Ulm University Medical Center

Institute of Immunology

Director: Prof. Dr. Hassan Jumaa

Generation and Analysis of a Novel, Cre-Inducible Knock-in Mouse Strain to Assess the Leukemogenic Potential of the Mutation H3.3-K27M in the context of Mll5-Deficiency

Dissertation submitted to obtain the doctoral degree of Human Biology of the Medical Faculty of Ulm University

Karolin Klumb

Oberwesel

2023

Acting Dean:	Prof. Dr. Thomas Wirth
1st Reviewer:	Prof. Dr. Hans Jörg Fehling
2nd Reviewer:	Prof. Dr. Michaela Feuring
Day of Graduation:	12.01.2024

Table of Contents

List of Abbreviations V
List of Figures VIII
List of Tables X
1. Introduction 1
1.1 Acute Myeloid Leukemia (AML) and Myelodysplastic Neoplasms (MDS) 1
1.2 -7/del(7q) in Myeloid Malignancies 2
1.3 A highly conspicuous novel gene in 7q22: *Mixed Lineage Leukemia 5* (MLL5) 4
1.4 What is known about the physiological function of MLL5? 6
1.5 The phenotype of *Mll5*-deficient mice shows striking resemblance to mouse mutants lacking the DNA repair and tumor suppressor protein ATM 8
1.6 Identification of potentially cooperating leukemogenic gene mutations associated with heterozygous loss of MLL5 10
1.7 Experimentally mimicing leukemogenesis in mice 18
1.8 Aim of the Project 21
2. Material and Methods 22
2.1 Material 22
2.2 Methods 37
3. Results 63
3.1 Introduction of a Cre inducible K27M mutation into the endogenous H3f3a gene locus of embryonic stem cells via gene targeting 63
3.2 Targeting of the H3f3a gene locus with more sophisticated targeting vectors 66
3.3 Characterization of targeted ES clones 67
3.4 Generation of H3K27M knock-in mice from targeted ES cells 70
3.5 Verification of recombination efficiency and specificity with vCre x RoYFP Mice 73
3.6 $vCre^{pos}$ x H3K27M knock-in mice correctly express the *H3f3a-K27M* allele specifically in hematopoietic cells 76
3.7 IRES-driven vYFP reporter expression faithfully reflects *H3f3a-K27M* expression and transcriptional activity of the endogenous *H3f3a* gene locus 78
3.8 The expression level of the *H3f3a-K27M-IRES-vYFP* knock-in allele seems significantly reduced in comparison with *H3f3a* wild-type alleles 80
3.9 Cre-mediated induction of K27M-mutated histone variant H3.3 results in marked reduction of global H3K27 trimethylation 84
3.10 Reduced H3K27me3 does not affect the phenotype of H3K27M mutant mice 86
3.11 H3K27M does not affect the survival and leukemia incidence in Mll5 and H3K27M compound mice 89
3.12 Introduction of IDH2-R172K does not cause leukemia in Mll5-deficient and H3K27M-mutant mice 91
4. Discussion 101
4.1 Generation of knock-in mice with a Cre-inducible H3f3a-K27M mutation 101

4.2 Phenotypic effects of hematopoietic cell compartment-specific *H3f3a-K27M* expression in knock-in mice 106
4.3 Potential role of *H3f3a-K27M* in leukemogenesis and future lines of investigation 108
5. Summary 113
6. References 114
Appendix 127
Full name of genes mentioned in Figure 1, Table 2, and Table 4 127
Plasmids used during the Project 131
Phenotypic analysis of cell subsets of thymus and spleen 140
Individual weight gain of mice transplanted with IDH2-R172K transduced bone marrow . 146
Survival of mice during *Dnmt3a*$^{flox/flox}$ breeding 146
Acknowledgements 147
Curriculum Vitae 148

List of Abbreviations

Abbreviation	
µF	Microfarad (Unit of Electrical Capacitance)
2-HG	2-Hydroxyglutarate
5-FU	5-Flourouracil
-7	Monosomy 7
aa	Amino Acid(s)
AITL	Angioimmunoblastic T-cell Lymphoma
AML	Acute Myeloid Leukemia
ATCC	American Type Culture Collection
ATM	Ataxia Telangiectasia Mutated
BM	Bone Marrow
BSA	Bovine Serum Albumin
CIP	Calf Intestinal Phosphatase
CMP	Common Myeloid Progenitor
del(7q)	Deletion of the long arm of chromosome 7
DN	Double Negative
DNMT3A	DNA Methyltransferase 3A
DP	Double Positive
DSB	Double Strand Break
DTA	Diphtheria Toxin A
EGFP	Enhanced Green Fluorescent Protein
ERT2	Estrogen Receptor Domain 2
ES Cells	Embryonic Stem Cells
EZH2	Enhancer of Zeste Homolog 2
FACS	Fluorescence Activated Cell Sorting
FLT3	FMS-like tyrosine kinase 3
gDNA	Genomic DNA
GMP	Granulocyte-Monocyte Progenitor
Gy	Gray (Unit of Ionizing Radiation)
H3K27	Histone 3 Lysine 27
H3K4	Histone 3 Lysine 4
HDAC	Histone Deacetylase
HI-FCS	Heat-Inactivated Fetal Calf Serum
HMT	Histone Methyltransferase
HSC	Hematopoietic Stem Cells
HSPCs	Hematopoietic Stem and Progenitor Cells
IDH1/2	Isocitrate Dehydrogenase 1/2
IFN-1	Interferon 1
IRES	Internal Ribosomal Entry Site
ITD	Internal Tandem Duplication
K	Lysine
kb	Kilo base

kDa	Kilo Dalton
KI	Knock-In
KMT2	Lysine N-methyl Transferases 2
KO	Knock-Out
LK	Lineage^{neg}, cKit^{pos}, Sca1^{neg} Cells
LMPP	Lymphoid-Myeloid Progenitor
LN	Lymph Nodes
LSK	Lineage^{neg}, cKit^{pos}, Sca1^{pos} Cells
LT-HSC	Long-Term Hematopoietic Stem Cells
M	Methionine
MDS	Myelodysplastic Neoplasms
me3	Trimethylation
MEF	Mouse Embryonic Fibroblasts
MEP	Megakaryocyte-Erythroid Progenitor
MFI	Mean Fluorescence Intensity
MLL	Mixed Lineage Leukemia
MSCV	Murine Stem Cell Virus
mV	Millivolt (Unit of Electrical Potential)
NAC	N-acetyl-L-cysteine
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
neg	Negative
Neo	Neomycin (G418)
NGS	Next Generation Sequencing
pA	Poly A tail
PBC	Peripheral Blood Cells
Pc-G	Polycomb-group of Proteins
PCR	Polymerase Chain Reaction
PDGFRA	Platelet Derived Growth Factor Receptor Alpha
Pen/Strep	Penicillin/Streptomycin
PHD	Plant Homeodomain Zinc Finger
pHGG	Pediatric High-grade Glioma
pos	Positive
PRC2	Polycomb Repressive Complex 2
PTM	Post-Translational Modifications
qRT-PCR	Quantitative Real-Time PCR
R	Arginine
RBC	Red Blood Cells
Ro	ROSA26 (Reverse Oriented Splice Acceptor, Clone26)
ROS	Reactive Oxygen Species
rSAP	Recombinant Shrimp Alkaline Phosphatase
RT	Room Temperature
RUNX	Runt-Related Transcription Factor
SA	Artificial Splice Acceptor
SAM	S-Adenosyl Methionine

SD	Standard Deviation
SEM	Standard Error of the Mean
SET	Su(var)3-9, Enhancer-of-zeste, and Trithorax
ST-HSC	Short-Term Hematopoietic Stem Cells
Trx-G	Trithorax-group of Proteins
TSG	Tumor Suppressor Gene
Tx	Transplantation
vCre	Vav Cre
vYFP	Venus Yellow Fluorescent Protein
WBC	White Blood Cells
WBI	Whole-Body γ-Irradiation
WES	Whole Exome Sequencing
WT	Wild-type
α-KG	α-Ketoglutarate

List of Figures

Figure 1: Two commonly deleted regions on chromosome 7 3
Figure 2: Members of the MLL/ Trithorax Family 4
Figure 3: Comparison of MLL1 and MLL5 protein domain structure 5
Figure 4: Phylogenetic tree of the MLL-family 7
Figure 5: Schematic representation of physiologic and pathophysiologic IDH2 activities within the cell 12
Figure 6: Schematic presentation of the nucleosome with a focus on histone 3 variants 15
Figure 7: Strategies to mimic leukemia development in murine models 19
Figure 8: Genotyping of H3f3aK27M 38
Figure 9: Genotyping of Mll5^{flox} 39
Figure 10: Genotyping of vCre 40
Figure 11: Simplified workflow gene targeting 49
Figure 12: Targeted ES cell colonies before picking 51
Figure 13: pFG-H3TG-CB control construct 51
Figure 14: Example for an RNA quality analysis using agarose gel electrophoresis 59
Figure 15: Targeting strategy (vector pKK-H3TG-AD) 64
Figure 16: Screening for homologous recombination after targeting with vector pKK-H3TG-AD 65
Figure 17: Structure of the targeting vectors used and comparison of targeting results 66
Figure 18: Flow cytometric analysis of targeted ES cells 68
Figure 19: Verification of H3f3a-K27M expression using RT-PCR and Western Blot 69
Figure 20: Sequencing results for pKK-H3TG-AG #10.a RT-PCR product 70
Figure 21: Quantitative and highly specific YFP marker expression in CD45^{pos} thymocytes reflects Cre-recombinase activity in vCre x RoYFP mice 74
Figure 22: Quantitative and highly specific YFP marker expression in CD45^{pos} splenocytes reflects Cre-recombinase activity in vCre x RoYFP mice 75
Figure 23: Quantitative and highly specific YFP marker expression in CD45^{pos} bone marrow cells reflects Cre-recombinase activity in vCre x RoYFP mice 76
Figure 24: Sequencing of full-length H3f3a transcript in thymus and spleen of a vCre^{pos} x H3K27M$^{KI/WT}$ mouse 77
Figure 25: Detection of H3f3a-K27M transcripts in vCre x H3K27M mice 78
Figure 26: H3f3a-K27M reporter expression (vYFP) in hematological organs of mice 79
Figure 27: Significantly different H3f3a expression levels in B- and T-lymphocytes are faithfully reflected by vYFP marker expression 80
Figure 28: Comparison of H3f3a transcription levels in vCre x H3K27M mice and appropriate controls 82
Figure 29: Western Blot detecting H3f3a-K27M mutant proteins 83
Figure 30: Cre-mediated H3.3A-K27M expression in hematopoietic cells from our knock-in mice reduces global H3K27me3 85
Figure 31: Representative results for flow cytometric of hematopoietic compartments in bone marrow of younger mice 86
Figure 32: Quantification of cell subsets in bone marrow of younger mice 87
Figure 33: Representative results for flow cytometric analysis of hematopoietic compartments in bone marrow of older mice 88
Figure 34: Quantification of cell subsets in bone marrow of aged mice 89
Figure 35: Comparison of survival curves from H3K27M-mutant and Mll5-deficient mice 90

Figure 36: Blood cell count at death monitor mice 91
Figure 37: Simplified workflow of retroviral bone marrow transduction followed by transplantation 92
Figure 38: Gating strategy for monthly blood analysis 94
Figure 39: Engraftment and expression of the retrovirus in donor derived cells in control mice 95
Figure 40: Summary of engraftment and expression of the retrovirus in donor derived cells . 96
Figure 41: Relative weight change of transplanted mice per genotype 97
Figure 42: Relative weight gain of all mice transplanted with IDH2-172K transduced $vCre^{pos}$ x $Mll5^{FL/WT}$ x $H3K27M^{KI/WT}$ bone marrow 98
Figure 43: Comparison of survival of transplanted mice 99
Figure 44: Blood cell count at death of mice transplanted with IDH2-R172K transduced bone marrow cells 100
Figure 45A: Plasmid map of pCL-Eco helper plasmid used during generation of retroviral supernatants 131
Figure 46A: Plasmid map of piCre (Cre-recombinase) used to activate K27M mutation in targeted ES cells 132
Figure 47A: Plasmid map of pCAGGS-FLPe-V1 (Flp-recombinase) used to delete Neo from targeted ES cells 133
Figure 48A: Plasmid map of retroviral vector pKK-ID-AFM used for transduction of BM cells 134
Figure 49A: Plasmid map of targeting construct obtained from the Jeffrey Mann laboratory 135
Figure 50A: Plasmid map of the control construct pFG-H3TG-CB 136
Figure 51A: Plasmid map of first targeting vector pKK-H3TG-AD 137
Figure 52A: Plasmid map of second targeting vector pKK-H3TG-AF 138
Figure 53A: Plasmid map of final targeting vector pKK-H3TG-AG 139
Figure 54A: Comparison of hematopoietic compartments in thymi of younger mice 140
Figure 55A: Quantification of cell subsets in thymi of younger mice 141
Figure 56A: Comparison of hematopoietic compartments in spleens of younger mice 142
Figure 57A: Quantification of cell subsets in spleens of younger mice 142
Figure 58A:Comparison of hematopoietic compartments in thymi of older mice 143
Figure 59A: Quantification of cell subsets in thymi of older mice 144
Figure 60A: Comparison of hematopoietic compartments in spleens of older mice 145
Figure 61A: Quantification of cell subsets in spleens of older mice 145
Figure 62A: Weight gain of individual mice transplanted with IDH2-R172K transduced bone marrow 146
Figure 63A. Survival of mice generated during intercross of vCre x $Mll5^{flox}$ mice with $Dnmt3a^{flox}$ mice 146

List of Tables

Table 1: Gene mutations identified by WES in leukemic cells from AML patients with 7q deletions, including loss of MLL5 10
Table 2: Frequency of recurrent gene mutations in the 7q-deleted patient pool 11
Table 3: Simultaneous occurrence of project-relevant mutations in individual patients form our del7q cohort 11
Table 4: Somatic gene mutations in another del(7q) AML patient with a H3.3-K27 substitution identified by WES in the laboratory of Prof. Dr. Christoph Plass (DKFZ Heidelberg) 17
Table 5: Exemplary quick ligation mix used for the ligation of pBSK-KK-AA 42
Table 6: Exemplary overnight ligation mix used for the ligation of pKK-H3TG-AF 42
Table 7: Composition of resolving gel, amount given for two gels 55
Table 8: Composition of stacking gel, amount given for two gels 56
Table 9: Comparison of homologous recombination frequencies observed with the three distinct targeting constructs 67
Table 10: Summary of blastocyst injection data 71
Table 11: Summary of breeding for germline transmission 71
Table 12: Mice transferred from Dresden to the TFZ Ulm 72
Table 13: Genotype distribution of H3K27M intercrosses 73
Table 14: Distribution of vCre alleles in offspring from heterozygous $vCre^{pos}$ and homozygous $H3K27M^{KI/KI}$ mice. 73
Table 15: Overview of transplanted genotypes 93

1. Introduction

1.1 Acute Myeloid Leukemia (AML) and Myelodysplastic Neoplasms (MDS)

Acute Myeloid Leukemia (AML) is a rapidly progressing cancer of the bone marrow and blood. Its defining hallmark is the uncontrolled clonal expansion and accumulation of myeloid precursors that prevent normal bone marrow (BM) hematopoiesis (DiNardo and Cortes, 2016). With a portion of ~20 % of newly diagnosed leukemias, AML is one of the most common blood cancers, with increasing numbers worldwide (Dong et al., 2020). Despite decades of intensive research and regardless of a number of new drugs that have been approved within the past years, treatment options are still limited and prognosis worsens with age: 60 – 65% of patients who are 60 years of age or younger and 85 – 95% of patients older than 60 years are incurable and die (Döhner et al., 2015). To develop novel therapeutic approaches, it is crucial to identify genetic aberrations, which causally contribute to AML, and to unravel the underlying molecular mechanisms.

Myelodysplastic neoplasms (formerly known as myelodysplastic syndromes, for this reason, abbreviated as MDS) are myeloid malignancies tightly related to AML. MDS are a heterogeneous group of BM defects that lead to insufficient hematopoiesis making it a pre-leukemic disease with a risk of progression to AML (Khwaja et al., 2016). One of the decisive factors to distinguish between AML and MDS is the blast count. Patients with a blast count greater than 20 % are assigned to AML and require immediate treatment. Furthermore, the type of cytogenetic and molecular aberration identified is used to differentiate between AML and MDS. For example, rearrangements in the core binding factor are considered “AML defining” regardless of the blast count (Estey et al., 2022; Khoury et al., 2022).

Therapy-related leukemia (t-AML/MDS) frequently occurs after treatment for primary malignancies with alkylating agents or radiation. Disease onset takes about 3 – 5 years and is related to poor prognosis with a median survival of 6 months (Bhatia, 2013). t-AML accounts for 7 – 10 % of total AML cases (Kayser et al., 2011) and is often accompanied by chromosome 5 and/or 7 abnormalities (Döhner et al., 2022), which is of particular interest in the context of this thesis. Alkylating agents transfer alkyl groups to the DNA, leading to DNA damage by abnormal base pairing and resulting in double-strand breaks (Bhatia, 2013).

Like other cancers, AML and MDS are genetically driven diseases. Typically, the genetic aberrations that eventually give rise to these myeloid malignancies are found in the hematopoietic stem and progenitor cells (HSPCs) (Döhner et al., 2015), which have the capacity to self-renew and thus to propagate neoplastic clones. Mutations within this cell compartment are responsible for disease development. The identification of disease-causing and disease-promoting mutations is one of the main goals of leukemia research. Knowledge of these genes and their pathophysiological influence on disease progression is key for the development of novel therapeutic approaches. The progressive improvements in Next Generation Sequencing (NGS) technology over the past decade allowed the complete sequencing of whole exomes/genomes of leukemic cells from individual patients. This led to the discovery, that despite the genetic heterogeneity of AML, only a limited number of recurrent mutations with clinical relevance are found in AML patients. Frequently identified driver

mutations include an internal tandem duplication within the *FLT3* gene (FLT3-ITD; ~20%) and specific mutations in *DNMT3A* (~20%), *RUNX1* (~10%), and *IDH2* (~10-20%) (Kantarjian et al., 2021). The relevance of these mutations with respect to leukemogenesis has been confirmed in respective genetically engineered mouse mutants (Kurtz et al., 2022; Mohanty and Heuser, 2021)

Leukemogenesis is considered a two-step process requiring mutations of two different classes to initiate malignant transformation in myeloid precursors (Gilliland and Griffin, 2002; Takahashi, 2011). Class I mutations induce proliferation and/or inhibit apoptosis. Typically, these mutations either activate (receptor) kinases or other oncogenes, or inactivate tumor suppressors. One common example of class I mutations are FLT3-ITDs. Among other effects, they cause constitutive activation of the STAT pathway and thereby promote cell survival (Mizuki et al., 2000). Class II mutations are typically loss-of-function mutations, for instance in certain transcription factors, leading to impaired differentiation of HSPCs. Another common example for class II mutations are translocations within the *MLL* gene family, that will be introduced below. Usually, the co-occurrence of just one mutation within each class is sufficient to initiate leukemogenesis (Gilliland and Griffin, 2002; Kitamura et al., 2016)

Identifying and understanding the underlying molecular mechanisms of recurrent genetic aberrations in leukemogenesis is a prerequisite for the development of novel therapeutic approaches. To this end, it is essential to identify genes and gene mutations that causally contribute to disease pathology. As part of my Ph.D. project, I have set out to analyze the pathologic impact of a number of mutations that have been identified in a small cohort of AML patients with a recurrent genetic aberration involving chromosome 7.

1.2 -7/del(7q) in Myeloid Malignancies

Among the most frequent chromosomal abnormalities in AML and MDS are monosomy 7 (-7) and specific deletions within the long arm of chromosome 7 (del(7q)) (Honda et al., 2015; Inaba et al., 2018). These deletions are associated with aggressive tumors, poor prognosis, and high relapse rates. Patients respond poorly to chemotherapy, are more susceptible to infections, and have shorter overall survival times (Fröhling et al., 2001; Grimwade et al., 2010; Kratz et al., 2001).

In many tumors, disease progression is triggered by the loss of a tumor suppressor gene (TSG). To determine whether a similar process also occurs in 7q deletions, potential candidate genes need to be identified and functionally characterized. 7q deletions in AML are often quite large and span several mega bases of genetic material. As a first step towards the identification of potential TSGs, commonly deleted regions were identified and carefully screened for the presence of candidate genes. Such minimally deleted regions were determined to reside within 7q21.3 – 22.2, and 7q32 – 35 (Curtiss et al., 2005; Dohner et al., 2006; Fischer et al., 1997; Le Beau et al., 1996). Especially relevant for my Ph.D. project is the segment spanning 7q21.3 – 7q22.2 (Figure 1).

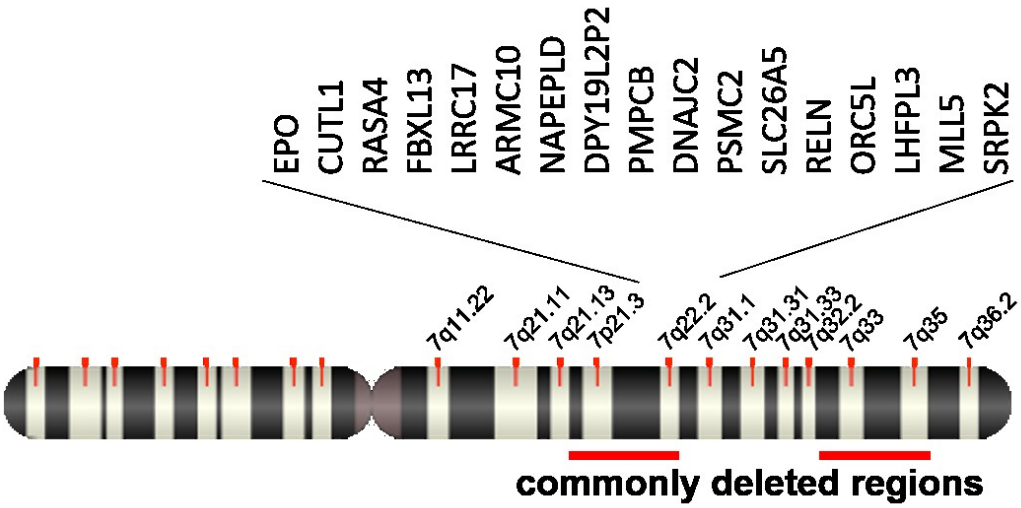


Figure 1: Two commonly deleted regions on chromosome 7
Highlighted in red are two regions on chromosome 7, which are recurrently deleted in leukemic cells from AML and MDS patients. For region 7q21.3 – 7q22.2, which is of particular relevance in the context of my Ph.D. project, the names of all identified genes are given. Full names of genes are given in the appendix on page 127. Based on (Curtiss et al., 2005; Dohner et al., 2006; Fischer et al., 1997; Le Beau et al., 1996).

One of the most frequently deleted sub-regions within the 7q21.3 – 22.2 segment is 7q22 (Le Beau et al., 1996). To investigate whether any of the genes located in this sub-region might act as a classical TSG, transcripts of the second, intact allele were isolated from leukemic cells and screened for potentially inactivating mutations. However, disappointingly, these comprehensive, painstaking studies failed to uncover biallelic inactivation or epigenetic silencing of any of the potential TSGs located in 7q22. Hence, it was hypothesized as an alternative possibility that the haplo-insufficiency of one or more TSGs located in 7q22 might be sufficient for leukemogenesis, not requiring biallelic gene defects, i.e. loss of heterozygosity. To evaluate this adjusted hypothesis, a mouse model with inducible deletion of the entire 7q22 segment ranging from *Fbxl13* to *Srpk2* (see Figure 1) was generated. Upon Mx1-Cre mediated heterozygous knock-out of this segment specifically in hematopoietic cells, the mice had normal blood counts and failed to show any leukemia-like features (Wong et al., 2010), thus providing no evidence for haplo-insufficient TSGs in 7q22. In fact, these findings strongly argue against the simple view that loss of hypothetical TSGs within 7q22 is sufficient to trigger leukemogenesis. Instead, the result would be consistent with the assumption that heterozygous loss of a particular gene in 7q22 might cooperate with one or more specific gene mutations located in other regions of the genome, providing an equally plausible explanation of why heterozygous deletions of 7q22 are frequently observed in AML patients. Furthermore, this alternative hypothesis would be perfectly in line with the previously described two-step model of leukemogenesis. If correct, heterozygous loss of a specific gene located in 7q22 would sensitize HSPCs for leukemogenesis upon additional mutations elsewhere in the genome. Which gene in 7q22 might be a promising candidate for this scenario?

1.3 A highly conspicuous novel gene in 7q22: *Mixed Lineage Leukemia 5* (MLL5)

A detailed characterization of the 7q22 region led to the identification of a previously unknown gene by two independent groups (Döhner et al., 2002; Emerling et al., 2002), one of them headed by our collaboration partner Prof. Dr.Konstanze Döhner at the University Hospital Ulm. Due to the striking resemblance to *MLL1*, the founding member of the Mixed Lineage Leukemia (MLL) family, the newly identified gene was assigned as the fifth member to the MLL family, hence named *MLL5*.

Members of the MLL family are also known as lysine N-methyl transferases 2 (KMT2). They were originally defined as proteins homologous to Trithorax in *Drosophila* (Zhang et al., 2017). Trithorax group (Trx-G) proteins are epigenetic regulators that maintain an active transcriptional state and control the expression of Hox genes and other genes involved in development. Antagonists for Trx-G are proteins of the Polycomb-group of proteins (Pc-G) that retain a repressed state of chromatin (Kingston and Tamkun, 2014).

Initially, the defining feature of the MLL/Trithorax family was the concomitant possession of two characteristic protein domains, the plant homeodomain zinc finger (PHD) and the *Drosophila* suppressor of variegation 3-9, enhancer of zeste, and trithorax (SET) domains. Some members of this family additionally exhibit DNA-binding AT-hooks and protein-binding domains such as bromodomains, that are involved in protein-histone association and chromatin remodeling (Basheer and Huntly, 2015; Rao and Dou, 2015). Figure 2 provides an overview of all human genes currently included in the MLL family with their respective original and revised names. Notably, each of the human genes has a corresponding mouse ortholog (*Mll1* – *Mll4*, *Setd1a*, *Setd1b*, *Mll5*, and *Setd5*).



Figure 2: Members of the MLL/ Trithorax Family
All members of the human MLL/Trithorax family are listed. Based on structural analysis MLL5 and SETD5 form a subgroup. Revised names are included (based on (Zhang et al., 2017).

What is known about the function of the two characteristic protein motifs of MLL family members, i.e. the SET and the PHD domain? The SET domain is a common element of chromatin regulators and typically confers histone methyl transferase (HMT) activity. The HMT activity of MLL-SET domains is specific for histone 3 lysine 4 (H3K4) (Dillon et al., 2005; Zhang et al., 2017). PHD fingers, on the other hand, are protein-protein interaction motifs that often participate in the control gene expression through interaction with covalently modified histones or recruitment of chromatin regulator- and transcription factor-complexes (Jain et al., 2020; Sanchez and Zhou, 2011).

As mentioned above, the newly identified *MLL5* gene exhibits striking similarities to *MLL1*. Despite differences in protein size and the presence of additional protein motifs in MLL1, highly significant homologies are apparent. For instance, *MLL5* has one PHD and one SET domain,

both defining features of MLL/Trithorax proteins. Over its entire amino acid sequence, *MLL5* is 38% identical and 51% similar to *MLL1*, notably with the highest degree of homology in the evolutionary conserved PHD and SET domains (Emerling et al., 2002), possibly indicating grossly similar physiological functions. A simplified comparison of the protein domain composition of MLL1 and MLL5 is shown in Figure 3.

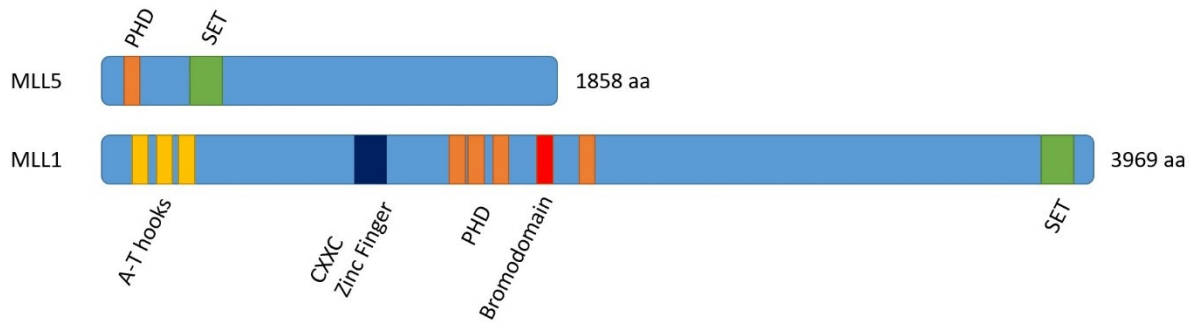


Figure 3: Comparison of MLL1 and MLL5 protein domain structure
The sequence of *MLL5* is 38 % identical and 51 % similar to MLL1. *MLL1* carries four PHD domains (orange), one SET domain (green), several DNA binding motifs (yellow and blue), and one bromodomain (red). *MLL5* lacks DNA binding and bromo motifs and carries just one single PHD and one SET domain (based on (Emerling et al., 2002)).

Of note, members of the MLL family are frequently mutated in various types of human cancer (Rao and Dou, 2015). In fact, *MLL1* translocations are among the most recurrent chromosomal aberrations in AML. The prevalence of this mutation is strongly dependent on the age of the patient, while 5 – 11 % of adult AML patients carry this mutation, the number rises up to 33 – 75 % in infants (Britten et al., 2019). The analysis of mice carrying engineered *MLL1* translocations has provided conclusive evidence that these mutations are oncogenic drivers during leukemogenesis (Milne, 2017). Mutations in *MLL2* have been reported to cause genome instability, which is associated with the progression of several cancers (Kantidakis et al., 2016). Like *MLL5*, *MLL3* is also located on chromosome 7 but in 7q36.1 just outside the commonly deleted regions depicted in Figure 1. *MLL3* has been shown to act as a haplo-insufficient tumor suppressor with loss of function mutations disrupting normal development of HSPCs, resulting in an MDS-like state (Chen et al., 2014). Remarkably, despite this strong phenotype, mutations in *MLL3* alone are not sufficient to cause leukemia but require additional genetic alterations (Chen et al., 2014). *MLL3* thus provides an instructive example for the validity of our working hypothesis formulated at the end of section 1.2, i.e. that heterozygous loss of a gene in 7q might cooperate with mutations elsewhere in the genome to trigger leukemia. Finally, also *MLL4* has been shown to be associated with leukemogenesis, playing complex roles, e.g. by inhibiting MLL1-fusion-induced leukemia on one hand (Santos et al., 2014) and by acting as a haplo-insufficient tumor suppressor in B-cell lymphoma on the other (Yang and Ernst, 2017).

Taken together, these findings highlight the important role of *MLL* genes in hematopoietic malignancies, especially in the context of the two-step theory of leukemogenesis. The homology of *MLL5* to these genes, especially to *MLL1*, might indicate involvement in similar processes making it a promising target for further evaluation in the context of leukemogenesis.

1.4 What is known about the physiological function of MLL5?

When *MLL5* was identified as a novel gene in 7q22, one could only speculate about its *in vivo* role. Ubiquitous expression in all analyzed tissues (Emerling et al., 2002) suggested some important housekeeping function. Furthermore, significant sequence homology with evolutionarily conserved parts of *MLL1* and in particular the presence of a SET domain suggested HMT activity, and thus a role as an epigenetic regulator of gene transcription, in line with known functions of other members of the MLL/Trithorax family. To obtain first experimental data on more specific functions of MLL5 and at the same time to test, whether hetero- or homozygous loss of *MLL5* might be sufficient to promote the development of leukemia, three independent groups, one of which was our laboratory, generated *MLL5* deficient mice via classical gene targeting. The highly overlapping findings reported in side-by-side publications (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009) along with an editorial (Liu et al., 2009) revealed informative, albeit complex phenotypic effects.

Homozygous *Mll5*$^{-/-}$ fetuses developed normally during gestation until day 18.5, indicating that *Mll5* is dispensable during embryogenesis. However, approximately 40 – 50 % of homozygous *Mll5*-deficient pups died during the first days of postnatal life for reasons still unknown. *Mll5*$^{-/-}$ pups surviving this critical period developed into outwardly healthy adults with an apparently normal life span. However, within litters, surviving *Mll5*$^{-/-}$ animals could be readily identified by their reduced size – a sign of mild growth retardation. Additionally, *Mll5*$^{-/-}$ female mice turned out to be sub-fertile while male mice were completely sterile due to a defect in spermiogenesis leading to the absence of mature sperm. Fortunately, in heterozygous mice fertility was not at all impaired, allowing maintenance of the constitutive *Mll5*$^{-/-}$ knock-out mouse strain by breeding heterozygotes. Apart from the described phenotype, *Mll5* deficient mice exhibited severe defects in lymphopoiesis and early hematopoiesis – characteristic phenotypic features, which provided first clear-cut indications for a potential role of *Mll5*-deficiency in leukemogenesis, as will be discussed in detail later in this chapter.

In the publication describing *Mll5*-deficient mice, our group also provided first, albeit circumstantial evidence that the SET domain of MLL5 – in contrast to the common view – might not be able to confer HMT activity (Madan et al., 2009). More targeted, later studies by other groups conclusively confirmed this initial suspicion. It could be shown that this lack of HMT activity arises from the inability of the MLL5 SET domain to interact neither with the methyl donor S-Adenosyl methionine (SAM) nor with the H3 substrate (Mas-y-Mas et al., 2016). Interaction with the essential cofactor SAM is one important prerequisite for the catalysis of methyl transfer to H3K4.

Further phylogenetic, structural, and functional analysis of the MLL5 SET domain revealed significant differences to the SET domains of the other MLL family members – with the exception of SETD5, which is a MLL5 paralog carrying a similar SET domain (Mas-y-Mas et al., 2016). Thus, together with SETD5, MLL5 is now considered an atypical, more distantly related MLL family member characterized by the selective loss of intrinsic HMT activity (Figure 4).

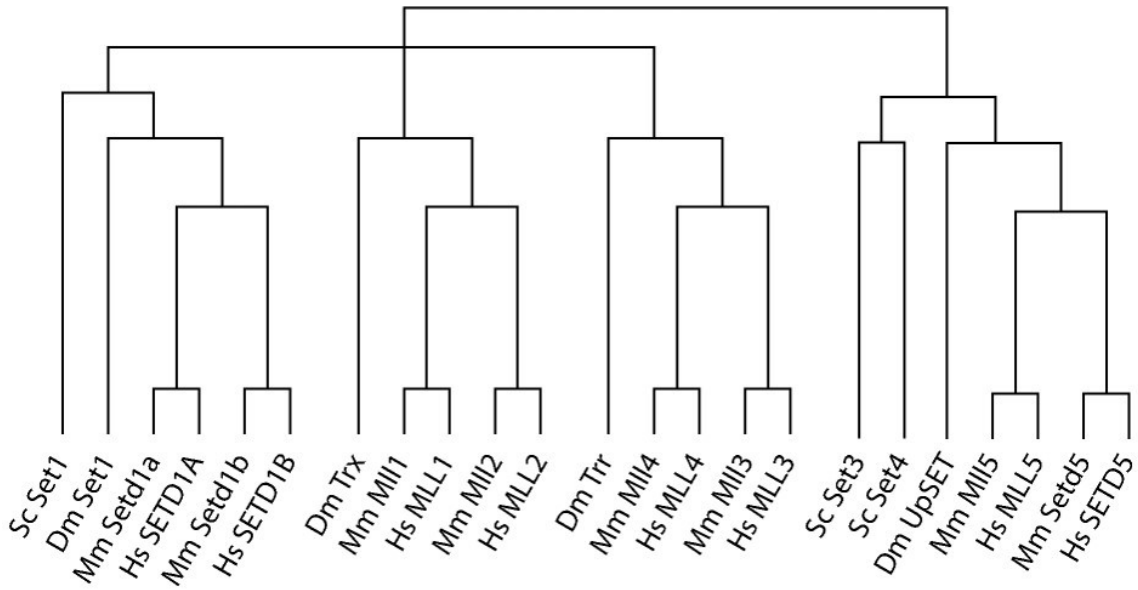


Figure 4: Phylogenetic tree of the MLL-family
Phylogenetic tree of the MLL family highlighting the more distant relationship between MLL5/ SETD5 and the other members of this family. Dm = *Drosophila melanogaster*, Hs = *Homo sapiens*, Mm = *Mus musculus*, Sc = *Saccharomyces cerevisiae*, Mll = Mixed-lineage leukemia, Trx = Trithorax, Trr = Trithorax-related. Based on (Crump and Milne, 2019; Mas-y-Mas et al., 2016).

Apart from an atypical SET domain, most other findings provide clear-cut support to the view that MLL5, like other, more classical MLL family members, represents some kind of epigenetic regulator of gene expression. For instance, the PHD domain of MLL5 has been shown to specifically bind to tri-methylated lysine at position 4 of histone H3 (H3K4me3) (Ali et al., 2013). Interestingly, this is exactly the amino acid residue that is methylated by the other, classical MLL family members. Thus, MLL5 seems to be recruited to actively transcribed genes that display H3K4me3 via recognition of this histone mark by the PHD domain (Ali et al., 2013).Further valuable insights into the potential epigenetic functions of MLL5 have been gained by comprehensive *in vivo* studies of its *Drosophila* ortholog “UpSET” (Rincon-Arano et al., 2012; Rincon-Arano et al., 2013). Like human MLL5, *Drosophila* UpSET is a nuclear protein localized in gene-rich euchromatin regions. Of note, *Drosophila* UpSET is part of a histone deacetylase (HDAC)-containing complex that restricts chromatin accessibility and histone acetylation to promotor regions. In the absence of UpSET, active chromatin marks and chromatin accessibility increase and spread to flanking regions. This leads to elevated levels of “transcriptional noise”, including the activation of repetitive elements and off-target genes (Rincon-Arano et al., 2012). These findings are in line with the observation that Set4, the *Saccharomyces cerevisiae* ortholog of MLL5/UpSET (Figure 4), is also part of a HDAC complex (Pijnappel et al., 2001). Set4 has been shown to play a role in the protection against oxidative stress upon which it interacts with chromatin and directly localizes to stress response genes. Set4 deficient cells showed elevated levels of reactive oxygen species (ROS) upon exposure to H_2O_2. These levels were decreased under Set4 overexpression. Interestingly, the SET domain of Set4 also does not have any HMT activity (Tran and Green, 2019; Tran et al., 2018). The elevated ROS levels observed in Set4 deficient yeast cells are of special interest as the same phenotype is observed in mice lacking *Mll5*, which will be discussed later on.

Although similarly convincing experimental data regarding the functional activities of MLL5 are not yet available, considerations of evolutionary conservation would suggest that MLL5 may be part of a mammalian HDAC complex, with analogous epigenetic functions in fine-tuning gene regulation (Mas-y-Mas et al., 2016). Thus, it can be assumed that MLL5 and its orthologues might act as transcriptional regulators through deacetylation rather than methylation. In line with this view, the loss of SETD5 in mice (i.e. the paralog of MLL5) increases levels of histone acetylation and thereby causes dysregulation of the transcriptome (Zhang et al., 2017).

1.5 The phenotype of *Mll5*-deficient mice shows striking resemblance to mouse mutants lacking the DNA repair and tumor suppressor protein ATM

Back to the phenotype of *Mll5*-deficient mouse mutants. Due to a potential, hypothesized involvement of MLL5-deficiency in leukemogenesis (see sections 1.2 and 1.3) rigorous analysis of hematopoietic compartments in *Mll5* knock-out (KO) mice seemed of particular importance. This analysis revealed a number of remarkable defects. For instance, in *Mll5*$^{-/-}$ mice the absolute number of thymocytes and splenocytes was reduced approximately 3-fold due to impaired lymphopoiesis. Interestingly, heterozygous *Mll5*$^{+/-}$ mice had intermediated numbers of thymocytes and splenocytes, indicating that *Mll5* haplo-insufficiency (as seen in del(7q22)-associated human leukemic cells) does have phenotypic consequences (Madan et al., 2009). This is of particular importance since the observed loss of *MLL5* in -7/del(7q) patients also occurs only heterozygously. A phenotype in heterozygous *Mll5* KO mice indicates the possibility of a MLL5-related phenotype in these patients as well.

Notably, in addition to impaired lymphopoiesis, *Mll5*-deficient mice exhibited prominent functional defects within their HSPCs, the cell compartment that gives rise to leukemic cells. These defects became visible in the failed ability of MLL5-deficient HSPCs to efficiently reconstitute lethally irradiated recipient mice (Madan et al., 2009). Further analysis revealed numerical and cell cycle abnormalities within the lineage^{neg}Sca1^{pos}cKit^{pos} (LSK) population, which contains short- and long-term hematopoietic stem cells (ST-HSCs and LT-HSCs). For instance, the total number of LT-HSCs defined as the LSK^{neg}Flt3^{neg}CD34^{neg} subset was reduced 3 – 4 fold and the LT-HSCs exhibited excessive proliferation accompanied by a concomitant reduction in the number of quiescent cells, suggesting an important role of MLL5 in the maintenance of HSC quiescence (Madan et al., 2009). Hyper-proliferation is a hallmark of leukemic cells (Hanahan and Weinberg, 2000). Thus, this observed phenotype could be considered a precursor to leukemia development.

Finally, mice deficient for *Mll5* were sensitive to radiation. Growth retardation, impaired fertility, and defects in lymphopoiesis, as observed in *Mll5*$^{-/-}$ mice, represent a phenotype combination often associated with genetic defects in the DNA repair machinery (see for instance *H2ax*$^{-/-}$ (Bassing et al., 2003; Celeste et al., 2002), *p53pb*$^{-/-}$ (Ward et al., 2003) or *Atm*$^{-/-}$ mice (Barlow et al., 1996; Ito et al., 2004)). To reveal potential defects in the response to DNA damage, *Mll5*$^{-/-}$ mice were thus exposed to whole-body γ-irradiation (WBI). In KO mice, a radiation dose that was readily tolerated by wild-type mice led invariably to death within 4 weeks. Interestingly, irradiated *Mll5*-deficient mice could be rescued with WT bone marrow grafts, indicating that the animals died due to defects in the DNA-damage response specifically in the HSPC compartment, and not because of the sensitivity of other tissues to WBI. Further comprehensive studies revealed, that this radiation sensitivity arises from defects in DNA double-strand break

(DSB) repair (Tasdogan et al., 2016). In fact, cells in the LSK compartment, including LT-HSCs, were shown to carry substantial amounts of unrepaired DSB in their genomic DNA, when Mll5 was absent. Again particularly noteworthy, heterozygous *Mll5*-deficient mice exhibited intermediary levels of DNA damage, highlighting that heterozygous loss of *Mll5* (as seen in human del(7q22) leukemic cells) is sufficient to cause readily detectable defects in DNA repair (Tasdogan et al., 2016).

Notably virtually all knockout-associated features of *Mll5*-deficient mice described so far are also observed in mouse mutants lacking the prominent tumor suppressor ATM (Ataxia telangiectasia mutated). For instance, $Atm^{-/-}$ mice are infertile, growth retarded and highly radiation sensitive (Barlow et al., 1996). Furthermore, like *Mll5* mouse mutants, $Atm^{-/-}$ mice exhibit impaired lymphopoiesis and highly similar functional defects in their HSPC compartment, including stem cell hyper-proliferation and lack of reconstitution potential (Ito et al., 2004; Ito et al., 2006). In marked contrast to understudied MLL5, due to its known function as eminent tumor suppressor, ATM is one of the most thoroughly characterized proteins. Its best known physiological function is to maintain genome stability by orchestrating the response to DNA damage, especially DSBs (Menolfi and Zha, 2020; Paull, 2015). Likewise, *Mll5* mutant mice have marked defects in DSB repair, as alluded to above. Findings made in *Atm*-deficient mice have thus been considered as paradigm for investigations into molecular mechanisms underlying the phenotype of *Mll5* mouse mutants (Tasdogan et al., 2016). In $Atm^{-/-}$ mice, comprehensive studies revealed that elevated levels of ROS in combination with ROS-mediated induction of the cell senescence inducer $p16^{Ink4a}$ are critical players for the HSPC malfunction (Ito et al., 2004; Ito et al., 2006). Analogous analysis of homo- and heterozygous Mll5 mouse mutants unveiled a dose-dependent accumulation of DNA breaks and high levels of intracellular ROS in HSPCs, comparable to *Atm*-deficient mice (Tasdogan et al., 2016). Oral administration of antioxidants such as N-acetyl-L-cysteine (NAC) reversed the observed defects in HSPCs both in *Atm* (Ito et al., 2004) and *Mll5* (Tasdogan et al., 2016) mutant mice, further supporting the role of increased ROS levels as drivers for HSPC malfunction.

Inefficient DNA damage repair and high intracellular levels of mutagenic ROS are known drivers of cancers, including leukemias, as both can trigger a variety of dangerous secondary mutations (Aparicio et al., 2014). In line with this notion, most $Atm^{-/-}$ mice develop malignant thymic lymphomas between 2 to 4 months of age, also $Atm^{+/-}$ heterozygotes are prone to tumorigenesis (Barlow et al., 1996). The finding that even heterozygous loss of *Mll5* in mice fuels accumulated DNA damage and increased levels of mutagenic ROS in cells of the HSPC compartment, i.e. in exactly those cells that give rise to leukemia, is so far the strongest experimental evidence for a causative role of heterozygous *MLL5*-deficiency in leukemogenesis.

However, despite these highly indicative phenotypes and in contrast to *Atm*-deficient mice, none of the three independently generated *Mll5* knock-out mouse strains exhibited any kind of leukemia or other tumors (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009). This result documents that the loss of *Mll5* alleles alone is not sufficient to transform cells, potentially because it requires additional genetic lesions. The hematopoietic phenotype of *Mll5*-deficient mice is thus in line with the previously described two-step theory of leukemogenesis (Section 1.1). The applicability of this theory to other MLL genes could already be proven for *MLL3* (Section 1.3, (Inaba et al., 2018)).

1.6 Identification of potentially cooperating leukemogenic gene mutations associated with heterozygous loss of MLL5

In an effort to identify specific mutations in one or more of the approximately 20,000 human genes, which may cooperate with heterozygous loss of *MLL5* in leukemia, the research group of Prof. Dr. Konstanze Döhner (Internal Medicine III, University Hospital Ulm) performed NGS of whole exomes (WES) from leukemic cells of 8 AML patients with 7q deletions and non-complex karyotype (2017/2018, unpublished information). Importantly, all patients had 7q deletions in leukemic cells that included the loss of *MLL5*. Table 1 provides a list of all genes that were found to have non-synonymous mutations in individual patients.

Table 1: Gene mutations identified by WES in leukemic cells from AML patients with 7q deletions, including loss of MLL5

Given are the patient ID, gene symbols of the mutated genes, and the total number of mutations found in the respective patient. A list with the full names of all genes is given in the appendix on page 127.

ID	Mutated genes	Total
331	CRTAC1, CSMD1, GAS6, IARS, METTL3, MGA, OR5R1, RNF17, RUNX1, SNAP91, STAG2, TP53, YEATS2	13
335	ADD1, ATP10D, BCOR, CCDC168, CHRNB1, H3F3A, IDH2, OCSTAMP, REPS2, RIPPLY3, SLC6A13, ZNF497	12
1110	ALS2CL, ARL4D, ASXL1, ATP13A2, BRCA2, CACNA1I, CBL, CCM2, CCND1, CCT6B, CD22, CNOT3, COL5A2, CUL3, DYNC1H1, E2F1, EIF4G1, FAT1, GAPVD1, GDF3, GMEB1GSDMA, LILRB4, MAPK8IP3, MDGA1, MS4A7, MYL9, NAALADL1, NRAP, NRG3, PCDH9, PIK3C2B, PIWIL2, PPWD1, PTPN11, RTN3, RUNX1, RYR1, SEC24C, SLC39A8, SLCO1C1, SNTB2, SPATA22, TAF1C, TCF20, TDRD3, TENM3, TRO, TRPM1, UNC5A, USP29, WBSCR28	53
2091	BCL9, BCOR, BRCA2, BUB3, CDH23, CROT, CTBP2, CYHR1, DDX47, DNMT3A, GCNT2, IDH2, MED12, MIB2, OR6C3, PPFIA1, PTPN11, RPL24, SHROOM2, TMEM35ULK4, VSTM2A,	22
2314	ABCA12, AGAP2, AHNAK2, DNAH3, FST, GNAQ, GPR21, GRID2IP, GSE1, HTR3C, IDH2, MAPRE1, NBEAL1, NOVA2, TEK, ZSCAN32	16
2599	ARID1B, DNMT3A, H3F3A, IDH2, KCTD16, MEGF6, RAD50, RUNX1, SMIM2	9
3500	BNIP3L, CSF3R, DHX57, IDH2, IGFN1, LCAT, POLR2H, RANBP3, RBM45, SRSF2, TUB	11
2717	ARV1, CRTC1, F8, FANCM, FLT3, HS6ST3, PAPD4, PNLIPRP2, PTPRD, RBL1, RBM26, RYR1, ZNF681	13

Based on the assumption that truly cooperating gene mutations should occur more frequently in the context of 7q-deficiency than pathologically less relevant mutations, we excluded all genes appearing just once in our list from further studies. The remaining 8 genes that were found to carry non-synonymous mutations in more than one patient are listed in Table 2 along with more information on the affected amino acids and the frequency of the respective mutations in AML as reported in the literature.

Table 2: Frequency of recurrent gene mutations in the 7q-deleted patient pool
Only gene mutations in exons that were found at least twice in our patient pool are listed. The estimated frequency of the mutation in AML given in the table is based on: [1] (Issa and DiNardo, 2021), [2] (Kunadt et al., 2022), [3] (Chaudry and Chevassut, 2017), [4] (Simon et al., 2020), [5] (Boileau et al., 2019b), [6] (Zhang et al., 2020), [7] (Iqbal et al., 2016), [8] (Stasik et al., 2021), [9] (Vosberg et al., 2018). Specific the amino acids are named by one letter code.

Mutated gene	Reported frequency of mutation in AML	Reported type of mutation in AML	Type of mutation in our cohort
5 x IDH2	12% [1]	R140 77 % [2] R172 19 % [2] other R-residues 4 % [2]	1x R140Q (20 %) 4x R172K (80 %)
3 x RUNX1	~ 10 % [4]	Loss-of-function	1x L144Q 1x frameshift T312 1x frameshift A338
2 x DNMT3A	22 % of *de novo* AML, 36 % of CN-AML [3]	R882 60 % (mostly R882H) [3]	1x R882H 1x E774G
2 x H3F3A	Total H3 1.6 % in *de novo* AML, 9% in secondary AML [5]	Q69H, A26P, R2Q, R8H and K27M/I [5]	2x K27M
2 x BCOR	3.8 – 5 % [6]		
2 x BRCA2	Described in secondary AML after breast cancer treatment [7]		
2 x PTPN11	~ 7 % [8]		
2 x RYR1	Sporadic cases described [9]		

Whether any of the identified gene mutations truly cooperates with heterozygous loss of *MLL5* in leukemogenesis needs to be tested in appropriate mouse models. Within the time-frame of a PhD thesis, such functional evaluation of all mutations listed in Table 2 seemed unrealistic. We thus decided to focus on **(i)** IDH2-R172K, **(ii)** DNMT3A-R882H or other DNMT3A inactivating mutations, and **(iii)** H3F3A-K27M based on considerations explained in detail for each selected mutation in the following sections. Interestingly, almost half of the -7/del(7q) AML patients whose exomes were sequenced (3 out of 8) carry combinations of the selected mutations (see Table 3) which might be of particular pathophysiological significance.

Table 3: Simultaneous occurrence of project-relevant mutations in individual patients form our del7q cohort
In three out of eight patients from our del(7q) cohort two of the genes selected for further study are simultaneously mutated.

ID	Project Relevant Mutations
335	IDH2-R172K, H3F3A-K27M
2091	IDH2-R172K, DNMT3A-R882H
2599	IDH2-R172K, DNMT3A-E477G, H3F3A-K27M

1.6.1 Isocitrate dehydrogenase 2 (*IDH2*)

Isocitrate dehydrogenase 2 (IDH2) is a metabolically active enzyme residing within the mitochondrion, where it catalyzes the decarboxylation of isocitrate to α-ketoglutarate (α-KG) as part of the Krebs-cycle. Recurrent IDH2 mutations have been identified in about 12 % of AML patients (Issa and DiNardo, 2021), but also in angioimmunoblastic T-cell lymphoma (AITL) (Lemonnier et al., 2016) as well as in solid cancers such as gliomas and certain cancers of the bone (Guo et al., 2021). The observed pathogenic mutations are always heterozygous and mainly limited to one of two conserved arginine (R) residues located within the enzymatically active sites of IDH2, i.e. R140 and R172. Each of these amino acid mutations endows IDH2 with a new, "neomorphic" activity, converting α-KG to the oncometabolite 2-Hydroxyglutarate (2-HG) (Issa and DiNardo, 2021). 2-HG accumulates in IDH2-mutant leukemic cells and competitively inhibits α-KG dependent dioxygenases such as lysine demethylases and DNA demethylating TET proteins, leading to global histone and DNA hypermethylation (Figueroa et al., 2010; Issa and DiNardo, 2021; Kunadt et al., 2022). Obligatory presence of a wild-type allele of IDH2 in heterozygous mutant cells ensures continued generation of α-KG, the normal product of IDH2 activity and the substrate of mutant IDH2 that is then reduced to the oncometabolite 2-HG. A simplified schematic representation of the pathogenic process is shown in Figure 5.

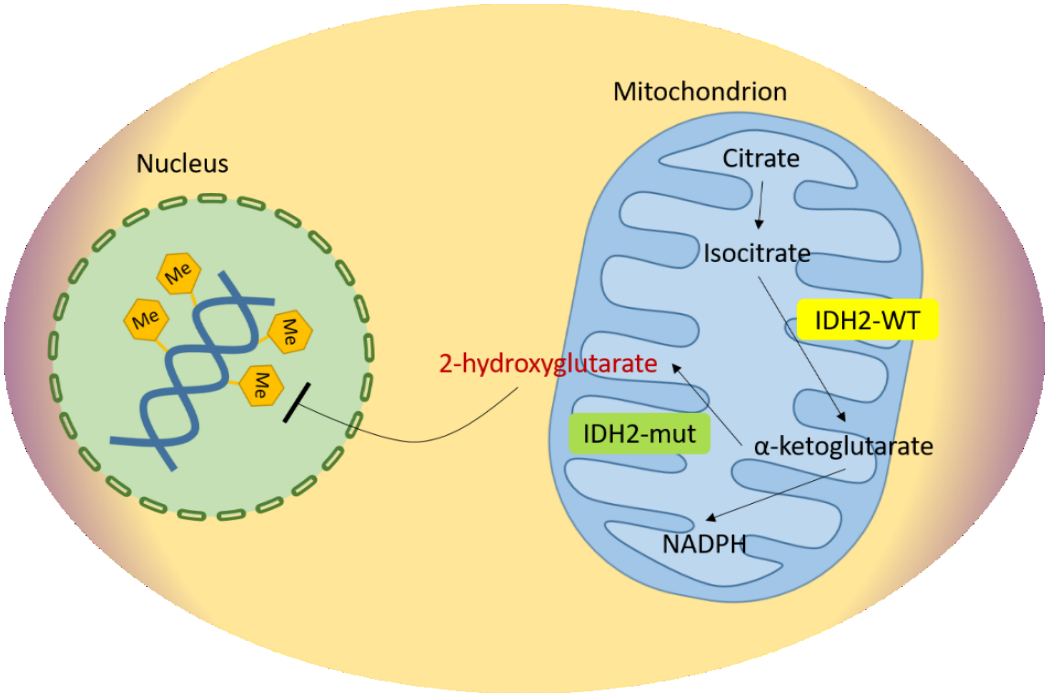


Figure 5: Schematic representation of physiologic and pathophysiologic IDH2 activities within the cell
Heterozygous IDH2-R140K or -R172K mutations result in the production of the oncometabolite 2-hydroxyglutarate (2-HG). 2-HG blocks demethylation of histones and/or DNA and thereby causes epigenetic changes. Me = Covalently attached methyl groups on lysine residues of histones or cytosines in DNA; mut = mutated; WT = wild-type; NADPH = Nicotinamide Adenine Dinucleotide Phosphate

Mutant IDH2-associated hyper methylation of histones and/or DNA leads to a block in HSPC differentiation (Figueroa et al., 2010; Kats et al., 2014; Kunadt et al., 2022). In line with the two-step theory of leukemogenesis, mutations resulting in a block in differentiation are considered

class II mutations, meaning an additional mutation within a class I gene would be sufficient to cause onset of leukemia (Kats et al., 2014; Takahashi, 2011).

In an unselected AML patient pool, mutations at *IDH2-R140* are four times more frequent than those at *IDH2-R172* (Green et al., 2011), with corresponding frequencies of ~ 10% and ~ 2.5%, respectively. The molecular basis for this intriguing dichotomy is still unknown, as both mutations seem to confer exactly the same pathogenic effect: acquisition of a neomorphic IDH2 activity that results in the production of the oncometabolite 2-HG. However, differences in the distribution of these two hotspot mutations among distinct malignancies suggest that they may have different functional consequences (Lemonnier et al., 2016; Meggendorfer et al., 2018). This is of particular interest in the context of our study, as curiously in our patient cohort R172 mutations are 4-fold more frequent than R140 mutations (Table 2), and thus exactly opposite to what is reported for AML in general.

Collectively, based on the above information, we decided to include the IDH2-R172K mutation in our study for the following reasons: **(i)** IDH2-R172K mutations are known to promote leukemogenesis by blocking HSPC differentiation. However, whether this mutation might cooperate with *Mll5*-deficiency has not yet been investigated. **(ii)** With 4 out of 8 patients affected, the IDH2-R172K mutation is by far the most frequent genetic alteration detected in our small patient cohort. A frequency of 50% dramatically exceeds the approximately 2.5% of R172K mutations expected for AML cases in general (for calculation, see above). **(iii)** Strikingly, in our del(7q) selected patient cohort, the reported ratio for R140/R172 mutations of 4:1 is neatly reversed with R172 being the 4x more frequently observed *IDH2* mutation. The high frequency of R172K mutations in our patient cohort together with the reversed ratio of R140/R172 mutations is highly suspicious and might hint to a potential cooperating relationship between -7/del(7q) and *IDH2-R172*. **(iv)** Finally, by giving rise to a neomorphic activity, the IDH2-R172 mutation confers a dominant effect, which provides a substantial technical advantage for experimentation. Due to the dominant effect of the mutation, the influence of *IDH2-R172K* mutations can be readily assessed via retroviral transduction of *Mll5*-deficient BM cells and subsequent transplantation of these cells into recipient mice. If an effect is observed, the less frequent IDH2-R140 mutation (Table 2) could easily be tested as well, with the intention to unveil potential functional differences between both IDH2 mutations.

1.6.2 DNA methyltransferase 3A (*DNMT3A*)

DNA methyltransferase 3A (DNMT3A) catalyzes the transfer of methyl groups to the C5 position of cytosine in CpG dinucleotides of DNA, an enzymatic reaction termed DNA methylation. Notably, DNMT3A is responsible exclusively for *de novo* DNA methylation, not for the maintenance of already existing DNA methylation marks during replication, which is taken care of by a different, homologous enzyme (Brunetti et al., 2017). CpG methylation provides an important epigenetic mark that profoundly effects gene expression typically by recruiting proteins involved in gene repression or by inhibiting the binding of transcription factors to DNA. DNA methylation marks are thus mostly associated with gene silencing.

Aberrant DNA methylation plays a prominent role in various cancers including leukemia (Brunetti et al., 2017). In AML, *DNMT3A* is one of the most frequently mutated genes. Inactivating mutations are found in ~37% of AML patients with a normal karyotype and

approximately 25% of all AML cases (Ley et al., 2010). These mutations are almost always heterozygous and are typically associated with high blast counts and poor prognosis (Ribeiro et al., 2012). Interestingly, about 60% of cases are heterozygous missense mutations at amino acid position R882 located within the catalytic domain (Gaidzik et al., 2013; Ley et al., 2010). Extensive studies of the R882H mutation, the prevalent R882 substitution, which is also found in one patient of our del(7q) cohort (Table 2), has revealed that this mutation gives rise to a hypomorphic protein with a residual activity of just 20% (Russler-Germain et al., 2014). Moreover, as the active enzyme consists of homo-dimers and homo-tetramers, DNMT3A-R882H additionally exerts a dominant-negative effect on the wild-type protein expressed from the non-mutated allele (Russler-Germain et al., 2014). Of note, distinct alterations in DNA methylation patterns in patients with heterozygous loss of DNMT3A function versus patients carrying dominant-negative DNMT3A-R882 mutations may indicate that both types of *DNMT3A* mutation possibly contribute to leukemogenesis via different mechanisms (Cole et al., 2017).

Both straight-forward *Dnmt3a* deletions (Celik et al., 2015; Challen et al., 2011; Mayle et al., 2015) as well as targeted Dnmt3a-R882 point mutations (Guryanova et al., 2016; Koya et al., 2016; Xu et al., 2014) have been extensively studied in various mouse models, with complex results. In essence, inactivation of Dnmt3a within the hematopoietic compartment was generally not sufficient on its own to induce overt leukemia. However, mutation-associated reduction of Dnmt3a activity resulted in a preleukemic state of HSPC expansion. Overt leukemia could be induced upon introduction of cooperating gene mutations (Guryanova et al., 2016; Meyer et al., 2016; Scourzic et al., 2016; Yang et al., 2016) or experimental manipulations resulting in proliferative stress, like non-competitive bone marrow transplantation (Celik et al., 2015; Mayle et al., 2015). Taken together, studies in mouse models support the view that *DNMT3A/Dnmt3a* acts as a haploinsufficient tumor suppressor in myeloid leukemias when cooperating mutations are present (Cole et al., 2017). This is in line with the observation that DNMT3A mutations in humans are strongly associated with clonal hematopoiesis (Brunetti et al., 2017) i.e. a dramatic expansion of individual hematopoietic cell clones. In fact, *DNMT3A* is by far the most frequently mutated gene in age-associated human clonal hematopoiesis, a condition considered to predispose to myeloid leukemia (Jaiswal and Ebert, 2019; Koeffler and Leong, 2017).

We have decided to include DNMT3A mutations in our study for the following reasons: **(i)** Two patients in our del(7q)i AML cohort feature DNMT3A mutations (Table 2). Patient #2091 carries the prevalent, well-characterized R882H substitution, and patient #2599 a less frequently observed E774G substitution. This latter mutation is predicted to reside at the interface of DNMT3A homodimers and thus may impair or prevent oligomerization, which is required for full enzymatic activity (Holz-Schietinger et al., 2011). Notably, both patients with DNMT3A mutations also carry the leukemogenic IDH2-R172K substitution (Table 3). **(ii)** DNMT3A is one of the most frequently mutated genes in AML. Comprehensive experimental data, as alluded to above, strongly suggest that these mutations are AML initiating events that require additional cooperating genetic aberrations for leukemogenesis. Importantly, whether DNMT3A mutations might cooperate with heterozygous loss of *Mll5* has not yet been investigated. **(iii)** The dominant-negative R882H mutation can be readily mimicked by retroviral transduction and subsequent transplantation of *Mll5*-deficient HSPCs. Moreover, prior to the start of this project, *Dnmt3a*$^{flox/flox}$ mice had been introduced into our mouse facility. By intercrossing these mice

with our vCre x *Mll5* $^{flox/flox}$ and/or vCre x H3K27M mice (see below) double- and triple-mutant mouse lines can be generated to comprehensively assess cooperativity between candidate mutant genes. Since the R882H mutation and deletional DNMT3A inactivation likely contribute via distinct mechanisms to leukemia, as pointed out above, ideally both types of mutation should be assessed in parallel.

1.6.3 Histone 3 family 3A (*H3F3A*)

Histones are the basic units of the nucleosome making them important players in the packaging of DNA and the regulation of epigenetic states. The nucleosome is made up of five core histones: H2A, H2B, H3, H4, and the stabilizing linker histone H1 as shown in Figure 6 (Martire and Banaszynski, 2020). Mammals have three classical and several non-classical isoforms of histone H3, that only differ in a few amino acids. The classical variants H3.1 and H3.2 are replication-dependent and are deposited onto newly replicated DNA (Martire and Banaszynski, 2020). H3.3 is a so-called replacement variant. It is expressed in all cell types throughout the cell cycle and replaces other H3 variants mainly at active gene loci in a highly conserved, replication-independent manner, contributing to a more open chromatin configuration. Curiously, H3.3 is also specifically deposited at repetitive heterochromatic regions, like telomeric and centromeric chromosomal sites, and at endogenous retroviral elements (Martire and Banaszynski, 2020; Shi et al., 2017; Szenker et al., 2011).

