

Identification of critical effector proteins in MLL-rearranged Acute Myeloid Leukemia

Doctoral thesis at the Medical University of Vienna
for obtaining the academic degree

Doctor of Philosophy

Submitted by:

Anna Skucha MSc

Supervisors:

Prof. Dr. Florian Grebien

Ludwig Boltzmann Institute for Cancer Research
Währinger Straße 13A, 1090 Vienna, Austria

Institute for Medical Biochemistry University of Veterinary Medicine Vienna
Veterinärplatz 1, 1210 Vienna

Prof. Dr. Giulio Superti-Furga

CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences
Lazarettgasse 14, AKH BT 25.3, 1090 Vienna, Austria

Center for Physiology and Pharmacology Medical University of Vienna
Währinger Straße 13A/HP, 1090 Vienna, Austria

Vienna 5/2018

Declaration

The following doctoral thesis has been prepared in a cumulative format. Anna Skucha has been assigned the first authorship on manuscript #1 and is listed as a co-author on manuscript #2. The work summarized in this thesis has been performed in the research group of Prof. Dr. Giulio Superti-Furga at the CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences and in the laboratory of Prof. Dr. Florian Grebien at the Ludwig Boltzmann Institute for Cancer Research, both in Vienna, Austria. All parts of this thesis were written by the author with scientific feedback and input provided by Prof. Dr. Florian Grebien. Detailed individual contributions are listed in the prologue and interlude parts of the thesis. The publications included in this thesis are:

Manuscript #1

MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity.

Anna Skucha, Jessica Ebner, Johannes Schmöllerl, Mareike Roth, Thomas Eder, Adrián César-Razquin, Alexey Stukalov, Sarah Vittori, Matthias Muhar, Bin Lu, Martin Aichinger, Julian Jude, André C. Müller, Balázs Györfy, Christopher R. Vakoc, Peter Valent, Keiryn L. Bennett, Johannes Zuber*, Giulio Superti-Furga*, Florian Grebien[#]

* equal contribution

[#] lead contact: Florian Grebien

Manuscript #2

Pharmacological targeting of the Wdr5-MLL interactions in C/EBP α N-terminal leukemia.

Florian Grebien*, Masoud Vedadi*, Matthäus Getlik*, Roberto Giambruno, Amit Grover, Roberto Avellino, Anna Skucha, Sarah Vittori, Ekatarina Kuznetsova, David Smil, Dalia Barsyte-Lovejoy, Fengling Li, Gennadiy Poda, Matthieu Schapira, Hong Wu, Aiping Dong, Guillermo Senisterra, Alexey Stukalov, Kilian V. M. Huber, Andreas Schönegger, Richard Marcellus, Martin Bilban, Christoph Bock, Peter J. Brown, Johannes Zuber, Keiryn L. Bennett, Rima Al-awar, Ruud Delwel, Claus Nerlov, Cheryl H. Arrowsmith*, Giulio Superti-Furga*

* equal contribution

[#] lead contact: Giulio Superti-Furga and Cheryl H. Arrowsmith

Table of Contents

Declaration.....	iii
Table of Contents.....	iv
List of figures and tables.....	vi
Abstract.....	vii
Zusammenfassung.....	ix
Abbreviations.....	xi
Acknowledgements.....	xiv
1. Introduction.....	1
1.1 Characteristics of Cancer Cells.....	1
1.2 Acute Myeloid Leukemia (AML).....	2
1.3 Multi-Partner Translocation (MPT) families in Acute Myeloid Leukemia.....	7
1.4 MLL rearranged leukemia.....	9
1.5 Structure of the MLL protein.....	10
1.6 Composition of the MLL protein complex.....	13
1.7 Translocations of the <i>MLL</i> gene in leukemia.....	15
1.8 Clinical consequences of MLL-fusion proteins expression in leukemia.....	16
1.9 Model systems establishing MLL-fusions as driver mutations.....	19
1.10 Transcriptional targets of MLL-fusion proteins.....	20
1.11 Protein complexes around MLL-fusion proteins and mechanisms of MLL-fusion-driven leukemic transformation.....	22
1.12 Signaling pathways involved in MLL-fusion-mediated leukemogenesis.....	25
1.13 Therapeutic strategies in MLL-rearranged leukemia.....	27
1.14 Experimental tools for functional characterization of MLL-fusions interactors.....	30
1.14.1 Biochemical characterization.....	30
1.14.2 Genetic characterization.....	33
1.15 Aims of this thesis.....	36
2. Results.....	37
2.1 Prologue.....	37
2.2 MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity.....	39
2.3 Interlude.....	67
2.4 Pharmacological targeting of the Wdr5-MLL interactions in C/EBP α N-terminal leukemia.....	69

