the primitive cell marker Sca1 in both conditions. Also in the secondary CFC assays, colony morphology, cell type and immunophenotype were not affected (data not shown) by miR-196b co-overexpression, suggesting a selection of *Cdx2* transformed clonogenic cells. Only after the third plating, miR-196b co-overexpression changed the colony morphology visibly (**Figure 24J**). While in the 3rd CFC assay *Cdx2* overexpressing cells formed prominently dense and less frequent burst-like or loose colonies, *Cdx2*/miR-196b co-overexpressing cells formed only small dense colonies. Moreover, May-Grünwald Giemsa stainings of cells from the 3rd CFC assay revealed that in addition to terminally differentiated macrophages and granulocytes in both conditions, premature myeloid cells were no longer visible upon miR-196b co-overexpression (**Figure 24K**).

In addition to the evaluation of cell numbers from CFC assays, cells transduced with *Cdx2* and *Cdx2*/miR-196b were plated into liquid culture medium and counted 7 and 14 days later. At both time points at least 95% of the cells were YFP or YFP/GFP positive. Cell counts after 7 days revealed a reduction of *Cdx2*/miR-196b co-transduced cells by 29.97% [23.17, 48.51] compared to *Cdx2* alone (Cell numbers on day 7 in % for *Cdx2*: median=100%, 95% CI [100, 100] vs *Cdx2*/miR-196b: median=70.03%, 95% CI [51.49, 76.83], n=3) (**Figure 25A**). More than 80% of *Cdx2* and *Cdx2*/miR-196b transduced cells co-expressed Gr1/Mac1, out of which over 85% were Sca1 positive (**Figure 25B**). After fourteen days, the effect of miR-196b in the context of *Cdx2* was reinforced with cell numbers reduced by 76% [70.5, 80.9] compared to *Cdx2* alone (Cell numbers on day 14 in % for *Cdx2*: median=100%, 95% CI [100, 100] vs *Cdx2*/miR-196b: median=24.3%, 95% CI [19.1, 29.5], n=2). Both *Cdx2* alone and miR-196b co-expressing cells had a primitive myeloid immunophenotype with

about 50% Mac1 and 50% Gr1/Mac positive cells out of which over 77% were Sca1 positive (data not shown).

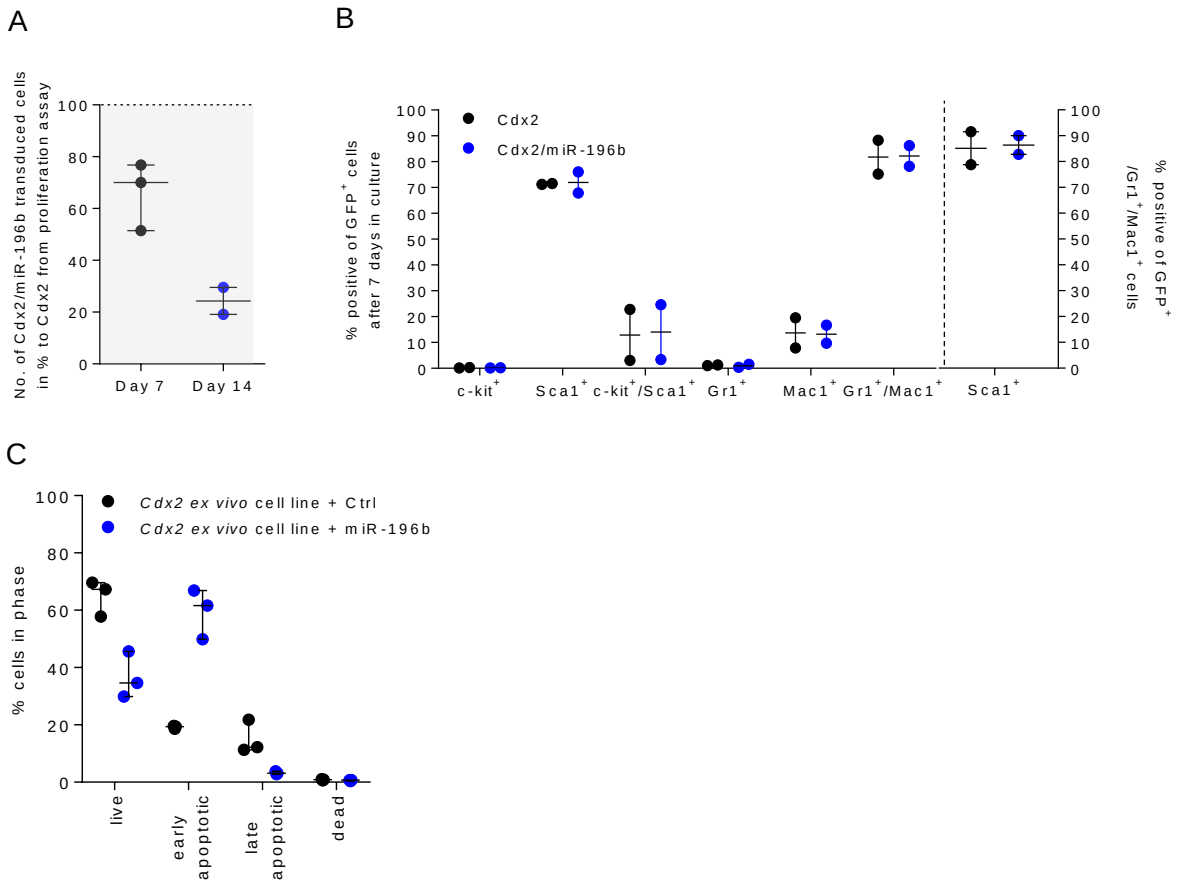


Figure 25. Effect of miR-196b overexpression in the context of *Cdx2* on cell growth, differentiation state and apoptosis of murine stem and progenitor cells in liquid cultures. (A) Number of miR-196b transduced cells after 7 (n=3) and 14 (n=2) days in liquid culture medium in % to control. No inference statistical test was applied. **(B)** Immunophenotypic analysis of *Cdx2*/miR-196b co-transduced murine hematopoietic stem and progenitor cells (HSPCs) compared to *Cdx2* single transduced cells after seven days in liquid culture medium. Analyzed were myeloid ($Gr1^+$ and $Mac1^+$) and primitive ($c\text{-}kit^+$ and $Sca1^+$) cell markers. n=2. **(C)** Apoptosis was investigated via Annexin V and 7-amino-actinomycin D (7AAD) stainings allowing to distinguish four apoptotic stages: live ($AnnexinV^-/7AAD^-$), early apoptotic ($AnnexinV^+/7AAD^-$), late apoptotic ($AnnexinV^+/7AAD^+$) and dead ($AnnexinV^-/7AAD^+$). n=3. **(A-C)** Presented are individual values with the median and 95% confidence interval. No inference statistical test was applied.

To investigate, whether increased apoptosis could be an underlying biological process for a reduced clonogenic and cell growth potential of *Cdx2*/miR-196b co-overexpressing murine HSPCs, miR-196b was stably overexpressed in a *Cdx2*-induced leukemic murine BM cell line, established by an IL-3 dependent long-term culture, and analyzed via flow cytometry. Annexin V was stained to detect phosphatidylserine, which, if detected on the membrane, is a marker for apoptosis and 7-amino-actinomycin D (7AAD) was stained to detect DNA, by which viable cells can be distinguished from dead cells. MiR-196b caused apoptosis in *Cdx2* transformed cells by increasing the proportion of early apoptotic cells three times, while the proportion of living cells was reduced twice (**Figure 25C**).

In vivo short-term repopulating and leukemia associated transformation potential of murine HSPCs overexpressing *Cdx2*/miR-196b

The CFU-S assay, which was performed by injecting 10.000 *Cdx2*/miR-196b transduced cells per lethally irradiated mouse after *in vitro* culture for one week (ΔCFU-S), showed a significant (p=0.012) 6-fold decrease of colonies on the spleen compared to *Cdx2* alone (**Figure 26A**). Next, transplanted *Cdx2*/miR-196b mice diseased with a significantly (p=0.0413) longer median latency than *Cdx2* mice (**Figure 26B**). While 8/10 *Cdx2* transplanted mice developed AML with a median latency of 193 days, only 4/10 *Cdx2*/miR-196b transplanted mice developed AML with a median latency of 246 days in the follow up period of 400 days (**Table 31**). The majority of *Cdx2*/miR-196b mice (6/10) did not become moribund. This suggests, that miR-196b has a great potential to attenuate leukemia development *in vivo*.

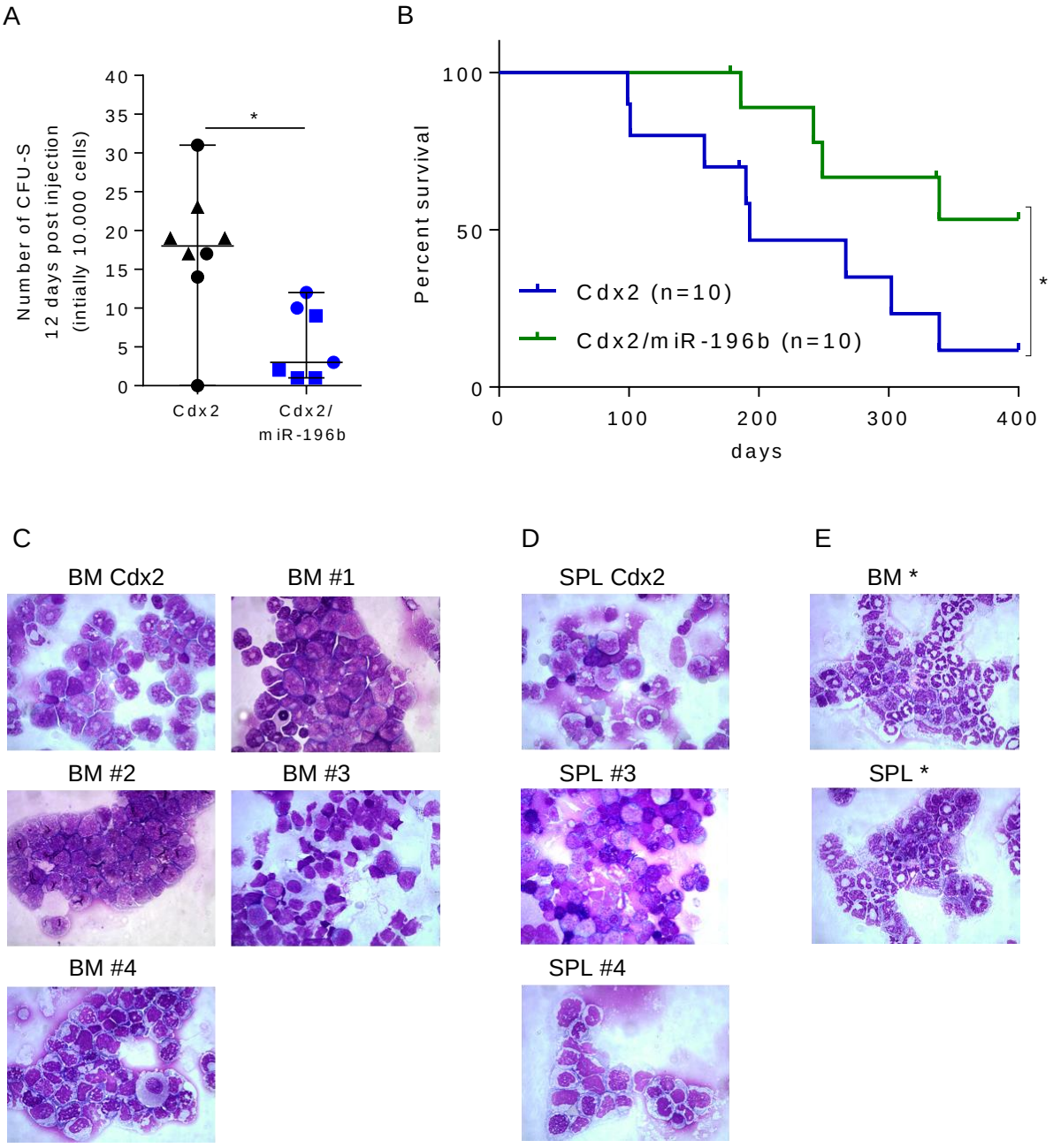


Figure legend on next page

Figure 26. Effect of *Cdx2*/miR-196b on short-term repopulation and leukemia associated transformation *in vivo*. (A) Total number of colonies on the spleen (CFU-S) 12 days post transplantation. 10.000 *Cdx2* or *Cdx2*/miR-196b transduced hematopoietic stem and progenitor cells (HSPCs) were kept in culture for 7 days and transplanted per lethally irradiated mouse. ΔCFU-S assays were conducted with three independently produced cell suspensions transplanted into 3-4 mice per experimental condition. Presented are individual values with the median and 95% confidence interval. Shape of the data points indicates data affiliation to the independent experimental replicate. Statistical significance was determined using the Mann-Whitney *U* test. **(B)** Survival cures of mice transplanted with *Cdx2* (n=10) or *Cdx2*/miR-196b (n=10) transduced HSPCs. Censoring criteria: alive on ≥400 days, diseased on <400 days, but no engraftment on the day of sacrifice or no leukemia related death. The experiment was conducted 4x with always 2-3 mice per experimental condition. Statistical significance was determined using log-rank test. Figure modified from Rawat & Götze et al. 2020, p. 287, CC BY-NC 4.0, https://creativecommons.org/licenses/by-nc/4.0 **(C-E)** Representative microscope photos of May-Grünwald stained (C) bone marrow (BM) and (D) spleen (SPL) cells from leukemic (#1-4) or (C) non-leukemic (*) mice on cytospins with a 600x original magnification. Numbers refer to the mouse number used in Table 31.

Table 31. Characteristics of leukemic mice transplanted with *Cdx2*/miR-196b co-transduced murine HSPCs. Transduced cells were transplanted and mice were followed up for 400 days. *Expression levels of miR-196b in YFP^+/GFP^+ bone marrow (BM) cells of *Cdx2*/miR-196b transplanted mice were determined via qRT-PCR, normalized on the housekeeping gene *RNU6b*, calculated relative to BM cells from sick mice overexpressing *Cdx2* and transformed into $2^{-\Delta\Delta CT}$-values. Normal range for (a) white blood cell (WBC) counts: 1.8-10.7x10^6 and (b) red blood cell (RBC) counts: 6.36-9.42x10^9. PB: Peripheral Blood. NA: Not analyzed. YFP is the reporter gene for *Cdx2*. GFP is the reporter gene for miR-196b.

Mouse no.	Day of death post transplant	GFP/YFP signal in BM (%)			$2^{-\Delta\Delta CT}$-values of miR-196b*	Diagnosis relevant findings			
		YFP	GFP	YFP+ GFP		WBC in PB x10^6/ml	RBC in PB x10^9/ml	Spleen weight mg	Blasts in BM >20%
#1	242	0.1	0.9	97.5	1	NA	NA	NA	YES
#2	186	0	0.8	99.2	6.7	139.28	1.59	1317	YES
#3	249	2.2	0.1	49.3	4000	13.38	8.03	432	YES
#4	339	0.1	0.5	98.2	7000	32.5	2.54	1759	YES

80% of *Cdx2* mice sacrificed for hematological analysis had blast counts greater than 20%, elevated peripheral white blood cell counts (WBCs) with a median number of 50.5x10^6 [16.94, 151] circulating WBC/ml (n=5) and reduced peripheral red blood cell counts (RBCs) with a median number of 4.41 x10^9 [0.56, 8.64] circulating RBC/ml (n=5). All *Cdx2* mice analyzed (n=7) suffered from splenomegaly with a median spleen weight of 714 mg [409, 1235]. 40% of *Cdx2*/miR-196b mice had more than 20% blast counts in the BM and as determined for mouse #3 and #4 an overrepresentation of myeloid cells in the spleen, a normally lymphoid organ (**Figure 26 C-D**). Moreover, all diseased *Cdx2*/miR-196b mice analyzed for hematological cell type composition in PB had an elevated median of WBCs (median=32.5x10^6/ml, 95% CI [13.38, 139.3], n=3), reduced median of RBCs (median=2.54x10^9/ml, 95% CI [1.59, 8.03], n=3) and splenomegaly (median= 1317, [432, 1759], n=3). Notably, survival was highest when miR-196b expression levels were high and

decreased as miR-196b expression levels were decreased in 3/4 AML mice. The majority (6/10) of *Cdx2*/miR-196b mice did not develop cause-specific AML. Four lived longer than 400 days, one became moribund after 337 days without reporter gene signals in the BM and another one after 178 days with reporter gene signals in the BM, but without elevated blast counts in the BM. The latter mouse with 39% Cdx2-YFP, 3% miR-196-GFP and 58% YFP/GFP signal in the BM showed an increased number of WBCs ($145.72x10^6$/ml) in PB together with reduced RBC counts ($4.85x10^9$/ml) compared to the normal range ($6.36-9.42x10^9$) and its spleen was enlarged (718mg) consisting of an unusually high proportion of neutrophils as also seen in the BM (**Figure 26E**).

In conclusion of *in vitro* and *in vivo* data, miR-196b has the potential to counteract *Cdx2* induced high *HOX*-gene driven leukemic transformation and therewith can prevent AML development in most cases.

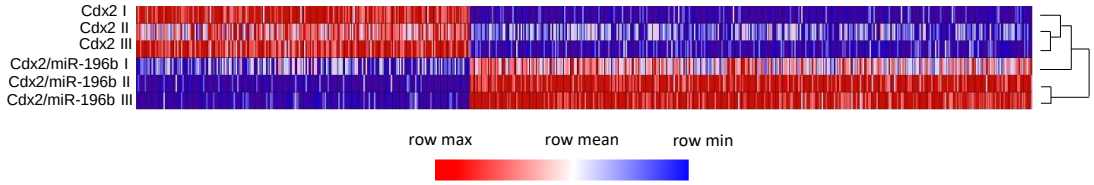
3.8 Effect of aberrant miR-196b expression on differential gene expression in murine leukemic cells

To understand the molecular mechanism through which aberrantly high miR-196b expression counteracts *Cdx2* induced leukemic transformation *in vitro* and i*n vivo*, RNA from *Cdx2*/miR-196b co-transduced HSPCs was paired-end sequenced and compared to results obtained in control cells. First, the generally altered genetic program was determined by identifying DEGs. Then, downregulated DEGs were specified to miR-196b targets by comparing them to a 'miR-196b specific target gene' list, which was created on the basis of four different miRNA target databases including experimentally validated and computed miR-196b targets (*see method section 2.2.24 section 'MiR-196b specific target gene list determination'*). Biological insights of generally and specifically altered genes were gained by searching for significant gene set enrichments in human disease and cellular process pathways. Finally, most phenotype relevant miR-196b specific candidate genes were postulated on the basis of gene list comparisons, gene function and expression values in CN-AML patients.

RNA-Seq revealed, that miR-196b in the context of *Cdx2* induced 524 DEGs out of which 329 (63%) genes were up- and 195 (37%) genes downregulated ($p<0.001$, FDR<0.1, fold>1.16) (**Figure 27A**). KEGG pathway analysis showed significant ($p<0.001$, FDR<0.06, score ≥ 10 genes) enrichment of upregulated genes in 'pathways in cancer' and 'apoptosis' (**Figure 27B**). Among apoptotic genes, all were pro-apoptotic, except *Rela*, which is mainly known to protect cells from apoptosis but its activation in the non-canonical pathway results in a moderate pro-apoptotic effect (Xiu et al. 2018, Zhou et al. 2015). Among the cancer genes there were a number of genes such as *Cdkn1a (p21)*, *Spi1 (Pu.1)*, *Dapk1*, *Hhip*, *Ednrb* and *Rxra* that can counteract AML (Zhang et al. 2015, Cook et al. 2004, Karlic et al. 2014, Kobune et al. 2012, Hsiao et al. 2008, Sanchez et al. 2014). In addition to the KEGG-based cancer related genes, the tumor suppressor gene *Fas* was also elevated in expression (Li, Huang, et al. 2012). Interestingly, downregulated DEGs after a less (FDR<0.3) and more (FDR<0.1) stringent cut off, showed significant ($p<0.001$, FDR<0.001) miR-196b target gene enrichment, which was not the case for upregulated DEGs, determined by the 'TargetScan microRNA 2017' database on the Enrichr platform (data not shown). A comparison of a 'more' and 'less' stringent downregulated DEG list to the 'miR-196b specific target gene' list revealed 31 and 98 miR-196b targets, respectively (data not shown). Reactome pathway analysis showed significant ($p<0.002$, FDR>0.25, score >2

genes) enrichment of those 98 miR-196b specific downregulated DEGs in pathways of ‘translation’, ‘senescence’, ‘cell cycle’, ‘transcription’ and ‘MAP kinase’ (**Figure 27C** top). Gene ontology database for molecular function 2015 confirmed the enrichment in ‘MAP kinase’- and ‘transcription’ pathways with statistical significance ($p<0.0021$, FDR>0.25) (**Figure 27C** middle). In line, a GSEA software based analysis of miR-196b specific target genes revealed enrichment in a ‘proliferation’ signature (**Figure 27C** bottom). Additionally, the GSEA software-based analysis of all RNA-Seq values showed significant (with $p<0.05$ and FDR<0.25) enrichment of up and downregulated genes in cancer relevant pathways (**Figure 27D**). Genes, which were upregulated in *Cdx2*/miR-196b transduced cells were enriched in 27 ‘oncogenic gene sets’. Two of the top five gene sets are associated with a pro-oncogenic phenotype: ‘RAPA_EARLY_UP.V1_DN’ in B-lymphoma cells and ‘TGFB_UP.V1’ in a panel of lung and breast epithelial cells (Peng et al. 2002, Padua et al. 2008). Interestingly, the proto-oncogene *Jun* and *Junb* was among the enriched genes of the ‘TGFB_UP.V1’ gene set (downregulated DEGs in RNA-Seq data with FDR<0.1). Notably 3/5 oncogenic gene sets are associated with an anti-cancer phenotype: ‘HINATA_NFKB_IMMU_INF’ in keratinocytes and dermal fibroblasts, ‘STK33_NOMO_UP’ in AML cells and ‘PRC2_EED_UP.V1_DN’ in embryonic fibroblasts (Hinata et al. 2003, Scholl et al. 2009, Bracken et al. 2006). Genes, which were downregulated upon miR-196b in the context of *Cdx2*, were enriched in 6 pro-oncogenic gene signatures such as *Hoxa7* in the MTOR pathway in T-CLL cells or *Tgfbr3* after knockdown of the tumor suppressor PTEN in cells of skin, lung and breast cancers (Wei et al. 2006, Vivanco et al. 2007). Downregulated genes upon miR-196b overexpression in the context of *Cdx2* were also enriched in leukemia-related gene sets. 19 downregulated genes were enriched in an LSC and 37 downregulated genes in a normal stem cell gene signature (LSC: p/FDR=0.0074, NES=-1.47 and HSC: p/FDR=0.0067, NES=-1.31) (**Figure 27 E-F**) (Ng et al. 2016, Gentles et al. 2010, Eppert et al. 2011).

A



B

Upregulated DEGs

Human KEGG 2016	Genes
Pathways in cancer	FZD1, LAMA5, JUN, CDKN1A, SPI1, DAPK1, HHIP, MMP2, FOS, LAMC1, RELA, EDNRB, RXRA, GNAS, ARHGEF1, PPARD
Osteoclast differentiation	JUN, SPI1, CTSK, NFATC2, CYBA, TREM2, FOS, JUNB, MAP3K14, RELA, TNFRSF1A
Apoptosis	JUN, CTSK, LMNA, CAPN1, FOS, MAP3K14, CTSD, RELA, BBC3, TNFRSF1A

Downregulated DEGs

Ribosome	RPS25, RPS27, RPL21, RPL31, RPL36A, RPL22L1, RPL26, RPS27A RPL39, RPS24, RPL7, MRPL33

C

Downregulated miR-196b specific target genes

Reactome 2016	Genes
SRP-dependent cotranslational protein targeting to membrane & Translation	RPS25, RPL27A, RPS3, SSR1, SRP9
Senescence-Associated Secretory Phenotype & Cellular senescence	CDKN1B, UBE2E1, MAPK1, IGFBP7 & MAPK8
MAP2K and MAPK activation	NRAS, RAP1A, MAPK1
TP53 Regulates Transcription of Cell Cycle Genes	CCNB1, NPM1, CDKN1B

GO Molecular Function 2015	Genes
Transcription coactivator activity	YY1, KDM5A, NPAT, DDX5, NPM1, NR2C2, YAF2
MAP kinase activity	MAPK8, MAPK1

GSEA c5.bp.v6.2.symbols	Genes
Positive regulation of cell proliferation	NRARP, DPP4, HMGA2, ACER2, TET1, CDKN1B, PPP1R16B, S1PR1, TGFBR3, SDCBP, VASH2, SSR1, PBX1, NPM1, BMPR1A, CCNB1, RPS3, NAP1L1, SOX11, AKIRN2, NRAS, CNOT7, GATA6, ANXA1

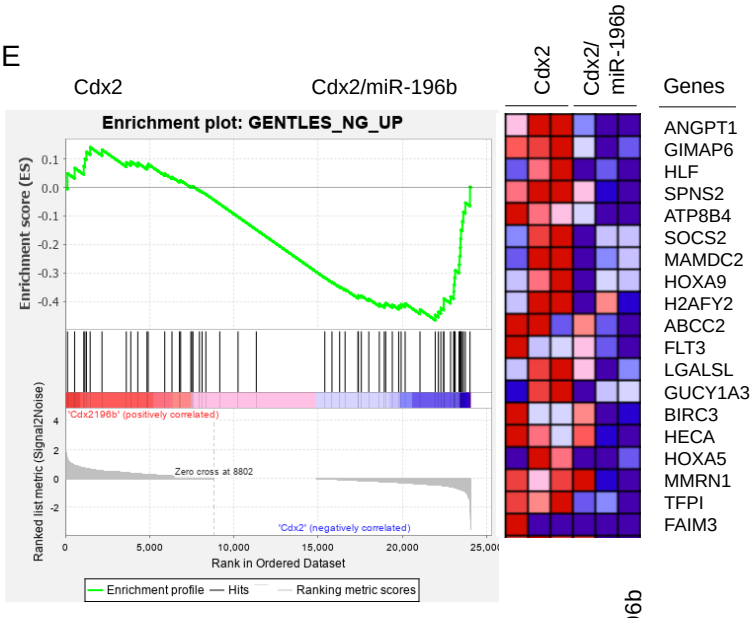
Figure continues on next page

Figure 27 continued

D

	GSEA c6.all.v6.1.symbols [oncogenic]	Genes
Upregulated Top 5 of 27	PRC2_EED_UP.V1_DN	91/176
	RAPA_EARLY_UP.V1_DN	61/167
	HINATA_NFKB_IMMU_INF	8/15
	TGFB_UP.V1_UP	56/174
	STK33_NOMO_UP	90/260
Downregulated all 6	BRCA1_DN.V1_UP	27/112
	GCNP_SHH_UP_LATE.V1_UP	93/169
	GCNP_SHH_UP_EARLY.V1_UP	63/158
	MTOR_UP.N4.V1_DN	56/140
	CYCLIN_D1_KE_.V1_DN	41/172
	PTEN_DN.V1_UP	35/158

E



F

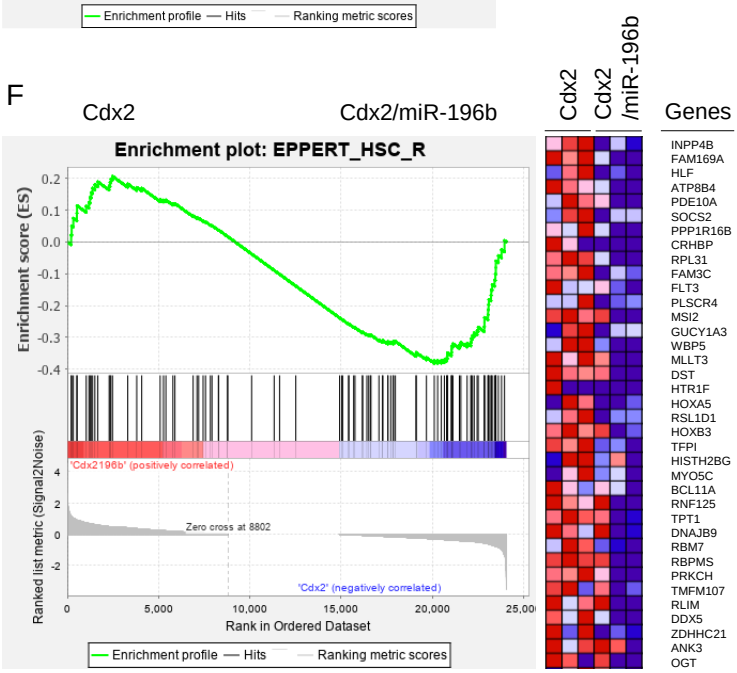


Figure legend on next page

Figure 27. Effect of *Cdx2*/miR-196b overexpression in murine HSPCs on differential gene expression and its relevance in biological processes. RNA was isolated 96 h post transduction (48 h post sort) of murine bone marrow cells and paired-end sequenced. **(A)** Differential gene expression (DEG) pattern identified via DESeq2 provided by the Basepair platform. Depicted is a heatmap of normalized count values of hierarchically clustered DEGs. Blue indicates low and red high expression. FDR<0.1; n=3. **(B)** Enrichment of down (DN)- and upregulated (UP) DEGs in human disease and cellular pathways revealed by the Kyoto Encyclopedia of Gene and Genomes (KEGG) database from 2016. For UP DEGs pathways with significance values of p<0.001 (FDR<0.06) and a score larger than nine genes in a set are presented. For DN DEGs pathways with significance values of p<0.05 (FDR<0.1) are presented. KEGG databases was accessed via the Enrichr platform. The order of pathways is sorted by the number of scores in descending order from top to bottom. Gene names are given to the right next to the pathway name. **(C)** Enrichment of miR-196b target genes selected on the basis of the created 'miR-196b specific target gene' list. Top: miR-196b specific DN DEG (FDR<0.3) enrichment using the Reactome database from 2016 accessed via the Enricr platform. Presented are pathways with significance values of p<0.002 (FDR>0.25) and a score larger than two genes in a gene set. Middle: miR-196b specific DN DEG (FDR<0.3) enrichment using the GO Molecular Function database from 2015 accessed via the Enricr platform. Presented are pathways with significance values of p<0.0021 (FDR>0.25). Bottom: Gene set enrichment analysis (GSEA) software based analysis within the gene matrix 'c5.bp.v7.1.symbols' of the RNA-Seq dataset filtered for miR-196b targets. Presented is a pathway with highly significant enrichment (p<0.001, FDR=1, NES=1.4) of a gene set. Gene names are given to the right next to the pathway name. **(D-F)** GSEA software based analysis of the unfiltered RNA-Seq dataset. **(D)** Enrichment in oncogenic gene matrix 'c6.all.v6.1.symbols [oncogenic]'. Presented are the top 5 (p<0.02, FDR<0.01, NES>1.5) out of 27 oncogenic pathways with significant (p<0.05, FDR<0.25) enrichment of upregulated genes and all six oncogenic pathways with significant (p<0.05, FDR<0.25) enrichment of downregulated genes. To the right next to the pathway name numbers indicate enriched genes out of the reference gene set. GSEA plots visualize enrichment in **(E)** a leukemic stem cell signature, which was reduced to exclusively upregulated genes identified by NG and Gentles et al provided by the MsigDB platform in 03/2019 (Ng et al. 2016, Gentles et al. 2010), and in **(F)** a healthy hematopoietic stem cell signature from Eppert et al (Eppert et al. 2011) provided on the MsigDB platform in 03/2019. The green curve represents the enrichment score. Black vertical lines represent each gene of the reference gene set. To the right of the enrichment plot normalized RNA-Seq expression count values of all significantly signature-matched genes are presented in a heatmap. Low and high expression levels are indicated by blue and red color shades, respectively. Figures E and F modified/extracted from Rawat & Götze et al. 2020, p. 288, CC BY-NC 4.0, https://creativecommons.org/licenses/by-nc/4.0

Candidate genes were categorized as 'proliferation', 'self-renewal' and 'CN-AML' relevant. From the 98 downregulated miR-196b specific DEGs mentioned before, *Mapk1*, *Mapk8* and *Nras* were selected for the category 'proliferation'. *Hoxa9*, *Hoxb8*, *Pbx1* and *Msi2* were selected for the category 'self-renewal' (**Figure 28A)**. Genes for the category' 'CN-AML specific' were selected on the basis of three different criteria that build on each other. First, genes were DEGs after stringent cut off (FDR<0.1) in *Cdx2*/miR-196b co-transduced cells, which did not overlap with downregulated DEGs identified in miR-196b alone transduced cells. Second, those genes were matched with genes from the 'miR-196b specific target gene list'. Third, matching genes had averaged (in case of >1 RNA variant) FPKM (fragment per kilo base per million mapped reads) expression values greater than 30 in CN-AML patients of the TCGA RNA-Seq expression data set (Cancer Genome Atlas Research et al. 2013). Thus, *Sdcbp*, *Eps15*, *Tmed2*, *Nme4* and *Ssr1* were selected for the category 'CN-AML specific' genes. In total, 13 candidate genes were identified of which the majority was also expressed in functionally validated leukemic stem cells of CN-AML patients (**Figure 28B**). In conclusion, miR-196b is able to target genes of different functional categories known to be associated with leukemogenesis.

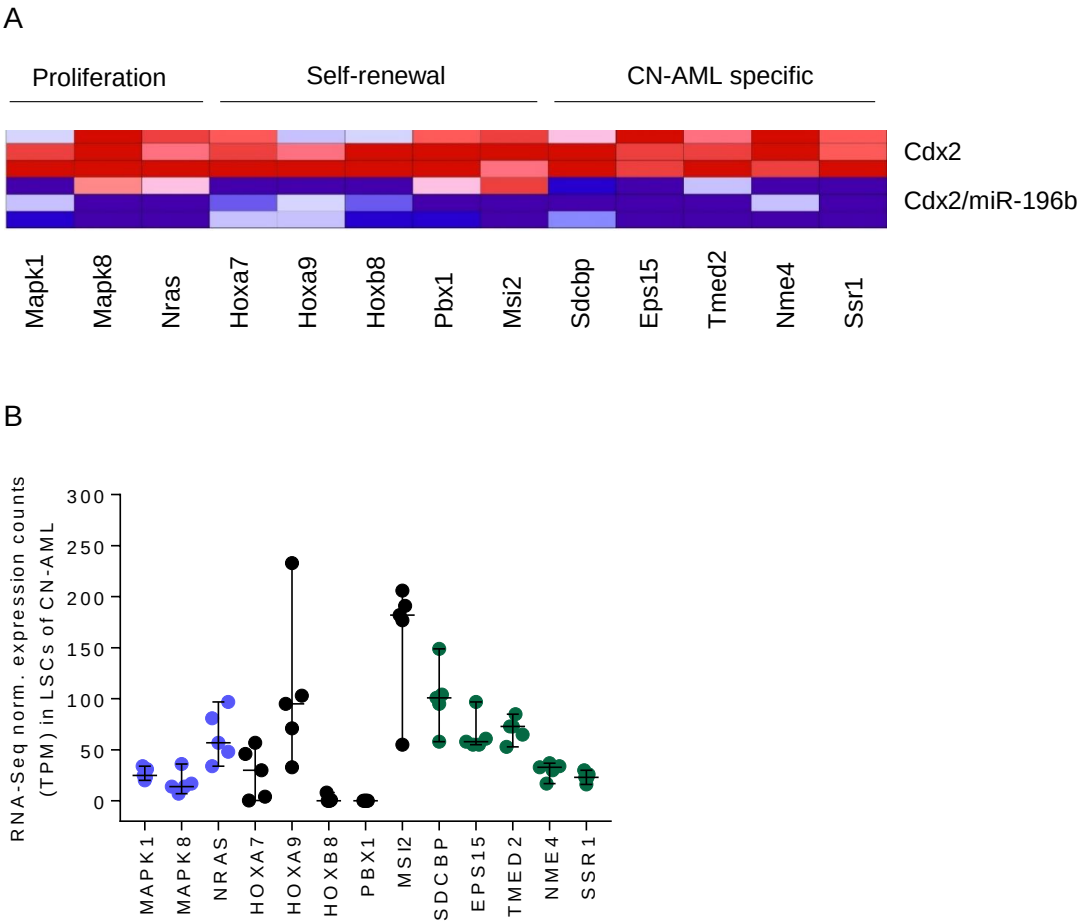


Figure 28. Identification of miR-196b target specific candidate genes in leukemic cells. (A) RNA-Sequencing (RNA-Seq) expression values of miR-196b specific candidate genes in *Cdx2*/miR-196b co-transduced murine hematopoietic stem and progenitor cells (n=3). Presented are normalized RNA-Seq expression count values in a heatmap generated using GenePattern. Blue indicates low and red high expression. **(B)** RNA-Seq expression values of miR-196b specific candidate genes in highly purified functionally validated leukemic stem cells (LSCs) of cytogenetically normal acute myeloid leukemia patients (CN-AML, n=5). Sequencing data was provided by Dr. Susann Ihme from our lab. For patient characteristics see Table 13. Presented are individual TPM expression values, averaged across values identified in leukemic granulocyte-macrophage progenitors and leukemic lymphoid-primed multi-progenitors classifying the LSC population, with the median and 95% confidence interval. Blue data points indicate genes involved in proliferation. Black data points indicate genes involved in self-renewal. Green data points indicate CN-AML specific genes. No inference statistical test was applied.