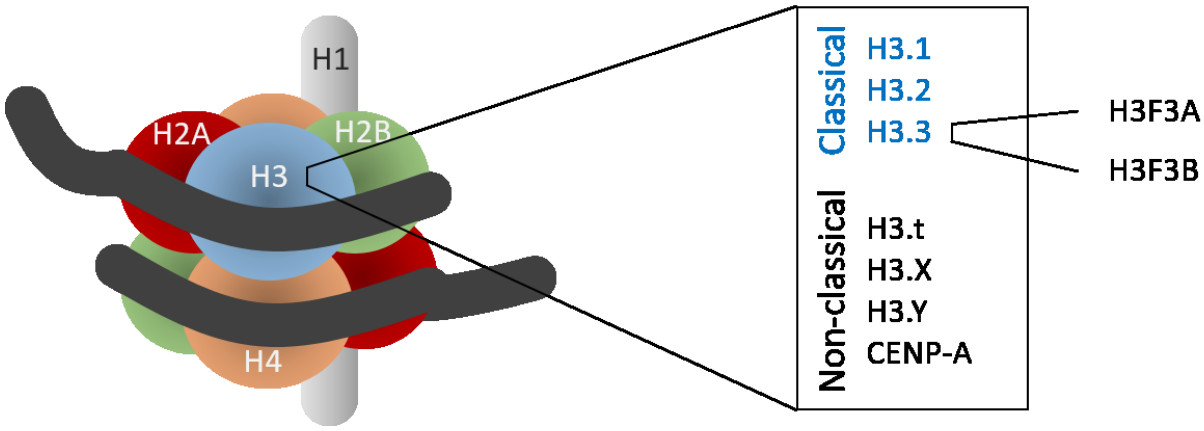


Figure 6: Schematic presentation of the nucleosome with a focus on histone 3 variants
Histone 3 (H3) is one of the core histones making up the nucleosome. There are three classical (H3.1, H3.2, H3.3) and several non-classical (e.g. H3.t, H3.X, H3.Y, CENP-A) variants. H3.3 is encoded by two unrelated genes (*H3F3A* and *H3F3B*) located on different chromosomes.

Strikingly, in our small cohort of eight -7/del(7q) AML patients, two individuals (#335 and #2599) turned out to be heterozygous for the same lysine (K) to methionine (M) amino acid substitution in H3.3 at position 27 (Table 1 -Table 3). Of note, K27 is a functionally important position in all H3 variants, including H3.3: Histones are subject to posttranslational modifications (PTM) such as acetylation, phosphorylation, ubiquitination, and methylation, which are involved in regulating gene expression (Millán-Zambrano et al., 2022). Many functionally important PTMs occur at specific K residues of histones. In all histone H3 variants, including H3.3, typical target sites are K4, K9, K27, and K36 (Millán-Zambrano et al., 2022). The observed K27M amino acid exchange would thus prevent all PTMs at position 27, in particular K27 actetylation and K27

methylation, which serve critical gene regulatory functions (Martire et al., 2019; Wiles and Selker, 2017).

Of note, at the time we identified this mutation in our patient cohort, there were no published records of H3K27 mutations in any case of leukemia. However, mutations in histones were well known to be associated with a variety of tumors. In fact, the realizations that mutant histones can be oncogenic and promote tumorigenesis led to creating the term “oncohistones” for the respective mutant protein version (Funato and Tabar, 2018; Mohammad and Helin, 2017). H3.3 oncohistones have first been described in 2012 in the context of pediatric glioblastomas, where a stunning 31 % of tumors showed a mutation either at H3.3A-G34 or H3.3A-K27 (Schwartzentruber et al., 2012) – exactly the amino acid position (K27) mutated in two of our patients. Meanwhile, a number of high-profile publications have confirmed this initial report and shown that H3.3A-K27 mutations, in particular K27M, play a key role as drivers of oncogenesis in childhood glioblastoma (Bender et al., 2013; Chan et al., 2013; Harutyunyan et al., 2019; Larson et al., 2019; Pathania et al., 2017; Sturm et al., 2012; Wu et al., 2012). One study even linked MLL5 and H3.3 by suggesting antagonistic roles of both mutations in adult glioblastomas that led to chromatin reorganization favoring self-renewal properties of cancer cells (Gallo et al., 2015); commented on in (Becher and Holland, 2015). Tumor-promoting mutations in histone variant H3.3 have also been detected at high frequencies in some types of bone cancer, albeit predominantly at position G24, which notably can influence the methylation pattern at K27. In summary, mutations in histone variant H3.3, including position K27, are obviously well established as promotors of oncogenesis (Kallappagoudar et al., 2015; Kumar and Pai, 2021; Yuen and Knoepfler, 2013), although notably not (yet) in leukemia.

Substantial experimental information is available on how mutations at H3.3-K27 may contribute to oncogenic transformation at the molecular level. The polycomb repressive complex 2 (PRC2) is the highly conserved HMT specific for H3K27. The enhancer of zeste homolog 2 (EZH2) is the SET domain containing catalytic subunit of the PRC2, responsible for the methylation of H3K27 (Simon and Lange, 2008). Trimethylation of H3K27 (H3K27me3) is a mark for inactive transcription (Millán-Zambrano et al., 2022). In WT cells, H3K27me3 results in an auto feedback loop allosterically activating the PRC2 complex and thereby propagating the repressive state (Deshmukh et al., 2022). Conversely, the aberrant H3K27M occupies the SET domain of EZH2 thereby inhibiting the PRC2 complex. This in turn causes a drastic reduction of global H3K27 di- and trimethylation (H3K27me2 and H3K27me3, respectively), i.e. not only at the mutated H3.3-K27M residues, but at H3K27 residues in the mutant bearing cell in general (Lewis and Allis, 2013; Lewis et al., 2013; Stafford et al., 2018). Globally reduced H3K27 methylation then gives rise to significant alterations in gene expression, which may drive oncogenesis. Independent support for the view that H3.3-K27M-mediated inhibition of EZH2 might directly contribute to leukemia comes from the fact, that inactivating mutations of EZH2 are common genetic aberrancies in myeloid neoplasms, including AML and MDS (Lund et al., 2014; Mechaal et al., 2019; Rinke et al., 2020).

Back to histone variant H3.3. Interestingly, this variant is encoded by two separate genes, namely *H3F3A* and *H3F3B*, which exhibit very different exon/intron structures and even different codon usage. However, the protein products of both genes are absolutely identical and thus cannot be readily distinguished from each other experimentally (Martire and Banaszynski, 2020). This can pose challenging analytical problems during experimental

approaches that require quantification of protein expression selectively from one of the two genes, for instance when trying to determine the potential effect of a mutation confined to *H3f3a* on its protein expression level, as in this Ph.D. project (e.g. see results 3.8).

Why did we select the H3F3A-K27M mutation for further study and decided to make the extra effort of generating a correspondent, Cre-inducible knock-in mouse mutant? At the start of this project, the *H3F3A-K27M* mutation had not been described in the context of leukemia. Hence, no information on a potential role in leukemogenesis was available. One possible reason for the lack of reports about *H3F3A-K27M* in the context of leukemia might arise from the fact, that in an unselected patient pool there was no accumulation of this mutation detected. In our case, the analysis was restricted to patients with 7q deletions. Here, the occurrence of exactly the same mutation at position K27 in two out of eight patients possibly indicates a particular association with heterozygous loss of genes in 7q, possibly *MLL5*. This *ad hoc* hypothesis received strong support, when we got access to exome sequencing data from another nineteen -7/del(7q) AML patients. Among these data generated in the laboratory of Prof. Dr. Christoph Plass (DKFZ Heidelberg), we identified again a H3F3A-K27 mutation, this time K27Q (Table 4; unpublished). Remarkably, this particular patient also carried the oncogenic IDH2-R172K mutations, like our two previous H3F3A-K27M patients (Table 3). We thus had identified in a total of 27 -7/del(7q) AML patients three times an H3F3A-K27 mutation that had not been reported before in the literature in the context of leukemia. This striking, novel observation seemed to merit experimental investigation. As the H3F3A-K27 mutation seemed to become conspicuous only when the patient pool was limited to -7/del(7q), we felt it would be worth-while to investigate the connection between MLL5 loss and H3K27M in the course of this project. However, at the time, no transgenic mouse model mimicking the H3F3A-K27M mutation was available. Thus the generation and characterization of an inducible *H3f3a-K27M* mutant mouse became a fundamental part of this project. In light of the contribution of H3.3-K27 mutations to brain tumors, as mentioned above, such mice should also provide an invaluable tool for investigators focusing on glioblastoma research.

Table 4: Somatic gene mutations in another del(7q) AML patient with a H3.3-K27 substitution identified by WES in the laboratory of Prof. Dr. Christoph Plass (DKFZ Heidelberg)
The identified mutation in *H3F3A* affects the same residue (K27) as in the patient cohort from Table 1, but the substitution involves a different amino acid: K27Q instead of K27M. The patient additionally carries the previously described *IDH2-R172K* mutation. Full names of genes are given on page 127.

ID	Gene Mutations	Total
A87Q	ANK3, APOBR, AP001055.1, BCOR, BCORL1, CEACAM5, H3F3A-K27Q, FAM9A, FAM174B, FLG2, GTF2A, IDH2-R172K, IFT81, KDM2B, LIPT1, LY6G6F, MAPRE1, NR0B1, PEX11G, PROKR2, PXDN, RIMS2, SCNN1B, TACC2, TRIM51	25

1.7 Experimentally mimicing leukemogenesis in mice

Proper experimental setups mimicking AML are crucial in understanding the disease mechanisms and investigating novel treatment options. However, the underlying genomic and molecular complexity of AML poses one of the major challenges during the development of suitable disease models, as few malignant early hematopoietic stem cells residing within the BM are able to initiate and sustain leukemic transformation (Banerjee et al., 2010).

Hematopoiesis is a tightly regulated process influenced by biochemical and physiochemical signals of the BM microenvironment and the corresponding niches (Lilly et al., 2011). In an attempt to experimentally mimic human AML, several *in vitro* and *in vivo* approaches have been developed over the past decades. The simplest models of AML are based on 2D and scaffold-free 3D cell culture of leukemia cell lines or patient derived primary cells. While this provides several advantages (e.g. easy-to-use, cost efficient, controllable, and high-throughput), they fail to replicate the complex and dynamic environment of AML (Dozzo et al., 2022). *In vitro* models have further been improved by the introduction of matrix, or scaffold-based 3D culture, where leukemic cells grow on biomaterial in a highly controllable manner. These models allow for a more accurate replication of pathophysiological mechanisms but still fail to fully replicate the BM microenvironment and stem cell niches, partially due to the lack of complete characterization of these niches (Bello et al., 2018).

Despite advances in 3D cell culture systems (i.e. BM on-chip technology (Sieber et al., 2018)), the use of murine *in vivo* models remains one of the most important pillars during leukemia research (Keinan et al., 2021; Kohnken et al., 2017). Apart from practical reasons including their small size, short reproductive cycle, and short life span, mice share several genetically and physiologically similarities with humans, thus providing a complex whole organ system to study a variety of diseases, including leukemia, and to develop and validate novel therapeutics (Kohnken et al., 2017; Sharpless and DePinho, 2006).

Mouse models in leukemia research can be divided into four major classes: **(i)** carcinogen induced, **(ii)** adoptive transfer of retrovirally transduced BM cells, **(iii)** transgenic mice, and **(iv)** xenograft models (Figure 7). When choosing a suitable mouse model, one needs to consider the disease etiology, including the affected cell type and the mutation(s) of interest. As briefly mentioned in the previous chapter, the physiological effect of the mutations selected for this project will be analyzed via retroviral transduction followed by adoptive BM transfer and via the generation of transgenic mice. Carcinogen induced and xenograft models are not of relevance during this project and therefore will not be further discussed.

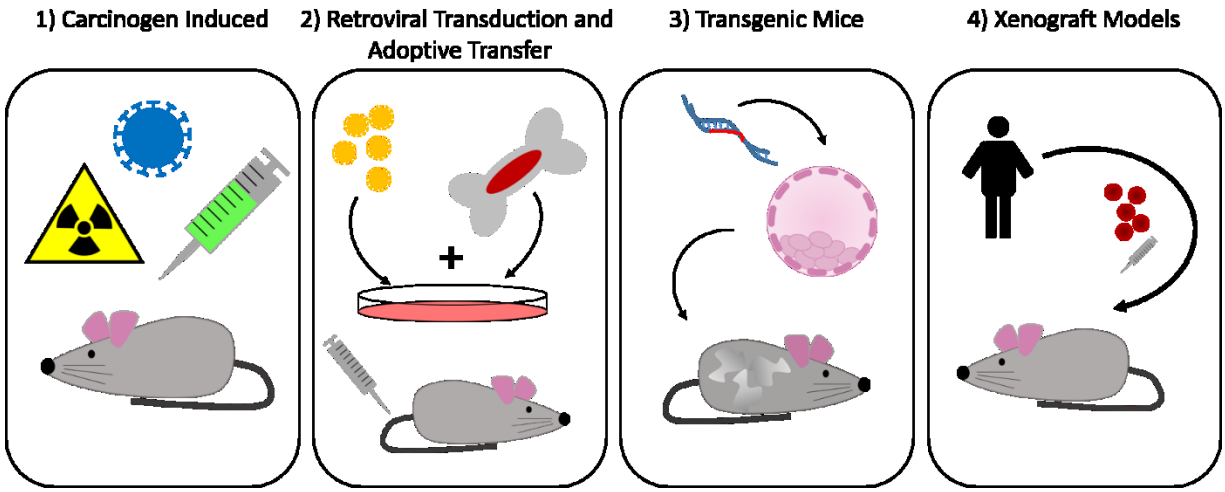


Figure 7: Strategies to mimic leukemia development in murine models
Murine models used in leukemia research: 1) Tumors are induced with carcinogens including radioactive irradiation, murine leukemia viruses or chemicals. 2) Adoptive transfer: Bone marrow is harvested from donor mice, leukemia associated mutations are introduced retrovirally or by gene editing, the bone marrow is then transplanted into lethally irradiated recipient mice. 3) Transgenic mice: mice expressing the potentially leukemia associated gene are generated via methods of gene targeting. Expression of the transgene can be constitutive or conditional upon enzymatic activation. 4) Xenograft models are generated by transferring patient cells into immunocompromised mice. Figure based on (Almosailleakh and Schwaller, 2019).

First, some information on murine models generated from retroviral transduction of BM cells and subsequent adoptive cell transfer. In brief, BM cells are harvested from donor mice and transfected with genetically modified retroviruses that carry the mutated gene of interest. Upon introduction of the virus into the host cell, its genome is transcribed into dsDNA that is integrated as a provirus into the host cell genome, leading to permanent expression of the introduced gene together with the endogenous gene loci (Ciuculescu et al., 2014). The transfected BM cells are then transplanted into lethally irradiated congenic recipient mice to monitor *in vivo* transformation ability of the potential oncogene. When choosing the retroviral vector backbone, the cell type (i.e. the sub population responsible for disease initiation) that should be transfected needs to be taken into account. Since AML is driven by HSPCs, we chose an MSCV (Murine stem cell virus) based vector carrying a promotor specific for hematopoietic cells (Hawley et al., 1994). *In vivo* expression of the potential oncogene can be monitored via a fluorescent marker, in our case mCherry (Shaner et al., 2004).

BM transfer models have one major advantage: they do not require complex breeding schemes to generate sophisticated transgenic mouse models. They provide the ability to quickly test new mutations by simply generating the required viral vector, also allowing for flexible combination of various oncogenic mutations. Nevertheless, leukemogenesis in the native immune system can still not fully be mimicked by adoptive transfer models (Gavrilescu and Van Etten, 2008).

Despite the availability of transgenic *Idh2-R172K* mutant mice (Akbay et al., 2014; Lemonnier et al., 2016), we still chose to evaluate IDH2-R172K using adoptive BM transfer because said mutation acts dominant over *Idh2-WT*, actually requiring remaining WT function to exhibit the oncogenic function, making IDH2 an ideal candidate for retroviral transduction experiments. The applicability of said method to analyze IDH2 mutations in the context of AML has been shown before (Chen et al., 2013; Ogawara et al., 2015). Furthermore, our lab generated

conditional *Mll5*$^{flox/flox}$ mice that can readily be used as BM donors without the time consuming need to intercross the aforementioned *Idh2* mutant mice.

Now let's talk about strategies for the generation of transgenic mice. One method to generate transgenic mice is the injection of the vector based gene of interest into oocytes. These oocytes are implanted into pseudo-pregnant mice resulting in random integration of the gene of interest, potentially leading to undesirable expression levels, lack of tissue specificity, and lethality (Almosailleakh and Schwaller, 2019; Dozzo et al., 2022). A more targeted approach is achieved by site directed manipulation of the transgene in mouse embryonic stem cells (in detail explained in methods section 2.2.8). Gene targeting can be used to generate knock-out (KO) mice, where the gene of interest is functionally inactivated, or knock-in (KI) mice introducing an altered version of the gene (i.e. introducing a mutation) (Rosenthal and Brown, 2007). The desired manipulation can either be constitutively active or conditionally activated upon introduction of recombinase activators.

One major drawback in using constitutively activated mouse models in AML is their inability to recapitulate the clinically more accurate, sporadic development of the disease (Dozzo et al., 2022). Furthermore, gene manipulation during development might be lethal or produce aberrant targets, eventually leading to inappropriate models of the human disease (Kohnken et al., 2017). To overcome these issues, conditional mouse models were developed. Here, mice carrying the target gene are intercrossed with recombinase effector mice, allowing for time and tissue specific expression of the transgene (Rosenthal and Brown, 2007). These recombinases, typically Cre-recombinase, recognize *loxP* sites and mediate deletion of DNA sequences located between two *loxP* sites (Nagy, 2000). To date, dozens of mouse strains with tissue specific Cre-recombinase are available. When choosing the appropriate Cre mouse, the evaluated disease should be taken into careful consideration. During this Ph.D. project we set out to generate a mouse model carrying the *H3F3A-K27M* mutation and to evaluate said mutation in the context of AML. Since AML is a hematopoietic disease, the Cre mouse used to knock-in the mutation should be active in the hematopoietic system. To this extend we had two different Cre mouse strains readily available in our mouse colony: vCre and CreERT2.

In vCre mice, the recombinase expression is under the vav promotor which is expressed in all adult nucleated cells of the hematopoietic tissue, but not in any non-hematopoietic cells (de Boer et al., 2003; Ogilvy et al., 1999). vCre is constitutively expressed as early as day 13.5 of embryonic development in hematopoietic cells of the fetal liver (McMahon et al., 2007). Of note, vCre transgenic mice cannot be bred to homozygosity as the Cre-encoding transgene has inserted and inactivated the *Commd10* gene locus on chromosome 18, one copy of which is required to ensure normal embryogenesis (de Boer et al., 2003; Goodwin et al., 2019). This needs to be taken into account during the design of breeding strategies and will be of relevance later in this project during the generation of the transgenic mice (see section 3.4).

In CreERT2 mice, a fusion protein consisting of Cre-recombinase and a mutated version of the estrogen receptor domain (ERT2) is inserted into the ubiquitous expressed *Rosa26* gene locus. Here, Cre is inactive in the absence of pharmacological induction. Oral administration of the estrogen receptor modulator tamoxifen activates Cre expression in all cells (Seibler et al., 2003). To limit expression of the transgene to the hematopoietic compartment, BM would have to be harvested from tamoxifen treated mice and transplanted into lethally irradiated recipient

mice. Despite the relatively higher experimental effort, this method would be closer to clinical reality since AML develops from few HSPCs that degenerate at a time point later in life.

In spite of a potential dominant effect of the *H3f3a-K27M* mutation (see chapter 1.6.3), and thereby the possibility to use adoptive BM transfer models as well, we still chose to generate a conditional transgenic mouse. Like all other histones, H3.3 is highly conserved, implying important functions. Additionally, position K27 is physiologically highly relevant during regulation of gene expression, as described in the previous chapter. Constitutive activation of the K27M mutation could lead to severe effects during embryogenesis, potentially even lethality. To overcome this problem, we chose to render the mutation inducible by applying the Cre-*loxP* system. During this project we are interested in a disease originating in HSCs, hence the aforementioned vCre mouse strain will be used to activate the mutation. Utilizing the Cre-*loxP* system provides another advantage: the same mouse model we set out to generate during this project, can be intercrossed with Cre-recombinase carrying mouse strains of different tissue specificity allowing the analysis of a completely different disease. By intercrossing with neural tissue specific Cre-strains (e.g. *Nestin*-Cre (Tronche et al., 1999)) the effect of *H3f3a-K27M* during glioblastoma could be analyzed.

1.8 Aim of the Project

Whole exome sequencing of DNA from leukemic cells of AML patients with heterozygous 7q-deletions (including the *MLL5* gene locus) led to the identification of *H3F3A-K27M*, *IDH2-R172K* and inactivating *DNMT3A* mutations as recurrent genetic aberrations that might cooperate with heterozygous *MLL5*-deficiency in leukemogenesis. The *H3F3A-K27M* mutation was of particular interest, as this mutation had not been reported previously to be associated with any form of leukemia. The main goal of this Ph.D. thesis was to establish appropriate mouse models, which would allow *in vivo* evaluation of potential pathophysiological consequences of these mutations with respect to leukemogenesis in the context of heterozygous *Mll5* deficiency. To this end, a novel knock-in mouse strain carrying a Cre-inducible K27M mutation of histone H3.3A had to be generated and comprehensively characterized. When generating this mouse strain, special efforts were made to allow non-invasive monitoring of individual cells expressing the mutant H3.3A-K27M protein. Thereafter, these mice were intercrossed with a conditional *Mll5*-deficient mouse strain previously generated by our group (Madan et al., 2009). In addition, methods for retroviral transduction and transplantation of hematopoietic stem and progenitor cells had to be established in the laboratory and optimized in order to introduce *Idh2-R172K*, *H3f3a-K27M* and dominant negative *Dnmt3a-R882H* mutations in *Mll5*-mutant mice. Mouse cohorts carrying individual and combined mutations were monitored for extended periods of time to unveil the potential cooperative effects of these mutations in the context of heterozygous *Mll5*-deficiency. Hence, the primary goal of this thesis was to establish a reliable experimental system that would allow in future experiments a comprehensive *in vivo* assessment of the combined effects of *H3f3a-K27M*, *Idh2-R172K,* and inactivating *Dnmt3a* mutations on leukemia induction in the context of *Mll5*-deficiency. Application of this experimental system in this thesis was supposed to provide a first set of *in vivo* data, which would reveal to what extent the newly diagnosed *H3F3A-K27M* mutation might contribute alone or in combination with *Mll5*-deficiency and/or *Idh2-R172K* to leukemogenesis.

2. Material and Methods

2.1 Material

All technical equipment, reagents, cell lines, and consumables used during the course of this project are listed in the following sections.

2.1.1 Technical equipment

Equipment	Manufacturer
Agarose Gel Electrophoresis Unit	Bio-Rad
Bacterial incubator	Heraeus
BD FACS Aria Ilu Cell Sorter	BD Biosciences
BD FACS Canto II Cell Analyzer	BD Biosciences
BD LSR Fortessa Cell Analyzer	BD Biosciences
Benchtop Centrifuge 5417R	Eppendorf
Bio Photometer	Eppendorf
Biofuge pico	Heraeus
Easy Sep Magnet	Stem Cell Technologies
Fusion-SL Advanced Imaging System	Vilber Lourmat
Gammacell Irradiation Unit (for cells)	Nuclear Data Inc.
Gel Documentation System	Biometra
Gene Pulser	Bio-Rad
GSR D1 Irradiator (for mice)	GSM (Gamma-Service Medical GmbH)
Hemavet 950	Drew Scientific
Heracell 240 CO_2 Incubator	Thermo Scientific
LaminAir HB 2460 (Laminar Flow Hood)	Heraeus Instruments
Light Box	LKB Wallac
Light Microscope	Zeiss
Megafuge 1.0 R	Heraeus Sepatec
Microliter Pipettes	Eppendorf
Microwave	Sharp
Mini-Protean Handcast 3 Cell System	Bio-Rad
Multifuge 3SR+	Heraeus Instruments
Multiskan FC	Thermo Scientific
Nanodrop Spectrophotometer	Implen
Nanophotometer	Serva
PCR Thermocycler	Biometra
pH Meter	Mettler Toledo
PowerPac HC High-current Power Supply	Bio-Rad
Precision Scale	Sartorius
ProgRes CS (CCD Research Camera)	Jenoptik
Rocking Shaker	Grant-Bio
Sorvall RC-5C Plus Superspeed Centrifuge	Sorvall
StepOne Plus Real-Time PCR System	Applied Biosystems

Table continued on next page.

Stereo Mikroscope	Nikon
Thermomixer	Eppendorf
Trans-Blot Turbo System	Bio-Rad
Tube Roller	Neo Lab
Vacuum Pump	Millipore
Vortexer	Heidolph
Water Bath	Grant

2.1.2 Consumables

Product	Reference	Supplier
24-Well Tissue Culture Plate	142485	Thermo Scientific
6-Well Plate (Nunclon Delta Surface)	140675	Thermo Scientific
96-Well Cell Culture Cluster Flat Bottom	3595	Costar
96-Well Cell Culture Cluster Round Bottom	3799	Costar
Bacteriological Dish (10 cm)	351029	Falcon
Cannula 26G x 1/2´´	4665457	Braun
Cell Strainer (40 µm, Nylon)	352340	Falcon
Cryogenic Vials (2 mL)	T309-2A	Simport Scientific
FACS Tubes (5 mL)	55.1579	Sarstedt
Gene Pulser Cuvettes (0.4 cm)	165-2088	Bio-Rad
Glass Pasteur Pipette (150 mm)	7477 15	Brand
Inoculation Loops	255L51	Copan
Inoculation Spreader	86.1569.001	Sarstedt
Insulin Syringes 29G (1 mL)	324827	BD Micro-Fine
Laboratory Film (Parafilm)	PM-996	Bemis
MicroAmp Fast 96-Well Reaction Plates (0.1 mL)	4344907	Applied Biosystems
MicroAmp Optical Adhesive Film	4311971	Applied Biosystems
Microcentrifuge Tubes (0.5 mL)	7060.1	Roth
Microcentrifuge Tubes (1.5 mL, 2 mL, 5 mL)	72.706.400 72.695.400 72.701.400	Sarstedt
Microcentrifuge Tubes, amber (1.5 mL)	0030 120.191	Eppendorf
Microscope Slides	03-0004	R. Langenbrinck
Microtitet Plates 96-Well U (for FACS)	701330	Brand
Microvette 200 tubes (EDTA)	20.1288	Sarstedt
Mr. Frosty	5100-0001	Nalgene
Neubauer Counting Chamber	6311120	Brand
Nylon Mesh Filter (70 µm)	03-150/38	Sefar
PCR tubes	781320	Brand
Petri Dish (3.5 cm)	351008	Falcon
Photometer Cuvettes	2712120	ratiolab

Table continued on next page.

Pipette Tips, filtered (0.1- 1000 µL)	70.1116.210 70.3030.265 70.3031.355 70.3050.355	Sarstedt
Serological Transfer Pipettes (5 mL, 10 mL, 25 mL)	86.1253.001 86.1254.001 86.1685.001	Sarstedt
Sterile Surgical Blades	BB536	Braun
Sterile Syringe Filter w/0.2 µm Polyethersulfone Membrane	514-0073	VWR
Syringes (1 mL)	303172	BD Plastipak
Syringes (5 mL, 10 mL)	4606108V 4606051V	Braun
Tissue Culture Dish, adherent (10 cm)	353003	Falcon
Tissue Culture Dish, adherent (15 cm)	168381	Thermo Scientific
Tissue Culture Flasks, suspension (25 cm^2)	353014	Falcon
Tissue Culture Flasks, suspension (75 cm^2)	83.3911.502	Sarstedt
Tubes, Conical (15 mL, 50 mL)	62.554.502 62.559.001	Sarstedt
Tubes, Round Bottom (13 mL)	62.515.006	Sarstedt
Vacuum Filtration Unit 0.2 µm (500 mL, 1000 mL)	83.3941.001 83.3942.001	Sarstedt
Whatman Filter Paper	10318486	Bio-Rad

2.1.3 Software

Software	**Version**	**Supplier**
Adobe Illustrator	v.27.0	Adobe Systems Incorporated
BD FACS Diva	v.9.0.1	BD Biosciences
BioDoc Analyze	v.2.1	Biometra
EndNote X9	X9.3.3	Clarivate Analytics
FlowJo	v.10.8.0	Tree Star Inc.
FusionCapt Advance	SL4	Vilber Lourmat
ImageJ	1.52a	National Institutes of Health (NIH)
Microsoft Office	2016	Microsoft Corporation
Primer Express	v.3.0.1	Applied Biosystems
Prism	v.8.0	Graphpad Software Inc.
PyRat	v.4.5.1-315	Scionics
SkanIt Software for Microplate Readers	v.7.0.0.50	Thermo Scientific
SnapGene	v.4.3.11	Dotmatics
StepOne Plus	v.2.3	Applied Biosystems

2.1.4 Reagents

Reagent	Reference Number	Supplier
1-Thio >= 97%	M6145	Sigma
2-Mercaptoethanol	31350-010	Gibco
3-[N-morpholino]propanesulfonic acid (MOPS)	M1254	Sigma
30 % Acrylamide/ Bis Solution	1610156	Bio-Rad
Acetic Acid	20104.334	VWR
Agarose	840004	Biozym
Albumin from Chicken Egg White	A5503	Sigma
Ammonium Persulfate (APS)	13376.02	Serva
Ampicillin	A5354	Sigma
Bacto Agar	214010	BD
Bacto Tryptone	211705	Gibco
Bacto Yeast Extract	212750	BD
Bovine Serum Albumin (BSA) 2% Solution	A1595	Merck
Bovine Serum Albumin, powder	A4503	Sigma
Bromphenol Blue	32768	Riedel-de-Haën
c0mplete Tablets Mini	04 693 159 001	Roche
Calcium Chloride (CaCl)	A119	Roth
Chloroform	32211	Sigma-Aldrich
Coomassie Brilliant Blue R-250	161-0400	Bio-Rad
Crystal Violet	101408	Merck
DeoxyribonucleotideTriphosphate (dNTP) Set 100 mM	10297-018	Invitrogen
Diethyl Pyrocarbonate (DEPC)	K028.1	Roth
Dimethyl Sulfoxide (DMSO)	D2650	Sigma
Disodium Hydrogen Phosphate (Na_2HPO_4)	30412	Riedel-de-Haën
DNA ladder, 1 kb plus	N3200	New England Biolabs
DNA ladder, 100 bp	N3231	New England Biolabs
DNA ladder, 50 bp	N3236	New England Biolabs
Dulbecco's Modified Eagle Medium (DMEM (1X) + GlutaMAX)	61965-026	Gibco
Ethanol	20821.330	VWR
Ethidium Bromide	17898	Invitrogen
Ethylenediaminetetraacetic acid (EDTA)	8043.2	Roth
Fetal Calf Serum (FCS)	F7524	Sigma
Formaldehyde	4979	Roth
Formamide	1.09684.100	Merck
Gel Loading Dye, Purple 6X	B7024	New England Biolabs
Gelatin (from Porcine Skin)	G1890	Sigma
Geneticin (G418)	11811-031	Gibco
Gentamicin	G1397	Sigma

Table continued on next page.

Glucose D(+), anhydrous	49150	Fluka
Glycerol	7530.1	Roth
Glycine	33226	Sigma-Aldrich
Glycogen	10901393001	Roche
Hydrochloric Acid 1 mol/L, 1N	30024.290	VWR
Hydrochloric Acid fuming 37 %	1.00317.100	Merck
Isopropanol (2-Propanol)	33539	Merck
Laemmli Sample Buffer 2X	S3401-10L	Sigma
Leukemia Inhibitory Factor (LIF)	ESG1107	Millipore
L-Glutamine 200 mM (100X)	25030-024	Gibco
Lipofectamine	P/N 52758	Invitrogen
Lumigen ECL Ultra	TMA-100	Lumigen Inc.
Methanol	32213	Honeywell
Minimum Essential Medium Non-Essential Amino Acids(MEM NEAA) 100 X	11140-035	Gibco
Mouse Leukemia Inhibitory Factor (mLIF), 10^7 U/mL	ESG1107	Millipore
Murine IL-3	213-13	Peprotech
Murine IL-6	216-16	Peprotech
Murine Stem Cell Factor (SCF)	250-03	Peprotech
Neomycin Trisulfate Salt Hydrate	N1876	Sigma
Nuclease-Free H_2O	P119C	Promega
Penicillin/Streptomycin (Pen/Strep)	15140-122	Gibco
Phenol	0038.2	Roth
Phosphate Buffered Saline (PBS)	14190-094	Gibco
Polybrene	TR-1003-G	Merck
Polythylenimine (PEI) C_2H_5N	43896	Alfa Aesar
Potassium Acetate (KAc)	4986	Roth
Precision Plus Protein Standards	161-0374	Bio-Rad
Protein Assay Dye (Bradford)	500-0006	Bio-Rad
Proteinase K	7528.2	Roth
Readiload	46-6029	Invitrogen
Restore Western Blot Stripping Buffer	21059	Thermo Scientific
RetroNectin	T100B	Takara
RIPA Lysis Buffer 10 X	20-188	Millipore
RNAlater TM Solution	AM7020	Invitrogen
RPMI	21875091	Gibco
Sample Buffer Laemmli 2 X	S3401-10VL	Sigma
Sodium Acetate (NaAc)	S2889	Sigma-Aldrich
Sodium Chloride (NaCl)	31424N	Sigma-Aldrich
Sodium Dihydrogen Phosphate (NaH_2PO_4)	A3559	AppliChem
Sodium Dodecylsulfate (SDS)	A2263	AppliChem
Sodium Hydroxide (NaOH)	1.09137	Merck
Sodium Pyruvate 100 mM (100 X)	11360-039	Gibco

Table continued on next page.

ssRNA ladder	N0362	New England Biolabs
TaqMan Gene Expression Master Mix	4369016	Applied Biosystems
Tetramethylethylenediamine (TEMED)	161-0800	Bio-Rad
TransIT-LT1	MIR 2300	Mirus
Tris-HCl	90903	Roth
Trizma®Base	T1503	Sigma
Trypan Blue	T8154	Sigma
Trypsin-EDTA 0.05 %	25300-054	Gibco
Tween 20	93773	Sigma
Ultrapure H_2O	34877	Sigma-Aldrich

2.1.5 Services

Two services were performed out of house.

Next-generation Sanger sequencing of genomic material was performed at GATC Biotech, Konstanz. Commercial primers were supplied by the company. The DNA as well as required custom primers (see section 2.1.13) were supplied by our laboratory.

The blastocyst injection required during gene targeting was performed at the Transgenic Core Facility of the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden under the supervision of Dipl. Ing. Ronald Naumann. Chimeric mice were transferred to our animal facility (Tierforschungszentrum, TFZ) for continuous breeding and housing.

2.1.6 Buffers and solutions

Unless otherwise stated, all buffers were prepared with double distilled H_2O (ddH_2O) delivered by the pharmacy of the University Clinics Ulm.

Buffers/ Solutions	**Composition**
Erythrocyte Lysis Buffer	150 mM NaCl, 10 m $NaHCO_3$, 100 mM EDTA-Na_2, pH 7.2 – 7.4
FACS Buffer	PBS + 5 % FCS, filtered
Gelatin Solution 0.15 %	1.5 g gelatin dissolved in 1 L MT-PBS
Geneticin (G418) Stock	5 g dissolved in 37 mL ES medium, aliquots stored at -20 °C
Mouse Toxicity (MT)-PBS	4mM $NaH_2PO_4 \cdot H_2O$, 16 mM $Na_2HPO_4 \cdot H_2O$, 150 mM NaCl; pH7.3
Proteinase K Solution	10 mM Tris-HCl (pH 8), 5 mM EDTA, 1% SDS, 300 mM NaAc (pH 8), 200 µg/mL proteinase K
TAE Buffer 50X	For 1 L: 242 g Trizma Base, 57.1 mL Acetic Acid, 100 mL of 0.5 M EDTA pH 8.0
TE Buffer	25 mM NaOH, 0.2 mM EDTA
Buffers for DNA extraction	
Alkaline Lysis Buffer	1 % 2.5 M NaOH, 1 % 20 mM EDTA
Neutralizing Buffer	10 % 400 mM Tris-HCl

Buffers used for plasmid Mini-Preparation	
GTE (Mini-Prep)	25 mM Tris-HCl (pH 8.0), 50 mM Glucose, 10 mM EDTA
Lysis Buffer Mini-Preparation (NaOH/SDS)	0.2 M NaOH, 1 % SDS
Neutralization Buffer Mini-Prep (KAc)	3 M KAc; pH 5.5 adjusted with acetic acid
TE/RNase	20 µL RNase, 980 µL TE Buffer
Buffers used for Western Blot	
Blocking Buffer 5 %	TBS-T, 5% Chicken Egg Albumin or Bovine Serum Albumin
Coomassie Staining	0.25 % Coomassie Brilliant Blue was dissolved in Methanol (final concentration of 40 %); 10 % Acetic Acid and 50% ddH_2O were added
Resolving Gel Buffer	1.5 M Trizma base, 0.4 % SDS; pH 8.8
$RIPA^+$ Buffer	10 mL 1X RIPA Lysis Buffer, 1 Tablet c0mplete Protease Inhibitor; Aliquots stored at – 20 °C
Running Buffer 10X	For 1 L: 33.3g Trizma Base, 144.2g Glycine, 1% SDS
Stacking Gel Buffer	0.5 M Trizma Base, 0.4 % SDS; pH 6.8
TBS-T 0.1 %	1X TBS, 0.1 % Tween-20
Tris Buffered Saline (TBS) 10X	For 1 L: 24 g Tris-HCl, 5.6 g Trizma Base, 88 g NaCl; pH 7.6
Buffers for RNA gel:	
10X FA Gel Buffer	200 mM MOPS, 50 mM NaAc, 10 mM EDTA; pH 7.0
1X FA Gel Running Buffer	For 1 L: 100 mL 10X FA buffer, 20 mL 37 % Formaldehyde, 880 mL RNase-free (DEPC treated) H_2O
5X RNA Loading Buffer	For 10 mL: 16 µL saturated Brophenol Blue Solution, 80 µL 500 mM EDTA pH8, 720 µL 37 % Formaldehyde, 2 mL 100 % Glycerol, 3.084 mL Formaide, 4 mL 10X FA Gel Buffer, RNase-free (DEPC treated) H_2O to 10 mL
RNase-free H_2O	1 mL DEPC per 1 L ddH_2O, overnight incubation at 60 °C water bath, autoclave the next day

2.1.7 Bacterial growth medium

Media intended for bacterial growth were prepared with ddH_2O, autoclaved, and stored at 4 °C until use. Agar was poured into 10 cm bacteriological dishes under a sterile conditions. After solidifying, the agar plates were packed in plastic bags provided with the dishes, sealed, and stored at 4 °C. All values presented in the following table correspond to 1 L of the medium.

Medium	**Composition**
2x YT Medium	16 g Bacto-Tryptone, 10 g Yeast Extract, 5 g NaCl; pH 7.0
LB Agar plates	15 g Bacto Agar added per 1 L LB medium; if required additional 50 µg/mL ampicillin after cooling the bottle to ~50 °C
LB Medium	10 g Bacto-Tryptone, 5 g Yeast Extract, 10 g NaCl; pH 7.0

2.1.8 Medium for mammalian cell culture

Medium for mammalian cells was equipped with heat-inactivated fetal calf serum (HI-FCS). Heat inactivation was achieved by 30 min incubation in a 56 °C water bath. Afterwards, the HI-FCS was filtered, aliquotped to 50 mL conical tubes, and stored at -20 °C. Details on the used compounds are given in section 2.1.4. All cell culture media were prepared under sterile conditions using a laminar flow hood and sterilized via vacuum filtration with a 0.22 µm filter system. The complete media were stored at 4 °C.

Medium		Composition
Mouse Embryonic Fibroblasts (MEF)	Culture	DMEM + GlutaMax, 7.5 % HI-FCS, 1X NEAA, 1 mM Sodiumpyruvat, 50 µg/mL Gentamicin, 150 µM/L 1-Thioglycerol 37 °C, 10 % CO_2
	Freezing	40 % medium, 50 % FCS, 10 % DMSO
Embryonic Stem Cells (ES)	Culture	DMEM + GlutaMax, 15 % HI-FCS, 1X NEAA, 1 mM Sodiumpyruvat, 50 µg/mL Gentamicin, 150 µM/L 1-Thioglycerol, 12 x 10^6 U/L LIF 37 °C, 10 % CO_2
	Freezing	70 % medium, 20 % FCS, 10 % DMSO
Primary Bone Marrow Cells (BM)	Culture	RPMI, 10 % HI-FCS, 1 % Pen/Strep Add directly before use: 1 µg/mL IL-3, 1 µg/mL IL-6, 5 µg/mL SCF 37 °C, 5 % CO_2
	Freezing	50 % medium, 40 % FCS, 10 % DMSO
Phoenix/Eco	Culture	DMEM + GlutaMax, 10 % HI-FCS, 1 % Pen/Strep 37 °C, 5 % CO_2
	Freezing	95 % medium, 5 % DMSO
	Transfection	DMEM + GlutaMax, 30 % HI-FCS, 1 % Pen/Strep
NIH/3T3	Culture	DMEM + GlutaMax, 10 % HI-FCS 37 °C, 5 % CO_2
	Freezing	95 % medium, 5 % DMSO
Geneticin (G418) Stock		5 g G418 and 0.4 mL 10 M NaOH added to 37 mL ESC medium, filter sterilized, 1.6 mL aliquots stored at -20 °C

2.1.9 Enzymes

2.1.9.1 Polymerases

Polymerase	Reference number	Supplier
AmpliTag Gold DNA Polymerase	4311820	Applied Biosystems
Expand Long Template PCR System	11681834001	Roche
Q5 High-Fidelity 2X Master Mix	M0492	New England Biolabs
Q5 High-Fidelity DNA Polymerase	M0491	New England Biolabs
CloneAmp HiFi PCR Premix	639298	Takara Clontech

2.1.9.2 Restriction enzymes

Unless otherwise stated, all enzymes were used for digestion in the respective buffer provided by the supplier under recommended conditions.

Enzyme	Reference number	Supplier
BamHI	R0136	New England Biolabs
BglII	R0144	New England Biolabs
BstEI	R0162	New England Biolabs
EcoRI	R0101	New England Biolabs
EcoRV	R0195	New England Biolabs
FseI	R0588	New England Biolabs
HindIII	R0104	New England Biolabs
NcoI	R0193	New England Biolabs
NotI-HF	R3189	New England Biolabs
NsiI	R0127	New England Biolabs
PshAI	R0593	New England Biolabs
PstI-HF	R3140	New England Biolabs
SalI	R0138	New England Biolabs
SnaBI	R0130	New England Biolabs
XhoI	R0146	New England Biolabs

2.1.9.3 Other enzymes

Enzyme	Reference number	Supplier
Alkaline Phosphatase Calf Intestinal (CIP)	M0290	New England Biolabs
DNaseI	79254	Qiagen
DNaseI	18047019	Invitrogen
Quick Ligation Kit	M2200	New England Biolabs
RNaseI	1006657	Qiagen
Shrimp Alkaline Phosphatase (rSAP)	M0371	New England Biolabs
T4 Ligase	M0202	New England Biolabs

2.1.10 Kits

Kit	Reference number	Supplier
DNeasy Blood & Tissue Kit	69504	Qiagen
EasySep™ Mouse Hematopoietic Progenitor Cell Isolation Kit	19856	Stem Cell
Foxp3/Transcription Factor Staining Buffer Kit	00-5523-00	eBioscience
Histone Extraction Kit	ab113476	Abcam
peqGOLD Gel Extraction Kit	12-2500-02	PeqLab
Plasmid Plus Midi Kit	12943	Qiagen
QIAprep Spin Miniprep Kit	27106	Qiagen
QIAshredder	79656	Qiagen

Table continued on next page.

RNeasy Midi Kit	75144	Qiagen
RNeasy Mini Kit	74104	Qiagen
SuperScript III First-Strand Synthesis	18080051	Invitrogen
Trans-Blot Turbo RTA Transfer Kit, PVDF	1704272	Bio-Rad

2.1.11 Antibodies

2.1.11.1 Flow Cytometry

Antigen	Conjugate	Clone	Reference	Supplier	Dilution
Biotinylated antibodies (* = used for lineage mix):					
CD3e*	Biotin	500A2	553239	BD	1:100
CD4*	Biotin	GK1.5	553728	BD	1:800
CD8a*	Biotin	53-6.7	13-0081-85	eBioscience	1:400
CD11b*	Biotin	M1/70	553309	BD	1:500
CD19*	Biotin	1D3	553784	BD	1:400
CD45.1	Biotin	A20	13-0453-82	Invitrogen	1:300
CD45.2	Biotin	104	13-0454-82	eBioscience	1:400
CD97 (Gr1)*	Biotin	RB6-8C5	13-5931-85	eBioscience	1:500
CD161 (NK1.1)*	Biotin	PK136	13-5941-82	eBioscience	1:800
TCRβ	Biotin	H57-597	13-5961-85	eBioscience	1:100
Ter119*	Biotin	TER-119	13-5921-85	BD	1:400
Directly labeled antibodies					
CD3e	APC	145-2C11	553066	BD	1:40
CD3e	PE Cy7	145-2C11	25-0031-82	Invitrogen	1:50
CD3e	APC Cy7	145-2C11	557596	BD	1:400
CD8a	APC	563-6.7	17-0081-83	eBioscience	1:800
CD8a	PerCP Cy5.5	145-2C11	25-0031-82	Invitrogen	1:50
CD16/32 (FcγRII/III)	PE	2.4G2	553145	BD	1:100
CD34	eFluor 660	RAM34	50-0341-82	Invitrogen	1:25
CD45.1	APC	A20	17-0453-82	Invitrogen	1:400
CD45.2	APC	104	17-0454-82	Invitrogen	1:100
CD117 (cKit)	APC Cy7	2B8	47-1171-82	eBioscience	1:400
CD135 (Flt3)	PE	A2F10	12-1351-82	Invitrogen	1:200
Sca1	PerCP Cy5.5	D7	45-5981-82	Invitrogen	1:200
TCRγδ	APC	eBioGL3	17-5711-81	eBioscience	1:200
Antibodies used for F_C-block					
CD 16/32	-	2.4G2	553142	BD	1:100
Mouse IgG	-		015-000-003	Jackson Immuno Research	1:20
Secondary fluorochrome conjugated antibodies					
Alex Fluor 488	For Rabbit		4412S	Cell Signaling	1:400
BV785	For Biotin		405249	BioLegend	1:200
Qdot BV605	For Biotin		Q100001MP	Invitrogen	1:400

Live/Death staining dyes					
Fixable Viability Dye	Pacific Blue		65-0863-14	Invitrogen	1:1000
Sytox	Pacific Blue		S34857	Invitrogen	1:2500
Antibodies used for Intracellular Staining					
FLAG	PE	L5	357309	BioLegend	1:200
H3K27me3	Rabbit/IgG	EPR18607	Ab192985	Abcam	1:1200
H3K4me3	Rabbit/IgG	G.532.8	MA5-11199	Invitrogen	1:200

2.1.11.2 Western Blot

Antigen	Isotype	Reference	Supplier	Size	Dilution
Primary antibodies					
FLAG M2	Mouse	F3165	Abcam	depends	1:5000
GAPDH	Rabbit	2118	Cell Signaling Technologies	37 kDa	1:5000
GFP and YFP	Rabbit	A6455	Life Technologies	25 kDa	1:1000
H3.3	Rabbit	ab183902	Abcam	15 kDa	1:5000
Secondary antibodies, HRP-conjugated					
α-Rabbit		32460	ThermoFisher	-	1:5000
α-Mouse		1858413	Pierce	-	1:20,000

2.1.12 Plasmids/ Vectors

Plasmid	Source
pBSK	Stratagene
pCAGGS-FLPe-V1	Institute for Immunology, Ulm
pCL Eco	Institute for Immunology, Ulm
pDB-AA	Institute for Immunology, Ulm
pEGFP	GeneCopoeia
pFG-AD	Institute for Immunology, Ulm
pFG-H3TG-CF	Institute for Immunology, Ulm
pH3F3-JM	Prof. Dr. Jeffrey Mann, The Royal Children's Hospital Australien (Tang et al., 2013)
piCre-AC	Institute for Immunology, Ulm
pIDH2-Ex-C0463-M02	GeneCopoeia
pIDH2-IRESmCherry	GeneCopoeia
pKK-ID-AFM	Institute for Immunology, Ulm
pMIBerry	Institute for Immunology, AG Jumaa, Ulm

2.1.13 Oligonucleotides

All self-designed oligonucleotides were synthesized by Biomers.net GmbH, Ulm. They were used as polylinkers during cloning procedures, or as primers during PCR. The oligonucleotides were delivered after cartridge cleanup as a lyophilized powder. Upon use, they were reconstituted using ultrapure H_2O to a stock concentration of 100 pm/µL. Unless stated, the oligonucleotides were further diluted to a working solution of 10 pm/µL.

Primer	Sequence (5' → 3')
Primers used for genotyping	
KK44	CGA AGT TAT AGA TCC GCC ACC A
KK45	ATG AAC GTG GAG AAA CGC CAA T
KK53	CCT TTC TGT GTT TGT GGC TTC G
KK54	TCC CAC TGC TCG ACG TTT TAG
vm17	AGG ACG AGC ACC ATA ATT ATG GTC
vm184	GGA CAT GTT CAG GGA CAG GCA G
vm185	CTC TGA TTC TGG CAA TTT CGG C
vm76	CCC TTA TTC TTA CAT AGA CTT G
Oligonucleotides and primers used for cloning and characterization of stem cell clones	
FB28	GGC CGC TTA AGA CGC GTG ACG TCA CGT G
FB29	AAT TCA CGT GAC GTC ACG CGT CTT AAG C
FB30	GAT CCA CCA TGG CTC GTA CAA AGC CTG CCC GCA AAT CCA
FB31	CCG GTG GAT TTG CGG GCA GTC TGC TTA GTA CGA GCC ATG GTG
FG16	GGC TTT TAA TGA CAG CTA CAC GA
HL51	CCT GAA AAC TTT GCC CCC TCC ATA
HL51_1	AAC AGG TAC CTG AAA ACT TTG CCC CCT CCA TAA AC
HL51_2	AAT TGT TTA TGG AGG GGG CAA AGT TTT CAG GTA CCT GTT
KK01	TCG AGA CAT ATG GTT TAA ACA TGA GAC AA
KK02	TTG TCT CAT GTT TAA ACC ATA TGT C
KK03	GAT CCA TTA CGG TTT ATG CAT AAA GGT ACC AGA TCT G
KK04	AAT TCA GAT CTG GTA CCT TTA TGC ATA AAC CGT AAT G
KK05	CTG ATG ACC AGG TCA CCA TTG ACT CTG C
KK06	TGG TCG CAA TGG GCG TGC TTG CCA TGA GTG ATG GGC
KK07	GGC CGC GAA GGA GTT CGA ACC ATG GCC GGC TAC CTG CGG GTC GTG CGC TCG CTC TGC A
KK08	GAG CGA GCG CAC GAC CCG CAG GTA GCC GGC CAT GGT TCG AAC TCC TTC GC
KK09	AAT TGA TGC ATG CGG CCG CAT TTA AAT GTC GAC CTC GAG AAT TCA CGC GTA
KK10	GAT CTA CGC GTG AAT TCT CGA GGT CGA CAT TTA AAT GCG GCC GCA TGC ATC
KK11	AGC TGG CCG GCC TAC GTA GAC GTC GAC A
KK12	AGC TTG TCG ACG TCT ACG TAG GCC GGC C
KK21	TGC TCG TCA AGA AGA CAG GG
KK22	CCT CGG TGT CAG CCA TCT TT

Table continued on next page.

KK25	CAA TTT CTC GCA CCA GAC GC
KK27	AGC GTC TGA TTT CAC GGA GG
KK34-FLAG	ATC CAG CTA GCA CGC CGC ATA CGC GGA GAA CGT GCT GCA GCT GCC GGC GGG GAT TAT AAA GAC CAC GAT GGA GAC TAT AAA GAT CAT GAC ATT GAC TAC AAG GAT GAC GAC GAC AAG AGC GCT TGA AAA GGC CGG
KK35-FLAG	CCT TTT CAA GCG CTC TTG TCG TCG TCA TCC TTG TAG TCA ATG TCA TGA TCT TTA TAG TCT CCA TCG TGG TCT TTA TAA TCC CCG CCG GCA GCT GCA GCA CGT TCT CCG CGT ATG CGG CGT GCT AGC TGG AT
KK39	ACA CCG GCC TTA TTC CAA GC
KK40	TCA ATT GTG TTC GCA GCC TTC
RT 173	CAT CAC TAT TGG CAA CGA GC
RT 174	ACG CAG CTC AGT AAC AGT CC

2.1.13.1 Sequencing

Next Generation Sequencing was performed by GATC Biotech, Konstanz. Primers with the additional comment 'GATC' are commercial standard primer provided by the company. All other primers were shipped for sequencing together with the sample DNA.

Primer	Sequence (5' → 3')	
KK05	CTG ATG ACC AGG TCA CCA TTG ACT CTG C	
KK21	TGC TCG TCA AGA AGA CAG GG	
KK25	CAA TTT CTC GCA CCA GAC GC	
KK27	AGC GTC TGA TTT CAC GGA GG	
KK30	GGC CTA TCT GGT TGG CCT TT	
KK36	GAC GGC GCG TGA ATT GGA TC	
KK39	ACA CCG GCC TTA TTC CAA GC	
KK40	TCA ATT GTG TTC GCA GCC TTC	
M13-FV	TGT AAA ACG ACG GCC AGT	GATC
M13-RV	CAG GAA ACA GCT ATG ACC	GATC
pEGFP_C2-FP	GAT CAC ATG GTC CTG CTG	GATC
vYFP-Fwd-FB	GCG GGC GAA GCA CTG CAG GCC	

2.1.13.2 qRT-PCR

Primers used for qRT-PCR were ordered as TaqMan gene expression assays from Applied Biosystems. Each assay included a gene-specific forward and reverse primer pair as well as a fluorophore (FAM) labeled probe. Custom primers were designed using the Primer Express Software. Expression was normalized against Actin β as endogenous control.

Assay	Reference Number	Supplier
ACTB (Actin β)	4352663	Applied Biosystems
H3f3a	Mm01612808_g1	Applied Biosystems
H3f3a WT-cDNA	Self-Designed	Applied Biosystems

2.1.14 Bacterial strains

The following competent Escherichia coli (*E. coli)* strains were used for transformation with plasmids. The original stocks were purchased from New England Biolabs and propagated as described in chapter 2.2.3.

Strain	Reference number	Supplier
NEB 5-α Competent *E.coli* (High Efficiency)	C2987I	New England Biolabs
NEB 10-beta Competent *E.coli* (High Efficiency)	C3019I	New England Biolabs

2.1.15 Mammalian cell lines

The following mammalian cell lines were used during this project. Information on the required media and culturing conditions are given in sections 0 and 2.2.6.

Cell line	Supplier	Use
NIH/3T3	ATCC	To test transformation efficiency of viral supernatants
Phoenix/Eco	ATCC	Generation of viral supernatants
ES/R1 Embryonic Stem Cells (ESC)	Max Planck Institute of Molecular Cell Biology and Genetics, Dresden	Stem cells used for gene targeting
Neo^R Mouse Embryonic Fibroblasts (MEF)	Institute for Immunology, Ulm	Feeder cells for ES cells

2.1.16 Transgenic mouse strains

All mouse work was performed according to German and European directives. The animal permit was granted by the Regierungspräsidium Tübingen with the number 1467.TschB: H. Housing and breeding of the required mice was done at the Tierforschungszentrum (TFZ) of Ulm University. The mouse strains used during this project are described in the following passage.

Mouse strain	Full scientific name
C57BL/6J	C57BL/6J
B6.SJL	B6.SJL-$Ptprc^{a}Pepc^{b}$/BoyJ
vCre	B6.Cg-Tg(Vav1-cre)A2Kio/J
vCre x RoYFP	B6N.Cg-Tg(Vav1-icre)A2Kio/J x B6.129X1-$Gt(ROSA)26Sor^{tm1(EYFP)Cos}$/HJF
vCre-Mll5 F/F	B6.Cg-Tg(Vav1-cre)A2Kio/Jx B6.129S2-$Mll5^{tm1Hjf}$/HJF
H3K27M	B6.129S1/X1-$H3f3a^{tm1(lox-K27M)Hjf}$
vCre x H3K27	B6N.Cg-Tg(Vav1-icre)A2Kio/J x B6.129S1/X1-$H3f3a^{tm1(lox-K27M)Hjf}$
vCre-Mll5 x H3K27	B6.Cg-Tg(Vav1-cre)A2Kio/J x B6.129S2-$Mll5^{tm1b(loxP)Hjf}$/HJF x B6.129S1/X1-$H3f3a^{tm1(lox-K27M)Hjf}$
vCre x DnmT3-F/F	B6.Cg-Tg(Vav1-cre)A2Kio/J x C57BL7&N-A^{tm1Sjk}/JDnmt3a$^{tm1c(KOMP)Wist/WtsiOr}$I
Neo^R	

C57BL/6J mice were used as wild type controls during some experiments, they carry CD45.2 positive lymphocytes. Furthermore, all of the transgenic mice used during this project were bred on a C57BL/6J background. B6.SJL mice (Shen et al., 1985) are congenic mice used as recipients during bone marrow transplantations. They harbor CD45.1 positive lymphocytes. Mice of both strains were intercrossed to get CD45.1/CD45.2 heterozygous lymphocytes. These intercrossed mice were used as donors for WT competitor cells during transplantation experiments (see 2.2.17).

H3K27M mice were generated during the course of this project (a detailed description of the process is given in chapter 2.2.8). They carry the whole targeting construct but, due to the absence of Cre recombinase, only express the *H3f3a* WT allele. To knock in the H3K27M mutation, these mice were intercrossed with vCre mice (de Boer et al., 2003) to vCre x H3K27M mice. In Cre-positive mice, the mutation is activated in all hematopoietic cells.

Mll5$^{flox/flox}$ mice have previously been generated in our group (Madan et al., 2009). Intercrossing with vCre (vCre-Mll5) activates the loss of *Mll5* exclusively in hematopoietic cells. To archive double mutant mice, vCre x H3K27M were intercrossed with vCre-Mll5 F/F mice. The resulting mice have a deletion of *Mll5* as well as a knock-in of the H3K27M mutation.

2.1.17 Statistical tests applied during the project

Statistical analysis was performed using GraphPad Prism software. Due to the small size of the sample groups, Anderson-Darling-Test for the calculation of normal distribution could not be applied, hence all applied tests were non-parametric.

The number of replicates can be found in the respective figure legends. Fold change data as well as the comparison of different cell subsets was determined via non-parametric, two-tailed Mann-Whitney-U test. Survival rate was displayed in Kaplan-Meier curves and significance was calculated using Log-Rank test. P-values <0.033 were considered significant ($p = 0.033$ (*); $p = 0.002$ (**); $p = < 0.001$ (***)).

2.2 Methods

2.2.1 DNA Extraction

Throughout this project, various techniques for DNA extraction were used. The specific extraction method varied depending on the purpose of extracted DNA.

During the weaning process, ear punch biopsies were taken to enable subsequent identification of the individual mice. These biopsies were then used to extract genomic DNA (gDNA) which was used for genotyping of the mouse cohort. For genotyping, the DNA was extracted using a two-buffer system. 100 µL of alkaline lysis buffer were added to the biopsies and incubated for 30 min at 95 °C on a shaker. The samples were then cooled on ice for 5 min and 100 µL of neutralization buffer were added. The detailed buffer compositions are given in section 2.1.6.

DNA from organs, cell suspensions, or cultured cells was isolated using phenol/chloroform extraction. In short, the material was incubated in two volumes of proteinase K solution (e.g. 200 µL of proteinase K for tail biopsies) for 3 – 24 h at 55 °C on a shaker. One volume of phenol was added and the sample was vortexed for 40 s. Immediately afterwards, one volume of chloroform was added, the sample was vortexed for 120 s and centrifuged at RT for 15 min at 13,000 rpm. The upper phase (containing the DNA) was transferred to a new microcentrifuge tube containing two volumes of chloroform. The sample was vortexed for 60 s and centrifuged under the previous conditions. Again, the upper phase was transferred to a new tube, containing 900 µL of 100% ethanol. The sample was briefly vortexed to precipitate DNA and centrifuged at RT for 10 min at 13,000 rpm. The supernatant was discarded and the pellet was washed with 600 mL of 70 % ethanol and centrifuged at RT for 5 min at 13,000 rpm. After removing the supernatant, the pellet was air-dried for 20 min and dissolved in 50 µL of TE buffer.

In both cases, isolated DNA was stored at 4 °C. For long-term storage of high-quality DNA, it was precipitated in 100 % ethanol and stored at -80 °C.

2.2.2 Genotyping

Genotyping of the mice was performed using polymerase chain reaction (PCR). For DNA samples isolated using the two-buffer system, 2 µL of isolated DNA were required for genotyping. After extraction of DNA with phenol/chloroform, 1 µL per sample was sufficient. In both cases, 24 µL of the genotyping master mix were used.

Specific details on the genotyping PCRs are given in the following sections. Unless otherwise idicated, gel electrophoresis of the PCR products was performed with 1.5 % agarose gels run at 110 V for 30 min.

2.2.2.1 H3-K27M genotyping PCR

During the process of gene targeting and the initial breeding to establish germline transmission of the new transgene, DNA was extracted via the phenol/chloroform technique. Later, for routine genotyping of mice, the two-buffer system was used.

Reaction Mix			PCR Conditions		
Q5 Master Mix		12.5 µL	95 °C	30 sec	
WT Forward (10 pm/µL)	KK 53	1.25 µL	92 °C	10 sec	
WT Reverse (10 pm/µL)	KK 54	1.25 µL	57 °C	20 sec	35x
KI Forward (10 pm/µL)	KK 44	1.25 µL	72 °C	25 sec	
KI Reverse (10 pm/µL)	KK 45	1.25 µL	72 °C	2 min	
H_2O		6.5 µL	4 °C	∞	
Total Volume		24 µL			

After PCR, 4 µL of purple loading dye were added to each sample. Gel electrophoresis was performed with a 2 % agarose gel run for 45 min at 110 V.