3. Discussion	101
3.1 The role of chromosomal translocations in leukemia	101
3.2 Functional proteomics identifies SETD2 as a critical effector of MLL-fusion proteins ..	102
3.3 The function of SETD2-mediated deposition of H3K36me3	103
3.4 Mutations of SETD2 in cancer	107
3.5 Mechanisms of action of SETD2 in cancer	108
3.6 Potential therapeutic opportunities	111
3.7 Interaction proteomics identifies Wdr5 as an interactor of the p30 isoform of C/EBP α in AML.....	112
3.8 Conclusions and future prospects	112
4. References.....	115
5. Annex.....	143
5.1 Curriculum Vitae	143
5.2 Scientific Publications	147

List of figures and tables

Figure 1. Comparison between normal and leukemic hematopoiesis

Figure 2. Mechanisms triggering the expansion of acute myeloid leukemia

Figure 3. Multi-Partner Translocation (MPT) families

Figure 4. MLL rearrangements associated with human acute leukemias

Figure 5. Detailed structure of the wild type MLL protein

Figure 6. The macromolecular complex around wild type MLL

Figure 7. Classification of patients with MLL-rearranged leukemia

Figure 8. Proposed molecular pathways responsible for development of MLL-rearranged leukemia

Figure 9. Therapeutic targeting of MLL-fusion proteins

Table 1. Current genomic classification of Acute Myeloid Leukemia

Abstract

Acute Myeloid Leukemia (AML) is driven by a number of recurrent mutations, which result in irrepressible growth and block of differentiation of leukemic cells. As the mutated proteins are often embedded in large macromolecular complexes we hypothesized that effector proteins empowering leukemogenesis might be encoded in stable physical and genetic interaction networks surrounding the initial oncoprotein. In this thesis, I aimed to identify common critical effectors of a subset of distinct MLL-fusion proteins. Moreover, we investigated the functional contribution of interactors of mutant C/EBP α proteins to AML development.

The Mixed Lineage Leukemia gene (MLL) is a frequent target of chromosomal rearrangements in human hematopoietic malignancies. Balanced translocations result in the fusion of the *MLL* gene to over 75 different fusion partner genes, leading to the production of novel chimeric proteins. Critical effectors of distinct MLL-fusion proteins have previously been identified, and some of them were shown to hold great potential for targeted therapies. However, it is not clear whether these effectors are conserved among all MLL-fusion proteins or if different molecular mechanisms of transformation exist for distinct MLL-fusion proteins. Characterization of the protein complexes nucleated by 7 MLL-fusion proteins (MLL-AF1p, MLL-AF4, MLL-AF9, MLL-CBP, MLL-EEN, MLL-ENL, MLL-GAS7) by affinity purification coupled to mass spectrometry (AP-MS) revealed a densely interconnected protein-protein interaction network. 128 proteins were found to interact with ≥ 5 of all 7 MLL-fusions. Systematic functional investigation of the conserved MLL-fusion interactome, using subtractive shRNA screens, identified the methyltransferase SETD2 as a critical effector of MLL-fusion proteins. Functional characterization of the role of SETD2 in MLL-rearranged leukemia through loss of function experiments *in vitro* and *in vivo* established a novel role for SETD2 in the maintenance of genomic integrity during initiation and progression of MLL-rearranged AML.

N-terminal frameshift mutations in the transcription factor C/EBP α , a master regulator of myeloid gene expression, have been found in 9% of patients presenting with AML, highlighting the role of the short isoform C/EBP α p30 in disease development. We identified the Wdr5/MLL containing complex as an interactor of the C/EBP α p30 variant. Wdr5 was responsible for the differentiation block and leukemia development of C/EBP α -mutated cells. Furthermore, we have characterized the new small molecule compound OICR-9429 as a potent Wdr5 antagonist. OICR-9429 disrupts the interaction between Wdr5 and MLL, thus restoring myeloid differentiation potential of C/EBP α p30-expressing cells.

In summary, our data highlight the functional relevance of combined proteomic-genomic cellular screening to identify critical effectors of genes involved in the development of acute myeloid leukemia in a comprehensive and comparative manner. Our studies contribute to further clarification of the molecular mechanisms driving leukemogenesis.