3.9 Effect of aberrant miR-196b expression on clonogenicity, cell growth and candidate gene levels in human leukemic cells

As forced miR-196b overexpression reduced *Cdx2* induced leukemic transformation, next it was investigated whether this miRNA would have a similar impact on already transformed cells, as this is the situation in patients. To study this, a human AML cell line model that represents AML patients without a leukemia driving fusion gene, but with high *HOX* and *CDX2* expression levels as known for CN-AML patients, was chosen, on the basis of a qRT-PCR analysis (Debernardi et al. 2003, Scholl et al. 2007). Using the CN-AML cell line model OCI-AML3, miR-196b was lentivirally overexpressed and compared to vector control transduced cells in various *in vitro* and *in vivo* assays (**Figure 29**).

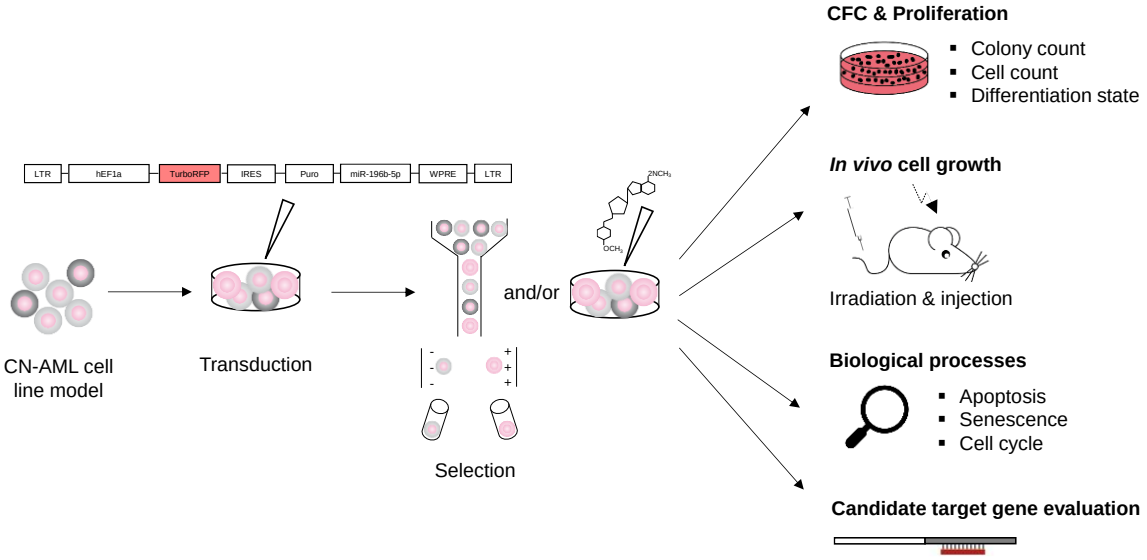


Figure 29. Experimental design for investigating the functional impact of miR-196b in a human CN-AML cell line model. OCI-AML3 cell line was lentivirally transduced and after a FACS and/or puromycin based selection for the reporter gene positive cells, *in vitro* and *in vivo* assays were performed.

In vitro, the effect of miR-196b on clonogenic and cell growth potential was investigated via CFC assays in semi-solid medium and proliferation assays in liquid cultures, respectively. *In vivo*, leukemic growth was investigated by analyzing the BM of NSG mice transplanted with miR-196b or vector control transduced OCI-AML3 cells for the human hematopoietic cell specific CD45 antigen as well as for the red fluorescent reporter protein (RFP) on a flow cytometer. To investigate underlying biological processes of a miR-196b induced impact on human OCI-AML3 cell growth, apoptosis, senescence and cell cycle was investigated - processes selected on the basis of gene enrichment for pathways in murine *Cdx2*/miR-196b co-overexpressing cells (Figure 27 B-C). In OCI-AML3 cells key candidate genes were evaluated via qRT-PCR, selected on the basis of the candidate gene identification in murine *Cdx2*/miR-196b BM cells (Figure 28A) and based on a well-known functional role in AML.

Selection of the human AML cell line model and proof of principle for lentiviral miR-196b overexpression

Among four AML cell lines, OCI-AML3 cells met all requirements for the research question under investigation by highly expressing not only *HOX* genes but also *CDX2* (**Figure 30 A-C**). Kasumi-1 was chosen as *HOX* gene negative control cell line, as it expressed *HOXA7, HOXA9* and *CDX2* the least (**Figure 30 A-C**). A lentiviral transduction of miR-196b successfully led to significant (p=0.0039) 29-times higher expression levels than in vector control transduced cells, 96 h post transduction (**Figure 30D**).

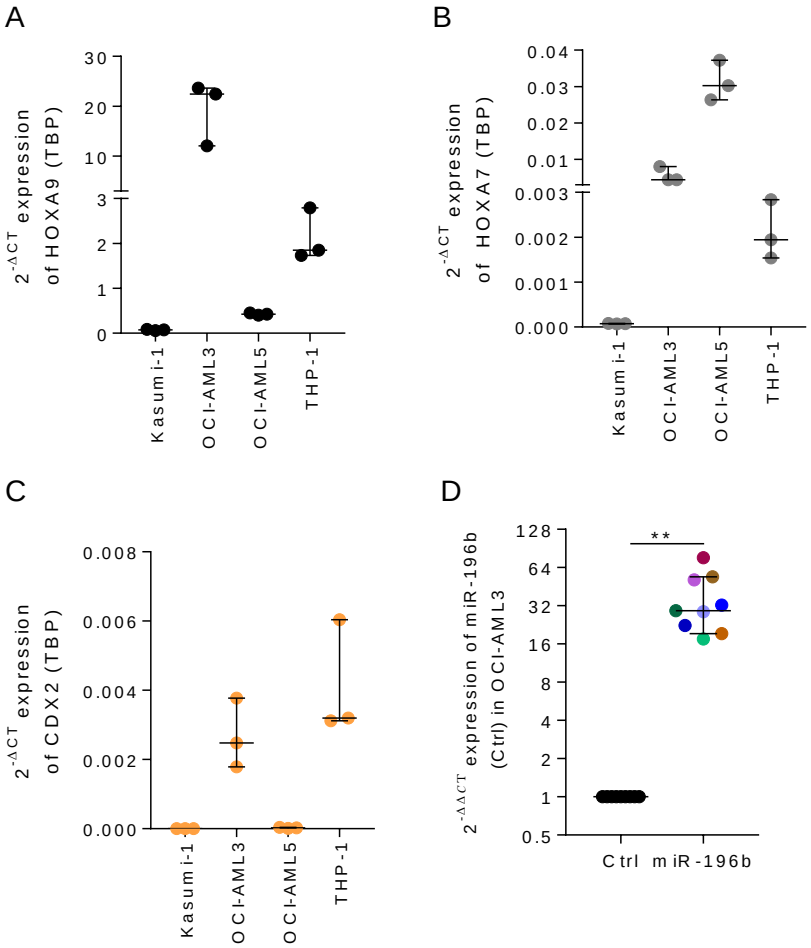
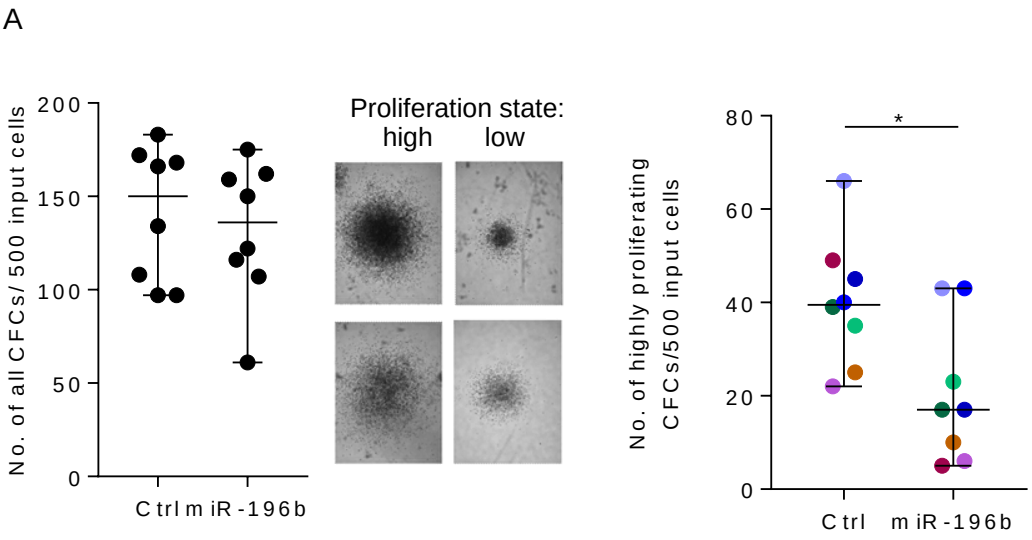


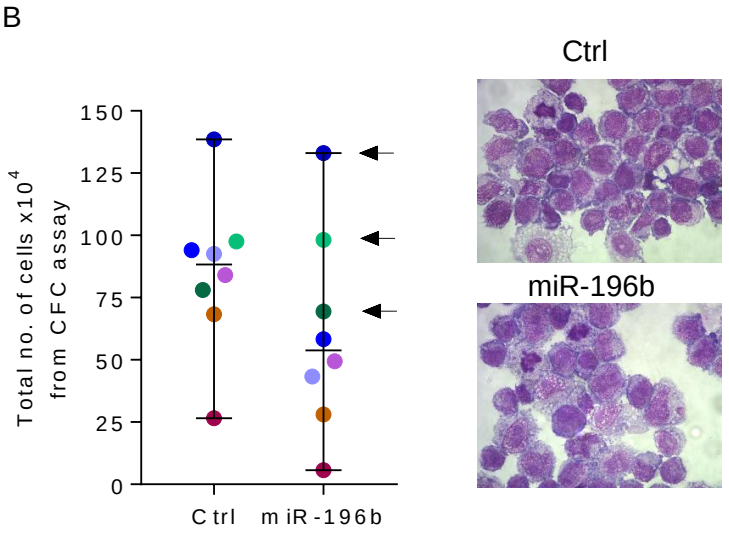
Figure 30. Expression of *HOX* genes, *CDX2* and miR-196b in human AML cell lines. Endogenous expression levels of **(A)** *HOXA9*, **(B)** *HOXA7* and **(C)** *CDX2* in four different AML cell lines were determined via qRT-PCR, normalized on the housekeeping gene *TBP* and transformed into $2^{-\Delta CT}$- values. No inference statistical test was applied. n=3. **(D)** Expression levels of miR-196b, 96 h post lentiviral transduction of OCI-AML3 cells were determined via qRT-PCR, normalized on *RNU6b*, calculated relative to the control condition and transformed into $2^{-\Delta\Delta CT}$-values. Colors of the individual values indicate data affiliations to all other assays in section 3.9. Statistical significance was determined using the One sample Wilcoxon signed rank test. n=9. In all panels **(A-D)**, individual values are presented with the median and 95% confidence interval. C_T: Cycle Threshold; TBP: TATA-box binding protein; RNU6b (RNU6-6P): RNA, U6 small nuclear 6, pseudogene.

Clonogenic and cell growth potential of miR-196b overexpression in human OCI-AML3 cells *in vitro*

Although miR-196b did not significantly affect the total colony number in CFC assays, there was a significant 2.3-fold reduction (p=0.0263) of highly proliferating colonies, reflected also by a significant (p=0.0156 using the One sample Wilcoxon signed rank test; n=8) reduction by 40% of total cell counts in percent to control (Ctrl: median=100%, 95% CI [100, 100] vs miR-196b: median=60.37%, 95% CI [21.23, 100.6]; n=8) (**Figure 31 A-C**). The cell type was not affected by miR-196b overexpression as shown by May-Grünwald Giemsa stainings (**Figure 31C** right). In both, the control and experimental condition, the relation of cytoplasm and chromatin condensation was typical for OCI-AML3 myeloblasts.

In proliferation assays, performed additionally to cell counts of CFC assays, cells were plated in liquid culture medium and counted 3 and 6 days later (**Figure 31D** left). Not after 3, but after 6 days, the number of cells in percent to control was significantly (p=0.0078 using the One sample Wilcoxon signed rank test) reduced upon miR-196b overexpression by 36% (Ctrl: median=100%, 95% CI [100, 100] vs miR-196b: median=65%, 95% CI [42.31, 82.35]; n=9). Notably, a reduction of cell counts between 25% and 43% could be observed even after day 3 in all but two experiments (indicated by arrows in the Figure). The cell differentiation state remained unaffected, determined via May-Grünwald Giemsa stainings (**Figure 31D** right).





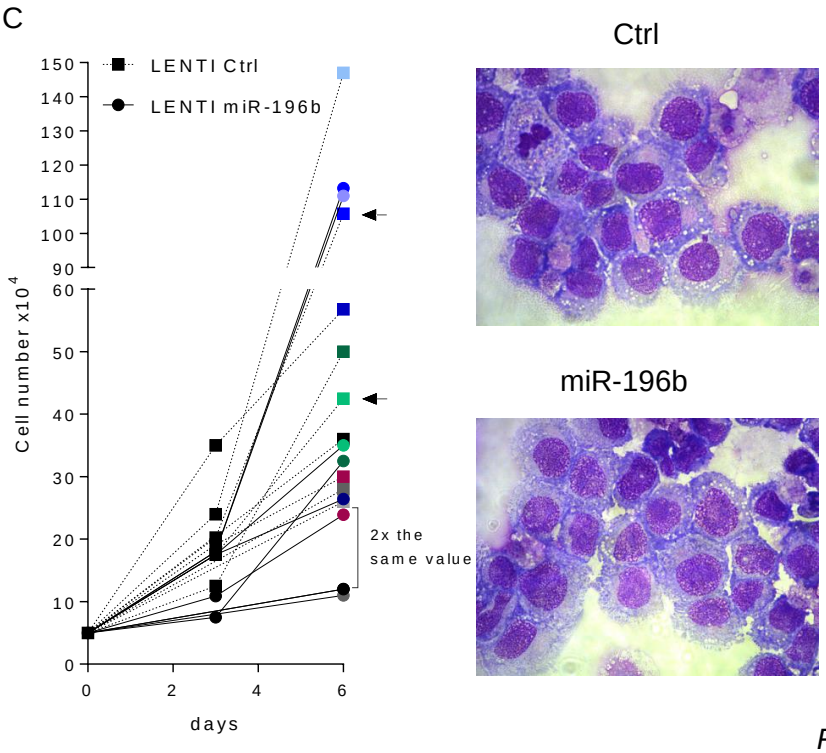


Figure legend on next page

Figure 31. Effect of miR-196b overexpression in OCI-AML3 cells on clonogenicity and cell growth *in vitro*. (A-C) Analysis of the clonogenic potential via colony forming cell assays (CFCs) with 500 OCI-AML3 cells constitutively expressing miR-196b or the control vector in methylcellulose. n=8. **(A)** Total number of colonies. To the right are representative microscope photos of different colony morphologies with a 40x original magnification. **(B)** Number of highly proliferative (>5000) colonies. **(C)** Total number of cells harvested from CFC assays. To the right are representative microscope photos from cells on cytospins after May-Grünwald-Giemsa staining with a 600x original magnification. **(D)** Analysis of the cell growth potential via proliferation assays. Presented are individual cell proliferation curves of OCI-AML3 cells constitutively expressing miR-196b or the control vector. Cell counts were determined on day 3 (n=6) and day 6 (n=9) after initially seeding $5x10^4$ cell in liquid culture medium. To the right are representative microscope photos from cells on cytospins after May-Grünwald-Giemsa staining with a 600x original magnification. **(A-D)** In the scatter blots, individual values are presented with the median and 95% confidence interval. Arrows indicate experimental replicates in which the effect was <20% upon miR-196b overexpression. Colors of the individual values indicate data affiliations of experimental replicates within the assays and to all other assays in section 3.9. Statistical significance was determined using the Mann-Whitney *U* test.

In Kasumi-1 cells, the clonogenic and cell growth potential was not affected by 125-285 times higher miR-196b expression levels (data not shown) compared to vector control cells, 96 h post transduction (**Figure 32**).

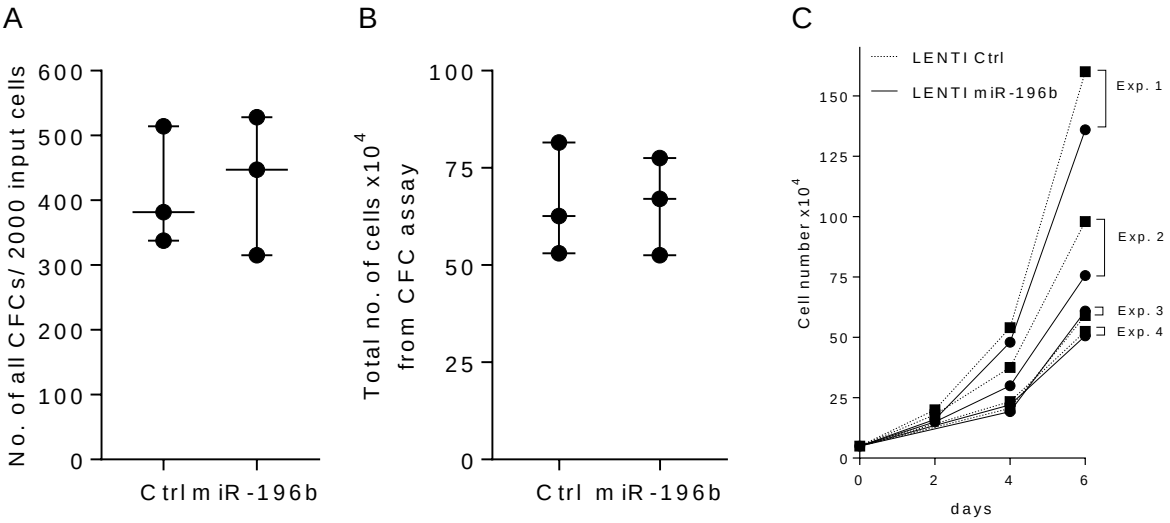


Figure 32. Effect of miR-196b overexpression in Kasumi-1 cells on clonogenicity and cell growth *in vitro*. (A) Total number of colonies and **(B)** total number of cells harvested from colony forming cell (CFC) assays generated by 2000 initially plated Kasumi-1 cells constitutively expressing miR-196b or the control vector in methylcellulose. Presented are individual values with the median and 95% confidence interval. n=3 **(C)** Analysis of the cell growth potential via proliferation assays. Presented are individual cell proliferation curves of Kasumi-1 cells constitutively expressing miR-196b or the control vector. Cell counts were determined on day 2 (n=2), 4 (n=4) and 6 (n=4). (**A-C**) No inference statistical test was applied.

In vivo cell growth potential of miR-196b overexpressing OCI-AML3 cells

To measure *in vivo* cell growth, $5x10^5$ miR-196b overexpressing OCI-AML3 cells were injected into sublethally irradiated and IvIG treated NSG mice. Four weeks post transplantation mice were sacrificed, because this was the time point, when mice transplanted with vector control transduced cells exhibited early signs of disease. Flow cytometry results of the BM revealed, that miR-196b significantly reduced leukemic cell

growth. The RFP signal of BM cells was significantly reduced by 37-63% (p=0.0313, n=6) and the CD45-APC-H7 signal was significantly reduced by 33-58% (p=0.0313, n=6) calculated in % to control using the One sample Wilcoxon signed rank (**Figure 33 A-B**).

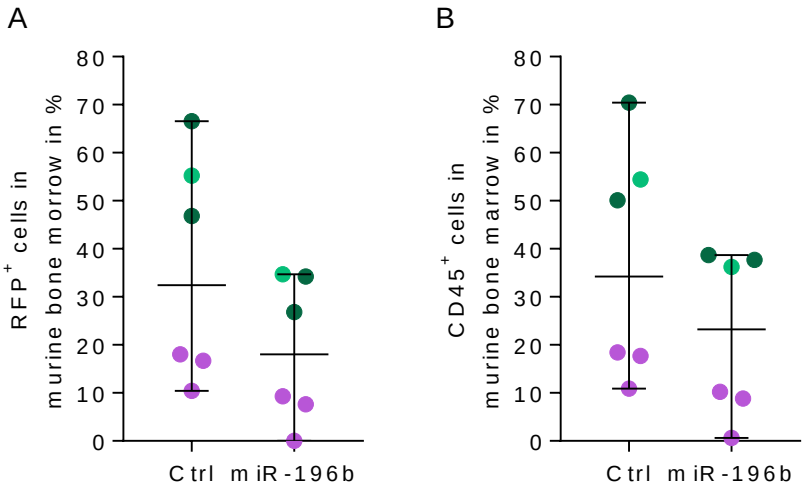


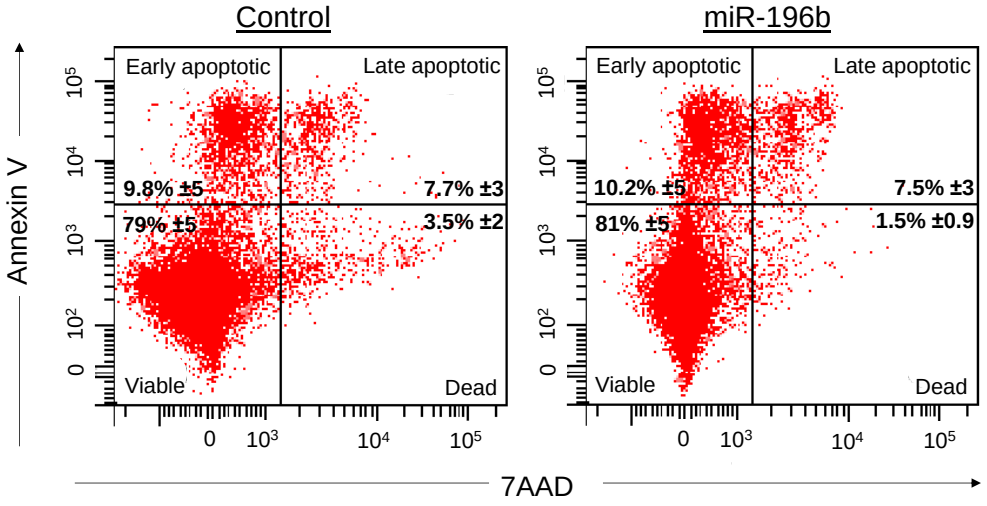
Figure 33. **Effect of miR-196b overexpression in OCI-AML3 cells on *in vivo* cell growth.** Bone marrow (BM) of NSG mice transplanted with $5x10^5$ miR-196b overexpressing cells was analyzed for engraftment four weeks post injection. Presented are individual values of percent (A) RFP^+ or (B) CD45-APC-$H7^+$ cells with the mean and standard deviation. Colors indicate affiliation of each mouse to one of three experimental replicates and allows data affiliations to all other assays in section 4.9. Statistical significance was determined using the Mann-Whitney *U* test. n=6 mice of three experimental replicates.

Biological processes and key candidate gene levels in miR-196b overexpressing OCI-AML3 cells

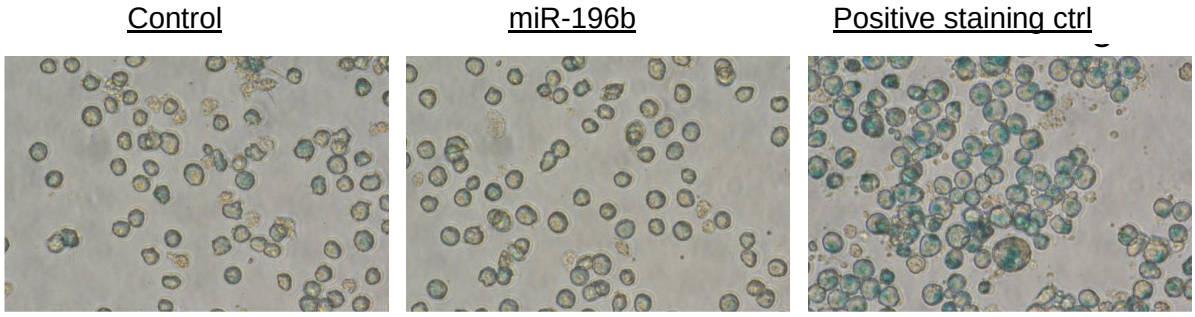
The mild effect of miR-196b in OCI-AML3 cells on cell growth and clonogenicity *in vitro* and *in vivo* could not be explained by apoptotic processes (Annexin V/7AAD staining), senescence (ß-galactosidase staining) or cell cycle (propiodium iodid staining) changes in OCI-AML3 cells cultured for 1.5 weeks, 1-2.5 weeks or 2 weeks after end of transduction, respectively (**Figure 34 A-C**). One possibility for these results might be, that the underlying biological process was no longer detectable due to a selection of cells that could bypass the effect of miR-196b. However, a qRT-PCR analysis of OCI-AML3 cells, 96 h post transduction with miR-196b, revealed changes in candidate mRNA expression levels (**Figure 34D**). Thus, while the proliferative gene *MAPK1* was not affected, its collaborator *MAPK3* was significantly (p=0.0313) reduced by 32% upon miR-196b overexpression. Among self-renewal candidate genes, *HOXA7* and *HOXA9* levels were significantly reduced by 64% (p=0.0156) and 42% (p=0.0156), respectively, whereas *PBX1* and *MSI2* levels remained unaffected.

In conclusion, miR-196b attenuated the leukemic phenotype of OCI-AML3 cells, but was not able to convert the malignant phenotype of the cells to a major extent.

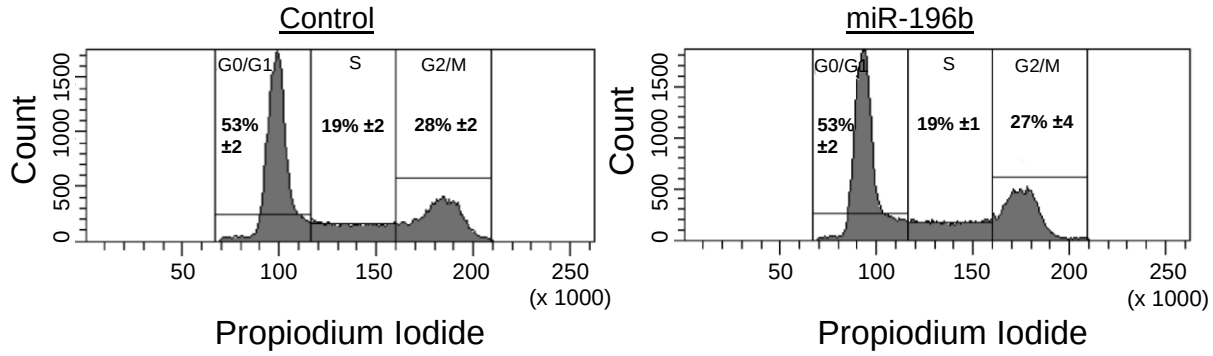
A



B



C



D

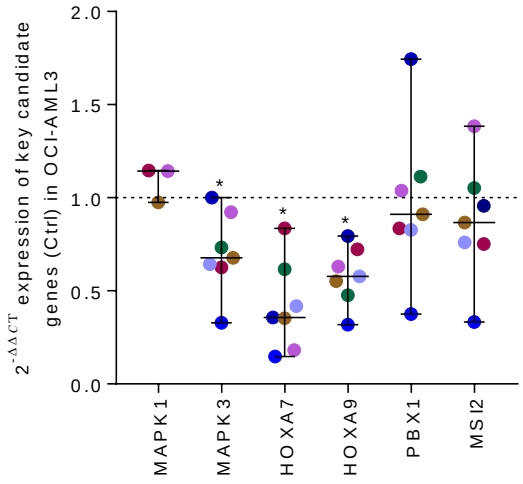


Figure legend on next page

Figure 34. Effect of miR-196b overexpression in OCI-AML3 cells on selected biological processes and candidate gene expression. (A) Apoptosis of cells, which had been cultured for 1.5 weeks after end of transduction, were analyzed via Annexin V /7AAD stainings acquired on a flow cytometer. Depicted are stainings representative for the results of 5 experimental replicates. In each quadrant the apoptotic stage, the mean of % positive cells and standard deviation are indicated. **(B)** Senescence of on retronectin plated fixed cells, which had been cultured up to 2.5 weeks after end of transduction, were analyzed via ß-galactosidase stainings evaluating the color change microscopically. Depicted are microscope photos with a 40x original magnification representative for the results of three experimental replicates. As positive control for the staining method served an AML cell line exposed to a 3 Gy irradiation. **(C)** Cell cycle phases of synchronized and fixed cells, which had been cultured for 2 weeks after end of transduction, were analyzed via propiodium iodide stainings acquired on a flow cytometer. Depicted are stainings after the exclusion of debris, representative for the results of 4 experimental replicates. In each segment the DNA content reveals the cell cycle phase. G0/G1: Exited cells from mitosis are quiescent (G0). They can be reactivated to grow and prepare DNA synthesis (G1). S: DNA is synthesized to enable DNA replication. G2/M: Cell grows and prepares for division (G2). The cell divides in four stages (telo-, ana-, meta- and prophase) of mitosis (M). Indicated are the mean of % positive cells and standard deviation. **(D)** Key candidate gene expression levels were determined 96 h post transduction of OCI-AML3 cells with miR-196b via qRT-PCR. Expression values were normalized on the housekeeping gene *TBP*, calculated relative to the control condition and transformed into $2^{-\Delta\Delta CT}$-values. Presented are individual values with the median and 95% confidence interval. Colors of the individual values indicate data affiliations of experimental replicates within the assays and to all other assays in section 3.9. The dashed line indicates the $2^{-\Delta\Delta CT}$-expression value in vector control cells. Statistical significance was determined using the One sample Wilcoxon signed rank test. n=3-7. C_T: Cycle Threshold; TBP: TATA-box binding protein.

4. Discussion

In this work, the functional and molecular role of miR-196b was investigated in *HOX* gene driven AML and healthy hematopoietic stem and progenitor cells by using gain-of-function and next generation sequencing approaches.

4.1 One of the functions of miR-196b in healthy BM cells is to promote differentiation into myeloid cells *in vivo*

MiR-196b is a miRNA known to be highly expressed in murine stem cells and upon further differentiation its level gradually decreases (Petriv et al. 2010, Connell et al. 2010, Popovic et al. 2009). Many miRNAs with this expression pattern have been assigned to function on stem cell level with relevance to leukemia when overexpressed (Han et al. 2010, Wurm et al. 2017, Li, Wu, et al. 2016).

This study showed high miR-196b expression levels in primitive healthy human BM $CD34^+$ cells and lower expression values in more mature BM MNCs, which harbor mainly mature lymphoid cells and monocytes. Single mature subpopulations were not analyzed, so that it cannot be ruled out that for example very low expression levels in lymphoid cells lower the average miR-196b expression value of the total BM MNC population (Witwer and Halushka 2016). However, the observed expression pattern is in line with the study from Petriv et al (2010), who showed the same trend in murine hematopoietic cells (Petriv et al. 2010). Further, Coskun et al. (2011) showed decreased miR-196b level upon $CD34^+$ maturation (Coskun et al. 2011). The expression of *HOX* genes follows the same pattern. This and the finding that miR-196b overexpression in murine $Lin^-/c\text{-}kit^+$ BM cells increased the number of CFCs with replatability *in vitro* by Popovic et al. (2009) led to the hypothesis that miR-196b gain-of-function would lead to an enhanced self-renewal potential in hematopoietic stem and progenitor cells. However, in this study miR-196b overexpression did neither affect the clonogenic nor the differentiation potential *in vitro*, but rather attenuated cell growth. This could be shown always in comparison to vector control and for two different vectors (Ph1 and LMP) making a vector specific effect (e.g. integration site, gene copy number) very unlikely and suggests, that miR-196b does not promote self-renewal. On one hand, this is contrary to Popovic's finding, presumably, because they have not used 5-FU, but a FACS-based method staining for murine $Lin^-/c\text{-}kit^+$ cells to select for stem and progenitor cells. Sort stress (e.g. shear forces) and other intrinsic conditions affecting silencing complex formation or the availability of 3' UTRs might have led to different functioning of miR-196b (Nam et al. 2014, Yuva-Aydemir et al. 2011, Carroll et al. 2013, Leung and Sharp 2010, Llufrio et al. 2018). In particular the cell cycle rate on day of gene transduction is known to affect the transformation process (Chen, Burkhardt, et al. 2019). On the other hand, the data is in line with the study from Velu et al. (2009), who observed a small reduction of colony number formation upon miR-196b overexpression in murine Lin^- cells (Velu et al. 2009). There can be many underlying mechanisms for the proliferative disadvantage. It seems to be very likely that the downregulation of stem cell relevant genes such as *Hoxa9*, *Msi2* and *Kdm5a* identified via RNA-Seq in this study leads to an attenuation of the self-renewal potential. This might be accompanied by a shift towards increased differentiation, indicated by increased, albeit not statistically confirmed, higher numbers of mature granulocyte colonies (Sauvageau et al. 1994, Kharas et al. 2010, Klose et al. 2007). However, neither May-Grünwald-Giemsa stainings nor immunophenotyping of myeloid and stem cell marker in cells from CFC and/or proliferation assays in this study confirmed this.

Consequently, on one hand, sample size needs to be increased to validate the data, and on the other hand, specific stem and progenitor subpopulations could be investigated to unmask the imminent effect (Nayak 2010, Altschuler and Wu 2010).

In vivo, experiments were continued with the LMP vector, due to its Pol II promoter occurring naturally in the miRNA expression process and due to its specific design for optimized expression in murine stem cells (Dickins et al. 2005, Lee et al. 2004). In addition, the LMP vector backbone has already been used in several miRNA gain-of function studies without reporting vector backbone specific effects such as lineage biases (He et al. 2005, Jiang et al. 2016, Chang et al. 2008, Li et al. 2015). The overexpression of miR-196b in murine HSPCs lead to a higher CFU-S frequency in animals transplanted with short-term cultured miR-196b transduced BM cells than with control vector transduced cells in ΔCFU-S experiments of this study. This is contrary to the *in vitro* data of this study. Nevertheless, it is a possible result, since with ΔCFU-S experiments short-term repopulating stem cells and therewith a different cell population is tested than with the aforementioned *in vitro* assays. In long-term transplantation experiments the overexpression of miR-196b did not lead to major hematopoietic effects, because in the few cases of leukemia retroviral insertional mutagenesis (RIM) or additional cooperating mutations induced by previous irradiation might have been the cause for an unphysiologically rapid expansion of transplanted cells in the BM niche *(see section 4.5 'technical aspects')*. Merely, immunophenotyping PB of mice transplanted with miR-196b for differentiation marker showed a myeloid bias, which did not result in a terminal disease. One reason for this observed myeloid skewing effect could be that miR-196b induces apoptosis in lymphoid cells leading to relatively more myeloid cells or miR-196b promotes myeloid cell differentiation in multipotent progenitor cells directly leading to more myeloid cells. This *in vivo* lineage differentiation experiment was conducted intra-individually by analyzing miR-196b transduced versus non-transduced PB within a given mouse. In this way, it could be made sure that the effect was not caused by inter-individual differences between mice. A shortcoming in this experiment is that during the transduction process the genetic insertion occurs uncontrollably at random. This issue could be overcome by the elegant method of hematopoietic cell specific miR-196b knock in via CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats-caspase 9) (Pavani et al. 2020). To avoid *ex vivo* cell manipulations the Cre-LoxP system could be used to generate miR-196b knock-in cells (Odell et al. 1994). Nevertheless, myeloid skewing is a known effect for miRNAs in hematopoiesis as for example shown for miR-101, which targets TET2 and EZH2 to create this phenotype (Cheng et al. 2013).

Taken together, under *in vitro* conditions miR-196b does not affect stemness of murine HSPCs and under *in vivo* conditions miR-196b seems to influence cell fate decisions towards myeloid differentiation, which, however, did not induce disease.