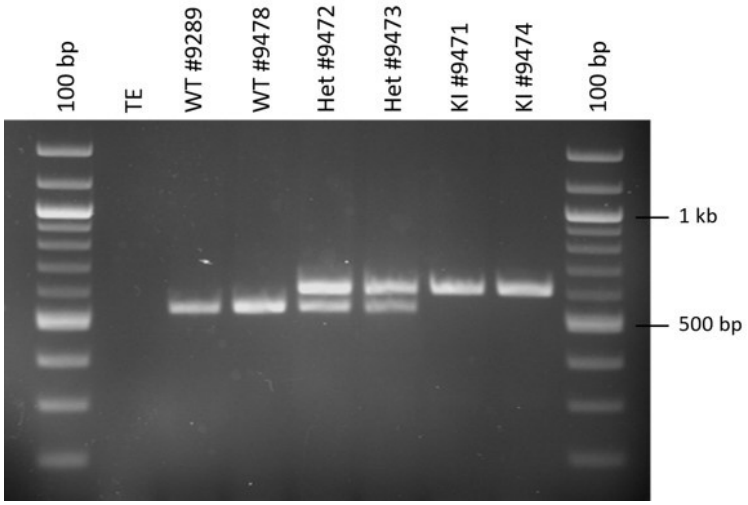


Figure 8: Genotyping of H3f3aK27M
Genotyping for H3f3a results in a 542 bp product from the wild type allele and 617 bp from the knock-in allele. TE = Tris/EDTA buffer; WT = wild type; Het = heterozygous KI/WT; KI = homozygous KI/KI; bp = base pair

2.2.2.2 Mll5-flox genotyping PCR

Reaction Mix		PCR Conditions		
$MgCl_2$ (25 M)	2.5 µL	95 °C	5 min	
vm17 (10 pm/µL)	1.0 µL	92 °C	30 sec	35x
vm76 (10 pm/µL)	1.0 µL	57 °C	30 sec	
dNTP (10 mM)	0.5 µL	72 °C	1min	
AmpliTaq Gold Buffer	2.5 µL	72 °C	7 min	
AmpliTaq Gold Polymerase	0.3 µL	4 °C	∞	
Rediload	2.5 µL			
H_2O	13.7 µL			
Total Volume	24 µL			

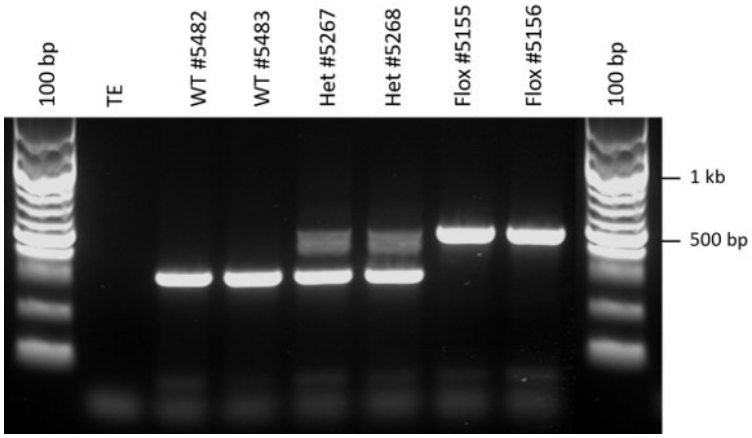


Figure 9: Genotyping of Mll5^{flox}
Amplification of the wild type results in a 333 bp product, amplification of the floxed allele in a 508 bp product.
TE = Tris/EDTA buffer; WT = wild type; Het = heterozygous KI/WT; bp = base pair

2.2.2.3 vCre genotyping PCR

Reaction Mix		PCR Conditions		
$MgCl_2$ (25 M)	2.5 µL	95 °C	5 min	
vm184 Forward (10 pm/µL)	1.0 µL	92 °C	30 sec	35x
vm185 Reverse (10 pm/µL)	1.0 µL	57 °C	30 sec	
dNTP (10 mM)	0.5 µL	72 °C	1min	
AmpliTaq Gold Buffer	2.5 µL	72 °C	7 min	
AmpliTaq Gold Polymerase	0.3 µL	4 °C	∞	
Rediload	2.5 µL			
H_2O	13.7 µL			
Total Volume	24 µL			

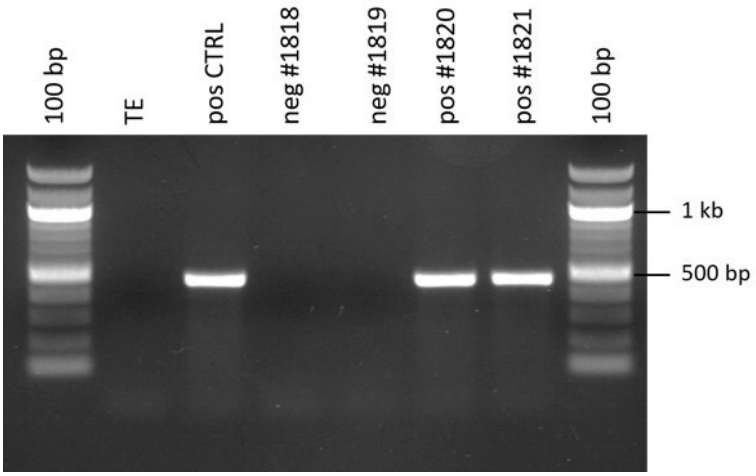


Figure 10: Genotyping of vCre
Cre flox positive samples yield a product of 480 bp. TE = Tris/EDTA buffer; neg = negative; pos = postivie; CTRL = control; bp = base pair

2.2.3 Preparation of competent cells

All cloning procedures were carried out in self-propagated $CaCl_2$-competent cells. Two strains of competent *Escherichia coli* (*E. coli*) were originally purchased from New England Biolabs: *DH5α*, used for smaller constructs <10 kb in size, and *NEB10β*, for constructs >10 kb. Preparation of cells was done according to Sambrook et al. (2001). All procedures were carried out in a laminar flow hood under sterile conditions.

A vial of competent cells was thawed on ice. 200 µL of bacterial suspension were transferred to 1 mL 2xYT medium in a 10 mL round bottom tube. After incubation for 1 h at 37 °C and 220 rpm on a shaker, 100 µL bacterial suspension were plated on a LB-Agar plate and incubated at 37 °C overnight in a bacterial incubator. The next day, a subculture was prepared by picking one single colony and streaking it onto a new LB-Agar plate, which again was incubated overnight at 37 °C. After 12 – 16 h a single colony was picked and pre-cultured overnight in 2 mL medium at 37 °C at 220 rpm. The next morning, the 2 mL subculture was transferred into 500 mL 2xYT medium in a 1 L Erlenmeyer flask. The bacterial suspension was incubated on a shaker at 37 °C at 220 rpm until an optical density (OD600) of 0.3 – 0.4, indicating a log growth phase, was reached. The culture was then transferred into two 250 mL screw cap bottles and incubated on ice for 15 min. After incubation, the suspension was centrifuged at 2,500 rpm at 4 °C for 10 min. The supernatant was discarded and the pellet was resuspended in 62.5 mL of ice-cold 50 mM $CaCl_2$ per bottle by gentle pipetting. The cell suspension was incubated for 45 min on ice. Afterward, the cells were centrifuged under the same conditions. The remaining pellets were resuspended in a total of 25 mL ice-cold 50 mM $CaCl_2$. 800 µL aliquots were stored at -80°C until use. Transformation efficiency was tested by transformation with pBSK.

2.2.4 Cloning procedure

All required cloning steps were carried out according to the standard cloning procedure provided by Sambrook et al. (2001). Details on the individual steps are given in the following sections.

2.2.4.1 Restriction digest

During the process of cloning, PCR products or plasmid fractions were incorporated into a vector plasmid. To obtain the necessary fragments, available restriction enzymes were utilized. The standard mix for digestion of 1 µg DNA is given below. However, the mix can be adapted to the intended use such as test digestion of plasmids, or vector and insert preparation. Every enzyme is equipped with a specific 10X reaction buffer that has to be adjusted according to the enzyme(s) used during the restriction digest.

Reagent	Amount
Plasmid DNA or PCR Product	1 µg
10X Buffer	2 µL
Restriction Enzyme	0.5 µL
Ultra-pure H_2O	To 20 µL

After pipetting, the digestion mix was incubated for 1 h at the enzyme-specific temperature, typically at 37 °C. To prevent relegation, the vector was treated with shrimp alkaline phosphatase (rSAP) for 1 h at 37 °C to remove the phosphate ends. The dephosphorylating reaction was terminated at 65 °C for 5 min. Subsequently, the digests were run on an agarose gel of the required concentration (0.8 – 2 %) to verify correct digestion pattern. If the fragments were required for cloning, they were cut out using a sterile scalpel blade under UV light. The purification of DNA from agarose gel is described in chapter 2.2.4.3.

2.2.4.2 Phosphorylation and annealing of oligonucleotides

In certain cloning procedures, the insertion of linker oligonucleotides into a plasmid or vector is required to introduce additional restriction sites. To enable the insertion of the oligonucleotides into the dephosphorylated vector, phosphorylation had to be performed. This was done using T4 polynucleotide kinase. The following mix was prepared individually for each primer. The reaction was incubated for 30 min at 37 °C.

Reagent	Amount
Primer (10 pmol/µL)	3 µL
T4 Polynucleotide Kinase (1000U/mL)	1 µL
Adenosine Triphosphate (ATP, 10 mM)	1.5 µL
10X T4 Kinase Buffer	2 µL
Ultra-pure H_2O	13.5 µL

Annealing was performed by boiling a 1:1 mixture of the (phosphorylated) oligonucleotides for 10 min in water. Afterwards, the sample was cooled down slowly to RT for roughly 20 min to allow the hybridization of oligonucleotides. The annealed linker was then used directly for ligation or stored at -20 °C.

2.2.4.3 Gel extraction

Extraction of DNA or PCR fragments from agarose gels was performed using the PeqLab Gold Kit according to the manufacturer's protocol. The yield of purified DNA was checked via a Nanodrop spectrophotometer or gel electrophoresis. The remaining DNA was stored at 4 °C until use.

2.2.4.4 Ligation of vector and insert

Two different ligation approaches were applied, depending on size and complexity of the plasmid. For smaller plasmids, quick ligase was used whereas larger constructs, such as the plasmids used for gene targeting, overnight ligation using the T4 ligation kit was performed.

Usually, four different reactions were set up for each ligation with various vector:insert ratios and negative controls. Examples for both approaches are given in Table 5 and Table 6.

Table 5: Exemplary quick ligation mix used for the ligation of pBSK-KK-AA

Reagent	Mix A	Mix B	Mix C	Mix D
Vector	4 µL	4 µL	4 µL	4 µL
Insert	8 µL	4 µL	-	-
2X Buffer	12 µL	10 µL	10 µL	10 µL
T4 Ligase	1 µL	1 µL	1 µL	-
H_2O	-	1 µL	5 µL	6 µL

Table 6: Exemplary overnight ligation mix used for the ligation of pKK-H3TG-AF

Reagent	Mix A	Mix B	Mix C	Mix D
Vector	6 µL	6 µL	6 µL	6 µL
Insert	10 µL	6 µL	-	-
10X Buffer	2 µL	2 µL	2 µL	2 µL
T4 Ligase	1 µL	1 µL	1 µL	-
H_2O	1 µL	5 µL	11 µL	12 µL

After ligation, the mixes were immediately placed on ice or stored at -20 °C until transformation.

2.2.4.5 Transformation of competent cells

Competent cells were thawed on ice. 200 µL of bacterial suspension per approach were transferred into pre-chilled 13 mL round bottom tubes. 20 µL of ligation mix, prepared as described above, were added to the cells. One positive control with 1 ng pBSK and one negative control without any DNA were included. The cells were mixed by gently flicking the tube, followed by incubation for 45 min on ice. Afterwards, a heat shock was performed in a 42 °C water bath for 60 sec followed by 5 min incubation on ice. Next, 900 µL of pre-warmed 2xYT medium were added to each tube. The cells were incubated for 1 h on a shaker at 37 °C and 220 rpm. After incubation, 100 µL of bacterial culture were plated on LB-Agar plates containing 50 µg/mL ampicillin and incubated overnight at 37 °C. The next day, colonies were counted and it was proceeded with the required experiments.

2.2.4.6 Mini-Preparation

After overnight incubation of agar plates, bacterial colonies were used for mini-preparation (mini-prep) to check for the correct ligation and insertion of the plasmids. Therefore, 2 mL of 2xYT medium supplied with 50 µg/mL ampicillin in a 13 mL round bottom tube were inoculated with one single bacterial colony (preferably picked from the center of the agar plate without any satellite colonies) and incubated overnight at 37 °C on a shaker at 220 rpm.

The next day, mini-prep was performed based on the alkaline lysis method introduced by Birnboim (1983). The composition of all buffers used is given in section 2.1.6. The next day, 1.5 mL of bacterial culture were transferred into sterile 1.5 mL microcentrifuge tubes and

centrifuged for 2 min at 13,000 rpm at RT. The remaining culture was stored at 4 °C until further use. After centrifugation, the supernatant was aspirated, and the pellet was resuspended in 150 µL GTE buffer by vortexing. The solution was incubated at RT until the pellet was completely dissolved. 300 µL of lysis buffer (NaOH/SDS) were added. The tubes were then virtuously shaken by hand and incubated for 5 min on ice. Afterwards, 150 µL of neutralization buffer (KAc) were added to precipitate the proteins. The samples were again mixed by hand followed by 10 min centrifugation at 13,000 rpm at RT. The supernatant was transferred to a new tube and centrifuged a second time under the previous conditions. Again, the clear supernatant was transferred to a new 1.5 mL microcentrifuge tube and precipitated under the addition of 600 µL isopropanol. After vortexing, the samples were centrifuged for 15 min at full speed. The supernatant was discarded and the pellet was washed once with 100 µL of 70 % Ethanol and centrifuged for 10 min at full speed. Then, the supernatant was discarded and the pellets were air-dried and afterward dissolved in 40 µL TE/RNase by vortexing. To ensure proper dissolving of the DNA, subsequent incubation for 1 – 16 h at 37 °C on a shaker was performed. Afterward, the DNA was stored at 4 °C until use.

Especially during the cloning of the targeting vectors, large quantities of mini-preps had to be performed to screen for correct integration of the insert into the plasmid. Here, the QIAGEN QIAprep Spin Miniprep Kit was used according to the manufacturers protocol.

For all mini-preps, correct integration and orientation of the plasmid was determined via restriction digest as described in section 2.2.4.1.

2.2.4.7 Midi-Preparation

If increased plasmid yield was required to enable downstream processes such as sequencing, continuous cloning procedures, or gene targeting, a midi-preparation (midi-prep) was performed.

To this extent, the remaining supernatants from the mini-prep were used for dilution streaks on ampicillin agar plates to grow single colonies. After overnight incubation at 37 °C, a single colony was picked and used to inoculate 100 mL of 2xYT medium containing 50 µg/mL ampicillin. Again, incubation was performed overnight at 37 °C and 220 rpm on a shaker. The next morning, midi-prep was performed using the QIAGEN Plasmid Plus Midi Kit according to the manufacturer's protocol. The DNA yield was determined by measuring the absorbance on a Nanodrop spectrophotometer. The final concentration of DNA was adjusted to 1 µg/µL with elution buffer. Isolated plasmids were stored at 4 °C.

Again, the correct orientation was determined using restriction digests. If PCR was part of the cloning procedure (e.g. during the introduction of a mutation), or for final targeting constructs, sequencing was performed at GATC Biotech, Konstanz.

2.2.5 Generation of plasmids

During this project, several plasmids were generated according to the standard cloning procedure described in chapter 2.2.4. An overview of the required steps is given in the following tables. Details about the oligonucleotides used are given in the material section 2.1.13.

2.2.5.1 Generation of a control construct to establish diagnostic PCR for gene targeting

Plasmid Name	Cloning Strategy
pFG-H3TG-CC	pH3f3-JM was digested with BamHI and NsiI. The fragment containing the H3.3 coding sequence was inserted into the BamHI/NsiI sites of pBSK.
pFG-H3TG-CD	The polylinker HL51_1 and HL_51_2 (containing the primer annealing site for HL51) were inserted into the HpaI/MfeI sites of pFG-H3TG-CC.
pFG-H3TG-CE	The polylinker FB30 and FB31 were inserted into the BamHI/AgeI sites of pFG-H3TG-CD to remove a mutation at K9.
pFG-H3TG-CF	pFG-AC was digested with AgeI/BglII to remove the fragment containing the K27M mutation. The fragment was inserted into the AgeI/BglII sites of pFG-H3TG-CE
pBSK-FB1	The polylinker FB28 and FB29 were inserted into the NotI/EcoRI sites of pBSK.
pFG-H3TG-CA	pH3f3-JM was digested with MluI, EcoRI, and AscI. The 7.0 kb EcoRI/MluI fragment was inserted into the MluI/EcoRI sites of pBSK-FB1.
pFG-H3TG-CB	The ~1.5 kb fragment containing the HL1 primer annealing site was transferred from pFG-H3TG-CF into pFG-H3TG-CA using BamHI/NsiI digests.

Cloning procedure was performed by Franziska Oliveri.

2.2.5.2 Generation of targeting constructs

Plasmid Name	Cloning Strategy
pKK-H3TG-AA	pH3f3-JM was digested with XhoI and PshAI to insert the oligonucleotides KK01 and KK02 and remove the EcoRI and BglII sites.
pKK-H3TG-AB	The polylinker KK03 and KK04 were inserted into the BamHI/EcoRI sites of pKK-H3TG-AA.
pKK-H3TG-AC	After digestion of pKK-H3TG-AA with NsiI and BglII the 1.2 kb fragment was transferred into the corresponding sited of pKK-H3TG-AB.
pKK-H3TG-AD	By digesting pFG-H3TG-CF with NsiI and BamHI the fragment carrying the annealing site for primer HL51 and the H3f3a-K27M encoding cDNA was transferred into the corresponding sites of the targeting vector. This is the first targeting vector. It does not contain IRES-vYFP or a 3xFLAG-tag.
pKK-H3TG-AE	The linker oligos KK11 and KK12 were inserted into the HindIII sites of pKK-H3TG-AD. PCR using primers FB30, KK11 and KK12 was used to ensure insertion in the correct orientation.
pKK-H3TG-AF	SnaBI and XhoI were used to excise the IRES-vYFP cassette from pDB-AA. This fragment was cloned into the SnaBI and SalI sites of pKK-H3TG-AE. Resulting clones were sequenced with vYFP-Fwd-FB, pEGFP_C2-FP and KK21.The resulting construct is a targeting vector with the fluorescent marker IRES-vYFP but no 3xFLAG-tag.
pBSK-KK-AA	pKK-H3TG-AF was digested with BamHI and HindIII to transfer the 3'-end of the H3f3a cDNA together with the IRESvYFP cassette into the BamHI/HindIII sites of pBSK.
pBSK-KK-AB	The 3xFLAG-tag oligos (KK34-FLAG and KK35-FLAG) were annealed and transferred into the EcoRV/FseI sites of pBSK-KK-AA. Resulting clones were sequenced with KK30, M13-FP, and M13-RP to ensure complete insertion.
pKK-H3TG-AG	BamHI and FseI were used to transfer the fragment carrying the 3'-end of H3f3a cDNA along with the 3xFLAG-tag and the IRESvYFP cassette from pBSK-KK-AA back into the corresponding sites of pKK-H3TG-AF. The construct was sequenced with KK25, KK36,and KK39.This is the 3rd targeting construct carrying IRESvYFP as well as the 3xFLAG-tag.

2.2.5.3 Cloning of retroviral plasmids

Plasmid Name	Cloning Strategy
pKK-ID-AA	pIDH2-Ex-C0463-M02 was digested with PstI and XhoI. The 1.3 kb fragment was inserted into the PstI/XhoI sites of pBSK.
pKK-ID-AB	pKK-ID-AA was amplified with KK05 and KK06 to introduce the R172K mutation using Q5 high-fidelity polymerase. The resulting fragment was digested with BstEII/NcoI and transferred to the corresponding sites of pKK-ID-AA. Correct introduction of the mutation was confirmed by sequencing with KK05.
pKK-ID-ACW	The polylinker KK07 and KK08 were inserted into the NotI/PstI sites of pKK-ID-AA to restore the full length IDH2 wild type cDNA. Correct insertion was confirmed by sequencing using primer pair M13.
pKK-ID-ACM	The polylinker KK07 and KK08 were inserted into the NotI/PstI sited of pKK-ID-AA to restore the full length IDH2 R172K mutant cDNA. Correct insertion was confirmed by sequencing using primer pari M13.
pMSCV-KK-1	The polylinker KK09 and KK10 were transferred into the EcoRI/BglII sites of pFG-AD.
pKK-ID-ADW	NotI/XhoI digest was used to excise the wild type IDH2 from pKK-ID-ACW and transfer it into the corresponding sites of retroviral vector pMSCV-KK-1.
pKK-ID-ADM	NotI/XhoI digest was used to excise the R172K mutant IDH2 from pKK-ID-ACW and transfer it into the corresponding sites of retroviral vector pMSCV-KK-1.
pKK-ID-AEW	The IRES-DsRed-Express2 fragment was excised from pMIBerry using EcoRI and SalI. It was inserted into the EcoRI/SalI sites of pKK-ID-ADW, thereby simultaneously deleting the NEO cassette.
pKK-ID-AEM	The IRES-DsRed-Express2 fragment was excised from pMIBerry using EcoRI and SalI. It was inserted into the EcoRI/SalI sites of pKK-ID-ADM, thereby simultaneously deleting the NEO cassette.
pKK-ID-AFW	pIDH2-IRESmCherry was digested with XhoI and HindIII to purify the IRES-mCherry fragment. The fragment was inserted into the XhoI/EcoRI sites of pKK-ID-AEW, replacing the IRES-DsRED-Express2 cassette.
pKK-ID-AFM	pIDH2-IRESmCherry was digested with XhoI and HindIII to purify the IRES-mCherry fragment. The fragment was inserted into the XhoI/EcoRI sites of pKK-ID-AEM, replacing the IRES-DsRED-Express2 cassette.

2.2.6 Mammalian cell culture methods

2.2.6.1 Trypsinization, cryopreservation and thawing of cells

Treatment with Trypsin (trypsinization) enzymatically dissociates adherent cells from the culturing dish/flask. First, the medium was removed and the cells were washed once with PBS to remove inhibiting agents present in the culture medium, especially FCS. The cells were then covered with a thin layer of Trypsin at the desired amount and potency (e.g. 2 mL of 0.05 % v/v Trypsin/EDTA per 10 cm plate), and incubated for 5 min at 37 °C. Afterwards, the reaction was

stopped by adding five volumes of cell culture medium. The detached cells were then harvested by pipetting, resuspended to remove cell aggregates, and used as required.

For cryopreservation cells were harvested and counted. The cell suspension was centrifuged for 5 min at 1,200 rpm at 4 °C. The supernatant was aspirated and the cells resuspended in their specific freezing medium (see section 0) at a final density of 5 x 10^6 cells/mL. 1 mL of cell suspension was then transferred into a 2 mL cryovial. To ensure a controlled cooling rate and thereby preventing cell loss, the cryovials were placed into pre-chilled freezing boxes with a cooling rate of -1 °C/min (Mr.Frosty, Nalgene). The boxes were then placed at -80 °C for one day and afterwards transferred to liquid nitrogen for long-term storage.

Thawing of the cells should be carried out as quickly as possible. Therefore, the cells were removed from the liquid nitrogen and placed in a 37 °C water bath. As soon as the suspension was thawed, a filtered Pasteur pipette was used to transfer the cells into 9 mL of the correlating medium. The vial was flushed once with the medium to collect remaining cells. The cells were then centrifuged for 5 min at 1,200 rpm. Subsequently, the cells were resuspended in fresh medium and plated at the desired density on the required plates.

2.2.6.2 Generation and culture of mouse embryonic fibroblasts

Mouse embryonic fibroblasts (MEFs) were used as feeder cells for embryonic stem cells (ES cells) and were previously isolated in our lab according to the following procedure:

After sacrificing pregnant Neomycin-resistant (Neo^R) female mice at 14.5 days post coitum, all procedures were carried out under sterile conditions using a laminar flow hood. The mice were disinfected with 70 % Ethanol. The uteri were then removed and placed into 15 mL sterile PBS in a 10 cm petri dish. All membranes and the placenta were removed from the embryos together with their heads, tails, limbs, and internal organs. To rinse off the remaining blood, the embryos were washed with fresh PBS and transferred into a sterile 10 mL beaker where they were minced using sterile scissors. 10 mL of 0.05% (v/v) Trypsin/EDTA were added, mixed by pipetting, and transferred to a sterile glass bottle. Glass beads and a magnetic stir bar were added and the mush was digested for 30 min at 37 °C under continuous stirring. Afterwards, 20 mL of MEF medium were added to stop the trypsinization. The suspension was then filtered through a 100 µm cell strainer into a 50 mL conical tube and adjusted to 20 mL MEF medium/embryo. 10 mL/10 cm plate of this cell suspension were plated and incubated in an incubator at 37 °C and 10 % CO_2. The next day, the medium was changed to remove dead cells and debris. As soon as a confluent monolayer was reached, the medium was removed, and cells were washed once with PBS and trypsinized. After stopping the reaction, cells were collected in a 50 mL conical tube and cryopreserved at a density of ½ embryo per cryo vial in 1 mL freezing medium.

Upon thawing, cells could be passaged up to five times at a 1:3 splitting ratio. Before their use in ES cell culture, MEFs were γ-irradiated with 38.5 Gy to stop the proliferation.

2.2.6.3 Maintaining embryonic stem cells

The ES cells used for gene targeting are R1 ES cells (129/SvJ, agouti) that were previously generated and described by Nagy et al. (1993). The cells were originally obtained from Dipl. Ing. Ronald Naumann (Max Planck Institute of Molecular Cell Biology and Genetics, Dresden).

One day before taking ES cells in culture, MEF feeder cells grown to a monolayer were harvested, irradiated at 38.5 Gy, and re-plated on 10 cm plates coated with 0.15 % gelatin (see section 2.1.6). The next day, ES cells were thawed and resuspended in 10 mL of ESC medium. 2 mL ES cell suspension were plated per 10 cm plate in a total of 10 mL ESC medium. Daily change of medium was required to maintain stem cell potential and inhibit differentiation. After reaching 60 – 70 % confluence, the cells were harvested by trypsinization and thoroughly pipetted to break up clumps. The cell suspension was transferred to a conical tube and incubated for 10 min at RT to allow for the settling of any remaining cell aggregates. Roughly 80 % of cell suspension were then transferred to a new tube and centrifuged for 5 min at 1,200 rpm. The supernatant was aspirated, the cells were resuspended in ES cell medium and split at a ratio of 1:10 and plated to irradiated MEFs or used for cryopreservation.

2.2.7 PCR clean-up

Before PCR products were used during cloning, they were cleaned up to remove any potential inhibiting agents. This procedure was also used to increase the concentration of DNA samples for their intended use. The indicated volumes are meant per 25 µL of PCR product/DNA solution.

The sample, together with 2 µL 3M NaAc and 1 µL Glycogen, was added to 95 µL of 100 % ethanol. After vortexing, the sample was centrifuged for 15 min at 14,000 rpm and 4 °C. The supernatant was discarded and the pellet was washed with 1 mL of 70 % ethanol followed by centrifugation under the previous conditions. Subsequently, the pellet was air-dried, resuspended in 20 µL nuclease-free H_2O, and stored at 4 °C.

2.2.8 Gene targeting

The transgenic H3K27M mouse was generated using classical gene targeting of ES cells. A simplified workflow is shown in Figure 11. The applied targeting strategy is based on one previously published by Tang et al. (2013). For further details on the generation of the targeting constructs refer to section 2.2.5.2. After cloning and verification of the construct, the targeting vector was electroporated into ES cells to enable integration into the host cell genome upon homologous recombination. Introduction of a neomycin resistance cassette allows for positive selection using G418, followed by expansion and characterization of promising clones. A detailed description of the individual processes is given in the following sections. Fully characterized cells were used for blastocyst injection and implantation into pseudo-pregnant female mice. These mice were set up for breeding to generate a stable transgenic mouse strain. Information on the breeding strategy is given in section 3.4.

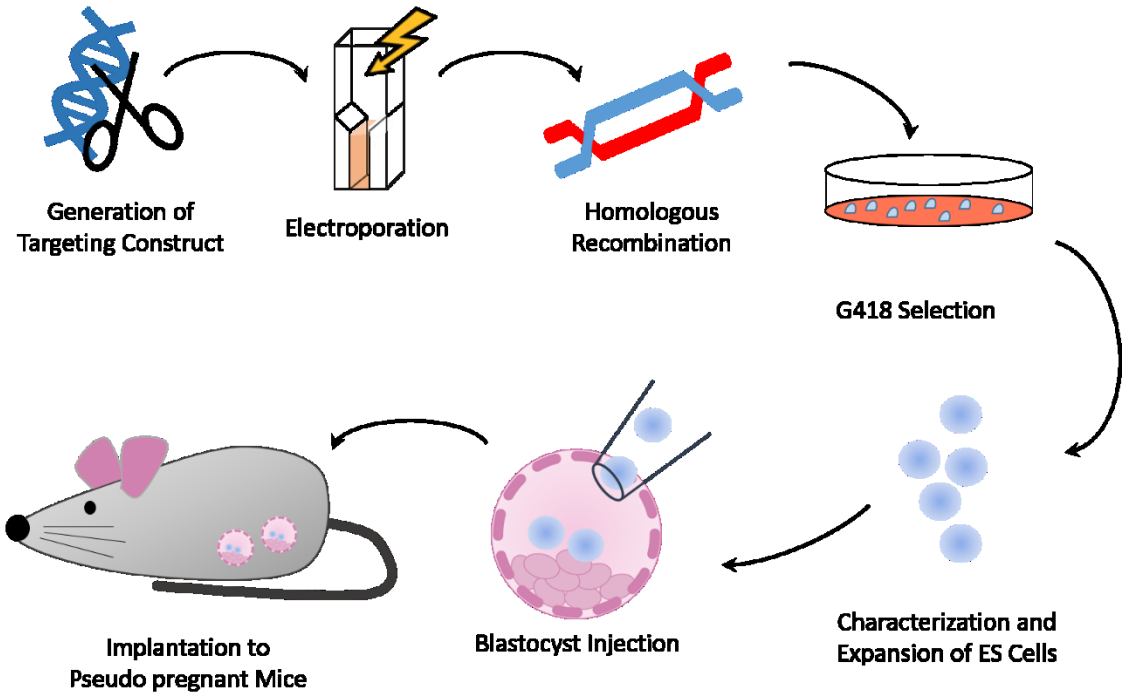


Figure 11: Simplified workflow gene targeting
First, the targeting construct was generated. The plasmid was introduced to the ES cell via electroporation and integrated in the cell genome through homologous recombination. Cells were selected using G418, expanded and characterized. Fully characterized cells were used for blastocyst injections. Blastocysts were then implanted into pseudo pregnant mice.

2.2.8.1 Vector linearization and preparation for electroporation

The targeting vector was introduced into the mouse genome via electroporation of ES cells. Homologous recombination requires linearized double-stranded DNA to be introduced into the cells. Therefore, the vector was linearized either with XhoI between the long arm of homology and the pBSK backbone, or EcoRI, between the short arm and the diphtheria toxin A (DTA) cassette. Restriction digestion was performed for 4 h in a 37 °C water bath using the following mix:

Reagent	Amount
Vector DNA	40 µg
Enzyme Buffer	40 µL
10X BSA	40 µL
Enzyme	4 µL
H_2O	Up to 400 µL

Afterwards, 5 µL of the digest were run on a 0.8 % agarose gel to check for complete digestion of the plasmid. The linearized targeting vector was then precipitated under sterile conditions as follows:

First, 900 µL sterile ethanol and 40 µL of 3 M NaAc were added to the remaining 395 µL vector digest. The solution was vortexed and incubated for 15 min at -20 °C followed by 10 min centrifugation at 13,000 rpm and 4 °C. Subsequently, the pellet was washed with 70 % ethanol

under the same conditions, air dried for 10 – 15 min, dissolved in 40 µL sterile TE buffer and immediately stored at 4 °C.

2.2.8.2 Electroporation of ES cells

The required amount of ES cells was cultured on irradiated MEFs as previously described. One-day prior to the electroporation, fresh MEF plates were prepared to seed the ES cells immediately after electroporation. 2 h before the procedure 10 mL of fresh ES medium were added to the cells.

The ES cells were harvested by trypsinization and washed once with PBS. Living cells were counted using a hemocytometer and trypan blue. For each condition 15 x 10^6 cells were transferred to a 15 mL conical tube. The conditions used were: **(i)** Linearized Vector, **(ii)** Electroporation using PBS only, and **(iii)** No Electroporation.

The cells were centrifuged at 1,500 rpm for 5 min at 4 °C. The supernatant was removed. Cells for condition **(i)** were suspended in 760 µL ice-cold PBS and 40 µL of the prepared DNA. Cells for **(ii)** and **(iii)** were resuspended in 800 µL ice-cold PBS only. After transferring the cells to pre-chilled gene pulser cuvettes (0.4 cm), they were incubated for 10 min on ice. Before electroporation, the cuvette was dried to ensure proper application of the current, and flicked to loosen cells at the bottom.

Electroporation was performed using a gene pulser at a capacity of 500 µF, a voltage of 230 mV, and an Extender of 500. Using a plugged Pasteur pipet, the electroporated cells were immediately transferred into 40 mL of fresh ES medium, plated on irradiated MEFs on four 10 cm plates, and incubated at 37 °C.

2.2.8.3 Geneticin (G418) selection and picking of recombinant ES cells

Upon introduction of the targeting vectors, markers for positive and negative selection were introduced as well. Negative selection of cells that randomly integrated construct together with its pBSK backbone was achieved by the DTA cassette without the requirement of additional factors. For positive selection of single cells that integrated the construct, Neomycin resistance was introduced to the cells. One day after electroporation, the regular ES cell medium was replaced by ES cell medium containing Geneticin (G418) at a concentration of 6.75 mg/mL.

After 10±1 days of G418 selection, the remaining cells were picked and transferred to a 96-well plate to enable screening for homologous recombination. All steps of this procedure were carried out under sterile conditions.

The plates were washed once with 10 mL PBS followed by the addition of 10 mL fresh PBS to the cells. Under a stereomicroscope, ES cell colonies were picked using a sterile, glass capillary attached to a mouth-controlled pipette. Morphology of the picked colonies was taken into account during picking. Only elevated, round colonies with clear margins and no flattened edges were chosen. Positive and negative examples are shown in Figure 12.The picked colonies were transferred individually to a 96-well round bottom plate containing 50 µL of 0.05 % Trypsin-EDTA/well. After 5 – 15 min, trypsinization was stopped by adding 150 µL of fresh ES cell medium.

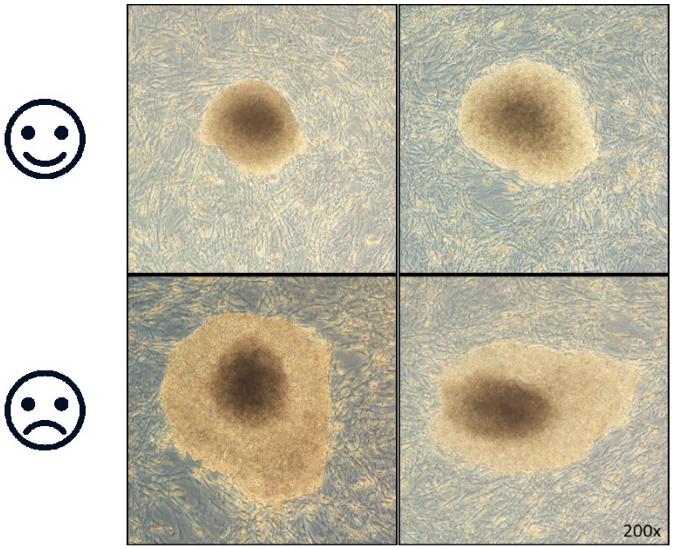


Figure 12: Targeted ES cell colonies before picking
Top row shows examples for ES cell colonies exhibiting the desired morphology (round, elevated, clear margins, no flattened edges). Bottom row shows negative examples. Pictures were taken at 200x magnification.

After stopping the trypsin reaction, the colonies were transferred into previously prepared 96-flat bottom plates (covered with 0.15 % gelatin and MEFs) using a multichannel pipette. The cells were then incubated at 37 °C and 10 % CO_2 until further use.

2.2.8.4 Establishment of a diagnostic PCR to screen for homologous recombination

A control construct (pFG-H3TG-CB) was cloned as described in section 2.2.5.1. It is ~10 kb in size, contains virtually no functional elements (as shown in Figure 13) and is randomly integrated into the cell's genome. Importantly, it carries the primer annealing sites for HL51 and FG16. This allows for reliable establishment of the diagnostic PCR that will be used to screen for homologous recombination in the targeted ES cells.

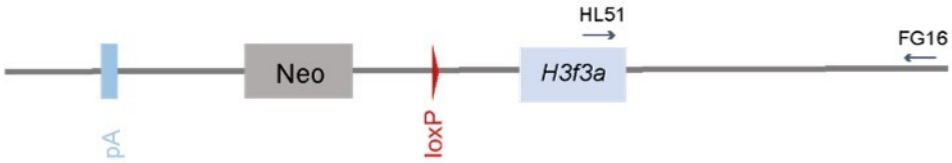


Figure 13: pFG-H3TG-CB control construct
The construct is roughly 10 kb and size. It carries an artificial splice acceptor (pA), Neomycin resitsance (Neo), and a primer annealing site containing fragment of the H3f3a gene. The location of both primer annealing sites (HL51 and FG16) is also indicated.

DNA from cells targeted with this control construct was isolated using the phenol/chloroform method. PCR was established using the Expand Long Template PCR System under the following conditions:

Reaction Mix		2X
Forward Primer	FG16	1.5 µL
Reverse Primer	HL51	1.5 µL
dNTPs		1.75 µL
Buffer 3		5.0 µL
Polymerase		0.75 µL
Rediload		5.0 µL
H_2O		32.5 µL
Total Volume		48 µL

PCR Conditions		
94 °C	2 min	
94 °C	10 sec	15x
56 °C	30 sec	
68 °C	2 min	
94 °C	15 sec	20x
56 °C	30 sec	
68 °C	2 min	
72 °C	7min	
4 °C	∞	

2.2.8.5 Propagation and screening of ES clones

As already stated, diagnostic PCR was performed to screen ES cells for homologous recombination. To reduce the number of samples, pools of 16 clones were screened first. Only for PCR-positive pools, the individual clones were screened.

One day before the actual experiment, 24-well plates were prepared and coated with MEFs. Five days after picking, the ES cells were split to collect cells for DNA extraction. To this end, the ES cells were trypsinized (washing with 200 µL PBS, 50 µL of 0.05 % Trypsin-EDTA/well.) The reaction was stopped by adding 100 µL of fresh ES cell medium. During incubation, the MEF medium was removed from the previously prepared plated and replaced by 1 mL of fresh ES medium supplemented with G418. Afterwards, the ES cells were carefully resuspended and 2 drops of cell suspension were transferred to MEF coated the 24-well plates. The remaining cells were pooled in 2 mL microcentrifuge tubes with cells from 16 individual clones per tube. Genomic DNA was isolated from each pool using the phenol/chloroform method described in section 2.2.1. Diagnostic PCR was performed under the conditions described above. Cells from negative pools were discarded. After two days, positive pools were harvested. Per clone, one drop of the cell suspension was transferred to a new MEF covered 24-well plate. The remaining cell suspension was used for DNA extraction and screening PCR as mentioned above.

Single clones that were identified positive for homologous recombination were grown until 60 – 70 % confluence. One half of the cells was frozen in liquid nitrogen as described in chapter 2.2.6.1, the other half was expanded to one 10 cm plate containing irradiated MEFs. After reaching 60 – 70 % confluence they were expanded to allow freezing of more cells as a backup. Per targeting construct, 2 – 3 clones were chosen at random to be expanded and excessively characterized as described in the following chapters.

2.2.9 Characterization of ES cell clones

Before sending the ES cell clones for blastocyst injection, they were carefully characterized. This served to ensure proper integration of the targeting cassette and correct expression of the introduced allele. Furthermore, Cre-mediated rearrangement to express mutant *H3f3a* together with the introduced markers was analyzed.

2.2.9.1 Preparation of vectors for electroporation

For functional analysis, recombinase expressing plasmids were introduced into the targeted ES cells. pCAGGS-flpe-VI was introduced to enzymatically remove the Neo cassette at the FRT sites. Activation of the K27M mutation is achieved by introducing Cre-recombinase into the cell by using piCre-AC. Both vectors are introduced via electroporation. Prior to electroporation, the vectors were precipitated and sterilized. Therefore, 20 µL of the vector were mixed with 380 µL H_2O, 900 µL sterile 100 % ethanol, and 40 µL sterile NaAc. The mix was vortexed and incubated for 15 min at -20 °C followed by centrifugation for 10 min at 4 °C and 13,000 rpm. The supernatant was discarded and the pellet was washed with 800 µL sterile 70 % ethanol. After a second centrifugation step under the same condition, the pellet was air-dried. The dried pellet was dissolved in 40 µL sterile TE buffer. The tube was closed with Parafilm and immediately stored at 4 °C.

2.2.9.2 Deletion of the Neomycin cassette and activation of K27M

Electroporation with the recombinase plasmids was performed as described in chapter 2.2.8.2, with only minor adaptions. The plasmids were not linearized, hence, the capacity was set to 250 Uf. Per clone 10 x 10^6 cells were electroporated. The pCAGGS-flpe-V plasmid introduces transient EGFP expression in clones that took up the plasmid and thereby are considered to have deleted thy neomycin resistance. After 24 h, the cells were harvested, sorted according to their EGFP expression and plated at a density of ~2000 cells/plate. The cells were incubated, picked, and transferred to a 96-well plate as previously described (see chapter 2.2.8.3). Four days after picking, the cells were harvested from the 96-well plate and transferred to 24-well plates in duplicates. After 24 h, the medium on one set of plates was replaced with ES medium containing G418 enabling the identification of non-resistant clones. The corresponding, non-G418 treated cells were then propagated.

The same selection procedure was applied for Cre-mediated activation of the mutation, as Cre-mediated recombination of the *loxP* flanked fragment includes deletion of the neo-cassette.

After identification of recombinant clones, they were expanded. Proteins, RNA, and DNA were extracted and used for characterization via Western Blot, FACS, and qRT-PCR, as described in the following chapters.

2.2.9.3 RT-PCR to characterize transcripts

For the characterization of targeted ES cells, two distinct RT-PCRs were performed on extracted RNA. The same RT-PCRs were also used during characterization of the newly generated transgenic mouse. Here, an additional RT-PCR for actin was performed as a positive control, e.g. to verify the presence of equal amounts of genetic material in all samples. The cDNA was synthesized from 0.5 µg RNA using the SuperScript III Kit.

The first RT-PCR serves as a control that both alleles are expressed. Due to the introduced elements, a size difference from the WT and from the targeted allele should be detectable. A schematic representation together with an exemplary results is given in Figure 19. The following running conditions were applied:

Reaction Mix		PCR Conditions		
KK22 (10 pm/µL)	1.5 µL	95 °C	2 min	
KK25 (10 pm/µL)	1.5 µL	94 °C	30 sec	
dNTPs	1.5 µL	57 °C	20 sec	30x
10X Buffer	5 µL	72 °C	30 sec	
$MgCl_2$	5 µL	72 °C	7 min	
AmpliTaq	1 µL	4 °C	∞	
Rediload	5 µL			
cDNA	2 µL			
H_2O	27.5 µL			
Total Volume	50 µL			

The second RT-PCR was performed to ensure full-length expression of the *H3f3a*-transgene. Additionally, these RT-PCR products were sequence to verify expression of the K27M mutation. To this extent, the CloneAmp PCR Premix was prepared six times per sample. The 692 bp product was purified from the gel using the peqGOLD Gel Extraction Kit. Sequencing was performed at GATC Biotech, Konstanz using primers KK27, KK39, and KK40.

Reaction Mix		PCR Conditions		
KK39 (10 pm/µL)	0.5 µL	98 °C	2 min	
KK40 (10 pm/µL)	0.5 µL	98 °C	10 sec	
CloneAmp Master Mix	12.5 µL	55 °C	15 sec	30x
cDNA	1 µL	72 °C	2 min	
H_2O	10.5 µL	72 °C	7 min	
Total Volume	25 µL	4 °C	∞	

RT-PCR for actin was performed under the following conditions:

Reaction Mix		PCR Conditions		
RT173 (10 pm/µL)	1.5 µL	95 °C	5 min	
RT174 (10 pm/µL)	1.5 µL	94 °C	30 sec	
dNTPs	1.5 µL	57 °C	20 sec	35x
10X Buffer	5 µL	72 °C	30 sec	
$MgCl_2$	5 µL	72 °C	7 min	
AmpliTaq	1 µL	4 °C	∞	
Rediload	5 µL			
cDNA	2 µL			
H_2O	27.5 µL			
Total Volume	50 µL			

2.2.10 Western Blot

2.2.10.1 Protein extraction

Extraction of total proteins from cells was done by using 1X RIPA buffer supplemented with c0mplete protease inhibitor (RIPA$^+$) at a ratio of 1 x 10^6 cells were mixed with 250 µL RIPA$^+$ buffer. The solution was incubated on ice for 30 – 60 min with several vortexing steps in between. After incubation, the cell suspension was centrifuged for 20 min at 13,000 rpm and 4 °C. The protein-containing supernatant was transferred to a new 1.5 mL microcentrifuge tube and immediately stored at -80 °C to prevent protein degradation.

For Western Blots performed on mouse cells, pure histones were extracted using the Histone Extraction Kit according to the manufacturer's manual.

2.2.10.2 Determination of protein concentration

The concentration of extracted proteins was determined using the Bradford assay. The protein-dye (Bradford) was diluted 1:5 with ddH_2O. Per sample one 2 mL tube was prepared to contain 1 mL diluted Bradford, 100 µL 150 mM NaCl, and 1 µL sample. Everything was mixed by inverting and incubated for 10 min at RT. Afterwards, the sample was transferred into plastic photometer cuvettes to measure the OD. Based on a BSA standard curve the protein concentration was calculated.

2.2.10.3 Sample preparation

The desired amount of protein was determined and mixed with 2X Laemmli and RIPA$^+$ buffer to a total volume of max. 30 µL per well. Afterward, the samples were denatured at 95 °C for 5 min and immediately loaded on the gel.

2.2.10.4 Gel casting

The gels were cast using the Mini-PROTEAN Handcast System from Bio-Rad. All parts were thoroughly cleaned with H_2O. Short- and spacer plates were additionally cleaned with 70 % ethanol to avoid any impurities on the glass. The system was assembled according to the instructions given by the manufacturer.

First, the resolving gel (see Table 7) was prepared in a conical tube with TEMED and APS being added last and the whole mixture was mixed by inverting and pipetting with a serological pipette. The gel chamber was filled up to the upper margin of the green holder. The gel was covered with of 1 mL of 100 % isopropanol. After complete polymerization, the isopropanol was removed and the gel was rinsed once with ddH_2O. The Remaining water was removed using Whatman paper. Afterwards, the stacking gel (see Table 8) was prepared. The gel chamber was filled to the top and the required combs were inserted.

Table 7: Composition of resolving gel, amount given for two gels

Percentage	**7.5%**	**10%**	**12%**	**15%**
30% Acrylamide	3.75 mL	5 mL	6 mL	7.5 mL
Gel Buffer	3.75 mL	3.75 mL	3.75 mL	3.75 mL
H_2O	7.5 mL	6.25 mL	5.25 mL	3.75 mL
TEMED	10 µL	10 µL	10 µL	10 µL
APS (10%)	80 µL	80 µL	80 µL	80 µL
Total Volume	15 mL	15 mL	15 mL	15 mL

Table 8: Composition of stacking gel, amount given for two gels

Percentage	4%	6%
30% Acrylamide	1 mL	1.5 mL
Gel Buffer	1.88 mL	1.88 mL
H_2O	4.63 mL	4.13 mL
TEMED	5 µL	5 µL
APS (10%)	40 µL	40 µL
Total Volume	7.5 mL	7.5 mL

After the polymerization of the stacking gel, the running chamber was set up and the gels were loaded with the freshly denatured protein samples.

2.2.10.5 Transfer and staining

Transfer of proteins from the gel to the PVDF membrane was done using the TransBlot Turbo system from Bio-Rad. Everything was performed according to the manufacturer's protocol. After transfer, the membrane was washed once for 5 min in TBS-T and incubated for at least 1 h with blocking buffer. Incubation with primary antibody was done in a 50 mL conical tube overnight at 4 °C on a roller. On the next morning, the membrane was removed from the antibody and washed three times for 5 min with TBS-T. Meanwhile, the secondary antibody was prepared. Incubation with the secondary antibody was performed for 1 h at RT. Prior to imaging the membrane was again washed for 3 x 5 min with TBS-T. Of note, for Western Blots of purified histones, 5 x 5 min washing in TBS-T was performed. Additionally, blocking was carried out in 5 % BSA instead of Chicken Egg Albumin.

All washing and incubation steps, except for overnight incubation with primary antibody, were performed at RT in a box on a rocking shaker at gentle motion.

2.2.10.6 Imaging and membrane stripping

Imaging of stained membranes was performed using the Fusion SL Advanced Imaging System by Viber Lourmat. Detection of the HRP-coupled secondary antibody was enabled by incubation of the membrane in 300 µL of Lumigen ECL Ultra solution.

If stripping of the membrane was required, the membrane was washed once with TBS-T and incubated for 7 min at RT on a rocking shaker in 10 mL stripping buffer. Afterwards, the membrane was again washed once with TBS-T and incubated for at least 1 h in blocking buffer.

2.2.10.7 Coomassie staining

Coomassie staining was performed to visualize total protein on the gels. It was performed either before transfer to check for the presence of all proteins, or after transfer to check for complete transfer of proteins to the membrane. All incubation and washing steps were performed in a box on a rocking shaker at RT.

The gel was incubated for 1 h in Coomassie Staining followed by several brief washing steps with methanol until all staining solution was removed. Finally, the gel was washed over night in ddH_2O. Using a light box, the gel was illuminated and photographed for documentation.

2.2.11 Generation of viral supernatants

Retroviral supernatants were used for the transduction of bone marrow cells. The supernatants were produced using Phoenix/Eco (Ph/Eco) cells transfected with retroviral expression plasmids.

24h before transfection, 5 x 10^6 Ph/Eco cells were plated on a 10 cm dish in 10 mL medium. The actual transfection was started on the afternoon of day one. The following quantities are meant per 10 cm plate.

First, the plasmids were prepared by mixing 10 µg of pCL-Eco helper plasmid and 5 µg retroviral expression plasmid in 1.5 mL pure DMEM in a 2 mL microcentrifuge tube. In a 15 mL conical tube, 3 µg TransIT/µg plasmid DNA were added (total of 45 µL) to 1.5 mL pure DMEM and incubated for 5 min at RT. After incubation the plasmid mix was transferred to the TransIT solution, immediately mixed, and incubated for 30-40 min at RT. Next, 12 mL of DMEM + 10 % of FCS w/o Pen/Strep were added to the transfection mix. The culture medium was removed from the Ph/Eco cells and replaced by the complete transfection mix. The cells were incubated overnight at 37 °C.

On the morning of day two, ~16 h after transfection, the medium was removed and replaced with 3 mL DMEM supplemented with 30 % FCS and 1% Pen/Strep. On day three the medium (= viral supernatant) was removed and pooled per condition in one conical tube and stored at 4 °C. 3 mL of fresh DMEM + 30 % FCS + 1 % Pen/Strep were added to the cells which were again incubated overnight at 37 °C.

On day four the second batch of viral supernatant was harvested and pooled with the one collected on day three. The supernatant was filtered through a 0.22 µM syringe filter. Aliquots of 1 mL were stored at -80 °C until use.

2.2.12 Determination of viral titer using antibiotic selection

The day before the actual experiment, NIH/3T3 cells were plated at a density of 2 x 10^5 cells/well in a 6-well plate.

On the first day, 20 mL of complete NIH/3T3 medium was equipped with polybrene to a final concentration of 12 µg/µL. The medium was used to prepare a 10-fold serial dilution in six 2 mL microcentrifuge tubes (tubes A – F). First, 1.35 mL of polybrene medium were added to each tube. 150 µL of viral supernatant are added to tube A. After gently mixing, 150 µL of the solution were transferred to the next tube. The process was repeated for all tubes. 1 mL of each tube was added to each well of NIH/3T3 cells to a total volume of 3 mL including one untransfected control. Afterward, the plates were centrifuged for 60 min at 1,200 x g and 32°C and incubated overnight at 37 °C.

On day two, the medium containing the viral supernatant was replaced by NIH/3T3 medium containing G418 at a final concentration of 6.75 mg/mL (1 mL G418 stock in 120 mL NIH/3T3 medium). This was not done for the cells transfected with mix A. On day three, these cells were used for FACS analysis of the expression marker. For all other cells, the G418 medium was replaced on day four and day six.

Depending on the condition of the cells, on day seven or eight the experiment was terminated. First, the medium was removed and the cells were carefully washed with PBS. 1 mL of crystal

violet solution (1 % in 10 % Ethanol) was added to the cells and incubated for 10 min at RT. 2 mL fresh PBS was added to the staining solution and immediately removed. Again, the cells were washed with fresh PBS and air-dried overnight. The following day, the viral titer was determined by counting the colonies and calculating the transducing units with the following equation:

$$\frac{number\ of\ colonies}{mL\ of\ initial\ supernatant\ mix} * dilution\ (10^{-x}) = TU/mL$$

2.2.13 RNA extraction

Various methods of RNA extraction were used depending on the desired purpose and the downstream application.

RNA isolated from a small number of cells, for example after sorting, was isolated using the Qiagen RNeasy Kit. Everything was done according to the manufacturer's protocol. Additional on-column DNA digest was performed using DNaseI.

For maximum recovery of RNA from mouse organs, the Qiagen RNeasy Midi Kit was used according to the provided protocol. Again, on-column DNA digest was performed using DNaseI.

In some cases, where maximum recovery was not required, RNA from organs was isolated using TRIzol according to the manufacturer's protocol.

2.2.14 RNA gel for quality check

After the isolation of RNA from organs, a denaturing agarose gel was run to determine whether or not the RNA was degraded during the process or subsequent handling. The procedure is based on the extended handbook of the Qiagen RNeasy Kits. In brief: 100 mL of a 1.2 % agarose gel was prepared from FA gel buffer (see section 2.1.6). After melting the agarose and allowing the mixture to cool down to ~65 °C, 1.8 mL of 37 % formaldehyde and 1 µL of 10 mg/mL ethidium bromide were added. Prior to running, the gel was eliquibrated for 30 min in 1X FA running buffer.

RNA samples were prepared by adding 1 volume of 5X RNA loading buffer to 4 volumes of RNA sample. RNA was denatured for 5 min at 65 °C, chilled on ice, and then loaded on the gel. The gel was run for 90 min at 60 V.

Intact RNA would show the 28S and 18S subunits at roughly 2:1 ratio. DNA contamination would be detectable as band a distinct band of much bigger size. Loss of RNA integrity is indicated by smearing and complete loss of bands.

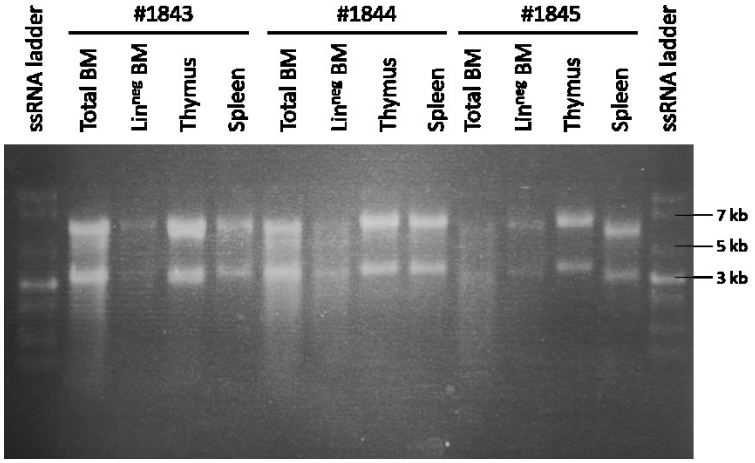


Figure 14: Example for an RNA quality analysis using agarose gel electrophoresis
Pictured are the results obtained from three C57BL/6J mice used for qRT-PCR presented in Figure 28. The 28S and 18S subunits are detectable in all samples, with slight hints on RNA degradation in total BM samples, as indicated by smearing. No DNA contamination can be seen.

2.2.15 Quantitative Real Time PCR (qRT-PCR)

2.2.15.1 cDNA synthesis

cDNA synthesis was performed using the SuperScriptIII First Stand Synthesis Kit from Invitrogen. The procedure was performed according to the protocol provided by the manufacturer. Typically, 0.5 or 1 µg RNA were used as starting material. If no on-column DNase digest was performed during RNA isolation, DNaseI digest using the Invitrogen Kit was done immediately before RNA synthesis.

2.2.15.2 qRT-PCR set-up and running conditions

All pipetting steps were carried out on ice with minimal surrounding light. The TaqMan gene expression assays, containing forward and reverse primers as well as a fluorophore labeled probe, were thawed at RT in the dark. All TaqMan assays used during this project were FAM coupled and Actin was used as endogenous control. For all reactions, samples were pipetted in triplicates in a 96-well reaction plate. For each gene, one triplicate with cDNA synthesized without reverse transcriptase and one negative control with no template was included. Prior to the experiments, cDNA dilution series were run to determine the required cDNA concentration to obtain C_T values between 20 and 30.

The reaction mix per well was composed of 2 µL template,1 µL TaqMan, 10 µL Gene Expression Master Mix, and 7 µL nuclease-free H_2O. After pipetting, the plate was closed with optical adhesive film and centrifuged for 2 min at 1,200 rpm.

qRT-PCR was run using the StepOne Plus System from Applied Biosystems. For all approaches, a Quantitation Comparative Experiment ($\Delta\Delta C_T$) was run using the standard ramp speed (~2 h) and the following cycling conditions:

Holding Stage:	50 °C	2 min
	95 °C	10 min
Cycling Stage (40x):	95 °C	15 sec
	60 °C	1 min

2.2.16 Flow cytometry

2.2.16.1 Preparation of cell suspensions

The mice were sacrificed using CO_2. Immunological organs were harvested and, after removing adjacent connective tissue, placed in 3.5 cm dishes filled with 1 mL of FACS buffer (5 % FCS in PBS). Spleens were homogenized between two frosted ends of microscopic slides; the slides were rinsed with one additional milliliter of FACS buffer. Everything was resuspended and filtered in one 2 mL microcentrifuge tube. Thymus and lymph nodes (LN) were meshed between two filters that were rinsed with 1 mL of FACS buffer, resuspended, and filtered into one 2 mL microcentrifuge tube. The harvested bones were cut open at both epiphysis, and flushed with 1 mL FACS buffer using a 23G needle and a syringe. The bone marrow (BM) was carefully resuspended and also filtered in a 2 mL microcentrifuge tube. During all steps, the organs as well as the cell suspensions were kept on ice.

For BM and spleen cell suspension erythrocyte lysis was performed. Therefore, the cells were centrifuged for 5 min at 2,000 rpm and 8 °C and resuspended in 400 µL of lysis buffer (see section 2.1.6). BM cells were incubated at RT for 3 minutes, and spleen cells for 5 minutes. After incubation, the reaction was stopped by filling the tube with FACS buffer. After subsequent centrifugation under the same conditions and removal of the supernatant, the cells were resuspended in 2 mL of fresh FACS buffer and filtered into a new tube to remove debris. The cell count for all cell suspensions was determined by using a Neubauer counting chamber and trypan blue solution.

2.2.16.2 Staining protocol

After counting, the required amount of cells is transferred into a 96-well microtiter plate. For most stains 1 x 10^6 cells were used. Only if cells within the lineage-negative compartment were analyzed, for example during stem cell precursor stainings in the BM, 3 x 10^6 cells were used. The standard staining volume was 100 µL. If 3 x 10^6 cells were stained, the volume was increased to 200 µL. Again, all steps described were carried out on ice.

First, F_C block was performed to prevent unspecific binding. In most cases, this was done by incubation for 30 min with unlabeled CD16/32 antibody. If CD16/32 (FcR III/II) was stained, purified mouse IgG was used instead. After F_C block the cells were washed once in FACS buffer by filling up the wells and centrifugation for 5 min at 1,500 rpm at 4 °C. The supernatant was removed and the cells were resuspended in the corresponding antibodies diluted in FACS buffer at the concentration given in section 2.1.11.1. Incubation was performed in the dark on ice for 30 min. Afterwards, the cells were washed again. If required, a secondary antibody was added using the same procedure. After the final staining, the cells were washed twice in FACS buffer before being resuspended in 200 µL of the live/death stain Sytox blue.

If required, enrichment of lineage-negative cells (Lin depletion) was performed on BM prior to staining. A detailed description of the process is given in section 2.2.17.2.

2.2.16.3 Intracellular staining

Staining of intracellular markers such as trimethylated histones, or the 3xFLAG tag was performed using the Foxp3/Transcription Factor Staining Buffer Kit according to the protocol. In brief: the cells were dyed with a fixable viability dye, surface staining was performed as described in the previous section and the cells were permeabilized using the kit. Afterwards,

staining of the intracellular markers was performed. Here, the antibodies were diluted in permeabilization buffer provided with the kit and incubation was performed for 30 min at RT in the dark. After permeabilization, all washing steps were done using permeabilization buffer and centrifugation at 800 rpm and RT. For analysis, the cells were resuspended in FACS buffer. Live/death staining was performed using Fixable Viability Dye instead of Sytox.

2.2.17 Competitive adoptive bone marrow transfer

2.2.17.1 Preparation of RetroNectin coated plates

Recombinant fibronectin (RetroNectin) coated plates were used to enhance the efficiency of retroviral transductions. RetroNectin was reconstituted with sterile PBS to a stock concentration of 1 mg/mL and aliquots were stored at -20 °C. Before coating the plates, the RetroNectin stock was thawed and diluted in sterile PBS to a working solution of 12 µg/mL (for instance 6 mL PBS + 72 µL RetroNectin). A non-treated 24-well plate intended for suspension cells was filled with 250 µL RetroNectin/well. At this stage, the plates can either be stored overnight at 4 °C if covered in plastic, or incubated at RT for 2 h.

After incubation, the RetroNectin was removed and the plates are blocked for 30 min at RT using 300 µL of 2% BSA (30% stock solution diluted in sterile PBS). Subsequently, the plates were washed twice with PBS. Again, the plates can either be used immediately or filled with fresh PBS and stored at 4 °C.

2.2.17.2 Retroviral transduction of bone marrow cells

On day one, the required mice were sacrificed, the hind legs were removed and placed in FACS buffer. All following procedures were performed under sterile conditions using a laminar flow hood.

The intact bones were briefly washed in 70 % ethanol. Afterwards, the bones were transferred to sterile FACS buffer. BM cells are prepared according to the procedure described in 2.2.16.1. Erythrocyte lysis was performed using 400 µL filter sterilized erythrocyte lysis buffer. The cells were then counted and lineage-positive cells were depleted using the EasySep™ Mouse Hematopoietic Progenitor Cell Isolation Kit from Stem Cell according to the manufacturer's protocol. The remaining cells were counted and seeded in complete BM medium at a density of 2 x 10^6 cells/mL in a 6-well plate.

The next day, previously prepared RetroNectin plates were coated with viral supernatant by adding 0.5 mL supernatant/24-well followed by centrifugation for 2 h at 2,000 x g and 22 °C. Afterward, the plates were washed once with PBS. Now BM cells were added to the freshly virus-coated plates at a density of 0.5 x 10^6 cells/well on a 24-well plate. The plates were then centrifuged for 5 min at 2,000 x g at 22 °C and incubated overnight.

On the third day, after 24 h, the cells were harvested, washed once with PBS, and re-plated in fresh 24-well plates.

On day five the cells were harvested and used for the desired experiments, such as BM transplantations.

2.2.17.3 Irradiation and transplantation of mice

Before injection, the mice were lethally irradiated with a split dose of 7+4 Gy using the GSR D1 Irradiator and Caesium-137 as nucleotide to ablate the host hematopoietic system, with three house between both radiations. Afterwards, the mice were transported into our mouse room to rest for 2 – 3 h before the actual transplantation. In the meantime, the donor cells were harvested, counted, resuspended in pre-warmed, fresh PBS, and filtered into a sterile microcentrifuge tube. Additionally, competitor mice were sacrificed and the BM was harvested under sterile conditions as described in the previous section. Per mouse, a total of 4 x 10^5 retrovirally transduced cells together with 5 x 10^5 WT competitor cells were injected in 200 µL volume. The Injection was performed using 29G Insulin syringes. For a period of one week before, and two weeks after irradiation and transplantation, the mice were fed with water containing neomycin at a concentration of 1.1 g/L to prevent death due to sepsis.

2.2.17.4 Monitoring of transplanted mice

The health status of all mice was checked daily. Weekly, the mice were weighed during changing of the cages. Starting one month after transplantation, blood was drawn from the tail vein once a month via punctuation and collected in EDTA tubes. The blood was used for FACS analysis as well as blood analysis using the automated cell counter Hemavet.

3. Results

3.1 Introduction of a Cre inducible K27M mutation into the endogenous H3f3a gene locus of embryonic stem cells via gene targeting

To be able to assess the potential pathophysiological effects of the H3K27M mutation under *in vivo* conditions, we set out to generate a Cre-recombinase inducible mouse model via classical gene targeting in murine embryonic stem cells (ES cells). To this end, we first had to construct an appropriate targeting vector, based on a strategy previously published by Tang et al. (2013). The published vector was generously provided by the Jeffrey Mann Laboratory (The Royal Children's Hospital, Victoria, Australia) and modified in several cloning steps (for details see section 2.2.5.2).