Zusammenfassung

Die Akute Myeloische Leukämie (AML) resultiert aus einer Vielzahl von wiederkehrenden Mutationen, welche zu einer Differenzierungsblockade und zu unkontrollierter Proliferation von leukämischen Zellen führen. Da die in der AML mutierten Onkoproteine oft in großen Proteinkomplexen vorliegen, stellen wir die Hypothese auf, dass wichtige Effektor-Proteine in der AML in stabilen physischen und genetischen Interaktionsnetzwerken lokalisiert sein könnten, welche das ursprüngliche Onkoprotein einschließen. Ein Ziel dieser Dissertation war es, neue Effektor-Proteine von Onkoproteinen aus der Gruppe der MLL-Fusionen zu finden. Des weiteren wurde der funktionelle Beitrag von Interaktionspartnern von mutiertem Versionen des C/EBP α Transkriptionsfaktor zur AML-Entwicklung untersucht.

Das MLL-Gen (*Mixed Lineage Leukemia*) ist in humanen hämatopoetischen Erkrankungen häufig in chromosomalen Translokationen involviert. Diese Translokationen führen zu Fusionen des MLL-Gens mit über 75 verschiedenen Fusionspartnern, was die Entstehung von neuen chimären Proteinen zur Folge hat. Einige bereits identifizierte, kritische Effektorproteine von bestimmten MLL-Fusionsproteinen zeigen ein großes Potential für zielgerichtete Therapien. Es ist jedoch nicht bekannt, ob diese Effektorproteine für die Gesamtheit aller MLL-Fusionsproteinen von Relevanz sind, oder ob für bestimmte MLL-Fusionsproteine verschiedene molekulare Mechanismen der Transformation existieren. Analyse der Interaktionspartner von sieben ausgewählten MLL-Fusionsproteinen (MLL-AF1p, MLL-AF4, MLL-AF9, MLL-CBP, MLL-EEN, MLL-ENL, MLL-GAS7) durch Affinitätschromatographie und Massenspektrometrie (AP-MS) resultierte in einem komplexen Protein-Protein Interaktionsnetzwerk mit >9650 Partnern. Es wurden 128 Proteine identifiziert, die mit mehr als fünf der sieben MLL-Fusionsproteine interagierten. Eine systematische Analyse des konservierten MLL-Fusions-Interaktoms wurde mittels subtraktiven shRNA Screens durchgeführt. Dies führte zur Identifizierung der Methyltransferase SETD2 (SET Domain Containing 2) als kritischer Effektor von MLL-Fusionsproteinen.

Funktionelle Charakterisierung von SETD2 in MLL-translokierter Leukämie durch Loss-of-Function Experimente *in vitro* und *in vivo* zeigte eine neue Funktion von SETD2 in der Erhaltung der genomischen Integrität während der Leukämie-Initiation und -Progression auf.

Der Transkriptionsfaktor C/EBP α ist ein wichtiger Regulator der Expression von myeloiden Genen. N-terminale Frameshift-Mutationen in C/EBP α wurden in 9% aller AML-Patienten gefunden. Diese Mutationen heben die Rolle einer verkürzten Form von C/EBP α (p30) in der Leukämie-Entstehung hervor. Wir konnten die Interaktion von C/EBP α p30 mit dem WDR5/MLL

Komplex mittels AP-MS nachweisen. Weiters konnte wir zeigen, dass Wdr5 für die Differenzierungsblockade und Leukämieentstehung in C/EBP α mutierten Zellen verantwortlich ist. Zusätzlich testeten wir die neue niedermolekulare Substanz OICR-9429, welche als potenter Wdr5-Antagonist wirken kann. OICR-9429-vermittelte Störung der Interaktion zwischen Wdr5 und MLL, konnte die Differenzierungsblockade von C/EBP α p30-exprimierenden Zellen aufheben. Zusammenfassend unterstreichen unsere Daten die funktionelle Relevanz von kombinierten zellulären Proteomik- und Genomik-Screens zur übergreifenden und vergleichenden Identifizierung von kritischen Effektorproteinen in der AML. Somit tragen unsere Studien zu einer weiteren Aufklärung von molekularen Mechanismen der Leukämieentstehung bei.