4.2 Transcriptional control of miR-196b in healthy human BM cells occurs via a specific promoter generating two primary miR-196b transcript variants

Promoter

Gene transcription is regulated via the association of specific proteins with a gene specific promoter. A promoter is defined as a DNA stretch harboring specific sequence elements named promoter features or motifs (Marbach-Bar et al. 2013). MiRNAs mostly use the expression strategy of their host genes, but can also have their own transcriptional program

(de Rie et al. 2017). In this study, a miR-196b associated promoter region was identified as well as its efficacy proven using luciferase assays in HEK293T cells. The proximity of this identified promoter region to the miR-196b gene and its long distance of 3.5 kb to the classical HOXA9 promoter are indicators for a miR-196b specific promoter region, which remains to be verified. In 2015, the miR-196b promoter region was reported in the context of embryogenesis, showing *in vivo* enrichment for Cdx2 and Hoxd13 at the promoter of miR-196b after electroporation of the promoter sequence into chicken eggs (Fantini et al. 2015). In line with an upstream regulatory role of Cdx2 for miR-196b expression regulation was a in this study determined 6-fold increase of miR-196b expression levels in 5-FU BM cells upon *Cdx2* overexpression (qRT-PCR data not shown). However, in this study, CDX2 was not identified as relevant TF on the miR-196b promoter region, since it was predicted by only one of six bioinformatics tools. The TF CEBPA, however, was predicted by all six TF prediction tools and EZH2 appeared in published ChiP-Seq data. Both TFs are known to play a role in AML and are therefore suggested as potentially most relevant TFs in miR-196b expression regulation (Pabst and Mueller 2009, Shi et al. 2013). Besides TFs, also epigenetic marks could be interesting targets in miR-196b expression regulation as successfully shown by Liu et al. (2013), who could increase miR-196b expression levels via the DNA methyltransferase 5-aza-2′-deoxycytidine and the histone deacetylase inhibitor 4-phenylbutyric acid acting, both decreasing the methylation status of the miR-196b promoter in the CML cell line model K562 (Liu et al. 2013). Moreover, the promoter feature analysis of the miR-196b promoter region in this study strongly indicates a functional division of this promoter region, which needs to be additionally taken into account when considering miR-196b expression regulation. One histone peak was concentrated near the miR-196b gene (proximal) and another one was further away (distal). Those histone peaks are H3K4me3 and H3K27ac marks, which demonstrate transcriptional activity (Bandeira and Boeck 2020). Predicted promoter features differ quite fundamentally from each other in those two promoter parts. While the proximal promoter region has especially unstable DNA features such as GC-boxes, CpG islands, a tandem repeat region and a poor evolutionary conservation, the distal part has rather stable DNA features such as a TATA-box and a high evolutionary conservation implicating a less tight expression regulation proximal and a more tight expression regulation distal (Carninci et al. 2006, Solovyev and Shahmuradov 2003, Yang et al. 2007, Esteller 2002, Quilez et al. 2016). Consequently, it can be speculated that each promoter part plays a specific role in miR-196b expression regulation, which needs to be considered when promoter manipulations want to be used for miR-196b expression level regulation.

Primary miR-196b transcript

MiRNA transcripts are either generated from DNA stretches between genes (intergenic) or originate from introns and sometimes from exons of genes (intragenic) (Boivin et al. 2018, Kim and Kim 2007). In this study, two miR-196b specific primary transcripts with the unusual characteristics of harboring both intergenic and intragenic parts were identified as well as their functionality proven via overexpression assays. For the miR-196b transcript identification, conventional PCR assays with a RP specific for an evolutionary conserved EST and a FP 201 bps upstream of the miR-196b stem-loop sequence was used. A homology analysis of the two miR-196b transcript variants revealed two mismatches on the 5’ end of the excised genetic stretch, but the remaining sequence at the beginning and the end were exactly the same. Genetic alterations at boundary sites of excised DNA stretches are known to occur during splicing of mRNAs (Ward and Cooper 2010). In some cases

splicing also occurs in pri-miRNAs as it has been demonstrated in plants (Bielewicz et al. 2013, Schwab et al. 2013). From this it can be assumed that the smaller transcript MF1309051 is a variant of MF139050 possibly presenting a pre-stage of the miR-196b precursor sequence, a phenomenon so far not mentioned in the literature. The conventional PCR method used for the miR-196b transcript identification is a standard method giving reliable results, but it cannot be claimed that the complete primary transcript has been identified. This is also indicated by a TSS found in Rie's expression atlas 10 kb away from the miR-196b gene using deep sequencing, which, however, has not been functionally validated (de Rie et al. 2017). Nevertheless, it shows that transcript and promoter determinations are tightly linked and in particular TSS determination is a key step.

In sum, this study identified an active promoter region in the direct vicinity of the miR-196b gene as well as primary transcript variants. The next question was whether analyses of miR-196b transcripts would reveal miR-196b expression regulation differences in leukemic cells originating from AML versus healthy cells.

4.3 Association of miR-196b and *HOXA9* expression levels in normal and leukemic hematopoiesis

Identified miR-196b transcript variants were located between and partly within the leukemia associated clustered *HOX* genes: *HOXA9* and *HOXA10*. Both the miR-196b and neighboring clustered *HOXA* genes such as *HOXA9, HOXA10, HOXA5* and *HOXA7* are highly evolutionary conserved (Cillo et al. 2001). In cancer, miR-196 plays a dichotomous role depending on cancer type and progression state (*see Table 4 and 5 in the introduction*). In AML, particularly *MLLr*-AML, several studies show that miR-196b is higher expressed than in healthy cells acting as oncomir and upon miR-196b depletion LSC activity is inhibited (Popovic et al. 2009, Meyer et al. 2018, Velu et al. 2014). In contrast, low miR-196b levels compared to healthy counterparts have been reported in T-ALL and CML (Bhatia et al. 2011, Liu et al. 2013). In CML an enhancement of miR-196b levels lead to a decrease of cell proliferation rate and a retardation of cell cycle (Liu et al. 2013). These published data suggest, that miR-196b plays different roles depending on the leukemia subtype. In this work, the role of miR-196b in CN-AML was investigated.

Expression of miR-196b in CN-AML patients

MiR-196b expression levels in CN-AML bulk cells were similar to those in stem and progenitor cells (BM $CD34^+$), but *HOXA9* expression levels were higher in CN-AML bulk cells in this comparison. BM progenitor subpopulations in the $CD34^+$ compartment were not analyzed, so that it cannot be ruled out that for example very high miR-196b levels in CMPs raise or very low levels in MEPs lower the mean miR-196b expression value in BM $CD34^+$ cells. Skvarova Kramarzova et al. (2014) showed for example that *HOXA* gene expression levels decreased with higher maturation status of BM progenitor subpopulations (Skvarova Kramarzova et al. 2014). However, the observed expression pattern is in line with the study from Lechmann et al. (2016) and Metzeler et al. (2013), who revealed that miR-196b is not among overexpressed miRNAs that associate with survival in AML and CN-AML, respectively (Lechman et al. 2016, Metzeler et al. 2013). Notably, within the CN-AML patients investigated in this study, miR-196b was 2-fold higher expressed in the favorable NPM1c subtype than in the less favorable FLT3 mutated subtype, suggesting that miR-196b could play a beneficial role in leukemia. The association of miR-196b with particularly

NPM1 mutated CN-AML was confirmed by several other publications (Jongen-Lavrencic et al. 2008, Becker et al. 2010, Cammarata et al. 2010). In LSCs, expression data generated in this study indicate slightly lower miR-196b expression levels in LSCs compared to the healthy matching counterpart, while *HOXA9* expression levels were significantly increased in LSC data originating from public data sets (Ng et al. 2016, Laurenti et al. 2013). However, in the LSC subpopulations (LLMPP and LGMP), which are responsible for leukemia initiation and maintenance, miR-196b was not differentially expressed in this study. These data are in line with data from de Leeuw et al. (2014), who identified miR-196b as one of eight miRNAs that were decreased in LSCs compared to HSCs (de Leeuw et al. 2014). Taken together, these LSC data indicate that miR-196b might be a negative modulator of leukemic activity on one hand. On the other hand, expression levels give no information about the activity of miRNAs, namely, whether they associate with the RISC complex and thereby cause gene silencing or not (Witwer and Halushka 2016). Thus, in future, functional assays such as the Competitive Repopulation Unit assay utilizing murine AML cell line models will help to generate a comprehensive view on the functional importance of miR-196b in LSCs. In addition to this interesting, but rather descriptive analysis of miR-196b in LSCs, mechanistic deductions could be made on the basis of associations between miR-196b and *HOXA9* as well as functional readouts, which will be described in the next sections.

Associations of the miR-196b transcript, its mature form and *HOXA9* in CN-AML

MiR-196b transcript – mature miRNA

The linear positive association between miR-196b transcript and its mature form, which was present in healthy human BM cells, was lost in CN-AML bulk cells. A plausible explanation could be that the miRNA maturation process varied in its efficiency between patients leading to uncorrelated levels of miR-196b transcript and the mature form. Reasons for varying efficiencies might be (a) mutations, (b) stability of the biosynthesis complex or (c) the methylation status. In relation to (a) a point mutation was found in the long transcript MF139050 (G to A at position 350) and MF139050 (A to T at position 132) outside the miR-196b stem-loop sequence in one patient after sequencing conventional PCR products of cDNA samples in this study. However, those mutations need to be verified using a proofreading polymerase for cDNA synthesis or even better by sequencing DNA from human CN-AML bulk cells. In relation to (b) an unstable biosynthesis complex during the maturation process was found for the miRNA let-7 in pancreatic cancer and, thus, might also be causative for the inadequate miR-196b maturation process in CN-AML (Dangi-Garimella et al. 2011, Paroo et al. 2009). In relation to (c) methylation marks on miRNA transcripts essential for the recognition and processing by DGCR8 to yield mature miRNA might have been disturbed on miR-196b transcript variants in CN-AML (Alarcón et al. 2015).

HOXA9 – miR-196b transcript

The linear positive association between *HOXA9* and the transcript MF139051, which was present in healthy BM cells, was slightly lost in CN-AML. It is possible, that under physiological conditions, HOXA9 regulates miR-196b transcription, whereas under pathological conditions HOXA9 molecules, with levels above the physiological threshold, might form different transcription complexes by associating with critical well-known HOXA9 co-factors such as PBX (Pbx1 and Pbx3) and Meis1 in AML, which would not lead to an

additional miR-196b upregulation (Lambert et al. 2019). The association of HOXA9 with different co-factors would need to be tested via CHiP-Seq for HOXA9 in a physiological vs a pathological context. Moreover, it can be speculated, that a third variable such as a common enhancer for *HOXA9* and miR-196b regulates the expression of both genes. Thus, if the enhancer's magnitude of action differs in healthy versus CN-AML conditions it will equally effect miR-196b and *HOXA9* transcription (Panigrahi and O'Malley 2021).

HOXA9 – mature miR-196b

While the positive association between *HOXA9* and the miR-196b transcript in BM $CD34^+$ cells was slightly lost in CN-AML bulk cells, the association between *HOXA9* and mature miR-196b was strongly lost in this comparison. This observation indicates that a potential feedback loop in healthy BM $CD34^+$ cells, in which HOXA9 drives miR-196b transcription and is target of miR-196b at the same time in order to maintain homeostasis of *HOXA9* expression levels, might be disturbed in CN-AML - to some extent on miR-196b transcript level and to a larger extent on mature miR-196b level, probably due to an ineffective maturation process.

Functional readouts of miR-196b overexpression in leukemic murine and human cells

Mouse

Gain of function studies were used to investigate, whether miR-196b would be able to counteract the leukemic phenotype in a *Cdx2*-induced leukemic mouse model, since miR-196b had no enhancing self-renewal effect in healthy cells, but stem cell relevant genes were reduced in RNA-Seq data generated in this study.

As expected, *in vitro*, miR-196b reduced the *Cdx2* induced clonogenic potential determined via CFC (re)plating assays, a standard and valid assay (Rawat et al. 2004, Rawat et al. 2008, Marley et al. 2003). Notably, despite a profound decrease in the number of CFCs, residual colonies were replatable up to two times indicating a delayed effect of miR-196b on a subpopulation of clonogenic progenitors. This effect might be due to an initially maintained self-renewal potential of *Cdx2*/miR-196b co-expressing cells, supported by May-Grünwald-Giemsa stainings revealing premature cells and similar *HOX* gene expression levels between *Cdx2*/miR-196b and *Cdx2* alone expressing cells (data not shown). However, in the following CFC assays May-Grünwald-Giemsa stained *Cdx2*/miR-196b expressing cells appeared to be differentiated indicating a gradual loss of the self-renewal potential over time. Further, miR-196b induced a proliferative disadvantage in *Cdx2* transformed cells, which might be an indicator for a role of miR-196b in the regulation of the cell cycle. In line, data from Liu et al. (2013) found a decrease in cell proliferation rate and retarded cell cycling upon enforced miR-196b levels in the CML cell line model K562 (Liu et al. 2013). In parallel, in the context of *Cdx2* expression, miR-196b overexpression lead to an increase of early apoptotic cells, a repertoire of cells, which might become late apoptotic and finally die at an advanced stage. Consequently, the number of cell is reduced over time affecting the proliferation rate accordingly. Thus, miR-196b might participate in apoptosis and/or the cell cycle. On genetic level, self-renewal and proliferation relevant genes were downregulated and are suspected to play a major role in the development of the miR-196b induced phenotype and, thus, are described in detail under section 4.6.

In vivo, miR-196b overexpression led to decreased ∆CFU-S frequencies in animals transplanted with *Cdx2*/miR-196b transduced BM compared to *Cdx2* alone. In line, in long-term transplantation experiments, miR-196b led to prolonged survival in the context of *Cdx2*. Only a minority of mice transplanted with *Cdx2*/miR-196b transduced BM became leukemic, possibly due to clonal selection of transformed cells with (a) miR-196b expression levels below a certain threshold, yet to be determined, which value that would be, (b) the ability to escape the otherwise inhibitory effect of miR-196b despite its genomic presence, or (c) merely showing a weakness of this irradiation-based transplantation assay (*see section 4.5 'technical aspects'*). However, the majority of mice lived longer when miR-196b was overexpressed in the context of *Cdx2*. This might have been caused by a disturbed ability of transplanted primitive cells to anchor and grow within the BM niche (engraftment) as Connell et al. (2010) observed a miR-196b induced disadvantage of BM engraftment in mice. Another explanation could be a disturbed ability of transplanted primitive cells to find their way back home to the BM niche (homing). Engraftment and homing processes require cell intrinsic mechanisms and in particular interactions with other cells or extracellular components (Caocci et al. 2017, Heazlewood et al. 2014). An interaction with other cell can be potentially be mediated via miRNAs, because miRNAs are known to be able to exit the cell via gap junctions and persist in blood plasma protected from degradation through packaging into vesicles, microvessels, exosomes or functional silencing complexes (Heazlewood et al. 2014, Arroyo et al. 2011, Caocci et al. 2017, Lim et al. 2011, Valadi et al. 2007, Wang, Zhang, Weber, et al. 2010). A function of miR-196b in cells of the BM niche can be derived from knowledge about its sister miRNA, miR-196a. MiR-196a was shown to enhance osteogenic differentiation and inhibit adipose tissue derived mesenchymal stem cell proliferation (Kim et al. 2009). The latter cell type shares characteristics with mesenchymal stroma cells in the BM niche. Thus, similarly to miR-196a, miR-196b could have an effect in cells of the BM niche, possibly mediated via an extracellular exchange of miR-196b between BM and BM niche cells. However, it can only be speculated about the exact mechanism explaining why mice transplanted with miR-196b in the context of *Cdx2* lived longer. Moreover, it would be important to clarify, whether the observed inhibitory role of miR-196b on *Cdx2*-induced leukemia in mice can be applied to other similar AML types. One alternative *Hox* gene driven mouse model is to overexpress the *MLL-AF9* fusion gene in murine HSPCs. In this mouse model the role of miR-196b is controversially discussed (Li, Huang, et al. 2012, Popovic et al. 2009, Velu et al. 2014, Meyer et al. 2018). If even within one murine leukemia model the effect of miR-196b overexpression is not uniform, then the function of miR-196b must be strongly context dependent including a dichotomous role. This dichotomy of miRNA functioning is a known phenomenon also for other miRNAs such as miR-126, miR-223 and miR-206 (Li et al. 2015, Gentner et al. 2015, Lin et al. 2011). MiR-126 knockout and overexpression both lead to enhanced leukemogenesis in murine HSPCs (Z. Li et al., 2015). MiR-223 function depends on the phenotype of AML. In the *Hoxa9*/*Meis1* mouse model for myeloid-differentiated AML the loss of miR-223 tendentially delays the onset of leukemia by significantly lowering the LIC frequency, which is not the case in the *MN1*-based mouse model for undifferentiated AML (Gentner et al., 2015). MiR-206 promotes or inhibits krüppel-like factor 4 translation depending on whether the epithelial cell is healthy or a tumor cell (C.-C. Lin et al., 2011). Those examples of context dependent miRNA functions strengthen the need to investigate the role of miR-196b also in alternative AML models, which are similar to the *Cdx2*-based leukemic mouse model used in this study. The leukemia mouse model based on MN1 overexpression seems to be particularly suitable, since MN1 overexpression causes an undifferentiated myeloid leukemic phenotype alike *Cdx2* (Heuser et al. 2007). In addition, the *MN1* gene acts as independent

prognostic marker for CN-AML patients and is associated with *HOXA9* upregulation (Lai et al. 2017). The *Hoxa9/Meis1*-based leukemia mouse model causes a differentiated AML phenotype and would therefore not be suitable for validating the result of this study (Gentner et al. 2015).

Taken together in both *in vitro* and *in vivo* conditions miR-196b inhibits the leukemic transformation in the context of *Cdx2*. It is not suspected, that this effect is an artefact as a result of random insertional mutagenesis (RIM), whose probability increases automatically with double gene transduction, because the anti-leukemic effect occurred reliably in a number of independent experiments. Nevertheless, a double gene transduction control with a combined transduction of the empty MIY and LMP-vector would strengthen the validity of the observed anti-leukemic effect additionally. In future, loss-of-function studies could be performed by overexpressing *Cdx2* in miR-196b knockout cells to see if this would lead to an opposite phenotype, namely a more aggressive AML development. For that, three general methods can be considered: genetic knockouts, antisense oligonucleotide inhibitors and sponges. A genetic knockout could be achieved via the CRISPR-Cas9 system, which might be the most specific and robust tool (Mali et al. 2013, Chang et al. 2016). Antisense miRNA oligonucleotides (AMOs) with special chemical modifications could alternatively be used to inhibit miR-196b. Particularly appreciated are those AMOs with a locked nucleic acid (LNA) modification or with the attachment of a cholesterol group (called antagomirs) (Lennox and Behlke 2011). Sponges can be stably expressed and act as dominant negatives for the miRNA of interest by binding to the same mRNA targets, but without causing an effect (Ebert and Sharp 2010). So far successful miR-196b inhibition in cancer was gained with antagomir and sponges (Velu et al. 2014, Stiegelbauer et al. 2017). For LNAs a strong impact of the negative control on apoptotic markers was reported and should rather not be used (Rebucci et al. 2015).

Human

Based on the promising mouse data, miR-196b was overexpressed in OCI-AML3 cells, a cell line model for CN-AML patients. As expected miR-196b overexpression reduced cell growth of OCI-AML3 cells *in vitro* and *in vivo*, which is in line with documented decreased proliferation rates in human choriocarcinoma cells after transient transfection with miR-196b mimics (Guo et al. 2019). However, the inhibitory effect of miR-196b in OCI-AML3 cells was moderate in this study. Causative might have been random varying effects between experimental replicates, the cell cycle status as for example shown to influence miRNA function of let-7 or stressors that have occurred during cell culturing, all masking a strong effect of miR-196b (Vasudevan et al. 2007). Overexpression levels of miR-196b slightly varying between independent experiments could not be linked to the strength of the effect. Nevertheless, miR-196b reduced cell growth. However, many underlying candidate mechanisms such as apoptosis, senescence, cell cycle status or monocyte/macrophage differentiation state were investigated without detecting a difference. An explanation for this could be that tested biological processes are either not the underlying mechanism, should be investigated at another time point or higher sensitivity of the assays are needed. To address the latter more sensitive assays such as the bromodeoxyuridine (BrdU) assay for cell cycle or a FACS based analysis for detecting the differentiation state via cell surface markers (e.g. CD11b and CD14) could be additionally used (Wald et al. 2008). Alternatively, the anti-cancer effect could be enhanced by identifying and then co-expressing a collaborator of miR-196b in OCI-AML3 cells as this cooperative action of miRNAs is known

to lead to higher repression of mRNA targets (Briskin, Wang & Bartel 2020; Friedman, Balaga & Linial 2013). Another analysis would be to systematically evaluate and then select the time point at which the miRNA induced effect is at its maximum. This time-dependent effect of miRNA expression levels was investigated by Bozkurt et al. (2020), who showed an effect of miRNA expression levels on invasion capacity and proliferation potential of breast cancer cells in a time-dependent manner over a range of four to seven days. They observed >50-fold expression level changes within 48 h. Nevertheless, mRNA levels of self-renewal genes such as *HOXA9* and *HOXA7* as well as proliferative genes such as *MAPK3* were significantly reduced upon miR-196b overexpression in OCI-AML3 cells indicating relevance of miR-196b in those biological processes. *In vivo*, miR-196b also reduced cell growth determined via the reporter gene RFP and CD45 on OCI-AML3 derived xenografted cells in NSG mice. CD45 was detected in addition to RFP in order to not run the risk that reduced RFP levels upon miR-196b overexpression was an effect of random transcriptional transgene silencing in daughter cells or a random translational inhibition of the reporter gene (Den Hartogh and Passier 2016, Ellis 2005, Komorowski et al. 2009, Nie and Htun 2006). The reduced cell growth induced by miR-196b in BM xenografts might have been caused by engraftment or homing issues as discussed for murine BM transplantation assays. Notably all BM xenografts, in the control and experimental condition, became sick. Thus, an about 29-fold miR-196b overexpression in OCI-AML3 cells did not prevent AML to occur, suggesting that an established strong leukemic program could circumvent the tumor suppressor activity of miR-196b. Taken together, in both *in vitro* and *in vivo* conditions miR-196b attenuates cell growth in the CN-AML cell line model OCI-AML3. Whether this effect can be applied to primary patient samples remains to be speculative, since CN-AML patients are a very heterogeneous AML cohort, whereas OCI-AML3 cells are very homogeneous and derived from one patient.

Human versus murine

Overall, the effect of miR-196b in the human OCI-AML3 cell line was rather less pronounced compared to the distinct effect in leukemic mouse cells. One reason could have been differences in the miR-196b overexpression level. While the overexpression in murine cells was about 4000-fold higher, it was only about 29-fold higher in human cells than in the respective control transduced cells. Consequently, the target gene repression mechanism in human cells might have been less effective than in murine cells (Bartel 2009, Denzler et al. 2016). This could be explained by either lower absolute miR-196b expression levels in OCI-AML3 cells or it may be that the critical miR-196b threshold value at which miR-196b expression is not sensed as naturally occurring fluctuation by the cell could not be consistently exceeded with the achieved miR-196b overexpression values in this study. Maybe the critical miR-196b threshold value in OCI-AML3 cells is higher than in murine BM cells due to elevated absolute native (endogeneous) miR-196b expression levels. Naturally occurring fluctuations are balanced out by the cell via a mechanism called dosage control (Ishikawa et al. 2020). Dosage control is a mechanism to keep protein levels in balance. It is a known fundamental quality of a cell's robustness against various perturbations. For example, the PU.1 transcription level can vary from 100-50% and still leads to normal HSC differentiation. Only at a PU.1 expression level below 20% normal differentiation is impaired and turns cells into a preleukemic phase (Rosenbauer et al. 2005). Since miRNAs act as mediators between mRNA and the protein silencing complex, just as certain proteins do between mRNA and protein complexes regulating posttranscriptional processes (e.g. translation and stability), protein dosage control mechanisms could also be a concept for

miRNAs, the non-protein-coding RNA elements in the cell (Rissland 2017). The miRNA dosage determines not only, if, or if not a miRNA effects the cell's behavior, miRNA dosage also determines the intensity of an effect by influencing mRNA binding site affinities as shown by Mayya and Duchaine (2015). Consequently, the concentration of miR-196b levels may dictate which genes are regulated and therewith influences the dimensions of the resulting phenotype. To directly address a miR-196b specific dosage-effect, one could use vectors with graded miR-196b precursor transcripts. However, an overload as mostly discussed for proteins should be strictly avoided in order to circumvent a so-called supraphysiological effect (Tomala and Korona 2013, Moriya 2015) (*see also section 4.4 'context dependent miR-196b function'*). Thus, miR-196b overexpression levels in OCI-AML3 cells should be maximally five times more than the overexpression levels achieved in this study, because with greater miR-196b fold changes an effect was observed in the Kasumi-1 control cell line that was otherwise not present (data not shown). To achieve higher miR-196b levels in OCI-AML3 cells, vectors with more miR-196b copies (multicopy plasmids) could be used, but without straining the size of the vector too much. Transduction efficiencies should not be increased since transduction efficiencies were already at 50% and should not be exceeded to avoid RIM. Alternatively, OCI-AML5 cells, which had 20-fold lower miR-196b expression levels than OCI-AML3 cells might serve the purpose as another CN-AML cell line model for patients with FLT3 mutation (data not shown). Another reason for the less pronounced effect in OCI-AML3 than in assays with the *Cdx2* induced leukemic mouse cells might be different cellular stages the cells were in, when miR-196b was overexpressed. While OCI-AML3 cells were already transformed, murine cells were just about to transform into leukemic cells. Thus, miR-196b might play a more important role during an early cellular stage of the multi-step transformation process such as on pre-LSC level at which it can protect the BM cell from a leukemic transformation (Pandolfi et al. 2013).

Hypothetical model of miR-196b in a negative feedback loop with *HOXA9*

Based on association studies and functional readouts a hypothetical model for the interaction between miR-196b and HOXA9 in healthy and leukemic BM was generated (**Figure 35**).

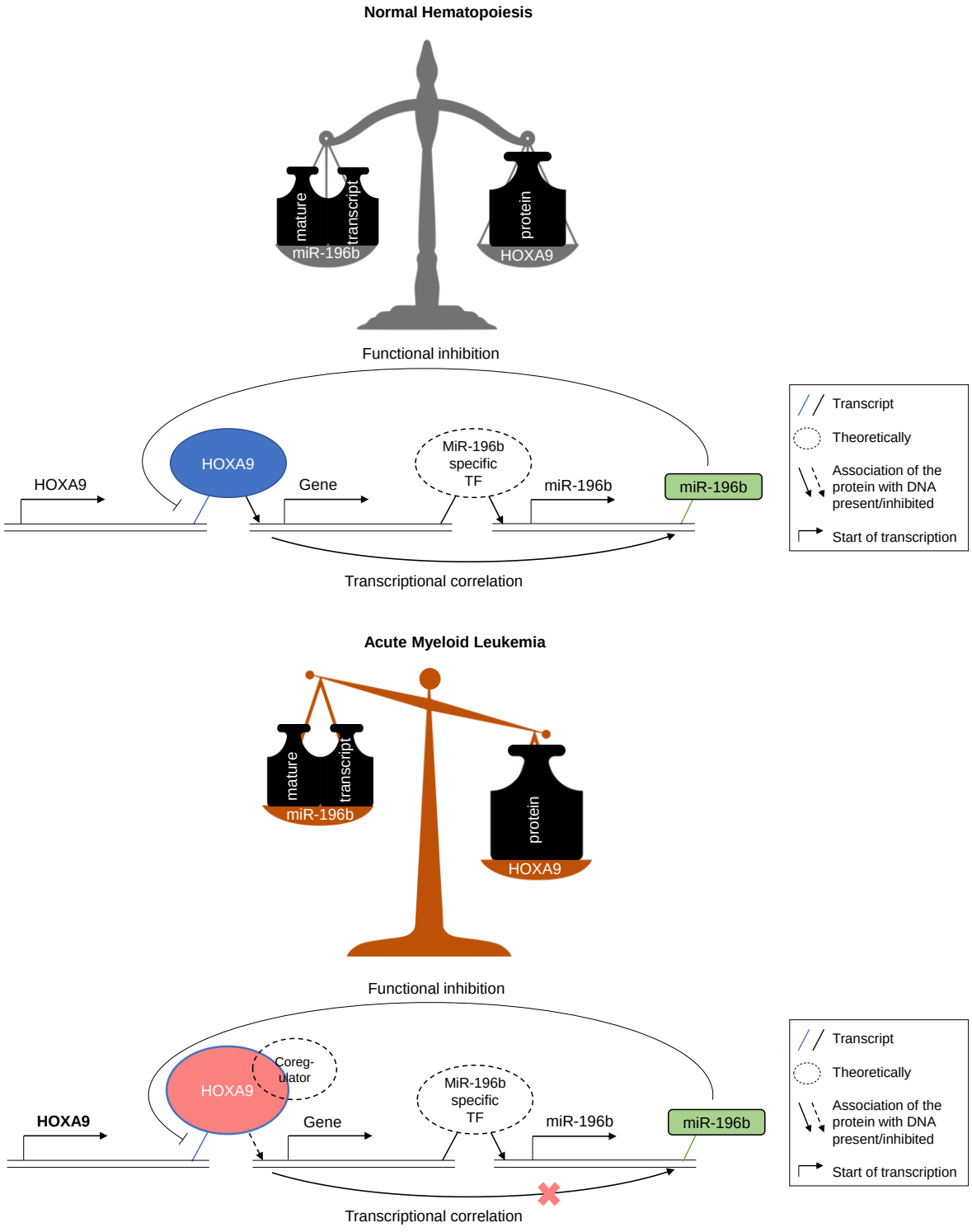


Figure 35. Hypothetical model of miR-196b in a negative feedback loop with HOXA9 in healthy and leukemic BM. The model is based on association studies between miR-196b and HOXA9 expression levels in healthy bone marrow (BM) bulk cells and in cytogenetically normal acute myeloid leukemia patient samples as well as on expression level changes in healthy and *Cdx2* transduced murine BM cells overexpressing miR-196b. In healthy BM cells, HOXA9 appears to fine-tune its own expression via miR-196b in a functional negative feedback loop. In AML, this mechanism seems to be disturbed.

Based on expression data in healthy human BM cells it appears, that with higher *HOXA9* levels, both the transcript and mature miR-196b levels were increased indicating a positive correlation. In addition, expression data in healthy murine BM cells revealed that miR-196b overexpression correlated with reduced *Hoxa9* levels. Taken together, under physiological BM cell conditions, HOXA9 leads to miR-196b upregulation and miR-196b degrades *HOXA9* mRNA. The HOXA9 induced upregulation of miR-196b probably occurs indirectly, for example by regulating the transcription of miR-196b specific TFs, since no HOXA9 binding sites were predicted for the miR-196b promoter region in a TFBS analysis in this study. This reverse interaction between miR-196b and HOXA9 might be one way of HOXA9 to buffer out its own stochastic fluctuations (noise) during transcription (Ebert and Sharp 2012). A very similar relationship was found between Runx1 and miR-27a, in which miR-27a attenuates Runx1, and in turn Runx1 upregulates miR-27a expression during megakaryopoiesis (Ben-Ami et al. 2009). Hence, in contrast to for example multiple E3 ubiquitin ligases, which have a rather broad target spectrum for proteosomal degradation, miRNAs have less targets and can therewith contribute more specifically to the fine-tuning of mRNA expression fluctuations in a cell (Ishikawa et al. 2020, Taggart and Li 2018). However, in CN-AML samples the correlation between *HOXA9* and the miR-196b transcript as well as between *HOXA9* and mature miR-196b was moderately to severely disturbed. Also, the correlation between miR-196b transcript and mature miR-196b was disturbed. These data point towards a misprocessing of miR-196b at both the transcript and maturation level. On miR-196b transcript level, it might be that under leukemic conditions the excess amount of HOXA9 binds to critical known coregulators in AML such as Meis1 through which the TF HOXA9 shifts its focus away from physiological gene transcription regulation to the transcription of oncogenes (Collins & Hess 2016). Thus, the HOXA9-mediated pathways for miR-196b transcription is attenuated and therewith the feedback loop to inhibit leukemic HOXA9 levels. Alternatively, other leukemic intrinsic cell signaling cascades might actively prevent miR-196b transcription. In any case, this is an interesting finding, that despite the local proximity of *HOXA9* and miR-196b on the genome their transcript processing in CN-AML seems to be very much independent from each other. What happens during miR-196b maturation in CN-AML remains a completely open question. However, the consequence of miR-196b misprocessing in CN-AML is, that probably mature miR-196b expression levels are too low in relation to increased leukemic *HOXA9* levels. Thus, the negative feedback loop in which miR-196b effectively inhibits excess *HOXA9* levels in healthy BM cells seems to be disturbed in CN-AML samples. This might lead to the observed 6-fold increased *HOXA9* levels in CN-AML compared to healthy BM bulk cells. One solution for this leukemia driving loop might be miR-196b overexpression as done in the Cdx2-based leukemia mouse model in this study might protect.

In summary, the impact of miR-196b on tumor pathogenesis will be dependent on the balance of miR-196b expression regulation and its target molecule associations (oncogenes or tumor suppressor genes), similar to what has been postulated for its sister miRNA miR196a associating with the High-mobility group A (HMGA) genes 1 and 2 in tumor pathogenesis (Chen et al. 2011).

4.4 MiR-196b functions in a cell context dependent manner

In healthy stem and progenitor cells, miR-196b is highly expressed, but cannot be detected in lymphoid cells (Petriv et al. 2010). Thus, the intrinsic context determined by the cell type affected miR-196b expression levels, which consequently may determine miR-196b

function. In leukemia, miR-196b can act inhibitory in T-ALL, CML, APL or as shown by the present study also in *Hox* gene driven AML (Bhatia et al. 2011, Liu et al. 2013, Kawasaki and Taira 2004). However, in other hematological malignancies such as the myelodysplastic syndrome or *MLLr*-driven leukemia miR-196b can be promoting, although HOXA9 is the major driver in both AML subtypes and mRNA target of miR-196b (Popovic et al. 2009, Velu et al. 2014, Meyer et al. 2018, Wen et al. 2017). An explanation for the different role of miR-196b in *MLLr* versus *Cdx2*-driven murine AML could be different stoichiometric ratios of miRNA and mRNA affecting miRNA affinities to mRNAs determining the miRNA's function. Stoichometry is influenced by mRNA and miRNA level changes. For example in the *MLLr*-driven murine AML model *Hoxa9* mRNA expression levels were 80-fold higher than in *Cdx2*-driven murine AML investigated by Faber et al. (2013). Those high *Hoxa9* levels might fish/consume most of miR-196b molecules without an inhibitory effect on the leukemic transformation. Consequently, less miR-199b molecules will be available for targeting other genetic leukemia drivers so that overall an enhanced phenotype upon miR-196b overexpression will be the result. Moreover, it might be that in the different cellular contexts different alternative 3'UTR isoforms of genes are available affecting the efficiency of miRNA binding to target mRNA. This could be another reason for different miR-196b post-transcriptional target regulation (Nam et al. 2014). On the other hand, if miRNA expression levels increase the stoichiometry is affected, too, accordingly to the miRNA expression level. Thus, the level of genetically engineered miR-196b overexpression decides over a small or a big subset of gene modulation through miR-196b and the following consequences on biological processes (Mayya and Duchaine 2015). In addition to the ratio of miR-196b and its mRNA targets, the sister miRNA of miR-196b, namely miR-196a, which differs from miR-196b by only a single nucleotide polymorphism, might influence target affinities of miR-196b and therewith its function by playing either a competitive or additive role to miR-196b (Alvarez-Saavedra and Horvitz 2010). Data in favor for the link between miR-196a and miR-196b are: a positive correlation between miR-196a and miR-196b in CN-AML (data not shown), similar expression pattern of both miRNAs among different AML subtypes, a very similar target spectrum revealed by enrichment analyses of *Cdx2*/miR-196b co-expressing murine cells (data not shown) as well as a tumor suppressor function of both miRNAs in solid cancers and leukemia documented in the literature (Li et al. 2010, Cao et al. 2018, Mueller and Bosserhoff 2011, Coskun et al. 2011). Those stoichiometric aspects can be used to explain all sorts of different effects upon genetically engineered overexpression in different cellular contexts including the different inhibitory strength of miR-196b in murine Cdx2 or human OCI-AML3 cells.

In conclusion, the role of miR-196b will be influenced by: the effect of relative miRNA and mRNA quantities (stoichometry) on miR-196b-target interactions, the activity state of miR-196b (determined by the incorporation of miR-196b into the RISC) and the subcellular location of miR-196b, because, those stoichometric aspects can place miR-196b in completely different regulatory networks (Denzler et al. 2016, Lai, Wolkenhauer, et al. 2016, Olejniczak et al. 2013).