The first vector we have generated (pKK-H3TG-AD) contains an insert of ~11.5 kb, whose central part is supposed to be inserted into exon 2 of the *H3f3a* gene locus upon homologous recombination, as depicted in Figure 15. This central DNA segment carries the following elements:

First, an artificial splice acceptor (SA) and a single *loxP* site, followed by a cDNA encoding wild-type *H3f3a*. Of note, this cDNA carries at most silent codon positions sequence alterations (wobble bases) that allow for the distinction between transcripts from the endogenous, unmodified *H3f3a* locus and the inserted WT cDNA. Introducing these silent codon alterations was necessary to reduce the degree of homology between the *H3f3a* coding cassette and the 3' sequence of the construct, thus preventing unwanted recombination events and thereby loss of upstream sequences. All sequences derived from the cDNA with altered wobble bases will from now on be denoted as *H3f3a-WT**** (marked with three asterisks). The *H3f3a-WT**** cDNA is followed by an SV40-derived polyA transcription termination signal (pA). Therefore, in the absence of recombinase, only the *H3f3a-WT**** will be transcribed. The next element is a FRT flanked Neomycin (Neo) resistance cassette, which serves as a marker for positive selection during gene targeting. If desired, the Neo resistance cassette can be deleted upon ectopic expression of a Flp-recombinase. The FRT-flanked Neo element is followed by a second *loxP* site, a cDNA encoding the mutated *H3f3a-K27M*, and another pA signal. Of note, the *H3f3a* sequence introduced here carries no altered wobble bases and thus differs from the *H3f3a* WT only at the mutated K27M codon position. Expression of Cre-recombinase is expected to induce loxP-mediated deletion of all sequences between the *loxP* sites replacing the *H3f3a-WT**** as well as the Neo-cassette with *H3f3a-K27M*, thus activating transcription from the mutant locus.

The described central elements are flanked by a long (5.5 kb) and a short (2.0 kb) arm of homology to promote homologous recombination of the targeting construct within the endogenous *H3f3a* gene locus of the ES cells. Adjacent to the short arm is a DTA cassette encoding diphtheria toxin A for negative selection of the majority of cells that will insert the targeting vector randomly without homologous recombination. The DTA cassette is protected from nonspecific, terminal degradation by the flanking pBSK backbone of the targeting construct.

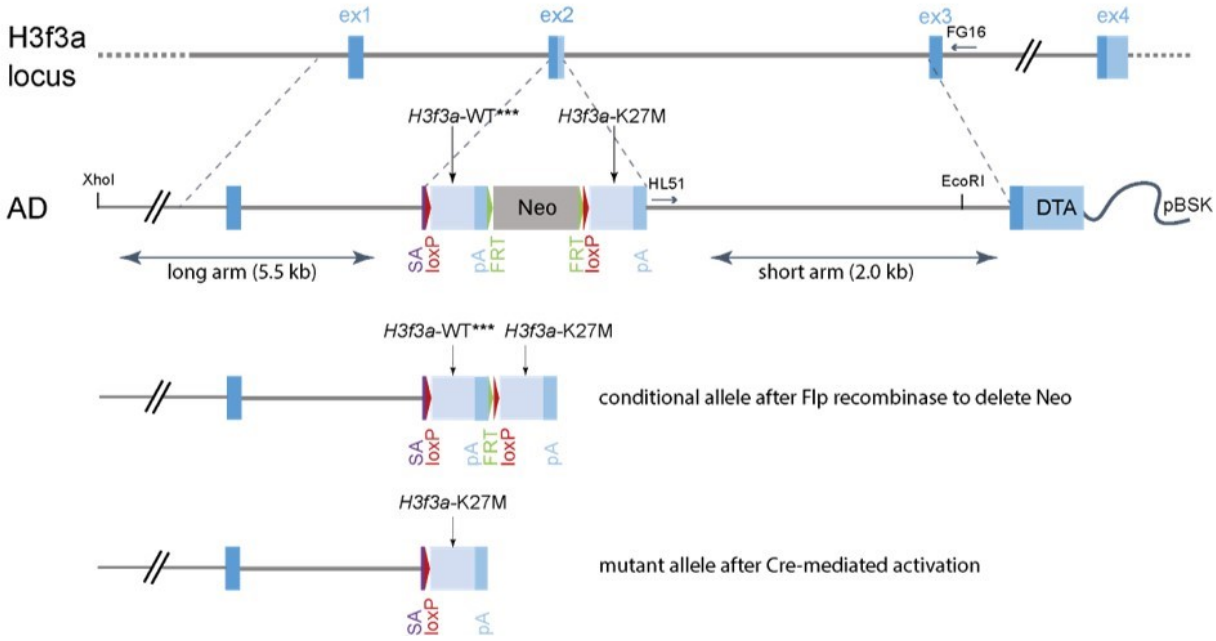


Figure 15: Targeting strategy (vector pKK-H3TG-AD)
Structure of the *H3f3a* gene locus and the first targeting vector AD (pKK-H3TG-AD). The vector carries a synthetic splice acceptor (SA) followed by *H3f3a-WT*** with silent codon alterations, and a Neomycin resistance cassette (Neo) flanked by FRT sites. The Neo cassette is used for positive selection and can be removed upon exposure to Flp recombinase giving rise to the allele shown in the third row. Cre-mediated recombination of the *loxP* sites results in simultaneous excision of *H3f3a-WT**** and Neo, leading to transcription of *H3f3a-K27M* (allele shown in the last row). The final elements are a diphtheria toxin A cassette (DTA) used for negative selection and the pBSK vector backbone. Depicted are also the location of primers FG16 and HL51 used to screen for homologous recombination events, and the location of enzymes XhoI and EcoRI used alternatively for linearization. ex = exon; pBSK = pBluescript vector

Homologous recombination of the targeting vector with the endogenous *H3f3a* locus is enhanced by linearization of the vector. The construct can be enzymatically digested either with EcoRI, between the short arm of homology and the DTA cassette or with XhoI, between the long arm of homology and the pBSK backbone (see Figure 15 for precise position of sites). To determine which strategy yields better results, alternate linearization with one or the other enzyme was performed. The linearized constructs were introduced into ES-R1 embryonic stem cells (Nagy et al., 1993)via electroporation. After the selection of transfected ES cells with G418, resistant colonies were picked and expanded (for technical details see section 2.2.8.3).

To screen for homologous recombination events, a diagnostic PCR using primers HL51 and FG16 was performed on genomic DNA isolated from G418-resistant clones. The locations of both primers are shown in Figure 15. Initially, PCR was performed with several pools comprising 16 clones each. ES cell clones of the PCR-positive pools were then individually screened to identify individual positive clones. Exemplary PCR results of the first targeting experiment with pKK-H3TG-AD are shown in Figure 16. Linearization of the targeting constructs was achieved with EcoRI or XhoI respectively. For each linearization strategy, a total of 384 clones were screened for homologous recombination. Linearization with EcoRI resulted in two positive clones equaling a frequency of homologous recombination of approximately 0.5 %. Targeting with the XhoI digested construct resulted in 14 distinct positive clones corresponding to a frequency of about 3.6 %.

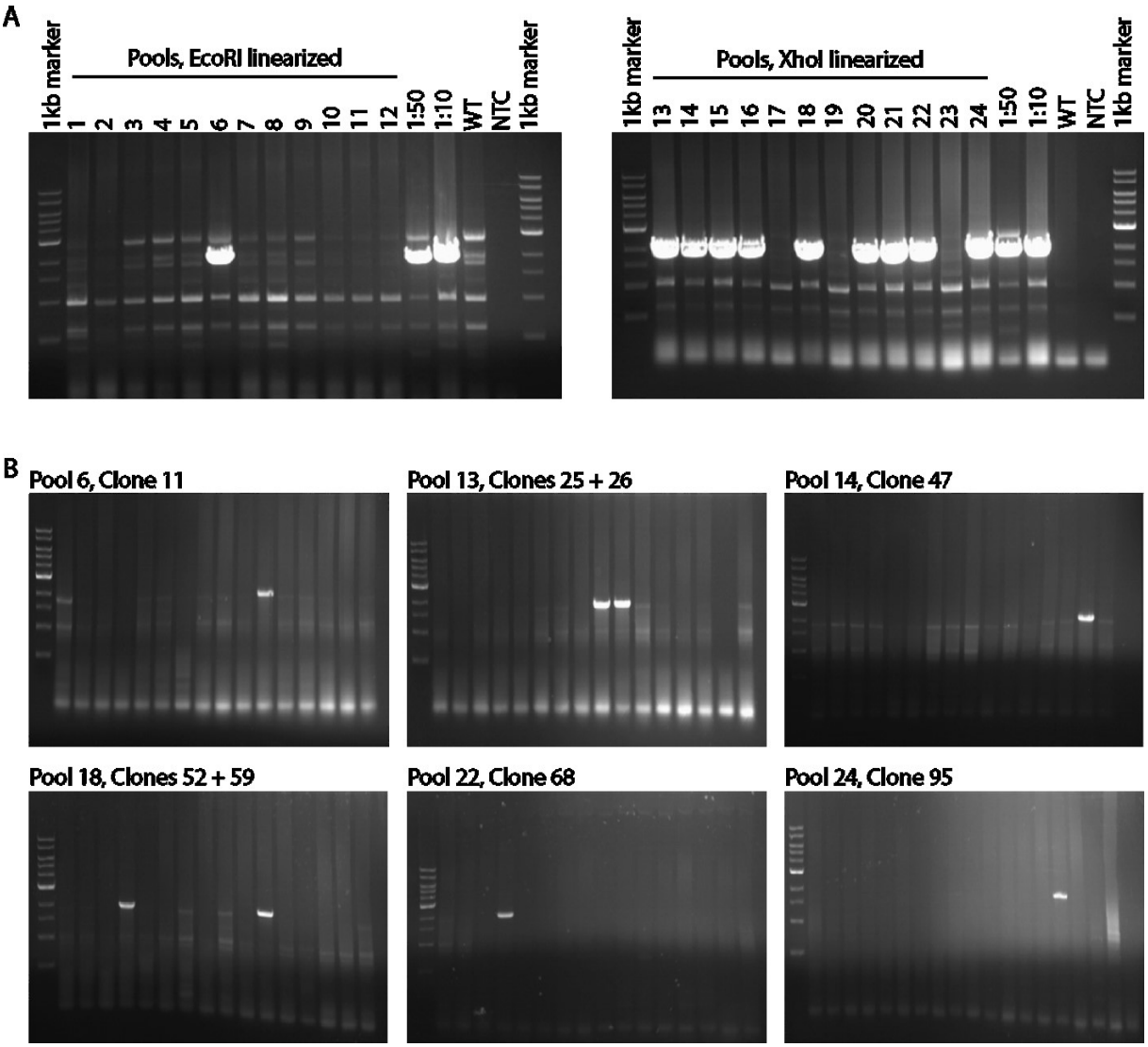


Figure 16: Screening for homologous recombination after targeting with vector pKK-H3TG-AD
(A) Results of the pool screening. The left picture shows pools 1-12 resulting from targeting after linearization of the vector with EcoRI. The right picture shows pools 13-24 after XhoI linearization. Samples marked with 1:10 and 1:50 refer to diluted gDNA obtained from a previous targeting with a control construct performed by Suresh Kumar (see methods section 2.2.5.1). **(B)** Exemplary results of screening individual clones. Only numbers of the positive clones are given. WT = wild-type (non-manipulated) ES cells; NTC = no template control

Successful targeting of the *H3f3a* locus with construct pKK-H3TG-AD encouraged us to incorporate two additional, useful sequence elements into the vector. While significantly increasing the non-homologous DNA section to be inserted into the target locus and thus possibly reducing targeting efficiency, this modification would substantially widen the range of applicable methods to the transgenic mouse, as detailed below.

3.2 Targeting of the H3f3a gene locus with more sophisticated targeting vectors

After successful gene targeting with pKK-H3TG-AD, two additional, more complex constructs were generated. Subsequent targeting experiments were performed under the same conditions as before, however, this time using exclusively XhoI for linearization, which had yielded the highest recombination frequency with vector pKK-H3TG-AD.

The improved, second targeting vector (pKK-H3TG-AF) carried the same elements as vector AD, but additionally, an internal ribosomal entry site (IRES) followed by a cDNA encoding the fluorescent marker venusYFP (vYFP) (Nagai et al., 2002). This marker would be expressed together with the *H3f3a-K27M*-encoded mutant histone from a bi-cistronic transcript only after Cre-mediated recombination of the *loxP* sites. Expression of vYFP would allow non-invasive detection of individual living cells when expressing the targeted *H3f3a-K27M* locus simply by flow cytometry or fluorescence microscopy. The third and final targeting construct pKK-H3TG-AG also carries the IRES-driven vYFP marker but in addition, a stretch of 30 in-frame codons at the 3' end of the *H3f3a-K27M* cDNA providing a carboxy-terminal 3xFLAG-tag, which serves an important purpose: as mentioned in the introduction, the H3.3 protein is encoded by two genes, *H3f3a,* and *H3f3b*. The 3xFLAG-tag allows the discrimination between the WT and the K27M mutated H3.3 at the protein level, for example in Western Blots. A schematic comparison of all three targeting constructs is given in Figure 17.

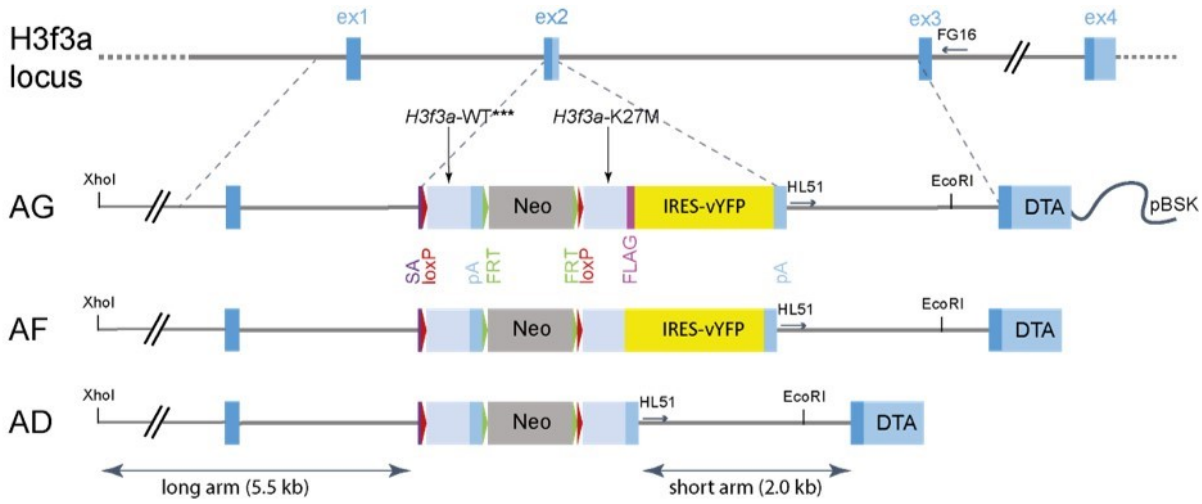


Figure 17: Structure of the targeting vectors used and comparison of targeting results
Schematic representation of the three targeting constructs. In constructs AF and AG, the main features remain unaltered in comparison to the previously described construct pKK-H3TG-AD. However, pKK-H3TG-AF additionally carries an internal ribosomal entry site (IRES) followed by the fluorescent marker vYFP that is expressed together with *H3f3a*-K27M after Cre-mediated recombination. pKK-H3TG-AG carries the IRES-driven fluorescent marker as well as a 3xFLAG-tag. AD = pKK-H3TG-AD; AF = pKK-H3TG-AF; AG = pKK-H3TG-AG; SA = synthetic splice acceptor; pA = SV40 polyA; Neo = Neomycin; DTA = diphtheria toxin A; pBSK = pBluescript vector

The frequency of homologous recombination obtained with in the respective targeting experiments is given in Table 9. Of note (and opposite to expectation), the highest frequency of homologous recombination (~21%) was achieved with the most complex targeting construct pKK-H3TG-AG harboring the 3x FLAG-tag as well the IRES-vYFP cassette.

Table 9: Comparison of homologous recombination frequencies observed with the three distinct targeting constructs
For pKK-H3TG-AD only the results after linearization with XhoI are given. AD = pKK-H3TG-AD; AF = pKK-H3TG-AF; AG = pKK-H3TG-AG

Targeting construct	Screened colonies	Positive clones	Frequency of homologous recombination
AG	560 in pools 48 single	10	21 %
AF	384	5	1.3 %
AD	384	14	3.6 %

3.3 Characterization of targeted ES clones

Several targeted ES clones obtained with constructs pKK-H3TG-AF and pKK-H3TG-AG were randomly selected for rigorous molecular evaluation to verify **(i)** correct Cre-inducible expression of the K27M-mutated *H3f3a* allele, **(ii)** concomitant, Cre-dependent expression of the fluorescent marker vYFP and – in the case of ES cell targeted with construct pKK-H3TG-AG – **(iii)** the possibility to detect the K27M-mutant protein with an anti-FLAG antibody. Only targeted ES cell clones, which pass all these criteria, would qualify to be used for the generation of corresponding mice. A complete set of results obtained in these comprehensive evaluation experiments is presented in Figure 18 - Figure 20 with a focus on pKK-H3TG-AG targeted ES clone #10 (and its sub clones), which was subsequently utilized for generating the H3K27M mouse strain employed and analyzed in this thesis project. Experiments equivalent to the ones shown were also performed for several other clones targeted with pKK-H3TG-AF and pKK-H3TG-AG, respectively, and delivered similar results (data not shown).

Figure 18 documents strong IRES-vYFP marker expression after Cre-mediated activation of the H3f3a-K27M-IRES-vYFP segment in sub clone AG #10a, fully in line based on the design of targeting vector pKK-H3TG-AG (Figure 17). Of note, neither the targeted clone AG #10 (before Cre-expression), nor sub clone AG #10.1, in which the Neo-cassette had been deleted upon ectopic Flp expression, exhibited any yellow fluorescence above background (Figure 18A, the red histogram represents IRES-vYFP expression in non-manipulated WT ES cells). This demonstrates that there is no significant, read-through transcription across the pA signal located immediately downstream of *H3f3a-WT****. As IRES-driven vYFP is translated from a bi-cistronic mRNA, marked expression of vYFP strongly suggested corresponding expression of the K27M-mutated *H3f3a* cDNA segment. These findings have been verified in four independent sub clones of AG #10 (AG #10.a – 10.d) as shown in Figure 18B.

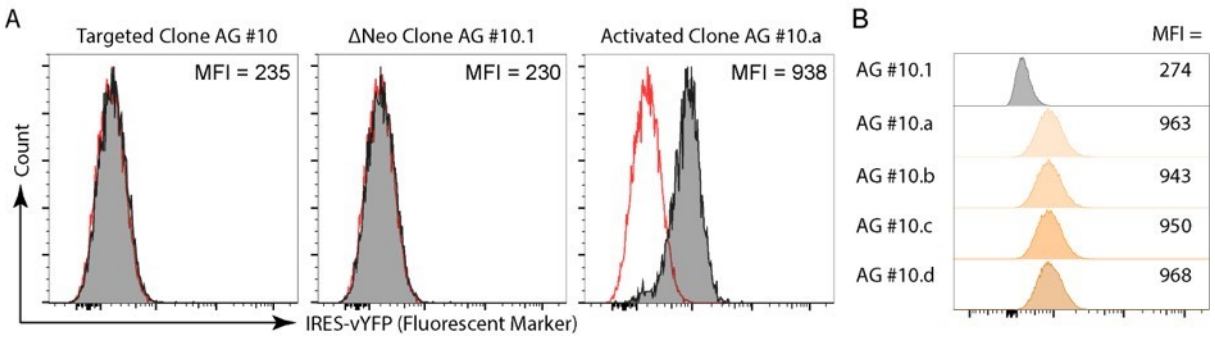


Figure 18: Flow cytometric analysis of targeted ES cells
(A) Expression of the fluorescent reporter IRES-vYFP can only be detected upon Cre-mediated activation. Red outline represents IRES-vYFP fluorescence in genetically non-manipulated (WT) ES cells. **(B)** Stacked histograms directly compare reporter expression after Neo-deletion (AF #10.1, grey) to four independent Cre-activated sub clones (AF #10.a – d; shades of yellow).The values given in all histograms indicate the mean fluorescence intensity (MFI) of the reporter

H3f3a-K27M expression after Cre-mediated activation of the targeted allele was confirmed by RT-PCR (Figure 19A and B). The presence of a 427 bp product exclusively in Cre-activates sub clones indicates robust transcription of the recombined allele encoding *H3f3a-K27M*. Although PCR conditions were not optimized for quantitative comparisons, the fact that PCR bands from the targeted allele (427 bp) and the wild-type allele (361 bp) had roughly similar intensities in all Cre-activated sub clones was a good sign seeming to suggest that the targeted allele is expressed as robustly as the wild-type allele at transcript level.

These findings and interpretations were corroborated by Western blotting experiments with commercially available antibodies directed against total histone variant H3.3, vYFP and the FLAG epitope (Figure 19C). The strong band obtained with anti-FLAG antibodies in all Cre-activated sub clones, like AG #10.a shown in Figure 19C, provided conclusive evidence for expression of the FLAG-tagged *H3f3a-K27M* allele also at the protein level. Moreover, this experiment confirmed the usefulness of the 3XFLAG-tag inserted in-frame at the 3'-end of the *H3f3a-K27M*-encoding cDNA to unequivocally identify K27M-mutated H3.3A protein, as the anti-H3.3 antibody would detect identical protein products from both the *H3f3a* and *H3f3b* genes (see introduction chapter 1.6.3).

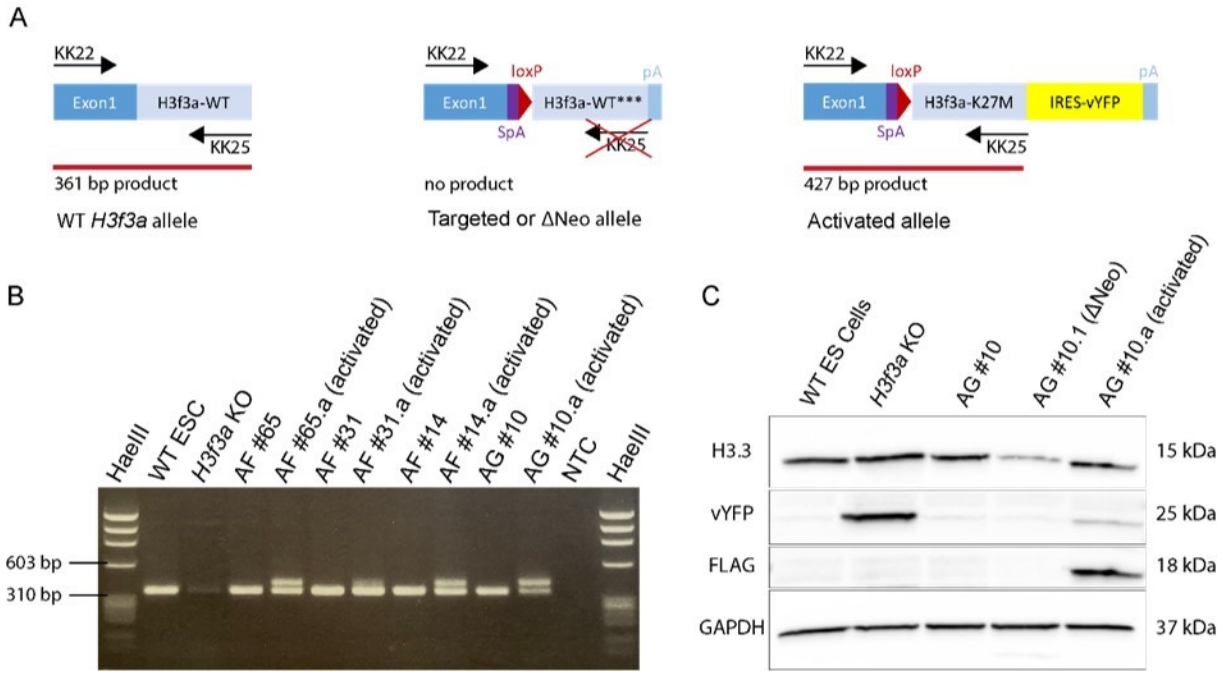


Figure 19: Verification of H3f3a-K27M expression using RT-PCR and Western Blot
(A) Schematic representation of RT-PCR performed with primers KK22 and KK25 to verify the expression of *H3f3a* transcripts. The primers bind the endogenous WT allele as well as the mutated allele. They do not bind the altered wobble bases of *H3f3a-WT*** cDNA. Amplification of the endogenous WT allele results in a 361 bp product. As targeting introduces the mutation only heterozygous, a product from the WT allele is expected in all samples. Due to altered wobble bases and thus lack of an annealing site for KK25, there is no product from the second allele in targeted or Neo deleted clones. Only after Cre-mediated recombination, the primers can anneal to the second (now K27M-mutatetd) allele and give rise to a product of 427 bp. **(B)** RT-PCR shows robust expression of the targeted allele upon Cre-mediated recombination in clones targeted with pKK-H3TG-AF and –AG. **(C)** Western Blot of 20 µL total protein extracts from ES cells. The same membrane was incubated with H3.3, GFP/YFP, FLAG, and GAPDH antibodies in that order. Stripping and blocking was performed between each incubation step. Cre-activated clone #10.a expresses both markers (IRES-YFP and FLAG). YFP expression is reduced as compared to the *H3f3a-KO*. Indicated on the right are the expected protein sizes. HaeIII = Marker, NTC = no template control; WT = wild-type; KO = knock out

The *H3f3a-KO* ES cell clone used in the experiments shown in Figure 19B and C was obtained as additional control also from the Jeffrey Mann Laboratory. These gene-targeted cells carry a vYFP expression cassette in both *H3f3a* gene loci, resulting in a bi-allelic *H3f3a* knock-out, as described in Tang et al. (2013). Exhibition of the vYFP fluorescence that was indicated during Western Blot (Figure 19C) was confirmed by flow cytometry (data not shown).

The weak band observed in RT-PCR on *H3f3a-KO* cells, potentially indicating *H3f3a* transcripts (Figure 19B), most likely arises from the experimental conditions rather than the ES cells, as the band was not observed in other, equal, experiments. The simplest explanation would be a spill over during loading of the gel or contamination during sample preparation. Despite the efforts to adapt the ES cells to feederless growth conditions prior to RNA extraction, they were originally co-cultured with embryonic fibroblasts. The observed *H3f3a* transcript might be attributed to remaining feeder cells. In contrast, the H3.3 signal observed in H3f3a-*KO* cells during Western Blot can easily be explained and further underlines the importance of the FLAG-tag. As mentioned in the introduction, H3f3a and H3f3b proteins cannot be readily

distinguished from each other, hence the observed band can be attributed solely on H3f3b proteins.

The Western Blot shown in Figure 19C further depicts a reduced H3.3 expression in the ΔNeo clone AF #10.1. The reason for this is unknown since, again, this inconsistency was not observed in other experiments. However, the observed size difference between H3.3 (15 kDa, with a detectable size of 15 – 17 kDa) and the FLAG-tag (additionally ~3.5 kDa) is fully in line with the expectations.

Finally, we had to verify the presence of the predicted AAG to ATG codon change at position 27 in the *H3f3a-K27M* transcript and whether any other mutation might have been introduced inadvertently into the critical *H3f3a-K27M* encoding gene section during gene targeting. To this end we designed primers (KK39 and KK40) located upstream and downstream of the *H3f3a*-K27M open reading frame. These primers were used for RT-PCR to amplify and isolate full-length cDNA fragments from total RNA of relevant targeted ES clones, including AG #10.a. Dideoxy-sequencing of the entire coding region revealed presence of the desired AAG to ATG mutation at codon 27 and, importantly, complete identity of the entire *H3f3*a sequence with the wild-type transcript sequence deposited at the NCBI gene bank (Gene ID: 15078) in all ES clones analyzed, including AG #10.a (Figure 20).

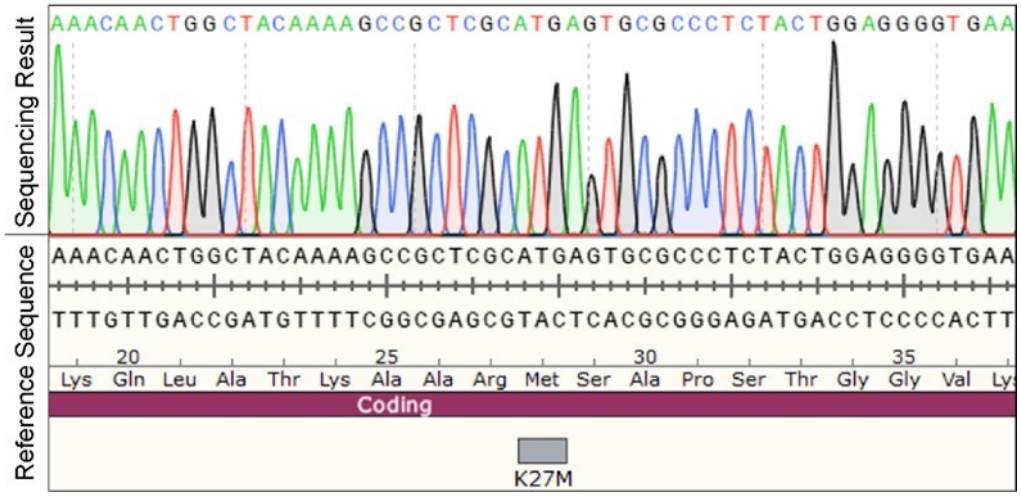


Figure 20: Sequencing results for pKK-H3TG-AG #10.a RT-PCR product
RT-PCR was performed, the product was extracted from the gel and sent for sequencing with primers KK27, KK39, and KK40. All primers delivered results equal to the one presented.

3.4 Generation of H3K27M knock-in mice from targeted ES cells

Three correctly targeted ES clones that had passed all rigorous verification steps described in the previous chapter (AF #65, AF #14, AG #10) were sent for blastocyst injections to the Transgenic Core Unit of the Max-Planck-Institute of Molecular Cell Biology and Genetics in Dresden (directed by Dipl.-Ing. Ronald Naumann), as, unfortunately, no blastocyst injection service was available on the campus of Ulm University.

Table 10 summarizes relevant data regarding blastocyst injections as obtained from Dresden. Briefly, after ES cell injection, a total of 328 chimeric embryos (blastocysts) were transferred to

pseudo-pregnant mice. From these, a total of 36 pups were born, ten of which showed significant chimerism.

Table 10: Summary of blastocyst injection data
Summary of ES cells sent for injection as well as the construct they were targeted with. Per clone, the number of transferred embryos, pups and chimeric mice are given. Data obtained from Ronald Naumann/ MPI Dresden.

Construct	Clone	Specification	Transferred embryos	Pups born	Chimeric mice
AF	#65	K27M mutation, IRES-vYFP	206	10	1 Hermaphrodite 1 Infertile
AF	#14		54	7	1 Female
AG	#10	K27M mutation, IRES-vYFP, 3xFLAG	68	19	4 Male 3 Female

Chimeras from AF #14 and AG #10 were set up for breeding with C57BL/6J mice in order to achieve germ line transmission and to establish stable lines of knock-in mice. Mice of the inbred strain C57BL/6J were chosen as breeding partners for two reasons. First, C57BL/6J mice are presumably the best characterized and most widely used inbred strain in immunological and hematological research, and second, while C57BL/6J mice have black coat color, intercross between 129/SvJ and C57BL/6J mice gives rise to offspring with agouti (brown) coat color, which makes an excellent marker to reliably indicate contribution of the 129/Sv ES cell genome to pups sired by chimeras. In other words, agouti coat color of pups indicates germline contribution, whereas pups with black fur are solely derived from the C57BL/6J genome and of no further interest. Provided the genetic modification does not interfere with embryogenesis, 50% of agouti offspring are thus expected to carry the targeted allele in their genome. Table 11 provides a summary of breeding results obtained at MPI/Dresden. Of note, genotyping of all chimeras listed in this table confirmed contribution of the ES cells.

Table 11: Summary of breeding for germline transmission
Chimeric mice born from blastocyst transfer were set up for breeding with C57BL/6J mice. The total number of litters and offspring with and without agouti coat color are given. Agouti (brown) coat color indicates germline contribution of the transgene. Data obtained from Ronald Naumann/MPI Dresden.

Chimera	Construct	Gender	Litter	Offspring	Agouti
53-B980	AF #14	Female	2	9	-
53-B981	AG #10	Female	3	26	-
53-B982	AG #10	Male	-	-	-
53-B983 and 53-B984	AG #10	Female	3	30	-
53-B985	AG #10	Male	-	-	-
53-B986	AG #10	Male	8	68	1
53-B987	AG #10	Male	6	36	-

Unfortunately, the first agouti offspring fathered by male #53-B986 was typed negative for the transgene. Nevertheless, the appearance of one agouti mouse was still positive news, since usually more agouti offspring will be obtained from the following litters. To maximize the success rate, sperm was isolated from male #53-B986 and used for *in vitro* fertilization (a strategy similar to the one described in (Li et al., 2015)) resulting in 120 more offspring. Luckily, several more agouti mice, derived from cells targeted with pKK-H3TG-AG (thus carrying the

IRES-vYFP marker and the FLAG-tag), were born. Tissue biopsies from the first four agouti offspring weaned were sent to Ulm for genotyping (data not shown). One male and one female were typed heterozygous for H3K27M and were promptly set up for breeding in Dresden to expand the line without any delay. Again, offspring of these breeders were typed in Ulm, and all 11 mice that were identified as heterozygous for H3K27M (see Table 12) were then transferred to the TFZ of Ulm University.

Table 12: Mice transferred from Dresden to the TFZ Ulm
Biopsies of these mice were sent to us for genotyping. 11 mice that were found to be heterozygous for the targeted gene locus (H3K27M) were sent to the TFZ Ulm. Given are the IDs, birth dates, and gender of the mice as well as the genotyping result.

ID	Date of Birth	Gender	Mutation
51-H248	12.04.2020	Male	H3K27M$^{ki/wt}$
51-H249	12.04.2020	Male	H3K27M$^{ki/wt}$
51-H250	12.04.2020	Male	H3K27M$^{ki/wt}$
51-H252	12.04.2020	Female	H3K27M$^{ki/wt}$
51-H265	19.04.2020	Male	H3K27M$^{ki/wt}$
51-H266	19.04.2020	Male	H3K27M$^{ki/wt}$
51-H267	19.04.2020	Male	H3K27M$^{ki/wt}$
51-H268	19.04.2020	Female	H3K27M$^{ki/wt}$
51-H269	19.04.2020	Female	H3K27M$^{ki/wt}$
51-H270	19.04.2020	Female	H3K27M$^{ki/wt}$
51-H271	19.04.2020	Female	H3K27M$^{ki/wt}$

After transfer of these mice with mixed 129/Sv-C57BL/6J genetic background into our facility, they were set up for further breeding following three strategies. First, as the genetic background can profoundly influence observed phenotypes, it seemed wise to initiate serial backcrossing onto C57BL/6J to assure that mutant mice with a defined genetic background as homogeneous as possible would gradually become available. During the course of this thesis project backcrossing with C57BL/6J was done for up to ten generations. Second, some of the available heterozygotes with mixed genetic background listed in Table 12 were directly intercrossed. On one hand, such breedings would readily reveal potential detrimental effects of the targeted *H3f3a* allele on embryogenesis or fertility, if deviations from the expected Mendelian genotype distribution were observed. On the other hand, these early intercrosses should promptly provide homozygous knock-in mice that would simplify subsequent breeding schemes, like intercrosses with vavCre, *Mll5*$^{flox/flox}$ and/or *Dnmt3a*$^{flox/flox}$ mice, by substantially reducing the number of offspring with unwanted genotypes. Table 13 depicts genotyping results for the first 211 pups from heterozygous parents of the first C57BL/6J backcross generation, i.e. from parents listed in Table 12. The data indicate Mendelian distribution and therefore do not provide any hint for dramatic negative effects of the targeted allele(s) on germ cell function or embryogenesis. This result was predicted as targeting with the original construct, as performed by the Jeffey Mann laboratory, also did not yield any abnormal changes in viability and fertility of heterozygous offspring (Tang et al., 2013).

Table 13: Genotype distribution of H3K27M intercrosses
Genotype distribution of KI and WT alleles obtained from intercrossing of heterozygous H3F3A-K27M mice. Data was generated from 11 breeding cages set up between 24.06.2020 and 26.07.2021.

Summary	WT/WT	KI/WT	KI/KI	Total number typed
Male	22	57	27	106
Female	25	51	29	105
Total	47 (22.27%)	108 (51.18%)	56 (26.54%)	211

Finally, some of the heterozygous H3K27M knock-in mice listed in Table 12 were intercrossed with heterozygous vCre mice to obtain experimental animals, in which the *H3f3a-K27M* mutant allele was supposed to be activated and expressed in all hematopoietic cells, as soon as possible. Only in such animals potential effects of the H3f3a-K27M mutation can be investigated. As soon as homozygous H3K27M$^{KI/KI}$ mice became available from the aforementioned heterozygous intercrosses, they were used preferentially for the intercross with vCre partners to inherit the targeted *H3f3a* allele to all offspring, thus reducing the number of pups born with an unwanted genotype by 50%. Table 14 reports on the frequency of Cre-positive offspring. Again, the genotype distribution of Cre can give an important first hint at a potentially lethal effect of the heterozygous knock-in allele, which now – after Cre-mediated recombination – is expected to express the mutant H3.3A histone with the K27M amino acid substitution in all hematopoietic cells. Notably, ~50 % of mice were Cre positive, strongly suggesting that expression of the K27M mutation in histone variant H3.3 in hematopoietic stem cells does not result in overall lethality.

Table 14: Distribution of vCre alleles in offspring from heterozygous vCre^{pos} and homozygous H3K27M$^{KI/KI}$ mice.
Data was obtained from 11 breeding cages set up between 04.12.2022 – 25.07.2021.

Summary	Cre-pos	Cre-neg	Total number typed
Male	50 (58.1%)	36 (41.9%)	86
Female	38 (46.9%)	43 (53.1%)	81
Total	88 (52.7%)	79 (47.3%)	167

3.5 Verification of recombination efficiency and specificity with vCre x RoYFP Mice

To confine expression of the mutant *H3f3a-K27M* knock-in allele to hematopoietic cells (for rational, see Introduction section 1.7), we used a specific strain of transgenic mice expressing Cre-recombinase under control of the vav promoter (de Boer et al., 2003). This promoter is expected to be active in all hematopoietic cells (Georgiades et al., 2002). To verify the performance of our vCre mice with regard to efficiency and specificity of recombination, we intercrossed vCre with Rosa26-EYFP (RoYFP) mice, a widely used Cre-reporter strain (Srinivas et al., 2001). Briefly, the *Rosa26* gene locus in mice is expressed at roughly equal intensity in all cells and tissues analyzed, giving rise to a non-coding RNA of unknown function. *Rosa26* has thus become a preferred genomic locus for the insertion of transgenes, including fluorescent reporter genes like YFP, as no known physiological function is disrupted upon insertion (Friedrich and Soriano, 1991; Zambrowicz et al., 1997). The RoYFP reporter mice

used in our experiments express an enhanced yellow fluorescent protein (YFP) upon Cre-mediated deletion of a loxP-flanked transcriptional “stopper” element (Srinivas et al., 2001).

To assess Cre-recombinase activity in vCre x RoYFP offspring carrying both one vCre and one RoYFP allele, we analyzed single cell suspensions from thymus, spleen and bone marrow for YFP reporter expression by flow cytometry. RoYFP reporter mice lacking a vCre allele (vCre^{neg} x RoYFP) served as negative controls. In all these experiments, we used the well-established CD45 cell surface marker to distinguish truly hematopoietic cells (CD45^{pos}) from minor, contaminating populations of mostly stromal cell origin (CD45^{neg}) (Trowbridge and Thomas, 1994). Figure 21 - Figure 23 show a selection of representative results obtained from these analyses in thymus, spleen, and bone marrow. Remarkably, virtually 100% of CD45^{pos} thymocytes were YFP-positive while the small fraction of CD45^{neg} thymocytes was completely negative for the fluorescent reporter (Figure 21, left panels). When fractioning thymocytes into distinct developmental stages based on CD4 and CD8 marker expression (Figure 21, right panels) it can be observed that the percentage of vYFP positive cells is slightly reduced by 2 – 4 %. However, the vast majority of CD45^{pos} thymocytes still shows strong expression of the fluorescent marker.

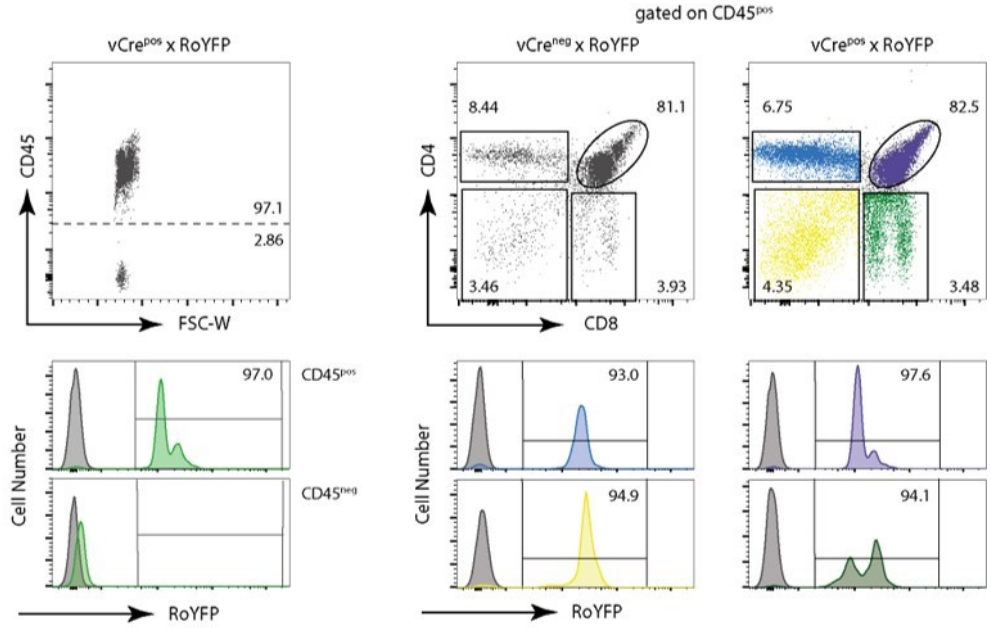


Figure 21: Quantitative and highly specific YFP marker expression in CD45^{pos} thymocytes reflects Cre-recombinase activity in vCre x RoYFP mice
Analysis of YFP expression in CD4/CD8 thymocyte subsets. The upper left dot plot gating strategy for CD45 in vCre^{pos} x RoYFP mice. Histograms below correspond to YFP expression in CD45^{pos} and CD45^{neg} cells of vCre^{neg} x RoYFP (grey) and vCre^{pos} x RoYFP (green) mice. Panels on the right show reporter expression levels in CD4/CD8 subsets gated on CD45^{pos} cells. Colored histograms correspond to the subsets of vCre^{pos} and grey to vCre^{neg} mice, respectively. Numbers indicated in the plots represent the percentage of YFP positive cells.

Also, virtually all CD45^{pos} splenocytes were strongly YFP-positive (Figure 22, left panel), which was equally true for splenic B- and T-lymphocytes, as identified by CD19 and CD3 marker expression, respectively (Figure 22, right panels). The weak fluorescent signal of widely varying intensity observed among the few CD45^{neg} splenocytes might be due to their non-specific absorbance of YFP protein liberated from apoptotic YFP-expressing cells. These weak, heterogeneous signals are certainly not derived from *Rosa26*-driven YFP expression,

otherwise fluorescence intensity would be less broadly distributed and at roughly similar strength as in CD45^{pos} cells.

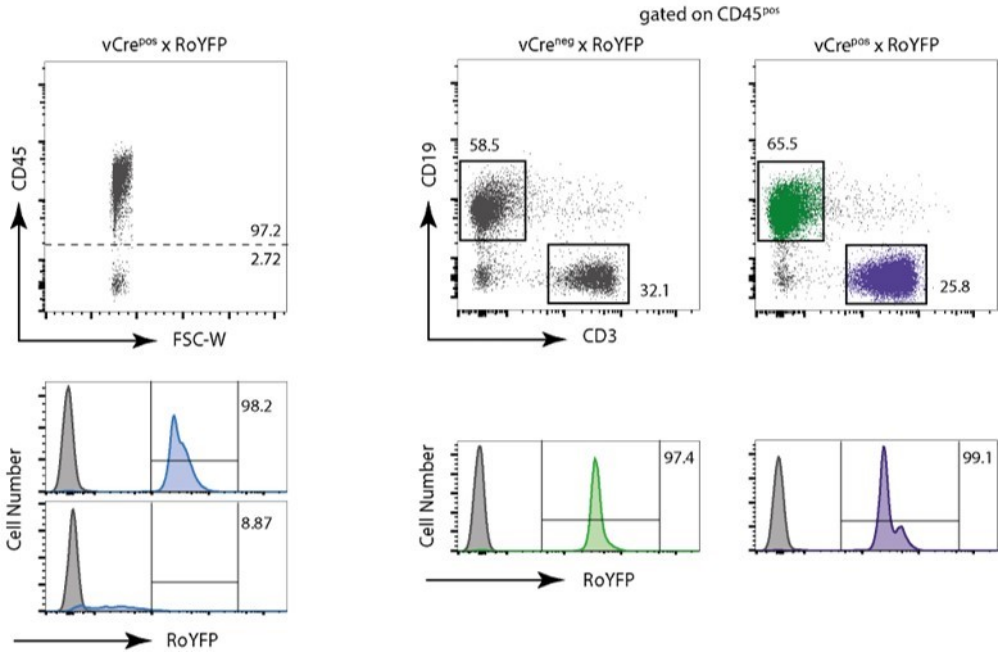


Figure 22: Quantitative and highly specific YFP marker expression in CD45^{pos} splenocytes reflects Cre-recombinase activity in vCre x RoYFP mice

Analysis of YFP expression in splenocytes. Left panels show gating for CD45. Below are histograms indicating the reporter expression in vCre^{neg} (grey) and vCre^{pos} (blue) CD45 sub-populations. The right panels show reporter expression in the CD3^{pos} and CD19^{pos} splenic T- and B-lymphocyte populations, respectively. Numbers indicated in the plots represent the percentage of YFP positive cells.

In line with these findings, also essentially all CD45^{pos} bone marrow cells in gated Lineage-negative, Kit-positive (LK) and Lin-negative, Kit-positive, Sca-1-positive compartments (LSK) were YFP marker positive (Figure 23). This is of particular importance as LK and LSK BM subpopulations comprise all hematopoietic stem and progenitor cell subsets from which AML typically initiates.

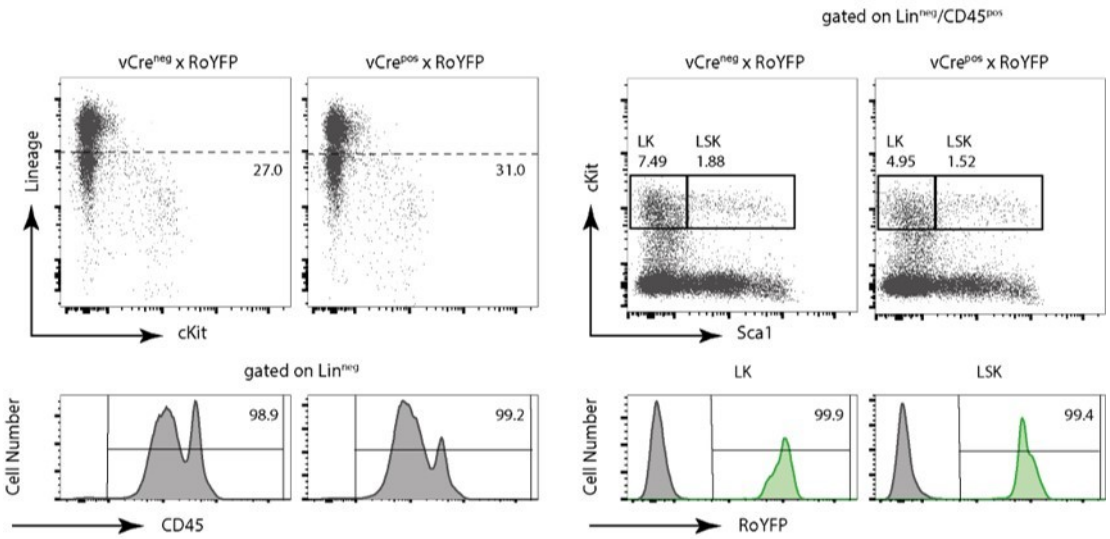


Figure 23: Quantitative and highly specific YFP marker expression in CD45^{pos} bone marrow cells reflects Cre-recombinase activity in vCre x RoYFP mice
Plots on the left show gating strategy for Lin^{neg} (negative for CD3e, CD4, CD8a, CD11b, CD19, Gr1, NK1.1, and Ter119) and CD45^{pos} cells. Panels on the right show YFP expression in LSK and LK progenitor cells. Numbers indicated in the plots represent the percentage of YFP positive cells.

In summary, the data convincingly demonstrate that our vCre-transgenic mice can reliably activate loxP-dependent transcription in all hematopoietic cells, including the HSPC subsets presumed to give rise to AML upon mutational transformation. Furthermore, even CD45^{neg} cells co-isolated from hematopoietic tissues were completely negative for the expression of the fluorescent marker, highlighting the strictly confined specificity of vav-Cre expression to genuine hematopoietic cells. Similar results can be expected with regard to activation of the *H3f3a*-K27M allele in vCre^{pos} x H3K27M mice. As our newly generated H3K27M mice also carry the Cre-inducible fluorescent marker vYFP to monitor expression of *H3f3a*-K27M transcripts in individual cells, vCre x RoYFP mice were frequently used as convenient positive control during this project.

3.6 vCre^{pos} x H3K27M knock-in mice correctly express the *H3f3a-K27M* allele specifically in hematopoietic cells

Although correct expression of the *H3f3a-K27M* knock-in allele had been rigorously evaluated at RNA and protein level in ES cell clone AG #10 used for generating our H3K27M strain of knock-in mice (see section 3.4), good scientific practice required analogous verification of *H3f3a-K27M* expression in relevant mouse tissues of vCre^{pos} x H3K27M$^{KI/WT}$ mice. To this end, the following RT-PCR experiments were performed:

First, total RNA isolated from thymus and spleen of two distinct vCre^{pos} x H3K27M$^{K/WT}$ mice was subject to RT-PCR using primer pair KK40/KK39 that anneals upstream and downstream of the *H3f3a-K27M* open reading frame in resultant cDNA, thus amplifying the entire coding region of *H3f3a-K27M*. PCR-fragments of predicted length (774 bp) were obtained from RNA of both organs and both mice (Figure 24A and data not shown), gel-purified and completely sequenced using primers annealing at different positions within the cDNA (primer KK27, KK36 and KK40). Comparison of the sequences obtained from spleen and thymus RNA with the published transcript sequence of *H3f3a* in the NCBI gene bank revealed full identity over the

entire coding region except for codon 27, which carried the introduced AAG to ATG mutation, as expected. The same positive results were obtained for *H3f3a-K27M* transcripts from both organs, i.e. thymus and spleen (Figure 24B) of both individual mice analyzed, demonstrating conclusively that *H3f3a* transcripts carrying the introduced K27M mutation are correctly expressed. Sequencing data also confirmed correct in-frame insertion of the 3xFLAG-encoding sequence between the 3' end of the *H3f3a-K27M*-encoding segment and the IRES element driving vYFP expression (data not shown).

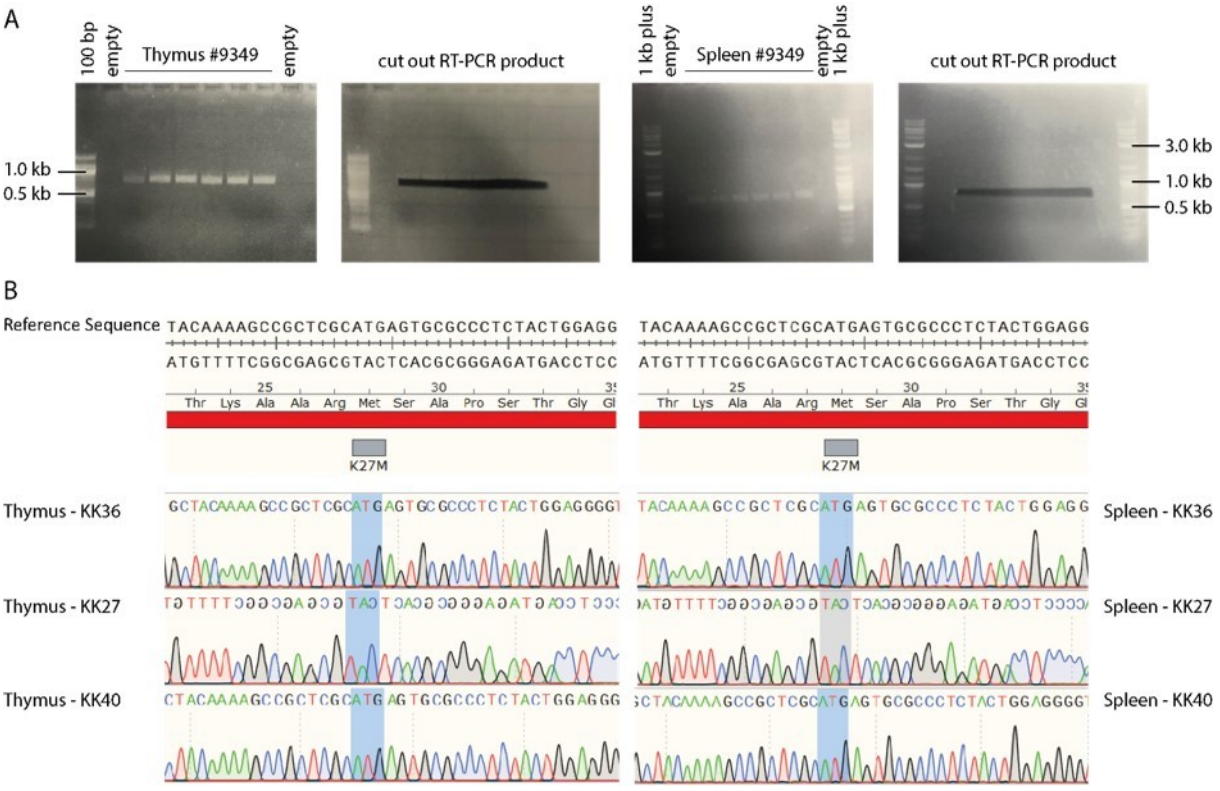


Figure 24: Sequencing of full-length H3f3a transcript in thymus and spleen of a vCre^{pos} x H3K27M$^{KI/WT}$ mouse
(A) RT-PCR products from thymus (left) and spleen (right) were cut out and extracted from the gel. **(B)** The isolated RT-PCR product was sequenced with primers KK27, KK36, and KK40 to verifiy in-frame expression of the H3K27M mutation. The top panel shows the reference sequence; lower three panels show the sequencing results for each primer. Highlighted in blue is the mutated codon at position 27.

To verify strict Cre-recombinase-dependent expression of *H3f3a-K27M* transcripts, an additional set of RT-PCR experiments was performed, again using primer pair KK39/KK40. However, this time RNA from brain of vCre^{pos} x H3K27M$^{KI/WT}$ mice was included in the analysis as negative control, as brain tissue should not express vav-driven Cre-recombinase. Furthermore, total RNA from thymus, spleen and brain of vCre^{neg} x H3K27M$^{KI/WT}$ mice (lacking the vCre transgene) was analyzed at the same time. Figure 25 shows a representative result for two vCre^{pos} x H3K27M$^{KI/WT}$ mice and one vCre^{neg} x H3K27M$^{KI/WT}$ animal. Data demonstrate that *H3f3a-K27M* expression is indeed strictly Cre-dependent, as *H3f3a-K27M* transcripts are only detectable in vCre-expressing tissues (thymus and spleen) of vCre^{pos} mice.

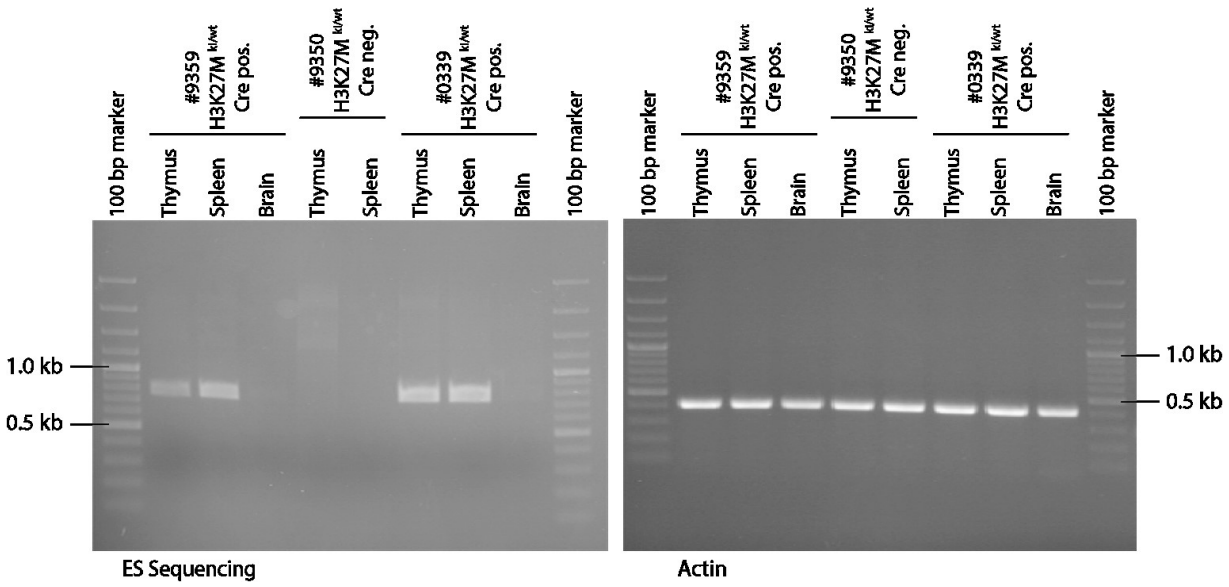


Figure 25: Detection of H3f3a-K27M transcripts in vCre x H3K27M mice
RT-PCR confirms that *H3f3a*-K27M is expressed in only in hematological organs of vCre^{pos} mice. RT-PCR for actin with primer pair RT173/RT174 was performed as positive control, e.g. indicating equal amounts of reverse transcribed RNA in all samples.

3.7 IRES-driven vYFP reporter expression faithfully reflects *H3f3a-K27M* expression and transcriptional activity of the endogenous *H3f3a* gene locus

Using flow cytometry, we also analyzed IRES-driven vYFP reporter expression in single cell suspensions prepared from relevant hematological organs of vCre^{pos} x H3K27M$^{KI/WT}$ and vCre^{pos} x H3K27M$^{KI/KI}$ mice, i.e. experimental animals hetero- and homozygous for the knock-in allele. Corresponding cell populations from C57BL/6J and heterozygous H3K27M mice lacking the vCre transgene were used as negative controls. Representative flow cytometric profiles are depicted in Figure 26. Highly reproducible data obtained in these experiments allow several clear-cut conclusions: **(i)** Robust reporter expression faithfully reflects Cre-mediated *H3f3a-K27M* activation in individual cells. **(ii)** Apparently, in mice carrying both a vCre transgene and a knock-in allele, essentially all hematopoietic cells in each organ analyzed express the *H3f3a*-K27M allele, as no distinct peak for vYFP-negative cells is visible in any cell population analyzed. This finding is fully in line with data obtained with vCre^{pos} x RoYFP mice (Figure 21 - Figure 23) and demonstrates that the targeted *H3f3a-K27M* locus is as accessible for Cre-recombinase as the previously assessed RoYFP locus. **(iii)** Essentially identical "background" MFI values in all corresponding cell populations from C57BL/6J and vCre^{neg} x H3K27M$^{KI/WT}$ control mice indicate strictly Cre-dependent expression of the *H3f3a*-K27M allele and no leaky vYFP reporter expression in the absence of Cre. This observation is in full agreement with RT-PCR data shown in Figure 25. **(iv)** Although vYFP reporter expression is robust in cells expressing the *H3f3a-K27M* allele, observed fluorescent intensities are not high enough to completely separate all vYFP^{neg} from all vYFP^{pos} cells in heterozygous K27M mice, since respective histograms are partly overlapping (Figure 26, panels with green histograms). However, if experimentally required, this imperfection can possibly be overcome by analyzing respective cell populations on a flow cytometer equipped with a 532 nm (green) laser for optimal vYFP excitation (Telford et al., 2012). In contrast, all analyses presented in this thesis were

done on a BD LSR Fortessa available in our institute with a more common, but suboptimal YFP-excitation laser at 488nm. **(v)** Hematopoietic cells in vCre^{pos} mice with two *H3f3a*-K27M alleles exhibit MFI values exactly twice as high as corresponding cell populations in heterozygous vCre^{pos} x H3K27M$^{KI/WT}$ mice (Figure 26, compare red and green histograms, respectively). This pattern of vYFP fluorescence is expected for bi-allelic *H3f3a* expression and provides additional evidence that vYFP reporter-derived fluorescence intensities perfectly correlate with *H3f3a*-K27M transcript levels. Moreover, in vCre^{pos} mice homozygous for the knock-in allele, vYFP^{pos} and vYFP^{neg} cell subsets can be efficiently separated by flow cytometry even with conventional laser settings (Figure 26, red histograms). Taken together, flow cytometric analysis of vYFP reporter expression in diverse hematopoietic cell populations of vCre^{pos} mice harboring one or two *H3f3a-K27M* knock-in alleles demonstrates robust, Cre-dependent expression of *H3f3a-K27M* transcripts in essentially all hematopoietic cells analyzed.

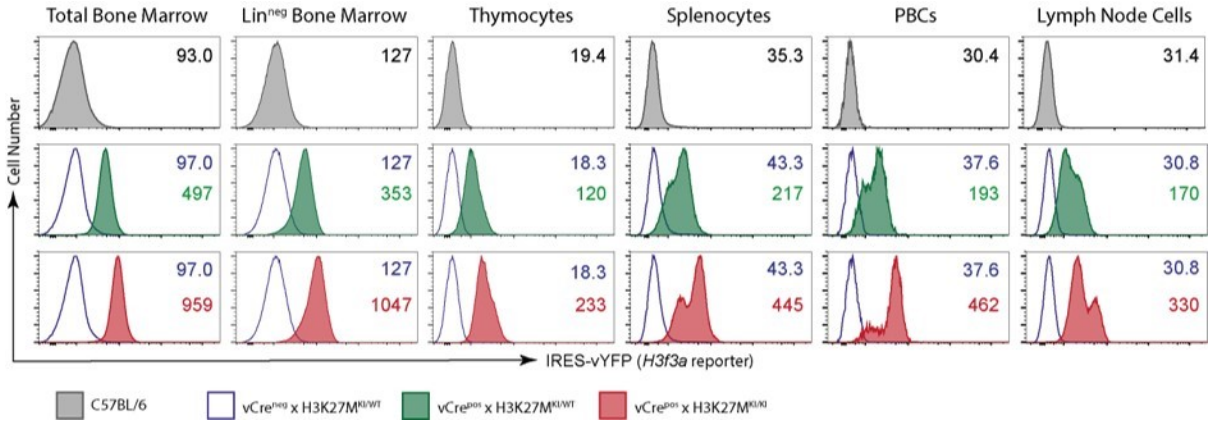


Figure 26: H3f3a-K27M reporter expression (vYFP) in hematological organs of mice
The IRES-driven vYFP reporter is faithfully expressed in hematological organs upon activation of the *H3f3a-K27M* allele with vCre. Comparison of histograms between C57BL/6J and vCre^{neg} x H3K27M$^{KI/WT}$ mice demonstrate that there is no significant leaky expression of the IRES-vYFP reporter in the absence of vCre. Data are representative of three independent experiments with identical results. Numbers indicated in the histogram represent the MFI of IRES-vYFP. PBCs = peripheral blood cells.

Interestingly, in some cell populations, like splenocytes, lymph node cells or cells from peripheral blood (PBCs), vYFP-derived profiles are not simply bell-shaped, but more complex, indicating substantial heterogeneity in vYFP expression, which is particularly obvious in mice with two *H3f3a-K27M* alleles and thus stronger vYFP expression (Figure 26, lower panels, red histograms). The observed pattern of vYFP fluorescence suggests that splenocytes, lymph node cells and PBCs contain cell populations with significantly differing *H3f3a* gene expression levels. To gain further insight into this phenomenon, splenocytes and lymph node cells from vCre^{pos} x H3K27M$^{KI/KI}$ mice were separated into T- and B-lymphocytes based on CD3/CD19 staining (Figure 27A). As documented in Figure 27B and C, MFI values for CD19- and CD3-positive cells differ substantially, indicating that combined expression from both *H3f3a-K27M-IRES-vYFP* alleles is approximately 60% lower in T-cells compared to B-cells. To verify whether this observation might be due to a transgenic artifact caused by targeted modification of the *H3f3a* gene locus or whether this finding reflects genuine, cell-type-specific differences in *H3f3a* gene expression, B- and T-lymphocytes were sorted from spleen and lymph nodes of genetically unmanipulated, wild-type C57BL/6 mice and used for isolation of total RNA to allow

precise quantification of natural, i.e. wild-type *H3f3a* transcript levels by qRT-PCR. Figure 27D presents combined results of three independent cell sorting/qRT-PCR experiments demonstrating approximately 60% lower *H3f3a* transcript levels also in genetically "normal" T- versus B-lymphocytes, which perfectly correlates with the observed differences in vYFP expression between both cell subsets (Figure 27B and C). The perfect correlation between *H3f3a* transcript levels in wild-type mice and vYFP expression levels (vYFP-derived MFI values) demonstrates that our newly generated knock-in mice provide a unique, novel tool to reliably visualize differences in *H3f3a* expression in individual living cells simply by flow cytometry. Beyond this thesis project, our mice should thus be of prime utility in studies aiming to investigate non-invasively potentially meaningful physiological variations in *H3f3a* expression at single cell level.