Abbreviations

ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
AP-MS	affinity purification coupled to mass spectrometry
APL	acute promyelocytic leukemia
ASB2	ankyrin repeat and SOCS box protein 2
ASXL1	additional sex combs-like protein 1
BCR	breakpoint cluster region
ABL1	abelson murine leukemia viral oncogene homolog 1
Bp	base pairs
BRD4	bromodomain containing protein 4
Cas9	CRISPR-associated protein 9
CBP	CREB binding protein
ccRCC	clear cell renal cell carcinoma
CDKIs	cyclins and cyclin dependent kinase inhibitors
C/EBP α	CCAAT/enhancer binding protein alpha
ChIP-seq	chromatin immunoprecipitation followed by sequencing
CLL	chronic lymphoblastic leukemia
CML	chronic myeloid leukemia
CRISPR	clustered, regularly interspaced, short palindromic repeats
CTD	carboxy-terminal domain
DNA	deoxyribonucleic acid
DNMT3A	DNA (cytosine-5)-methyltransferase 3A
dNTP	deoxyribonucleotide triphosphate
dsRNA	double strand Ribonucleic acid
EAF1/2	ELL associated factor 1/2
ECS(ASB)E3	elonginBC-Cullin-SOCS box ankyrin repeat and SOCS box E3 ligase
EFS	event free survival
ETP-ALL	early T-cell precursor acute lymphoblastic leukemia
EVI-1	ecotropic viral integration site-1
EZH2	enhancer of zeste homolog 2
FACS	fluorescence-activated cell sorting
FISH	fluorescence in situ hybridization
FLT3	Fms-like tyrosine kinase 3
FOXO	forkhead box protein family
Frat1/2	frequently rearranged in advanced T-cell lymphomas 1
GAS7	growth arrest-specific protein 7
GFP	green fluorescent protein
GMP	granulocyte-macrophage progenitors
GRO-seq	global run-on sequencing
GSK3	glycogen synthase kinase 3
HA-tag	human influenza hemagglutinin tag
HCF-1	host cell factor1
HDAC1/2	histone deacetylase 1/2
HLPC	high performance liquid chromatography
HOX	homeobox protein hox cluster
HR	homologous recombination
HSCs	hematopoietic stem cells
HTRX	human trithorax-like protein

IDH1/2	isocitrate dehydrogenase NADP 1/2
IL1	interleukin-1
IMAC	ion metal chromatography
IRAKs	IL1 receptor associated kinases
ITD	internal tandem duplication
ITGB3	integrin beta 3
iTRAQ	isobaric tags for relative and absolute quantitation
JMJD1C	jumonji domain-containing protein 1C
Kb	kilo bases
KMT2A	lysine N-methyltransferase 2A
KRAS	kirsten rat sarcoma viral oncogene homolog
LC-MS/MS	liquid chromatography – tandem mass spectrometry
LEDGF	lens epithelium derived growth factor
LICs	leukemia initiating cells
MBD	methyl-CpG-binding domain
MEN1	multiple endocrine neoplasia I
MGR15	MORF-Related Gene 15 Protein
miRNA	micro RNA
MLL	mixed lineage leukemia
MMR	mismatch repair
MOF	histone acetyltransferase KAT8
MOZ	histone acetyltransferase KAT6A
MPT	multi-partner translocation family
MSH6	MutS Protein Homolog 6
mTOR	mammalian target of rapamycin
MudPIT	multidimensional protein identification technology
NGS	next generation sequencing
NHEJ	non-homologous end joining
NOD/SCID	non-obese diabetic severe combined immunodeficient
NPM1	nucleophosmin 1
NRAS	neuroblastoma RAS viral oncogene homolog
NUP98	nucleoporin 98
P-TEFb	positive transcription elongation factor b
PI3K	phosphoinositide 3-kinase
PAFc	polymerase associated factor complex
PAM	protospacer adjacent motive
PBX3	pre-B-cell leukemia transcription factor 3
PHD	plant homology domains
piRNA	piwi-interacting RNAs
RNA Pol II	RNA polymerase II
PRMT1	Protein arginine N-methyltransferase 1
Pro-seq	Precision nuclear run-on sequencing
PTD	partial tandem duplication
PTEN	phosphatase and tensin homolog
PTMs	post translational modifications
Rac 1	Ras-related C3 botulinum toxin substrate 1
RAS-RAF	rat sarcoma virus protein – RAF proto oncogene serine/t
RISC	RNA-induced silencing complex
RNA Pol III	RNA polymerase III
RNA-seq	RNA sequencing
RNAi	RNA interference