4.5 Technical aspects for investigating miR-196b function

To ensure effective vector packaging and gene transfer, two independent retroviral vectors for *Cdx2* and miR-196b were used to co-transduce murine HSPCs, which are cells that can be easily transduced. This transduction method served the purpose of the study, but transduction efficiency was still limited, probably due to viral interference, which does not

occur with lentiviral vectors (Frimpong and Spector 2000). Lentiviral vectors (higher virion stability) were used in harder to transduce leukemic AML cells. With both the retroviral and lentiviral vector transduction efficiencies were between 10-50%. Higher transduction rates were avoided to prevent too many genomic integrations which, preferentially with retroviruses, can lead to higher frequencies of RIM which may cause the second hit in cancer induction (Dufait et al. 2012, TaKaRa 2021). However, also double transduction presumably bears an increased risk for RIM. In this case, not miR-196b would have caused the described effects in the context of *Cdx2* expression, but rather the side effects of a double transduction. An effect due to RIM can only be reliably excluded with a thermal asymmetric interlaced (TAIL)-PCR (Lin et al. 2014). TAIL-PCR is a tool to identify unknown flanking regions of a known target sequence within the nuclear genome via a set of nested sequence-specific primers together with a shorter arbitrary degenerate primer (Liu and Whittier 1995). Another aspect of vector-based assays is the method of transducing them. An innovative method is the usage of RetroNectin, also used in this study for retroviral transduction. RetroNectin coated wells of a culture plate co-localizes cell and virus via integrin binding, has low cytotoxicity and preserves stem cell quality (Tonks et al. 2005). That means, cells with a lot of RetroNectin-relevant integrins (VL4-4 or VLA-5) will probably get preferably transduced with this method. Taken this further, this integrin feature of the RetrNnectin based transduction method might lead to the enrichment of a certain cell type out of the heterogeneous HSPC cell mix to be transduced. In addition, integrin binding might trigger signaling cascades in the cell affecting its intracellular condition. This is not an irrelevant thought, considering that RectroNectin is a recombinant fibronectin fragment and fibronectin is well known to be an essential player during hematopoiesis. On the other hand it is unclear how long lasting the Retroectin-induced signaling cascade in the transduced cells is and thus effects miRNA function (Wirth et al. 2020). However, those speculations might help to understand variations between experimental replicates. A good alternative to the viral-based miRNA integration are exogenously delivered ‘synthetic miRNAs (also known as miRNA mimics) [, that can be] used to mimic the function of the endogenous miRNAs’ in order to restore the loss of function in leukemic cells Lam et al. 2017, p. 1). Mimics are a very attractive tool for use in therapy ‘due to their better safety profile and lower production cost’ (Lam et al. 2017, p. 9), once the transduction hurdle in certain settings can get resolved (Hanna et al. 2019). Therefore, the use of miR-196b mimics in basic research would give the results more translational relevance. Also the usage of transgenic or knock-in mouse models in a cell type specific fashion might be a good alternative approach to the irradiation-based transplantation assays of murine HSPCs used in this study (Chopra and Bohlander 2019). Transplanted cells in mice, that have been irradiated, will undergo a very rapid expansion shortly after the transplantation procedure, which can lead to additional cooperating mutations skewing the result. However, transgenic or knock-in mouse models are more difficult to establish than BM transplantation assays.

5-FU generated murine HSPCs

5-Fluorouracil (5-FU) (nucleotide analog) is an anticancer drug, whose 5-FU metabolites are misincorporated into RNA and DNA and inhibit the enzyme thymidylate synthase. Consequently DNA synthesis and repair are decreased causing damage in rapidly dividing cells such as cancer cells (Noordhuis et al. 2004, Longley et al. 2003). In research, also in this study, the cytotoxic agent 5-FU was used to eliminate many of the rapidly cycling, more committed progenitor cells of the BM in order to enrich for stem and very primitive progenitor cells in mice. However, 5-FU also acts as an injury-like activator of HSCs, which is a

reversible effect, but this takes some time (Wilson et al. 2008). Additionally, 5-FU induces the transcription of several genes in murine HSCs including important hematopoietic stem cell regulators such as *c-kit* or serine protease granzyme B (Randall and Weissman 1997, Carnevalli et al. 2014). Very critically this means that the murine stem and progenitor cells used for miR-196b overexpression experiments in this study might have not been in a completely unbiased physiological state at the time of miRNA transduction, which might have influenced the outcome of this study. To avoid these uncertainties about 5-FU based chemical manipulations for the generation of murine HSPCs cell sorts for Lin^-/low $Sca\text{-}1^+$ positive cells could be used instead.

4.6 Identification of context dependent mRNA targets is essential for understanding miRNA function

MRNA targets will determine the predominant effect of miRNAs in a given context. If oncogenic molecules are targeted, the miRNA will play a tumor suppressor function, if the miRNA mainly targets tumor suppressors, it has an oncogenic effect. However, the aim of this study was not to identify new miR-196b targets, but to identify the functional role of miR-196b and speculate on targets that can give rise to that phenotype.

RNA-Seq analysis of murine HSPCs with elevated *Hox* gene expression levels induced by *Cdx2* revealed that miR-196b reduced the expression of pro-apoptotic, MAPK-associated, proliferation and self-renewal genes. In parallel, miR-196b overexpression lead to increased levels of potential leukemia mediator genes and tumor suppressors. Interestingly, there were also view upregulated proto-oncogenes (e.g. *Jun* and *Fos*) for which two plausible explanations are considered: either (a) the cells tried to resist the miR-196b induced inhibitory effect on cell growth through the acquisition of additional proto-oncogene expression levels - a distant analogy to the ability of bacteria to resist antibiotics (Reygaert 2018) or (b) the overexpressed proto-oncogenes are not relevant for the cell, since on protein level they are not phosphorylated and thus not activated as this was for example shown for the proto-oncogene Syk in Hoxa9/Meis1-driven leukemia (Mohr et al. 2017). In any case, the role of upregulated proto-oncogenes remains elusive. The focus in this study, however, lied specifically on the downregulation of self-renewal and proliferative genes as they are known to be aberrantly expressed in leukemia with different emphasis depending on the leukemia subtype. Consequently, in human OCI-AML3 cells, the effect of miR-196b overexpression was specifically tested for genes of those categories via RT-PCR. Among them were well-known HOX (category: self-renewal) and MAPK-pathway (category: proliferation) relevant genes.

HOX and *MAPK* targets of miR-196b

HOX genes are key regulators in self-renewal (Alharbi et al. 2013). Important contributors in acute myeloid leukemogenesis are *Hoxa9* and *Hoxa7*, which were found to be downregulated in RNA-Seq data of *Cdx2* overexpressing mouse cells and in qRT-PCR data of human OCI-AML3 cells upon miR-196b overexpression in this study (Novak et al. 2012, Ayton and Cleary 2003). Both *HOX* genes are known to induce severe perturbations during normal hematopoietic development in human and murine experimental models. (Kroon et al. 1998, Nakamura et al. 1996, Thorsteinsdottir et al. 1999, Calvo et al. 2002). Thus, *HOXA9* and *HOXA7* inhibition via miR-196b can be essential for its tumor-suppressive effect in this study.

MAPK-pathways are the main pathways regulating cell proliferation (cell division) (Morrison 2012, Geest and Coffer 2009). Downregulated genes associated with proliferation were *Nras*, *Mapk1* (*Erk2*), *Rap1a* and *Mapk8*. One particular MAPK-pathway is the ERK-pathway made of the Ras/Raf/MEK/ERK cascade, which has been found to be particularly relevant in leukemia. In particular ERK isoforms are currently tested as targets in anticancer treatment as it contributes to cell growth, survival and transformation of hematopoietic cells (Weisberg et al. 2020, Liu et al. 2018, Yap et al. 2011). Thus, it was an attractive discovery that the MAPK1-related *MAPK3* gene was downregulated by miR-196b in the human AML cell line OCI-AML3. The RAS subfamily comprises HRAS, KRAS and NRAS, which are frequently mutated in cancer (Goitre et al. 2014). Here, *Nras* was downregulated in *Cdx2*/miR-196b transduced murine HSPCs. ERK isoforms (Erk1/Erk2) have been found to contribute to cell growth, survival and transformation of hematopoietic cells and are currently tested as targets in anticancer treatment. Whereas in murine *Cdx2*-induced leukemic cells *Erk2* (*Mapk1*) was downregulated upon miR-196b overexpression, it was *ERK1* (*MAPK3*) in the human OCI-AML3 cell line. To check whether this mRNA/miRNA link would be functionally relevant, translated biologically active protein levels should be determined via Western Blot. The affinity of miR-196b to MAPK pathway proteins was also found by Meyer et al. (2018) showing KRAS and MAPK3 enrichment in THP1 and/or HEK293T cells after biotinylated miR-196b pull down. Notably, this information is not directly in line with the main massage of their paper saying that miR-196b maintains leukemia stemness, but KRAS and MAPK3 can be found as target genes of miR-196b in supplementary data (Meyer et al. 2018). In addition, Guo et al. (2019) found direct targeting (using dual luciferase reporter assays) and downregulation of the ERK pathway regulator MAP3K1 by miR-196b in human choriocarcinoma cells via western blot analysis (Guo et al. 2019).

Non-*Hox* targets of miR-196b

Outstanding non-*Hox* target genes of miR-196b in *Cdx2* overexpressing mouse cells were the self-renewal relevant genes *Npm1*, *Pbx1* and *Msi2.* In OCI-AML3 cells, however, those genes remained unaffected (for NPM1 data were not shown in the result section). Three possible explanations are: First, the miRNA/mRNA pairing was not strong enough due to reasons such as mutated 3' UTR regions or other altered pairing circumstances beyond the seed region (Zhang et al. 2016, Broughton et al. 2016). Second, the miRNA affects those targets on translational level rather than via mRNA degradation. Third, *PBX1* and *MSI2* are simply not targets of miR-196b in OCI-AML3 cells. Moreover, in literature it was documented that miR-196b inhibits the proto-oncogene c-Myc in B-ALL, so that attention was drawn to c-Myc associated DEGs in RNA-Seq data of *Cdx2/miR-196b* co-overexpressing mouse cells generated in this study (Bhatia et al. 2010). Indeed, the Myc binding protein (*Mycbp*) was downregulated in murine RNA-Seq data, possibly preventing the activation of c-Myc and its associated oncogenic effect in the *Cdx2* induced leukemic mouse model used in this study (FDR<0.3, data not shown). Furthermore, *Mllt3* (alias Af9), which was not reported as miR-196b specific target, but is known to function in stem cell regulation was found to be downregulated in the murine RNA-Seq data of this study (FDR<0.3, data not shown) (Calvanese et al. 2019). Thus, in order to gain a true mechanistic insight into the complex network of miR-196b functioning, it is probably useful not only to focus on direct but also on indirect targets. To experimentally validate a direct or indirect miRNA-target interaction the luciferase reporter assays could be conducted to show whether one particular miRNA inhibits the luciferase signal proportionally to the target input amount (Kuhn et al. 2008,

Elton and Yalowich 2015). Other potential direct non-*HOX* targets with a relation to high expression levels in CN-AML patients and cancer relevance were *SDCBP* and *EPS15* (Qian et al. 2013, Kotecha et al. 2012).

To assess how relevant each of the suspected miR-196b targets is for the inhibition of the leukemic program in HOX gene driven leukemic cells, one could apply stable shRNA- or a CRISPR-Cas9 mediated knockout of those targets and see if the same phenotype occurs as with the miRNA overexpression approach used in this study (Bassett et al. 2014). Alternatively, si-RNAs specific for potential target mRNAs can be used as for example done for the tumor suppressor of cytokine signaling 2 gene, which was identified as target of miR-196b in hepatocellular carcinoma cells (Ren et al. 2019).

Beyond RNA-Seq

RNA-Seq based mRNA target suggestions for miR-196b used in this study offer a great starting point for further gene network investigations, but have some biases. For example, smaller but relevant changes tend to get overseen in RNA-Seq analysis, which focuses mainly on statistically strong mRNA difference with FDR-values <0.1. However, it is important to keep in mind, that small differences on mRNA level can lead to large changes of the protein level if a certain threshold is crossed (Mukherji et al. 2011). In order to keep this error as low as possible in this study, differentially expressed genes with an FDR value of <0.1 but also <0.3 were used. Mainly miRNA-mediated repression occurs via transcript degradation, but there is accumulating evidence that there are gene targets whose translation is repressed (Guo, Ingolia, et al. 2010, Pillai et al. 2005). Thus, an overlap of transcriptomic RNA-Seq data with proteomic profiling data as done for investigating the function of miR-126 in AML LSCs would allow to capture the full capacity of miR-196b in regulating a complex network of genes (Lechman et al. 2016). In future, instead of identifying potential target genes of miR-196b in RNA-Seq data upon miR-196b overexpression, target gens should be detected in a natural context utilizing methods such as HITS-CLIP (high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation), PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation), iCLIP (individual-nucleotide resolution crosslinking immunoprecipitation) or CLASH (cross-linking, ligation, and sequencing of hybrids) (Chi et al. 2009, Hafner et al. 2010, Helwak and Tollervey 2014, Broughton et al. 2016). All these experimental approaches have the advantage over a computational analysis that not only canonical 5’ seed region bindings but also non-canonical bindings, which make up around 60% of the bindings will be revealed and false-positive results of in silico approaches which can also be up to 60% will be avoided (Helwak et al. 2013, Rajewsky 2006).

4.7 Therapy

This study revealed an inhibitory role of miR-196b on *Cdx2* induced murine leukemia BM cells with minimal effects on healthy cells. Thus, an enforced expression of miR-196b could in theory be a possibility to target Hox and ParaHox driven leukemic cells without harming normal HSCs as a more specific treatment alternative to the already investigated drug azacitidine in the Cdx2 and Hox gene context (Vu et al. 2020). Furthermore, this study provides first evidence that miR-196b plays a yet unrecognized role in CN-AML. A delivery of miR-196b mimics might represent a novel treatment strategy, which needs to be evaluated in preclinical leukemia models as has been already discussed for this miRNA in

colorectal cancer (Stiegelbauer et al. 2017). The advantage of miRNAs over siRNAs as therapeutic agents is, that miRNAs 'inhibit the expression of a number of genes, which often work together as a network within the same cellular pathway' (Lam et al. 2015, p. 13), particularly interesting for particularly interesting for multigenetic diseases such as neurodegenerative disorders or cancer (Lam et al. 2015, Krek et al. 2005). Nevertheless, it can occur that miR-196b alone might not be sufficient to achieve a tumor suppressive effect. In this case, it makes sense to make use of the synergy effect. That means, that several miRNAs need to work together in order to inhibit a central oncogene. For example, in cancer, at least 28 miRNAs are involved in inhibiting the oncogene *Cdkn1a* (Wu et al. 2010). The synergy effect is also useful to achieve a tissue-specific anti-tumor effect while healthy cells remain unaffected as already investigated for glioma cells (Yao et al. 2014). Another approach to increase the inhibitory effect of miR-196b might be to reduce the methylation status of its promoter to increase miR-196b levels as it was shown to be effective in CML (Liu et al. 2013). In addition, the aspect of progression seems to be particularly relevant for the effect of miR-196b as indicated by (a) strong effects in *Cdx2*-driven transforming cells into leukemic cells compared to mild effects of miR-196b in already transformed leukemic OCI-AML3 cells and by (b) the finding from Kuchenbauer et al. (2008), who investigated miRNA expression levels in a murine *Hox*-based leukemia progression model by deep sequencing preleukemic cells generated via the overexpression of nucleoporin 98 (*NUP98*) and *HOXD13* (ND13) versus leukemic cells generated via the overexpression of ND13 and *Meis1*, and found miR-196b to be the most prominent differentially expressed miRNA in this preleukemic/leukemic comparison (Kuchenbauer et al. 2008). Based on these data, miR-196b interventions should probably rather focus on the preleukemic state. Despite all, always the individual molecular context of the AML patient should be monitored before treatment with miR-196b, because the inhibitory effect of miR-196b cannot be generalized to all HOX gene driven AML types. This suggests the comparison of data from this study with published data indicating that miR-196b plays an inhibitory role in CN-AML but an oncogenic role in *MLLr*-AML, although both AMLs are HOX gene driven.

Moving beyond miR-196b related treatment possibilities, the results of the study by Díaz-Beyá et al. (2014) suggest that miR-196b may also have a prognostic value in intermediate risk group AML patients out of which 72% had a normal karyotype, which is in line with the overall survival analysis from Li et al. (2012) (Díaz-Beyá et al. 2014, Li, Huang, et al. 2012). Albeit that in those studies high miR-196b levels were associated with worse outcome, which is contrary to the findings of this study, it might be worth to investigate whether miR-196b is secreted into the serum in a progression-dependent manner and can thus function as biomarker in CN-AML patients. This could be for example investigated via the fast and low cost but highly sensitive method for the detection of miRNAs in human serum published by Ma et al. (2017). With this method a miRNA of interest is detected by hybridizing to the complementary loop domain of a hairpin probe (HP), which leads to an exonuclease I mediated release of 2-Aminopurine (2-AP) from the overhanging nucleotides, which former built the double stranded stem of the HP. Freed 2-AP gives a fluorescence signal, which is directly proportional to the amount of investigated miRNA present. If the target miRNA is not present, 2-AP remains incorporated into the double stranded stem sequence of the HP so that no fluorescent signal occurs. However, the fact that miR-196b expression levels did not associate with survival in CN-AML TCGA data reanalyzed in this study argues against a biomarker function (n=83, p=0.9073, data not shown).

5. Summary

Acute myeloid leukemia (AML) accounts for 50% of leukemia-related deaths and thus it is necessary to identify tumor suppressors, that can counteract proto-oncogenes such as the clustered homeobox A9 (*HOXA9*) gene and the non-clustered caudal-type *HOX* gene *CDX2*, which is known to act upstream of *HOX* genes reinforcing their expression. The micro ribonuclease acid (miRNA) 196a/b was shown to cleave *HOXB8* and *HOXA9* mRNA, repectively, in leukemia cell lines suggesting a tumor-suppressive function, particularly in HOX-driven AMLs such as mixed lineage leukemia (*MLL*)-rearranged leukemia and cytogenetically normal (CN)-AML. However, investigations of miR-196b in *MLLr*-leukemia suggest an oncogenic role. In this study, the role of miR-196b in CN-AML was investigated by using murine and human cell line models as well as primary cells. It was found, that miR-196b originates from a non-coding transcript in the HOXA9-HOXA10 locus, whose expression is likely to be regulated via a close by promoter-like region. Two non-coding transcript variants encoding the miR-196b stem-loop sequence and parts of the *HOXA9* gene were identified in healthy human bone marrow (BM) cells via standard polymerase chain reaction (PCR): the longer transcript MF139050 with 486 bps and the shorter transcript variant with 396 bps. The shorter transcript lacks a 90 bps long fragment, but the remaining sequence is 99.47% homologous to the longer transcript. Both transcripts are highly evolutionary conserved to other vertebrates and in particular to mammals. Promoter activity of the promoter-like region (802 bps) directly upstream of the miR-196b transcripts was shown via a luciferase reporter assay. This promoter-like region was located several nucleotides upstream of the classical HOXA9 promoter indicating a miR-196b associated promoter region, which is independent of HOXA9 transcription regulation. Further, it was found that miR-196b was not differentially expressed in CN-AML patient samples compared to healthy cells using quantitative real time PCR. However, the relative expression of miR-196b to its transcript and target *HOXA9* was grossly perturbed in favor for *HOXA9*. Functional data revealed that the retroviral overexpression of miR-196b reduced the clonogenic potential investigated in semi-solid medium and significantly delayed leukemogenesis in a Hox-driven mouse model induced by retroviral *Cdx2* overexpression in primary murine hematopoietic stem and progenitor cells. Lentiviral overexpression of miR-196b impaired leukemic growth of HOX-driven human AML cell line cells in semi-solid medium and in the BM of transplanted mice. Overexpression of miR-196b in healthy BM cells had no biologically significant *in vitro* effect and no obvious pathological *in vivo* effect. Those results underscore a growth inhibitory role of miR-196b in leukemic cells but not normal cells by downregulating predicted target gene sets of mitogen-activated protein kinase (MAPK) pathways (e.g. *MAPK1/3*) or of leukemic stem cell relevant genes such as *HOXA9 and HOXA7*. Consequently, a restoration of miR-196b expression could have a substantial anti-leukemic effect in the large cohort of CN-AML patients. However, therapeutic strategies aiming to achieve a miR-196b mediated anti-leukemic function in AML will have to consider the individual molecular context.

6. References

Adolfsson, J., R. Månsson, N. Buza-Vidas, A. Hultquist, K. Liuba, C. T. Jensen, D. Bryder, L. Yang, O.-J. Borge, L. A. M. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson and S. E. W. Jacobsen (2005). Identification of Flt3+ Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential: A Revised Road Map for Adult Blood Lineage Commitment. *Cell, 121*(2), 295-306.

Agarwal, V., G. W. Bell, J.-W. Nam and D. P. Bartel (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife, 4*, Article e05005.

Aggarwal, R., J. Lu, V. J. Pompili and H. Das (2012). Hematopoietic stem cells: transcriptional regulation, ex vivo expansion and clinical application. *Current molecular medicine, 12*(1), 34-49.

Akashi, K., D. Traver, T. Miyamoto and I. L. Weissman (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature, 404*(6774), 193-197.

Alarcón, C. R., H. Lee, H. Goodarzi, N. Halberg and S. F. Tavazoie (2015). N6-methyladenosine marks primary microRNAs for processing. *Nature, 519*(7544), 482-485.

Alcalay, M., E. Tiacci, R. Bergomas, B. Bigerna, E. Venturini, S. P. Minardi, N. Meani, D. Diverio, L. Bernard, L. Tizzoni, S. Volorio, L. Luzi, E. Colombo, F. Lo Coco, C. Mecucci, B. Falini, P. G. Pelicci and P. for the Gruppo Italiano Malattie Ematologiche Maligne dell'Adulto Acute Leukemia Working (2005). Acute myeloid leukemia bearing cytoplasmic nucleophosmin (NPMc+ AML) shows a distinct gene expression profile characterized by up-regulation of genes involved in stem-cell maintenance. *Blood, 106*(3), 899-902.

Alenzi, F. Q., B. Q. Alenazi, S. Y. Ahmad, M. L. Salem, A. A. Al-Jabri and R. K. H. Wyse (2009). The haemopoietic stem cell: between apoptosis and self renewal. *The Yale journal of biology and medicine, 82*(1), 7-18.

Alexander, T., C. Nolte and R. Krumlauf (2009). Hox Genes and Segmentation of the Hindbrain and Axial Skeleton. *Annual Review of Cell and Developmental Biology, 25*(1), 431-456.

Alharbi, R. A., R. Pettengell, H. S. Pandha and R. Morgan (2013). The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia, 27*, 1000-1008.

Almada, A. E., X. Wu, A. J. Kriz, C. B. Burge and P. A. Sharp (2013). Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature, 499*(7458), 360-363.

Altschuler, S. J. and L. F. Wu (2010). Cellular heterogeneity: do differences make a difference? *Cell, 141*(4), 559-563.

Alvarez-Saavedra, E. and H. R. Horvitz (2010). Many families of C. elegans microRNAs are not essential for development or viability. *Current biology, 20*(4), 367-373.

Andersson, R., Y. Chen, L. Core, J. T. Lis, A. Sandelin and T. H. Jensen (2015). Human Gene Promoters Are Intrinsically Bidirectional. *Molecular cell, 60*(3), 346-347.

Andreeff, M., V. Ruvolo, S. Gadgil, C. Zeng, K. Coombes, W. Chen, S. Kornblau, A. E. Barón and H. A. Drabkin (2008). HOX expression patterns identify a common signature for favorable AML. *Leukemia, 22*(11), 2041-2047.

Arber, D. A., A. Orazi, R. Hasserjian, J. Thiele, M. J. Borowitz, M. M. Le Beau, C. D. Bloomfield, M. Cazzola and J. W. Vardiman (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood, 127*(20), 2391-2405.

Argiropoulos, B. and R. K. Humphries (2007). Hox genes in hematopoiesis and leukemogenesis. *Oncogene, 26*(47), Article 6766.

Arroyo, J. D., J. R. Chevillet, E. M. Kroh, I. K. Ruf, C. C. Pritchard, D. F. Gibson, P. S. Mitchell, C. F. Bennett, E. L. Pogosova-Agadjanyan, D. L. Stirewalt, J. F. Tait and M. Tewari (2011). Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proceedings of the National Academy of Sciences, 108*(12), 5003-5008.

Aryal, S., Y. Zhang, S. Wren, C. Li and R. Lu (2021). Molecular regulators of HOXA9 in acute myeloid leukemia. *The FEBS Journal*, Article 16268.

Ayton, P. M. and M. L. Cleary (2003). Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes & development, 17*(18), 2298-2307.

Bach, C., S. Buhl, D. Mueller, M.-P. García-Cuéllar, E. Maethner and R. K. Slany (2010). Leukemogenic transformation by HOXA cluster genes. *Blood, 115*(14), 2910-2918.

Bai, X., L. Meng, H. Sun, Z. Li, X. Zhang and S. Hua (2017). MicroRNA-196b Inhibits Cell Growth and Metastasis of Lung Cancer Cells by Targeting Runx2. *Cellular Physiology and Biochemistry, 43*(2), 757-767.

Bajic, V. B., M. R. Brent, R. H. Brown, A. Frankish, J. Harrow, U. Ohler, V. V. Solovyev and S. L. Tan (2006). Performance assessment of promoter predictions on ENCODE regions in the EGASP experiment. *Genome Biology, 7*(1), S3.

Bakocevic, N., C. Claser, S. Yoshikawa, L. A. Jones, S. Chew, C. C. Goh, B. Malleret, A. Larbi, F. Ginhoux, M. C. de Lafaille, H. Karasuyama, L. Renia and L. G. Ng (2014). CD41 is a reliable identification and activation marker for murine basophils in the steady state and during helminth and malarial infections. *European Journal of Immunology, 44*(6), 1823-1834.

Bandeira, M. A. and M. E. Boeck (2020). Histone modifications at the transcription start site of non-coding RNA revealed regulatory patterns of activation in Caenorhabditis elegans embryo. *bioRxiv*.

Banerjee, A., F. Schambach, C. S. DeJong, S. M. Hammond and S. L. Reiner (2010). Micro-RNA-155 inhibits IFN-γ signaling in CD4+ T cells. *European Journal of Immunology, 40*(1), 225-231.

Barski, A., R. Jothi, S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones and K. Zhao (2009). Chromatin poises miRNA- and protein-coding genes for expression. *Genome Research, 19*(10), 1742-1751.

Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell, 116*(2), 281-297.

Bartel, D. P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell, 136*(2), 215-233.

Bartelmez, S. H., R. G. Andrews and I. D. Bernstein (1991). Uncovering the heterogeneity of hematopoietic repopulating cells. *Experimental Hematology, 19*(9), 861-862.

Bassett, A. R., G. Azzam, L. Wheatley, C. Tibbit, T. Rajakumar, S. McGowan, N. Stanger, P. A. Ewels, S. Taylor, C. P. Ponting, J.-L. Liu, T. Sauka-Spengler and T. A. Fulga (2014). Understanding functional miRNA–target interactions in vivo by site-specific genome engineering. *Nature Communications, 5*, Article 4640.

Bayraktar, R., M. T. S. Bertilaccio and G. A. Calin (2019). The Interaction Between Two Worlds: MicroRNAs and Toll-Like Receptors. *Frontiers in immunology, 10*, Article 1053.

Becker, H., G. Marcucci, K. Maharry, M. D. Radmacher, K. Mrózek, D. Margeson, S. P. Whitman, Y.-Z. Wu, S. Schwind, P. Paschka, B. L. Powell, T. H. Carter, J. E. Kolitz, M. Wetzler, A. J. Carroll, M. R. Baer, M. A. Caligiuri, R. A. Larson and C. D. Bloomfield (2010). Favorable prognostic impact of NPM1 mutations in older patients with cytogenetically normal de novo acute myeloid leukemia and associated gene- and microRNA-expression signatures: a Cancer and Leukemia Group B study. *Journal of clinical oncology, 28*(4), 596-604.

Beckmann, J., S. Scheitza, P. Wernet, J. C. Fischer and B. Giebel (2007). Asymmetric cell division within the human hematopoietic stem and progenitor cell compartment: identification of asymmetrically segregating proteins. *Blood, 109*(12), 5494-5501.

Behre, G., L. T. Smith and D. G. Tenen (1999). Use of a promoterless Renilla luciferase vector as an internal control plasmid for transient co-transfection assays of Ras-mediated transcription activation. *Biotechniques, 26*(1), 24-28.

Ben-Ami, O., N. Pencovich, J. Lotem, D. Levanon and Y. Groner (2009). A regulatory interplay between miR-27a and Runx1 during megakaryopoiesis. *Proceedings of the National Academy of Sciences, 106*(1), 238-243.

Bessière, C., M. Saraswat, M. Grapotte, C. Menichelli, J. A. Ramilowski, J. Severin, Y. Hayashizaki, M. Itoh, A. Hasegawa, H. Suzuki, P. Carninci, M. J. L. d. Hoon, W. W. Wasserman, L. Bréhélin and C.-H. Lecellier (2019). Evidence of transcription at polyT short tandem repeats. *bioRxiv*.

Bezman, N. A., T. Chakraborty, T. Bender and L. L. Lanier (2011). miR-150 regulates the development of NK and iNKT cells. *The Journal of experimental medicine, 208*(13), 2717-2731.

Bhatia, S., D. Kaul and N. Varma (2011). Functional genomics of tumor suppressor miR-196b in T-cell acute lymphoblastic leukemia. *Molecular and Cellular Biochemistry, 346*(1), 103-116.

Bhatlekar, S., J. Z. Fields and B. M. Boman (2018). Role of HOX Genes in Stem Cell Differentiation and Cancer. *Stem Cells International, 2018*, Article 3569493.

Bielewicz, D., M. Kalak, M. Kalyna, D. Windels, A. Barta, F. Vazquez, Z. Szweykowska-Kulinska and A. Jarmolowski (2013). Introns of plant pri-miRNAs enhance miRNA biogenesis. *EMBO reports, 14*(7), 622-628.

Bill, M., D. Nicolet, J. Kohlschmidt, C. J. Walker, K. Mrózek, A.-K. Eisfeld, D. Papaioannou, X. Rong-Mullins, Z. Brannan, J. E. Kolitz, B. L. Powell, K. J. Archer, A. M. Dorrance,

A. J. Carroll, R. M. Stone, J. C. Byrd, R. Garzon and C. D. Bloomfield (2020). Mutations associated with a 17-gene leukemia stem cell score and the score's prognostic relevance in the context of the European LeukemiaNet classification of acute myeloid leukemia. *Haematologica, 105*(3), 721-729.

Bissels, U., A. Bosio and W. Wagner (2012). MicroRNAs are shaping the hematopoietic landscape. *Haematologica, 97*(2), 160-167.

Bloomston, M., W. L. Frankel, F. Petrocca, S. Volinia, H. Alder, J. P. Hagan, C.-G. Liu, D. Bhatt, C. Taccioli and C. M. Croce (2007). MicroRNA Expression Patterns to Differentiate Pancreatic Adenocarcinoma From Normal Pancreas and Chronic Pancreatitis. *JAMA, 297*(17), 1901-1908.

Bochner, B. S. and R. P. Schleimer (2001). Mast cells, basophils, and eosinophils: distinct but overlapping pathways for recruitment. *Immunological Reviews, 179*(1), 5-15.

Boeva, V. (2016). Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells. *Frontiers in Genetics, 7*, Article 24.

Boivin, V., G. Deschamps-Francoeur and M. S. Scott (2018). Protein coding genes as hosts for noncoding RNA expression. *Seminars in Cell & Developmental Biology, 75*(2018), 3-12.

Boldin, M. P., K. D. Taganov, D. S. Rao, L. Yang, J. L. Zhao, M. Kalwani, Y. Garcia-Flores, M. Luong, A. Devrekanli, J. Xu, G. Sun, J. Tay, P. S. Linsley and D. Baltimore (2011). miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *Journal of Experimental Medicine, 208*(6), 1189-1201.

Bonnet, D. and J. E. Dick (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine, 3*(7), 730-737.

Borchert, G. M., W. Lanier and B. L. Davidson (2006). RNA polymerase III transcribes human microRNAs. *Nature Structural & Molecular Biology, 13*(12), 1097-1101.

Borrow, J., A. M. Shearman, V. P. Stanton, R. Becher, T. Collins, A. J. Williams, I. Dubé, F. Katz, Y. L. Kwong, C. Morris, K. Ohyashiki, K. Toyama, J. Rowley and D. E. Housman (1996). The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP96 and class I homeoprotein HOXA9. *Nature Genetics, 12*(2), 159-167.

Bottini, S., N. Hamouda-Tekaya, R. Mategot, L.-E. Zaragosi, S. Audebert, S. Pisano, V. Grandjean, C. Mauduit, M. Benahmed, P. Barbry, E. Repetto and M. Trabucchi (2017). Post-transcriptional gene silencing mediated by microRNAs is controlled by nucleoplasmic Sfpq. *Nature Communications, 8*, Article 1189.

Bracken, A. P., N. Dietrich, D. Pasini, K. H. Hansen and K. Helin (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes & development, 20*(9), 1123-1136.

Braconi, C., N. Huang and T. Patel (2010). MicroRNA-dependent regulation of DNA methyltransferase-1 and tumor suppressor gene expression by interleukin-6 in human malignant cholangiocytes. *Hepatology (Baltimore, Md.), 51*(3), 881-890.

Braekeleer, E. D., N. Douet-Guilbert, A. Basinko, M.-J. L. Bris, F. Morel and M. D. Braekeleer (2014). Hox gene dysregulation in acute myeloid leukemia. *Future Oncology, 10*(3), 475-495.

Braig, S., D. W. Mueller, T. Rothhammer and A.-K. Bosserhoff (2010). MicroRNA miR-196a is a central regulator of HOX-B7 and BMP4 expression in malignant melanoma. *Cellular and Molecular Life Sciences, 67*(20), 3535-3548.

Brauchle, E., S. Thude, S. Y. Brucker and K. Schenke-Layland (2014). Cell death stages in single apoptotic and necrotic cells monitored by Raman microspectroscopy. *Scientific Reports, 4*, Article 4698.

Broughton, J. P., M. T. Lovci, J. L. Huang, G. W. Yeo and A. E. Pasquinelli (2016). Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Molecular cell, 64*(2), 320-333.

Burda, P., P. Laslo and T. Stopka (2010). The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. *Leukemia, 24*(7), 1249-1257.

Bürglin, T. (2005). Homeodomain proteins. *Encylopedia of Molecular Cell Biology and Molecular Medicine*. R. Meyers. Weinheim, Wiley-VCH Verlag GmbH & Co. *2*.

Butler, J. E. F. and J. T. Kadonaga (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & Development, 16*(20), 2583-2592.

Byrd, J. C., K. Mrózek, R. K. Dodge, A. J. Carroll, C. G. Edwards, D. C. Arthur, M. J. Pettenati, S. R. Patil, K. W. Rao, M. S. Watson, P. R. K. Koduru, J. O. Moore, R. M. Stone, R. J. Mayer, E. J. Feldman, F. R. Davey, C. A. Schiffer, R. A. Larson and C. D. Bloomfield (2002). Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia. *Blood, 100*(13), 4325-4336.

Calvanese, V., A. T. Nguyen, T. J. Bolan, A. Vavilina, T. Su, L. K. Lee, Y. Wang, F. D. Lay, M. Magnusson, G. M. Crooks, S. K. Kurdistani and H. K. A. Mikkola (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. *Nature, 576*(7786), 281-286.

Calvo, K. R., D. B. Sykes, M. P. Pasillas and M. P. Kamps (2002). Nup98-HoxA9 immortalizes myeloid progenitors, enforces expression of Hoxa9, Hoxa7 and Meis1, and alters cytokine-specific responses in a manner similar to that induced by retroviral co-expression of Hoxa9 and Meis1. *Oncogene, 21*(27), 4247-4256.

Cammarata, G., L. Augugliaro, D. Salemi, C. Agueli, M. L. Rosa, L. Dagnino, G. Civiletto, F. Messana, A. Marfia, M. G. Bica, L. Cascio, P. M. Floridia, A. M. Mineo, M. Russo, F. Fabbiano and A. Santoro (2010). Differential expression of specific microRNA and their targets in acute myeloid leukemia. *American Journal of Hematology, 85*(5), 331-339.

Cancer Genome Atlas Research, N., T. J. Ley, C. Miller, L. Ding, B. J. Raphael, A. J. Mungall, A. G. Robertson, K. Hoadley, T. J. Triche, Jr., P. W. Laird, J. D. Baty, L. L. Fulton, R. Fulton, S. E. Heath, J. Kalicki-Veizer, C. Kandoth, J. M. Klco, D. C. Koboldt, K.-L. Kanchi, S. Kulkarni, T. L. Lamprecht, D. E. Larson, L. Lin, C. Lu, M. D. McLellan, J. F. McMichael, J. Payton, H. Schmidt, D. H. Spencer, M. H. Tomasson, J. W. Wallis, L. D. Wartman, M. A. Watson, J. Welch, M. C. Wendl, A. Ally, M. Balasundaram, I. Birol, Y. Butterfield, R. Chiu, A. Chu, E. Chuah, H.-J. Chun, R. Corbett, N. Dhalla, R. Guin, A. He, C. Hirst, M. Hirst, R. A. Holt, S. Jones, A.