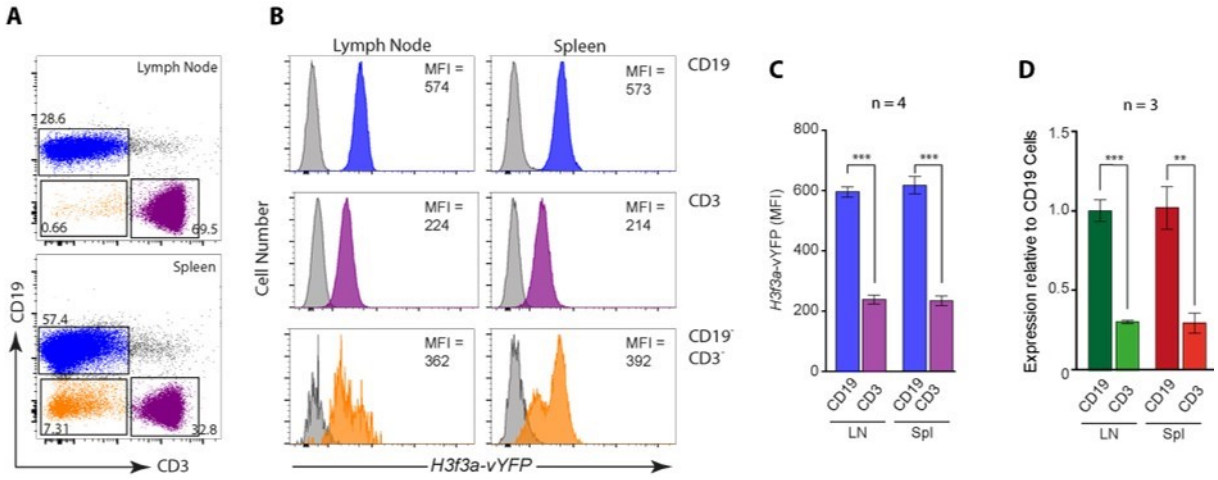


Figure 27: Significantly different H3f3a expression levels in B- and T-lymphocytes are faithfully reflected by vYFP marker expression

(A) Gating strategy for CD3/CD19 cell subsets in spleen and lymph nodes. The numbers given are percentages of the parent population. **(B)** Histograms for vYFP expression in gated cell subsets from C57BL/6J (grey) and vCre^{pos} x H3K27M$^{KI/KI}$ (colored) mice. MFI values are indicated for cells from vCre^{pos} x H3K27M$^{KI/KI}$ mice only. **(C)** Quantification of vYFP-derived MFI values in indicated cell populations from vCre^{pos} x H3K27M$^{KI/KI}$ mice. **(D)** qRT-PCR results for *H3f3a* expression in sorted C57BL/6J cell populations. The values are normalized against the CD19 expression in lymph nodes and spleen, respectively. Data was collected from three independent sorts. **All panels:** Blue = CD19^{pos} cells from vCre x H3K27M$^{KI/KI}$, orange = CD3^{neg}/CD19^{neg} cells from vCre x H3K27M$^{KI/KI}$, purple = CD3^{pos} cells from vCre x H3K27M$^{KI/KI}$, dark green = CD19^{pos} wild-type cells from lymph node, light green = CD3^{pos} wild-type cells from lymph node, dark red = CD19pos cells from wild-type spleen, light red = CD3pos cells from wild-type spleen; MFI = mean fluorescence intensitiy, LN = lymph node, Spl = spleen. **(C+D)** Significance was calculated using unpaired, parametric t-test. Data shown as mean ± SEM.

3.8 The expression level of the *H3f3a-K27M-IRES-vYFP* knock-in allele seems significantly reduced in comparison with *H3f3a* wild-type alleles

Although RT-PCR data (Figure 24 and Figure 25) and MFI values for vYFP expression (Figure 26 and Figure 27) indicate substantial expression of the *H3f3a-K27M-IRES-vYFP* knock-in allele in all hematopoietic cells of vCre^{pos} mice, it was of interest to know how *H3f3a-K27M* expression compares to *H3f3a-WT* expression from the non-manipulated gene locus. In an attempt to address this issue, we performed qRT-PCR experiments with RNA isolated from unfractionated, total BM cells, thymocytes, and splenocytes of mice from six different strains specified in Figure 28. Corresponding RNA samples from one mouse of each strain were always analyzed together in a total of three independent qRT-PCR experiments. Combined

data from all three experiments are presented in Figure 28, providing a complex set of informative data: **(i)** *H3f3a* expression levels determined in cell populations from genetically non-manipulated C57BL/6J mice were arbitrarily set at 1.0 (100%). Mice heterozygous for the vCre transgene were included as an additional control, in case Cre-recombinase expression itself has an effect on *H3f3a* transcript levels, for instance, because strong Cre activity can cause non-specific DNA damage (Schmidt-Supprian and Rajewsky, 2007), indirectly altering gene expression in unpredictable ways. While *H3f3a* transcript levels were highly similar in BM cells and thymocytes of wild-type C57BL/6J and vCre-transgenic mice, as expected, if Cre expression itself had no effect, an unexplained, approximately 40% reduction of *H3f3a* expression was consistently observed in splenocytes from vCre mice in comparison with C57BL/6 controls (Figure 28, rightmost panel), underlining the importance of including vCre mice in the analysis. **(ii)** The primers used in our qRT-PCR experiments would not recognize *H3f3a-WT**** transcripts expressed from the uninduced knock-in allele, as these transcripts carry nucleotide alterations at wobble bases (see chapter 3), including the primer annealing sites. After targeting, transcription levels of *H3f3a* are therefore expected to be reduced by 50 % in heterozygous and by 100 % in homozygous mice due to the expression of *H3f3a-WT**** from one or both alleles. This was indeed observed in BM cells and thymocytes from H3K27M$^{KI/WT}$ and H3K27M$^{KI/KI}$ mice, respectively. Why transcript levels of *H3f3a* in splenocytes of H3K27M$^{KI/WT}$ are reduced by approximately 85 % rather than the predicted 50 %, when compared to C57BL/6J mice, is currently an unexplained aberrancy. It further highlights the importance of appropriate controls included in these experiments, as, based on this result, vCre^{pos} mice appear to be the more suitable control. **(iii)** Upon Cre-mediated activation of the targeted *H3f3a* allele in vCre^{pos} x H3K27M mice, the transcript levels should reach values comparable to those observed in C57BL/6J or vCre^{pos} mice. However, this is clearly not the case. While a modest increase in *H3f3a* transcript levels is observed in all samples after Cre-mediated activation of knock-in loci, the expression always remains much below the predicted wild-type level. At face value, these data seem to indicate that transcript levels of K27M-mutant *H3f3a* from the targeted gene locus are – dependent on the cell populations analyzed – at least 80 – 90 % lower than those from the non-manipulated wild-type allele. However, have compelling reason to doubt that qRT-PCR data obtained actually reflect transcript levels from the knock-in allele correctly, as will be debated in detail in chapter 4.1 (“Discussion” section). For instance, the inclusion of the 3xFLAG-tag and IRES-vYFP reporter element in the targeting construct more than doubles the size of the transcript from the targeted allele and certainly boosts RNA secondary structure. Both effects might negatively influence the efficiency of reverse transcription and/or PCR amplification, potentially resulting in a significant underestimation of transcript levels.

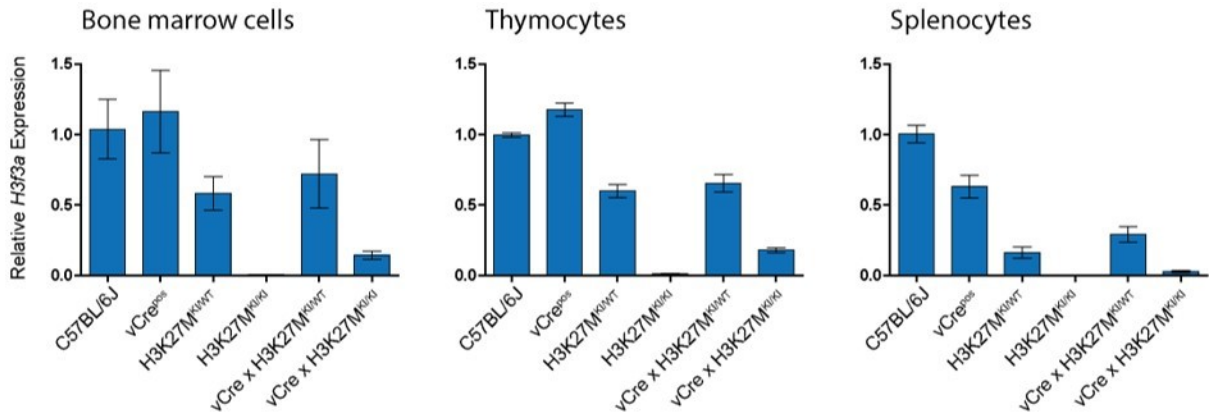
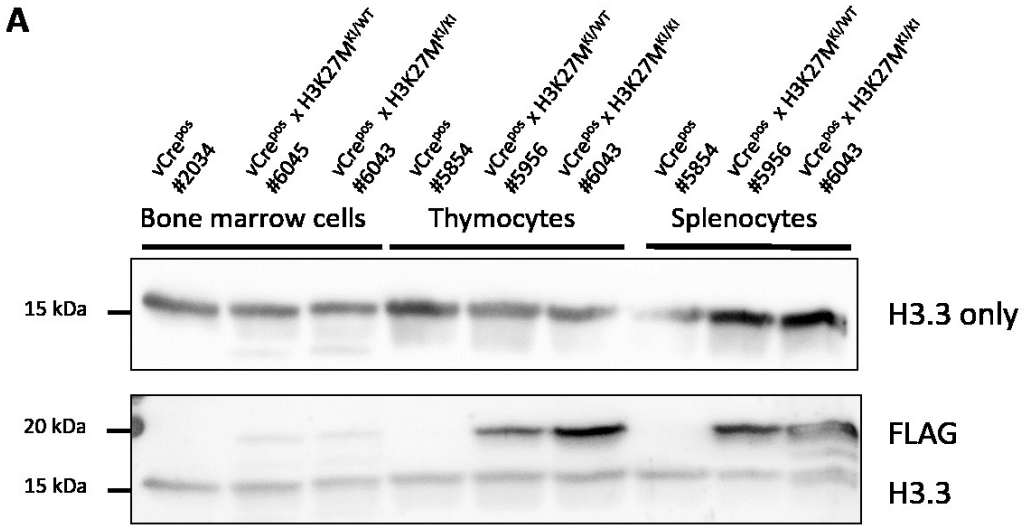


Figure 28: Comparison of H3f3a transcription levels in vCre x H3K27M mice and appropriate controls
Expression of *H3f3a* in unfractionated BM cells, thymocytes and splenocytes of mice with relevant genotypes. All samples are normalized against the mean obtained with correspondent cell populations from three C57BL/6J mice. Samples from 3 individual mice for each genotype were analyzed in 3 independent qRT-PCR experiments (n = 3). Data are shown as mean ± SEM. Statistical significance between C57BL/6J and vCre^{pos} mice was calculated using Mann-Whitney test and did not reveal any significant changes in *H3f3a* expression.

To complement qRT-PCR results with an independent method, we performed Western Blotting hoping to be able to directly compare output from wild-type and knock-in *H3f3a* alleles at the protein level. However, also this approach has potential pitfalls, as total histone H3.3 protein levels are determined by the activity of two separate genes, *H3f3a* and *H3f3b,* which give rise to polypeptides of identical amino acid sequence (see “Introduction”, chapter 1.6.3). As a consequence, histone H3.3-specific bands in Western blots always represent a mixture of H3.3A and H3.3B proteins of unknown stoichiometry, and modification of one gene locus, e.g. *H3f3a*, could bring about compensatory expression changes at the other, i.e. *H3f3b* locus, which would go unnoticed. With these intricacies in mind, we nevertheless performed Western Blots with histone extracts from total thymocytes, splenocytes, and BM cells from relevant mouse strains, using antibodies directed specifically either against histone variant H3.3 or against the FLAG epitope. As shown in Figure 29A, top panel, the anti-H3.3 reagent marks a specific band at ~15 kDa corresponding to the size of wild-type histone H3.3. In hematopoietic cell samples from vCre^{pos} x H3K27M$^{KI/WT}$ and vCre^{pos} x H3K27M$^{KI/KI}$ mice an additional band should appear at 18.5 kDa corresponding to mutant H3.3A-K27M, whose molecular weight is increased by approximately 3.5 kDa due to the attached FLAG-tag. Published data suggests different expression levels for *H3f3a* and *H3f3b* genes in a tissue-specific manner (Jang et al., 2015). Especially in the thymus and spleen of wild-type mice, the expression of *H3f3a* is increased 3 – 4 fold as compared to *H3f3b*. Hence, the second 18.5 kDa band resulting from the corresponding *H3f3a-K27M-IRESvYFP-FLAG* allele should be clearly detectable. Unfortunately, there was no second band measurable at that size in any of our blots. At its worst, this could mean that K27M-mutant histone variant H3.3 is not expressed at all, or at least that its expression is below the level of detection. Alternatively, the anti-H3.3 antibody used might not be able to efficiently detect mutant H3.3 with a K27M amino acid change, as this amino acid position may be part of the recognized epitope. This latter interpretation is strongly supported by the fact that the anti-FLAG reagent reliably detects a specific band at ~18.5 kDa, and as expected exclusively in mice carrying the vCre-induced *H3f3a-K27M* alleles (Figure 29A, lower panel). This finding provides conclusive evidence for the presence of significant amounts of H3.3A-K27M protein in knock-in mice with vCre expression. Furthermore, the

increased intensity from homozygous *H3f3a-K27M* mice is clearly detectable, as well as the predicted, tissue-specific changes in expression levels. To quantify the observed relative differences in FLAG-expression, ImageJ was used to determine the mean grey value of FLAG-specific bands detected in thymocytes, as representative organ. The absolute results from three independent experiments are shown in Figure 29B. Figure 29C presents the calculated relative expression changes of the FLAG-tag in H3K27M$^{KI/WT}$ as compared to H3K27M$^{KI/KI}$ thymocytes.

This positive result also highlights how important it was to include a FLAG-tag in our targeting construct to enable the detection of the H3.3-K27M protein. However, despite these positive data, also our Western blotting experiments did not allow a reliable estimate of H3.3A-K27M protein expression in direct comparison with wild-type H3.3. Of note, the differences in band intensity between H3.3 and FLAG (Figure 29A, bottom blot) do not represent quantitative differences, as two different secondary antibodies were used.



B

	H3K27M$^{KI/WT}$	H3K27M$^{KI/KI}$
Experiment 1	73	142
Experiment 2	69	124
Experiment 3	55	109

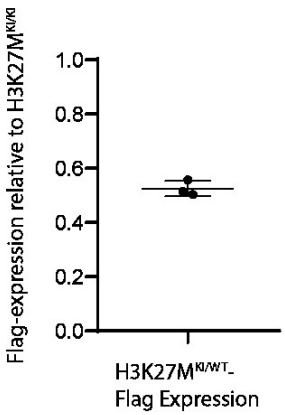


Figure 29: Western Blot detecting H3f3a-K27M mutant proteins
(A) 2.5 µg of acid extracted histones were run 6 + 15 % acrylamide gel and blotted on a PVDF membrane. Top panel: Over-night hybridization with total H3.3 antibody reveals equal protein levels in all samples. Lower panel: Without stripping, the same membrane was incubated over-night with anti-FLAG antibody resulting in bands specific for the H3.3-K27M FLAG-tagged protein. **(B)** Absolute mean grey values for the FLAG-specific band observed in thymocytes. The results are from three independent experiments. **(C)** Expression of the FLAG-tag in H3K27M$^{KI/WT}$ mice relative to H3K27M$^{KI/KI}$. Pictured are the individual values with mean ± SD.

3.9 Cre-mediated induction of K27M-mutated histone variant H3.3 results in marked reduction of global H3K27 trimethylation

Unfavorable, albeit suspicious qRT-PCR results discussed above (Figure 28) and our inability to reliably compare mutant H3.3A-K27M with wild-type H3.3 protein levels raised concerns that H3.3A-K27M protein expression from the knock-in allele might be inadequate to mimic functional effects of heterozygous *H3F3A-K27M* mutations observed in our 7q- AML patients ("Introduction", Table 2). We thus looked for known H3.3-K27M-driven pathophysiologically relevant events in our mice as surrogate expression marker. In roughly 30 % of children with high-grade glioma, tumor cells carry heterozygous *H3F3A-K27M* driver mutations (see "Introduction", section 1.6.3). An intense investigation of underlying molecular mechanisms revealed that the H3.3A-K27M protein acts as a dominant negative inhibitor of EZH2 in the PRC2, resulting in global histone H3K27 undermethylation (Lewis and Allis, 2013). Although mutant H3.3A-K27M protein comprises only 3 – 17 % of total H3 histones in these tumors (Lewis and Allis, 2013), K27-specific trimethylation is reduced in the majority of H3 proteins in affected cells, thus the term "global undermethylation". Furthermore, ample evidence suggests that global H3K27 undermethylation substantially contributes to cell transformation (Kallappagoudar et al., 2015; Larson et al., 2019). To determine whether Cre-mediated expression of *H3f3a-K27M* in our mice affects global H3K27 trimethylation, intracellular FACS was performed with an antibody specific for H3K27me3. A second common methylation site (H3K4) was analyzed with an anti-H3K4me3-specific antibody as control. The results obtained were striking and clear-cut (Figure 30). In all hematopoietic cell populations analyzed (unfractionated BM cells, thymocytes, and splenocytes) MFI values for H3K27me3 were reduced by about 25 % in cells with one active *H3f3a-K27M* allele and by roughly 50 % in cells with two active *H3f3a-K27M* alleles. Notably, trimethylation at position H3K4 was completely unaffected in all cell populations from mutant knock-in mice, demonstrating that the observed epigenetic effects of the H3.3A-K27M mutation were restricted to H3K27. Besides, intracellular flow cytometry demonstrated robust 3xFLAG-tag, and thus H3.3A-K27M protein expression in all hematopoietic cells after Cre-mediated activation of the knock-in allele. Why MFI values for H3.3A-K27M-FLAG-tag expression are increased not just 2 fold, but rather ~4 – 5 fold in essentially all cells with two activated knock-in alleles is an interesting and unexpected observation that remains to be investigated.

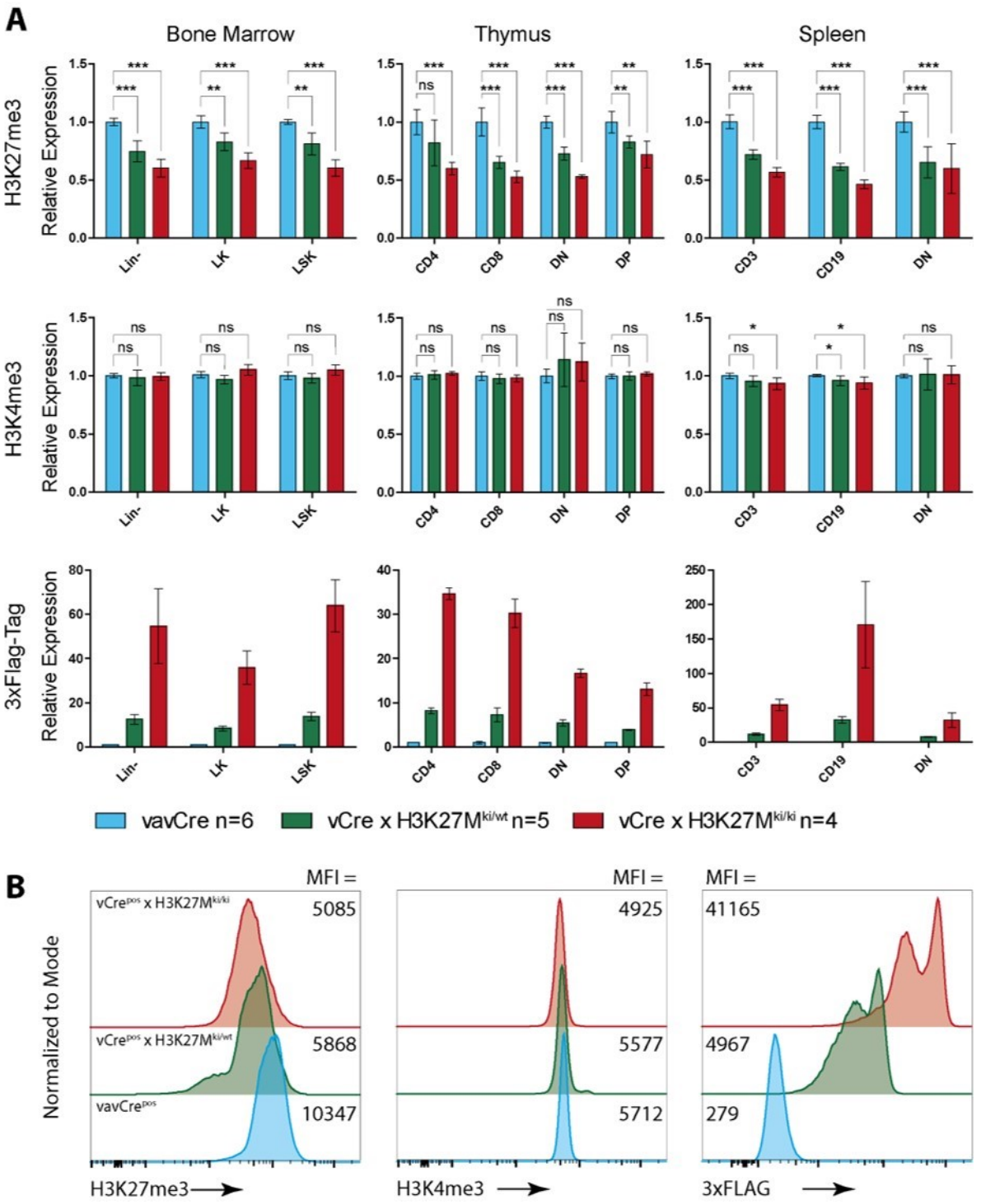


Figure 30: Cre-mediated H3.3A-K27M expression in hematopoietic cells from our knock-in mice reduces global H3K27me3

The color code remains the same in all panels: blue = vCre^{pos} mice, green = vCre^{pos} x H3K27M$^{KI/WT}$, red = vCre^{pos} x H3K27M$^{KI/KI}$. **(A)** The upper panels depict global trimethylation of the H3K27 methylation site. The average MFI values were obtained from three independent experiments. All data are normalized to MFI values obtained with cells from vCre mice (blue bar). Panels in the second row show results of analogous flow cytometric analyses for H3K4 trimethylation. The significance of observed expression differences was determined using Mann-Whitney-U t-tests. The third row shows the expression of the 3xFLAG-Tag. The number of mice analyzed for each genotype is given below the bar graphs. **(B)** Representative histograms of total splenocytes from one mouse per genotype. Equivalent results were obtained in all mice analyzed.

3.10 Reduced H3K27me3 does not affect the phenotype of H3K27M mutant mice

To determine whether the massive reduction of H3K27 trimethylation in vCre x H3K27M mice (see Figure 30) results in any phenotypic changes, flow cytometric analysis of lymphatic cell subsets of mice with relevant genotypes, was performed. This analysis was conducted in mice of two distinct age groups: younger mice aged 7 to 19 weeks, and older mice aged 68 to 95 weeks.

BM cells, thymocytes, and splenocytes of C57BL6/J, vavCre^{pos}, H3K27M$^{KI/WT}$, and H3K27M$^{KI/KI}$ mice as controls, as well as vCre^{pos} x H3K27M$^{KI/WT}$ and vCre^{pos} x H3K27M$^{KI/KI}$ knock-in mice, were analyzed. Representative results from the BM of younger mice are shown in Figure 31, and those of older mice in Figure 33.

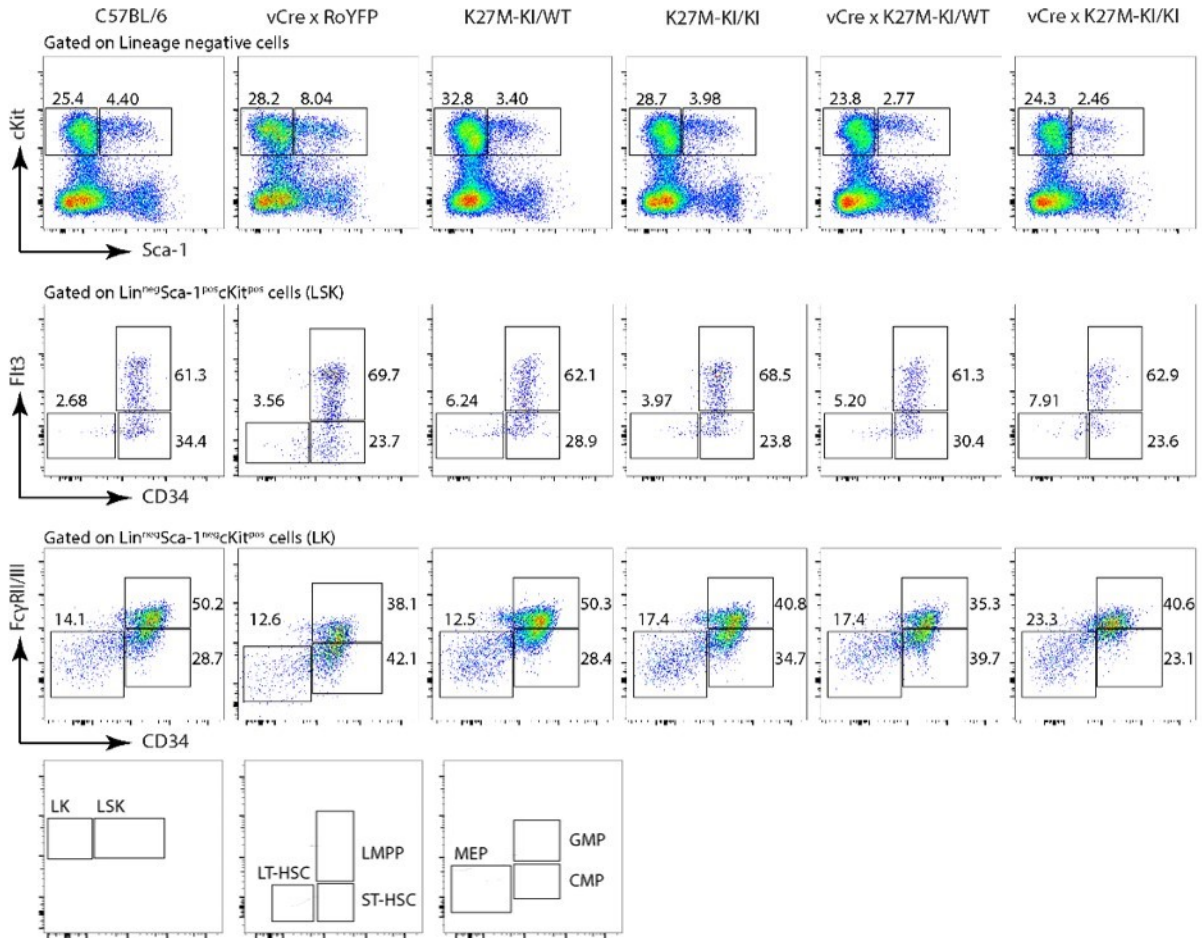


Figure 31: Representative results for flow cytometric of hematopoietic compartments in bone marrow of younger mice

The top row is gated on lineage-negative cells and gives the percentages of Lin^{neg}Sca-1^{neg}cKit^{pos} cells (LK) and Lin^{neg}Sca-1^{pos}cKit^{pos} cells (LSK). The second row is gated on LSK cells and shows percentages for early, totipotent stem cells: LT-HSC = Long-term hematopoietic stem cells, ST-HSC = short-term hematopoietic stem cells, LMPP = Lymphoid-myeloid primed progenitors. The third row is gated on LK cells and shows percentages for multipotent progenitor cells: MEP = Megakaryocyte-erythroid progenitor, GMP = Granulocyte-monocyte progenitor, CMP = common myeloid progenitor. A gating scheme indicating the respective cell populations is presented in the last row.

Depending on the availability of suitable genotypes at the time, the experiment was performed with 2 – 7 mice/genotype. For quantification and statistical analysis, the obtained percentages of each subset were summarized in Figure 32 and Figure 34, respectively. Mann-Whitney-U t-test was performed to indicate significant effects as compared to C57BL/6J and vCre^{pos} mice. Unless otherwise stated, the results were not significant. However, the sample cohort, especially for homozygous H3f3a-K27M mice, was relatively small. The inclusion of further mice could potentially still reveal significant changes in certain cell compartments.

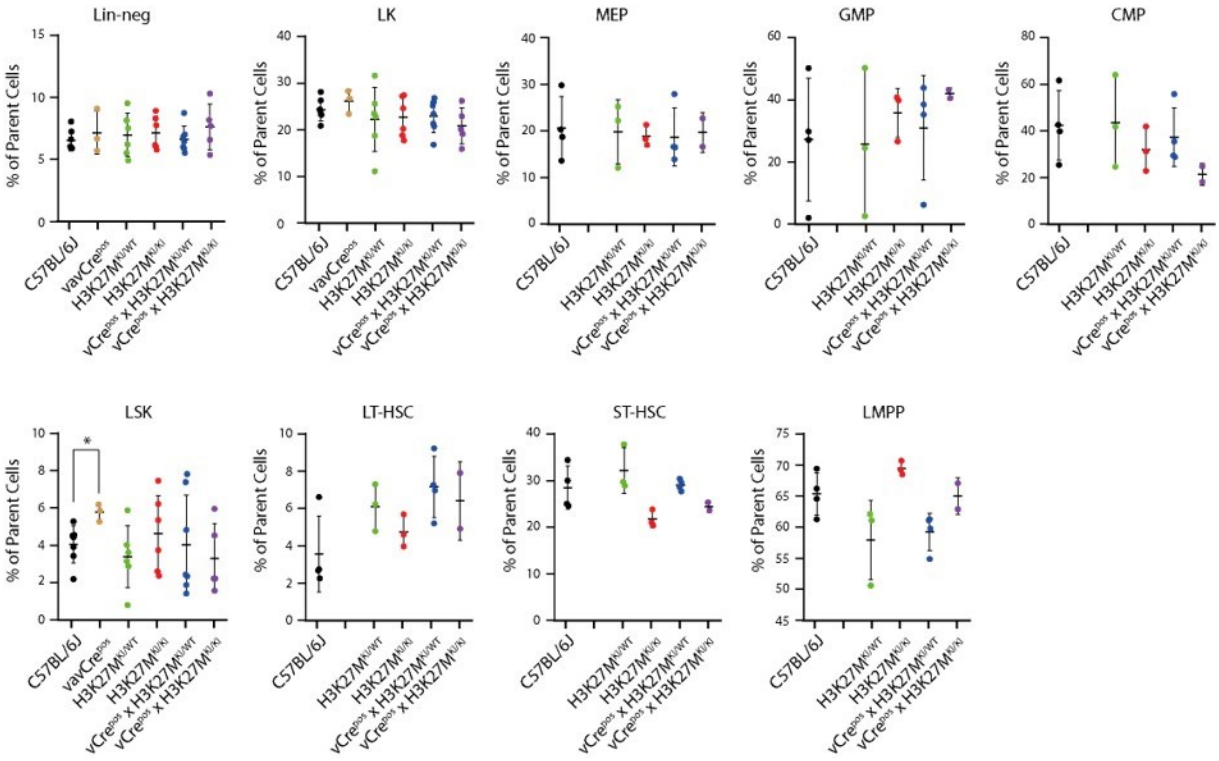


Figure 32: Quantification of cell subsets in bone marrow of younger mice
Comparison of bone marrow subsets does not reveal any significant phenotypic changes as compared to C57BL/6J mice. Just the LSK compartment shows a minor significant difference between vCre^{pos} and C57BL/6J mice. Each dot represents one mouse. Depending on the genotype n = 2 – 7. Shown are the mean ± SD. The significance was determined using Mann-Whitney-U test. Only significant values are given.

As stated above, analogous experiment were also performed with older compound mice, shown in Figure 33 and Figure 34. Due to a specific, time consuming breeding pattern, older H3K27M$^{KI/KI}$ and vCre^{pos} x H3K27M$^{KI/KI}$ mice were not available, yet. Hence, these genotypes could not be included into the analysis.

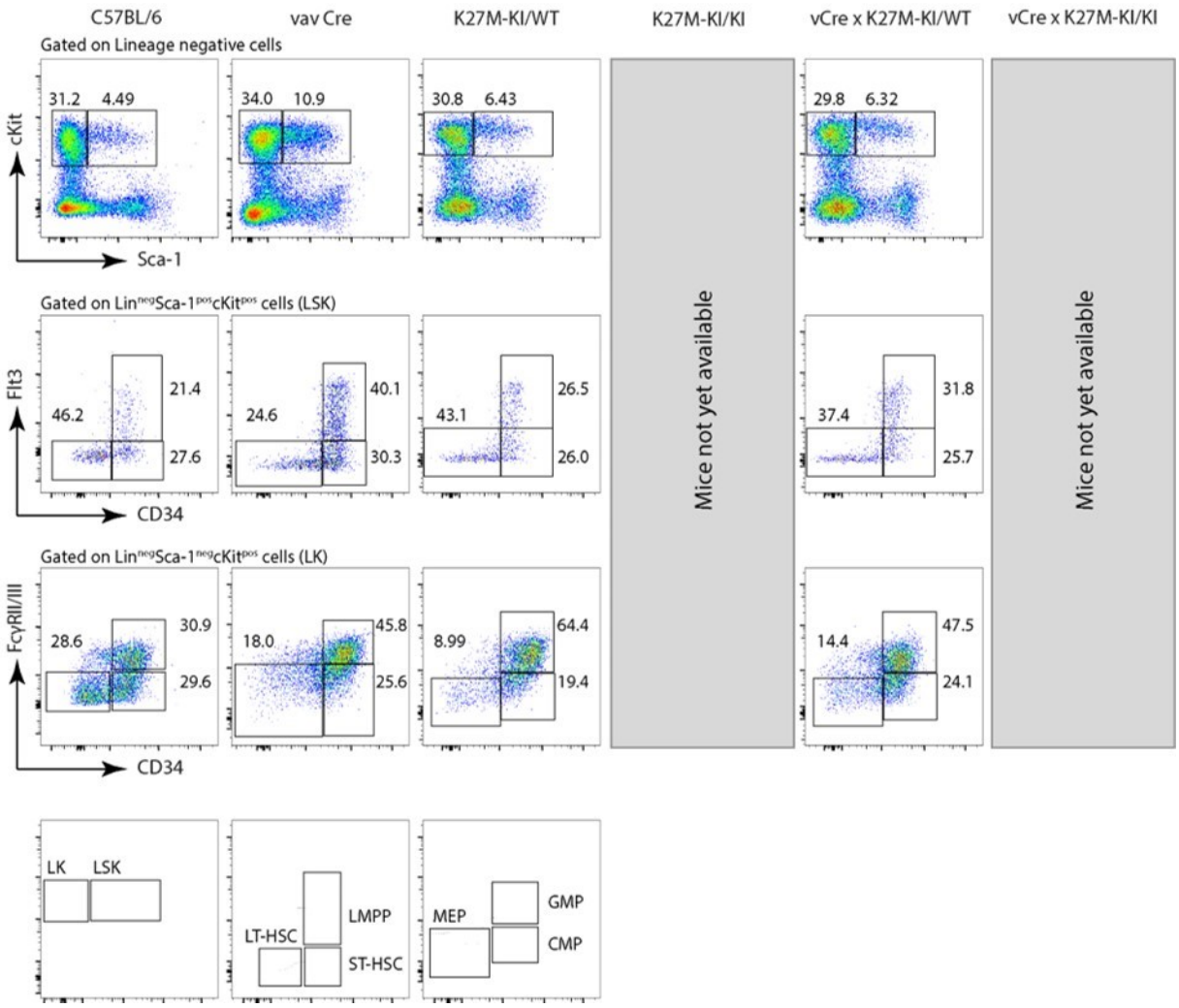


Figure 33: Representative results for flow cytometric analysis of hematopoietic compartments in bone marrow of older mice

Analysis was performed analogous to the results for younger mice as presented in Figure 31. The top row is gated on lineage-negative cells and gives the percentages of $Lin^{neg}Sca\text{-}1^{neg}cKit^{pos}$ cells (LK) and $Lin^{neg}Sca\text{-}1^{pos}cKit^{pos}$ cells (LSK). The second row is gated on LSK cells and shows percentages for early, totipotent stem cells: LT-HSC = Long-term hematopoietic stem cells, ST-HSC = short-term hematopoietic stem cells, LMPP = Lymphoid-myeloid primed progenitors. The third row is gated on LK cells and shows percentages for multipotent progenitor cells: MEP = Megakaryocyte-erythroid progenitor, GMP = Granulocyte-monocyte progenitor, CMP = common myeloid progenitor. A gating scheme indicating the specific cell types is presented in the last row.

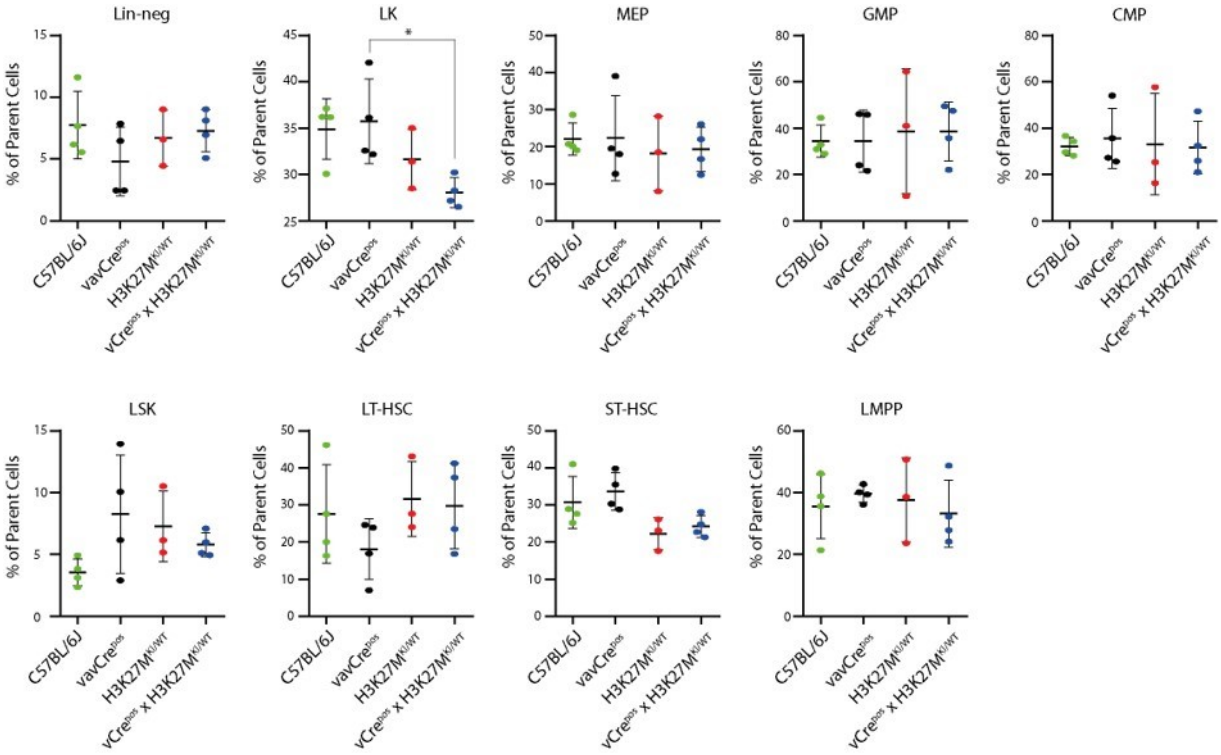


Figure 34: Quantification of cell subsets in bone marrow of aged mice
Comparison of bone marrow subsets does not reveal any significant phenotypic changes as compared to C57BL/6J mice. Just the LSK compartment shows a minor significant difference between vCre^{pos} and C57BL/6J mice. Each dot represents one mouse. Depending on the genotype n = 3 – 4. Shown are the mean ± SD. The significance was determined using Mann-Whitney-U test. Only significant values are given.

Analysis of thymus and spleen delivered similar, non-significant results and are presented in the appendix pages 140 - 145 Taken together, despite the strong reduction of H3K27me3, apart from minor differences of unknown significance, these findings do not suggest any phenotypic effects in the analyzed mouse cohorts. However, it cannot be excluded that larger cohorts, especially of even older vCre^{pos} x H2K27$^{KI/KI}$ mice would result in marked differences. This will be subject of further analysis.

3.11 H3K27M does not affect the survival and leukemia incidence in Mll5 and H3K27M compound mice

One major part of this project was the evaluation of the cooperativity of *Mll5*-deficiency and the *H3f3a-K27M* mutation in regard to leukemogenesis. To this extent vCre x Mll5$^{FL/WT}$ and vCre x H3K27M$^{KI/WT}$ mice were intercrossed, to generate double or triple mutants, respectively. These mice were kept for up to 20 months to analyze the onset of spontaneous leukemia. Survival curves of the relevant genotypes are given in Figure 35. vCre^{pos} mice (wild-type for both transgenes) and vCre^{neg} x H3K27M$^{KI/WT}$ mice were included as a control.

Figure 35 top panel compares vCre^{pos}, vCre^{neg} x H3K27M$^{KI/WT}$ mice to vCre^{pos} x H3K27M$^{KI/WT}$ and vCre^{pos} x H3K27M$^{KI/KI}$ mice. Log-Rank-Test was performed and revealed no significant differences in the survival of these mice. Figure 35 bottom panel compares vCre^{pos} x H3K27M$^{KI/WT}$ and x MLL5$^{FL/WT}$ double mutant mice to intercrosses heterozygous for both transgenes, i.e. triple mutants. Again, Log-Rank-Test did not show a significant increase in

mortality of these mice upon a combination of both mutations. However, again, it cannot be excluded that extended monitoring periods or larger mouse cohorts would potentially unveil an impact on survival and leukemia, eventually.

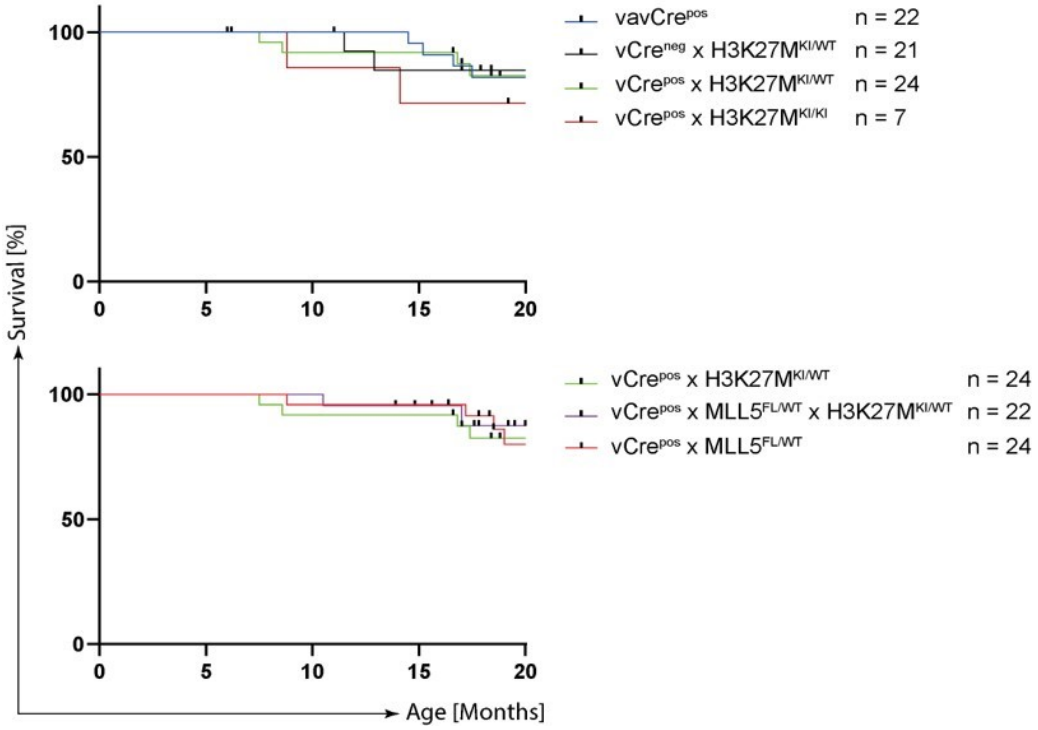


Figure 35: Comparison of survival curves from H3K27M-mutant and Mll5-deficient mice
Mice of the indicated genotypes were kept for 20 months. None of the included cohorts showed increased mortality as compared to the controls. Log-Rank-Test was performed and did not determine significant changes in overall survival. The number of included individual mice is indicated next to the respective genotypes.

As additional tool to screen for leukemia and pre-leukemic lesions, peripheral blood was collected at the time of death and blood parameters were determined using the Hemavet system (Drew Scientific). Clear-cut sign for leukemia would be increased levels of white blood cells (WBCs) or anemia. The obtained results for each analyzed moue are given in Figure 36. Of note, the grey shaded area indicates the normal range of parameters as provided by the manufacturer. However, the type automated cell analyzer can have an influence on the observed values (O'Connell et al., 2015). Furthermore, age and background strain specific differences have been published before (Fukuda et al., 2017; Silva-Santana et al., 2020). Hence, the seemingly overall increased levels of red blood cells (RBC) are not of major concern as they are consistent through all analyzed genotypes.

However, only a minority of mice exhibit abnormal blood parameters. Two vCre^{pos} mice show increased WBC counts in combination with decreased RBC counts, indicating (pre-)leukemic transformation in ~10 % of analyzed mice. Moreover, of all other mice, only one vCre^{pos} x H3K27M$^{KI/WT}$ mouse showed similar signs of disease, leading to the conclusion that these cases occur spontaneously and not due to *H3f3a-K27M* alone or in combination with the loss of *Mll5*.

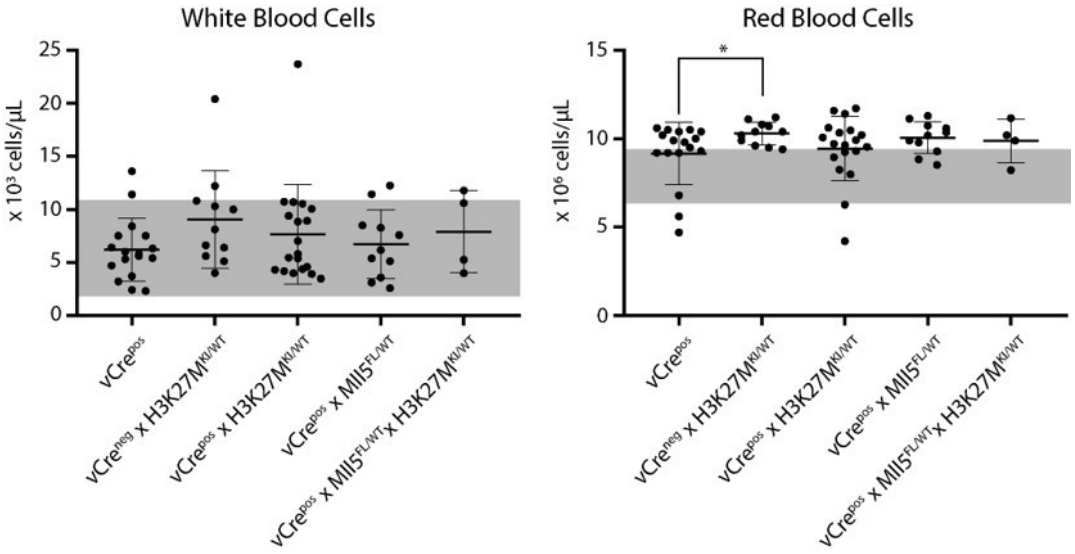


Figure 36: Blood cell count at death monitor mice
At the time of death, peripheral blood was collected from the portal vein and analyzed using the Hemavet system (Drew scientific). Grey shades indicate the normal range as provided by the manufacturer. Significance was determined using Mann-Whitney-U test, only significant values are given. For $vCre^{pos}$ n= 17; $vCre^{neg}$ x $H3K27M^{KI/WT}$ n = 11; $vCre^{pos}$ x $H3K27M^{KI/WT}$ n = 18; $vCre^{pos}$ x $Mll5^{FL/WT}$ n = 11; $vCre^{pos}$ x $Mll5^{FL/WT}$ x $H3K27M^{KI/WT}$ n = 4.

3.12 Introduction of IDH2-R172K does not cause leukemia in Mll5-deficient and H3K27M-mutant mice

In another line of investigation we set out to assess the effect of the IDH2-R172K mutation (see Introduction 1.6.1) via retroviral transduction of *H3K27M*-mutated and/or *Mll5*-deficient lineage-depleted bone marrow cells and the subsequent transplantation into lethally irradiated recipient mice. To this end, the successful generation of retroviral vectors as well as optimization and establishment of all experimental procedures, including the viral transduction of lineage-depleted BM cells (HSCs) as well as the transplantation of said cells was an essential part of this Ph.D. project.

A detailed description of the workflow is given in the methods section 2.2.17, while a simplified scheme is shown in Figure 37. In brief: $CD45.2^{pos}$ donor mice were sacrificed, the bone marrow was harvested, lineage depleted, transduced with retroviral supernatant, and injected into the tail vein of lethally irradiated, congenic $CD45.1^{pos}$ recipient mice.

The transplantations were performed competitively, meaning together with the transduced bone marrow cells, CD45.1/CD45.2 heterozygous wild-type whole bone marrow cells were injected. These cells are short-lived and help bridge the time the transplanted lineage-negative cells require to differentiate into functional cells and to repopulate the hematopoietic system (Duran-Struuck and Dysko, 2009). This is of special importance as *Mll5*-deficient BM cells have impaired reconstitution potential (Madan et al., 2009), i.e. mice would die within two weeks from radiation.

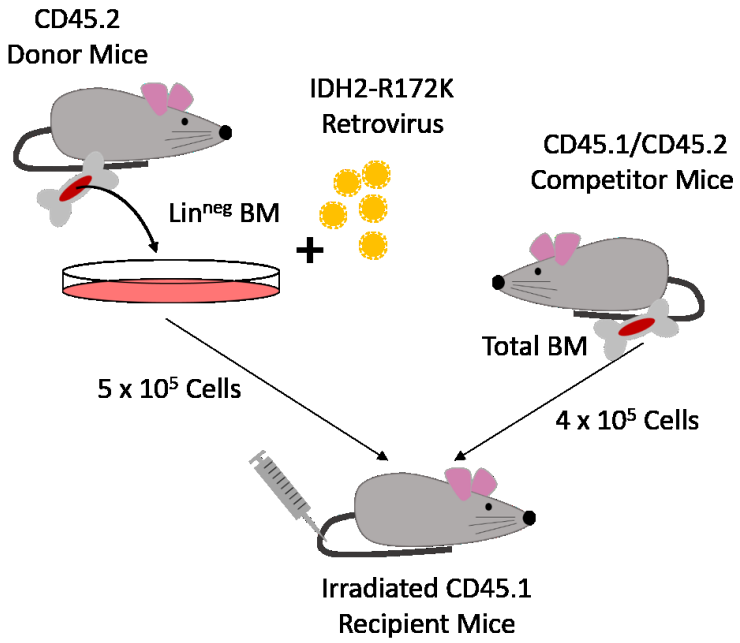


Figure 37: Simplified workflow of retroviral bone marrow transduction followed by transplantation
CD45.2 donor mice were sacrificed. The bone marrow was harvested, pooled according to genotype, and taken in culture. The next day, retroviral supernatants were used to transfect the cells with IDH2-R172K. On day five (transplantation day) CD45.1/CD45.2 heterozygous competitor mice were sacrificed to harvest the whole bone marrow. Together with 4 x 10^5 WT competitor cells, 5 x 10^5 transfected Lin^{neg} cells of the donor mice were injected into lethally irradiated CD45.1 recipient mice.

Donor mice of various genotypes were sacrificed during this project. As previously stated we used $vCre^{pos}$ x RoYFP mice as control since they exhibit the same vYFP expression pattern as our activated H3K27M mice. Furthermore, we included a pool of $vCre^{neg}$ control mice to account for any phenotypic effects caused be Cre itself. This pool includes mice of the following genotypes: C57BL/6J (n = 8), $vCre^{neg}$ x $Mll5^{FL/FL}$ (n = 3), $vCre^{neg}$ x $H3K27M^{KI/WT}$ (n = 4), and $vCre^{neg}$ x $Mll5^{FL/WT}$ x $H3K27M^{KI/WT}$ (n = 10). The experimental cohort includes homo- and heterozygous *Mll5*-deficient mice as well as heterozygous H3K27M mice. Of note, homozygous H3K27M mice were not available in sufficient numbers at the time these transplantation assays were conducted. Furthermore, $vCre^{pos}$ x $Mll5^{FL/WT}$ x $H3K27M^{KI/WT}$ mice were used as they correspond most closely to the findings observed in patients, where both genetic alterations occur heterozygously. The number of mice transplanted with the respective genotype is given in Table 15. Only mice surviving two weeks past the irradiation were included, as death within this time occurs due to radiation-induced bone marrow failure rather than as a result of the transplantation.

Table 15: Overview of transplanted genotypes
Only mice that survived two weeks past irradiation were included. The cohort of Creneg cotrols includes mice of the following genotypes: C57BL/6J (n = 8), vCre^{neg} x Mll5$^{FL/FL}$ (n = 3), vCre^{neg} x H3K27M$^{KI/WT}$ (n = 4), vCre^{neg} x Mll5$^{FL/WT}$ x H3K27M$^{KI/WT}$ (n = 10)

Transplanted Genotype	Number of Transplanted Mice
vCre^{neg} controls + IDH2-R172K	25
vCre^{pos} x RoYFP + IDH2-R172K	21
vCre^{pos} x MLL5$^{FL/WT}$ + IDH2-R172K	8
vCre^{pos} x MLL5$^{FL/FL}$ + IDH2-R172K	8
vCre^{pos} x H3K27M$^{KI/WT}$ + IDH2-R172K	27
vCre^{pos} x MLL5$^{FL/WT}$ x H3K27M$^{KI/WT}$ + IDH2-R172K	28

After successful transplantation, blood was drawn once a month from the tail vein to evaluate engraftment and expression of the introduced IDH2-R172K mutation via the expression of corresponding fluorescent marker using flow cytometry. Additionally, this blood was used to measure blood parameters at the Hemavet System to monitor leukemia development.

Lymphocytes originating from donor, recipient, and competitor (= helper) mice can be discriminated by their distinct CD45 expression. All donor mice used during these transplantations were on a C57BL/6J background and therefore can be identified based on their CD45.2 expression. The recipient mice used were congenic B6.SJL mice expressing CD45.1 on their surface. The helper cells were harvested from C57BL/6J x B6.SJL intercrosses, hence being CD45.1/CD45.2 heterozygous. An example of the applied gating strategy is presented in Figure 38. The panels of the upper row show gating for living, single lymphocytes. The first dot plot in the bottom row shows separation based on CD45 expression. The histograms on the left represent the expression of the fluorescent markers: vYFP (measured in the FITC channel) as a marker for K27M and PE-Texas Red as a marker for the virally introduced IDH2-R172K mutation. Here, it needs to be pointed out, that due to the relatively small shift in vYFP expression in H3K27M heterozygous mice (see Figure 26), marker expression does not allow for the indication of reliable percentages for vYFP-positive cells. However, based on FACS analysis of bone marrow from vCre^{pos} x RoYFP mice (Figure 23) it can be assumed that nearly 100 % of CD45.2 lymphocytes have been activated by vCre to express the H3K27M mutation. In contrast, the expression of PE-Texas Red is much more intense, allowing us to directly determine the IDH2-R172K expression within the donor cells.

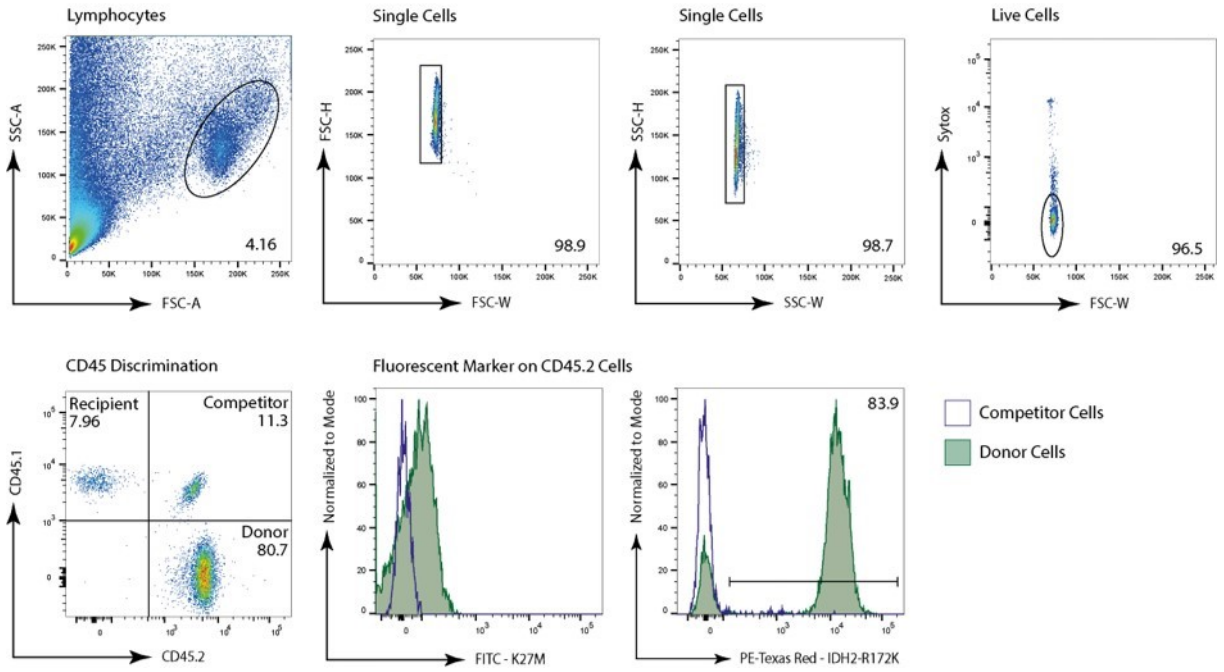


Figure 38: Gating strategy for monthly blood analysis
Staining was performed on lysed blood cells to determine engraftment and expression of the mutation. The upper row shows gating for living single cells. Bottom left dot plot shows discrimination of donor, recipient, and competitor mice according to CD45 expression. Histograms show expression of the fluorescent markers in competitor cells (blue outline) and donor cells (green) Values given in the plots represent percentages of the parent population.

The results of the monthly blood analysis are summarized in Figure 39 for the Cre^{neg} and $vCre^{pos}$ x RoYFP control cohorts and in Figure 40 for all other genotypes. Mice transplanted during the individual transplantations were grouped according to the transplanted genotype. Panels on the left indicate the percentage of circulating lymphocytes derived from the donor cells. The panels on the right show how many of these cells express the IDH2-R172K mutation according to the expression of the fluorescent marker. Each line represents one mouse and the color code remains the same within one genotype.

The engraftment as well as the expression of the retrovirally introduced IDH2-R172K mutation are highly ambiguous. In all transplanted genotypes, with the exception of $vCre^{pos}$ x $H3K27M^{KI/WT}$ (where all but one mouse show efficient engraftment > 60 %; Figure 40 fourth row) and the $vCre^{neg}$ mice (with only three mice, showing engraftment < 10 %; Figure 39 left panel), about half of the mice indicate a loss of cells of donor origin, while the other half exhibit high engraftment levels. One reason for this effect might be incomplete ablation of the recipient's BM, possibly due to minor experimental variations during irradiation (Duran-Struuck and Dysko, 2009). Strikingly, 50 % of the mice transplanted with *Mll5*-deficient bone marrow show good engraftment levels. This is somewhat unexpected, as $Mll5^{-/-}$ HSPCs exhibit severe defects in their ability to reconstitute lethally irradiated mice (Madan et al., 2009). However, this further underlines the importance of including the wild-type helper cells to bridge to the time until the transduced donor cells differentiate into fully functional cells.

Similar inconsistent results are shown for the expression levels of the retrovirus. In $vCre^{neg}$ mice, nearly all animals exhibit marker expression in less than 20 % of circulating lymphocytes at death (Figure 39, top right panel). The same holds true for $vCre^{pos}$ x $Mll5^{FL/WT}$ mice, were

expression in all mice is < 10 % (Figure 40 second row, right panel). In all other transplanted genotypes, about half of the mice show high levels of IDH2-R172K expression, while the other half show a decrease in expression, with percentages around 1 – 2 % at 12 months after transplantation.

Interestingly, good engraftment does not necessarily correlate with a high expression of the IDH2-R172K mutation, as can be nicely seen in the vCre^{pos} x H3K27M$^{KI/WT}$ cohort in Figure 40, fourth row. The reasons for this observation are manifold as they can be attributed for example to experimental aberrations (e.g. low-quality supernatants) or an unfavorable integration site of the virus into the cell genome. Furthermore, the fluorescence intensity of the IDH2-R172K marker varied among positive cells, indicating differences in the virus copy number within individual cells (Halene et al., 1999).