RNF20	RING finger protein 20
RPA	replication protein A
RRM2	Ribonucleoside-diphosphate reductase subunit M2
RUNX1	runt related transcription factor 1
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-methionine
SEC	super elongation complex
SET	suppressor of variegation, enhancer of zeste and thithorax
SETD2	SET domain-containing protein 2
sgRNA	single guide RNA
shRNA	short hairpin RNA
SILAC	stable isotope labeling with amino acids
siRNA	short interfering RNA
SIRT1	sirtuin1
Sterp-tag	streptavidin tag
SWI/SNF	SWItch/Sucrose Non-Fermentable complex
TAP	tandem affinity purifications
TERT	telomerase reverse transcriptase
TKD	tyrosine kinase domain
TMT	tandem mass tags
TRD	transcription repression domain
TrxG	trithorax-group of proteins
TT-seq	transient transcriptome sequencing
WDR5	WD repeat-containing protein 5
RbBP5	retinoblastoma-binding protein 5
ASH2L	Set1/Ash2 histone methyltransferase complex subunit ASH2
WHO	world health organization
WRAD	protein complex of Wdr5-RbBP5-ASH2L-DPY30
Wt	wild type
ZMYND11	zinc finger MYND domain-containing protein 11
ZNF521	zinc finger protein 521

Acknowledgements

Over the course of my PhD studies, I have met a lot of people whose support was essential for the success of my project and thus I would like to acknowledge.

First and foremost, I would like to thank **Prof. Dr. Florian Grebien** for the chance of working on this exciting project, for his support and always fair supervision. I have really enjoyed being part of his group, truly appreciated his help and all chances I was given.

I would like to thank **Prof. Dr. Giulio Superti-Furga** for excellent mentorship throughout this journey and for providing me with a great work environment. It was a terrific adventure and I believe that his attitude makes CeMM a very special place to work in. I will always be proud of being a CeMMie.

Moreover, I would like to thank my PhD committee: Dr. Johannes Zuber – for critical comments throughout my PhD studies but also for hosting me in his laboratory. The knowledge I gained during my stay at IMP allowed me to conduct my research project independently afterwards. Dr. Christoph Bock – for his critical advises on my project.

I would also like to acknowledge all the members of Grebien's, Superti-Furga's and Zuber's laboratories. Not only they helped me with my research but they were all exceptionally kind and friendly.

Last but not least, I would like to thank my family, my boyfriend and my friends who love and support me, regardless of my academic title.

1. Introduction

1.1 Characteristics of Cancer Cells

Cancer represents a leading cause of human death worldwide (Torre, Bray et al., 2015). While evermore research is conducted to elucidate mechanisms underlying the initiation and progression of various types of cancer, a detailed understanding of the neoplastic disease remains one of the biggest challenges of biomedical and life sciences. Regardless of the type and tissue of origin, every cancer cell can be characterized by several capabilities, which it acquires during its lifetime. These “Hallmarks of Cancer” initially included the following features: ability to resist cell death, sustained proliferative signaling, averting the effect of growth suppression, induction of angiogenesis, replicative immortality and activated invasion and metastasis (Hanahan & Weinberg, 2000). This list was extended by the ability of cancer cells to reprogram their energy metabolism and to circumvent recognition and destruction by the immune system (Hanahan & Weinberg, 2000). These hallmarks cooperate, allowing the tumor cells to establish a firm microenvironment that facilitates expansion of the neoplasm.

The initial reference indicating that a cell can pass a genetic mutation onto its progeny over the course of cell divisions, later called the “somatic mutation theory of cancer” dates back to over 100 years ago (Calkins, 1914). Together with the increasing evidence that the mutational status of a cancer cell plays an important role in disease development, the scientific world became interested in sequencing the human genome in the late 80’s of the 20th century (Dulbecco, 1986, Lewin, 1987, Palca, 1988). Currently, widely available and affordable access to large-scale sequencing experiments, including single-cell technologies, has empowered base-pair resolution analysis of cancer genomes. This contributes additional knowledge about cancer heterogeneity and complexity, which is expected to significantly improve the management of cancer patients. On average, each tumor requires 2-8 “driver” mutations, which impose a selective growth advantage on to cancer cells in a direct or indirect way. Driver mutations are accompanied by a much larger number of “passenger mutations”, which alone may not have any direct effect on the cancer’s growth advantage and therefore might not play any roles in oncogenesis (Vogelstein, Papadopoulos et al., 2013). The exact number of mutations per tumor depends on the location and type of the neoplasm as well as on additional impact of environmental carcinogens and varies between solid and liquid cancer. For instance, the prevalence of mutations in malignant melanoma or lung cancer at frequencies above 100 mutations per mega base of DNA is significantly higher than in pediatric cancers, which often present with a frequency of 0.1

mutations per mega base (Alexandrov, Nik-Zainal et al., 2013, The Cancer Genome Atlas Research Network, 2015, Hayward, Wilmott et al., 2017, Lawrence, Stojanov et al., 2013, Vogelstein et al., 2013).