Karsan, D. Lee, H. I. Li, M. A. Marra, M. Mayo, R. A. Moore, K. Mungall, J. Parker, E. Pleasance, P. Plettner, J. Schein, D. Stoll, L. Swanson, A. Tam, N. Thiessen, R. Varhol, N. Wye, Y. Zhao, S. Gabriel, G. Getz, C. Sougnez, L. Zou, M. D. M. Leiserson, F. Vandin, H.-T. Wu, F. Applebaum, S. B. Baylin, R. Akbani, B. M. Broom, K. Chen, T. C. Motter, K. Nguyen, J. N. Weinstein, N. Zhang, M. L. Ferguson, C. Adams, A. Black, J. Bowen, J. Gastier-Foster, T. Grossman, T. Lichtenberg, L. Wise, T. Davidsen, J. A. Demchok, K. R. M. Shaw, M. Sheth, H. J. Sofia, L. Yang, J. R. Downing and G. Eley (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *The New England journal of medicine, 368*(22), 2059-2074.

Cao, Y., B. Wu, D. Wang and Z. Bi (2018). miR-196 acts as a tumor suppressor in osteosarcoma by targeting HOXA9. *Int J Clin Exp Pathol 11*(9), 4579-4584.

Caocci, G., M. Greco and G. La Nasa (2017). Bone Marrow Homing and Engraftment Defects of Human Hematopoietic Stem and Progenitor Cells. *Mediterranean journal of hematology and infectious diseases, 9*(1), Article e2017032.

Carnevalli, L. S., R. Scognamiglio, N. Cabezas-Wallscheid, S. Rahmig, E. Laurenti, K. Masuda, L. Jöckel, A. Kuck, S. Sujer, A. Polykratis, M. Erlacher, M. Pasparakis, M. A. G. Essers and A. Trumpp (2014). Improved HSC reconstitution and protection from inflammatory stress and chemotherapy in mice lacking granzyme B. *The Journal of experimental medicine, 211*(5), 769-779.

Carninci, P., A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C. A. M. Semple, M. S. Taylor, P. G. Engström, M. C. Frith, A. R. R. Forrest, W. B. Alkema, S. L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S. M. Grimmond, C. A. Wells, V. Orlando, C. Wahlestedt, E. T. Liu, M. Harbers, J. Kawai, V. B. Bajic, D. A. Hume and Y. Hayashizaki (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genetics, 38*(6), 626-635.

Carroll, A. P., P. A. Tooney and M. J. Cairns (2013). Context-specific microRNA function in developmental complexity. *Journal of Molecular Cell Biology, 5*(2), 73-84.

Catovsky, D., E. Matutes, V. Buccheri, V. Shetty, J. Hanslip, N. Yoshida and R. Morilla (1991). A classification of acute leukaemia for the 1990s. *Annals of Hematology, 62*(1), 16-21.

Challen, G. A., N. Boles, K. K.-Y. Lin and M. A. Goodell (2009). Mouse hematopoietic stem cell identification and analysis. *Cytometry. Part A : the journal of the International Society for Analytical Cytology, 75*(1), 14-24.

Chang, H., B. Yi, R. Ma, X. Zhang, H. Zhao and Y. Xi (2016). CRISPR/cas9, a novel genomic tool to knock down microRNA in vitro and in vivo. *Scientific Reports, 6*(1), Article 22312.

Chang, T.-C., D. Yu, Y.-S. Lee, E. A. Wentzel, D. E. Arking, K. M. West, C. V. Dang, A. Thomas-Tikhonenko and J. T. Mendell (2008). Widespread microRNA repression by Myc contributes to tumorigenesis. *Nature genetics, 40*(1), 43-50.

Chekmenev, D. S., C. Haid and A. E. Kel (2005). P-Match: transcription factor binding site search by combining patterns and weight matrices. *Nucleic Acids Research, 33*(suppl_2), W432-W437.

Chen, C.-Z., L. Li, H. F. Lodish and D. P. Bartel (2004). MicroRNAs Modulate Hematopoietic Lineage Differentiation. *Science, 303*(5654), 83-86.

Chen, C., D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak and K. J. Guegler (2005). Real-time quantification of microRNAs by stem–loop RT–PCR. *Nucleic Acids Research, 33*(20), Article e179.

Chen, C., Y. Zhang, L. Zhang, S. M. Weakley and Q. Yao (2011). MicroRNA-196: critical roles and clinical applications in development and cancer. *Journal of Cellular and Molecular Medicine, 15*(1), 14-23.

Chen, E. Y., C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. V. Meirelles, N. R. Clark and A. Ma'ayan (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics, 14*, Article 128.

Chen, L., L. Heikkinen, C. Wang, Y. Yang, H. Sun and G. Wong (2019). Trends in the development of miRNA bioinformatics tools. *Briefings in Bioinformatics, 20*(5), 1836-1852.

Chen, S.-L., Z.-Y. Qin, F. Hu, Y. Wang, Y.-J. Dai and Y. Liang (2019). The Role of the HOXA Gene Family in Acute Myeloid Leukemia. *Genes, 10*(8), Article 621.

Chen, X., D. B. Burkhardt, A. A. Hartman, X. Hu, A. E. Eastman, C. Sun, X. Wang, M. Zhong, S. Krishnaswamy and S. Guo (2019). MLL-AF9 initiates transformation from fast-proliferating myeloid progenitors. *Nature Communications, 10*(1), Article 5767.

Chen, Z., E. Stelekati, M. Kurachi, S. Yu, Z. Cai, S. Manne, O. Khan, X. Yang and E. J. Wherry (2017). miR-150 Regulates Memory CD8 T Cell Differentiation via c-Myb. *Cell reports, 20*(11), 2584-2597.

Cheng, H., Z. Zheng and T. Cheng (2020). New paradigms on hematopoietic stem cell differentiation. *Protein & Cell, 11*(1), 34-44.

Cheng, J., S. Guo, S. Chen, S. J. Mastriano, C. Liu, A. C. D'Alessio, E. Hysolli, Y. Guo, H. Yao, C. M. Megyola, D. Li, J. Liu, W. Pan, C. A. Roden, X.-L. Zhou, K. Heydari, J. Chen, I.-H. Park, Y. Ding, Y. Zhang and J. Lu (2013). An extensive network of TET2-targeting MicroRNAs regulates malignant hematopoiesis. *Cell reports, 5*(2), 471-481.

Chi, S. W., J. B. Zang, A. Mele and R. B. Darnell (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature, 460*(7254), 479-486.

Chomczynski, P. and N. Sacchi (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry, 162*(1), 156-159.

Chong, G. O., H.-S. Jeon, H. S. Han, J. W. Son, Y. H. Lee, D. G. Hong, H. J. Park, Y. S. Lee and Y. L. Cho (2017). Overexpression of microRNA-196b Accelerates Invasiveness of Cancer Cells in Recurrent Epithelial Ovarian Cancer Through Regulation of Homeobox A9. *Cancer genomics & proteomics, 14*(2), 137-141.

Chopra, M. and S. K. Bohlander (2019). The cell of origin and the leukemia stem cell in acute myeloid leukemia. *Genes, Chromosomes and Cancer, 58*(12), 850-858.

Chotinantakul, K. and W. Leeanansaksiri (2012). Hematopoietic stem cell development, niches, and signaling pathways. *Bone marrow research, 2012*, Article 270425.

Chung, K. Y., G. Morrone, J. J. Schuringa, M. Plasilova, J.-H. Shieh, Y. Zhang, P. Zhou and M. A. S. Moore (2006). Enforced Expression of NUP98-HOXA9 in Human

CD34+Cells Enhances Stem Cell Proliferation. *Cancer Research, 66*(24), Article 11781.

Cichocki, F., M. Felices, V. McCullar, S. R. Presnell, A. Al-Attar, C. T. Lutz and J. S. Miller (2011). Cutting edge: microRNA-181 promotes human NK cell development by regulating Notch signaling. *Journal of immunology (Baltimore, Md. : 1950), 187*(12), 6171-6175.

Cillo, C., M. Cantile, A. Faiella and E. Boncinelli (2001). Homeobox genes in normal and malignant cells. *Journal of Cellular Physiology, 188*(2), 161-169.

Cobb, B. S., A. Hertweck, J. Smith, E. O'Connor, D. Graf, T. Cook, S. T. Smale, S. Sakaguchi, F. J. Livesey, A. G. Fisher and M. Merkenschlager (2006). A role for Dicer in immune regulation. *The Journal of experimental medicine, 203*(11), 2519-2527.

Congdon, K. L. and T. Reya (2008). Divide and conquer: how asymmetric division shapes cell fate in the hematopoietic system. *Current opinion in immunology, 20*(3), 302-307.

Connell, R. M., A. A. Chaudhuri, D. S. Rao, W. S. J. Gibson, A. B. Balazs and D. Baltimore (2010). MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output. *Proceedings of the National Academy of Sciences, 107*(32), Article 14235.

Cook, W. D., B. J. McCaw, C. Herring, D. L. John, S. J. Foote, S. L. Nutt and J. M. Adams (2004). PU.1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain. *Blood, 104*(12), 3437-3444.

Coskun, E., E. K. von der Heide, C. Schlee, A. Kühnl, N. Gökbuget, D. Hoelzer, W.-K. Hofmann, E. Thiel and C. D. Baldus (2011). The role of microRNA-196a and microRNA-196b as ERG regulators in acute myeloid leukemia and acute T-lymphoblastic leukemia. *Leukemia Research, 35*(2), 208-213.

Cusan, M. (2012). *Studie über das leukämogene Potential der Hoxb4-ΔProlin Mutante in einem murinen Knochenmarkstransplantationsmodell und den Effekt der Hemmung von Histondeacetylasen in vitro und in vivo.* Doctoral dissertation, Ludwig-Maximilian-University Munich.

Dahlin, J. S. and J. Hallgren (2015). Mast cell progenitors: Origin, development and migration to tissues. *Molecular Immunology, 63*(1), 9-17.

Dangi-Garimella, S., M. J. Strouch, P. J. Grippo, D. J. Bentrem and H. G. Munshi (2011). Collagen regulation of let-7 in pancreatic cancer involves TGF-β1-mediated membrane type 1-matrix metalloproteinase expression. *Oncogene, 30*(8), 1002-1008.

Dard, A., Y. Jia, J. Reboulet, F. Bleicher, C. Lavau and S. Merabet (2019). The human HOXA9 protein uses paralog-specific residues of the homeodomain to interact with TALE-class cofactors. *Scientific Reports, 9*(1), Article 5664.

Dassé, E., G. Volpe, D. S. Walton, N. Wilson, W. Del Pozzo, L. P. O'Neill, R. K. Slany, J. Frampton and S. Dumon (2012). Distinct regulation of c-myb gene expression by HoxA9, Meis1 and Pbx proteins in normal hematopoietic progenitors and transformed myeloid cells. *Blood Cancer Journal, 2*(6), Article e76.

de Leeuw, D. C., F. Denkers, M. C. Olthof, A. P. Rutten, W. Pouwels, G. Jan Schuurhuis, G. J. Ossenkoppele and L. Smit (2014). Attenuation of microRNA-126 Expression That Drives $CD34^{+}38^{-}$ Stem/Progenitor Cells in Acute Myeloid Leukemia Leads to Tumor Eradication. *Cancer Research, 74*(7), Article 2094.

de Pater, E., P. Kaimakis, C. S. Vink, T. Yokomizo, T. Yamada-Inagawa, R. van der Linden, P. S. Kartalaei, S. A. Camper, N. Speck and E. Dzierzak (2013). Gata2 is required for HSC generation and survival. *The Journal of experimental medicine, 210*(13), 2843-2850.

de Rie, D., I. Abugessaisa, T. Alam, E. Arner, P. Arner, H. Ashoor, G. Åström, M. Babina, N. Bertin, A. M. Burroughs, A. J. Carlisle, C. O. Daub, M. Detmar, R. Deviatiiarov, A. Fort, C. Gebhard, D. Goldowitz, S. Guhl, T. J. Ha, J. Harshbarger, A. Hasegawa, K. Hashimoto, M. Herlyn, P. Heutink, K. J. Hitchens, C. C. Hon, E. Huang, Y. Ishizu, C. Kai, T. Kasukawa, P. Klinken, T. Lassmann, C.-H. Lecellier, W. Lee, M. Lizio, V. Makeev, A. Mathelier, Y. A. Medvedeva, N. Mejhert, C. J. Mungall, S. Noma, M. Ohshima, M. Okada-Hatakeyama, H. Persson, P. Rizzu, F. Roudnicky, P. Sætrom, H. Sato, J. Severin, J. W. Shin, R. K. Swoboda, H. Tarui, H. Toyoda, K. Vitting-Seerup, L. Winteringham, Y. Yamaguchi, K. Yasuzawa, M. Yoneda, N. Yumoto, S. Zabierowski, P. G. Zhang, C. A. Wells, K. M. Summers, H. Kawaji, A. Sandelin, M. Rehli, Y. Hayashizaki, P. Carninci, A. R. R. Forrest, M. J. L. de Hoon and F. C. The (2017). An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nature Biotechnology, 35*(9), 872-878.

Debernardi, S., D. M. Lillington, T. Chaplin, S. Tomlinson, J. Amess, A. Rohatiner, T. A. Lister and B. D. Young (2003). Genome-wide analysis of acute myeloid leukemia with normal karyotype reveals a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. *Genes, Chromosomes and Cancer, 37*(2), 149-158.

Debernardi, S., S. Skoulakis, G. Molloy, T. Chaplin, A. Dixon-McIver and B. D. Young (2007). MicroRNA miR-181a correlates with morphological sub-class of acute myeloid leukaemia and the expression of its target genes in global genome-wide analysis. *Leukemia, 21*, Article 912.

Den Hartogh, S. C. and R. Passier (2016). Concise Review: Fluorescent Reporters in Human Pluripotent Stem Cells: Contributions to Cardiac Differentiation and Their Applications in Cardiac Disease and Toxicity. *STEM CELLS, 34*(1), 13-26.

Denli, A. M., B. B. J. Tops, R. H. A. Plasterk, R. F. Ketting and G. J. Hannon (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature, 432*(7014), 231-235.

Denzler, R., S. E. McGeary, A. C. Title, V. Agarwal, D. P. Bartel and M. Stoffel (2016). Impact of MicroRNA Levels, Target-Site Complementarity, and Cooperativity on Competing Endogenous RNA-Regulated Gene Expression. *Molecular cell, 64*(3), 565-579.

Díaz-Beyá, M., S. Brunet, J. Nomdedéu, R. Tejero, T. Díaz, M. Pratcorona, M. Tormo, J. M. Ribera, L. Escoda, R. Duarte, D. Gallardo, I. Heras, M. P. Queipo de Llano, J. Bargay, M. Monzo, J. Sierra, A. Navarro, J. Esteve and A. M. L. g. C. on behalf of the Cooperative (2014). MicroRNA expression at diagnosis adds relevant prognostic information to molecular categorization in patients with intermediate-risk cytogenetic acute myeloid leukemia. *Leukemia, 28*(4), 804-812.

Dick, J. E. (2003). Stem cells: Self-renewal writ in blood. *Nature, 423*(6937), 231-233.

Dickins, R. A., M. T. Hemann, J. T. Zilfou, D. R. Simpson, I. Ibarra, G. J. Hannon and S. W. Lowe (2005). Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nature Genetics, 37*(11), 1289-1295.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson and T. R. Gingeras (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21.

Döhner, H., E. Estey, D. Grimwade, S. Amadori, F. R. Appelbaum, T. Büchner, H. Dombret, B. L. Ebert, P. Fenaux, R. A. Larson, R. L. Levine, F. Lo-Coco, T. Naoe, D. Niederwieser, G. J. Ossenkoppele, M. Sanz, J. Sierra, M. S. Tallman, H.-F. Tien, A. H. Wei, B. Löwenberg and C. D. Bloomfield (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood, 129*(4), 424-447.

Döhner, H., E. H. Estey, S. Amadori, F. R. Appelbaum, T. Büchner, A. K. Burnett, H. Dombret, P. Fenaux, D. Grimwade, R. A. Larson, F. Lo-Coco, T. Naoe, D. Niederwieser, G. J. Ossenkoppele, M. A. Sanz, J. Sierra, M. S. Tallman, B. Löwenberg and C. D. Bloomfield (2010). Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood, 115*(3), 453-474.

Döhner, H., D. J. Weisdorf and C. D. Bloomfield (2015). Acute Myeloid Leukemia. *New England Journal of Medicine, 373*(12), 1136-1152.

Dong, F., H. Bai, X. Wang, S. Zhang, Z. Wang, M. Xie, S. Zhang, J. Wang, S. Hao, T. Cheng and H. Ema (2019). Mouse acute leukemia develops independent of self-renewal and differentiation potentials in hematopoietic stem and progenitor cells. *Blood advances, 3*(3), 419-431.

Dörner, T., A. Radbruch and G. R. Burmester (2009). B-cell-directed therapies for autoimmune disease. *Nature Reviews Rheumatology, 5*(8), 433-441.

Doulatov, S., F. Notta, E. Laurenti and John E. Dick (2012). Hematopoiesis: A Human Perspective. *Cell Stem Cell, 10*(2), 120-136.

Drew, E., H. Merkens, S. Chelliah, R. Doyonnas and K. M. McNagny (2002). CD34 is a specific marker of mature murine mast cells. *Experimental Hematology, 30*(10), 1211-1218.

Dufait, I., T. Liechtenstein, A. Lanna, C. Bricogne, R. Laranga, A. Padella, K. Breckpot and D. Escors (2012). Retroviral and Lentiviral Vectors for the Induction of Immunological Tolerance. *Scientifica, 2012*, Article 694137.

Duttke, S. H. C., S. A. Lacadie, M. M. Ibrahim, C. K. Glass, D. L. Corcoran, C. Benner, S. Heinz, J. T. Kadonaga and U. Ohler (2015). Human promoters are intrinsically directional. *Molecular cell, 57*(4), 674-684.

Duursma, A. M., M. Kedde, M. Schrier, C. le Sage and R. Agami (2008). miR-148 targets human DNMT3b protein coding region. *RNA (New York, N.Y.), 14*(5), 872-877.

Dwivedi, S., P. Purohit and P. Sharma (2019). MicroRNAs and Diseases: Promising Biomarkers for Diagnosis and Therapeutics. *Indian Journal of Clinical Biochemistry, 34*(3), 243-245.

Dzierzak, E. and S. Philipsen (2013). Erythropoiesis: development and differentiation. *Cold Spring Harbor perspectives in medicine, 3*(4), Article a011601.

Ebert, M. S. and P. A. Sharp (2010). MicroRNA sponges: progress and possibilities. *RNA (New York, N.Y.), 16*(11), 2043-2050.

Ebert, M. S. and P. A. Sharp (2012). Roles for microRNAs in conferring robustness to biological processes. *Cell, 149*(3), 515-524.

Edmaier, K. E. (2013). *Dissecting the Role of Lef-1 in Normal Hematopoiesis.* Doctoral dissertation, University of Ulm.

Ellis, J. (2005). Silencing and Variegation of Gammaretrovirus and Lentivirus Vectors. *Human Gene Therapy, 16*(11), 1241-1246.

Ellwanger, D. C., F. A. Büttner, H.-W. Mewes and V. Stümpflen (2011). The sufficient minimal set of miRNA seed types. *Bioinformatics (Oxford, England), 27*(10), 1346-1350.

Elton, T. S. and J. C. Yalowich (2015). Experimental procedures to identify and validate specific mRNA targets of miRNAs. *EXCLI journal, 14*, 758-790.

Ender, C. and G. Meister (2010). Argonaute proteins at a glance. *Journal of Cell Science, 123*(11), Article 1819.

Eppert, K., K. Takenaka, E. R. Lechman, L. Waldron, B. Nilsson, P. van Galen, K. H. Metzeler, A. Poeppl, V. Ling, J. Beyene, A. J. Canty, J. S. Danska, S. K. Bohlander, C. Buske, M. D. Minden, T. R. Golub, I. Jurisica, B. L. Ebert and J. E. Dick (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nature Medicine, 17*(9), 1086-1093.

Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene, 21*(35), 5427-5440.

Evans, T., M. Reitman and G. Felsenfeld (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proceedings of the National Academy of Sciences of the United States of America, 85*(16), 5976-5980.

Faber, J., A. V. Krivtsov, M. C. Stubbs, R. Wright, T. N. Davis, M. van den Heuvel-Eibrink, C. M. Zwaan, A. L. Kung and S. A. Armstrong (2009). HOXA9 is required for survival in human MLL-rearranged acute leukemias. *Blood, 113*(11), 2375-2385.

Faber, K., L. Bullinger, C. Ragu, A. Garding, D. Mertens, C. Miller, D. Martin, D. Walcher, K. Döhner, H. Döhner, R. Claus, C. Plass, S. M. Sykes, S. W. Lane, C. Scholl and S. Fröhling (2013). CDX2-driven leukemogenesis involves KLF4 repression and deregulated PPARγ signaling. *The Journal of clinical investigation, 123*(1), 299-314.

Fantini, S., V. Salsi, L. Reggiani, A. Maiorana and V. Zappavigna (2016). The miR-196b miRNA inhibits the GATA6 intestinal transcription factor and is upregulated in colon cancer patients. *Oncotarget, 8*(3), 4747-4759.

Fantini, S., V. Salsi, A. Vitobello, F. M. Rijli and V. Zappavigna (2015). MicroRNA-196b is transcribed from an autonomous promoter and is directly regulated by Cdx2 and by posterior Hox proteins during embryogenesis. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1849*(8), 1066-1080.

Ficara, F., M. J. Murphy, M. Lin and M. L. Cleary (2008). Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. *Cell stem cell, 2*(5), 484-496.

Fliedner, T. M., D. Graessle, C. Paulsen and K. Reimers (2002). Structure and Function of Bone Marrow Hemopoiesis: Mechanisms of Response to Ionizing Radiation Exposure. *Cancer Biotherapy and Radiopharmaceuticals, 17*(4), 405-426.

Fontana, L., E. Pelosi, P. Greco, S. Racanicchi, U. Testa, F. Liuzzi, C. M. Croce, E. Brunetti, F. Grignani and C. Peschle (2007). MicroRNAs 17-5p–20a–106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. *Nature Cell Biology, 9*(7), 775-787.

Forsberg, E. C., S. S. Prohaska, S. Katzman, G. C. Heffner, J. M. Stuart and I. L. Weissman (2005). Differential Expression of Novel Potential Regulators in Hematopoietic Stem Cells. *PLoS Genetics, 1*(3), 0281-0294.

Forsberg, E. C., T. Serwold, S. Kogan, I. L. Weissman and E. Passegué (2006). New Evidence Supporting Megakaryocyte-Erythrocyte Potential of Flk2/Flt3+ Multipotent Hematopoietic Progenitors. *Cell, 126*(2), 415-426.

Frimpong, K. and S. A. Spector (2000). Cotransduction of nondividing cells using lentiviral vectors. *Gene Therapy, 7*(18), 1562-1569.

Frisch, B. J. and L. M. Calvi (2014). Hematopoietic stem cell cultures and assays. *Methods in molecular biology (Clifton, N.J.), 1130*, 315-324.

Frith, M. C., U. Hansen and Z. Weng (2001). Detection of cis-element clusters in higher eukaryotic DNA. *Bioinformatics, 17*(10), 878-889.

Frith, M. C., J. L. Spouge, U. Hansen and Z. Weng (2002). Statistical significance of clusters of motifs represented by position specific scoring matrices in nucleotide sequences. *Nucleic Acids Res, 30*(14), 3214-3224.

Fu, Y., M. C. Frith, P. M. Haverty and Z. Weng (2004). MotifViz: an analysis and visualization tool for motif discovery. *Nucleic Acids Res, 32*(Web Server issue), W420-423.

Gangal, R. and P. Sharma (2005). Human pol II promoter prediction: time series descriptors and machine learning. *Nucleic Acids Research, 33*(5), 1739-1739.

Garzon, R., M. Garofalo, M. P. Martelli, R. Briesewitz, L. Wang, C. Fernandez-Cymering, S. Volinia, C.-G. Liu, S. Schnittger, T. Haferlach, A. Liso, D. Diverio, M. Mancini, G. Meloni, R. Foa, M. F. Martelli, C. Mecucci, C. M. Croce and B. Falini (2008). Distinctive microRNA signature of acute myeloid leukemia bearing cytoplasmic mutated nucleophosmin. *Proceedings of the National Academy of Sciences of the United States of America, 105*(10), 3945-3950.

Garzon, R., S. Volinia, C.-G. Liu, C. Fernandez-Cymering, T. Palumbo, F. Pichiorri, M. Fabbri, K. Coombes, H. Alder, T. Nakamura, N. Flomenberg, G. Marcucci, G. A. Calin, S. M. Kornblau, H. Kantarjian, C. D. Bloomfield, M. Andreeff and C. M. Croce (2008). MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood, 111*(6), 3183-3189.

Ge, J., Z. Chen, R. Li, T. Lu and G. Xiao (2014). Upregulation of microRNA-196a and microRNA-196b cooperatively correlate with aggressive progression and unfavorable prognosis in patients with colorectal cancer. *Cancer cell international, 14*(1), Article 128.

Geest, C. R. and P. J. Coffer (2009). MAPK signaling pathways in the regulation of hematopoiesis. *Journal of Leukocyte Biology, 86*(2), 237-250.

Gentles, A. J., S. K. Plevritis, R. Majeti and A. A. Alizadeh (2010). Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. *JAMA, 304*(24), 2706-2715.

Gentner, B., N. Pochert, A. Rouhi, F. Boccalatte, T. Plati, T. Berg, S. M. Sun, S. M. Mah, M. Mirkovic-Hösle, J. Ruschmann, A. Muranyi, S. Leierseder, B. Argiropoulos, D. T. Starczynowski, A. Karsan, M. Heuser, D. Hogge, F. D. Camargo, S. Engelhardt, H. Döhner, C. Buske, M. Jongen-Lavrencic, L. Naldini, R. K. Humphries and F. Kuchenbauer (2015). MicroRNA-223 dose levels fine tune proliferation and differentiation in human cord blood progenitors and acute myeloid leukemia. *Experimental hematology, 43*(10), 858-868.

Georgantas, R. W., 3rd, R. Hildreth, S. Morisot, J. Alder, C.-g. Liu, S. Heimfeld, G. A. Calin, C. M. Croce and C. I. Civin (2007). CD34+ hematopoietic stem-progenitor cell microRNA expression and function: a circuit diagram of differentiation control. *Proceedings of the National Academy of Sciences of the United States of America, 104*(8), 2750-2755.

German Cancer Research Center. (2020). *Leukämie bei Erwachsenen: Häufigkeit und Krankheitsentstehung*. *Fact sheet* Retrieved 07.09.2020, from https://www.krebsinformationsdienst.de/tumorarten/leukaemien/risikofaktoren.php.

Gerrits, A., M. A. Walasek, S. Olthof, E. Weersing, M. Ritsema, E. Zwart, R. van Os, L. V. Bystrykh and G. de Haan (2012). Genetic screen identifies microRNA cluster 99b/let-7e/125a as a regulator of primitive hematopoietic cells. *Blood, 119*(2), 377-387.

Gershenzon, N. I. and I. P. Ioshikhes (2005). Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. *Bioinformatics, 21*(8), 1295-1300.

Ghani, S., P. Riemke, J. Schönheit, D. Lenze, J. Stumm, M. Hoogenkamp, A. Lagendijk, S. Heinz, C. Bonifer, J. Bakkers, S. Abdelilah-Seyfried, M. Hummel and F. Rosenbauer (2011). Macrophage development from HSCs requires PU.1-coordinated microRNA expression. *Blood, 118*(8), 2275-2284.

Ghini, F., C. Rubolino, M. Climent, I. Simeone, M. J. Marzi and F. Nicassio (2018). Endogenous transcripts control miRNA levels and activity in mammalian cells by target-directed miRNA degradation. *Nature Communications, 9*(1), Article 3119.

Gilliland, D. G. and J. D. Griffin (2002). The roles of FLT3 in hematopoiesis and leukemia. *Blood, 100*(5), 1532-1542.

Goardon, N., E. Marchi, A. Atzberger, L. Quek, A. Schuh, S. Soneji, P. Woll, A. Mead, K. A. Alford, R. Rout, S. Chaudhury, A. Gilkes, S. Knapper, K. Beldjord, S. Begum, S. Rose, N. Geddes, M. Griffiths, G. Standen, A. Sternberg, J. Cavenagh, H. Hunter, D. Bowen, S. Killick, L. Robinson, A. Price, E. Macintyre, P. Virgo, A. Burnett, C. Craddock, T. Enver, S. E. W. Jacobsen, C. Porcher and P. Vyas (2011). Coexistence of LMPP-like and GMP-like Leukemia Stem Cells in Acute Myeloid Leukemia. *Cancer Cell, 19*(1), 138-152.

Godin, I. (2016). Ontogeny of the Hematopoietic System. *Kaufman's Atlas of Mouse Development Supplement*. R. Baldock, J. Bard, D. R. Davidson and G. Morriss-Kay. Boston, Academic Press: 111-120.

Goitre, L., L. Trabalzini and S. F. Retta (2014). *Ras Signaling Methods and Protocols* Totowa, NJ, Human Press.

Golub, T. R., D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield and E. S. Lander (1999). Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. *Science, 286*(5439), Article 531.

Goyama, S., G. Yamamoto, M. Shimabe, T. Sato, M. Ichikawa, S. Ogawa, S. Chiba and M. Kurokawa (2008). Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell Stem Cell, 3*(2), 207-220.

Grabe, N. (2002). AliBaba2: context specific identification of transcription factor binding sites. *In Silico Biol, 2*(1), S1-15.

Greaves, M. (2015). When One Mutation Is All It Takes. *Cancer Cell, 27*(4), 433-434.

Griffiths-Jones, S. (2018). *miRBase: the microRNA sequence database*. 22. Retrieved 14.09.2020, from http://www.mirbase.org/cgi-bin/browse.pl.

Grishok, A., A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D. L. Baillie, A. Fire, G. Ruvkun and C. C. Mello (2001). Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control C. elegans Developmental Timing. *Cell, 106*(1), 23-34.

Gross, D. S. and W. T. Garrard (1988). NUCLEASE HYPERSENSITIVE SITES IN CHROMATIN. *Annual Review of Biochemistry, 57*(1), 159-197.

Guan, Y., M. Mizoguchi, K. Yoshimoto, N. Hata, T. Shono, S. O. Suzuki, Y. Araki, D. Kuga, A. Nakamizo, T. Amano, X. Ma, K. Hayashi and T. Sasaki (2010). MiRNA-196 Is Upregulated in Glioblastoma But Not in Anaplastic Astrocytoma and Has Prognostic Significance. *Clinical Cancer Research, 16*(16), Article 4289.

Guerriero, I., D. D'Angelo, P. Pallante, M. Santos, M. Scrima, D. Malanga, C. De Marco, M. Ravo, A. Weisz, C. Laudanna, M. Ceccarelli, G. Falco, A. Rizzuto and G. Viglietto (2016). Analysis of miRNA profiles identified miR-196a as a crucial mediator of aberrant PI3K/AKT signaling in lung cancer cells. *Oncotarget, 8*(12), 19172-19191.

Guo, H., N. T. Ingolia, J. S. Weissman and D. P. Bartel (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature, 466*(7308), 835-840.

Guo, S., J. Lu, R. Schlanger, H. Zhang, J. Y. Wang, M. C. Fox, L. E. Purton, H. H. Fleming, B. Cobb, M. Merkenschlager, T. R. Golub and D. T. Scadden (2010). MicroRNA miR-125a controls hematopoietic stem cell number. *Proceedings of the National Academy of Sciences, 107*(32), Article 14229.

Guo, Z., L. Sui, J. Qi, Q. Sun, Y. Xu, N. Zou, Y. Xie and Y. Kong (2019). miR-196b inhibits cell migration and invasion through targeting MAP3K1 in hydatidiform mole. *Biomedicine & Pharmacotherapy, 113*, Article 108760.

Haberle, V. and A. Stark (2018). Eukaryotic core promoters and the functional basis of transcription initiation. *Nature reviews. Molecular cell biology, 19*(10), 621-637.

Hafner, M., M. Landthaler, L. Burger, M. Khorshid, J. Hausser, P. Berninger, A. Rothballer, M. Ascano, Jr., A.-C. Jungkamp, M. Munschauer, A. Ulrich, G. S. Wardle, S. Dewell, M. Zavolan and T. Tuschl (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell, 141*(1), 129-141.

Han, Y.-C., C. Y. Park, G. Bhagat, J. Zhang, Y. Wang, J.-B. Fan, M. Liu, Y. Zou, I. L. Weissman and H. Gu (2010). microRNA-29a induces aberrant self-renewal capacity

in hematopoietic progenitors, biased myeloid development, and acute myeloid leukemia. *The Journal of experimental medicine, 207*(3), 475-489.

Hanekamp, D., J. Cloos and G. J. Schuurhuis (2017). Leukemic stem cells: identification and clinical application. *International Journal of Hematology, 105*(5), 549-557.

Hanna, J., G. S. Hossain and J. Kocerha (2019). The Potential for microRNA Therapeutics and Clinical Research. *Frontiers in Genetics, 10*, Article 478.

Harris, N. L., E. S. Jaffe, J. Diebold, G. Flandrin, H. K. Muller-Hermelink, J. Vardiman, T. A. Lister and C. D. Bloomfield (1999). The World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues. Report of the Clinical Advisory Committee meeting, Airlie House, Virginia, November, 1997. *Ann Oncol, 10*(12), 1419-1432.

Hashimi, S. T., J. A. Fulcher, M. H. Chang, L. Gov, S. Wang and B. Lee (2009). MicroRNA profiling identifies miR-34a and miR-21 and their target genes JAG1 and WNT1 in the coordinate regulation of dendritic cell differentiation. *Blood, 114*(2), 404-414.

Hayes, J., P. P. Peruzzi and S. Lawler (2014). MicroRNAs in cancer: biomarkers, functions and therapy. *Trends in Molecular Medicine, 20*(8), 460-469.

He, L., J. M. Thomson, M. T. Hemann, E. Hernando-Monge, D. Mu, S. Goodson, S. Powers, C. Cordon-Cardo, S. W. Lowe, G. J. Hannon and S. M. Hammond (2005). A microRNA polycistron as a potential human oncogene. *Nature, 435*(7043), 828-833.

Heazlewood, S. Y., A. Oteiza, H. Cao and S. K. Nilsson (2014). Analyzing hematopoietic stem cell homing, lodgment, and engraftment to better understand the bone marrow niche. *Annals of the New York Academy of Sciences, 1310*(1), 119-128.

Hebbes, T. R., A. L. Clayton, A. W. Thorne and C. Crane-Robinson (1994). Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *The EMBO journal, 13*(8), 1823-1830.

Helwak, A., G. Kudla, T. Dudnakova and D. Tollervey (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell, 153*(3), 654-665.

Helwak, A. and D. Tollervey (2014). Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). *Nature protocols, 9*(3), 711-728.

Heuser, M., B. Argiropoulos, F. Kuchenbauer, E. Yung, J. Piper, S. Fung, R. F. Schlenk, K. Dohner, T. Hinrichsen, C. Rudolph, A. Schambach, C. Baum, B. Schlegelberger, H. Dohner, A. Ganser and R. K. Humphries (2007). MN1 overexpression induces acute myeloid leukemia in mice and predicts ATRA resistance in patients with AML. *Blood, 110*(5), 1639-1647.

Heuser, M., Y. Ofran, N. Boissel, S. Brunet Mauri, C. Craddock, J. Janssen, A. Wierzbowska and C. Buske (2020). Acute myeloid leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology, 31*(6), 697-712.

Heuser, M., H. Yun, T. Berg, E. Yung, B. Argiropoulos, F. Kuchenbauer, G. Park, I. Hamwi, L. Palmqvist, C. K. Lai, M. Leung, G. Lin, A. Chaturvedi, B. K. Thakur, M. Iwasaki, M. Bilenky, N. Thiessen, G. Robertson, M. Hirst, D. Kent, N. K. Wilson, B. Göttgens, C. Eaves, M. L. Cleary, M. Marra, A. Ganser and R. K. Humphries (2011). Cell of

origin in AML: susceptibility to MN1-induced transformation is regulated by the MEIS1/AbdB-like HOX protein complex. *Cancer cell, 20*(1), 39-52.

Hinata, K., A. M. Gervin, Y. Jennifer Zhang and P. A. Khavari (2003). Divergent gene regulation and growth effects by NF-κB in epithelial and mesenchymal cells of human skin. *Oncogene, 22*(13), 1955-1964.

Hoffman, A. E., T. Zheng, C. Yi, D. Leaderer, J. Weidhaas, F. Slack, Y. Zhang, T. Paranjape and Y. Zhu (2009). microRNA miR-196a-2 and breast cancer: a genetic and epigenetic association study and functional analysis. *Cancer research, 69*(14), 5970-5977.