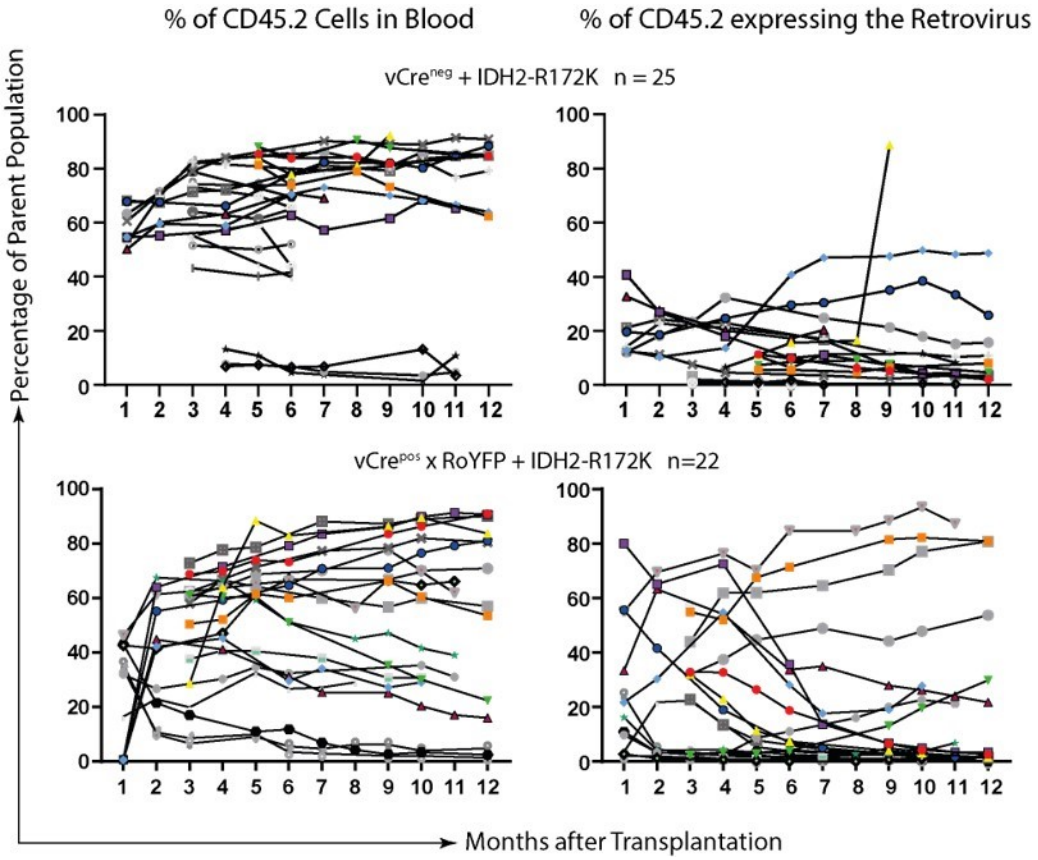


Figure 39: Engraftment and expression of the retrovirus in donor derived cells in control mice
Each line represents one individual mouse, color code is true per genotype. Left panels indicate the percentage of donor derived circulating lymphocytes, right panels show how many of these cells express the retrovirally introduced IH2-R172K mutation. Included genotypes are: C57BL/6J (n = 8), vCre^{neg} x Mll5$^{FL/FL}$ (n = 3), vCre^{neg} x H3K27M$^{KI/WT}$ (n = 4), vCre^{neg} x Mll5$^{FL/WT}$ x H3K27M$^{KI/WT}$ (n = 10)

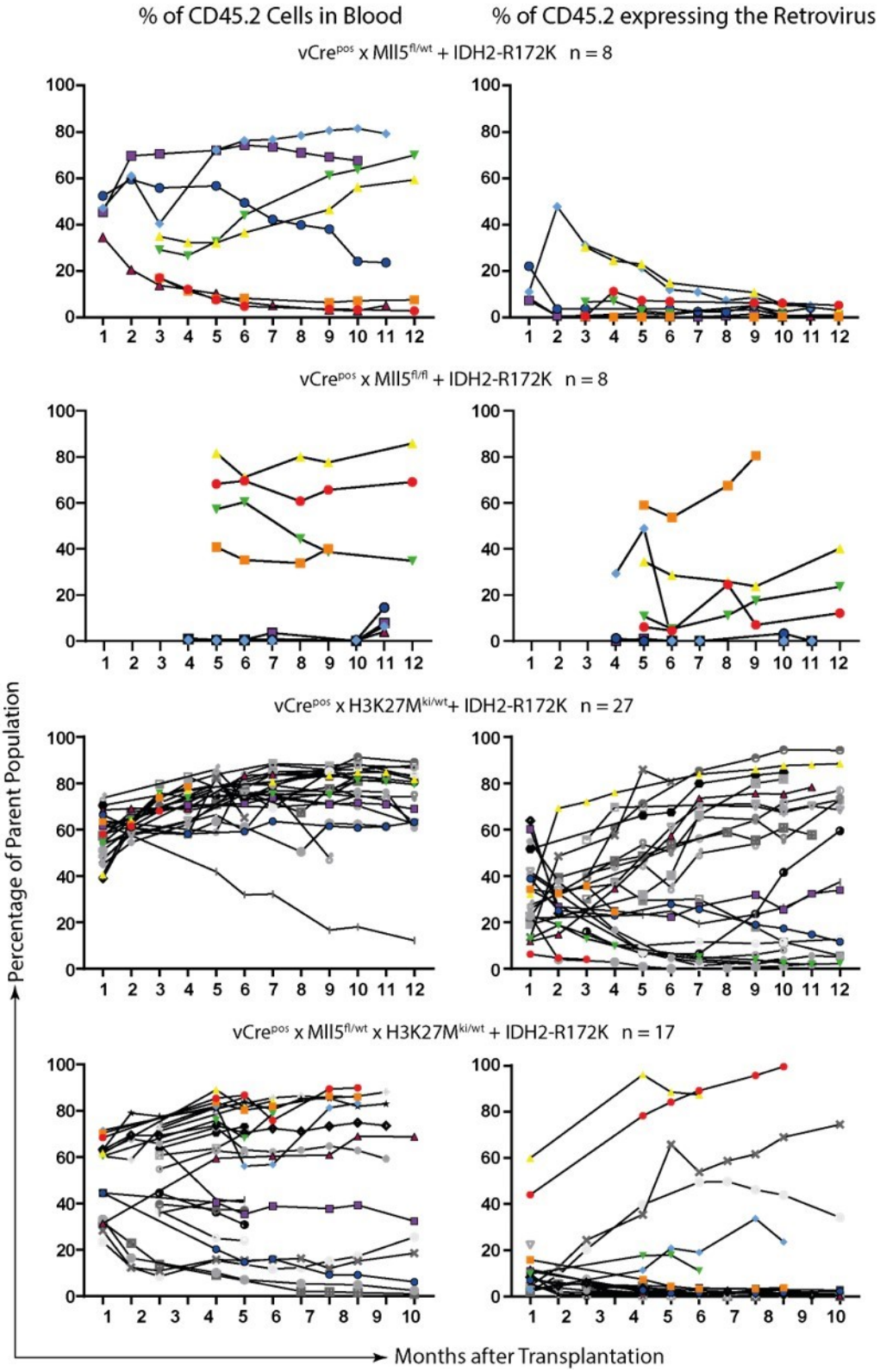


Figure 40: Summary of engraftment and expression of the retrovirus in donor derived cells
Each row of graphs represents one transplanted genotype. The transplanted genotype and the number of recipient mice are given above each set of graphs. Each line represents one individual mouse, color code is true per genotype. Left panels indicate the percentage of donor derived circulating lymphocytes, right panels show how many of these cells express the retrovirally introduced IH2-R172K mutation.

Post transplantation care of mice includes daily visual health check-ups, as well as weekly weighing to track potential weight loss due to disease progression. To enable easier comparison of weight gain/loss, the relative weight as compared to the starting weight was calculated. The results are shown in Figure 41, grouped according to the transplanted genotype. Graphs showing the weight history for each individual mouse are given in the appendix on page 146.

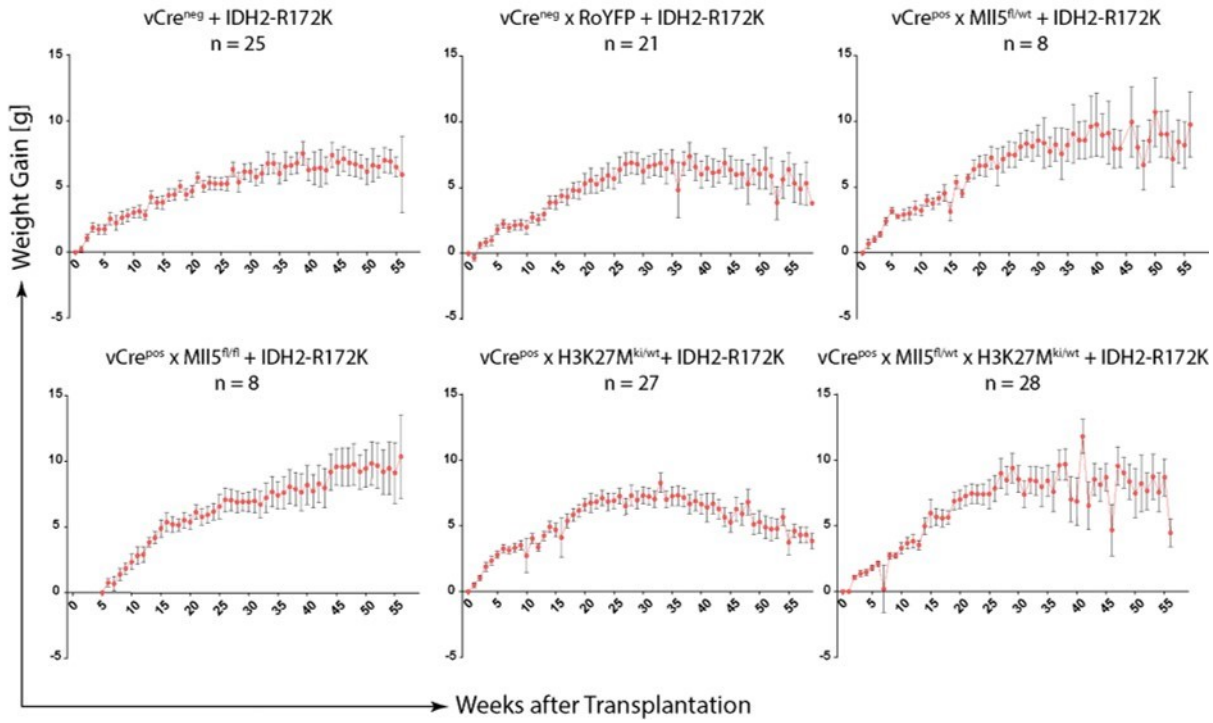


Figure 41: Relative weight change of transplanted mice per genotype
Transplanted mice were weighted weekly. The graphs represent the relative weight difference as compared to the starting weight. The red dots represent the mean value, error bars indicated mean ± SEM. The genotype as well as the number of mice is given above each graph.

Notably, in all transplanted genotypes, except for vCre^{pos} x Mll5$^{FL/FL}$ mice, mice gained weight up to 40 – 45 weeks post-irradiation/transplantation, before they gradually start losing weight again. Partially, these effects can be explained due to the fact that individual mice became sick and lost weight, hence the mean weight also slightly decreased.

The assumption that weight loss indicates the onset of any disease and therefore would be an important prognostic factor, turned out to be true. Individual mice, that rapidly lost weight, were sacrificed. Upon necropsy, these mice showed the aforementioned signs of disease, including aberrant blood cell counts, light-colored BM, and splenomegaly. Furthermore, these mice frequently developed severe skin lesions on their back and upper body regions. Exemplarily, Figure 42 shows the weight curves for all individual mice transplanted with IDH2-172K transduced vCre^{pos} x Mll5$^{FL/WT}$ x H3K27M$^{KI/WT}$ bone marrow cells. Four mice (yellow, orange, green, and red-black graphs) rapidly lost > 10 % of their body weight within 5 – 10 weeks and were sacrificed.

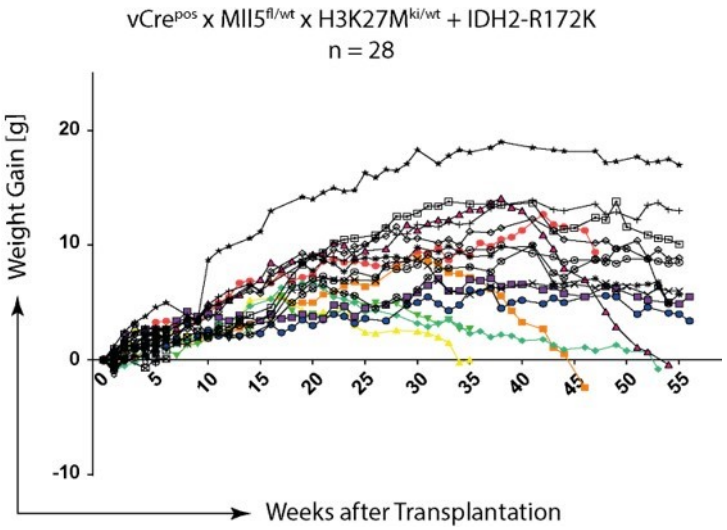


Figure 42: Relative weight gain of all mice transplanted with IDH2-172K transduced vCre^{pos} x Mll5$^{FL/WT}$ x H3K27M$^{KI/WT}$ bone marrow
Each line represents one individual mouse. Mice with a weight loss > 10% were sacrificed and to analyze lymphatic cells via flow cytometry and blood cell counts via an automatic cell analyzer system.

Twelve to fourteen months past transplantation, all remaining mice were sacrificed irrespectively of their health status. Lymphatic organs (BM, thymus, spleen, blood, and lymph nodes) were harvested and analyzed using flow cytometry to analyze disease progression. Furthermore, whole blood was analyzed using the Hemavet system.

The survival curve for all transplanted genotypes is given in Figure 43. The upper graph compares survival of the control cohorts (vCre^{neg} and vCre^{pos} x RoYFP) with mice transplanted with transduced bone marrow cells hetero- or homozygous for the loss of *Mll5*. Neither of these genotypes show any increase mortality. Interestingly, in the vCre^{pos} x Mll5$^{FL/FL}$ cohort not one single mouse died prematurely – potentially due to the generally low expression levels of the retrovirus. The lower graph compares the aforementioned control groups to mice transplanted with bone marrow heterozygous for the loss of *Mll5* and/or *H3f3a-K27M*. Again, no significant changes in the overall survival were observed.

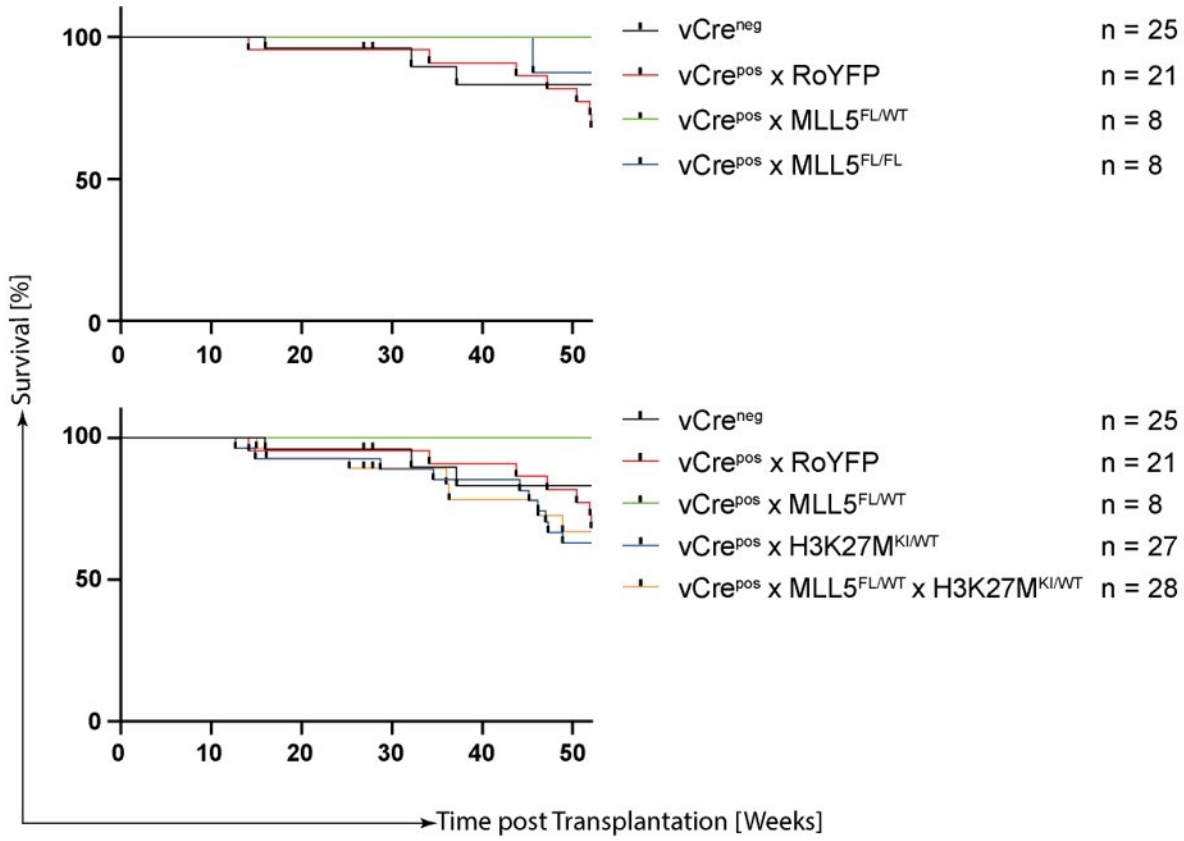


Figure 43: Comparison of survival of transplanted mice
Mice were transplanted with IDH2-R172K transduced bone marrow of the indicated genotype. The number of mice is indicated. Top panel compares survival of vCre^{pos} x RoYFP control mice to *Mll5* and *H3f3a* single and double mutants. Lower panel compares control mice to homozygous and heterozygous loss of *Mll5*. Significance was determined using Log-rank test and did not reveal any significant differences in survival.

Overall, the observed survival rates do not suggest that the cooperation of loss of *Mll5* and the specific mutations in *H3f3a* and *Idh2* leads to an increased death rate due to leukemia. However, high survival rates are proof for the successful transplantation procedure as graft failure would have resulted in much higher numbers of deceased mice.

As another tool to determine a potential onset of leukemia, analysis of total blood cell count at death was performed and revealed a significant increase of WBC in mice transplanted with H3K27M$^{KI/WT}$ + IDH2-R172K bone marrow as compared to vCre^{pos} x RoYFP mice, but not to the vCre^{neg} pool (Figure 44, top left). Furthermore, mice of these cohorts show a (albeit not significant) decrease in RBC counts (Figure 44, top right). Both phenotypes can hint at pre-leukemic lesions that could potentially develop into leukemia upon longer observation periods. Interestingly, these changes were not observed in H3K27M$^{KI/WT}$ mice in the absence of the IDH2 mutation (Figure 36), indicating a potential negative effect of IDH2-R172K.

However, in mice double mutant for *H3f3a* and *Mll5*, these effects were partially reversed (i.e. less mice with increased WBC count and reduced RBC count), hinting at a potential, somewhat antagonistic role of loss of *Mll5* and the *Idh2-R172K* mutation.

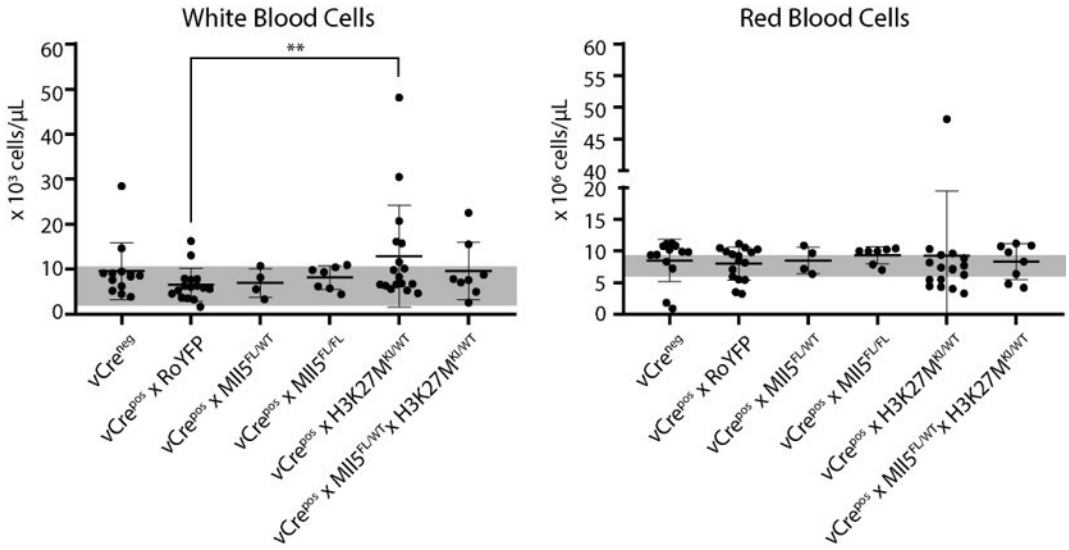


Figure 44: Blood cell count at death of mice transplanted with IDH2-R172K transduced bone marrow cells
At the time of death, peripheral blood was collected from the portal vein and analyzed using the Hemavet system (Drew scientific). Grey shades indicate the normal range as provided by the manufacturer. Significance was determined using Mann-Whitney-U test, only significant values are given. For $vCre^{neg}$ n = 13 $vCre^{pos}$ x RoYFP n = 14; $vCre^{pos}$ x $Mll5^{FL/WT}$ n = 4; $vCre^{pos}$ x $Mll5^{FL/FL}$ n = 7; $vCre^{pos}$ x $H3K27M^{KI/WT}$ n = 17; $vCre^{pos}$ x $Mll5^{FL/WT}$ x $H3K27M^{KI/WT}$ n = 8.

4. Discussion

4.1 Generation of knock-in mice with a Cre-inducible H3f3a-K27M mutation

Although acute myeloid leukemia (AML) is a genetically highly heterogeneous disease, only a limited number of driver mutations have been identified. The identification of causative mutations is essential in the search for potential treatment options. Patients with a particularly severe course of the disease often have deletions in chromosome 7 (Inaba et al., 2018), consistently including loss of *MLL5* (see Table 1,page 10). Despite the seemingly obvious assumption that loss of *MLL5* must be contributing in some ways to leukemogenesis, its role in AML development and progression remains unclear. Although the hematopoietic compartment of mice with *Mll5* deficiency shows a phenotype resembling pre-leukemic lesions, these mice do not develop spontaneous leukemia even after prolonged observation (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009). Still, this finding does not exclude an influence of MLL5 deficiency on leukemogenesis, as it is possible that additional cooperating mutations may be necessary. A widely established theory for the development of leukemia is the so-called two-step hypothesis. This hypothesis states that at least one mutation within each of two classes is necessary to cause the disease: Class I mutations induce the proliferation of aberrant HSPC clones, while class II mutations inhibit the differentiation of these cells (Gilliland and Griffin, 2002).

To identify mutations potentially cooperating with the loss of *MLL5*, the whole exome of patients with heterozygous chromosome 7 deletions (including *MLL5*) was analyzed by DNA sequencing (Konstanze Döhner and colleagues, Internal Medicine III, University Hospital Ulm). Again, exome sequencing of these patients revealed only a limited number of recurrent mutations (Table 2 - Table 4, pages 11 and 17). Of particular interest in the context of *MLL5* deficiency appear the mutations identified in *IDH2*, *DNMT3A*, and *H3F3A*, as explained in detail in introductory chapter 1.6, pages 10 ff.. To further investigate these mutations alongside loss of *Mll5* in mice, mouse models were established utilizing retroviral transduction of bone marrow cells followed by competitive transplantation, and the generation of a novel *H3f3a-K27M* mutant mouse via traditional gene targeting.

Although the *H3F3A-K27M* mutation has been described and intensely studied in connection with pediatric high-grade gliomas (Bender et al., 2013; Chan et al., 2013; Harutyunyan et al., 2019; Lu et al., 2018; Panditharatna et al., 2022; Wang et al., 2021), its role in leukemogenesis had not been explored at the beginning of this project, making it a promising target to study its contribution to AML, in particular in the context of *MLL5* deficiency. To ensure proper expression of the engineered mutation at physiological levels, one central part of this project was the generation of a novel, inducible knock-in mouse model carrying the previously identified mutation in *H3F3A-K27M* exclusively activated in hematopoietic cells, as such a mouse model was not available at the time. To this end, we adapted a targeting strategy based on Cre-mediated conditional allelic replacement that had previously been published by Tang et al. (2013). In addition to the relevant *H3f3a-K27M* mutation, we introduced the IRES-driven fluorescence marker vYFP (Nagai et al., 2002), which allows for non-invasive tracking of *H3K27M* expressing cells on single cell levels, and a protein tag (3xFLAG) to distinguish proteins encoded by *H3f3a* and *H3f3b* from each other (see introduction 1.6.3).

Notably, soon after the beginning of this project, an independently generated inducible *H3f3a-K27M* mouse model has been published to investigate the role of H3.3-K27M in brainstem gliomas (Larson et al., 2019), with more details on the strategy explained by Larson and Baker (2019). The report demonstrated – albeit exclusively for the nervous system – that an *H3f3a-K27M* mutation significantly enhances self-renewal of neuronal stem cells and accelerates the formation of hindbrain tumors. Furthermore, they were able to show that the cooperativity with certain mutations shifts the spectrum of expected tumors. These findings seemed very encouraging, as they provided a proof-of-concept for our “leukemia-focused” study.

Unlike our mouse model, where only exon 2 is manipulated, Larson et al. inserted the selection markers into intron 1 (upstream of exon 2), together with a FLAG/HA protein tag into exon 4. The resulting mice were intercrossed with *Nestin*-Cre mice, which constitutively express Cre in all neural stem and progenitor cells (Tronche et al., 1999). Knock-in mice were also intercrossed with *Nestin*-CreERT2 mice, allowing conditional Cre-mediated activation following intraperitoneal injections with tamoxifen. Notably, as significant improvement, our model would not only possess a protein tag but also the unique feature of an additional fluorescence marker (vYFP), which enables precise tracking of viable cells expressing *H3f3a-K27M*, thereby expanding the range of potential applications for this model.

Upon gene targeting, we noticed a remarkable aberrancy from apparently well-established rules regarding the effectiveness of homologous recombination when using the three different constructs we had generated for this project: **(i)** pKK-H3TG-AD (H3K27M only), **(ii)** pKK-H3TG-AF (H3K27M and vYFP), and **(iii)** pKK-H3TG-AG (H3K27M, vYFP and FLAG tag). Generally, the frequency of homologous recombination considered to be is influenced by three main factors: **(i)** positively be the degree of homology between sequences present within the targeting construct and the genomic target site, negatively by the size of the non-homologous insert (Deng and Capecchi, 1992), **(ii)** isogenicity of the introduced DNA (te Riele et al., 1992), and **(iii)** locus-dependent differences in chromatin structure (Müller, 1999).

In our hands, targeting with pKK-H3TG-AD and pKK-H3TG-AF resulted in homologous recombination in 3.6 % and 1.3 % of analyzed ES cell clones, respectively, which is in line with the idea that the longer non-homologous inserts (4.6 kb versus 6 kb) negatively affect targeting frequencies. However, unexpectedly, the most complex construct with the largest non-homologous insert (6.1 kb) resulted in dramatically higher targeting efficiency with homologous recombination observed in 21 % of analyzed ES cell clones (Table 9, page 67). Notably, this observed frequency closely matches the 26 % reported by Tang et al. (2013), on whose strategy our targeting construct is based. One potential explanation for this massively increased frequency of homologous recombination with the most complex construct AG could be linked to position effects that influence accessibility of the target locus and unknown “steric requirements” of the homologous recombination reaction. Whatever the deeper molecular reason for this counter-intuitive result maybe, our finding demonstrates that it is not always possible to reliably predict the efficiency of homologous recombination from modified variants of a basic targeting construct. Our example thus highlights that one should not be deterred to increase the complexity of a targeting vector, if deemed scientifically useful, due to theoretical considerations that might predict a too low success rate in homologous recombination. Eventually, only the experiment itself will tell!

As targeting with pKK-H3TG-AD served more or less as proof of principle and also to (re-)establish the required procedures and methods in our laboratory, positively targeted ES cell clones were not further characterized. Instead, selected ES clones targeted with pKK-H3TG-AF and pKK-H3TG-AG were thoroughly evaluated using a variety of methods. Notably, only results obtained with the final construct pKK-H3TG-AG will be discussed here, as only ES cells derived from targeting with this construct were subsequently used for the generation of the *H3f3a-K27M* mutated mouse.

FACS analysis of targeted ES clones was performed and revealed reliable expression of the fluorescent reporter vYFP exclusively upon Cre-mediated recombination in several independently activated clones (Figure 18B, page 68), indicating marked expression of the targeted K27M-mutated allele. Mean fluorescence intensity (MFI) values for **(i)** entirely non-manipulated, **(ii)** targeted (but not yet Cre-activated) and **(iii)** Neo deleted cells were equally low (Figure 18A, page 68), thereby ruling out unintentional activation or leaky reporter gene expression. Whether the expression level of the K27M-mutated allele equals expression of the non-manipulated allele cannot be determined by flow cytometry as appropriate control ES carrying just an IRES-vYFP insertion in the 3'-untranslated region of an otherwise unmodified *H3f3a* allele are not available.

To confirm expression of the targeted allele at transcript level, semi-quantitative RT-PCR was conducted (Figure 19A and B). Results indicated expression of the WT and the targeted, K27M-mutated allele at roughly comparable levels, suggesting that the targeted manipulation of the *H3f3a* allele did not markedly affect its transcriptional activity. In addition, a second RT-PCR of the full-length *H3f3a* transcript was performed followed by sequencing of the product (Figure 20, page 70). Corresponding results provide unequivocal evidence that the targeted allele is expressed with the predicted sequence over its entire length encoding indeed the introduced K to M amino acid exchange at position 27. Furthermore, Western Blot experiments (Figure 19C, page 69) indicated similar H3.3 protein levels across cells of all genotypes, indicating roughly wild-type expression of total H3.3 protein in manipulated ES cells.

Moreover, the Western Blot served as another method to verify the expression of the introduced markers on a protein level – especially of the 3xFLAG tag. As previously stated, this is the only possibility to discriminate H3.3A-K27M mutant protein from the *H3f3a* and *H3f3b* encoded wild-type. The fact that a prominent band with the predicted increase in protein size (15 kDa from the non-targeted protein vs. 18.5 kDa including 3XFLAG tag) was readily detectable by Western Blotting (Figure 19C, page 69) provided conclusive evidence that the K27M-mutated H3f3a allele was not only expressed as transcript, but also robustly at the protein level, albeit exclusively after Cre-mediated activation, as predicted. At the same time, this affirmative result highlighted that the 3xFLAG tag perfectly fulfilled its intended purpose. Additionally, the YFP protein was detected at the expected size of 25 kDa.

After successful characterization of targeted ES cells, clones that were obtained with vectors pKK-H3TG-AF and pKK-H3TG-AG were used for blastocyst injection with the goal to generate knock-in mice carrying the inducible *H3f3a-K27M* mutation (Table 10 and Table 11, page 71). The observed low number of agouti offspring (Table 11, page 71), is unexpected as the ES cells used (Nagy et al., 1993) are generally known for robust contribution to chimerism and were specifically provided by the transgenic core facility of the MPI/Dresden. Despite this,

germline transmission was accomplished and we obtained a stable line of mice heterozygous for H3K27M derived from ES cells that had been targeted with pKK-H3TG-AG. The resulting heterozygous H3K27M mice were set up for breeding and did not reveal any hint on adverse effects of the targeted allele on germ cell function or embryogenesis, as anticipated genotypes were obtained at Mendelian frequencies (Table 13, page 73). This finding was consistent with published data (Tang et al., 2013).

Heterozygous H3K27M mice were intercrossed with vCre mice to activate the mutation exclusively in hematopoietic cells (de Boer et al., 2003; Ogilvy et al., 1999). The genotype distribution after intercrossing did not indicate any overall lethality due to activation of the targeted locus (Table 14, page 73). Prior to intercrossing of H3K27M mice with vCre mice, we had analyzed vCre x RoYFP mice as controls (carrying the Cre-inducible vYFP-encoding cDNA in the ubiquiteously expressed ROSA26 gene locus), confirming virtually 100 % efficient and hematopoietic cell-specific activation of Cre-recombinase in the particular transgenic strain of vCre mice used in our experiments (Figure 21 - Figure 23, pages 74 - 76).

To confirm the results obtained from the characterization of targeted ES cells, analogous experiments were performed to verify the *H3f3a-K27M* expression in vCre x H3K27M mice. First, RT-PCR was performed on RNA extracted from thymocytes and splenocytes of $vCre^{pos}$ x $H3K27M^{KI/WT}$ mice to confirm full-length expression of the full-length *H3f3a* transgene, including the in-frame expression of the mutated codon 27 (Figure 24, page 77). Furthermore, it was shown that *H3f3a-K27M* transcripts are exclusively detected in hematopoietic tissues of vCre positive mice (Figure 25, page 78) – as predicted and perfectly in line with the results obtained from the ES cells.

Co-expression of the fluorescent marker vYFP was again monitored by flow cytometry. As expected from the analysis of vCre x RoYFP mice, marker expression in vCre x H3K27M mice was only detectable in hematopoietic tissues after Cre-mediated recombination (Figure 26, page 79), thus confirming the strictly Cre-dependent expression of *H3f3a-K27M* as demonstrated before by RT-PCR (Figure 25, page 78). Additionally, cells from homozygous mice exhibit MFI values approximately twice as high as those observed from heterozygous ones (Figure 26, page 79), providing further evidence for the reliability of the IRES-driven vYFP reporter. Moreover, within specific hematopoietic cell subsets, histograms of vYFP-derived signals consistently exhibited narrow, monophasic peaks with cell-type-specific MFI values, indicating that *H3f3a* expression is uniform within a specific cell type, but may vary between distinct cell types. Representative examples are shown in Figure 27A – C, on page 80: B lymphocytes ($CD19^{pos}$ cells) from lymph nodes as well as spleen exhibit vYFP-specific MFI values more than twice as high as gated T lymphocytes ($CD3^{pos}$ cells), suggesting more than 2-fold stronger *H3f3a* expression in B versus T cells. This notion was neatly confirmed in three separate qRT-PCR experiments with RNA of T-and B-lymphocytes sorted from lymph nodes and spleen of wild-type mice (Figure 27D, page 80). These data provide another piece of conclusive evidence that our vYFP reporter in K27M knock-in mice faithfully reflects endogenous *H3f3a* expression levels in wild-type mice. Concomitantly, our data provide strong circumstantial evidence that the genetic manipulation associated with the introduction of the K27M mutation, the IRES-vYFP reporter and the 3xFLAG tag did not alter natural differences in *H3f3a* expression levels, at least in T and B lymphocytes that were scrutinized here. Our K27M knock-in mice thus not only model a recurrent oncogenic *H3f3a* mutation, but they can

also be used to monitor cell type-specific *H3f3a* expression levels non-invasively, i.e. in individual living cells at distinct stages of development. Information on endogenous *H3f3a* expression levels in different cells is a prerequisite for revealing potential cell-type-specific functions of *H3f3a*. Notably, similar differences in *H3f3a* expression levels between distinct cell types, including T and B lymphocytes have also been demonstrated in an independent line of knock-in mice with epitope-tagged wild-type H3.3a and an IRES-GFP reporter (Bachu et al., 2019).

Ideally, after Cre-mediated activation of the knock-in allele, the expression levels of *H3f3a-K27M* should be similar to those of *H3f3a-WT*. However, extensive qRT-PCR experiments seemed to reveal a drastic reduction in expression of the targeted allele when compared to *H3f3a* expression in wild-type or vCre mice (Figure 28, page 82; please also refer to detailed methodological information provided in section 3.8, page 80 ff.). If correct, such low expression of the mutant *H3f3a* allele, as insinuated by the qRT-PCR data, would have serious consequences with regard to the usefulness of our knock-in mouse model. One could rightfully argue that apparently low H3.3A-K27M expression levels in our mice were unable to faithfully mimic critical pathophysiological effects of the respective mutation in leukemic cells of AML patients. However, we have compelling reason to doubt that the qRT-PCR data presented in chapter 3.8 allow a meaningful direct comparison of expression levels between wild-type and knock-in *H3f3a* alleles, as originally thought. Rather, the following lines of evidence strongly suggest the qRT-PCR data massively underestimate *H3f3a-K27M* transcript levels: **(i)** semi-quantitative RT-PCR data did not reveal any dramatic reduction of *H3f3a-K27M* transcript levels compared to transcript from the wild-type allele in heterozygous cells (Figure 19B, page 69), although the amplification product of the targeted allele is 66 bp (~20 %) longer than its wild-type counterpart, potentially causing less efficient amplification. **(ii)** Western blotting of protein extracts from diverse hematopoietic cells (thymocytes, splenocytes, bone marrow cells) of $vCre^{pos}$ x $H3K27M^{KI/WT}$ and $vCre^{pos}$ x $H3K27M^{KI/KI}$ mice revealed robust H3.3A-K27M protein expression (Figure 29A – C, page 83). Unfortunately, it is not possible to directly compare the expression levels of wild-type and K27M-mutated H3.3A proteins in heterozygous cells by Western blotting, as available H3.3 antibodies fail to detect their cognate epitopes in the K27M-utated variant (data not shown). **(iii)** Flow cytometry reveals robust, copy number-dependent expression of the IRES-driven vYFP reporter (Figure 26, page 79), indicating robust *H3f3a-K27M* transcript levels. Of note, intensities of vYFP-derived fluorescent signals in cells from our knock-in mice are grossly similar to signal intensities observed in an independently generated *H3f3a-IRES-GFP* knock-in mouse (Figure 3 in Bachu et al. (2019), also arguing against too low expression specifically in our mice. **(iv)** Comparison of vYFP-derived fluorescent intensities in T and B lymphocytes demonstrates that physiologic differences in *H3f3a* expression levels are perfectly preserved in our knock-in mice (Figure 26, page 79), supporting the view that regulation of *H3f3a* expression remains unaltered in knock-in alleles. **(v)** Most important, expression of the *H3f3a-K27M* mutant allele results in marked global undermethylation of amino acid K27 in all H3 family members, exactly as predicted by previous studies (Larson et al., 2019; Lewis and Allis, 2013; Lewis et al., 2013). Such a profound phenotypic effect seems unlikely, if the mutant allele were indeed expressed at levels as low as suggested by our qRT-PCR data.

Why should our qRT-PCR assays massively underestimate *H3f3a-K27M* mutant transcript levels? We are currently investigating this issue. Before we have clear-cut results, we can only speculate, but some preliminary data provide first hints. For instance, troubleshooting experiments revealed that the method used for priming cDNA can have a profound impact on the detected *H3f3a* levels. We observed that oligo dT priming resulted in massively reduced amplification signals specifically for knock-in versus wild-type transcripts, when oligo dT primers were compared to random primers or custom primers complementary to defined sites within the transcript (data not shown). Moreover, due to the insertion of an IRES-vYFP cassette transcript length increases by approximately 1.4 kb, which may markedly reduce the efficiency of reverse transcription. In addition, presence of the IRES-vYFP segment significantly alters the nucleotide sequence of the transcript and certainly has a profound influence on RNA secondary structure. Notably, in particular the IRES element is rich in long stretches of GC repeats, which are known to impair reverse transcription due to stable secondary structures and the requirement for high melting temperatures (McDowell et al., 1998). Ostensibly low transcript levels of the knock-in allele may thus be an artifact of technical imperfection. To finally resolve this issue, we may have to resort to methods of transcript quantification that are independent of reverse transcription and polymerase chain reaction, for instance using “old-fashioned” Northern blotting of oligo-dT-enriched RNA or labor-intensive, but highly sensitive RNA-protection assays.

4.2 Phenotypic effects of hematopoietic cell compartment-specific *H3f3a-K27M* expression in knock-in mice

H3K27 is one of the most prominent sites of post-translational histone modifications. Trimethylation at this site (H3K27me3) has been demonstrated to play an essential role in regulating gene expression (see Introduction 1.6.3). Moreover, the H3K27M mutation causes dominant-negative inhibition of EZH2, the enzymatic component of the polycomp repressor complex (PRC2) responsible of H3K27 methylation, resulting in a substantial decrease in global H3K27me3 levels, as described in the context of neurological tumors, where H3K27M has been identified as one of the key drivers in oncogenesis (Bender et al., 2013; Harutyunyan et al., 2019; Lewis and Allis, 2013; Lewis et al., 2013; Stafford et al., 2018). Larson et al. used their *Nestin*-Cre x H3K27M mouse model to conclusively demonstrate that the H3K27M mutation leads to decreased levels of H3K27me3 in neural stem cells, which, however, were not associated with equally global alterations in gene transcription. Rather, differences in transcriptomes of H3.3-K27 and WT neural stem cells were relatively modest and selective for specific genes. Importantly, *in vivo* introduction of the *H3f3a-K27M* mutation in neuronal stem cells did not result in overt tumorigenesis, demonstrating that this mutation alone is not sufficient for full cell transformation, fully in line with our findings in the hematopoietic system. However, *H3f3a-K27M* expression significantly enhanced neuronal stem cell self-renewal and markedly accelerated tumorigenesis in combination with two other recurrent glioma-typical mutations, loss of p53 and activation of the platelet-derived growth factor receptor alpha (PDGFRA) (Larson et al., 2019). These exciting findings fully support our working hypothesis that the *H3f3a-K27M* mutation needs additional mutations in other genes to contribute to leukemogenesis.

The results obtained from the previously discussed qRT-PCR sparked doubt in us, whether the observed expression levels of *H3f3a-K27M* after Cre-mediated activation are sufficient to mimic

pathophysiologic effects potentially associated with this mutation in our knock-in mice. As discussed above, for as yet unknown reason, the detected transcript levels for *H3f3a-K27M* are apparently far below those representing *H3f3a-WT* (Figure 28, page 82). To test, whether a well-documented effect of the H3.3-K27M mutation, namely global reduction of H3K27me3, as observed in different neurological tumors (Bender et al., 2013; Harutyunyan et al., 2019; Larson et al., 2019) can also be detected in our mice (where the mutation is activated in a different cell compartment i.e. hematopoietic cells), we performed intracellular FACS analysis of cells isolated from bone marrow, thymus, and spleen of $vCre^{pos}$ control, as well as hetero- and homozygous vCre x K27M mice (Figure 30, page 85). In line with previous reports (Larson et al., 2019; Lewis et al., 2013) we detected a highly significant, dose-dependent decrease in MFI values for H3K27me3 in all analyzed cell subsets, with ~25 % reduction in heterozygous and ~50 % reduction in homozygous mice, respectively (Figure 30A, top row and Figure 30B, left panels, page 85). Analysis of a second common methylation site (H3K4) revealed that the observed effects are confined to the tri-methylation of H3K27 (Figure 30A, second row and Figure 30B, middle panels, page 85). This is of special interest as H3.3 is frequently deposited at active promotors co-enriched for H3K4me3 and H3K27me3 (Delbarre et al., 2010). Analogous results were obtained upon immunohistochemical analysis of H3K27M mutated pediatric high-grade gliomas (Bender et al., 2013). Effects of the H3.3-K27M mutation on post-translational histone H3 modifications were also investigated in another, most recent study using high-resolution, single molecule imaging, which allowed the analysis of individual nucleosomes in cells from pediatric gliomas (Furth et al., 2022). Again, no change in global H3K24me3 levels were observed, like in our study, however, the authors could show that K27M mutated nucleosomes interact with MLL1, the histone methyl transferase specific for H3K4, resulting in an enrichment of H3K4e3 marks in individual nucleosomes harboring the H3K27M variant. Despite these technically advanced studies, how H3K27M mutations contribute to oncogenic transformation at the molecular level, has remained a mystery, likely requiring many additional investigations, including in appropriate mouse models like ours.

Of note, intracellular FACS analysis with an antibody specific for the FLAG epitope confirmed robust, copy number-dependent expression of 3xFLAG-tagged H3.3A-K27M protein in hematopoietic cells after Cre-mediated activation of the targeted *H3f3a* allele (Figure 30, right-most histograms; page 85), as documented before by Western blotting (Figure 29, page 83). The impressive shift in FLAG-tag-specific MFI values once again strongly argues against ostensibly low expression of the mutant allele, as insinuated from our qRT-PCR data. However, to precisely compare H3.3A-K27M with H3.3A-WT protein levels based on FLAG-tag quantification, we would need a separate knock-in mouse as control, with no other genetic modification than an in-frame insertion of 3xFLAG-encoding codons at the 3' end of the *H3f3a* coding region in exon 4. Why the striking increase of FLAG-specific signals in homozygous versus heterozygous knock-in mice consistently exceeds the expected 2-fold range, is an interesting, but currently unexplained observation (compare MFI values for $vCre^{pos}$ x $H3K27M^{KI/WT}$ and $vCre^{pos}$ x $H3K27M^{KI/KI}$ mice in Figure 30B, right-most histograms, page 85).

Remarkably, despite massive reduction of global H3K27 trimethylation, comprehensive FACS analysis of both younger (aged 7 – 9 weeks) and older mice (aged 68 – 95 weeks) failed to reveal any prominent phenotypic effect in all analyzed hematopoietic organs (Figure 31 - Figure 34, pages 86 – 89and Figure 54A - Figure 61A, pages 140 – 145). Likewise, macroscopic

inspection during necropsy of analyzed mice also did not reveal any unusual pathologies (data not shown). Mice of relevant genotypes were kept for up to 90 weeks and revealed no differences in survival as compared to vCre^{pos} and vCre^{pos} x Mll5$^{FL/WT\ or\ FL/FL}$ control mice; the same was true for vCre^{pos} x Mll5$^{FL/WT}$ x H3K27M$^{KI/WT}$ double mutants (Figure 35, page 90). In accordance with these findings, the analysis of blood parameters did not reveal any consistent, cohort-specific aberrations in cell counts (Figure 36, page 91). The fact that a prominent, genome-wide epigenetic alteration, like global reduction of H3K27 trimethylation as observed in our mice, is apparently well tolerated by the organism, not triggering immediately obvious pathophysiological effects, seems a rather unexpected and thus remarkable observation.

When analyzing our transgenic mice, the absence of observable phenotypic changes, apart from the significant reduction in H3K27me3 levels, can potentially be attributed to functional redundancy of certain genes or compensatory mechanisms that ensure proper function (Müller, 1999). In this study, the K27M mutation has only been introduced in *H3f3a*, one of two genes encoding the protein H3.3. Therefore, it is plausible that, despite the substantial decrease in H3K27me3, the H3.3-WT proteins encoded by the second gene *H3f3b* are sufficient to ensure proper physiological functionality. Additionally, in our mouse model, the mutation is activated upon Cre-mediated recombination from an early developmental stage, leaving sufficient time for the organism to adapt and to compensate for potential phenotypic effects. However, as the major goal of this Ph.D. thesis was the establishment of mouse models to evaluate leukemogenesis in the context of *Mll5*-deficiency, we did not further investigate potential compensatory mechanism that might prevent phenotypic changes in our vCre x H3K27M mice. Nevertheless, experimental efforts to unveil and then circumnavigate such potential mechanisms could be a worth-while line of future investigation.

4.3 Potential role of *H3f3a-K27M* in leukemogenesis and future lines of investigation

The primary purpose of creating the H3f3a-K27M mutant mouse was to assess the impact of this mutation on leukemogenesis in the context of *Mll5*-deficiency. To simultaneously evaluate the cooperativity of both mutations together with *Idh2-R172K*, another major part of this project was the establishment of retroviral transduction and bone marrow transplantation procedures within our laboratory. After optimizing the methods for virus production, lineage depletion, transduction efficiency and transplantation, we were able to successfully conduct competitive bone marrow transplantations of *Idh2-R172K* transduced bone marrow in more than 100 mice, with only 3 mice dying within the first two weeks due to failed engraftment of the transplanted cells. Details on transplanted genotypes and mouse numbers are given in Table 15 on page 93.

Our methods for viral transduction and bone marrow transplantation were proven to be effective, as FACS analysis demonstrated the engraftment of donor cells exceeding 50 % in over two-thirds of the transplanted mice (Figure 39 and Figure 40, left panels, pages 95 and 96). Batch to batch differences in engraftment between the individual transplantations can be attributed to minor differences in the preparation and injection of donor cells, as well as incomplete ablation of the host cells during irradiation (Kwarteng and Heinonen, 2016). However, when analyzing the fluorescent marker IRES-mCherry (measured in the PE-Texas Red channel) linked to *Idh2-R172K* expression, highly varying expression patterns among the

genotypes can be observed (Figure 39 and Figure 40, right panels, pages 95 and 96). Furthermore, the fluorescence intensity is inhomogeneous in specific cell populations of some individual mice, expressing IRES-mCherry at varying ratios (data not shown). This observed phenomenon might be attributed to the design of the vector rather than physiological differences in the expression of the introduced *Idh2-R172K*. During the ectopic expression of the viral plasmid, it remains possible that the second transgene, in our case the reporter, might be transcribed less efficiently (Ciuculescu et al., 2014), resulting in misleading results regarding the expression levels of *Idh2-R172K*.

Despite the generally good engraftment and expression of the *Idh2* mutation, neither the transplanted mice, nor the H3K27M-mutated and/or *Mll5*-deficient compound mice showed any changes in overall survival or a distinct phenotype hinting to leukemic transformation. For *Mll5*-deficient mice, this has been shown before (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009). However, this still doesn't exclude a potential role of *H3f3a-K27M* in leukemogenesis. It might be worth testing a slightly different approach. Leukemic transformation in AML usually occurs late in life. In our model, the *H3f3a-K27M* mutation is active from an early developmental stage. In both models applied during this study (i.e. transgenic compound mice and retroviral transduction/bone marrow transplantation) it is possible to intercross the mutant mice with CreERT2 instead of vCre mice. This would allow conditional induction of recombinase activity upon administration of tamoxifen in adult mice. As leukemogenic gene mutations are typically somatic and usually arise in single stem/progenitor cells of adult or aging individuals, the "CreERT2/tamoxifen-approach" would also resemble the human disease more closely. Furthermore, as we not only generated H3K27M knock-in mice, but corresponding retroviral expression plasmids as well, we could also modify the experimental set up by using double transduced (*Idh2-R172K* and *H3f3a-K27M*) BM cells isolated from readily available vCre-expressing *Mll5*-deficient donor mice.

In another line of investigation, it would also be worth to evaluate the influence of *H3f3a-K27M* in murine models of secondary AML. First studies demonstrated that the incidence for H3 mutations in these cases is 5 – 6 times higher than in primary AML (Boileau et al., 2019b). Hence, it might be worth-while to include treatment of relevant mouse mutants with alkylating agents in the experimental scheme or to perform serial transplantations to evaluate a potential role of the mutation in the context of secondary AML.

Generally, the complete failure of recurrent AML development in all our mice might be attributable to experimental procedures, rather than lack of cooperativity of the analyzed mutations. For one, there is the possibility that integration of the virus occurred with huge batch to batch variations – also explaining the observed differences in *Idh2-R172K* expression – leading to non-physiological expression levels that potentially affect the disease phenotype (Almosailleakh and Schwaller, 2019). Variation in expression levels could also explain why only few mice showed signs of pathology (typically slightly increased white blood cell counts or anemia). Furthermore, recombinant, viral-derived DNA might have integrated into precursors differentiating into inappropriate cell types. Development of the various leukemia subtypes strongly correlates with the driving oncogene and the cell of origin (Keinan et al., 2021). During FACS analysis, it could be seen that the *Idh2-R172K* expression levels varied among the different cell subsets, with generally higher levels in B cells (data not shown). Technically, within the HSPC compartment there could have been some sort of selection process favoring those

cells that eventually differentiate into B-lymphocytes i.e. not the compartment that is aberrant during transformation into AML. To overcome this issue in future experiments, it might be helpful to sort the transduced bone marrow cells prior to injection and thereby exclude non-transduced and lineage committed cells that escaped lineage depletion. Additionally, treatment of donor mice with 5-flourouracil (5-FU) induces cell cycling of bone marrow cells prior to harvesting the cells – a state promoting retroviral integration (Stubbs and Krivtsov, 2017).

For better and more complete evaluation of potential disease phenotypes, it might be useful to include histological analysis of blood smears and bone marrow cells. Murine hematopathology is highly variable and often context-dependent, hence histological analysis might help to identify more subtle pathophysiologic changes, and serve as additional quality control for the results obtained with the automated cell counter (O'Connell et al., 2015). Such histological analyses, might possibly detect minor changes in blood homeostasis, well before the manifestation of disease. An additional factor that needs to be considered is disease latency for AML. It might well be possible that the observation period of 12 months was not sufficient to detect significant and consistent pathophysiological effects of the introduced mutations. Furthermore, analysis of much larger cohorts of experimental animals should be considered to increase the statistical relevance of the results. In the course of our study, only very few mice got sick at all, leaving room for speculation whether this observation can be attributed to the introduced mutations, any experimental variation, or merely to chance.

Another possible reason for the absence of leukemia phenotypes might be credited to the differences between the murine and human bone marrow microenvironment, which plays a major part during disease development and is still not completely understood (Bello et al., 2018). Despite all efforts, murine leukemia models still cannot completely mimic the complexity of human tumors (Richmond and Su, 2008). Therefore, it is important to understand the limitations of murine models when examining human diseases. For instance, one could consider the use of humanized mouse models to focus the analysis on human hematopoietic HSPCs, like in a recent study using human $CD34^{pos}CD38^{neg}$ hematopoietic stem cells isolated from cord blood and transduced with H3K27M-encoding virus. Subsequent transplantation into NSG mice revealed expansion of myeloid precursors via shifting the balance between self-renewal and quiescence programs (Boileau et al., 2019a; Dakik et al., 2022) – a phenotype not observed in any of our mice.

Transplantation of $H3K27M^{KI/WT}$ bone marrow cells transduced with IDH2-R172K leads to a significant increase in white and a decrease in red blood cells (Figure 44, page 100), which can potentially be attributed to cooperating effects of both mutations, as a similar effect is not observed in mice carrying only the H3K27M mutation (Figure 36, page 91). In mice that additionally exhibit heterozygous loss of *Mll5*, these effects on cell counts appear to be partially reversed. In line with this observation, an antagonistic role of H3.3 and MLL5 in glioblastoma primary cultures has been described before (Gallo et al., 2015). This publication reports that overexpression of MLL5 leads to lower levels of H3.3 due to repression of *H3F3B* (but not *H3F3A*) gene activity. In contrast this could mean that the loss of MLL5 leads to increased levels of H3.3 protein originating from the *H3F3B* gene locus. Imagining a similar (but as yet unproven) scenario in mice, overabundant wild-type H3.3 from the “de-repressed” *H3f3b* locus might then dilute K27M-mutant H3.3 of the *H3f3a* knock-in allele, potentially responsible for the

lack of phenotype observed in our mice as the protein would potentially compensate for any defects mediated by H3.3-K27M.

Interestingly, when looking at the expression of *IDH2-R172K* in transplanted mice, expression of the introduced mutation appears to be linked to *Mll5*-deficiency as well. In the majority of mice that have been transplanted with Mll5$^{FL/WT}$ bone marrow, the expression levels are below 10 % (Figure 40, first and last row, page 96), an effect that cannot be observed in bone marrow cells with a complete loss of Mll5 (Figure 40, second row, page 96). However, these cohorts, especially the one for homozygous *Mll5*-deficient mice, are relatively small. To make a statistical relevant statement about the influence of *Mll5* on *Idh2-R172K* expression, again larger cohorts would be required. Furthermore, despite mimicking the clinical state within the patients less closely, mice transplanted with bone marrow homozygous for *H3f3a-K27M* and loss of *Mll5* should be studied as well. Due to the time consuming breeding and long observation period, these mice could not be included in this study. Nevertheless, the total loss of *Mll5* does not necessarily guarantee a leukemia favoring genotype. For two other members of the MLL-family (*MLL1* and *MLL4*) it has been shown that one wild-type allele is required to maintain the leukemic phenotype in the context of MLL-AF9 fusion gene induced leukemia (Santos et al., 2014; Thiel et al., 2010).

Some initial evidence for an antagonistic role of MLL5 with respect to leukemogenesis was gathered during pilot experiments conducted in a different, still unpublished line of investigation in our laboratory. Here, wild-type and *Mll5*-deficient bone marrow cells were immortalized with MLL-AF9 retroviruses and transplanted into irradiated recipient mice. Observed survival curves suggested that the loss of *Mll5* delayed MLL-AF9 mediated leukemic death. However, additional experiments of this kind need to be performed to substantiate this initial observation.

After all, knowledge of the role of *H3f3a-K27M* in leukemogenesis remains limited. Despite all our efforts, we failed to establish a direct link between this mutation and a potential role in leukemogenesis. However, by establishing a sophisticated, novel knock-in mouse model and all relevant methodologies within our laboratory, we have set the stage for an in-depth evaluation of the *H3f3a-K27M* mutation, its role in leukemogenesis and potentially cooperating genetic lesions beyond those addressed so far.

Apart from the aforementioned methodological adaptions to further improve the analysis of H3K27M mice in the context of leukemogenesis, several other lines of investigation are worth following.

As alluded to in introductory section 1.6.2, initially this study was supposed to include an evaluation of hetero- and homozygous loss of *Dnmt3a* as well as the DNMT3A-R882H mutation in the context of *Mll5* deficiency. To this end, we introduced *Dnmt3a*$^{flox/flox}$ mice into our mouse colony, initiated appropriate breeding schemes and also generated retroviral constructs, carrying a fluorescent marker together with the R882H mutation However, for reasons that are not entirely clear, intercross of the *Dnmt3a* mutant mouse strain with vCre, Mll5-flox and H3K27M mice turned out to be more challenging than anticipated. Small litter sizes, recurrent killing of newborn pups by parents, under-representation of desired genotypes among offspring and unexplained deaths among mice with the desired genotypes shortly after weaning describe just some of the difficulties that were regularly encountered, effectively impeding statistically meaningful phenotypic analyses. Despite these problems, some representative survival curves

are given in the appendix on page 146. Due to the already enormous amount of time and effort required for the comprehensive study of the other leukemia-associated gene mutations, we decided to discontinue the investigation of the influence of *Dnmt3a* mutations on leukemia in the context of *Mll5* deficiency for the time being. Nevertheless, we have laid the foundation for such a study by having generated at least some double- and triple mutant mice that will serve as breeders to expand corresponding mouse colonies, and by having completed the construction of the required retroviral plasmids. Thus, appropriate breeding schemes can now be readily set up to generate hopefully in relatively short time sufficient numbers of complex mouse mutants as required for the intended experiments. Furthermore, adoptive transfer of Dntm3a-R882H transduced bone marrow cells can be conducted at any time point, as the required retroviral supernatants have also been produced.

In yet another line of investigation, it might be interesting to include inactivating mutations of the transcription factor RUNX1 into the analyses. The whole exome sequencing of eight patients with 7q deletions revealed RUNX1 loss-of-function mutations in 3 cases (Table 1 and Table 2), once even in combination with mutations in all the other selected genes (*DNMT3A*, *H3F3A*, and *IDH2*, Table 1). Analysis of a large cohort of samples from AML patients could detect a significant enrichment of K27 mutations in several H3 isoforms in the context of alterations in RUNX1, acting as strong disease accelerators in an AML mouse model with RUNX1 fusions (Lehnertz et al., 2017). Mouse models with *loxP* flanked exon of *Runx1* are available and can be directly intercrossed with our vCre x Mll5 and vCre x H3K27M mice to inactivate Runx1 in all hematopoietic cells (Taniuchi et al., 2002) and subsequently analyze to cooperativity of these mutations.

5. Summary

Acute myeloid leukemia (AML) is a genetically heterogeneous, life-threatening disease with low cure rate. Understanding the effect of disease-promoting genetic aberrations is a prerequisite for the development of new treatment approaches. Leukemic cells from patients with severe forms of AML frequently exhibit specific deletions within the long arm of chromosome 7, which are of unknown pathophysiological significance. Studies to unveil a potential tumor suppressor located in this region lead to the discovery of *MLL5* – an atypical member of the Mixed-Lineage-Leukemia (MLL) family of epigenetic regulators, which are known to be involved in oncogenesis, including leukemia development. Indeed, previous analysis of *Mll5* deficient mice revealed phenotypic features indicative of a pre-leukemic state. However, neither hetero- nor homozygous inactivation of *Mll5* in mice results in overt leukemia or any other tumors. So far, clear scientific proof for an involvement of MLL5 in leukemogenesis is thus still missing.

In line with the 2-hit model of tumorigenesis, we hypothesized that additional mutations in one or more other genes might be required to induce leukemia in the context of *MLL5* deficiency. Upon whole exome sequencing (WES) of leukemic cells from AML patients with heterozygous *MLL5* deletion, several potentially cooperating gene mutations were identified. Based on technical and scientific considerations, mutations in three genes were selected for *in vivo* evaluation: H3F3A-K27M, IDH2-R172K, DNMT3A-R882H and inactivation of DNMT3A.

The main goal of this project was the generation and establishment of suitable mouse models, that would allow simultaneous analysis of the aforementioned mutations in the context of *Mll5*-deficiency with regard to leukemogenesis. To this end, classical gene targeting of murine embryonic stem cells was utilized to generate a novel, Cre-inducible knock-in mouse strain carrying a point mutation that would convert lysine at position 27 into methionine in *H3f3a*, i.e. one of the two genes encoding the canonical histone variant H3.3. To allow non-invasive monitoring of *H3f3a-K27M* expression, the fluorescent marker IRES-vYFP was introduced along with the mutation; in addition, the protein tag 3xFLAG was attached at the C-terminus of the mutated H3.3 variant. Intercrossing of these newly generated mice with vCre-expressing mice resulted in reliable activation of the introduced mutation exclusively in the hematopoietic cell compartment, as conclusively demonstrated by a number of independent methods. Importantly, intracellular FACS analysis revealed a marked reduction in global trimethylation of histone H3 at position K27 (H3K27me3), a known effect of the K27M mutation well-documented in neurological tumors. However, despite these aberrant, pro-oncogenic methylation patterns, no pathologic phenotype, including leukemia was evident in the context of H3K27M alone or in combination with *Mll5*-deficiency. However, highly reproducible results proved the usability of our mouse model to analyze the expression of *H3f3a-K27M* in various clinical contexts.

As additional approach, competitive bone marrow transplantation assays were successfully established and optimized in our laboratory to assess some of the above-mentioned gene mutations, in particular IDH2-R172K, in the context of *Mll5*-deficiency and/or the H3.3-K27M mutation. Again, no evidence for a potential cooperation of any of these mutations regarding leukemogenesis was obtained as yet. However, for a complete evaluation, several additional, long-term studies are required, for which our conscientiously established mouse models will serve as perfect tools.

6. References

Akbay, E.A., J. Moslehi, C.L. Christensen, S. Saha, J.H. Tchaicha, S.H. Ramkissoon, K.M. Stewart, J. Carretero, E. Kikuchi, H. Zhang, T.J. Cohoon, S. Murray, W. Liu, K. Uno, S. Fisch, K. Jones, S. Gurumurthy, C. Gliser, S. Choe, M. Keenan, J. Son, I. Stanley, J.A. Losman, R. Padera, R.T. Bronson, J.M. Asara, O. Abdel-Wahab, P.C. Amrein, A.T. Fathi, N.N. Danial, A.C. Kimmelman, A.L. Kung, K.L. Ligon, K.E. Yen, W.G. Kaelin, Jr., N. Bardeesy, and K.K. Wong. 2014. D-2-hydroxyglutarate produced by mutant IDH2 causes cardiomyopathy and neurodegeneration in mice. *Genes Dev* 28:479-490.

Ali, M., H. Rincón-Arano, W. Zhao, S.B. Rothbart, Q. Tong, S.M. Parkhurst, B.D. Strahl, L.-W. Deng, M. Groudine, and T.G. Kutateladze. 2013. Molecular basis for chromatin binding and regulation of MLL5. *Proceedings of the National Academy of Sciences* 110:11296-11301.

Almosailleakh, M., and J. Schwaller. 2019. Murine Models of Acute Myeloid Leukaemia. *Int J Mol Sci* 20:

Aparicio, T., R. Baer, and J. Gautier. 2014. DNA double-strand break repair pathway choice and cancer. *DNA Repair (Amst)* 19:169-175.

Bachu, M., T. Tamura, C. Chen, A. Narain, V. Nehru, N. Sarai, S.B. Ghosh, A. Ghosh, R. Kavarthapu, M.L. Dufau, and K. Ozato. 2019. A versatile mouse model of epitope-tagged histone H3.3 to study epigenome dynamics. *J Biol Chem* 294:1904-1914.

Banerjee, P., L. Crawford, E. Samuelson, and G. Feuer. 2010. Hematopoietic stem cells and retroviral infection. *Retrovirology* 7:8.

Barlow, C., S. Hirotsune, R. Paylor, M. Liyanage, M. Eckhaus, F. Collins, Y. Shiloh, J.N. Crawley, T. Ried, D. Tagle, and A. Wynshaw-Boris. 1996. Atm-deficient mice: a paradigm of ataxia telangiectasia. *Cell* 86:159-171.

Basheer, F., and B.J. Huntly. 2015. BET bromodomain inhibitors in leukemia. *Exp Hematol* 43:718-731.

Bassing, C.H., H. Suh, D.O. Ferguson, K.F. Chua, J. Manis, M. Eckersdorff, M. Gleason, R. Bronson, C. Lee, and F.W. Alt. 2003. Histone H2AX: A Dosage-Dependent Suppressor of Oncogenic Translocations and Tumors. *Cell* 114:359-370.

Becher, O.J., and E.C. Holland. 2015. Glioma Stem-like Cells Keep Their H3.3 Variant Levels at Bay. *Cancer Cell* 28:679-680.

Bello, A.B., H. Park, and S.-H. Lee. 2018. Current approaches in biomaterial-based hematopoietic stem cell niches. *Acta Biomaterialia* 72:1-15.

Bender, S., Y. Tang, A.M. Lindroth, V. Hovestadt, D.T. Jones, M. Kool, M. Zapatka, P.A. Northcott, D. Sturm, W. Wang, B. Radlwimmer, J.W. Højfeldt, N. Truffaux, D. Castel, S. Schubert, M. Ryzhova, H. Seker-Cin, J. Gronych, P.D. Johann, S. Stark, J. Meyer, T. Milde, M. Schuhmann, M. Ebinger, C.M. Monoranu, A. Ponnuswami, S. Chen, C. Jones, O. Witt, V.P. Collins, A. von Deimling, N. Jabado, S. Puget, J. Grill, K. Helin, A. Korshunov, P. Lichter, M. Monje, C. Plass, Y.J. Cho, and S.M. Pfister. 2013. Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. *Cancer Cell* 24:660-672.