1.2 Acute Myeloid Leukemia (AML)

Acute myeloid leukemia (AML) is the most common type of acute leukemia in adults. It represents a cancer affecting the white blood cell lineage, and is often associated with poor prognosis. The disease affects both pediatric and adult patients. There have been significant improvements in survival rates for most of cancer types due to earlier detection and advances in treatment. Also in AML, the overall five-year relative survival rates have risen significantly within the last 50 years, from 14% in 1960 to 62.7% in 2012. The number of newly diagnosed AML cases and AML-related deaths in the United States in 2017 were estimated to reach >20,000 and >4,500, respectively (American Cancer Society, 2017).

Many of the mutations contributing to the development of acute myeloid leukemia were initially discovered by examination of chromosomes of AML cells. Advances in cytogenetic analysis allowed for efficient investigation of the pathogenesis of AML since the 70's of the 20th century. Since then other methods have been developed. For instance, conventional karyotyping allows for microscopic studies of fixed metaphase chromosomes. Fluorescence in situ hybridization (FISH) is usually utilized to determine specific localization of genes on chromosomes using denatured probes. The probe can be later detected by its fluorescence under UV light. For instance, a fusion of the *PML* gene on chromosome 15 with the *RARA* gene on chromosome 17 was first detected in cytogenetic studies of AML samples by FISH. This method was also used to classify patients' risk profiles to design adequate treatment plans. However, it was observed that over 50% of AML patients presented with a normal karyotype, which was correlated with an intermediate risk of relapse. Furthermore, advanced methods of molecular diagnostics have been established and helped to identify genes that are involved in the development of hematopoietic malignances. DNA or RNA-based methodologies, such as polymerase chain reaction (PCR), DNA microarrays and DNA sequencing, allow the identification of previously annotated mutations. The presence of unknown and/or novel mutations can be investigated by single strand conformational polymorphism (SSCP) or denaturing gradient gel electrophoresis (DGGE). In recent years, high throughput next generation sequencing (NGS) methods have emerged as the most powerful tools to identify mutational abnormalities in the human genome. Applications of whole-genome sequencing (WGS) whole exome sequencing (WES), transcriptomics (RNA-seq)

and variations of these approaches contributed valuable knowledge in cancer genomics (Cotton, 1993).

Table 1. Proposed Genomic Classification of Acute Myeloid Leukemia (AML).		
Genomic Subgroup	Frequency in the Study Cohort (N = 1540)	Most Frequently Mutated Genes*
	<i>no. of patients (%)</i>	<i>gene (%)</i>
AML with <i>NPM1</i> mutation	418 (27)	<i>NPM1</i> (100), <i>DNMT3A</i> (54), <i>FLT3</i> ^{ITD} (39), <i>NRAS</i> (19), <i>TET2</i> (16), <i>PTPN11</i> (15)
AML with mutated chromatin, RNA-splicing genes, or both†	275 (18)	<i>RUNX1</i> (39), <i>MLL</i> ^{PTD} (25), <i>SRSF2</i> (22), <i>DNMT3A</i> (20), <i>ASXL1</i> (17), <i>STAG2</i> (16), <i>NRAS</i> (16), <i>TET2</i> (15), <i>FLT3</i> ^{ITD} (15)
AML with <i>TP53</i> mutations, chromosomal aneuploidy, or both‡	199 (13)	Complex karyotype (68), -5/5q (47), -7/7q (44), <i>TP53</i> (44), -17/17p (31), -12/12p (17), +8/8q (16)
AML with inv(16)(p13.1;q22) or t(16;16)(p13.1;q22); <i>CBFB-MYH11</i>	81 (5)	<i>inv(16)</i> (100), <i>NRAS</i> (53), +8/8q (16), +22 (16), <i>KIT</i> (15), <i>FLT3</i> ^{TKD} (15)
AML with biallelic <i>CEBPA</i> mutations	66 (4)	<i>CEBPA</i> ^{biallelic} (100), <i>NRAS</i> (30), <i>WT1</i> (21), <i>GATA2</i> (20)
AML with t(15;17)(q22;q12); <i>PML-RARA</i>	60 (4)	<i>t(15;17)</i> (100), <i>FLT3</i> ^{ITD} (35), <i>WT1</i> (17)
AML with t(8;21)(q22;q22); <i>RUNX1-RUNX1T1</i>	60 (4)	<i>t(8;21)</i> (100), <i>KIT</i> (38), -Y (33), -9q (18)
AML with <i>MLL</i> fusion genes; t(x;11)(x;q23)§	44 (3)	<i>t(x;11q23)</i> (100), <i>NRAS</i> (23)
AML with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); <i>GATA2</i> , <i>MECOM(EVI1)</i>	20 (1)	<i>inv(3)</i> (100), -7 (85), <i>KRAS</i> (30), <i>NRAS</i> (30), <i>PTPN11</i> (30), <i>ETV6</i> (15), <i>PHF6</i> (15), <i>SF3B1</i> (15)
AML with <i>IDH2</i> ^{R172} mutations and no other class-defining lesions	18 (1)	<i>IDH2</i> ^{R172} (100), <i>DNMT3A</i> (67), +8/8q (17)
AML with t(6;9)(p23;q34); <i>DEK-NUP214</i>	15 (1)	<i>t(6;9)</i> (100), <i>FLT3</i> ^{ITD} (80), <i>KRAS</i> (20)
AML with driver mutations but no detected class-defining lesions	166 (11)	<i>FLT3</i> ^{ITD} (39), <i>DNMT3A</i> (16)
AML with no detected driver mutations	62 (4)	
AML meeting criteria for ≥2 genomic subgroups	56 (4)	