Hornstein, E., J. H. Mansfield, S. Yekta, J. K.-H. Hu, B. D. Harfe, M. T. McManus, S. Baskerville, D. P. Bartel and C. J. Tabin (2005). The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. *Nature, 438*(7068), 671-674.

Horton, S. J. and B. J. P. Huntly (2012). Recent advances in acute myeloid leukemia stem cell biology. *Haematologica, 97*(7), 966-974.

Hoss, A. G., V. K. Kartha, X. Dong, J. C. Latourelle, A. Dumitriu, T. C. Hadzi, M. E. Macdonald, J. F. Gusella, S. Akbarian, J.-F. Chen, Z. Weng and R. H. Myers (2014). MicroRNAs located in the Hox gene clusters are implicated in huntington's disease pathogenesis. *PLoS genetics, 10*(2), Article e1004188.

Hsiao, P. C., M. C. Liu, L. M. Chen, C. Y. Tsai, Y. T. Wang, J. Chen and L. S. Hsu (2008). Promoter methylation of p16 and EDNRB gene in leukemia patients in Taiwan. *Chin J Physiol, 51*(1), 27-31.

Hu, W., J. Dooley, S. S. Chung, D. Chandramohan, L. Cimmino, S. Mukherjee, C. E. Mason, B. de Strooper, A. Liston and C. Y. Park (2015). miR-29a maintains mouse hematopoietic stem cell self-renewal by regulating Dnmt3a. *Blood, 125*(14), 2206-2216.

Huang, F., J. Tang, X. Zhuang, Y. Zhuang, W. Cheng, W. Chen, H. Yao and S. Zhang (2014). MiR-196a promotes pancreatic cancer progression by targeting nuclear factor kappa-B-inhibitor alpha. *PloS one, 9*(2), Article e87897.

Huang, H.-Y., Y.-C.-D. Lin, J. Li, K.-Y. Huang, S. Shrestha, H.-C. Hong, Y. Tang, Y.-G. Chen, C.-N. Jin, Y. Yu, J.-T. Xu, Y.-M. Li, X.-X. Cai, Z.-Y. Zhou, X.-H. Chen, Y.-Y. Pei, L. Hu, J.-J. Su, S.-D. Cui, F. Wang, Y.-Y. Xie, S.-Y. Ding, M.-F. Luo, C.-H. Chou, N.-W. Chang, K.-W. Chen, Y.-H. Cheng, X.-H. Wan, W.-L. Hsu, T.-Y. Lee, F.-X. Wei and H.-D. Huang (2020). miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. *Nucleic Acids Research, 48*(D1), D148-D154.

Huang, W. (2017). MicroRNAs: Biomarkers, Diagnostics, and Therapeutics. *Bioinformatics in MicroRNA Research*. J. Huang, G. M. Borchert, D. Dou et al. New York, NY, Springer: 57-67.

Hui, A. B. Y., W. Shi, P. C. Boutros, N. Miller, M. Pintilie, T. Fyles, D. McCready, D. Wong, K. Gerster, I. Jurisica, L. Z. Penn and F.-F. Liu (2009). Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. *Laboratory Investigation, 89*, Article 597.

Humphreys, D. T., B. J. Westman, D. I. K. Martin and T. Preiss (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail

function. *Proceedings of the National Academy of Sciences of the United States of America, 102*(47), 16961-16966.

Ihme, S. (2014). *Characterization of the Leukemia Initiating Cell in Human Acute Myeloid Leukemia.* Doctoral dissertation, University of Ulm.

Ikuta, K., T. Kina, I. MacNeil, N. Uchida, B. Peault, Y.-h. Chien and I. L. Weissman (1990). A developmental switch in thymic lymphocyte maturation potential occurs at the level of hematopoietic stem cells. *Cell, 62*(5), 863-874.

Ishikawa, F., S. Yoshida, Y. Saito, A. Hijikata, H. Kitamura, S. Tanaka, R. Nakamura, T. Tanaka, H. Tomiyama, N. Saito, M. Fukata, T. Miyamoto, B. Lyons, K. Ohshima, N. Uchida, S. Taniguchi, O. Ohara, K. Akashi, M. Harada and L. D. Shultz (2007). Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. *Nature Biotechnology, 25*(11), 1315-1321.

Ishikawa, K., A. Ishihara and H. Moriya (2020). Exploring the Complexity of Protein-Level Dosage Compensation that Fine-Tunes Stoichiometry of Multiprotein Complexes. *PLoS genetics, 16*(10), Article e1009091.

Itahana, K., J. Campisi and G. P. Dimri (2007). Methods to detect biomarkers of cellular senescence: the senescence-associated beta-galactosidase assay. *Methods Mol Biol, 371*, 21-31.

Iwakawa, H.-o. and Y. Tomari (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends in Cell Biology, 25*(11), 651-665.

Jiang, C.-F., Z.-M. Shi, D.-M. Li, Y.-C. Qian, Y. Ren, X.-M. Bai, Y.-X. Xie, L. Wang, X. Ge, W.-T. Liu, L.-L. Zhen, L.-Z. Liu and B.-H. Jiang (2018). Estrogen-induced miR-196a elevation promotes tumor growth and metastasis via targeting SPRED1 in breast cancer. *Molecular Cancer, 17*(1), Article 83.

Jiang, S., C. Li, V. Olive, E. Lykken, F. Feng, J. Sevilla, Y. Wan, L. He and Q.-J. Li (2011). Molecular dissection of the miR-17-92 cluster's critical dual roles in promoting Th1 responses and preventing inducible Treg differentiation. *Blood, 118*(20), 5487-5497.

Jiang, X., C. Hu, S. Arnovitz, J. Bugno, M. Yu, Z. Zuo, P. Chen, H. Huang, B. Ulrich, S. Gurbuxani, H. Weng, J. Strong, Y. Wang, Y. Li, J. Salat, S. Li, A. G. Elkahloun, Y. Yang, M. B. Neilly, R. A. Larson, M. M. Le Beau, T. Herold, S. K. Bohlander, P. P. Liu, J. Zhang, Z. Li, C. He, J. Jin, S. Hong and J. Chen (2016). miR-22 has a potent anti-tumour role with therapeutic potential in acute myeloid leukaemia. *Nature communications, 7*, Article 11452.

Jonas, S. and E. Izaurralde (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews Genetics, 16*(7), 421-433.

Jongen-Lavrencic, M., S. M. Sun, M. K. Dijkstra, P. J. M. Valk and B. Löwenberg (2008). MicroRNA expression profiling in relation to the genetic heterogeneity of acute myeloid leukemia. *Blood, 111*(10), Article 5078.

Jung, N., B. Dai, A. J. Gentles, R. Majeti and A. P. Feinberg (2015). An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis. *Nature Communications, 6*(1), Article 8489.

Juven-Gershon, T. and J. T. Kadonaga (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental biology, 339*(2), 225-229.

Kanhere, A. and M. Bansal (2005). Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes. *Nucleic acids research, 33*(10), 3165-3175.

Kanno, S., K. Nosho, K. Ishigami, I. Yamamoto, H. Koide, H. Kurihara, K. Mitsuhashi, M. Shitani, M. Motoya, S. Sasaki, T. Tanuma, H. Maguchi, T. Hasegawa, Y. Kimura, I. Takemasa, Y. Shinomura and H. Nakase (2017). MicroRNA-196b is an independent prognostic biomarker in patients with pancreatic cancer. *Carcinogenesis, 38*(4), 425-431.

Kao, H.-W., D.-C. Liang, J.-H. Wu, M.-C. Kuo, P.-N. Wang, C.-P. Yang, Y.-S. Shih, T.-H. Lin, Y.-H. Huang and L.-Y. Shih (2014). Gene mutation patterns in patients with minimally differentiated acute myeloid leukemia. *Neoplasia (New York, N.Y.), 16*(6), 481-488.

Karagkouni, D., M. D. Paraskevopoulou, S. Chatzopoulos, I. S. Vlachos, S. Tastsoglou, I. Kanellos, D. Papadimitriou, I. Kavakiotis, S. Maniou, G. Skoufos, T. Vergoulis, T. Dalamagas and A. G. Hatzigeorgiou (2018). DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic acids research, 46*(D1), D239-D245.

Karlic, H., H. Herrmann, F. Varga, R. Thaler, R. Reitermaier, S. Spitzer, V. Ghanim, K. Blatt, W. R. Sperr, P. Valent and M. Pfeilstöcker (2014). The role of epigenetics in the regulation of apoptosis in myelodysplastic syndromes and acute myeloid leukemia. *Critical Reviews in Oncology/Hematology, 90*(1), 1-16.

Karolchik, D., G. P. Barber, J. Casper, H. Clawson, M. S. Cline, M. Diekhans, T. R. Dreszer, P. A. Fujita, L. Guruvadoo, M. Haeussler, R. A. Harte, S. Heitner, A. S. Hinrichs, K. Learned, B. T. Lee, C. H. Li, B. J. Raney, B. Rhead, K. R. Rosenbloom, C. A. Sloan, M. L. Speir, A. S. Zweig, D. Haussler, R. M. Kuhn and W. J. Kent (2014). The UCSC Genome Browser database: 2014 update. *Nucleic acids research, 42*(Database issue), D764-D770.

Kawasaki, H. and K. Taira (2004). MicroRNA-196 inhibits HOXB8 expression in myeloid differentiation of HL60 cells. *Nucleic Acids Symp Ser (Oxf), 48*, 211-212.

Keene, M. A., V. Corces, K. Lowenhaupt and S. C. Elgin (1981). DNase I hypersensitive sites in Drosophila chromatin occur at the 5' ends of regions of transcription. *Proceedings of the National Academy of Sciences, 78*(1), 143-146.

Kent, W. J. (2002). BLAT—The BLAST-Like Alignment Tool. *Genome Research, 12*(4), 656-664.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, Haussler and David (2002). The Human Genome Browser at UCSC. *Genome Research, 12*(6), 996-1006.

Kharas, M. G., C. J. Lengner, F. Al-Shahrour, L. Bullinger, B. Ball, S. Zaidi, K. Morgan, W. Tam, M. Paktinat, R. Okabe, M. Gozo, W. Einhorn, S. W. Lane, C. Scholl, S. Fröhling, M. Fleming, B. L. Ebert, D. G. Gilliland, R. Jaenisch and G. Q. Daley (2010). Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. *Nature Medicine, 16*(8), 903-908.

Kim, V. N. and J.-W. Nam (2006). Genomics of microRNA. *Trends in Genetics, 22*(3), 165-173.

Kim, Y.-K. and V. N. Kim (2007). Processing of intronic microRNAs. *The EMBO journal, 26*(3), 775-783.

Kim, Y. J., S. W. Bae, S. S. Yu, Y. C. Bae and J. S. Jung (2009). miR-196a Regulates Proliferation and Osteogenic Differentiation in Mesenchymal Stem Cells Derived From Human Adipose Tissue. *Journal of Bone and Mineral Research, 24*(5), 816-825.

Klepin, H. D., E. Estey and T. Kadia (2019). More Versus Less Therapy for Older Adults With Acute Myeloid Leukemia: New Perspectives on an Old Debate. *American Society of Clinical Oncology Educational Book,*(39), 421-432.

Klose, R. J., Q. Yan, Z. Tothova, K. Yamane, H. Erdjument-Bromage, P. Tempst, D. G. Gilliland, Y. Zhang and W. G. Kaelin (2007). The Retinoblastoma Binding Protein RBP2 Is an H3K4 Demethylase. *Cell, 128*(5), 889-900.

Kobune, M., S. Iyama, S. Kikuchi, H. Horiguchi, T. Sato, K. Murase, Y. Kawano, K. Takada, K. Ono, Y. Kamihara, T. Hayashi, K. Miyanishi, Y. Sato, R. Takimoto and J. Kato (2012). Stromal cells expressing hedgehog-interacting protein regulate the proliferation of myeloid neoplasms. *Blood Cancer Journal, 2*(9), Article e87.

Kodandapani, R., F. Pio, C.-Z. Ni, G. Piccialli, M. Klemsz, S. McKercher, R. A. Maki and K. R. Ely (1996). A new pattern for helix–turn–helix recognition revealed by the PU.l ETS–domain–DNA complex. *Nature, 380*(6573), 456-460.

Komorowski, M., J. Miekisz and A. M. Kierzek (2009). Translational repression contributes greater noise to gene expression than transcriptional repression. *Biophysical journal, 96*(2), 372-384.

Kondo, M., I. L. Weissman and K. Akashi (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell, 91*(5), 661-672.

Kong, K. Y., K. S. Owens, J. H. Rogers, J. Mullenix, C. S. Velu, H. L. Grimes and R. Dahl (2010). MIR-23A microRNA cluster inhibits B-cell development. *Experimental hematology, 38*(8), 629-640.

Koralov, S. B., S. A. Muljo, G. R. Galler, A. Krek, T. Chakraborty, C. Kanellopoulou, K. Jensen, B. S. Cobb, M. Merkenschlager, N. Rajewsky and K. Rajewsky (2008). Dicer Ablation Affects Antibody Diversity and Cell Survival in the B Lymphocyte Lineage. *Cell, 132*(5), 860-874.

Kotecha, R. S., J. Ford, A. H. Beesley, D. Anderson, C. H. Cole and U. R. Kees (2012). Molecular characterization of identical, novel MLL-EPS15 translocation and individual genomic copy number alterations in monozygotic infant twins with acute lymphoblastic leukemia. *Haematologica, 97*(9), 1447-1450.

Kottaridis, P. D., R. E. Gale and D. C. Linch (2003). Flt3 mutations and leukaemia. *British Journal of Haematology, 122*(4), 523-538.

Krek, A., D. Grün, M. N. Poy, R. Wolf, L. Rosenberg, E. J. Epstein, P. MacMenamin, I. da Piedade, K. C. Gunsalus, M. Stoffel and N. Rajewsky (2005). Combinatorial microRNA target predictions. *Nature Genetics, 37*(5), 495-500.

Krishnaraju, K., B. Hoffman and D. A. Liebermann (2001). Early growth response gene 1 stimulates development of hematopoietic progenitor cells along the macrophage lineage at the expense of the granulocyte and erythroid lineages. *Blood, 97*(5), 1298-1305.

Kroon, E., J. Krosl, U. Thorsteinsdottir, S. Baban, A. M. Buchberg and G. Sauvageau (1998). Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *The EMBO Journal, 17*(13), 3714-3725.

Kuchenbauer, F., R. D. Morin, B. Argiropoulos, O. I. Petriv, M. Griffith, M. Heuser, E. Yung, J. Piper, A. Delaney, A.-L. Prabhu, Y. Zhao, H. McDonald, T. Zeng, M. Hirst, C. L. Hansen, M. A. Marra and R. K. Humphries (2008). In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Research, 18*(11), 1787-1797.

Kucherenko, M. M. and H. R. Shcherbata (2018). miRNA targeting and alternative splicing in the stress response – events hosted by membrane-less compartments. *Journal of Cell Science, 131*(4), Article jcs202002.

Kuhn, D. E., M. M. Martin, D. S. Feldman, A. V. Terry, Jr., G. J. Nuovo and T. S. Elton (2008). Experimental validation of miRNA targets. *Methods (San Diego, Calif.), 44*(1), 47-54.

Kühn, M. W. M., E. Song, Z. Feng, A. Sinha, C.-W. Chen, A. J. Deshpande, M. Cusan, N. Farnoud, A. Mupo, C. Grove, R. Koche, J. E. Bradner, E. de Stanchina, G. S. Vassiliou, T. Hoshii and S. A. Armstrong (2016). Targeting Chromatin Regulators Inhibits Leukemogenic Gene Expression in NPM1 Mutant Leukemia. *Cancer discovery, 6*(10), 1166-1181.

Kuleshov, M. V., M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen and A. Ma'ayan (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research, 44*(W1), W90-W97.

Kuo, G., C.-Y. Wu and H.-Y. Yang (2019). MiR-17-92 cluster and immunity. *Journal of the Formosan Medical Association, 118*(1, Part 1), 2-6.

Kurkewich, J. L., J. Hansen, N. Klopfenstein, H. Zhang, C. Wood, A. Boucher, J. Hickman, D. E. Muench, H. L. Grimes and R. Dahl (2017). The miR-23a~27a~24-2 microRNA cluster buffers transcription and signaling pathways during hematopoiesis. *PLoS genetics, 13*(7), Article e1006887.

Labbaye, C., I. Spinello, M. T. Quaranta, E. Pelosi, L. Pasquini, E. Petrucci, M. Biffoni, E. R. Nuzzolo, M. Billi, R. Foà, E. Brunetti, F. Grignani, U. Testa and C. Peschle (2008). A three-step pathway comprising PLZF/miR-146a/CXCR4 controls megakaryopoiesis. *Nature Cell Biology, 10*(7), 788-801.

Lai, C. K., G. L. Norddahl, T. Maetzig, P. Rosten, T. Lohr, L. Sanchez Milde, N. von Krosigk, T. R. Docking, M. Heuser, A. Karsan and R. K. Humphries (2017). Meis2 as a critical player in MN1-induced leukemia. *Blood Cancer Journal, 7*(9), Article e613.

Lai, M., A. Gonzalez-Martin, A. B. Cooper, H. Oda, H. Y. Jin, J. Shepherd, L. He, J. Zhu, D. Nemazee and C. Xiao (2016). Regulation of B-cell development and tolerance by different members of the miR-17~92 family microRNAs. *Nature communications, 7*, Article 12207.

Lai, X., O. Wolkenhauer and J. Vera (2016). Understanding microRNA-mediated gene regulatory networks through mathematical modelling. *Nucleic acids research, 44*(13), 6019-6035.

Lam, J. K. W., M. Y. T. Chow, Y. Zhang and S. W. S. Leung (2015). siRNA Versus miRNA as Therapeutics for Gene Silencing. *Molecular Therapy - Nucleic Acids, 4*, Article e252.

Lambert, M., M. Alioui, S. Jambon, S. Depauw, I. V. Seuningen and M.-H. David-Cordonnier (2019). Direct and Indirect Targeting of HOXA9 Transcription Factor in Acute Myeloid Leukemia. *Cancers, 11*(6), Article 837.

Lapidot, T., C. Sirard, J. Vormoor, B. Murdoch, T. Hoang, J. Caceres-Cortes, M. Minden, B. Paterson, M. A. Caligiuri and J. E. Dick (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature, 367*(6464), 645-648.

Laura, C., M. Sharon, U. Paolo, W. Yulei, I. Masayuki, L. Silvia, A. Achille, S. C. Federico, C.-S. Carmelo, L. C. Michael, V. Anna, G. Bernhard and F. Francesca (2019). MicroRNA-127-3p controls murine hematopoietic stem cell maintenance by limiting differentiation. *Haematologica, 104*(9), 1744-1755.

Laurenti, E., S. Doulatov, S. Zandi, I. Plumb, J. Chen, C. April, J.-B. Fan and J. E. Dick (2013). The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. *Nature immunology, 14*(7), 756-763.

Laurenti, E. and B. Göttgens (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature, 553*(7689), 418-426.

Laverdière, I., M. Boileau, A. L. Neumann, H. Frison, A. Mitchell, S. W. K. Ng, J. C. Y. Wang, M. D. Minden and K. Eppert (2018). Leukemic stem cell signatures identify novel therapeutics targeting acute myeloid leukemia. *Blood Cancer Journal, 8*(6), Article 52.

Lawrence, H. J., J. Christensen, S. Fong, Y.-L. Hu, I. Weissman, G. Sauvageau, R. K. Humphries and C. Largman (2005). Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood, 106*(12), 3988-3994.

Lechman, E. R., B. Gentner, S. W. K. Ng, E. M. Schoof, P. van Galen, J. A. Kennedy, S. Nucera, F. Ciceri, K. B. Kaufmann, N. Takayama, S. M. Dobson, A. Trotman-Grant, G. Krivdova, J. Elzinga, A. Mitchell, B. Nilsson, K. G. Hermans, K. Eppert, R. Marke, R. Isserlin, V. Voisin, G. D. Bader, P. W. Zandstra, T. R. Golub, B. L. Ebert, J. Lu, M. Minden, J. C. Y. Wang, L. Naldini and J. E. Dick (2016). miR-126 Regulates Distinct Self-Renewal Outcomes in Normal and Malignant Hematopoietic Stem Cells. *Cancer cell, 29*(2), 214-228.

Lechman, E. R., B. Gentner, P. van Galen, A. Giustacchini, M. Saini, F. E. Boccalatte, H. Hiramatsu, U. Restuccia, A. Bachi, V. Voisin, G. D. Bader, J. E. Dick and L. Naldini (2012). Attenuation of miR-126 activity expands HSC in vivo without exhaustion. *Cell stem cell, 11*(6), 799-811.

Lee, C. and C. H. Huang (2013). LASAGNA-Search: an integrated web tool for transcription factor binding site search and visualization. *Biotechniques, 54*(3), 141-153.

Lee, R. C., R. L. Feinbaum and V. Ambros (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell, 75*(5), 843-854.

Lee, S. W., K. C. Park, J. G. Kim, S. J. Moon, S. B. Kang, D. S. Lee, H. J. Sul, J. S. Ji and H. Y. Jeong (2016). Dysregulation of MicroRNA-196b-5p and MicroRNA-375 in Gastric Cancer. *Journal of gastric cancer, 16*(4), 221-229.

Lee, T.-Y., W.-C. Chang, J. B.-K. Hsu, T.-H. Chang and D.-M. Shien (2012). GPMiner: an integrated system for mining combinatorial cis-regulatory elements in mammalian gene group. *BMC genomics, 13 Suppl 1*(Suppl 1), S3-S3.

Lee, Y., K. Jeon, J.-T. Lee, S. Kim and V. N. Kim (2002). MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO journal, 21*(17), 4663-4670.

Lee, Y., M. Kim, J. Han, K.-H. Yeom, S. Lee, S. H. Baek and V. N. Kim (2004). MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal, 23*(20), 4051-4060.

Lehnertz, B., C. Pabst, L. Su, M. Miller, F. Liu, L. Yi, R. Zhang, J. Krosl, E. Yung, J. Kirschner, P. Rosten, T. M. Underhill, J. Jin, J. Hébert, G. Sauvageau, R. K. Humphries and F. M. Rossi (2014). The methyltransferase G9a regulates HoxA9-dependent transcription in AML. *Genes & development, 28*(4), 317-327.

Lenhard, B., A. Sandelin and P. Carninci (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nature Reviews Genetics, 13*(4), 233-245.

Lennox, K. A. and M. A. Behlke (2011). Chemical modification and design of anti-miRNA oligonucleotides. *Gene Therapy, 18*(12), 1111-1120.

Leung, A. K. L. and P. A. Sharp (2010). MicroRNA functions in stress responses. *Molecular cell, 40*(2), 205-215.

Levine, M. and R. Tjian (2003). Transcription regulation and animal diversity. *Nature, 424*(6945), 147-151.

Li, H. S., N. Greeley, N. Sugimoto, Y.-J. Liu and S. S. Watowich (2012). miR-22 controls Irf8 mRNA abundance and murine dendritic cell development. *PloS one, 7*(12), Article e52341.

Li, N., W. Wang, B. Xu and H. Gong (2016). miR-196b regulates gastric cancer cell proliferation and invasion via PI3K/AKT/mTOR signaling pathway. *Oncology Letters, 11*, 1745-1749.

Li, Q., Y. Wu, J. Zhang, T. Yi and W. Li (2016). MicroRNA-130a regulates cell malignancy by targeting RECK in chronic myeloid leukemia. *American journal of translational research, 8*(2), 955-967.

Li, Q., Z. Yang, M. Chen and Y. Liu (2016). Downregulation of microRNA-196a enhances the sensitivity of non-small cell lung cancer cells to cisplatin treatment. *International Journal of Molecular Medicine, 37*, 1067-1074.

Li, X., J. Zhang, L. Gao, S. McClellan, M. A. Finan, T. W. Butler, L. B. Owen, G. A. Piazza and Y. Xi (2012). MiR-181 mediates cell differentiation by interrupting the Lin28 and let-7 feedback circuit. *Cell death and differentiation, 19*(3), 378-386.

Li, Y., M. Zhang, H. Chen, Z. Dong, V. Ganapathy, M. Thangaraju and S. Huang (2010). Ratio of miR-196s to HOXC8 messenger RNA correlates with breast cancer cell migration and metastasis. *Cancer research, 70*(20), 7894-7904.

Li, Z., P. Chen, R. Su, Y. Li, C. Hu, Y. Wang, S. Arnovitz, M. He, S. Gurbuxani, Z. Zuo, A. G. Elkahloun, S. Li, H. Weng, H. Huang, M. B. Neilly, S. Wang, E. N. Olson, R. A. Larson, M. M. Le Beau, J. Zhang, X. Jiang, M. Wei, J. Jin, P. P. Liu and J. Chen

(2015). Overexpression and knockout of miR-126 both promote leukemogenesis. *Blood, 126*(17), 2005-2015.

Li, Z., H. Huang, P. Chen, M. He, Y. Li, S. Arnovitz, X. Jiang, C. He, E. Hyjek, J. Zhang, Z. Zhang, A. Elkahloun, D. Cao, C. Shen, M. Wunderlich, Y. Wang, M. B. Neilly, J. Jin, M. Wei, J. Lu, P. J. M. Valk, R. Delwel, B. Lowenberg, M. M. Le Beau, J. Vardiman, J. C. Mulloy, N. J. Zeleznik-Le, P. P. Liu, J. Zhang and J. Chen (2012). miR-196b directly targets both HOXA9/MEIS1 oncogenes and FAS tumour suppressor in MLL-rearranged leukaemia. *Nature communications, 3*, Article 688.

Li, Z., J. Lu, M. Sun, S. Mi, H. Zhang, R. T. Luo, P. Chen, Y. Wang, M. Yan, Z. Qian, M. B. Neilly, J. Jin, Y. Zhang, S. K. Bohlander, D.-E. Zhang, R. A. Larson, M. M. Le Beau, M. J. Thirman, T. R. Golub, J. D. Rowley and J. Chen (2008). Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. *Proceedings of the National Academy of Sciences, 105*(40), Article 15535.

Liao, R., J. Sun, L. Zhang, G. Lou, M. Chen, D. Zhou, Z. Chen and S. Zhang (2008). MicroRNAs play a role in the development of human hematopoietic stem cells. *Journal of cellular biochemistry, 104*(3), 805-817.

Liao, Y., G. K. Smyth and W. Shi (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics, 30*(7), 923-930.

Lim, H.-C. and C.-S. Kim (2014). Oral signs of acute leukemia for early detection. *J Periodontal Implant Sci, 44*(6), 293-299.

Lim, P. K., S. A. Bliss, S. A. Patel, M. Taborga, M. A. Dave, L. A. Gregory, S. J. Greco, M. Bryan, P. S. Patel and P. Rameshwar (2011). Gap Junction–Mediated Import of MicroRNA from Bone Marrow Stromal Cells Can Elicit Cell Cycle Quiescence in Breast Cancer Cells. *Cancer Research, 71*(5), Article 1550.

Lin, C.-C., L.-Z. Liu, J. B. Addison, W. F. Wonderlin, A. V. Ivanov and J. M. Ruppert (2011). A KLF4-miRNA-206 autoregulatory feedback loop can promote or inhibit protein translation depending upon cell context. *Molecular and cellular biology, 31*(12), 2513-2527.

Lin, J., Q. Zhang, L. Q. Zhu, Q. H. Yu and Q. Yang (2014). The copy number and integration site analysis of IGF-1 transgenic goat. *Int J Mol Med, 34*(3), 900-910.

Liu, F., X. Yang, M. Geng and M. Huang (2018). Targeting ERK, an Achilles' Heel of the MAPK pathway, in cancer therapy. *Acta Pharmaceutica Sinica B, 8*(4), 552-562.

Liu, R. and D. J. States (2002). Consensus promoter identification in the human genome utilizing expressed gene markers and gene modeling. *Genome research, 12*(3), 462-469.

Liu, S.-Q., S. Jiang, C. Li, B. Zhang and Q.-J. Li (2014). miR-17-92 Cluster Targets Phosphatase and Tensin Homology and Ikaros Family Zinc Finger 4 to Promote TH17-mediated Inflammation. *Journal of Biological Chemistry, 289*(18), 12446-12456.

Liu, X.-h., K.-h. Lu, K.-m. Wang, M. Sun, E.-b. Zhang, J.-s. Yang, D.-d. Yin, Z.-l. Liu, J. Zhou, Z.-j. Liu, W. De and Z.-x. Wang (2012). MicroRNA-196a promotes non-small cell lung cancer cell proliferation and invasion through targeting HOXA5. *BMC cancer, 12*, Article 348.

Liu, Y., P. He, F. Liu, L. Shi, H. Zhu, J. Zhao, Y. Wang, X. Cheng and M. Zhang (2014). Prognostic significance of NPM1 mutations in acute myeloid leukemia: A meta-analysis. *Molecular and clinical oncology, 2*(2), 275-281.

Liu, Y., W. Zheng, Y. Song, W. Ma and H. Yin (2013). Low expression of miR-196b enhances the expression of BCR-ABL1 and HOXA9 oncogenes in chronic myeloid leukemogenesis. *PloS one, 8*(7), Article e68442.

Liu, Y. G. and R. F. Whittier (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics, 25*(3), 674-681.

Livak, K. J. and T. D. Schmittgen (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods, 25*(4), 402-408.

Llufrio, E. M., L. Wang, F. J. Naser and G. J. Patti (2018). Sorting cells alters their redox state and cellular metabolome. *Redox biology, 16*, 381-387.

Longley, D. B., D. P. Harkin and P. G. Johnston (2003). 5-Fluorouracil: mechanisms of action and clinical strategies. *Nature Reviews Cancer, 3*(5), 330-338.

Lopez-Serra, P. and M. Esteller (2012). DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. *Oncogene, 31*(13), 1609-1622.

Love, M. I., W. Huber and S. Anders (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), Article 550.

Luthra, R., R. R. Singh, M. G. Luthra, Y. X. Li, C. Hannah, A. M. Romans, B. A. Barkoh, S. S. Chen, J. Ensor, D. M. Maru, R. R. Broaddus, A. Rashid and C. T. Albarracin (2008). MicroRNA-196a targets annexin A1: a microRNA-mediated mechanism of annexin A1 downregulation in cancers. *Oncogene, 27*, Article 6667.

Luzna, P., J. Gregar, I. Uberall, L. Radova, V. Prochazka and J. Ehrmann, Jr. (2011). Changes of microRNAs-192, 196a and 203 correlate with Barrett's esophagus diagnosis and its progression compared to normal healthy individuals. *Diagnostic pathology, 6*, Article 114.

Lynch, M., W. Sung, K. Morris, N. Coffey, C. R. Landry, E. B. Dopman, W. J. Dickinson, K. Okamoto, S. Kulkarni, D. L. Hartl and W. K. Thomas (2008). A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Sciences, 105*(27), Article 9272.

Lytle, J. R., T. A. Yario and J. A. Steitz (2007). Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5′ UTR as in the 3′ UTR. *Proceedings of the National Academy of Sciences, 104*(23), Article 9667.

Ma, C., H. Liu, K. Wu, M. Chen, L. Zheng and J. Wang (2017). An Exonuclease I-Based Quencher-Free Fluorescent Method Using DNA Hairpin Probes for Rapid Detection of MicroRNA. *Sensors (Basel, Switzerland), 17*(4), Article 760.

Ma, R., W. Yan, G. Zhang, H. Lv, Z. Liu, F. Fang, W. Zhang, J. Zhang, T. Tao, Y. You, T. Jiang and X. Kang (2012). Upregulation of miR-196b confers a poor prognosis in glioblastoma patients via inducing a proliferative phenotype. *PloS one, 7*(6), Article e38096.

Madeira, F., Y. m. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn and R. Lopez (2019). The EMBL-EBI search and

sequence analysis tools APIs in 2019. *Nucleic Acids Research, 47*(W1), W636-W641.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville and G. M. Church (2013). RNA-Guided Human Genome Engineering via Cas9. *Science, 339*(6121), Article 823.

Mallo, M., D. M. Wellik and J. Deschamps (2010). Hox genes and regional patterning of the vertebrate body plan. *Developmental Biology, 344*(1), 7-15.

Mann, R. S. (1995). The specificity of homeotic gene function. *Bioessays, 17*(10), 855-863.

Mansfield, J. H., B. D. Harfe, R. Nissen, J. Obenauer, J. Srineel, A. Chaudhuri, R. Farzan-Kashani, M. Zuker, A. E. Pasquinelli, G. Ruvkun, P. A. Sharp, C. J. Tabin and M. T. McManus (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nature Genetics, 36*(10), 1079-1083.

Marbach-Bar, N., A. Ben-Noon, S. Ashkenazi, A. T.-B. Harush, T. Avnit-Sagi, M. D. Walker and R. Dikstein (2013). Disparity between microRNA levels and promoter strength is associated with initiation rate and Pol II pausing. *Nature Communications, 4*(1), Article 2118.

Marcucci, G., M. D. Radmacher, K. Mrózek and C. D. Bloomfield (2009). MicroRNA expression in acute myeloid leukemia. *Current Hematologic Malignancy Reports, 4*(2), 83-88.

Marinescu, V. D., I. S. Kohane and A. Riva (2005). The MAPPER database: a multi-genome catalog of putative transcription factor binding sites. *Nucleic acids research, 33*(Database issue), D91-D97.

Marley, S. B., J. L. Lewis and M. Y. Gordon (2003). Progenitor cells divide symmetrically to generate new colony-forming cells and clonal heterogeneity. *British Journal of Haematology, 121*(4), 643-648.

Maru, D. M., R. R. Singh, C. Hannah, C. T. Albarracin, Y. X. Li, R. Abraham, A. M. Romans, H. Yao, M. G. Luthra, S. Anandasabapathy, S. G. Swisher, W. L. Hofstetter, A. Rashid and R. Luthra (2009). MicroRNA-196a is a potential marker of progression during Barrett's metaplasia-dysplasia-invasive adenocarcinoma sequence in esophagus. *The American journal of pathology, 174*(5), 1940-1948.

Mayya, V. K. and T. F. Duchaine (2015). On the availability of microRNA-induced silencing complexes, saturation of microRNA-binding sites and stoichiometry. *Nucleic acids research, 43*(15), Article 7556.

McGhee, J. D., W. I. Wood, M. Dolan, J. D. Engel and G. Felsenfeld (1981). A 200 base pair region at the 5′ end of the chicken adult β-globin gene is accessible to nuclease digestion. *Cell, 27*(1, Part 2), 45-55.

Menendez-Gonzalez, J. B., M. Vukovic, A. Abdelfattah, L. Saleh, A. Almotiri, L.-a. Thomas, A. Agirre-Lizaso, A. Azevedo, A. C. Menezes, G. Tornillo, S. Edkins, K. Kong, P. Giles, F. Anjos-Afonso, A. Tonks, A. S. Boyd, K. R. Kranc and N. P. Rodrigues (2019). Gata2 as a Crucial Regulator of Stem Cells in Adult Hematopoiesis and Acute Myeloid Leukemia. *Stem Cell Reports, 13*(2), 291-306.

Metzeler, K. H., K. Maharry, J. Kohlschmidt, S. Volinia, K. Mrózek, H. Becker, D. Nicolet, S. P. Whitman, J. H. Mendler, S. Schwind, A. K. Eisfeld, Y. Z. Wu, B. L. Powell, T. H.

Carter, M. Wetzler, J. E. Kolitz, M. R. Baer, A. J. Carroll, R. M. Stone, M. A. Caligiuri, G. Marcucci and C. D. Bloomfield (2013). A stem cell-like gene expression signature associates with inferior outcomes and a distinct microRNA expression profile in adults with primary cytogenetically normal acute myeloid leukemia. *Leukemia, 27*(10), 2023-2031.

Meunier, J., F. Lemoine, M. Soumillon, A. Liechti, M. Weier, K. Guschanski, H. Hu, P. Khaitovich and H. Kaessmann (2013). Birth and expression evolution of mammalian microRNA genes. *Genome Research, 23*(1), 34-45.

Meyer, S. E., D. E. Muench, A. M. Rogers, T. J. Newkold, E. Orr, E. O'Brien, J. P. Perentesis, J. G. Doench, A. Lal, P. J. Morris, C. J. Thomas, J. Lieberman, E. McGlinn, B. J. Aronow, N. Salomonis and H. L. Grimes (2018). miR-196b target screen reveals mechanisms maintaining leukemia stemness with therapeutic potential. *The Journal of experimental medicine, 215*(8), 2115-2136.

Micevic, G., V. Muthusamy, W. Damsky, N. Theodosakis, X. Liu, K. Meeth, E. Wingrove, M. Santhanakrishnan and M. Bosenberg (2016). DNMT3b Modulates Melanoma Growth by Controlling Levels of mTORC2 Component RICTOR. *Cell reports, 14*(9), 2180-2192.