Bhatia, S. 2013. Therapy-related myelodysplasia and acute myeloid leukemia. *Semin Oncol* 40:666-675.

Birnboim, H.C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol* 100:243-255.

Boileau, M., S. Jessa, S. Worme, D. Faury, N. Jabado, F. Mercier, C. Kleinman, and K. Eppert. 2019a. Single Cell Transcriptomic Analysis of the Histone H3 K27M Mutation in Pre-Leukemic Hematopoietic Stem Cells. *Blood* 134:3773-3773.

Boileau, M., M. Shirinian, T. Gayden, A.S. Harutyunyan, C.C.L. Chen, L.G. Mikael, H.M. Duncan, A.L. Neumann, P. Arreba-Tutusaus, N. De Jay, M. Zeinieh, K. Rossokhata,

Y. Zhang, H. Nikbakht, C. Mouawad, R. Massoud, F. Frey, R. Nasr, J. El Cheikh, M. El Sabban, C.L. Kleinman, R. Mahfouz, M.D. Minden, N. Jabado, A. Bazarbachi, and K. Eppert. 2019b. Mutant H3 histones drive human pre-leukemic hematopoietic stem cell expansion and promote leukemic aggressiveness. *Nat Commun* 10:2891.

Britten, O., D. Ragusa, S. Tosi, and Y.M. Kamel. 2019. MLL-Rearranged Acute Leukemia with t(4;11)(q21;q23)-Current Treatment Options. Is There a Role for CAR-T Cell Therapy? *Cells* 8:

Brunetti, L., M.C. Gundry, and M.A. Goodell. 2017. DNMT3A in Leukemia. *Cold Spring Harb Perspect Med* 7:

Celeste, A., S. Petersen, P.J. Romanienko, O. Fernandez-Capetillo, H.T. Chen, O.A. Sedelnikova, B. Reina-San-Martin, V. Coppola, E. Meffre, M.J. Difilippantonio, C. Redon, D.R. Pilch, A. Olaru, M. Eckhaus, R.D. Camerini-Otero, L. Tessarollo, F. Livak, K. Manova, W.M. Bonner, M.C. Nussenzweig, and A. Nussenzweig. 2002. Genomic instability in mice lacking histone H2AX. *Science* 296:922-927.

Celik, H., C. Mallaney, A. Kothari, E.L. Ostrander, E. Eultgen, A. Martens, C.A. Miller, J. Hundal, J.M. Klco, and G.A. Challen. 2015. Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation. *Blood* 125:619-628.

Challen, G.A., D. Sun, M. Jeong, M. Luo, J. Jelinek, J.S. Berg, C. Bock, A. Vasanthakumar, H. Gu, Y. Xi, S. Liang, Y. Lu, G.J. Darlington, A. Meissner, J.P. Issa, L.A. Godley, W. Li, and M.A. Goodell. 2011. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet* 44:23-31.

Chan, K.M., D. Fang, H. Gan, R. Hashizume, C. Yu, M. Schroeder, N. Gupta, S. Mueller, C.D. James, R. Jenkins, J. Sarkaria, and Z. Zhang. 2013. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev* 27:985-990.

Chaudry, S.F., and T.J.T. Chevassut. 2017. Epigenetic Guardian: A Review of the DNA Methyltransferase DNMT3A in Acute Myeloid Leukaemia and Clonal Haematopoiesis. *BioMed Research International* 2017:5473197.

Chen, C., Y. Liu, C. Lu, J.R. Cross, J.P.t. Morris, A.S. Shroff, P.S. Ward, J.E. Bradner, C. Thompson, and S.W. Lowe. 2013. Cancer-associated IDH2 mutants drive an acute myeloid leukemia that is susceptible to Brd4 inhibition. *Genes Dev* 27:1974-1985.

Chen, C., Y. Liu, A.R. Rappaport, T. Kitzing, N. Schultz, Z. Zhao, A.S. Shroff, R.A. Dickins, C.R. Vakoc, J.E. Bradner, W. Stock, M.M. LeBeau, K.M. Shannon, S. Kogan, J. Zuber, and S.W. Lowe. 2014. MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. *Cancer Cell* 25:652-665.

Ciuculescu, M.F., C. Brendel, C.E. Harris, and D.A. Williams. 2014. Retroviral transduction of murine and human hematopoietic progenitors and stem cells. *Methods Mol Biol* 1185:287-309.

Cole, C.B., D.A. Russler-Germain, S. Ketkar, A.M. Verdoni, A.M. Smith, C.V. Bangert, N.M. Helton, M. Guo, J.M. Klco, S. O'Laughlin, C. Fronick, R. Fulton, G.S. Chang, A.A. Petti, C.A. Miller, and T.J. Ley. 2017. Haploinsufficiency for DNA methyltransferase 3A predisposes hematopoietic cells to myeloid malignancies. *J Clin Invest* 127:3657-3674.

Crump, N.T., and T.A. Milne. 2019. Why are so many MLL lysine methyltransferases required for normal mammalian development? *Cell Mol Life Sci* 76:2885-2898.

Curtiss, N.P., J.M. Bonifas, J.O. Lauchle, J.D. Balkman, C.P. Kratz, B.M. Emerling, E.D. Green, M.M. Le Beau, and K.M. Shannon. 2005. Isolation and analysis of candidate myeloid tumor suppressor genes from a commonly deleted segment of 7q22. *Genomics* 85:600-607.

Dakik, H., M. Boileau, S. Worme, S. Jessa, I. Iasenza, A. Neumann, Y.L. Sun, M. Saby, M. Craig, C.L. Kleinman, and K. Eppert. 2022. The Role of Histone H3-K27M Mutation in Preleukemic HSC Expansion. *Blood* 140:5839-5839.

de Boer, J., A. Williams, G. Skavdis, N. Harker, M. Coles, M. Tolaini, T. Norton, K. Williams, K. Roderick, A.J. Potocnik, and D. Kioussis. 2003. Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *Eur J Immunol* 33:314-325.

Delbarre, E., B.M. Jacobsen, A.H. Reiner, A.L. Sørensen, T. Küntziger, and P. Collas. 2010. Chromatin environment of histone variant H3.3 revealed by quantitative imaging and genome-scale chromatin and DNA immunoprecipitation. *Mol Biol Cell* 21:1872-1884.

Deng, C., and M.R. Capecchi. 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol Cell Biol* 12:3365-3371.

Deshmukh, S., A. Ptack, B. Krug, and N. Jabado. 2022. Oncohistones: a roadmap to stalled development. *Febs j* 289:1315-1328.

Dillon, S.C., X. Zhang, R.C. Trievel, and X. Cheng. 2005. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* 6:227.

DiNardo, C.D., and J.E. Cortes. 2016. Mutations in AML: prognostic and therapeutic implications. *Hematology Am Soc Hematol Educ Program* 2016:348-355.

Döhner, H., A.H. Wei, F.R. Appelbaum, C. Craddock, C.D. DiNardo, H. Dombret, B.L. Ebert, P. Fenaux, L.A. Godley, R.P. Hasserjian, R.A. Larson, R.L. Levine, Y. Miyazaki, D. Niederwieser, G. Ossenkoppele, C. Röllig, J. Sierra, E.M. Stein, M.S. Tallman, H.F. Tien, J. Wang, A. Wierzbowska, and B. Löwenberg. 2022. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood* 140:1345-1377.

Döhner, H., D.J. Weisdorf, and C.D. Bloomfield. 2015. Acute Myeloid Leukemia. *N Engl J Med* 373:1136-1152.

Dohner, K., M. Habdank, F.G. Rucker, S. Miller, S. Frohling, S.W. Scherer, L. Bullinger, and H. Dohner. 2006. Molecular Characterization of Distinct Hot Spot Regions on Chromosome 7q in Myeloid Leukemias. *Blood* 108:2349-2349.

Döhner, K., R. Obermiller, D. Lipka, K. Hofmann, G. Fazekas, S. Fröling, P. Lichter, S.W. Scherer, and H. Döhner. 2002. Identification and characterization of a novel gene located in the commonly deleted region 7q22.q31.1 in myeloid leukemias. *Blood* 100:

Dong, Y., O. Shi, Q. Zeng, X. Lu, W. Wang, Y. Li, and Q. Wang. 2020. Leukemia incidence trends at the global, regional, and national level between 1990 and 2017. *Exp Hematol Oncol* 9:14.

Dozzo, A., A. Galvin, J.-W. Shin, S. Scalia, C.M. O'Driscoll, and K.B. Ryan. 2022. Modelling acute myeloid leukemia (AML): What's new? A transition from the classical to the modern. *Drug Delivery and Translational Research*

Duran-Struuck, R., and R.C. Dysko. 2009. Principles of bone marrow transplantation (BMT): providing optimal veterinary and husbandry care to irradiated mice in BMT studies. *J Am Assoc Lab Anim Sci* 48:11-22.

Emerling, B.M., J. Bonifas, C.P. Kratz, S. Donovan, B.R. Taylor, E.D. Green, M.M. Le Beau, and K.M. Shannon. 2002. MLL5, a homolog of Drosophila trithorax located within a segment of chromosome band 7q22 implicated in myeloid leukemia. *Oncogene* 21:4849-4854.

Estey, E., R.P. Hasserjian, and H. Döhner. 2022. Distinguishing AML from MDS: a fixed blast percentage may no longer be optimal. *Blood* 139:323-332.

Figueroa, M.E., O. Abdel-Wahab, C. Lu, P.S. Ward, J. Patel, A. Shih, Y. Li, N. Bhagwat, A. Vasanthakumar, H.F. Fernandez, M.S. Tallman, Z. Sun, K. Wolniak, J.K. Peeters, W. Liu, S.E. Choe, V.R. Fantin, E. Paietta, B. Löwenberg, J.D. Licht, L.A. Godley, R. Delwel, P.J. Valk, C.B. Thompson, R.L. Levine, and A. Melnick. 2010. Leukemic IDH1

and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* 18:553-567.

Fischer, K., S. Fröhling, S.W. Scherer, J. McAllister Brown, C. Scholl, S. Stilgenbauer, L.C. Tsui, P. Lichter, and H. Döhner. 1997. Molecular cytogenetic delineation of deletions and translocations involving chromosome band 7q22 in myeloid leukemias. *Blood* 89:2036-2041.

Friedrich, G., and P. Soriano. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* 5:1513-1523.

Fröhling, S., K. Nakabayashi, S.W. Scherer, H. Döhner, and K. Döhner. 2001. Mutation analysis of the origin recognition complex subunit 5 (ORC5L) gene in adult patients with myeloid leukemias exhibiting deletions of chromosome band 7q22. *Hum Genet* 108:304-309.

Fukuda, T., E. Asou, K. Nogi, and K. Goto. 2017. Evaluation of mouse red blood cell and platelet counting with an automated hematology analyzer. *J Vet Med Sci* 79:1707-1711.

Funato, K., and V. Tabar. 2018. Histone Mutations in Cancer. *Annual Review of Cancer Biology* 2:337-351.

Furth, N., D. Algranati, B. Dassa, O. Beresh, V. Fedyuk, N. Morris, L.H. Kasper, D. Jones, M. Monje, S.J. Baker, and E. Shema. 2022. H3-K27M-mutant nucleosomes interact with MLL1 to shape the glioma epigenetic landscape. *Cell Reports* 39:

Gaidzik, V.I., R.F. Schlenk, P. Paschka, A. Stölzle, D. Späth, A. Kuendgen, M. von Lilienfeld-Toal, W. Brugger, H.G. Derigs, S. Kremers, R. Greil, A. Raghavachar, M. Ringhoffer, H.R. Salih, M. Wattad, H.G. Kirchen, V. Runde, G. Heil, A.L. Petzer, M. Girschikofsky, M. Heuser, S. Kayser, G. Goehring, M.V. Teleanu, B. Schlegelberger, A. Ganser, J. Krauter, L. Bullinger, H. Döhner, and K. Döhner. 2013. Clinical impact of DNMT3A mutations in younger adult patients with acute myeloid leukemia: results of the AML Study Group (AMLSG). *Blood* 121:4769-4777.

Gallo, M., F.J. Coutinho, R.J. Vanner, T. Gayden, S.C. Mack, A. Murison, M. Remke, R. Li, N. Takayama, K. Desai, L. Lee, X. Lan, N.I. Park, D. Barsyte-Lovejoy, D. Smil, D. Sturm, M.M. Kushida, R. Head, M.D. Cusimano, M. Bernstein, I.D. Clarke, J.E. Dick, S.M. Pfister, J.N. Rich, C.H. Arrowsmith, M.D. Taylor, N. Jabado, D.P. Bazett-Jones, M. Lupien, and P.B. Dirks. 2015. MLL5 Orchestrates a Cancer Self-Renewal State by Repressing the Histone Variant H3.3 and Globally Reorganizing Chromatin. *Cancer Cell* 28:715-729.

Gavrilescu, L.C., and R.A. Van Etten. 2008. Murine retroviral bone marrow transplantation models for the study of human myeloproliferative disorders. *Curr Protoc Pharmacol* Chapter 14:Unit14.10.

Georgiades, P., S. Ogilvy, H. Duval, D.R. Licence, D.S. Charnock-Jones, S.K. Smith, and C.G. Print. 2002. VavCre transgenic mice: a tool for mutagenesis in hematopoietic and endothelial lineages. *Genesis* 34:251-256.

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

Goodwin, L.O., E. Splinter, T.L. Davis, R. Urban, H. He, R.E. Braun, E.J. Chesler, V. Kumar, M. van Min, J. Ndukum, V.M. Philip, L.G. Reinholdt, K. Svenson, J.K. White, M. Sasner, C. Lutz, and S.A. Murray. 2019. Large-scale discovery of mouse transgenic integration sites reveals frequent structural variation and insertional mutagenesis. *Genome Res* 29:494-505.

Green, C.L., C.M. Evans, L. Zhao, R.K. Hills, A.K. Burnett, D.C. Linch, and R.E. Gale. 2011. The prognostic significance of IDH2 mutations in AML depends on the location of the mutation. *Blood* 118:409-412.

Grimwade, D., R.K. Hills, A.V. Moorman, H. Walker, S. Chatters, A.H. Goldstone, K. Wheatley, C.J. Harrison, A.K. Burnett, and o.b.o.t.N.C.R.I.A.L.W. Group. 2010.

Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* 116:354-365.

Guo, J., R. Zhang, Z. Yang, Z. Duan, D. Yin, and Y. Zhou. 2021. Biological Roles and Therapeutic Applications of IDH2 Mutations in Human Cancer. *Front Oncol* 11:644857.

Guryanova, O.A., K. Shank, B. Spitzer, L. Luciani, R.P. Koche, F.E. Garrett-Bakelman, C. Ganzel, B.H. Durham, A. Mohanty, G. Hoermann, S.A. Rivera, A.G. Chramiec, E. Pronier, L. Bastian, M.D. Keller, D. Tovbin, E. Loizou, A.R. Weinstein, A.R. Gonzalez, Y.K. Lieu, J.M. Rowe, F. Pastore, A.S. McKenney, A.V. Krivtsov, W.R. Sperr, J.R. Cross, C.E. Mason, M.S. Tallman, M.E. Arcila, O. Abdel-Wahab, S.A. Armstrong, S. Kubicek, P.B. Staber, M. Gönen, E.M. Paietta, A.M. Melnick, S.D. Nimer, S. Mukherjee, and R.L. Levine. 2016. DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling. *Nat Med* 22:1488-1495.

Halene, S., L. Wang, R.M. Cooper, D.C. Bockstoce, P.B. Robbins, and D.B. Kohn. 1999. Improved expression in hematopoietic and lymphoid cells in mice after transplantation of bone marrow transduced with a modified retroviral vector. *Blood* 94:3349-3357.

Hanahan, D., and R.A. Weinberg. 2000. The Hallmarks of Cancer. *Cell* 100:57-70.

Harutyunyan, A.S., B. Krug, H. Chen, S. Papillon-Cavanagh, M. Zeinieh, N. De Jay, S. Deshmukh, C.C.L. Chen, J. Belle, L.G. Mikael, D.M. Marchione, R. Li, H. Nikbakht, B. Hu, G. Cagnone, W.A. Cheung, A. Mohammadnia, D. Bechet, D. Faury, M.K. McConechy, M. Pathania, S.U. Jain, B. Ellezam, A.G. Weil, A. Montpetit, P. Salomoni, T. Pastinen, C. Lu, P.W. Lewis, B.A. Garcia, C.L. Kleinman, N. Jabado, and J. Majewski. 2019. H3K27M induces defective chromatin spread of PRC2-mediated repressive H3K27me2/me3 and is essential for glioma tumorigenesis. *Nature Communications* 10:1262.

Hawley, R.G., F.H. Lieu, A.Z. Fong, and T.S. Hawley. 1994. Versatile retroviral vectors for potential use in gene therapy. *Gene Ther* 1:136-138.

Heuser, M., D.B. Yap, M. Leung, T.R. de Algara, A. Tafech, S. McKinney, J. Dixon, R. Thresher, B. Colledge, M. Carlton, R.K. Humphries, and S.A. Aparicio. 2009. Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. *Blood* 113:1432-1443.

Holz-Schietinger, C., D.M. Matje, M.F. Harrison, and N.O. Reich. 2011. Oligomerization of DNMT3A controls the mechanism of de novo DNA methylation. *J Biol Chem* 286:41479-41488.

Honda, H., A. Nagamachi, and T. Inaba. 2015. -7/7q- syndrome in myeloid-lineage hematopoietic malignancies: attempts to understand this complex disease entity. *Oncogene* 34:2413-2425.

Inaba, T., H. Honda, and H. Matsui. 2018. The enigma of monosomy 7. *Blood* 131:2891-2898.

Iqbal, J., A. Nussenzweig, J. Lubinski, T. Byrski, A. Eisen, L. Bordeleau, N.M. Tung, S. Manoukian, C.M. Phelan, P. Sun, and S.A. Narod. 2016. The incidence of leukaemia in women with BRCA1 and BRCA2 mutations: an International Prospective Cohort Study. *Br J Cancer* 114:1160-1164.

Issa, G.C., and C.D. DiNardo. 2021. Acute myeloid leukemia with IDH1 and IDH2 mutations: 2021 treatment algorithm. *Blood Cancer J* 11:107.

Ito, K., A. Hirao, F. Arai, S. Matsuoka, K. Takubo, I. Hamaguchi, K. Nomiyama, K. Hosokawa, K. Sakurada, N. Nakagata, Y. Ikeda, T.W. Mak, and T. Suda. 2004. Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells. *Nature* 431:997-1002.

Ito, K., A. Hirao, F. Arai, K. Takubo, S. Matsuoka, K. Miyamoto, M. Ohmura, K. Naka, K. Hosokawa, Y. Ikeda, and T. Suda. 2006. Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nat Med* 12:446-451.

Jain, K., C.S. Fraser, M.R. Marunde, M.M. Parker, C. Sagum, J.M. Burg, N. Hall, I.K. Popova, K.L. Rodriguez, A. Vaidya, K. Krajewski, M.-C. Keogh, M.T. Bedford, and B.D. Strahl. 2020. Characterization of the plant homeodomain (PHD) reader family for their histone tail interactions. *Epigenetics & Chromatin* 13:3.

Jaiswal, S., and B.L. Ebert. 2019. Clonal hematopoiesis in human aging and disease. *Science* 366:

Jang, C.W., Y. Shibata, J. Starmer, D. Yee, and T. Magnuson. 2015. Histone H3.3 maintains genome integrity during mammalian development. *Genes Dev* 29:1377-1392.

Kallappagoudar, S., R.K. Yadav, B.R. Lowe, and J.F. Partridge. 2015. Histone H3 mutations--a special role for H3.3 in tumorigenesis? *Chromosoma* 124:177-189.

Kantarjian, H., T. Kadia, C. DiNardo, N. Daver, G. Borthakur, E. Jabbour, G. Garcia-Manero, M. Konopleva, and F. Ravandi. 2021. Acute myeloid leukemia: current progress and future directions. *Blood Cancer J* 11:41.

Kantidakis, T., M. Saponaro, R. Mitter, S. Horswell, A. Kranz, S. Boeing, O. Aygün, G.P. Kelly, N. Matthews, A. Stewart, A.F. Stewart, and J.Q. Svejstrup. 2016. Mutation of cancer driver MLL2 results in transcription stress and genome instability. *Genes Dev* 30:408-420.

Kats, L.M., M. Reschke, R. Taulli, O. Pozdnyakova, K. Burgess, P. Bhargava, K. Straley, R. Karnik, A. Meissner, D. Small, S.M. Su, K. Yen, J. Zhang, and P.P. Pandolfi. 2014. Proto-oncogenic role of mutant IDH2 in leukemia initiation and maintenance. *Cell Stem Cell* 14:329-341.

Kayser, S., K. Döhner, J. Krauter, C.-H. Köhne, H.A. Horst, G. Held, M. von Lilienfeld-Toal, S. Wilhelm, A. Kündgen, K. Götze, M. Rummel, D. Nachbaur, B. Schlegelberger, G. Göhring, D. Späth, C. Morlok, M. Zucknick, A. Ganser, H. Döhner, R.F. Schlenk, and f.t.G.-A. AMLSG. 2011. The impact of therapy-related acute myeloid leukemia (AML) on outcome in 2853 adult patients with newly diagnosed AML. *Blood* 117:2137-2145.

Keinan, N., Y.e. Scharff, O. Goldstein, M. Chamo, S. Ilic, and R. Gazit. 2021. Syngeneic leukemia models using lentiviral transgenics. *Cell Death & Disease* 12:193.

Khoury, J.D., E. Solary, O. Abla, Y. Akkari, R. Alaggio, J.F. Apperley, R. Bejar, E. Berti, L. Busque, J.K.C. Chan, W. Chen, X. Chen, W.J. Chng, J.K. Choi, I. Colmenero, S.E. Coupland, N.C.P. Cross, D. De Jong, M.T. Elghetany, E. Takahashi, J.F. Emile, J. Ferry, L. Fogelstrand, M. Fontenay, U. Germing, S. Gujral, T. Haferlach, C. Harrison, J.C. Hodge, S. Hu, J.H. Jansen, R. Kanagal-Shamanna, H.M. Kantarjian, C.P. Kratz, X.Q. Li, M.S. Lim, K. Loeb, S. Loghavi, A. Marcogliese, S. Meshinchi, P. Michaels, K.N. Naresh, Y. Natkunam, R. Nejati, G. Ott, E. Padron, K.P. Patel, N. Patkar, J. Picarsic, U. Platzbecker, I. Roberts, A. Schuh, W. Sewell, R. Siebert, P. Tembhare, J. Tyner, S. Verstovsek, W. Wang, B. Wood, W. Xiao, C. Yeung, and A. Hochhaus. 2022. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia* 36:1703-1719.

Khwaja, A., M. Bjorkholm, R.E. Gale, R.L. Levine, C.T. Jordan, G. Ehninger, C.D. Bloomfield, E. Estey, A. Burnett, J.J. Cornelissen, D.A. Scheinberg, D. Bouscary, and D.C. Linch. 2016. Acute myeloid leukaemia. *Nature Reviews Disease Primers* 2:16010.

Kingston, R.E., and J.W. Tamkun. 2014. Transcriptional regulation by trithorax-group proteins. *Cold Spring Harb Perspect Biol* 6:a019349.

Kitamura, T., N. Watanabe-Okochi, Y. Enomoto, F. Nakahara, T. Oki, Y. Komeno, N. Kato, N. Doki, T. Uchida, Y. Kagiyama, K. Togami, K.C. Kawabata, K. Nishimura, Y. Hayashi, R. Nagase, M. Saika, T. Fukushima, S. Asada, T. Fujino, Y. Izawa, S. Horikawa, T. Fukuyama, Y. Tanaka, R. Ono, S. Goyama, T. Nosaka, J. Kitaura, and D. Inoue.

2016. Novel working hypothesis for pathogenesis of hematological malignancies: combination of mutations-induced cellular phenotypes determines the disease (cMIP-DD). *J Biochem* 159:17-25.

Koeffler, H.P., and G. Leong. 2017. Preleukemia: one name, many meanings. *Leukemia* 31:534-542.

Kohnken, R., P. Porcu, and A. Mishra. 2017. Overview of the Use of Murine Models in Leukemia and Lymphoma Research. *Front Oncol* 7:22.

Koya, J., K. Kataoka, T. Sato, M. Bando, Y. Kato, T. Tsuruta-Kishino, H. Kobayashi, K. Narukawa, H. Miyoshi, K. Shirahige, and M. Kurokawa. 2016. DNMT3A R882 mutants interact with polycomb proteins to block haematopoietic stem and leukaemic cell differentiation. *Nat Commun* 7:10924.

Kratz, C.P., B.M. Emerling, S. Donovan, M. Laig-Webster, B.R. Taylor, P. Thompson, S. Jensen, A. Banerjee, J. Bonifas, W. Makalowski, E.D. Green, M.M. Le Beau, and K.M. Shannon. 2001. Candidate gene isolation and comparative analysis of a commonly deleted segment of 7q22 implicated in myeloid malignancies. *Genomics* 77:171-180.

Kumar, V.C., and R. Pai. 2021. Genes of the month: H3.3 histone genes: H3F3A and H3F3B. *J Clin Pathol* 74:753-758.

Kunadt, D., S. Stasik, K.H. Metzeler, C. Röllig, C. Schliemann, P.A. Greif, K. Spiekermann, M. Rothenberg-Thurley, U. Krug, J. Braess, A. Krämer, A. Hochhaus, S. Scholl, I. Hilgendorf, T.H. Brümmendorf, E. Jost, B. Steffen, G. Bug, H. Einsele, D. Görlich, C. Sauerland, K. Schäfer-Eckart, S.W. Krause, M. Hänel, M. Hanoun, M. Kaufmann, B. Wörmann, M. Kramer, K. Sockel, K. Egger-Heidrich, T. Herold, G. Ehninger, A. Burchert, U. Platzbecker, W.E. Berdel, C. Müller-Tidow, W. Hiddemann, H. Serve, M. Stelljes, C.D. Baldus, A. Neubauer, J. Schetelig, C. Thiede, M. Bornhäuser, J.M. Middeke, F. Stölzel, and S.A.L. the A. M. L. Cooperative Group. 2022. Impact of IDH1 and IDH2 mutational subgroups in AML patients after allogeneic stem cell transplantation. *Journal of Hematology & Oncology* 15:126.

Kurtz, K.J., S.E. Conneely, M. O'Keefe, K. Wohlan, and R.E. Rau. 2022. Murine Models of Acute Myeloid Leukemia. *Front Oncol* 12:854973.

Kwarteng, E.O., and K.M. Heinonen. 2016. Competitive Transplants to Evaluate Hematopoietic Stem Cell Fitness. *J Vis Exp*

Larson, J.D., and S.J. Baker. 2019. Engineering Inducible Knock-In Mice to Model Oncogenic Brain Tumor Mutations from Endogenous Loci. *Methods Mol Biol* 1869:207-230.

Larson, J.D., L.H. Kasper, B.S. Paugh, H. Jin, G. Wu, C.H. Kwon, Y. Fan, T.I. Shaw, A.B. Silveira, C. Qu, R. Xu, X. Zhu, J. Zhang, H.R. Russell, J.L. Peters, D. Finkelstein, B. Xu, T. Lin, C.L. Tinkle, Z. Patay, A. Onar-Thomas, S.B. Pounds, P.J. McKinnon, D.W. Ellison, J. Zhang, and S.J. Baker. 2019. Histone H3.3 K27M Accelerates Spontaneous Brainstem Glioma and Drives Restricted Changes in Bivalent Gene Expression. *Cancer Cell* 35:140-155.e147.

Le Beau, M.M., R. Espinosa, 3rd, E.M. Davis, J.D. Eisenbart, R.A. Larson, and E.D. Green. 1996. Cytogenetic and molecular delineation of a region of chromosome 7 commonly deleted in malignant myeloid diseases. *Blood* 88:1930-1935.

Lehnertz, B., Y.W. Zhang, I. Boivin, N. Mayotte, E. Tomellini, J. Chagraoui, V.-P. Lavallée, J. Hébert, and G. Sauvageau. 2017. H3K27M/I mutations promote context-dependent transformation in acute myeloid leukemia with RUNX1 alterations. *Blood* 130:2204-2214.

Lemonnier, F., R.A. Cairns, S. Inoue, W.Y. Li, A. Dupuy, S. Broutin, N. Martin, V. Fataccioli, R. Pelletier, A. Wakeham, B.E. Snow, L. de Leval, A. Pujals, C. Haioun, A. Paci, E.R. Tobin, R. Narayanaswamy, K. Yen, S. Jin, P. Gaulard, and T.W. Mak. 2016. The IDH2 R172K mutation associated with angioimmunoblastic T-cell lymphoma produces 2HG in T cells and impacts lymphoid development. *Proc Natl Acad Sci U S A* 113:15084-15089.

Lewis, P.W., and C.D. Allis. 2013. Poisoning the "histone code" in pediatric gliomagenesis. *Cell Cycle* 12:3241-3242.

Lewis, P.W., M.M. Müller, M.S. Koletsky, F. Cordero, S. Lin, L.A. Banaszynski, B.A. Garcia, T.W. Muir, O.J. Becher, and C.D. Allis. 2013. Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric Glioblastoma. *Science* 340:857-861.

Ley, T.J., L. Ding, M.J. Walter, M.D. McLellan, T. Lamprecht, D.E. Larson, C. Kandoth, J.E. Payton, J. Baty, J. Welch, C.C. Harris, C.F. Lichti, R.R. Townsend, R.S. Fulton, D.J. Dooling, D.C. Koboldt, H. Schmidt, Q. Zhang, J.R. Osborne, L. Lin, M. O'Laughlin, J.F. McMichael, K.D. Delehaunty, S.D. McGrath, L.A. Fulton, V.J. Magrini, T.L. Vickery, J. Hundal, L.L. Cook, J.J. Conyers, G.W. Swift, J.P. Reed, P.A. Alldredge, T. Wylie, J. Walker, J. Kalicki, M.A. Watson, S. Heath, W.D. Shannon, N. Varghese, R. Nagarajan, P. Westervelt, M.H. Tomasson, D.C. Link, T.A. Graubert, J.F. DiPersio, E.R. Mardis, and R.K. Wilson. 2010. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* 363:2424-2433.

Li, M.W., B.J. Willis, K.D. Evans, R.S. Araiza, A.Y. Lee, and K.C. Lloyd. 2015. Rescue of germline transmission from chimeras by IVF after sperm analysis. *Transgenic Res* 24:99-108.

Lilly, A.J., W.E. Johnson, and C.M. Bunce. 2011. The haematopoietic stem cell niche: new insights into the mechanisms regulating haematopoietic stem cell behaviour. *Stem Cells Int* 2011:274564.

Liu, H., T.D. Westergard, and J.J. Hsieh. 2009. MLL5 governs hematopoiesis: a step closer. *Blood* 113:1395-1396.

Lu, V.M., M.A. Alvi, K.L. McDonald, and D.J. Daniels. 2018. Impact of the H3K27M mutation on survival in pediatric high-grade glioma: a systematic review and meta-analysis. *J Neurosurg Pediatr* 23:308-316.

Lund, K., P.D. Adams, and M. Copland. 2014. EZH2 in normal and malignant hematopoiesis. *Leukemia* 28:44-49.

Madan, V., B. Madan, U. Brykczynska, F. Zilbermann, K. Hogeveen, K. Dohner, H. Dohner, O. Weber, C. Blum, H.R. Rodewald, P. Sassone-Corsi, A.H.F.M. Peters, and H.J. Fehling. 2009. Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog MLL5. *Blood* 113:1444-1454.

Martire, S., and L.A. Banaszynski. 2020. The roles of histone variants in fine-tuning chromatin organization and function. *Nature Reviews Molecular Cell Biology* 21:522-541.

Martire, S., A.A. Gogate, A. Whitmill, A. Tafessu, J. Nguyen, Y.C. Teng, M. Tastemel, and L.A. Banaszynski. 2019. Phosphorylation of histone H3.3 at serine 31 promotes p300 activity and enhancer acetylation. *Nat Genet* 51:941-946.

Mas-y-Mas, S., M. Barbon, C. Teyssier, H. Déméné, J.E. Carvalho, L.E. Bird, A. Lebedev, J. Fattori, M. Schubert, C. Dumas, W. Bourguet, and A. le Maire. 2016. The Human Mixed Lineage Leukemia 5 (MLL5), a Sequentially and Structurally Divergent SET Domain-Containing Protein with No Intrinsic Catalytic Activity. *PLoS One* 11:e0165139.

Mayle, A., L. Yang, B. Rodriguez, T. Zhou, E. Chang, C.V. Curry, G.A. Challen, W. Li, D. Wheeler, V.I. Rebel, and M.A. Goodell. 2015. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood* 125:629-638.

McDowell, D.G., N.A. Burns, and H.C. Parkes. 1998. Localised sequence regions possessing high melting temperatures prevent the amplification of a DNA mimic in competitive PCR. *Nucleic Acids Res* 26:3340-3347.

McMahon, K.A., S.Y.L. Hiew, S. Hadjur, H. Veiga-Fernandes, U. Menzel, A.J. Price, D. Kioussis, O. Williams, and H.J.M. Brady. 2007. Mll Has a Critical Role in Fetal and Adult Hematopoietic Stem Cell Self-Renewal. *Cell Stem Cell* 1:338-345.

Mechaal, A., S. Menif, S. Abbes, and I. Safra. 2019. EZH2, new diagnosis and prognosis marker in acute myeloid leukemia patients. *Adv Med Sci* 64:395-401.

Meggendorfer, M., L.V. Cappelli, W. Walter, C. Haferlach, W. Kern, B. Falini, and T. Haferlach. 2018. IDH1R132, IDH2R140 and IDH2R172 in AML: different genetic landscapes correlate with outcome and may influence targeted treatment strategies. *Leukemia* 32:1249-1253.

Menolfi, D., and S. Zha. 2020. ATM, ATR and DNA-PKcs kinases—the lessons from the mouse models: inhibition ≠ deletion. *Cell & Bioscience* 10:8.

Meyer, S.E., T. Qin, D.E. Muench, K. Masuda, M. Venkatasubramanian, E. Orr, L. Suarez, S.D. Gore, R. Delwel, E. Paietta, M.S. Tallman, H. Fernandez, A. Melnick, M.M. Le Beau, S. Kogan, N. Salomonis, M.E. Figueroa, and H.L. Grimes. 2016. DNMT3A Haploinsufficiency Transforms FLT3ITD Myeloproliferative Disease into a Rapid, Spontaneous, and Fully Penetrant Acute Myeloid Leukemia. *Cancer Discov* 6:501-515.

Millán-Zambrano, G., A. Burton, A.J. Bannister, and R. Schneider. 2022. Histone post-translational modifications — cause and consequence of genome function. *Nature Reviews Genetics* 23:563-580.

Milne, T.A. 2017. Mouse models of MLL leukemia: recapitulating the human disease. *Blood* 129:2217-2223.

Mizuki, M., R. Fenski, H. Halfter, I. Matsumura, R. Schmidt, C. Müller, W. Grüning, K. Kratz-Albers, S. Serve, C. Steur, T. Büchner, J. Kienast, Y. Kanakura, W.E. Berdel, and H. Serve. 2000. Flt3 mutations from patients with acute myeloid leukemia induce transformation of 32D cells mediated by the Ras and STAT5 pathways. *Blood* 96:3907-3914.

Mohammad, F., and K. Helin. 2017. Oncohistones: drivers of pediatric cancers. *Genes Dev* 31:2313-2324.

Mohanty, S., and M. Heuser. 2021. Mouse Models of Frequently Mutated Genes in Acute Myeloid Leukemia. *Cancers (Basel)* 13:

Müller, U. 1999. Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mech Dev* 82:3-21.

Nagai, T., K. Ibata, E.S. Park, M. Kubota, K. Mikoshiba, and A. Miyawaki. 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology* 20:87-90.

Nagy, A. 2000. Cre recombinase: the universal reagent for genome tailoring. *Genesis* 26:99-109.

Nagy, A., J. Rossant, R. Nagy, W. Abramow-Newerly, and J.C. Roder. 1993. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc Natl Acad Sci U S A* 90:8424-8428.

O'Connell, K.E., A.M. Mikkola, A.M. Stepanek, A. Vernet, C.D. Hall, C.C. Sun, E. Yildirim, J.F. Staropoli, J.T. Lee, and D.E. Brown. 2015. Practical murine hematopathology: a comparative review and implications for research. *Comp Med* 65:96-113.

Ogawara, Y., T. Katsumoto, Y. Aikawa, Y. Shima, Y. Kagiyama, T. Soga, H. Matsunaga, T. Seki, K. Araki, and I. Kitabayashi. 2015. IDH2 and NPM1 Mutations Cooperate to Activate Hoxa9/Meis1 and Hypoxia Pathways in Acute Myeloid Leukemia. *Cancer Research* 75:2005-2016.

Ogilvy, S., D. Metcalf, L. Gibson, M.L. Bath, A.W. Harris, and J.M. Adams. 1999. Promoter elements of vav drive transgene expression in vivo throughout the hematopoietic compartment. *Blood* 94:1855-1863.

Panditharatna, E., J.G. Marques, T. Wang, M.C. Trissal, I. Liu, L. Jiang, A. Beck, A. Groves, N.V. Dharia, D. Li, S.E. Hoffman, G. Kugener, M.L. Shaw, H.M. Mire, O.A. Hack, J.M. Dempster, C. Lareau, L. Dai, L.H. Sigua, M.A. Quezada, A.J. Stanton, M. Wyatt, Z. Kalani, A. Goodale, F. Vazquez, F. Piccioni, J.G. Doench, D.E. Root, J.N. Anastas, K.L. Jones, A.S. Conway, S. Stopka, M.S. Regan, Y. Liang, H.S. Seo, K. Song, P. Bashyal, W.P. Jerome, N.D. Mathewson, S. Dhe-Paganon, M.L. Suvà, A.M.

Carcaboso, C. Lavarino, J. Mora, Q.D. Nguyen, K.L. Ligon, Y. Shi, S. Agnihotri, N.Y.R. Agar, K. Stegmaier, C.D. Stiles, M. Monje, T.R. Golub, J. Qi, and M.G. Filbin. 2022. BAF Complex Maintains Glioma Stem Cells in Pediatric H3K27M Glioma. *Cancer Discov* 12:2880-2905.

Pathania, M., N. De Jay, N. Maestro, A.S. Harutyunyan, J. Nitarska, P. Pahlavan, S. Henderson, L.G. Mikael, A. Richard-Londt, Y. Zhang, J.R. Costa, S. Hébert, S. Khazaei, N.S. Ibrahim, J. Herrero, A. Riccio, S. Albrecht, R. Ketteler, S. Brandner, C.L. Kleinman, N. Jabado, and P. Salomoni. 2017. H3.3(K27M) Cooperates with Trp53 Loss and PDGFRA Gain in Mouse Embryonic Neural Progenitor Cells to Induce Invasive High-Grade Gliomas. *Cancer Cell* 32:684-700.e689.

Paull, T.T. 2015. Mechanisms of ATM Activation. *Annual Review of Biochemistry* 84:711-738.

Pijnappel, W.W., D. Schaft, A. Roguev, A. Shevchenko, H. Tekotte, M. Wilm, G. Rigaut, B. Séraphin, R. Aasland, and A.F. Stewart. 2001. The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev* 15:2991-3004.

Rao, R.C., and Y. Dou. 2015. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* 15:334-346.

Ribeiro, A.F., M. Pratcorona, C. Erpelinck-Verschueren, V. Rockova, M. Sanders, S. Abbas, M.E. Figueroa, A. Zeilemaker, A. Melnick, B. Löwenberg, P.J. Valk, and R. Delwel. 2012. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. *Blood* 119:5824-5831.

Richmond, A., and Y. Su. 2008. Mouse xenograft models vs GEM models for human cancer therapeutics. *Dis Model Mech* 1:78-82.

Rincon-Arano, H., J. Halow, J.J. Delrow, S.M. Parkhurst, and M. Groudine. 2012. UpSET recruits HDAC complexes and restricts chromatin accessibility and acetylation at promoter regions. *Cell* 151:1214-1228.

Rincon-Arano, H., S.M. Parkhurst, and M. Groudine. 2013. UpSET-ting the balance: modulating open chromatin features in metazoan genomes. *Fly (Austin)* 7:153-160.

Rinke, J., A. Chase, N.C.P. Cross, A. Hochhaus, and T. Ernst. 2020. EZH2 in Myeloid Malignancies. *Cells* 9:

Rosenthal, N., and S. Brown. 2007. The mouse ascending: perspectives for human-disease models. *Nat Cell Biol* 9:993-999.

Russler-Germain, D.A., D.H. Spencer, M.A. Young, T.L. Lamprecht, C.A. Miller, R. Fulton, M.R. Meyer, P. Erdmann-Gilmore, R.R. Townsend, R.K. Wilson, and T.J. Ley. 2014. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell* 25:442-454.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 2001. Molecular cloning: A laboratory manual. CSHL, New York.

Sanchez, R., and M.M. Zhou. 2011. The PHD finger: a versatile epigenome reader. *Trends Biochem Sci* 36:364-372.

Santos, M.A., R.B. Faryabi, A.V. Ergen, A.M. Day, A. Malhowski, A. Canela, M. Onozawa, J.E. Lee, E. Callen, P. Gutierrez-Martinez, H.T. Chen, N. Wong, N. Finkel, A. Deshpande, S. Sharrow, D.J. Rossi, K. Ito, K. Ge, P.D. Aplan, S.A. Armstrong, and A. Nussenzweig. 2014. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. *Nature* 514:107-111.

Schmidt-Supprian, M., and K. Rajewsky. 2007. Vagaries of conditional gene targeting. *Nat Immunol* 8:665-668.

Schwartzentruber, J., A. Korshunov, X.-Y. Liu, D.T.W. Jones, E. Pfaff, K. Jacob, D. Sturm, A.M. Fontebasso, D.-A.K. Quang, M. Tönjes, V. Hovestadt, S. Albrecht, M. Kool, A. Nantel, C. Konermann, A. Lindroth, N. Jäger, T. Rausch, M. Ryzhova, J.O. Korbel, T. Hielscher, P. Hauser, M. Garami, A. Klekner, L. Bognar, M. Ebinger, M.U. Schuhmann, W. Scheurlen, A. Pekrun, M.C. Frühwald, W. Roggendorf, C. Kramm, M.

Dürken, J. Atkinson, P. Lepage, A. Montpetit, M. Zakrzewska, K. Zakrzewski, P.P. Liberski, Z. Dong, P. Siegel, A.E. Kulozik, M. Zapatka, A. Guha, D. Malkin, J. Felsberg, G. Reifenberger, A. von Deimling, K. Ichimura, V.P. Collins, H. Witt, T. Milde, O. Witt, C. Zhang, P. Castelo-Branco, P. Lichter, D. Faury, U. Tabori, C. Plass, J. Majewski, S.M. Pfister, and N. Jabado. 2012. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482:226-231.

Scourzic, L., L. Couronné, M.T. Pedersen, V. Della Valle, M. Diop, E. Mylonas, J. Calvo, E. Mouly, C.K. Lopez, N. Martin, M. Fontenay, A. Bender, S. Guibert, P. Dubreuil, P. Dessen, N. Droin, F. Pflumio, M. Weber, P. Gaulard, K. Helin, T. Mercher, and O.A. Bernard. 2016. DNMT3A(R882H) mutant and Tet2 inactivation cooperate in the deregulation of DNA methylation control to induce lymphoid malignancies in mice. *Leukemia* 30:1388-1398.

Seibler, J., B. Zevnik, B. Küter-Luks, S. Andreas, H. Kern, T. Hennek, A. Rode, C. Heimann, N. Faust, G. Kauselmann, M. Schoor, R. Jaenisch, K. Rajewsky, R. Kühn, and F. Schwenk. 2003. Rapid generation of inducible mouse mutants. *Nucleic Acids Res* 31:e12.

Shaner, N.C., R.E. Campbell, P.A. Steinbach, B.N.G. Giepmans, A.E. Palmer, and R.Y. Tsien. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* 22:1567-1572.

Sharpless, N.E., and R.A. DePinho. 2006. The mighty mouse: genetically engineered mouse models in cancer drug development. *Nature Reviews Drug Discovery* 5:741-754.

Shen, F.W., Y. Saga, G. Litman, G. Freeman, J.S. Tung, H. Cantor, and E.A. Boyse. 1985. Cloning of Ly-5 cDNA. *Proc Natl Acad Sci U S A* 82:7360-7363.

Shi, L., H. Wen, and X. Shi. 2017. The Histone Variant H3.3 in Transcriptional Regulation and Human Disease. *J Mol Biol* 429:1934-1945.

Sieber, S., L. Wirth, N. Cavak, M. Koenigsmark, U. Marx, R. Lauster, and M. Rosowski. 2018. Bone marrow-on-a-chip: Long-term culture of human haematopoietic stem cells in a three-dimensional microfluidic environment. *Journal of Tissue Engineering and Regenerative Medicine* 12:479-489.

Silva-Santana, G., J.C. Bax, D.C.S. Fernandes, D.T.L. Bacellar, C. Hooper, A. Dias, C.B. Silva, A.M. de Souza, S. Ramos, R.A. Santos, T.R. Pinto, M.A. Ramão, and A.L. Mattos-Guaraldi. 2020. Clinical hematological and biochemical parameters in Swiss, BALB/c, C57BL/6 and B6D2F1 Mus musculus. *Animal Model Exp Med* 3:304-315.

Simon, J.A., and C.A. Lange. 2008. Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutat Res* 647:21-29.

Simon, L., J.-F. Spinella, C.-Y. Yao, V.-P. Lavallée, I. Boivin, G. Boucher, E. Audemard, M.-E. Bordeleau, S. Lemieux, J. Hébert, and G. Sauvageau. 2020. High frequency of germline RUNX1 mutations in patients with RUNX1-mutated AML. *Blood* 135:1882-1886.

Srinivas, S., T. Watanabe, C.S. Lin, C.M. William, Y. Tanabe, T.M. Jessell, and F. Costantini. 2001. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* 1:4.

Stafford, J.M., C.H. Lee, P. Voigt, N. Descostes, R. Saldaña-Meyer, J.R. Yu, G. Leroy, O. Oksuz, J.R. Chapman, F. Suarez, A.S. Modrek, N.S. Bayin, D.G. Placantonakis, M.A. Karajannis, M. Snuderl, B. Ueberheide, and D. Reinberg. 2018. Multiple modes of PRC2 inhibition elicit global chromatin alterations in H3K27M pediatric glioma. *Sci Adv* 4:eaau5935.

Stasik, S., J.-N. Eckardt, M. Kramer, C. Röllig, A. Krämer, S. Scholl, A. Hochhaus, M. Crysandt, T.H. Brümmendorf, R. Naumann, B. Steffen, V. Kunzmann, H. Einsele, M. Schaich, A. Burchert, A. Neubauer, K. Schäfer-Eckart, C. Schliemann, S. Krause, R. Herbst, M. Hänel, N. Frickhofen, R. Noppeney, U. Kaiser, C.D. Baldus, M. Kaufmann, Z. Rácil, U. Platzbecker, W.E. Berdel, J. Mayer, H. Serve, C. Müller-Tidow, G.

Ehninger, M. Bornhäuser, J. Schetelig, J.M. Middeke, C. Thiede, and o.b.o.t.S.A. Leukemia. 2021. Impact of PTPN11 mutations on clinical outcome analyzed in 1529 patients with acute myeloid leukemia. *Blood Advances* 5:3279-3289.

Stubbs, M.C., and A.V. Krivtsov. 2017. Murine Retrovirally-Transduced Bone Marrow Engraftment Models of MLL-Fusion-Driven Acute Myelogenous Leukemias (AML). *Curr Protoc Pharmacol* 78:14.42.11-14.42.19.

Sturm, D., H. Witt, V. Hovestadt, D.A. Khuong-Quang, D.T. Jones, C. Konermann, E. Pfaff, M. Tönjes, M. Sill, S. Bender, M. Kool, M. Zapatka, N. Becker, M. Zucknick, T. Hielscher, X.Y. Liu, A.M. Fontebasso, M. Ryzhova, S. Albrecht, K. Jacob, M. Wolter, M. Ebinger, M.U. Schuhmann, T. van Meter, M.C. Frühwald, H. Hauch, A. Pekrun, B. Radlwimmer, T. Niehues, G. von Komorowski, M. Dürken, A.E. Kulozik, J. Madden, A. Donson, N.K. Foreman, R. Drissi, M. Fouladi, W. Scheurlen, A. von Deimling, C. Monoranu, W. Roggendorf, C. Herold-Mende, A. Unterberg, C.M. Kramm, J. Felsberg, C. Hartmann, B. Wiestler, W. Wick, T. Milde, O. Witt, A.M. Lindroth, J. Schwartzentruber, D. Faury, A. Fleming, M. Zakrzewska, P.P. Liberski, K. Zakrzewski, P. Hauser, M. Garami, A. Klekner, L. Bognar, S. Morrissy, F. Cavalli, M.D. Taylor, P. van Sluis, J. Koster, R. Versteeg, R. Volckmann, T. Mikkelsen, K. Aldape, G. Reifenberger, V.P. Collins, J. Majewski, A. Korshunov, P. Lichter, C. Plass, N. Jabado, and S.M. Pfister. 2012. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* 22:425-437.

Szenker, E., D. Ray-Gallet, and G. Almouzni. 2011. The double face of the histone variant H3.3. *Cell Research* 21:421-434.

Takahashi, S. 2011. Current findings for recurring mutations in acute myeloid leukemia. *J Hematol Oncol* 4:36.

Tang, M.C., S.A. Jacobs, L.H. Wong, and J.R. Mann. 2013. Conditional allelic replacement applied to genes encoding the histone variant H3.3 in the mouse. *Genesis* 51:142-146.

Taniuchi, I., M. Osato, T. Egawa, M.J. Sunshine, S.C. Bae, T. Komori, Y. Ito, and D.R. Littman. 2002. Differential requirements for Runx proteins in CD4 repression and epigenetic silencing during T lymphocyte development. *Cell* 111:621-633.

Tasdogan, A., S. Kumar, G. Allies, J. Bausinger, F. Beckel, H. Hofemeister, M. Mulaw, V. Madan, K. Scharffetter-Kochanek, M. Feuring-Buske, K. Doehner, G. Speit, A.F. Stewart, and H.J. Fehling. 2016. DNA Damage-Induced HSPC Malfunction Depends on ROS Accumulation Downstream of IFN-1 Signaling and Bid Mobilization. *Cell Stem Cell* 19:752-767.

te Riele, H., E.R. Maandag, and A. Berns. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc Natl Acad Sci U S A* 89:5128-5132.

Telford, W.G., T. Hawley, F. Subach, V. Verkhusha, and R.G. Hawley. 2012. Flow cytometry of fluorescent proteins. *Methods* 57:318-330.

Thiel, A.T., P. Blessington, T. Zou, D. Feather, X. Wu, J. Yan, H. Zhang, Z. Liu, P. Ernst, G.A. Koretzky, and X. Hua. 2010. MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. *Cancer Cell* 17:148-159.

Tran, K., and E.M. Green. 2019. SET domains and stress: uncovering new functions for yeast Set4. *Curr Genet* 65:643-648.

Tran, K., Y. Jethmalani, D. Jaiswal, and E.M. Green. 2018. Set4 is a chromatin-associated protein, promotes survival during oxidative stress, and regulates stress response genes in yeast. *J Biol Chem* 293:14429-14443.

Tronche, F., C. Kellendonk, O. Kretz, P. Gass, K. Anlag, P.C. Orban, R. Bock, R. Klein, and G. Schütz. 1999. Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nature Genetics* 23:99-103.

Trowbridge, I.S., and M.L. Thomas. 1994. CD45: an emerging role as a protein tyrosine phosphatase required for lymphocyte activation and development. *Annu Rev Immunol* 12:85-116.

Vosberg, S., L. Hartmann, K.H. Metzeler, N.P. Konstandin, S. Schneider, A. Varadharajan, A. Hauser, S. Krebs, H. Blum, S.K. Bohlander, W. Hiddemann, J. Tischer, K. Spiekermann, and P.A. Greif. 2018. Relapse of acute myeloid leukemia after allogeneic stem cell transplantation is associated with gain of WT1 alterations and high mutation load. *Haematologica* 103:e581-e584.

Wang, J., T.Y.-T. Huang, Y. Hou, E. Bartom, X. Lu, A. Shilatifard, F. Yue, and A. Saratsis. 2021. Epigenomic landscape and 3D genome structure in pediatric high-grade glioma. *Science Advances* 7:eabg4126.

Ward, I.M., K. Minn, J. van Deursen, and J. Chen. 2003. p53 Binding protein 53BP1 is required for DNA damage responses and tumor suppression in mice. *Mol Cell Biol* 23:2556-2563.

Wiles, E.T., and E.U. Selker. 2017. H3K27 methylation: a promiscuous repressive chromatin mark. *Curr Opin Genet Dev* 43:31-37.

Wong, J.C., Y. Zhang, K.H. Lieuw, M.T. Tran, E. Forgo, K. Weinfurtner, P. Alzamora, S.C. Kogan, K. Akagi, L. Wolff, M.M. Le Beau, N. Killeen, and K. Shannon. 2010. Use of chromosome engineering to model a segmental deletion of chromosome band 7q22 found in myeloid malignancies. *Blood* 115:4524-4532.

Wu, G., A. Broniscer, T.A. McEachron, C. Lu, B.S. Paugh, J. Becksfort, C. Qu, L. Ding, R. Huether, M. Parker, J. Zhang, A. Gajjar, M.A. Dyer, C.G. Mullighan, R.J. Gilbertson, E.R. Mardis, R.K. Wilson, J.R. Downing, D.W. Ellison, J. Zhang, and S.J. Baker. 2012. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet* 44:251-253.

Xu, J., Y.-Y. Wang, Y.-J. Dai, W. Zhang, W.-N. Zhang, S.-M. Xiong, Z.-H. Gu, K.-K. Wang, R. Zeng, Z. Chen, and S.-J. Chen. 2014. DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. *Proceedings of the National Academy of Sciences* 111:2620-2625.

Yang, L., B. Rodriguez, A. Mayle, H.J. Park, X. Lin, M. Luo, M. Jeong, C.V. Curry, S.B. Kim, D. Ruau, X. Zhang, T. Zhou, M. Zhou, V.I. Rebel, G.A. Challen, B. Gottgens, J.S. Lee, R. Rau, W. Li, and M.A. Goodell. 2016. DNMT3A Loss Drives Enhancer Hypomethylation in FLT3-ITD-Associated Leukemias. *Cancer Cell* 29:922-934.

Yang, W., and P. Ernst. 2017. SET/MLL family proteins in hematopoiesis and leukemia. *International Journal of Hematology* 105:7-16.

Yuen, B.T., and P.S. Knoepfler. 2013. Histone H3.3 mutations: a variant path to cancer. *Cancer Cell* 24:567-574.

Zambrowicz, B.P., A. Imamoto, S. Fiering, L.A. Herzenberg, W.G. Kerr, and P. Soriano. 1997. Disruption of overlapping transcripts in the ROSA βgeo 26 gene trap strain leads to widespread expression of β-galactosidase in mouse embryos and hematopoietic cells. *Proceedings of the National Academy of Sciences* 94:3789-3794.

Zhang, A., Y. Liu, S. Wei, B. Gong, C. Zhou, Y. Wang, H. Wei, Y. Mi, and J. Wang. 2020. BCOR Mutations in Acute Myeloid Leukemia: Clonal Evolution and Prognosis. *Blood* 136:4-4.

Zhang, X., W. Novera, Y. Zhang, and L.W. Deng. 2017. MLL5 (KMT2E): structure, function, and clinical relevance. *Cell Mol Life Sci* 74:2333-2344.

Zhang, Y., J. Wong, M. Klinger, M.T. Tran, K.M. Shannon, and N. Killeen. 2009. MLL5 contributes to hematopoietic stem cell fitness and homeostasis. *Blood* 113:1455-1463.