Table 1. Current genomic classification of Acute Myeloid Leukemia. Classification of AML based on mutations identified in 1540 AML patients (Reproduced with permission from Papaemmanuil et. al, 2016, Copyright Massachusetts Medical Society).

A first comprehensive mutational landscape of AML was released in 2013, when the Cancer Genome Atlas Research Network analyzed the genomes of 200 de novo AML patients by whole exome sequencing. Among many, mutations in genes such as *FLT3*, *NPM1*, *DNMT3A*, *NRAS* and *TET2* were recurrently found. Moreover, it became apparent that some patients carried more than one driver mutation (Ley, Miller et al., 2013). Overall, AML was described to have significantly less mutations than other cancers. The list of mutations described in 2013 was extended in 2016, resulting in the most recent version of the driver mutation landscape of acute myeloid leukemia (Papaemmanuil, Gerstung et al., 2016). 5234 driver mutations involving 76 genes were identified in 1540 samples from AML patients using a targeted sequencing approach. In line with previous findings, the *FLT3*, *NPM1*, *DNMT3A* and *NRAS* genes were described to be the most frequently mutated genes across all adult AML samples. All identified mutations were classified into 11 functionally related categories of AML, with patients harboring *NPM1* mutations comprising the

largest subgroup of the investigated cohort (Table 1). Moreover, this study identified patterns of co-occurred as well as mutually exclusive mutations. For instance, NPM1 mutations preferentially co-existed with NRAS G12/G13 mutations, but not with NRAS Q61 mutations (Chen, Shen et al., 2013, Ley et al., 2013, Papaemmanuil et al., 2016). These analyses indicated that over 40% of AML patients do not have mutations in genes encoding conventional signaling pathway components and many mutations remain “undruggable” which highlights the high complexity of this disease (Chen et al., 2013, Papaemmanuil et al., 2016).

The abovementioned characteristics of AML relate to adult patients. As acute leukemia remains the most frequently diagnosed form of cancer in children, several research groups became interested in characterizing the mutational landscape of AML in pediatric patients (Steliarova-Foucher, Colombet et al., 2017). Interestingly, whole genome-, mRNA-, targeted DNA- and microRNA-sequencing revealed that many of the most common mutations in adults were significantly under-represented in pediatric AML (Farrar, Schuback et al., 2016). A recent study summarizes results of nearly 1000 samples from infants (<3 years), children (3-14 years), adolescents and young adults (15-39 years) (Bolouri, Farrar et al., 2018). Similar to adult AML patients, pediatric AML showed lowest mutations rates of among all cancers. Only five mutations and/or structural abnormalities were identified in more than 5% of patients. Those involve mutations in *FLT3*, *NPM1*, *WT1*, *CEBPA* and *KIT* as well as translocations involving *RUNX1*, *CBFB*, *MLL*, loss of the Y chromosome and trisomy 8. Furthermore, *NRAS* and *WT1* were mutated far more often in younger patients, while mutations in *DNMT3A*, *IDH1/2*, *TP53* and *NPM1* were more frequently found in adults. In fact, despite their high prevalence in adults, no mutations in the protein-coding region of *DNMT3A* were found in younger patients. Newly characterized mutations in *FLT3* (*FLT3.N*) and internal tandem duplications of *MYC* (*MYC* ITDs) were exclusively found in younger patients. Both, prevalence and positions of mutations in *GATA2*, *CBL*, *WT1* were also found to be distinct in samples from young patients (Bolouri et al., 2018, Ma, Liu et al., 2018).