Migliaccio, A. R., R. A. Rana, M. Sanchez, R. Lorenzini, L. Centurione, L. Bianchi, A. M. Vannucchi, G. Migliaccio and S. H. Orkin (2003). GATA-1 as a regulator of mast cell differentiation revealed by the phenotype of the GATA-1low mouse mutant. *The Journal of experimental medicine, 197*(3), 281-296.

Mildner, A., E. Chapnik, O. Manor, S. Yona, K.-W. Kim, T. Aychek, D. Varol, G. Beck, Z. B. Itzhaki, E. Feldmesser, I. Amit, E. Hornstein and S. Jung (2013). Mononuclear phagocyte miRNome analysis identifies miR-142 as critical regulator of murine dendritic cell homeostasis. *Blood, 121*(6), 1016-1027.

Min, B., M. A. Brown and G. Legros (2012). Understanding the roles of basophils: breaking dawn. *Immunology, 135*(3), 192-197.

Min, I. M., G. Pietramaggiori, F. S. Kim, E. Passegué, K. E. Stevenson and A. J. Wagers (2008). The Transcription Factor EGR1 Controls Both the Proliferation and Localization of Hematopoietic Stem Cells. *Cell Stem Cell, 2*(4), 380-391.

Minegishi, N., N. Suzuki, T. Yokomizo, X. Pan, T. Fujimoto, S. Takahashi, T. Hara, A. Miyajima, S.-i. Nishikawa and M. Yamamoto (2003). Expression and domain-specific function of GATA-2 during differentiation of the hematopoietic precursor cells in midgestation mouse embryos. *Blood, 102*(3), 896-905.

Mo, J.-S., K. Jahengir Alam, I.-H. Kang, W. Cheol Park, G.-S. Seo, S.-C. Choi, H.-S. Kim, H.-B. Moon, K.-J. Yun and S.-C. Chae (2014). MicroRNA 196B regulates FAS-mediated apoptosis in colorectal cancer cells. *Oncotarget, 6*(5), 2843-2855.

Mohr, S., C. Doebele, F. Comoglio, T. Berg, J. Beck, H. Bohnenberger, G. Alexe, J. Corso, P. Ströbel, A. Wachter, T. Beissbarth, F. Schnütgen, A. Cremer, N. Haetscher, S. Göllner, A. Rouhi, L. Palmqvist, M. A. Rieger, T. Schroeder, H. Bönig, C. Müller-Tidow, F. Kuchenbauer, E. Schütz, A. R. Green, H. Urlaub, K. Stegmaier, R. K. Humphries, H. Serve and T. Oellerich (2017). Hoxa9 and Meis1 Cooperatively Induce Addiction to Syk Signaling by Suppressing miR-146a in Acute Myeloid Leukemia. *Cancer cell, 31*(4), 549-562.

Montagner, S., L. Dehó and S. Monticelli (2014). MicroRNAs in hematopoietic development. *BMC immunology, 15*, 14-14.

Mootha, V. K., C. M. Lindgren, K.-F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstråle, E. Laurila, N. Houstis, M. J. Daly, N. Patterson, J. P. Mesirov, T. R. Golub, P. Tamayo, B. Spiegelman, E. S. Lander, J. N. Hirschhorn, D. Altshuler and L. C. Groop (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics, 34*(3), 267-273.

Moriya, H. (2015). Quantitative nature of overexpression experiments. *Molecular biology of the cell, 26*(22), 3932-3939.

Morrison, D. K. (2012). MAP kinase pathways. *Cold Spring Harbor perspectives in biology, 4*(11), Article a011254.

Morrison, S. J., N. Uchida and I. L. Weissman (1995). The biology of hematopoietic stem cells. *Annual Review of Cell and Developmental Biology, 11*, 35-71.

Morrison, S. J., A. M. Wandycz, K. Akashi, A. Globerson and I. L. Weissman (1996). The aging of hematopoietic stem cells. *Nature Medicine, 2*(9), 1011-1016.

Morrison, S. J. and I. L. Weissman (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity, 1*(8), 661-673.

Mueller, D. W. and A.-K. Bosserhoff (2011). MicroRNA miR-196a controls melanoma-associated genes by regulating HOX-C8 expression. *International Journal of Cancer, 129*(5), 1064-1074.

Mukherji, S., M. S. Ebert, G. X. Y. Zheng, J. S. Tsang, P. A. Sharp and A. van Oudenaarden (2011). MicroRNAs can generate thresholds in target gene expression. *Nature genetics, 43*(9), 854-859.

Nagy, Á., Á. Ősz, J. Budczies, S. Krizsán, G. Szombath, J. Demeter, C. Bödör and B. Győrffy (2019). Elevated HOX gene expression in acute myeloid leukemia is associated with NPM1 mutations and poor survival. *Journal of advanced research, 20*, 105-116.

Nakamura, T., D. A. Largaespada, J. D. Shaughnessy, N. A. Jenkins and N. G. Copeland (1996). Cooperative activation of Hoxa and Pbx1-related genes in murine myeloid leukaemias. *Nature Genetics, 12*(2), 149-153.

Nakanishi, K. (2016). Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *WIREs RNA, 7*(5), 637-660.

Nam, J.-W., O. S. Rissland, D. Koppstein, C. Abreu-Goodger, C. H. Jan, V. Agarwal, M. A. Yildirim, A. Rodriguez and D. P. Bartel (2014). Global analyses of the effect of different cellular contexts on microRNA targeting. *Molecular cell, 53*(6), 1031-1043.

Naoe, T. and H. Kiyoi (2013). Gene mutations of acute myeloid leukemia in the genome era. *International Journal of Hematology, 97*(2), 165-174.

Narang, V., W.-K. Sung and A. Mittal (2005). Computational modeling of oligonucleotide positional densities for human promoter prediction. *Artificial Intelligence in Medicine, 35*(1), 107-119.

Nayak, B. K. (2010). Understanding the relevance of sample size calculation. *Indian journal of ophthalmology, 58*(6), 469-470.

Ng, S. W. K., A. Mitchell, J. A. Kennedy, W. C. Chen, J. McLeod, N. Ibrahimova, A. Arruda, A. Popescu, V. Gupta, A. D. Schimmer, A. C. Schuh, K. W. Yee, L. Bullinger, T. Herold, D. Görlich, T. Büchner, W. Hiddemann, W. E. Berdel, B. Wörmann, M. Cheok, C. Preudhomme, H. Dombret, K. Metzeler, C. Buske, B. Löwenberg, P. J. M. Valk, P. W. Zandstra, M. D. Minden, J. E. Dick and J. C. Y. Wang (2016). A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature, 540*(7633), 433-437.

Nguyen, T. K., S. J. Morse and A. G. Fleischman (2016). Transduction-Transplantation Mouse Model of Myeloproliferative Neoplasm. *Journal of visualized experiments : JoVE,*(118), Article 54624.

Nie, M. and H. Htun (2006). Different modes and potencies of translational repression by sequence-specific RNA–protein interaction at the 5′-UTR. *Nucleic Acids Research, 34*(19), 5528-5540.

Noordhuis, P., U. Holwerda, C. L. Van der Wilt, C. J. Van Groeningen, K. Smid, S. Meijer, H. M. Pinedo and G. J. Peters (2004). 5-Fluorouracil incorporation into RNA and DNA in relation to thymidylate synthase inhibition of human colorectal cancers. *Annals of Oncology, 15*(7), 1025-1032.

Novak, R. L., D. P. Harper, D. Caudell, C. Slape, S. H. Beachy and P. D. Aplan (2012). Gene expression profiling and candidate gene resequencing identifies pathways and mutations important for malignant transformation caused by leukemogenic fusion genes. *Experimental hematology, 40*(12), 1016-1027.

Ntini, E., A. I. Järvelin, J. Bornholdt, Y. Chen, M. Boyd, M. Jørgensen, R. Andersson, I. Hoof, A. Schein, P. R. Andersen, P. K. Andersen, P. Preker, E. Valen, X. Zhao, V. Pelechano, L. M. Steinmetz, A. Sandelin and T. H. Jensen (2013). Polyadenylation site–induced decay of upstream transcripts enforces promoter directionality. *Nature Structural & Molecular Biology, 20*(8), 923-928.

O'Brien, J., H. Hayder, Y. Zayed and C. Peng (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Frontiers in Endocrinology, 9*, Article 402.

O'Carroll, D., I. Mecklenbrauker, P. P. Das, A. Santana, U. Koenig, A. J. Enright, E. A. Miska and A. Tarakhovsky (2007). A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes & development, 21*(16), 1999-2004.

O'Connell, R. M., K. D. Taganov, M. P. Boldin, G. Cheng and D. Baltimore (2007). MicroRNA-155 is induced during the macrophage inflammatory response. *Proceedings of the National Academy of Sciences of the United States of America, 104*(5), 1604-1609.

Odell, J. T., J. L. Hoopes and W. Vermerris (1994). Seed-specific gene activation mediated by the Cre/lox site-specific recombination system. *Plant physiology, 106*(2), 447-458.

Oh, I.-H. and R. K. Humphries (2012). Concise Review: Multidimensional Regulation of the Hematopoietic Stem Cell State. *STEM CELLS, 30*(1), 82-88.

Olejniczak, S. H., G. La Rocca, J. J. Gruber and C. B. Thompson (2013). Long-lived microRNA-Argonaute complexes in quiescent cells can be activated to regulate mitogenic responses. *Proceedings of the National Academy of Sciences of the United States of America, 110*(1), 157-162.

Olsen, P. H. and V. Ambros (1999). The lin-4 Regulatory RNA Controls Developmental Timing in Caenorhabditis elegans by Blocking LIN-14 Protein Synthesis after the Initiation of Translation. *Developmental Biology, 216*(2), 671-680.

Omichinski, J. G., G. M. Clore, O. Schaad, G. Felsenfeld, C. Trainor, E. Appella, S. J. Stahl and A. M. Gronenborn (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science, 261*(5120), 438-446.

Ooi, A. G. L., D. Sahoo, M. Adorno, Y. Wang, I. L. Weissman and C. Y. Park (2010). MicroRNA-125b expands hematopoietic stem cells and enriches for the lymphoid-balanced and lymphoid-biased subsets. *Proceedings of the National Academy of Sciences, 107*(50), Article 21505.

Orkin, S. H. (1995). Transcription factors and hematopoietic development. *Journal of Biological Chemistry, 270*(10), 4955-4958.

Ovcharenko, I., M. A. Nobrega, G. G. Loots and L. Stubbs (2004). ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic acids research, 32*(Web Server issue), W280-W286.

Ozsolak, F., L. L. Poling, Z. Wang, H. Liu, X. S. Liu, R. G. Roeder, X. Zhang, J. S. Song and D. E. Fisher (2008). Chromatin structure analyses identify miRNA promoters. *Genes & development, 22*(22), 3172-3183.

Pabst, T. and B. U. Mueller (2009). Complexity of CEBPA Dysregulation in Human Acute Myeloid Leukemia. *Clinical Cancer Research, 15*(17), Article 5303.

Padua, D., X. H. F. Zhang, Q. Wang, C. Nadal, W. L. Gerald, R. R. Gomis and J. Massagué (2008). TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell, 133*(1), 66-77.

Pan, Y., X. Shu, L. Sun, L. Yu, L. Sun, Z. Yang and Y. Ran (2017). miR196a5p modulates gastric cancer stem cell characteristics by targeting Smad4. *International Journal of Oncology, 50*, 1965-1976.

Pandolfi, A., L. Barreyro and U. Steidl (2013). Concise review: preleukemic stem cells: molecular biology and clinical implications of the precursors to leukemia stem cells. *Stem cells translational medicine, 2*(2), 143-150.

Panigrahi, A. and B. W. O'Malley (2021). Mechanisms of enhancer action: the known and the unknown. *Genome Biology, 22*(1), Article 108.

Paroo, Z., X. Ye, S. Chen and Q. Liu (2009). Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell, 139*(1), 112-122.

Pavani, G., M. Laurent, A. Fabiano, E. Cantelli, A. Sakkal, G. Corre, P. J. Lenting, J.-P. Concordet, M. Toueille, A. Miccio and M. Amendola (2020). Ex vivo editing of human hematopoietic stem cells for erythroid expression of therapeutic proteins. *Nature Communications, 11*(1), Article 3778.

Pelt-verkuil, E., A. van Belkum and J. Hays (2008). *Principles and Technical Aspects of PCR Amplification.* Dordrecht, Springer.

Peng, T., T. R. Golub and D. M. Sabatini (2002). The immunosuppressant rapamycin mimics a starvation-like signal distinct from amino acid and glucose deprivation. *Molecular and cellular biology, 22*(15), 5575-5584.

Peng, Y. and C. M. Croce (2016). The role of MicroRNAs in human cancer. *Signal Transduction and Targeted Therapy, 1*(1), Article 15004.

Petersen, C. P., M.-E. Bordeleau, J. Pelletier and P. A. Sharp (2006). Short RNAs Repress Translation after Initiation in Mammalian Cells. *Molecular Cell, 21*(4), 533-542.

Petriv, O. I., F. Kuchenbauer, A. D. Delaney, V. Lecault, A. White, D. Kent, L. Marmolejo, M. Heuser, T. Berg, M. Copley, J. Ruschmann, S. Sekulovic, C. Benz, E. Kuroda, V. Ho, F. Antignano, T. Halim, V. Giambra, G. Krystal, C. J. F. Takei, A. P. Weng, J. Piret, C. Eaves, M. A. Marra, R. K. Humphries and C. L. Hansen (2010). Comprehensive microRNA expression profiling of the hematopoietic hierarchy. *Proceedings of the National Academy of Sciences, 107*(35), Article 15443.

Pevny, L., M. C. Simon, E. Robertson, W. H. Klein, S.-F. Tsai, V. D'Agati, S. H. Orkin and F. Costantini (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature, 349*(6306), 257-260.

Pillai, R. S., S. N. Bhattacharyya, C. G. Artus, T. Zoller, N. Cougot, E. Basyuk, E. Bertrand and W. Filipowicz (2005). Inhibition of Translational Initiation by Let-7 MicroRNA in Human Cells. *Science, 309*(5740), Article 1573.

Pineault, N., C. D. Helgason, H. J. Lawrence and R. K. Humphries (2002). Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Experimental Hematology, 30*(1), 49-57.

Piper, D. E., A. H. Batchelor, C. P. Chang, M. L. Cleary and C. Wolberger (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell, 96*(4), 587-597.

Pitchiaya, S., L. A. Heinicke, J. I. Park, E. L. Cameron and N. G. Walter (2017). Resolving Subcellular miRNA Trafficking and Turnover at Single-Molecule Resolution. *Cell reports, 19*(3), 630-642.

Ponger, L. c. and D. Mouchiroud (2002). CpGProD: identifying CpG islands associated with transcription start sites in large genomic mammalian sequences. *Bioinformatics, 18*(4), 631-633.

Popovic, R., L. E. Riesbeck, C. S. Velu, A. Chaubey, J. Zhang, N. J. Achille, F. E. Erfurth, K. Eaton, J. Lu, H. L. Grimes, J. Chen, J. D. Rowley and N. J. Zeleznik-Le (2009). Regulation of mir-196b by MLL and its overexpression by MLL fusions contributes to immortalization. *Blood, 113*(14), 3314-3322.

Prabhakar, M., W. B. Ershler and D. L. Longo (2009). Bone marrow, thymus and blood: changes across the lifespan. *Aging Health, 5*(3), 385-393.

Qian, X.-L., Y.-Q. Li, B. Yu, F. Gu, F.-F. Liu, W.-D. Li, X.-M. Zhang and L. Fu (2013). Syndecan binding protein (SDCBP) is overexpressed in estrogen receptor negative breast cancers, and is a potential promoter for tumor proliferation. *PloS one, 8*(3), Article e60046.

Quilez, J., A. Guilmatre, P. Garg, G. Highnam, M. Gymrek, Y. Erlich, R. S. Joshi, D. Mittelman and A. J. Sharp (2016). Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. *Nucleic acids research, 44*(8), 3750-3762.

Raghav, P. K. and Gangenahalli (2018). Hematopoietic Stem Cell Molecular Targets and Factors Essential for Hematopoiesis. *Journal of Stem Cell Researh & Therapy, 8*(11), 1-16.

Rajewsky, N. (2006). microRNA target predictions in animals. *Nature Genetics, 38*(6), S8-S13.

Randall, T. D. and I. L. Weissman (1997). Phenotypic and Functional Changes Induced at the Clonal Level in Hematopoietic Stem Cells After 5-Fluorouracil Treatment. *Blood, 89*(10), 3596-3606.

Rawat, V. P. S., M. Cusan, A. Deshpande, W. Hiddemann, L. Quintanilla-Martinez, R. K. Humphries, S. K. Bohlander, M. Feuring-Buske and C. Buske (2004). Ectopic expression of the homeobox gene Cdx2 is the transforming event in a mouse model of t(12;13)(p13;q12) acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America, 101*(3), 817-822.

Rawat, V. P. S., M. Götze, A. Rasalkar, N. M. Vegi, S. Ihme, S. Thoene, A. Pastore, D. Bararia, H. Döhner, K. Döhner, M. Feuring-Buske, L. Quintanilla-Fend and C. Buske (2020). The microRNA miR-196b acts as a tumor suppressor in Cdx2-driven acute myeloid leukemia. *Haematologica, 105*(6), e285-e289.

Rawat, V. P. S., R. K. Humphries and C. Buske (2012). Beyond Hox: the role of ParaHox genes in normal and malignant hematopoiesis. *Blood, 120*(3), 519-527.

Rawat, V. P. S., S. Thoene, V. M. Naidu, N. Arseni, B. Heilmeier, K. Metzeler, K. Petropoulos, A. Deshpande, L. Quintanilla-Martinez, S. K. Bohlander, K. Spiekermann, W. Hiddemann, M. Feuring-Buske and C. Buske (2008). Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood, 111*(1), 309-319.

Rebucci, M., A. Sermeus, E. Leonard, E. Delaive, M. Dieu, M. Fransolet, T. Arnould and C. Michiels (2015). miRNA-196b inhibits cell proliferation and induces apoptosis in HepG2 cells by targeting IGF2BP1. *Molecular Cancer, 14*(1), Article 79.

Regelin, M., J. Blume, J. Pommerencke, R. Vakilzadeh, K. Witzlau, M. Łyszkiewicz, N. Ziętara, N. Saran, A. Schambach and A. Krueger (2015). Responsiveness of Developing T Cells to IL-7 Signals Is Sustained by miR-17~92. *The Journal of Immunology, 195*(10), Article 4832.

Reich, M., T. Liefeld, J. Gould, J. Lerner, P. Tamayo and J. P. Mesirov (2006). GenePattern 2.0. *Nature Genetics, 38*(5), 500-501.

Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz and G. Ruvkun (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature, 403*(6772), 901-906.

Ren, W., S. Wu, Y. Wu, T. Liu, X. Zhao and Y. Li (2019). MicroRNA-196a/-196b regulate the progression of hepatocellular carcinoma through modulating the JAK/STAT pathway via targeting SOCS2. *Cell Death & Disease, 10*(5), Article 333.

Reya, T. (2003). Regulation of hematopoietic stem cell self-renewal. *Recent Progress in Hormone Research, 58*, 283-295.

Reygaert, W. C. (2018). An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS microbiology, 4*(3), 482-501.

Rice, K. L. and J. D. Licht (2007). HOX deregulation in acute myeloid leukemia. *The Journal of clinical investigation, 117*(4), 865-868.

Rissland, O. S. (2017). The organization and regulation of mRNA-protein complexes. *Wiley interdisciplinary reviews. RNA, 8*(1), Article e1369.

Roden, C. and J. Lu (2016). MicroRNAs in Control of Stem Cells in Normal and Malignant Hematopoiesis. *Current Stem Cell Reports, 2*(3), 183-196.

Romania, P., V. Lulli, E. Pelosi, M. Biffoni, C. Peschle and G. Marziali (2008). MicroRNA 155 modulates megakaryopoiesis at progenitor and precursor level by targeting Ets-1 and Meis1 transcription factors. *British Journal of Haematology, 143*(4), 570-580.

Ronshaugen, M., F. Biemar, J. Piel, M. Levine and E. C. Lai (2005). The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes & development, 19*(24), 2947-2952.

Rosa, A., M. Ballarino, A. Sorrentino, O. Sthandier, F. G. De Angelis, M. Marchioni, B. Masella, A. Guarini, A. Fatica, C. Peschle and I. Bozzoni (2007). The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation. *Proceedings of the National Academy of Sciences of the United States of America, 104*(50), 19849-19854.

Rosenbauer, F., S. Koschmieder, U. Steidl and D. G. Tenen (2005). Effect of transcription-factor concentrations on leukemic stem cells. *Blood, 106*(5), 1519-1524.

Rothenberg, E. V., H. Hosokawa and J. Ungerbäck (2019). Mechanisms of Action of Hematopoietic Transcription Factor PU.1 in Initiation of T-Cell Development. *Frontiers in Immunology, 10*, Article 228.

Roy, A. L. and D. S. Singer (2015). Core promoters in transcription: old problem, new insights. *Trends in biochemical sciences, 40*(3), 165-171.

Saito, Y., H. Kitamura, A. Hijikata, M. Tomizawa-Murasawa, S. Tanaka, S. Takagi, N. Uchida, N. Suzuki, A. Sone, Y. Najima, H. Ozawa, A. Wake, S. Taniguchi, L. D. Shultz, O. Ohara and F. Ishikawa (2010). Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. *Science translational medicine, 2*(17), Article 17ra19.

Sanchez, P. V., S. T. Glantz, S. Scotland, M. T. Kasner and M. Carroll (2014). Induced differentiation of acute myeloid leukemia cells by activation of retinoid X and liver X receptors. *Leukemia, 28*(4), 749-760.

Sauvageau, G., P. M. Lansdorp, C. J. Eaves, D. E. Hogge, W. H. Dragowska, D. S. Reid, C. Largman, H. J. Lawrence and R. K. Humphries (1994). Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proceedings of the National Academy of Sciences of the United States of America, 91*(25), 12223-12227.

Sayed, D. and M. Abdellatif (2011). MicroRNAs in Development and Disease. *Physiological Reviews, 91*(3), 827-887.

Schimanski, C. C., K. Frerichs, F. Rahman, M. Berger, H. Lang, P. R. Galle, M. Moehler and I. Gockel (2009). High miR-196a levels promote the oncogenic phenotype of colorectal cancer cells. *World journal of gastroenterology, 15*(17), 2089-2096.

Schmittgen, T. D. and K. J. Livak (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols, 3*(6), 1101-1108.

Schoch, C., S. Schnittger, M. Klaus, W. Kern, W. Hiddemann and T. Haferlach (2003). AML with 11q23/MLL abnormalities as defined by the WHO classification: incidence, partner chromosomes, FAB subtype, age distribution, and prognostic impact in an unselected series of 1897 cytogenetically analyzed AML cases. *Blood, 102*(7), 2395-2402.

Schoenfelder, S. and P. Fraser (2019). Long-range enhancer–promoter contacts in gene expression control. *Nature Reviews Genetics, 20*(8), 437-455.

Scholl, C., D. Bansal, K. Döhner, K. Eiwen, B. J. P. Huntly, B. H. Lee, F. G. Rücker, R. F. Schlenk, L. Bullinger, H. Döhner, D. G. Gilliland and S. Fröhling (2007). The homeobox gene CDX2 is aberrantly expressed in most cases of acute myeloid leukemia and promotes leukemogenesis. *The Journal of clinical investigation, 117*(4), 1037-1048.

Scholl, C., S. Fröhling, I. F. Dunn, A. C. Schinzel, D. A. Barbie, S. Y. Kim, S. J. Silver, P. Tamayo, R. C. Wadlow, S. Ramaswamy, K. Döhner, L. Bullinger, P. Sandy, J. S. Boehm, D. E. Root, T. Jacks, W. C. Hahn and D. G. Gilliland (2009). Synthetic Lethal Interaction between Oncogenic KRAS Dependency and STK33 Suppression in Human Cancer Cells. *Cell, 137*(5), 821-834.

Schotte, D., J. C. K. Chau, G. Sylvester, G. Liu, C. Chen, V. H. J. van der Velden, M. J. C. Broekhuis, T. C. J. M. Peters, R. Pieters and M. L. d. Boer (2008). Identification of new microRNA genes and aberrant microRNA profiles in childhood acute lymphoblastic leukemia. *Leukemia, 23*, Article 313.

Schotte, D., E. A. M. Lange-Turenhout, D. J. P. M. Stumpel, R. W. Stam, J. G. C. A. M. Buijs-Gladdines, J. P. P. Meijerink, R. Pieters and M. L. Den Boer (2010). Expression of miR-196b is not exclusively MLL-driven but is especially linked to activation of HOXA genes in pediatric acute lymphoblastic leukemia. *Haematologica, 95*(10), 1675-1682.

Schwab, R., C. Speth, S. Laubinger and O. Voinnet (2013). Enhanced microRNA accumulation through stemloop-adjacent introns. *EMBO reports, 14*(7), 615-621.

Seita, J. and I. L. Weissman (2010). Hematopoietic stem cell: self-renewal versus differentiation. *Wiley interdisciplinary reviews. Systems biology and medicine, 2*(6), 640-653.

Senyuk, V., Y. Zhang, Y. Liu, M. Ming, K. Premanand, L. Zhou, P. Chen, J. Chen, J. D. Rowley, G. Nucifora and Z. Qian (2013). Critical role of miR-9 in myelopoiesis and EVI1-induced leukemogenesis. *Proceedings of the National Academy of Sciences, 110*(14), Article 5594.

Shang, Y., L.-Q. Wang, Q.-Y. Guo and T.-L. Shi (2015). MicroRNA-196a overexpression promotes cell proliferation and inhibits cell apoptosis through PTEN/Akt/FOXO1 pathway. *International journal of clinical and experimental pathology, 8*(3), 2461-2472.

Shao, L., Z. Chen, D. Peng, M. Soutto, S. Zhu, A. Bates, S. Zhang and W. El-Rifai (2018). Methylation of the HOXA10 Promoter Directs miR-196b-5p–Dependent Cell Proliferation and Invasion of Gastric Cancer Cells. *Molecular Cancer Research, 16*(4), 696-706.

Shen, S., J. Pan, X. Lu and P. Chi (2016). Role of miR-196 and its target gene HoxB8 in the development and proliferation of human colorectal cancer and the impact of neoadjuvant chemotherapy with FOLFOX4 on their expression. *Oncology letters, 12*(5), 4041-4047.

Shen, Y. N., I.-S. Bae, G. H. Park, H. S. Choi, K.-H. Lee and S. H. Kim (2018). MicroRNA-196b enhances the radiosensitivity of SNU-638 gastric cancer cells by targeting RAD23B. *Biomedicine & Pharmacotherapy, 105*, 362-369.

Shi, J., E. Wang, J. Zuber, A. Rappaport, M. Taylor, C. Johns, S. W. Lowe and C. R. Vakoc (2013). The Polycomb complex PRC2 supports aberrant self-renewal in a mouse model of MLL-AF9;Nras(G12D) acute myeloid leukemia. *Oncogene, 32*(7), 930-938.

Shimabe, M., S. Goyama, N. Watanabe-Okochi, A. Yoshimi, M. Ichikawa, Y. Imai and M. Kurokawa (2009). Pbx1 is a downstream target of Evi-1 in hematopoietic stem/progenitors and leukemic cells. *Oncogene, 28*(49), 4364-4374.

Shivdasani, R. A., Y. Fujiwara, M. A. McDevitt and S. H. Orkin (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *The EMBO journal, 16*(13), 3965-3973.

Simpson, L. J., S. Patel, N. R. Bhakta, D. F. Choy, H. D. Brightbill, X. Ren, Y. Wang, H. H. Pua, D. Baumjohann, M. M. Montoya, M. Panduro, K. A. Remedios, X. Huang, J. V. Fahy, J. R. Arron, P. G. Woodruff and K. M. Ansel (2014). A microRNA upregulated in asthma airway T cells promotes TH2 cytokine production. *Nature Immunology, 15*(12), 1162-1170.

Singh-Jasuja, H., A. Thiolat, M. Ribon, M.-C. Boissier, N. Bessis, H.-G. Rammensee and P. Decker (2013). The mouse dendritic cell marker CD11c is down-regulated upon cell activation through Toll-like receptor triggering. *Immunobiology, 218*(1), 28-39.

Skvarova Kramarzova, K., K. Fiser, E. Mejstrikova, K. Rejlova, M. Zaliova, M. Fornerod, H. A. Drabkin, M. M. van den Heuvel-Eibrink, J. Stary, J. Trka and J. Starkova (2014). Homeobox gene expression in acute myeloid leukemia is linked to typical underlying molecular aberrations. *Journal of hematology & oncology, 7*, Article 94.

Small, D. (2006). FLT3 Mutations: Biology and Treatment. *ASH Education Program Book, 2006*(1), 178-184.

So, C. W., H. Karsunky, P. Wong, I. L. Weissman and M. L. Cleary (2004). Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. *Blood, 103*(8), 3192-3199.

Solovyev, V., P. Kosarev, I. Seledsov and D. Vorobyev (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology, 7*(1), S10.

Solovyev, V. V. and I. A. Shahmuradov (2003). PromH: promoters identification using orthologous genomic sequences. *Nucleic Acids Research, 31*(13), 3540-3545.

Solovyev, V. V., I. A. Shahmuradov and A. A. Salamov (2010). Identification of promoter regions and regulatory sites. *Methods Mol Biol, 674*, 57-83.

Song, S. J., K. Ito, U. Ala, L. Kats, K. Webster, S. M. Sun, M. Jongen-Lavrencic, K. Manova-Todorova, J. Teruya-Feldstein, D. E. Avigan, R. Delwel and P. P. Pandolfi (2013). The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote

hematopoietic stem cell self-renewal and transformation. *Cell stem cell, 13*(1), 87-101.

Spencer, D. H., M. A. Young, T. L. Lamprecht, N. M. Helton, R. Fulton, M. O'Laughlin, C. Fronick, V. Magrini, R. T. Demeter, C. A. Miller, J. M. Klco, R. K. Wilson and T. J. Ley (2015). Epigenomic analysis of the HOX gene loci reveals mechanisms that may control canonical expression patterns in AML and normal hematopoietic cells. *Leukemia, 29*(6), 1279-1289.

Srikant, R., Q. Vu and R. Agrawal (1997). Mining generalized association rules. *Future Generation Computer Systems, 13*(2-3), 161-180.

Stiegelbauer, V., P. Vychytilova-Faltejskova, M. Karbiener, A.-M. Pehserl, A. Reicher, M. Resel, E. Heitzer, C. Ivan, M. Bullock, H. Ling, A. Deutsch, A. Wulf-Goldenberg, J. B. Adiprasito, H. Stoeger, J. Haybaeck, M. Svoboda, M. Stotz, G. Hoefler, O. Slaby, G. A. Calin, A. Gerger and M. Pichler (2017). miR-196b-5p Regulates Colorectal Cancer Cell Migration and Metastases through Interaction with HOXB7 and GALNT5. *Clinical Cancer Research, 23*(17), Article 5255.

Stirewalt, D. L. and J. P. Radich (2003). The role of FLT3 in haematopoietic malignancies. *Nature Reviews Cancer, 3*, Article 650.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander and J. P. Mesirov (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, 102*(43), Article 15545.

Sun, M., X.-h. Liu, J.-h. Li, J.-s. Yang, E.-b. Zhang, D.-d. Yin, Z.-l. Liu, J. Zhou, Y. Ding, S.-q. Li, Z.-x. Wang, X.-f. Cao and W. De (2012). MiR-196a Is Upregulated in Gastric Cancer and Promotes Cell Proliferation by Downregulating p27<sup>kip1</sup&gt. *Molecular Cancer Therapeutics, 11*(4), 842.

Sun, Y., B. Zhou, F. Mao, J. Xu, H. Miao, Z. Zou, L. T. Phuc Khoa, Y. Jang, S. Cai, M. Witkin, R. Koche, K. Ge, G. R. Dressler, R. L. Levine, S. A. Armstrong, Y. Dou and J. L. Hess (2018). HOXA9 Reprograms the Enhancer Landscape to Promote Leukemogenesis. *Cancer cell, 34*(4), 643-658.

Suzuki, M., T. Moriguchi, K. Ohneda and M. Yamamoto (2009). Differential contribution of the Gata1 gene hematopoietic enhancer to erythroid differentiation. *Molecular and cellular biology, 29*(5), 1163-1175.

Suzuki, N., O. Ohneda, N. Minegishi, M. Nishikawa, T. Ohta, S. Takahashi, J. D. Engel and M. Yamamoto (2006). Combinatorial Gata2 and Sca1 expression defines hematopoietic stem cells in the bone marrow niche. *Proceedings of the National Academy of Sciences of the United States of America, 103*(7), 2202-2207.

Suzuki, N., N. Suwabe, O. Ohneda, N. Obara, S. Imagawa, X. Pan, H. Motohashi and M. Yamamoto (2003). Identification and characterization of 2 types of erythroid progenitors that express GATA-1 at distinct levels. *Blood, 102*(10), 3575-3583.

Swerdlow, S. H., E. Campo, N. L. Harris, E. S. Jaffe, S. A. Pileri, H. Stein, J. Thiele and J. W. Vardiman, Eds. (2008). *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Geneva, Switzerland, WHO Press.

Szafranska, A. E., M. Doleshal, H. S. Edmunds, S. Gordon, J. Luttges, J. B. Munding, R. J. Barth, E. J. Gutmann, A. A. Suriawinata, J. Marc Pipas, A. Tannapfel, M. Korc, S.

A. Hahn, E. Labourier and G. J. Tsongalis (2008). Analysis of MicroRNAs in Pancreatic Fine-Needle Aspirates Can Classify Benign and Malignant Tissues. *Clinical Chemistry, 54*(10), Article 1716.

Taggart, J. C. and G.-W. Li (2018). Production of Protein-Complex Components Is Stoichiometric and Lacks General Feedback Regulation in Eukaryotes. *Cell Systems, 7*(6), 580-589.

Takahashi, S. (2011). Current findings for recurring mutations in acute myeloid leukemia. *Journal of hematology & oncology, 4*, Article 36.

TaKaRa. (2021). *RetroNectin reagent overview*. Retrieved 06.02.2021, from https://www.takarabio.com/learning-centers/gene-function/t-cell-transduction-and-culture/technology-overview.

Tanzer, A., C. T. Amemiya, C.-B. Kim and P. F. Stadler (2005). Evolution of microRNAs located within Hox gene clusters. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 304B*(1), 75-85.

Taussig, D. C., J. Vargaftig, F. Miraki-Moud, E. Griessinger, K. Sharrock, T. Luke, D. Lillington, H. Oakervee, J. Cavenagh, S. G. Agrawal, T. A. Lister, J. G. Gribben and D. Bonnet (2010). Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34(-) fraction. *Blood, 115*(10), 1976-1984.

Taylor, P. R., L. Martinez-Pomares, M. Stacey, H. H. Lin, G. D. Brown and S. Gordon (2004). MACROPHAGE RECEPTORS AND IMMUNE RECOGNITION. *Annual Review of Immunology, 23*(1), 901-944.

Team for Medical and Editorial Content in The American Cancer Society. (2018). *Risk Factors for Acute Myeloid Leukemia*. Retrieved 13.09.2020, 2020, from https://www.cancer.org/cancer/acute-myeloid-leukemia/causes-risks-prevention/risk-factors.html#references.

Team for Medical and Editorial Content in The American Cancer Society. (2020). *Cancer Statistics*. Retrieved 12.09.2020, 2020, from https://cancerstatisticscenter.cancer.org/?_ga=2.239921670.408351700.1599855160-1162872903.1591976898#!/.

Team R&D Systems and Novus Biologicals (2015). Cell Surface & Intracellular Markers Expressed at Different Stages of B Cell Development. B Cells, Biotechne.

Team Research and Development in Miltenyi Biotec (2019). Cell subsets frequencies, and marker expression Pan T-cell (mouse) handbook, Miltenyi Biotech.

Team Research and Development in Miltenyi Biotec (2019). Cell subsets, frequencies, and marker expression. Granulocytes, Miltenyi Biotech.

Tellez, C. S., D. E. Juri, K. Do, M. A. Picchi, T. Wang, G. Liu, A. Spira and S. A. Belinsky (2016). miR-196b Is Epigenetically Silenced during the Premalignant Stage of Lung Carcinogenesis. *Cancer Research, 76*(16), Article 4741.

Terwijn, M., W. Zeijlemaker, A. Kelder, A. P. Rutten, A. N. Snel, W. J. Scholten, T. Pabst, G. Verhoef, B. Löwenberg, S. Zweegman, G. J. Ossenkoppele and G. J. Schuurhuis (2014). Leukemic stem cell frequency: a strong biomarker for clinical outcome in acute myeloid leukemia. *PloS one, 9*(9), Article e107587.