Appendix

Full name of genes mentioned in Figure 1, Table 2, and Table 4

Abbreviation	Full Gene Name
ABCA12	ATP-binding cassette sub-family A member 12
ADD1	Alpha-adducin
AGAP2	ArfGAP with GTPase Domain, Ankyrin Repeat And PH Domain 2
AHNAK2	Neuroblast Differentiation-associated Protein AHNAK Nucleoprotein 2
ALS2CL	Alsin Rho Guanine Nucleotide Exchange Facto C-Terminal Like
ANK3	Ankyrin 3
AP001055.1	No full name available
APOBR	Apolipoprotein B Receptor
ARID1B	AT-rich Interactive Domain-containing Protein 1B
ARL4D	ADP Ribosylation Factor Like GTPase 4D
ARMC10	Armadillo Repeat Containing 10
ARV1	Androgen Receptor Splice Variant 1
ASXL1	Additional Sex Combs-Like 1
ATP10D	ATPase Phospholipid Transporting 10D
ATP13A2	Probable cation-transporting ATPase 13A2
BCL9	B-cell CLL/lymphoma 9
BCOR	B-cell lymphoma 6 Corepressor
BCORL1	BCL6 Corepressor Like 1
BNIP3L	B-cell lymphoma 2 Interacting Protein 3 Like
BRCA2	Breast Cancer Gene 1
BUB3	BUB3 Mitotic Checkpoint Protein
CACNA1I	Calcium Voltage-Gated Channel Subunit Alpha1 I
CBL	Casitas B-lineage Lymphoma
CCDC168	Coiled-Coil Domain Containing 168
CCM2	Cerebral Cavernous Malformations 2
CCND1	Cyclin D1
CCT6B	Chaperonin Containing T-Complex 1 Subunit 6B
CD22	Cluster of Differentiation 22
CDH23	Cadherin Related 23
CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5
CHRNB1	Cholinergic Receptor Nicotinic Beta 1 Subunit
CNOT3	CCR4-NOT Transcription Complex Subunit 3
COL5A	Collagen Type V Alpha 1 Chain
CROT	Carnitine O-Octanoyltransferase
CRTAC1	Cartilage Acidic Protein 1
CRTC1	cAMP Response Element-binding Protein Regulated Transcription Coactivator 1
CSF3R	Colony Stimulating Factor 3 Receptor
CSMD1	CUB And Sushi Multiple Domains 1
CTBP2	C-Terminal Binding Protein 2
CUL3	Cullin 3
CUTL1	Cut Like Homeobox 1
CYHR1	Cysteine/Histidine-Rich 1
DDX47	DEAD-Box Helicase 47
DHX57	DExH-Box Helicase 57

DNAH3	Dynein Axonemal Heavy Chain 3
DNAJC2	DnaJ Heat Shock Protein Family (Hsp40) Member C2
DNMT3A	DNA Methyltransferase 3 A
DOCK1	Dedicator Of Cytokinesis 1
DPY19L2P2	DPY19L2 Pseudogene 2
DYNC1H1	Dynein Cytoplasmic 1 Heavy Chain 1
E2F1	E2F Transcription Factor 1
EIF4G1	Eukaryotic Translation Initiation Factor 4 Gamma 1
EPO	Erythropoietin
F8	Coagulation Factor VIII
FAM174B	Family With Sequence Similarity 174 Member B
FAM9A	Family With Sequence Similarity 9 Member A
FANCM	Fanconi Anemia Group M
FAT1	FAT Atypical Cadherin 1
FBXL13	F-Box And Leucine Rich Repeat Protein 13
FGD5	FYVE, RhoGEF And PH Domain Containing 5
FLG2	Filaggrin 2
FLT3	FMS-like Tyrosine Kinase 3
FST	Follistatin
GAPVD1	GTPase Activating Protein And VPS9 Domains 1
GAS6	Growth-arrest-specific gene-6
GCNT2	Glucosaminyl (N-Acetyl) Transferase 2 (I Blood Group)
GDF3	Growth Differentiation Factor-3
GMEB1	Glucocorticoid Modulatory Element Binding Protein 1
GNAQ	Guanine nucleotide-binding protein G(q) subunit alpha
GPR21	G Protein-Coupled Receptor 21
GRID2IP	Grid2 Interacting Protei
GSDMA	Gasdermin A
GSE1	Genetic Suppressor Element 1 Coiled-Coil Protein
GTF2A	Transcription initiation factor IIA subunit 1
H3F3A	Histone 3.3 Family A
HS6ST3	Heparan Sulfate 6-O-Sulfotransferase 3
HTR3C	5-Hydroxytryptamine Receptor 3C
IARS	Soleucyl-TRNA Synthetase
IDH2	Isocitrate Dehydrogenase 2
IFT81	Intraflagellar Transport 81
IGFN1	Immunoglobulin Like And Fibronectin Type III Domain Containing 1
KCTD16	Potassium Channel Tetramerization Domain Containing 16
KDM2B	Lysine Demethylase 2B
LCAT	Lecithin-Cholesterin-Acyltransferase
LHFPL3	LHFPL Tetraspan Subfamily Member 3
LILRB4	Leukocyte Immunoglobulin Like Receptor B4
LIPT1	Lipoyltransferase 1
LRRC17	Leucine Rich Repeat Containing 17
LY6G6F	Lymphocyte Antigen 6 Family Member G6F
MAP2	Microtubule-associated Protein 2
MAPK8IP3	Mitogen-Activated Protein Kinase 8 Interacting Protein 3
MAPRE1	Microtubule Associated Protein RP/EB Family Member 1
MDGA1	MAM Domain Containing Glycosylphosphatidylinositol Anchor 1
MED12	Mediator Complex Subunit 12
MEGF6	Multiple EGF Like Domains 6)

METTL3	Methyltransferase-like 3
MGA	MAX Gene-Associated Protein
MIB2	E3 Ubiquitin-protein Ligase
MLL5	Mixed Lineage Leukemia 5
MS4A7	Membrane-spanning 4-domain Subfamily A member 7
MYL9	Myosin Light Chain 9
NAALADL1	N-Acetylated Alpha-Linked Acidic Dipeptidase Like 1
NAPEPLD	N-Acyl Phosphatidylethanolamine Phospholipase D
NBEAL1	Neurobeachin Like 1
NOVA2	NOVA Alternative Splicing Regulator 2
NR0B1	Nuclear Receptor Subfamily 0 Group B Member 1
NRAP	Nebulin-related-anchoring Protein
NRG3	Neuregulin 3
NYAP2	Neuronal Tyrosine-Phosphorylated Phosphoinositide-3-Kinase Adaptor 2
OCSTAMP	Osteoclast Stimulatory Transmembrane Protein
OR5R1	Olfactory Receptor Family 5 Subfamily R Member 1
OR6C3	Olfactory Receptor Family 6 Subfamily C Member 3
ORC5L	Origin Recognition Complex Subunit 5
OTOG	Otogelin
PAPD4	Terminal Nucleotidyltransferase 2
PAPLOB	poly (A) polymerase beta (Testis Specific)
PCDH9	Protocadherin 9
PEX11G	Peroxisomal Biogenesis Factor 11 Gamma
PIK3C2B	Phosphatidylinositol-4-phosphate 3-kinase C2 Domain-containing Beta Polypeptide
PIWIL2	Piwi Like RNA-Mediated Gene Silencing 2
PMPCB	Peptidase, Mitochondrial Processing Subunit Beta
PNLIPRP2	Pancreatic Lipase Related Protein 2
POLR2H	RNA Polymerase II, I And III Subunit H
PPFIA1	PTPRF Interacting Protein Alpha 1
PPWD1	Peptidylprolyl Isomerase Domain And WD Repeat Containing 1
PROKR2	Prokineticin Receptor 2
PSMC2	Proteasome 26S Subunit, ATPase 2
PTPN11	Protein Tyrosine Phosphatase Non-receptor Type 11
PTPRD	Protein Tyrosine Phosphatase Receptor Type D
PTPRJ	Protein Tyrosine Phosphatase Receptor Type J
PTPRQ	Protein Tyrosine Phosphatase Receptor Type Q
PXDN	Peroxidasin
RAD50	RAD50 Double Strand Break Repair Protein
RANBP3	RAN Binding Protein 3
RASA4	Ras GTPase-activating protein 4
RBL1	Retinoblastoma-like Protein 1
RBM26	RNA binding Motif Protein 26
RBM45	RNA binding Motif Protein 45
RELN	Reelin
REPS2	RALBP1 Associated Eps Domain Containing 2
RIMS2	Regulating Synaptic Membrane Exocytosis 2
RIPPLY3	Ripply Transcriptional Repressor 3
RNF17	Ring Finger Protein 17
RNFT2	Ring Finger Protein, Transmembrane 2
RPL24	Ribosomal Protein L24

RTN3	Reticulon 3
RUNX1	Runt-related Transcription Factor 1
RYR1	Ryanodine Receptor 1
SCNN1B	Sodium Channel Epithelial 1 Subunit Beta
SEC24C	SEC24 Homolog C
SHROOM2	Shroom Family Member 2
SLAMF1	Signaling Lymphocytic Activation Molecule Family Member 1
SLC26A5	Solute Carrier Family 26 Member 5
SLC39A8	Solute Carrier Family 39 Member 8
SLC6A13	Solute Carrier Family 6 Member 13
SLCO1C1	Solute Carrier Organic Anion Transporter Family Member 1C1
SMIM2	Small Integral Membrane Protein 2
SNAP91	Synaptosome Associated Protein 91
SNTB2	Syntrophin Beta 2
SPATA22	Spermatogenesis Associated 22
SRPK2	SRSF Protein Kinase 2
SRSF2	Serine And Arginine Rich Splicing Factor 2
STAG2	Stromal Antigen 2
TACC2	Transforming Acidic Coiled-Coil Containing Protein 2
TAF1C	TATA-Box Binding Protein Associated Factor RNA Polymerase I Subunit C
TCF20	Transcription Factor 20
TDRD3	Tudor Domain Containing 3
TEK	TEK Receptor Tyrosine Kinase
TENM3	Teneurin Transmembrane Protein 3
TMEM35	Transmembrane Protein 35
TP53	Tumor Protein 53
TRIM51	Tripartite Motif-Containing 51
TRO	Trophinin
TRPM1	Transient Receptor Potential Cation Channel Subfamily M Member 1
TUB	Tubby Protein Homolog
ULK4	Unc-51 Like Kinase 4
UNC5A	Unc-5 Netrin Receptor A
USP29	Ubiquitin Specific Peptidase 29
VSTM2A	V-Set And Transmembrane Domain Containing 2A
WBSCR28	Williams-Beuren Syndrome Chromosomal Region 28
YEATS2	YEATS Domain Containing 2
ZNF497	Zinc Finger Protein 497
ZNF681	Zinc Finger Protein 681
ZSCAN32	Zinc Finger And SCAN Domain Containing 32

Plasmids used during the Project

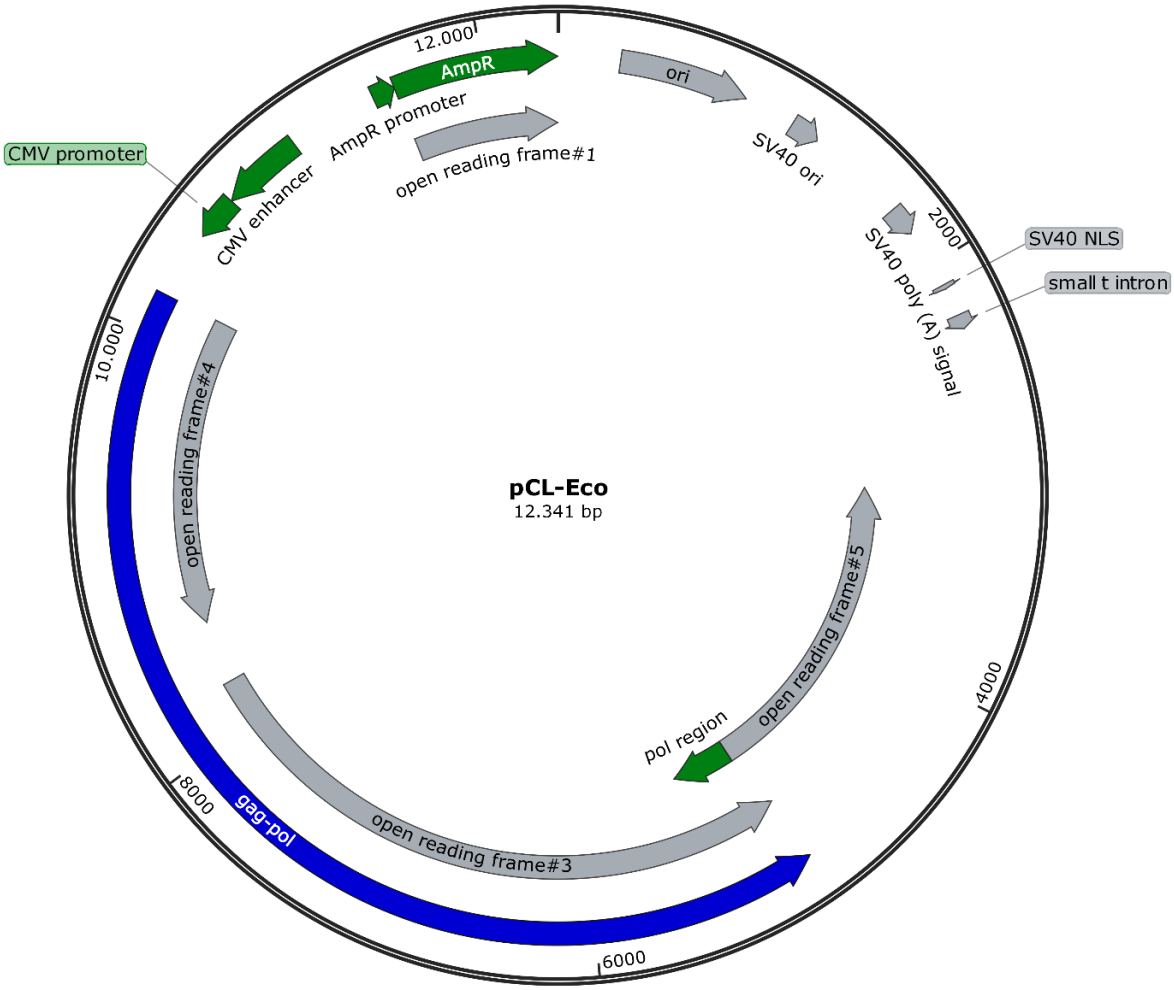


Figure 45A: Plasmid map of pCL-Eco helper plasmid used during generation of retroviral supernatants

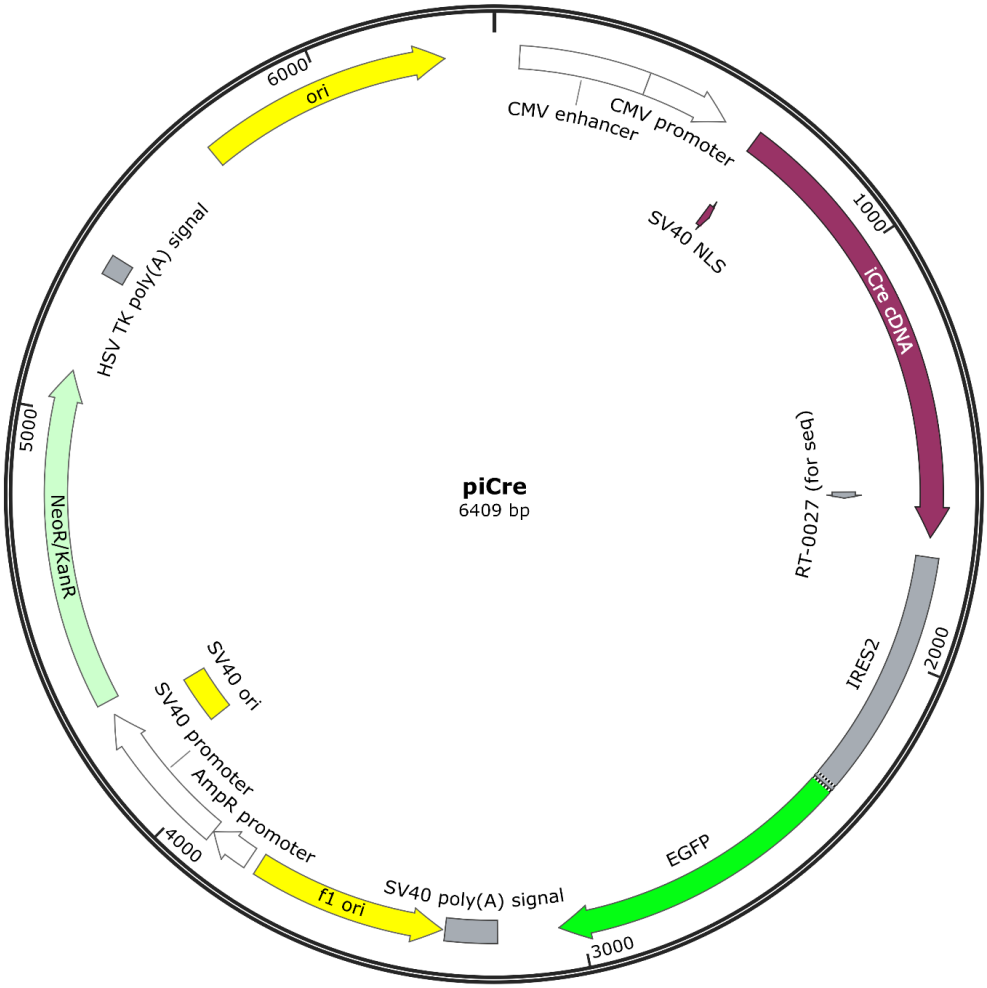


Figure 46A: Plasmid map of piCre (Cre-recombinase) used to activate K27M mutation in targeted ES cells

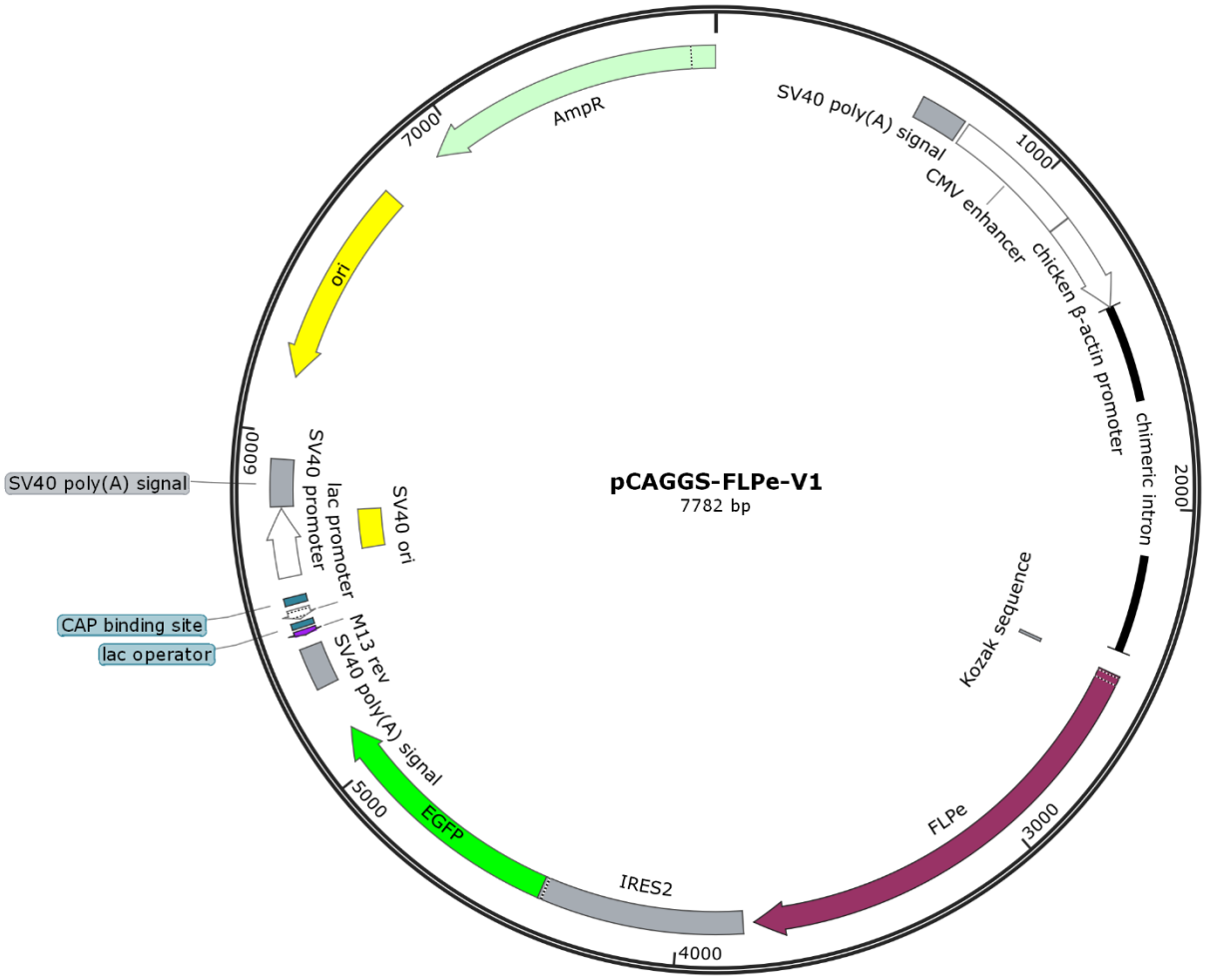


Figure 47A: Plasmid map of pCAGGS-FLPe-V1 (Flp-recombinase) used to delete Neo from targeted ES cells

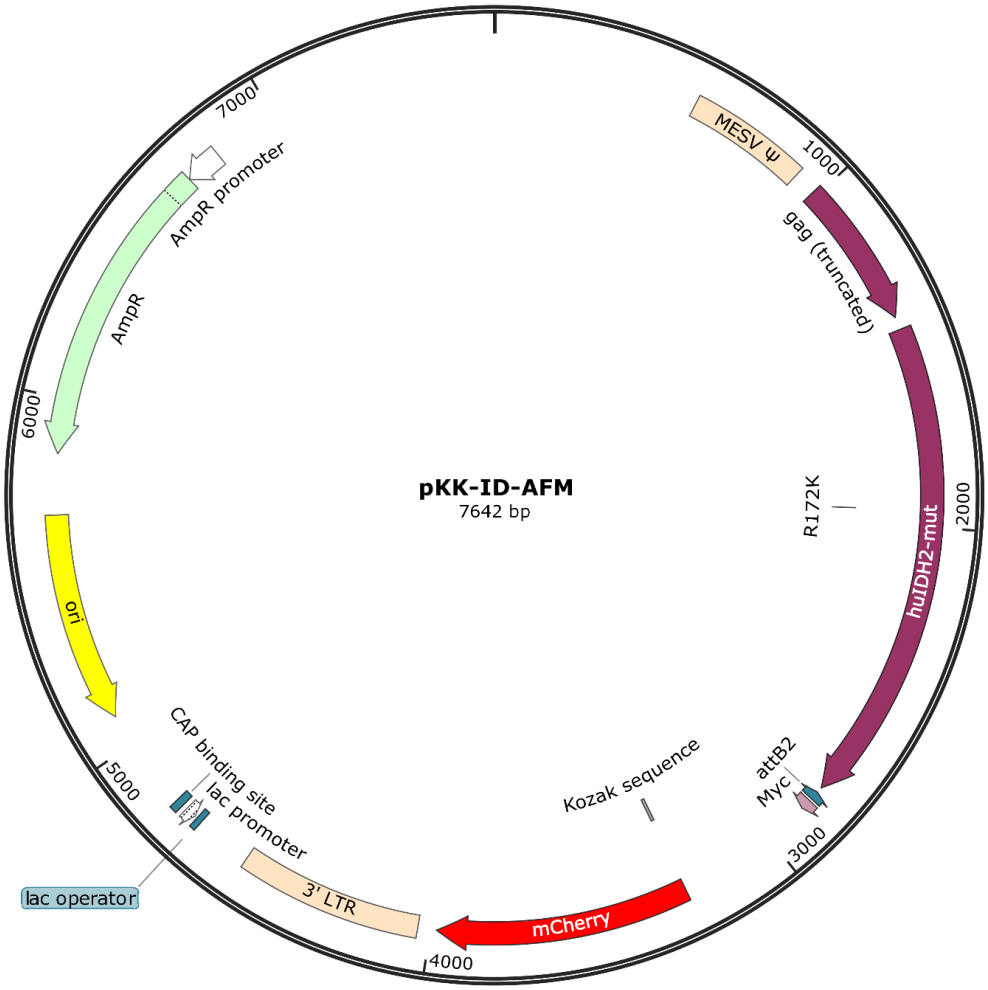


Figure 48A: Plasmid map of retroviral vector pKK-ID-AFM used for transduction of BM cells

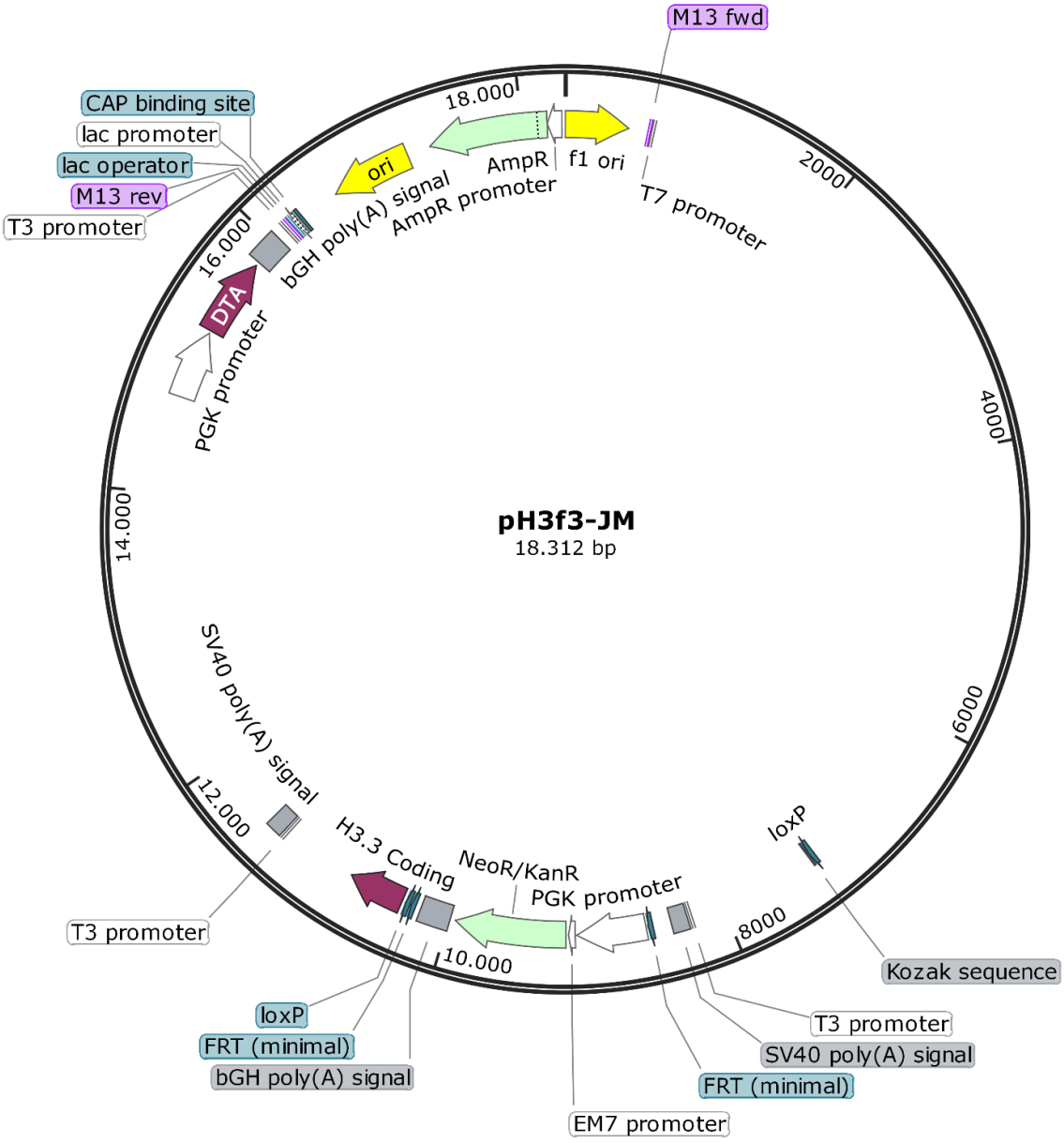


Figure 49A: Plasmid map of targeting construct obtained from the Jeffrey Mann laboratory

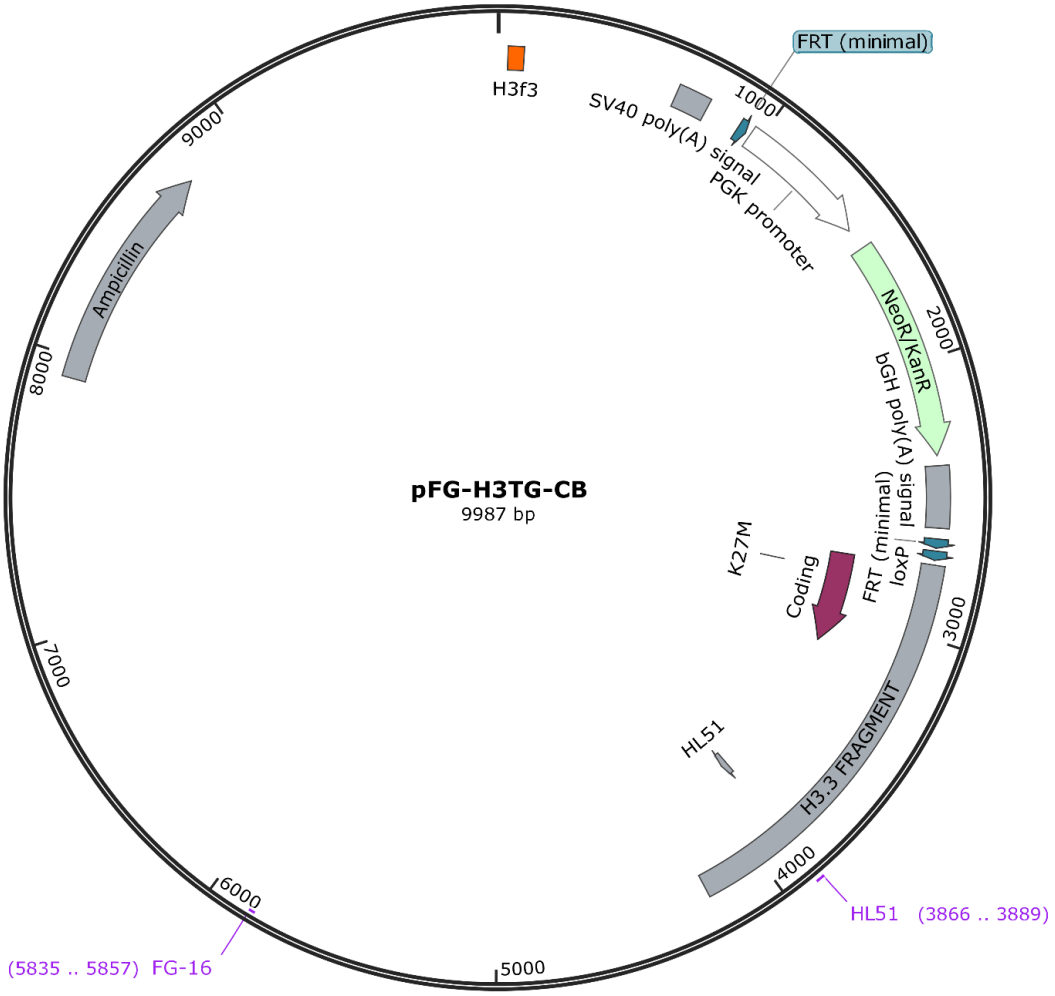


Figure 50A: Plasmid map of the control construct pFG-H3TG-CB

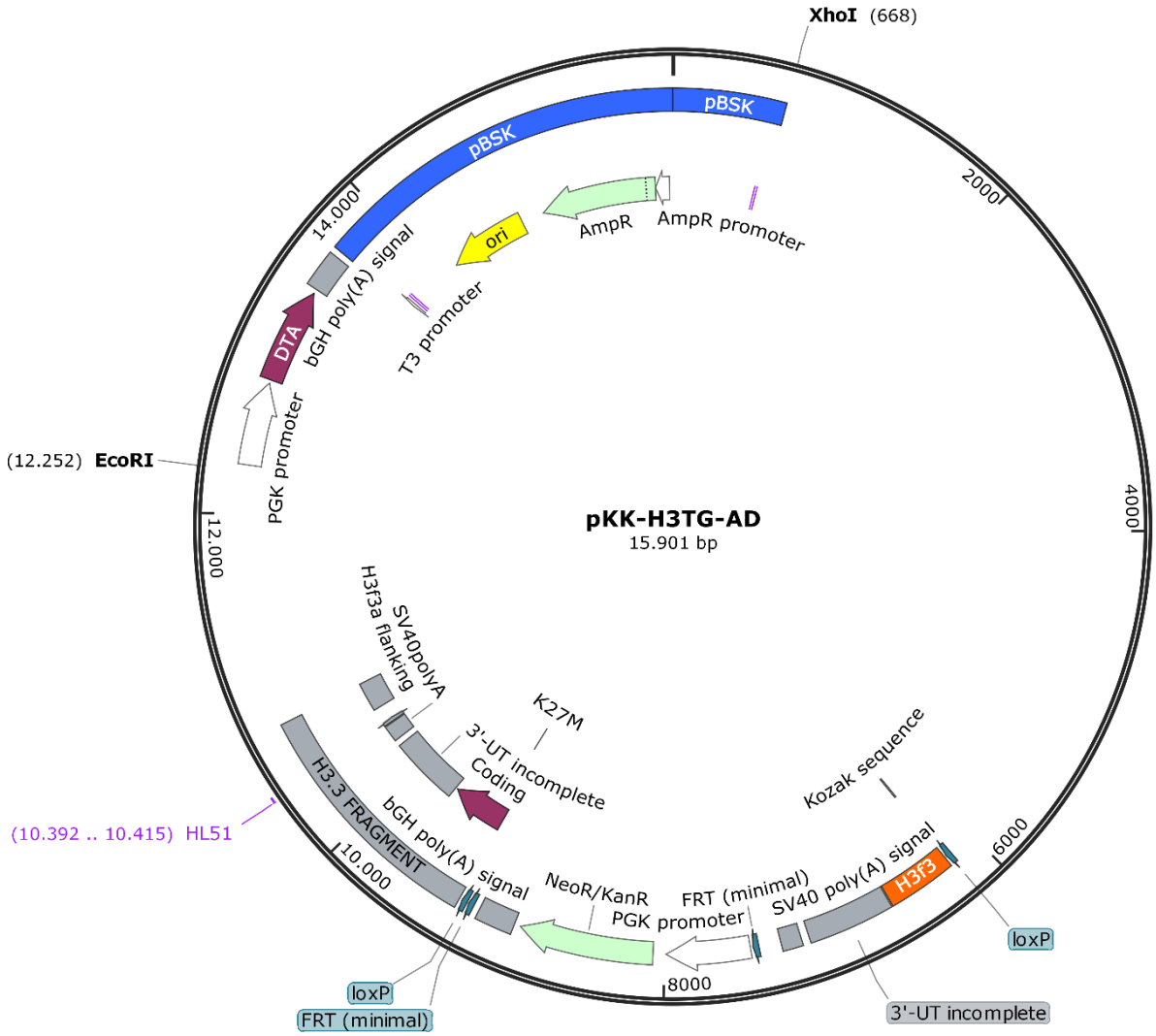


Figure 51A: Plasmid map of first targeting vector pKK-H3TG-AD

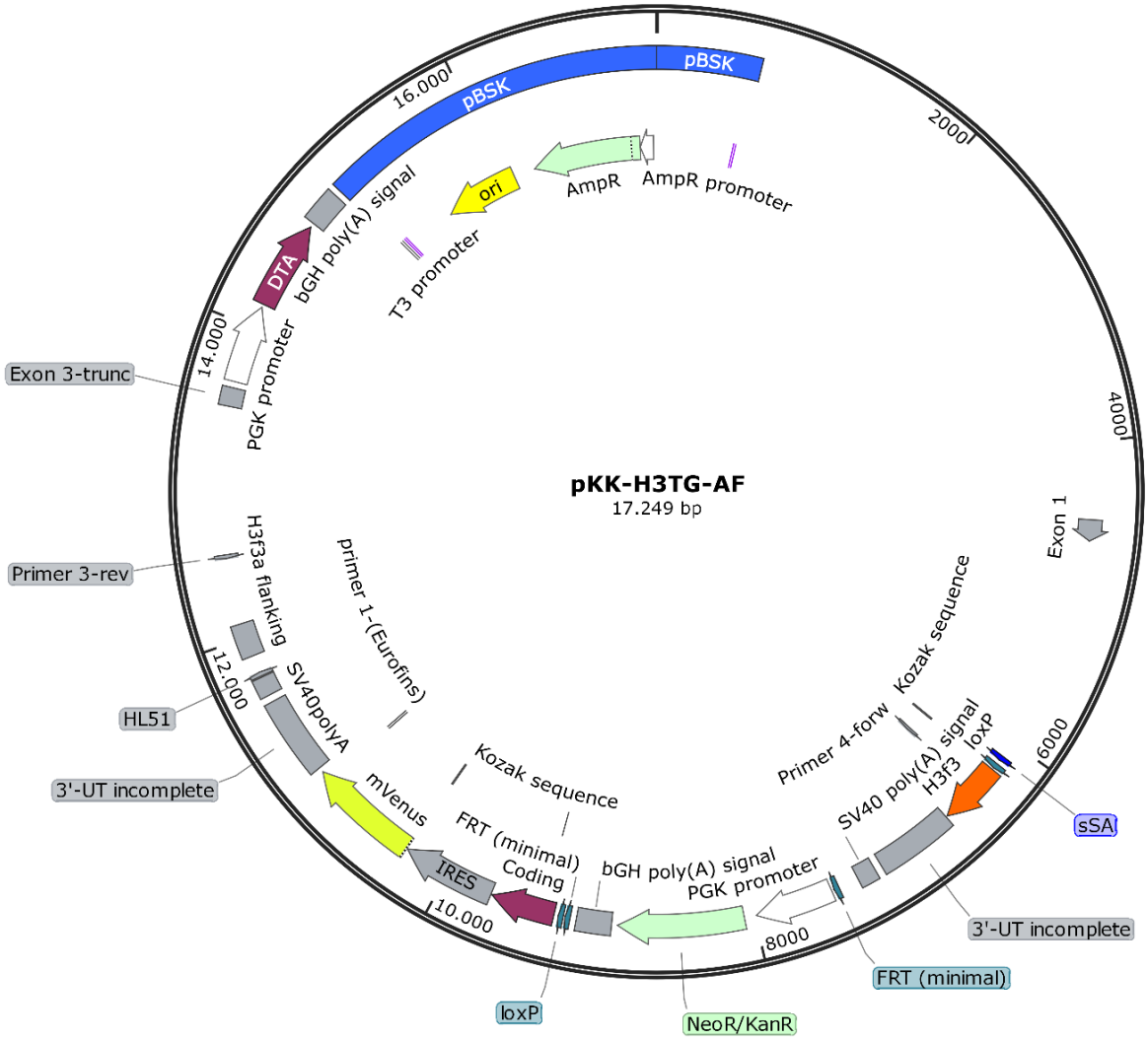


Figure 52A: Plasmid map of second targeting vector pKK-H3TG-AF

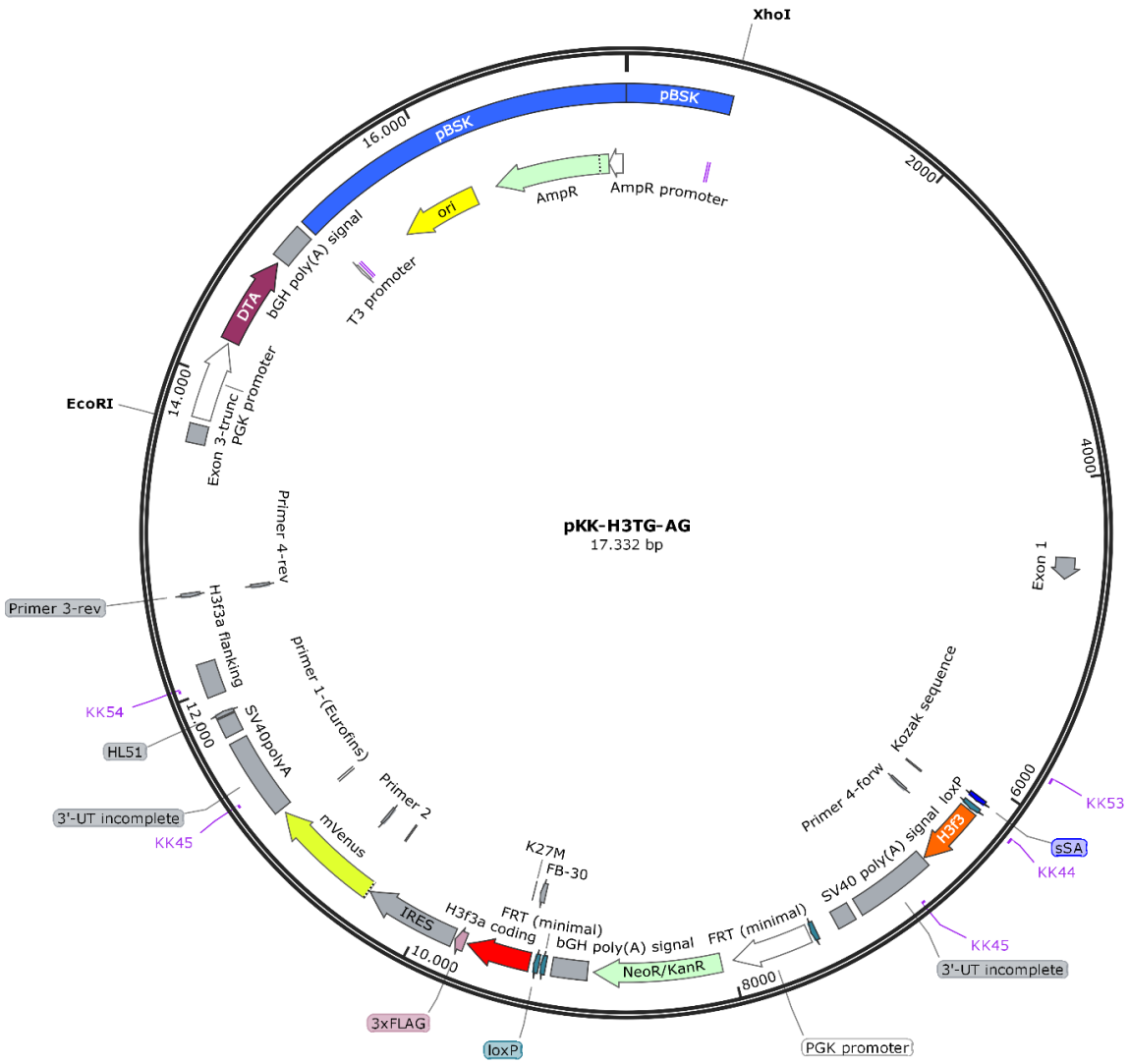


Figure 53A: Plasmid map of final targeting vector pKK-H3TG-AG

Phenotypic analysis of cell subsets of thymus and spleen

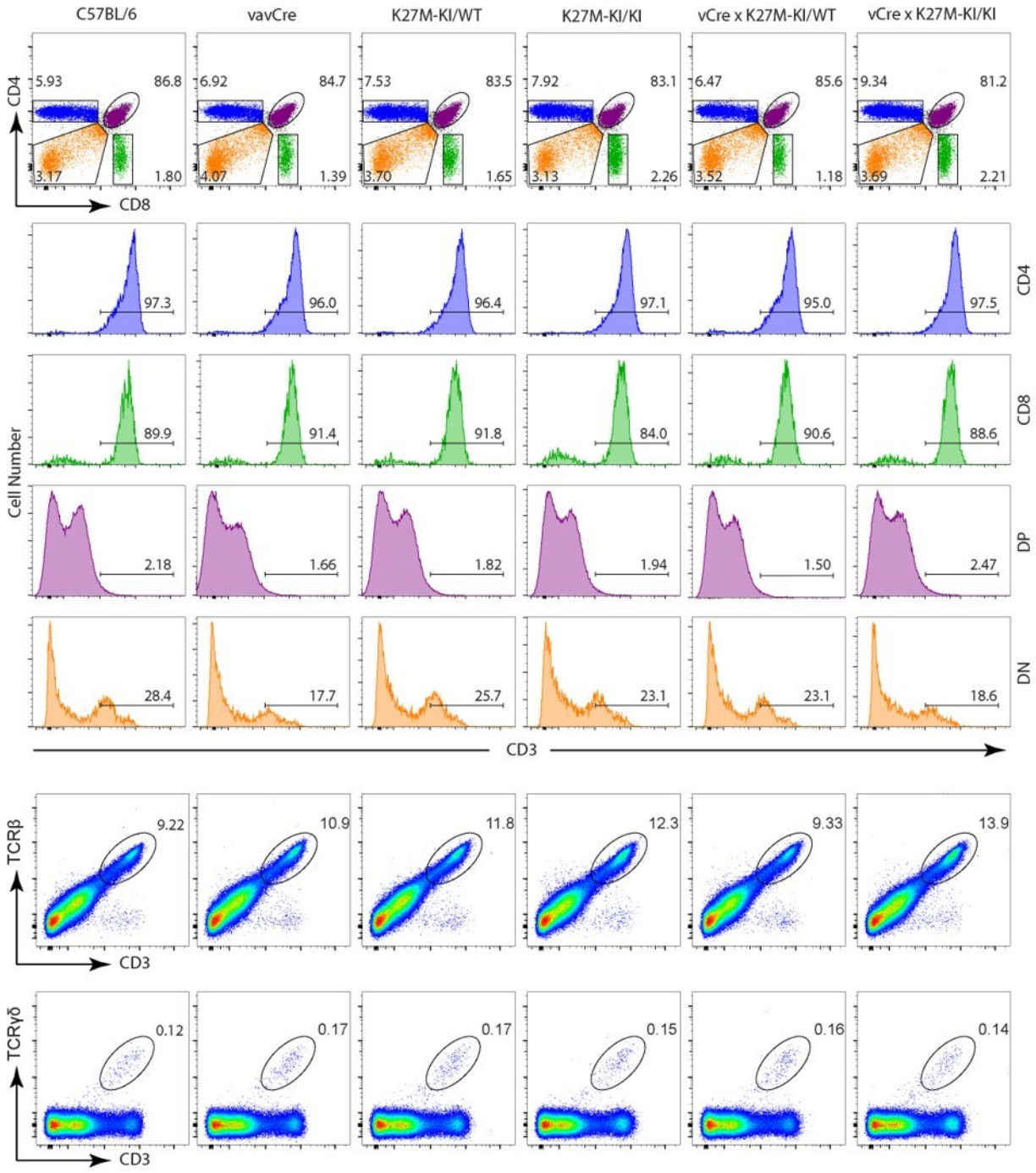


Figure 54A: Comparison of hematopoietic compartments in thymi of younger mice
FACS analysis of cell compartments in the thymi of younger mice. The top row shows gating for CD4/CD8 T-cells. The histograms below show the CD3 expression of the respective cell types. In blue are $CD4^{pos}$ cells, green are $CD8^{pos}$ cells, purple are double positive cells, and orange are double negative cells. The lowest two rows show the percentage of TCR-CD3 complexes.

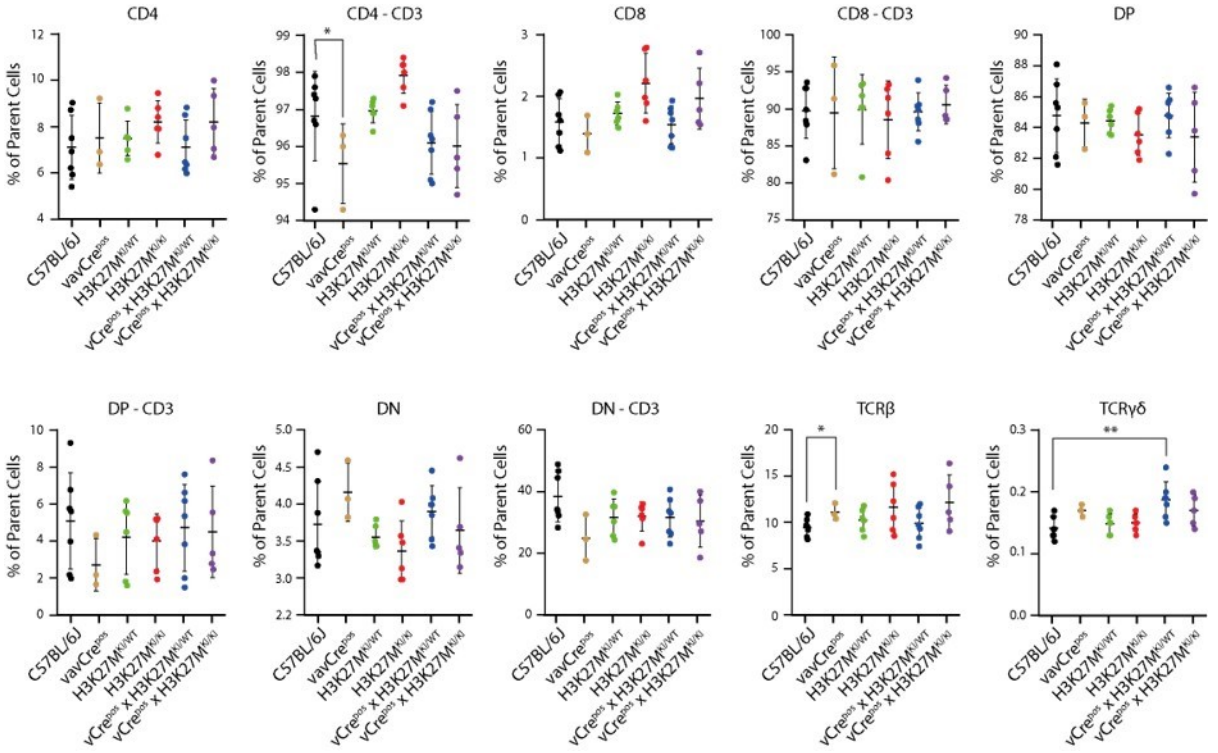


Figure 55A: Quantification of cell subsets in thymi of younger mice

Comparison of several thymic cell compartments of mice with various genotypes. Each dot represents one mouse. Shown are the mean and SD. The significance was determined using Mann-Whitney-U test. Only significant values are given.

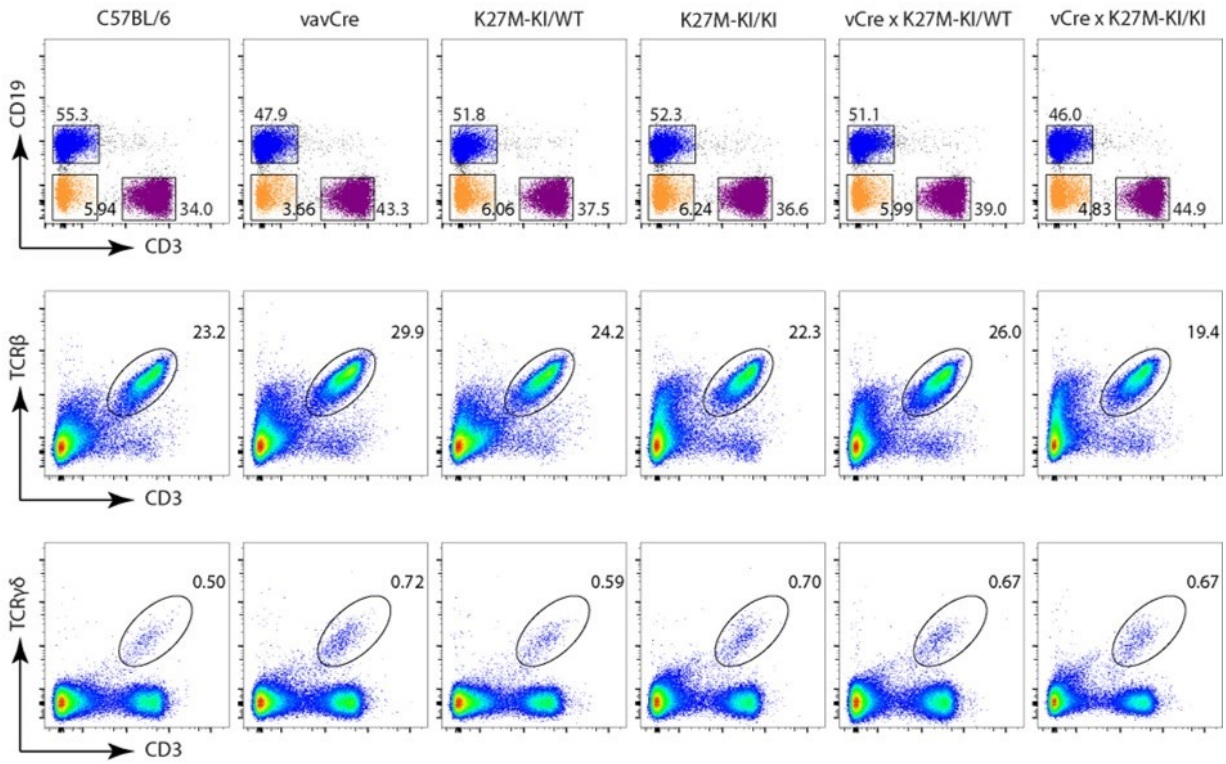


Figure 56A: Comparison of hematopoietic compartments in spleens of younger mice

FACS analysis of cell compartments in the spleen of mice. The top row shows gating for CD3/CD19 T-cells. Blue are $CD19^{pos}$ cells, purple are $CD3^{pos}$ cells, and orange are DN cells. The lowest two rows show the percentage of TCR-CD3 complexes.

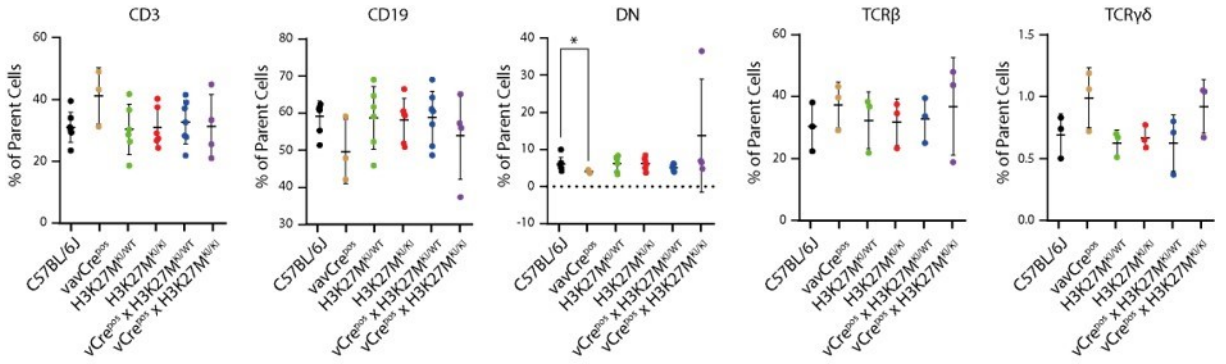


Figure 57A: Quantification of cell subsets in spleens of younger mice

Comparison of several splenic cell compartments of mice with various genotypes. Each dot represents one mouse. Shown are the mean and SD. The significance was determined using Mann-Whitney-U test. Only significant values are given.

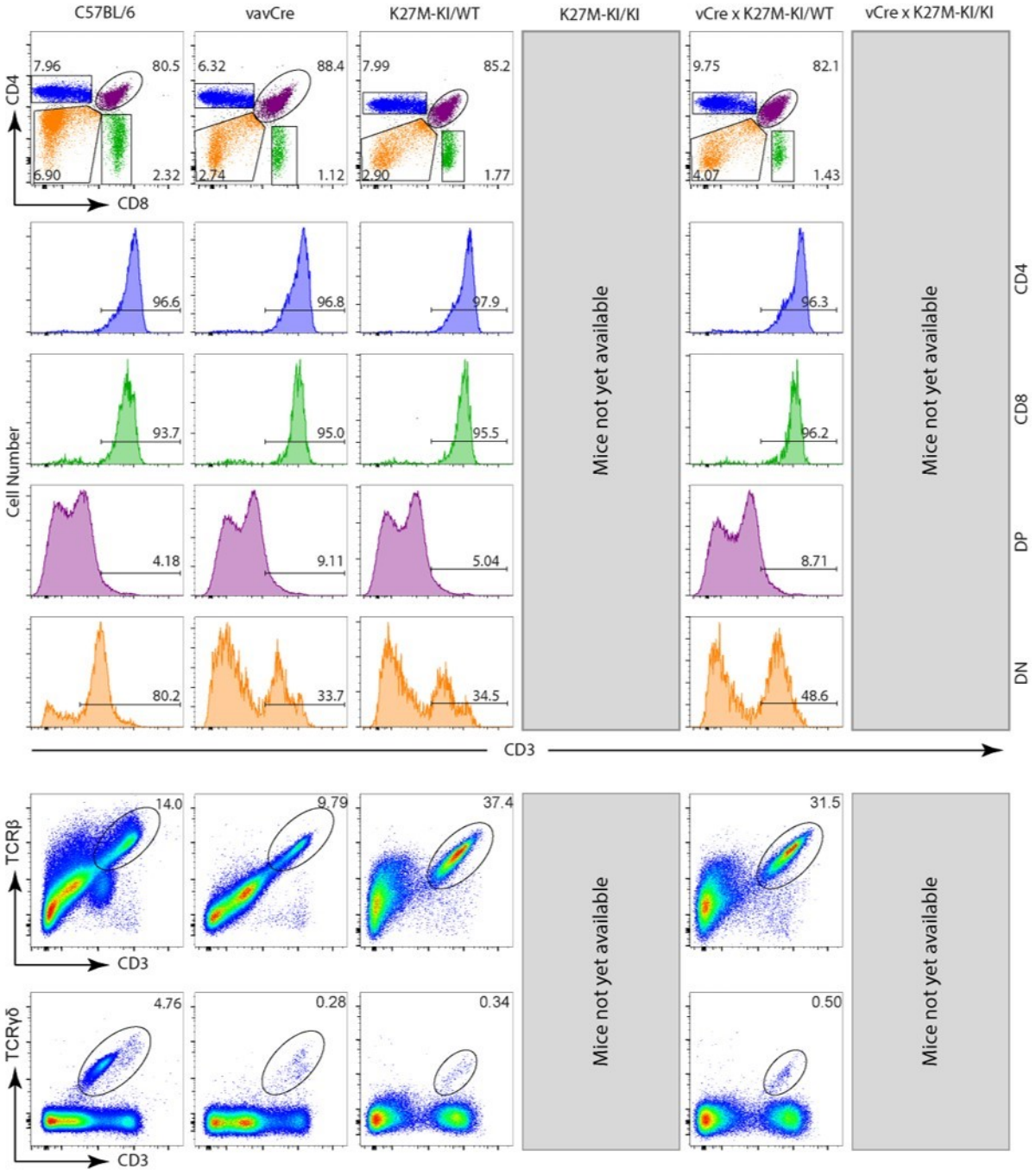


Figure 58A:Comparison of hematopoietic compartments in thymi of older mice

FACS analysis of cell compartments in the thymus of aged monitor mice. The top row shows gating for CD4/CD8 T-cells. The histograms below show the CD3 expression of the respective cell types. In blue are $CD4^{pos}$ cells, green are $CD8^{pos}$ cells, purple are double positive cells, and orange are double negative cells. The lowest two rows show the percentage of TCR-CD3 complexes. $H3K27M^{KI/KI}$ mice are not shown since they did not reach the age, yet.

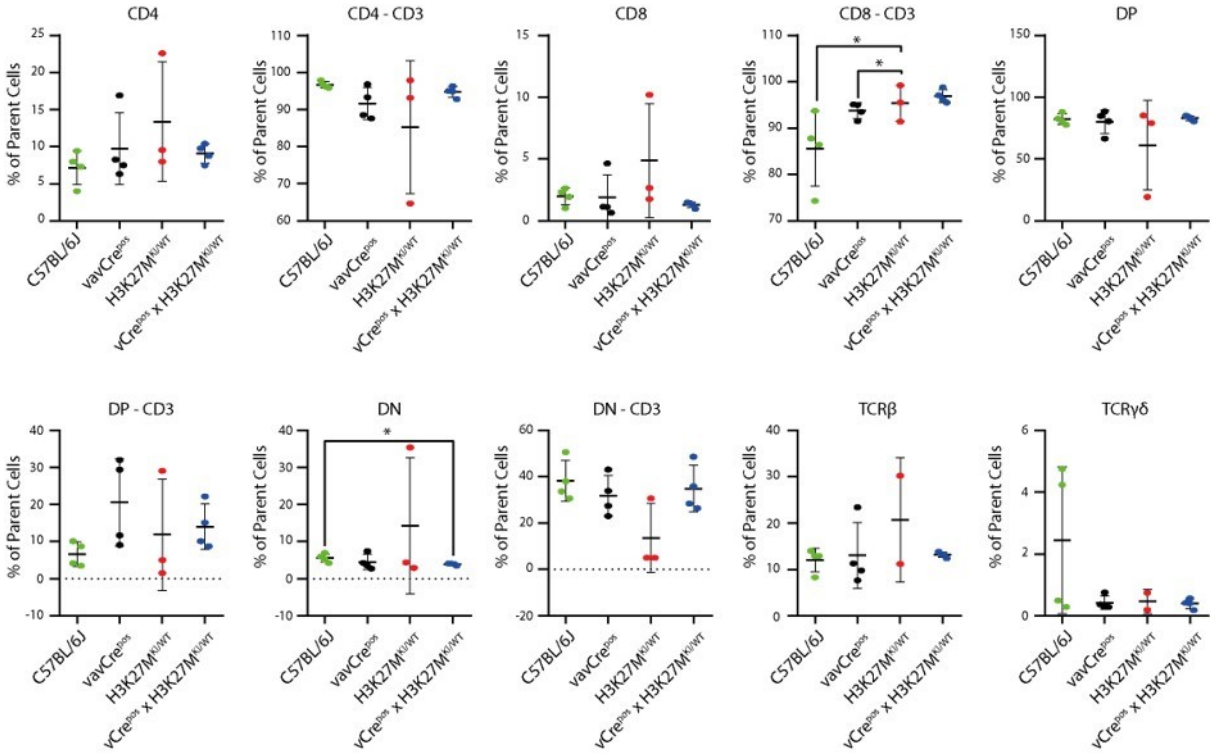


Figure 59A: Quantification of cell subsets in thymi of older mice

Comparison of several thymic cell compartments of older monitor mice with various genotypes. Each dot represents one mouse. Shown are the mean and SD. The significance was determined using Mann-Whitney-U test. Only significant values are given.

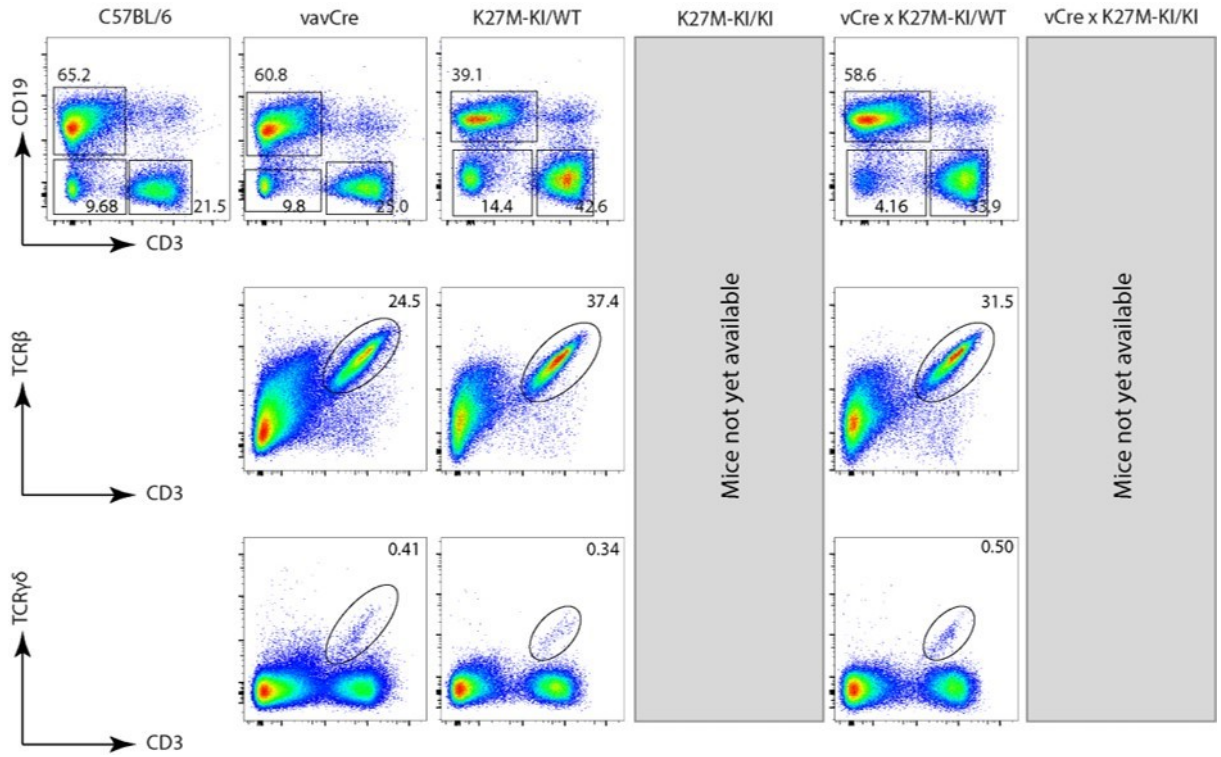


Figure 60A: Comparison of hematopoietic compartments in spleens of older mice
FACS analysis of cell compartments in the spleen of older monitor mice. The top row shows gating for CD3/CD19 T-cells. The second and third row show the percentage of TCR-CD3 complexes.

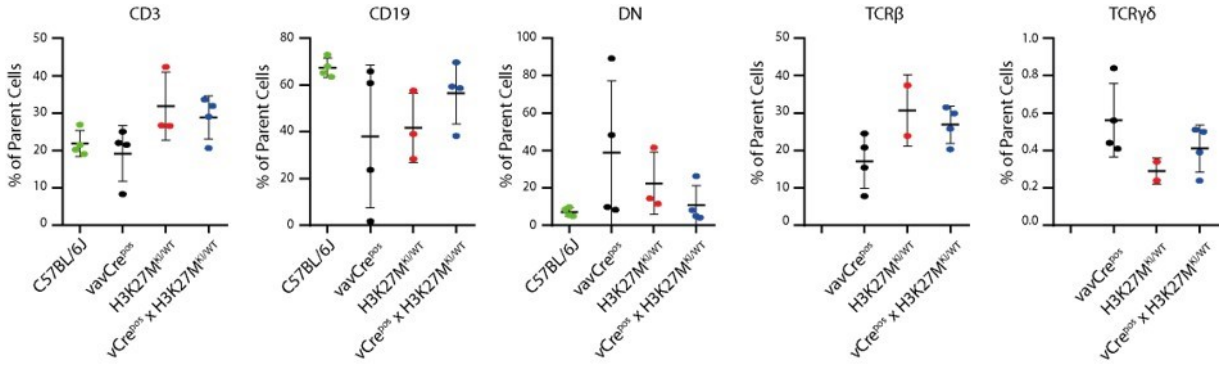


Figure 61A: Quantification of cell subsets in spleens of older mice
Comparison of several splenic cell compartments of older monitor mice with various genotypes. Each dot represents one mouse. Shown are the mean and SD. The significance was determined using Mann-Whitney-U test. Only significant values are given. Note: C57BL/6J mice are not included because to cohort was too small.

Individual weight gain of mice transplanted with IDH2-R172K transduced bone marrow

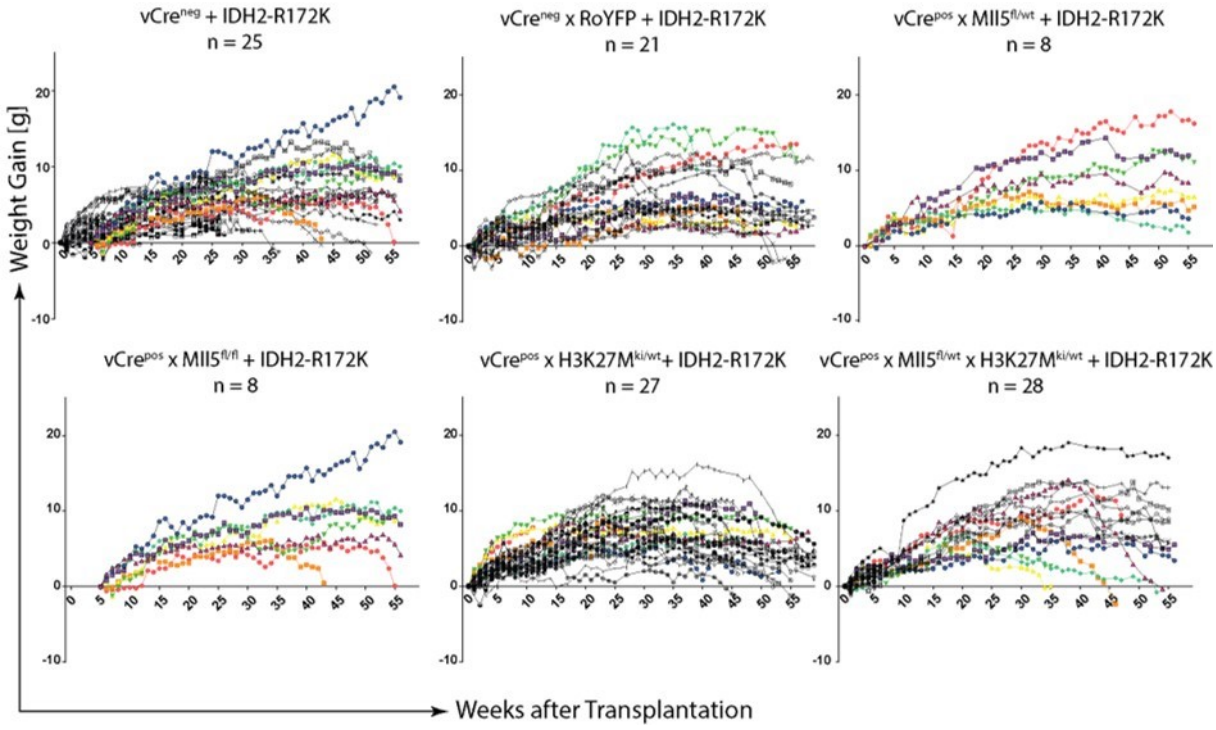


Figure 62A: Weight gain of individual mice transplanted with IDH2-R172K transduced bone marrow
Individual data sets used to generate the mean relative weight gain presented in Figure 41. Each line represents one mouse transplanted with the indicated genotype.

Survival of mice during *Dnmt3a*$^{flox/flox}$ breeding

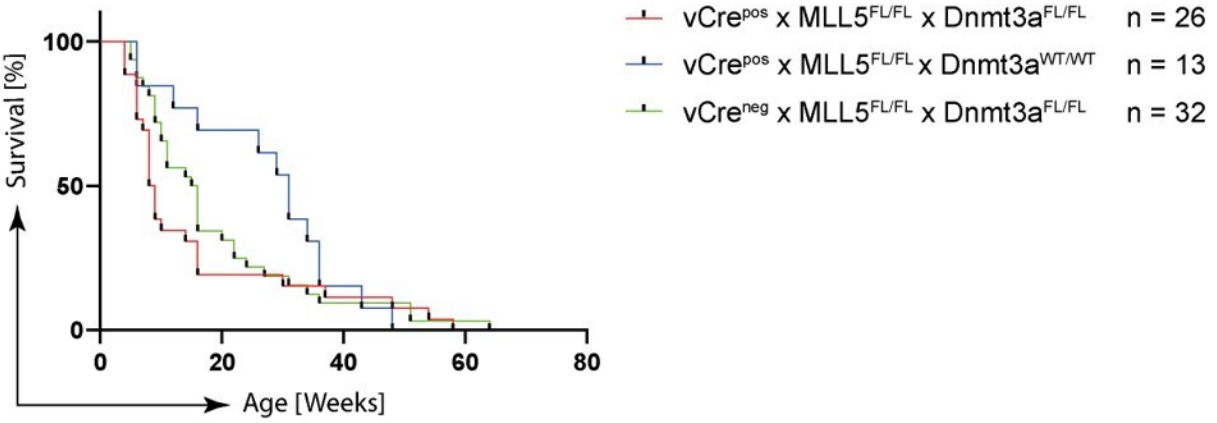


Figure 63A. Survival of mice generated during intercross of vCre x Mll5^{flox} mice with Dnmt3a^{flox} mice
vCre-Mll5flox/flox and Dnmt3aflox/flox mice were supposed to be bred to generate double deficient mice. Due to unknown reasons, barely any offspring were generated and mice of all genotypes died prematurely.

Acknowledgements

Acknowledgments removed for reasons of data protection.

Curriculum Vitae removed for reasons of data protection.