The, E. P. C., E. Birney, J. A. Stamatoyannopoulos, A. Dutta, R. Guigó, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, J. A. Stamatoyannopoulos, R. E. Thurman, M. S. Kuehn, C. M. Taylor, S. Neph, C. M. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. A. Greenbaum, R. M. Andrews, P. Flicek, P. J. Boyle, H. Cao, N. P. Carter, G. K. Clelland, S. Davis, N. Day, P. Dhami, S. C. Dillon, M. O. Dorschner, H. Fiegler, P. G. Giresi, J. Goldy, M. Hawrylycz, A. Haydock, R. Humbert, K. D. James, B. E. Johnson, E. M. Johnson, T. T. Frum, E. R. Rosenzweig, N. Karnani, K. Lee, G. C. Lefebvre, P. A. Navas, F. Neri, S. C. J. Parker, P. J. Sabo, R. Sandstrom, A. Shafer, D. Vetrie, M. Weaver, S. Wilcox, M. Yu, F. S. Collins, J. Dekker, J. D. Lieb, T. D. Tullius, G. E. Crawford, S. Sunyaev, W. S. Noble, I. Dunham, A. Dutta, R. Guigó, F. Denoeud, A. Reymond, P. Kapranov, J. Rozowsky, D. Zheng, R. Castelo, A. Frankish, J. Harrow, S. Ghosh, A. Sandelin, I. L. Hofacker, R. Baertsch, D. Keefe, P. Flicek, S. Dike, J. Cheng, H. A. Hirsch, E. A. Sekinger, J. Lagarde, J. F. Abril, A. Shahab, C. Flamm, C. Fried, J. Hackermüller, J. Hertel, M. Lindemeyer, K. Missal, A. Tanzer, S. Washietl, J. Korbel, O. Emanuelsson, J. S. Pedersen, N. Holroyd, R. Taylor, D. Swarbreck, N. Matthews, M. C. Dickson, D. J. Thomas, M. T. Weirauch, J. Gilbert, J. Drenkow, I. Bell, X. Zhao, K. G. Srinivasan, W.-K. Sung, H. S. Ooi, K. P. Chiu, S. Foissac, T. Alioto, M. Brent, L. Pachter, M. L. Tress, A. Valencia, S. W. Choo, C. Y. Choo, C. Ucla, C. Manzano, C. Wyss, E. Cheung, T. G. Clark, J. B. Brown, M. Ganesh, S. Patel, H. Tammana, J. Chrast, C. N. Henrichsen, C. Kai, J. Kawai, U. Nagalakshmi, J. Wu, Z. Lian, J. Lian, P. Newburger, X. Zhang, P. Bickel, J. S. Mattick, P. Carninci, Y. Hayashizaki, S. Weissman, E. T. Dermitzakis, E. H. Margulies, T. Hubbard, R. M. Myers, J. Rogers, P. F. Stadler, T. M. Lowe, C.-L. Wei, Y. Ruan, M. Snyder, E. Birney, K. Struhl, M. Gerstein, S. E. Antonarakis, T. R. Gingeras, J. B. Brown, P. Flicek, Y. Fu, D. Keefe, E. Birney, F. Denoeud, M. Gerstein, E. D. Green, P. Kapranov, U. Karaöz, R. M. Myers, W. S. Noble, A. Reymond, J. Rozowsky, K. Struhl, A. Siepel, J. A. Stamatoyannopoulos, C. M. Taylor, J. Taylor, R. E. Thurman, T. D. Tullius, S. Washietl, D. Zheng, L. A. Liefer, K. A. Wetterstrand, P. J. Good, E. A. Feingold, M. S. Guyer, F. S. Collins, E. H. Margulies, G. M. Cooper, G. Asimenos, D. J. Thomas, C. N. Dewey, A. Siepel, E. Birney, D. Keefe, M. Hou, J. Taylor, S. Nikolaev, J. I. Montoya-Burgos, A. Löytynoja, S. Whelan, F. Pardi, T. Massingham, J. B. Brown, H. Huang, N. R. Zhang, P. Bickel, I. Holmes, J. C. Mullikin, A. Ureta-Vidal, B. Paten, M. Seringhaus, D. Church, K. Rosenbloom, W. J. Kent, E. A. Stone, M. Gerstein, S. E. Antonarakis, S. Batzoglou, N. Goldman, R. C. Hardison, D. Haussler, W. Miller, L. Pachter, E. D. Green, A. Sidow, Z. Weng, N. D. Trinklein, Y. Fu, Z. D. Zhang, U. Karaöz, L. Barrera, R. Stuart, D. Zheng, S. Ghosh, P. Flicek, D. C. King, J. Taylor, A. Ameur, S. Enroth, M. C. Bieda, C. M. Koch, H. A. Hirsch, C.-L. Wei, J. Cheng, J. Kim, A. A. Bhinge, P. G. Giresi, N. Jiang, J. Liu, F. Yao, W.-K. Sung, K. P. Chiu, V. B. Vega, C. W. H. Lee, P. Ng, A. Shahab, E. A. Sekinger, A. Yang, Z. Moqtaderi, Z. Zhu, X. Xu, S. Squazzo, M. J. Oberley, D. Inman, M. A. Singer, T. A. Richmond, K. J. Munn, A. Rada-Iglesias, O. Wallerman, J. Komorowski, G. K. Clelland, S. Wilcox, S. C. Dillon, R. M. Andrews, J. C. Fowler, P. Couttet, K. D. James, G. C. Lefebvre, A. W. Bruce, O. M. Dovey, P. D. Ellis, P. Dhami, C. F. Langford, N. P. Carter, D. Vetrie, P. Kapranov, D. A. Nix, I. Bell, S. Patel, J. Rozowsky, G. Euskirchen, S. Hartman, J. Lian, J. Wu, A. E. Urban, P. Kraus, S. Van Calcar, N. Heintzman, T. Hoon Kim, K. Wang, C. Qu, G. Hon, R. Luna, C. K. Glass, M. G. Rosenfeld, S. F. Aldred, S. J. Cooper, A. Halees, J. M. Lin, H. P. Shulha, X. Zhang, M. Xu, J. N. S. Haidar, Y. Yu, E. Birney*, S. Weissman, Y. Ruan, J. D. Lieb, V. R. Iyer, R. D. Green, T. R. Gingeras, C. Wadelius, I. Dunham, K. Struhl, R. C. Hardison, M. Gerstein, P. J. Farnham, R. M. Myers, B. Ren, M. Snyder, D. J. Thomas, K. Rosenbloom, R. A. Harte, A. S. Hinrichs, H. Trumbower, H. Clawson, J. Hillman-Jackson, A. S. Zweig, K. Smith, A. Thakkapallayil, G. Barber, R. M. Kuhn, D. Karolchik, D. Haussler, W. J. Kent, E. T. Dermitzakis, L. Armengol, C. P. Bird, T. G. Clark, G. M. Cooper, P. I. W. de Bakker, A. D. Kern, N. Lopez-Bigas, J. D. Martin, B. E. Stranger, D. J. Thomas, A. Woodroffe, S. Batzoglou, E. Davydov, A. Dimas, E. Eyras, I. B. Hallgrímsdóttir,

R. C. Hardison, J. Huppert, A. Sidow, J. Taylor, H. Trumbower, M. C. Zody, R. Guigó, J. C. Mullikin, G. R. Abecasis, X. Estivill, E. Birney, G. G. Bouffard, X. Guan, N. F. Hansen, J. R. Idol, V. V. B. Maduro, B. Maskeri, J. C. McDowell, M. Park, P. J. Thomas, A. C. Young, R. W. Blakesley, D. M. Muzny, E. Sodergren, D. A. Wheeler, K. C. Worley, H. Jiang, G. M. Weinstock, R. A. Gibbs, T. Graves, R. Fulton, E. R. Mardis, R. K. Wilson, M. Clamp, J. Cuff, S. Gnerre, D. B. Jaffe, J. L. Chang, K. Lindblad-Toh, E. S. Lander, M. Koriabine, M. Nefedov, K. Osoegawa, Y. Yoshinaga, B. Zhu and P. J. de Jong (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature, 447*, Article 799.

Thomson, D. W. and M. E. Dinger (2016). Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics, 17*(5), 272-283.

Thorsteinsdottir, U., J. Krosl, E. Kroon, A. Haman, T. Hoang and G. Sauvageau (1999). The oncoprotein E2A-Pbx1a collaborates with Hoxa9 to acutely transform primary bone marrow cells. *Molecular and cellular biology, 19*(9), 6355-6366.

Tomala, K. and R. Korona (2013). Evaluating the fitness cost of protein expression in Saccharomyces cerevisiae. *Genome biology and evolution, 5*(11), 2051-2060.

Tonks, A., A. J. Tonks, L. Pearn, Z. Mohamad, A. K. Burnett and R. L. Darley (2005). Optimized Retroviral Transduction Protocol Which Preserves the Primitive Subpopulation of Human Hematopoietic Cells. *Biotechnology Progress, 21*(3), 953-958.

Tsai, F.-Y., G. Keller, F. C. Kuo, M. Weiss, J. Chen, M. Rosenblatt, F. W. Alt and S. H. Orkin (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature, 371*(6494), 221-226.

Tsai, K.-W., L.-Y. Hu, C.-W. Wu, S.-C. Li, C.-H. Lai, H.-W. Kao, W.-L. Fang and W.-c. Lin (2010). Epigenetic regulation of miR-196b expression in gastric cancer. *Genes, Chromosomes and Cancer, 49*(11), 969-980.

Tsai, M.-M., C.-S. Wang, C.-Y. Tsai, C.-G. Huang, K.-F. Lee, H.-W. Huang, Y.-H. Lin, H.-C. Chi, L.-M. Kuo, P.-H. Lu and K.-H. Lin (2016). Circulating microRNA-196a/b are novel biomarkers associated with metastatic gastric cancer. *European Journal of Cancer, 64*, 137-148.

Tsunoda, T. and T. Takagi (1999). Estimating transcription factor bindability on DNA. *Bioinformatics, 15*(7-8), 622-630.

U.S. Cancer Statistics Working Group. (2017). *Top 10 Cancers by Rates of Cancer Deaths in the United States (2017)*. Retrieved 11.09.2020, from https://gis.cdc.gov/cancer/USCS/DataViz.html.

Undi, R. B., R. Kandi and R. K. Gutti (2013). MicroRNAs as Haematopoiesis Regulators. *Advances in hematology, 2013*, Article 695754.

Valadi, H., K. Ekström, A. Bossios, M. Sjöstrand, J. J. Lee and J. O. Lötvall (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nature Cell Biology, 9*(6), 654-659.

Valen, E. and A. Sandelin (2011). Genomic and chromatin signals underlying transcription start-site selection. *Trends in Genetics, 27*(11), 475-485.

Valent, P., W. Kern, G. Hoermann, J. D. Milosevic Feenstra, K. Sotlar, M. Pfeilstöcker, U. Germing, W. R. Sperr, A. Reiter, D. Wolf, M. Arock, T. Haferlach and H.-P. Horny (2019). Clonal Hematopoiesis with Oncogenic Potential (CHOP): Separation from CHIP and Roads to AML. *International journal of molecular sciences, 20*(3), Article 789.

Vasudevan, S. (2012). Posttranscriptional Upregulation by MicroRNAs. *WIREs RNA, 3*(3), 311-330.

Vasudevan, S., Y. Tong and J. A. Steitz (2007). Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation. *Science, 318*(5858), Article 1931.

Velu, C. S., A. M. Baktula and H. L. Grimes (2009). Gfi1 regulates miR-21 and miR-196b to control myelopoiesis. *Blood, 113*(19), 4720-4728.

Velu, C. S., A. Chaubey, J. D. Phelan, S. R. Horman, M. Wunderlich, M. L. Guzman, A. G. Jegga, N. J. Zeleznik-Le, J. Chen, J. C. Mulloy, J. A. Cancelas, C. T. Jordan, B. J. Aronow, G. Marcucci, B. Bhat, B. Gebelein and H. L. Grimes (2014). Therapeutic antagonists of microRNAs deplete leukemia-initiating cell activity. *The Journal of Clinical Investigation, 124*(1), 222-236.

Vergoulis, T., I. S. Vlachos, P. Alexiou, G. Georgakilas, M. Maragkakis, M. Reczko, S. Gerangelos, N. Koziris, T. Dalamagas and A. G. Hatzigeorgiou (2012). TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Research, 40*(D1), D222-D229.

Vian, L., M. Di Carlo, E. Pelosi, F. Fazi, S. Santoro, A. M. Cerio, A. Boe, V. Rotilio, M. Billi, S. Racanicchi, U. Testa, F. Grignani and C. Nervi (2014). Transcriptional fine-tuning of microRNA-223 levels directs lineage choice of human hematopoietic progenitors. *Cell death and differentiation, 21*(2), 290-301.

Vigorito, E., K. L. Perks, C. Abreu-Goodger, S. Bunting, Z. Xiang, S. Kohlhaas, P. P. Das, E. A. Miska, A. Rodriguez, A. Bradley, K. G. C. Smith, C. Rada, A. J. Enright, K.-M. Toellner, I. C. M. Maclennan and M. Turner (2007). microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity, 27*(6), 847-859.

Villatoro, A., J. Konieczny, V. Cuminetti and L. Arranz (2020). Leukemia Stem Cell Release From the Stem Cell Niche to Treat Acute Myeloid Leukemia. *Frontiers in Cell and Developmental Biology, 8*, Article 607.

Vinces, M. D., M. Legendre, M. Caldara, M. Hagihara and K. J. Verstrepen (2009). Unstable tandem repeats in promoters confer transcriptional evolvability. *Science (New York, N.Y.), 324*(5931), 1213-1216.

Vivanco, I., N. Palaskas, C. Tran, S. P. Finn, G. Getz, N. J. Kennedy, J. Jiao, J. Rose, W. Xie, M. Loda, T. Golub, I. K. Mellinghoff, R. J. Davis, H. Wu and C. L. Sawyers (2007). Identification of the JNK Signaling Pathway as a Functional Target of the Tumor Suppressor PTEN. *Cancer Cell, 11*(6), 555-569.

Vu, T., J. Straube, A. H. Porter, M. Bywater, A. Song, V. Ling, L. Cooper, G. Pali, C. Bruedigam, S. Jacquelin, J. Green, G. Magor, A. Perkins, A. M. Chalk, C. R. Walkley, F. H. Heidel, P. Mukhopadhyay, N. Cloonan, S. Gröschel, J.-P. Mallm, S. Fröhling, C. Scholl and S. W. Lane (2020). Hematopoietic stem and progenitor cell-restricted Cdx2 expression induces transformation to myelodysplasia and acute leukemia. *Nature communications, 11*(1), 3021-3021.

Wald, D. N., H. M. Vermaat, S. Zang, A. Lavik, Z. Kang, G. Peleg, S. L. Gerson, K. D. Bunting, M. L. Agarwal, B. L. Roth and W. Tse (2008). Identification of 6-benzylthioinosine as a myeloid leukemia differentiation-inducing compound. *Cancer research, 68*(11), 4369-4376.

Walker, A. and G. Marcucci (2012). Molecular prognostic factors in cytogenetically normal acute myeloid leukemia. *Expert Review of Hematology, 5*(5), 547-558.

Walzer, T., M. Bléry, J. Chaix, N. Fuseri, L. Chasson, S. H. Robbins, S. Jaeger, P. André, L. Gauthier, L. Daniel, K. Chemin, Y. Morel, M. Dalod, J. Imbert, M. Pierres, A. Moretta, F. Romagné and E. Vivier (2007). Identification, activation, and selective in vivo ablation of mouse NK cells via NKp46. *Proceedings of the National Academy of Sciences, 104*(9), 3384.

Wan, S.-M., S.-K. Yi, J. Zhong, C.-H. Nie, N.-N. Guan, W.-Z. Zhang and Z.-X. Gao (2016). Dynamic mRNA and miRNA expression analysis in response to intermuscular bone development of blunt snout bream (Megalobrama amblycephala). *Scientific Reports, 6*(1), Article 31050.

Wang, H.-L., R. Zhou, J. Liu, Y. Chang, S. Liu, X.-B. Wang, M.-F. Huang and Q. Zhao (2017). MicroRNA-196b inhibits late apoptosis of pancreatic cancer cells by targeting CADM1. *Scientific Reports, 7*(1), Article 11467.

Wang, H., S.-Y. Lin, F.-F. Hu, A.-Y. Guo and H. Hu (2020). The expression and regulation of HOX genes and membrane proteins among different cytogenetic groups of acute myeloid leukemia. *Molecular Genetics & Genomic Medicine, 8*(9), Article e1365.

Wang, J., J. Chen, P. Chang, A. LeBlanc, D. Li, J. L. Abbruzzesse, M. L. Frazier, A. M. Killary and S. Sen (2009). MicroRNAs in Plasma of Pancreatic Ductal Adenocarcinoma Patients as Novel Blood-Based Biomarkers of Disease. *Cancer Prevention Research, 2*(9), Article 807.

Wang, K., S. Zhang, J. Weber, D. Baxter and D. J. Galas (2010). Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic acids research, 38*(20), 7248-7259.

Wang, K. C., Y. W. Yang, B. Liu, A. Sanyal, R. Corces-Zimmerman, Y. Chen, B. R. Lajoie, A. Protacio, R. A. Flynn, R. A. Gupta, J. Wysocka, M. Lei, J. Dekker, J. A. Helms and H. Y. Chang (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature, 472*(7341), 120-124.

Wang, L. D. and A. J. Wagers (2011). Dynamic niches in the origination and differentiation of haematopoietic stem cells. *Nature Reviews Molecular Cell Biology, 12*, Article 643.

Wang, X. (2016). Improving microRNA target prediction by modeling with unambiguously identified microRNA-target pairs from CLIP-ligation studies. *Bioinformatics (Oxford, England), 32*(9), 1316-1322.

Wang, Y., Z. Li, C. He, D. Wang, X. Yuan, J. Chen and J. Jin (2010). MicroRNAs expression signatures are associated with lineage and survival in acute leukemias. *Blood cells, molecules & diseases, 44*(3), 191-197.

Wang, Y. X., X. Y. Zhang, B. F. Zhang, C. Q. Yang, X. M. Chen and H. J. Gao (2010). Initial study of microRNA expression profiles of colonic cancer without lymph node metastasis. *Journal of Digestive Diseases, 11*(1), 50-54.

Ward, A. J. and T. A. Cooper (2010). The pathobiology of splicing. *The Journal of pathology, 220*(2), 152-163.

Wee, L. M., C. F. Flores-Jasso, W. E. Salomon and P. D. Zamore (2012). Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell, 151*(5), 1055-1067.

Wei, G., D. Twomey, J. Lamb, K. Schlis, J. Agarwal, R. W. Stam, J. T. Opferman, S. E. Sallan, M. L. den Boer, R. Pieters, T. R. Golub and S. A. Armstrong (2006). Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell, 10*(4), 331-342.

Weisberg, E., C. Meng, A. Case, M. Sattler, H. L. Tiv, P. C. Gokhale, S. Buhrlage, J. Wang, N. Gray, R. Stone, S. Liu, S. V. Bhagwat, R. V. Tiu, S. Adamia and J. D. Griffin (2020). Evaluation of ERK as a therapeutic target in acute myelogenous leukemia. *Leukemia, 34*(2), 625-629.

Weller, C. L., S. J. Collington, J. K. Brown, H. R. P. Miller, A. Al-Kashi, P. Clark, P. J. Jose, A. Hartnell and T. J. Williams (2005). Leukotriene B4, an activation product of mast cells, is a chemoattractant for their progenitors. *The Journal of experimental medicine, 201*(12), 1961-1971.

Wen, J., Y. Huang, H. Li, X. Zhang, P. Cheng, D. Deng, Z. Peng, J. Luo, W. Zhao, Y. Lai and Z. Liu (2017). Over-expression of miR-196b-5p is significantly associated with the progression of myelodysplastic syndrome. *International Journal of Hematology, 105*(6), 777-783.

Whyatt, D. J., E. deBoer and F. Grosveld (1993). The two zinc finger-like domains of GATA-1 have different DNA binding specificities. *The EMBO journal, 12*(13), 4993-5005.

Wilson, A., E. Laurenti, G. Oser, R. C. van der Wath, W. Blanco-Bose, M. Jaworski, S. Offner, C. F. Dunant, L. Eshkind, E. Bockamp, P. Lió, H. R. MacDonald and A. Trumpp (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell, 135*(6), 1118-1129.

Wirth, F., A. Lubosch, S. Hamelmann and I. A. Nakchbandi (2020). Fibronectin and Its Receptors in Hematopoiesis. *Cells, 9*(12), Article 2717.

Witwer, K. W. and M. K. Halushka (2016). Toward the promise of microRNAs - Enhancing reproducibility and rigor in microRNA research. *RNA biology, 13*(11), 1103-1116.

Wong, P., M. Iwasaki, T. C. P. Somervaille, F. Ficara, C. Carico, C. Arnold, C.-Z. Chen and M. L. Cleary (2010). The miR-17-92 microRNA polycistron regulates MLL leukemia stem cell potential by modulating p21 expression. *Cancer research, 70*(9), 3833-3842.

Wu, S., S. Huang, J. Ding, Y. Zhao, L. Liang, T. Liu, R. Zhan and X. He (2010). Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3′ untranslated region. *Oncogene, 29*(15), 2302-2308.

Wujcik, D. (2003). Molecular biology of leukemia. *Seminars in Oncology Nursing, 19*(2), 83-89.

Wurm, A. A., D. G. Tenen and G. Behre (2017). The Janus-faced Nature of miR-22 in Hematopoiesis: Is It an Oncogenic Tumor Suppressor or Rather a Tumor-Suppressive Oncogene? *PLoS Genet 13*(1), Article e1006505.

Xiao, C., D. P. Calado, G. Galler, T.-H. Thai, H. C. Patterson, J. Wang, N. Rajewsky, T. P. Bender and K. Rajewsky (2007). MiR-150 Controls B Cell Differentiation by Targeting the Transcription Factor c-Myb. *Cell, 131*(1), 146-159.

Xiao, M., J. Li, W. Li, Y. Wang, F. Wu, Y. Xi, L. Zhang, C. Ding, H. Luo, Y. Li, L. Peng, L. Zhao, S. Peng, Y. Xiao, S. Dong, J. Cao and W. Yu (2017). MicroRNAs activate gene transcription epigenetically as an enhancer trigger. *RNA biology, 14*(10), 1326-1334.

Xie, M., M. Li, A. Vilborg, N. Lee, M.-D. Shu, V. Yartseva, N. Šestan and J. A. Steitz (2013). Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell, 155*(7), 1568-1580.

Xiu, Y., Q. Dong, Q. Li, F. Li, N. Borcherding, W. Zhang, B. Boyce, H.-H. Xue and C. Zhao (2018). Stabilization of NF-κB-Inducing Kinase Suppresses MLL-AF9-Induced Acute Myeloid Leukemia. *Cell reports, 22*(2), 350-358.

Xu, C. and L. Gu (2017). The diagnostic effect of serum miR196b as biomarker in colorectal cancer. *Biomedical Reports, 6*, 39-45.

Xu, H. and Q. Wen (2018). Downregulation of miR‑135a predicts poor prognosis in acute myeloid leukemia and regulates leukemia progression via modulating HOXA10 expression. *Mol Med Rep, 18*(1), 1134-1140.

Xu, W., A. San Lucas, Z. Wang and Y. Liu (2014). Identifying microRNA targets in different gene regions. *BMC Bioinformatics, 15*(7), S4.

Yang, B., S.-Z. Li, L. Ma, H.-L. Liu, J. Liu and J.-J. Shao (2016). Expression and mechanism of action of miR-196a in epithelial ovarian cancer. *Asian Pacific Journal of Tropical Medicine, 9*(11), 1105-1110.

Yang, C., E. Bolotin, T. Jiang, F. M. Sladek and E. Martinez (2007). Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene, 389*(1), 52-65.

Yang, G., D. Han, X. Chen, D. Zhang, L. Wang, C. Shi, W. Zhang, C. Li, X. Chen, H. Liu, D. Zhang, J. Kang, F. Peng, Z. Liu, J. Qi, X. Gao, J. Ai, C. Shi and S. Zhao (2014). MiR-196a exerts its oncogenic effect in glioblastoma multiforme by inhibition of IκBα both in vitro and in vivo. *Neuro-oncology, 16*(5), 652-661.

Yang, J.-p., J.-k. Yang, C. Li, Z.-q. Cui, H.-j. Liu, X.-f. Sun, S.-m. Geng, S.-k. Lu, J. Song, C.-y. Guo and B.-h. Jiao (2017). Downregulation of ZMYND11 induced by miR-196a-5p promotes the progression and growth of GBM. *Biochemical and Biophysical Research Communications, 494*(3), 674-680.

Yang, J.-S., T. Maurin, N. Robine, K. D. Rasmussen, K. L. Jeffrey, R. Chandwani, E. P. Papapetrou, M. Sadelain, D. O'Carroll and E. C. Lai (2010). Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America, 107*(34), 15163-15168.

Yang, L., D. Bryder, J. r. Adolfsson, J. Nygren, R. Månsson, M. Sigvardsson and S. E. W. Jacobsen (2005). Identification of Lin–Sca1+kit+CD34+Flt3– short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood, 105*(7), 2717-2723.

Yang, L., F. Peng, J. Qin, H. Zhou and B. Wang (2017). Downregulation of microRNA-196a inhibits human liver cancer cell proliferation and invasion by targeting FOXO1. *Oncology reports, 38*(4), 2148-2154.

Yao, W., G. Guo, Q. Zhang, L. Fan, N. Wu and Y. Bo (2014). The application of multiple miRNA response elements enables oncolytic adenoviruses to possess specificity to glioma cells. *Virology, 458-459*, 69-82.

Yap, J. L., S. Worlikar, A. D. MacKerell, Jr., P. Shapiro and S. Fletcher (2011). Small-molecule inhibitors of the ERK signaling pathway: Towards novel anticancer therapeutics. *ChemMedChem, 6*(1), 38-48.

Yekta, S., I. h. Shih and D. P. Bartel (2004). MicroRNA-Directed Cleavage of HOXB8 mRNA. *Science, 304*(5670), 594-596.

Yekta, S., C. J. Tabin and D. P. Bartel (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nature reviews. Genetics, 9*(10), 789-796.

Yokoyama, A. (2015). Molecular mechanisms of MLL-associated leukemia. *International Journal of Hematology, 101*(4), 352-361.

Yu, C., A. B. Cantor, H. Yang, C. Browne, R. A. Wells, Y. Fujiwara and S. H. Orkin (2002). Targeted deletion of a high-affinity GATA-binding site in the GATA-1 promoter leads to selective loss of the eosinophil lineage in vivo. *The Journal of experimental medicine, 195*(11), 1387-1395.

Yu, J., F. Wang, G.-H. Yang, F.-L. Wang, Y.-N. Ma, Z.-W. Du and J.-W. Zhang (2006). Human microRNA clusters: Genomic organization and expression profile in leukemia cell lines. *Biochemical and Biophysical Research Communications, 349*(1), 59-68.

Yu, Z., X. Lin, M. Tian and W. Chang (2018). microRNA196b promotes cell migration and invasion by targeting FOXP2 in hepatocellular carcinoma. *Oncology Reports, 39*, 731-738.

Yuasa, H., Y. Oike, A. Iwama, I. Nishikata, D. Sugiyama, A. Perkins, M. L. Mucenski, T. Suda and K. Morishita (2005). Oncogenic transcription factor Evi1 regulates hematopoietic stem cell proliferation through GATA-2 expression. *Embo j, 24*(11), 1976-1987.

Yuva-Aydemir, Y., A. Simkin, E. Gascon and F.-B. Gao (2011). MicroRNA-9: functional evolution of a conserved small regulatory RNA. *RNA biology, 8*(4), 557-564.

Zagozdzon, R. and J. Golab (2015). Cancer stem cells in haematological malignancies. *Contemporary oncology (Poznan, Poland), 19*(1A), A1-A6.

Zaidan, N. and K. Ottersbach (2018). The multi-faceted role of Gata3 in developmental haematopoiesis. *Open biology, 8*(11), Article 180152.

Zardo, G., A. Ciolfi, L. Vian, L. M. Starnes, M. Billi, S. Racanicchi, C. Maresca, F. Fazi, L. Travaglini, N. Noguera, M. Mancini, M. Nanni, G. Cimino, F. Lo-Coco, F. Grignani and C. Nervi (2012). Polycombs and microRNA-223 regulate human granulopoiesis by transcriptional control of target gene expression. *Blood, 119*(17), 4034-4046.

Zeijlemaker, W., A. Kelder, Y. J. M. Oussoren-Brockhoff, W. J. Scholten, A. N. Snel, D. Veldhuizen, J. Cloos, G. J. Ossenkoppele and G. J. Schuurhuis (2016). A simple

one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. *Leukemia, 30*(2), 439-446.

Zhai, P.-F., F. Wang, R. Su, H.-S. Lin, C.-L. Jiang, G.-H. Yang, J. Yu and J.-W. Zhang (2014). The regulatory roles of microRNA-146b-5p and its target platelet-derived growth factor receptor α (PDGFRA) in erythropoiesis and megakaryocytopoiesis. *The Journal of biological chemistry, 289*(33), 22600-22613.

Zhan, M., C. P. Miller, T. Papayannopoulou, G. Stamatoyannopoulos and C.-Z. Song (2007). MicroRNA expression dynamics during murine and human erythroid differentiation. *Experimental hematology, 35*(7), 1015-1025.

Zhang, C., C. Yao, H. Li, G. Wang and X. He (2014). Combined elevation of microRNA-196a and microRNA-196b in sera predicts unfavorable prognosis in patients with osteosarcomas. *International journal of molecular sciences, 15*(4), 6544-6555.

Zhang, J., W. Liu, J. Du, Y. Jin, M. Zhao, L. Li and Y. Wang (2018). Prognostic impact of miR-196a/b expression in adult acute myeloid leukaemia: a single-centre, retrospective cohort study. *The Journal of international medical research, 46*(9), 3675-3683.

Zhang, J., F. Varas, M. Stadtfeld, S. Heck, N. Faust and T. Graf (2007). CD41-YFP mice allow in vivo labeling of megakaryocytic cells and reveal a subset of platelets hyperreactive to thrombin stimulation. *Experimental Hematology, 35*(3), 490-499.

Zhang, L., C.-S. Yang, X. Varelas and S. Monti (2016). Altered RNA editing in 3′ UTR perturbs microRNA-mediated regulation of oncogenes and tumor-suppressors. *Scientific Reports, 6*(1), Article 23226.

Zhang, X., W. Ma, J. Cui, H. Yao, H. Zhou, Y. Ge, L. Xiao, X. Hu, B. H. Liu, J. Yang, Y. Y. Li, S. Chen, C. J. Eaves, D. Wu and Y. Zhao (2015). Regulation of p21 by TWIST2 contributes to its tumor-suppressor function in human acute myeloid leukemia. *Oncogene, 34*(23), 3000-3010.

Zhang, Y., M. Li, H. Wang, W. E. Fisher, P. H. Lin, Q. Yao and C. Chen (2008). Profiling of 95 MicroRNAs in Pancreatic Cancer Cell Lines and Surgical Specimens by Real-Time PCR Analysis. *World Journal of Surgery, 33*(4), Article 698.

Zhang, Y., W. Wu, Y. Cheng, D. C. King, R. S. Harris, J. Taylor, F. Chiaromonte and R. C. Hardison (2009). Primary sequence and epigenetic determinants of in vivo occupancy of genomic DNA by GATA1. *Nucleic acids research, 37*(21), 7024-7038.

Zhong, X., A. Prinz, J. Steger, M.-P. Garcia-Cuellar, M. Radsak, A. Bentaher and R. K. Slany (2018). HoxA9 transforms murine myeloid cells by a feedback loop driving expression of key oncogenes and cell cycle control genes. *Blood advances, 2*(22), 3137-3148.

Zhou, J., Y. Q. Ching and W.-J. Chng (2015). Aberrant nuclear factor-kappa B activity in acute myeloid leukemia: from molecular pathogenesis to therapeutic target. *Oncotarget, 6*(8), 5490-5500.

Zhu, Y., D. Wang, F. Wang, T. Li, L. Dong, H. Liu, Y. Ma, F. Jiang, H. Yin, W. Yan, M. Luo, Z. Tang, G. Zhang, Q. Wang, J. Zhang, J. Zhou and J. Yu (2013). A comprehensive analysis of GATA-1-regulated miRNAs reveals miR-23a to be a positive modulator of erythropoiesis. *Nucleic acids research, 41*(7), 4129-4143.

Ziller, M. (2021). *Overexpression of HoxA9 in CD150- hematopoietic stem cells leads to rapid progression of acute myeloid leukemia.* Doctoral thesis, Charité-University hospital of Berlin.

Zou, W. and N. P. Restifo (2010). T(H)17 cells in tumour immunity and immunotherapy. *Nature reviews. Immunology, 10*(4), 248-256.

Acknowledgement

First, and foremost I would like to thank Prof. Dr. C. Buske, who gave me the doctoral position, funded by a stipend from the International Graduate School in Molecular Medicine Ulm. He, together with Dr. V. P. S. Rawat gave me continuing support and mentorship through the entire time. Thank you for believing in me.

I also want to take this opportunity to thank Dr. Shiva Bamezai for his scientific support in writing this thesis. Him together with Dr. Alex Pulikkottil and Dr. Tamoghna Mandal I would also like to thank for all the scientific and other philosophical discussions in meetings or the lunch break. We all became friends for which I am very grateful. I would also like to thank my friend Anne Richter for her great support in creating the illustrations.

Special thanks go to my father, Rainer Hollnagel, who taught me that it is fundamental to understand the things and question them openly in order to help shape the here and now. My first scientific lessons were learned in school with my biology teacher Mrs. Nonnenmacher, who taught me that everything in nature has a reason and it is on me to understand and utilize that knowledge. I owe my present scientific thrill to these people. Thank you!

I would like to dedicate my work to my grandfather, Gerhardt Hollnagel. He was a great thinker and role model with his detailed and calm way of thinking in mathematics.

In particular, I would like to thank my mother, Andrea Hollnagel, and mother-in-law, Monika Götze, for their loving commitment in caring for my children giving me space for my doctoral thesis as it was coming to an end. I would also like to thank my sisters Rosmarie Hollnagel, Susanna Berger and Karolin Hollnagel for their constant moral support.

Last but not least I want to thank my husband, Andreas Götze, who supported me throughout the whole time by giving me security, trust and space to do my research work the way I feel it should be done.

Curriculum vitae

Maria Götze

Education:

2013-2019: **PhD** at Institute of Experimental Cancer Research, University of Ulm

2010-2012: **Master** in Integrative Neuroscience at Otto-von-Guericke University, Magdeburg (Grade: 1.6)

2007-2010: **Bachelor** in Psychology at Technical University Chemnitz (Grade: 1.6)

2006-2007: **Work** and Travel in Australia

2004-2006: **A-level** at Gymnasium Einsiedel, Chemnitz (Grade: 1.6)

2002-2003: **Class 10** at Tettenhall College, England

Research Experience:

2013-2019: PhD student at Prof. Dr. Buske's lab at the Institute of Experimental Cancer Research. Projects:
- Role of **miR-196b** in normal and malignant hematopoiesis
- A **single proto-oncogene Cdx2** converts normal hematopoietic stem- & progenitors into leukemic cells by simultaneously initiating an aberrant self-renewal and proliferation program
- Characterization of PHF19 as **epigenetic regulator** in normal and acute myeloid hematopoiesis.

2013-2013: Research fellow in Prof. Dr. Schumacher's lab at the Institute of **Molecular and Cellular Anatomy**. Project:
- Functional relevance of CALEB/NGC for the differentiation of dendrites, dendritic spines and the formation of synapses.

2010-2012: Master's degree project:
- Inflammatory and metabolic regulation of TTP, TDP-43 and FUS in **microglia**.

Research assistant at the Institute of **Developmental Biology** to generate and evaluate brain slice stainings

Participation in neuroscientific projects (á 4 weeks):
- ß-Amyloid in **Alzheimer**
- **perinatal stress**
- **mRNA binding proteins**
- **axonal regeneration**

2010-2010: Research fellow at King's College in London, England. In the field of **imaging neuroscience**

2007-2010: Bachelor's degree project in the field of psychology:

- Does the graphical representation influence the understanding of simple **dynamic systems**?

Research assistant at the institute of **research methodologies and evaluation**

Intern for other psychology relevant projects (á 4 weeks):

- Evaluation of interviews with employees aiming at **omptimizing a process line** in a chemical company. Done at the institute of work and health in Dresden
- Effect of **pediatric interactive therapy system** at the rehabilitation center Affoltern am Albis, Switzerland

Publications:

- Bamezai, Demir, …. **Götze** et al. 2021. 'TET1 promotes growth of T-cell acute Lymphoblastic leukemia and can be antagonized via RARP inhibition'. Leukemia.
- Rawat and **Götze** et al. 2020. 'The microRNA miR-196b acts as tumor suppressor in Cdx2-driven acute myeloid leukemia'. Haematologica
- Daria, Kirsten, …, **Hollnagel** et al. 2016. 'GPR56 contributes to the development of acute myeloid leukemia in mice'. Leukemia

Date, Signature