These findings highlight differences in mutational burden between childhood and adult AML patients and emphasize the need for pediatric leukemia-specific development of precision therapies.

In AML, malignant myeloid progenitor cells proliferate irrepressibly and accumulate in the bone marrow due to impaired differentiation, influencing vital functions of the patient's hematopoietic system (Figure 1). Different molecular mechanisms were proposed to provoke these aberrant features (Figure 2).

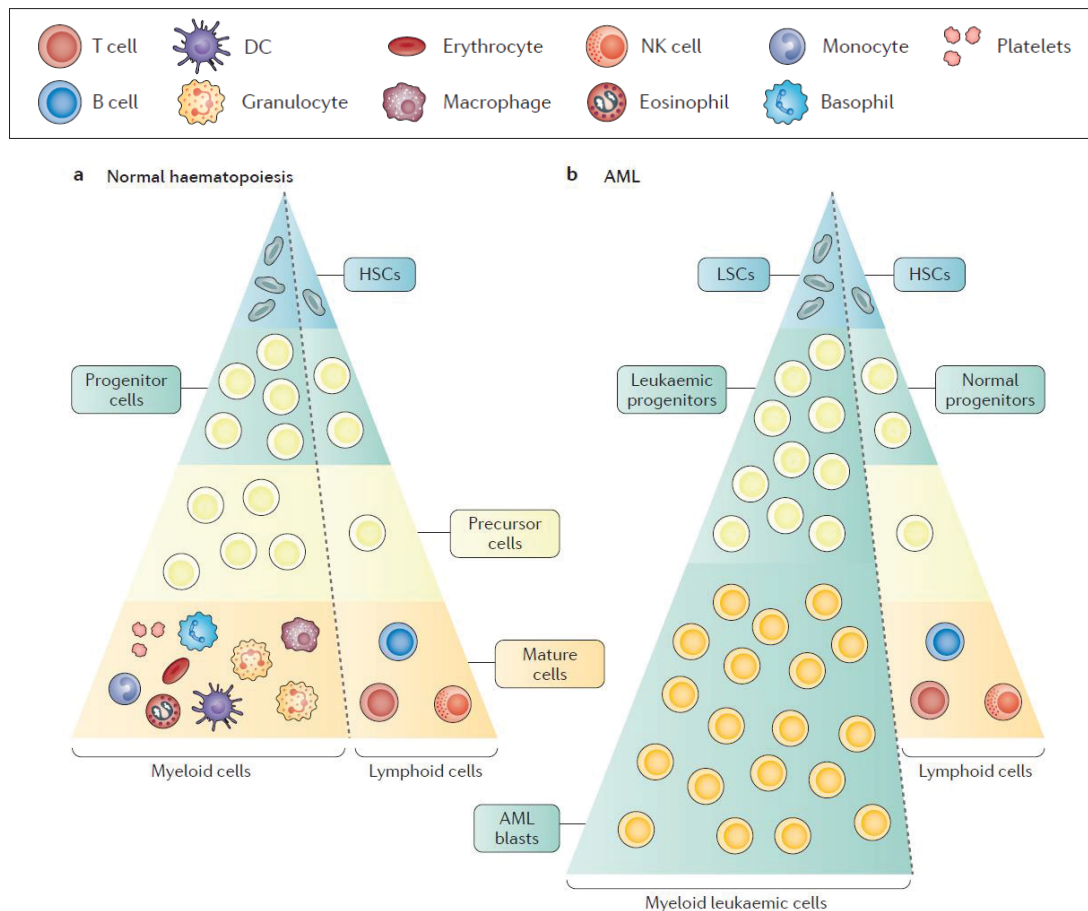


Figure 1. Comparison between normal (left) and leukemic (right) hematopoiesis. (a) In the hierarchical structure of normal hematopoiesis, hematopoietic stem cells (HSCs) have self-renewal capacity and give rise to hematopoietic progenitor cells. Populations of cells undergoing differentiation become more and more committed to mature cell lineages and lose self-renewal potential. **(b)** Similarly, in acute myeloid leukemia, leukemic stem cells (LSCs) give rise to AML progenitor cells. Leukemic progenitor cells are characterized by a block in differentiation, uncontrolled proliferation and expansion of myeloid blasts. DC - dendritic cells, NK - natural killer (Taken from: Khwaja et al., 2016; Reprinted with permission from Springer Nature).

For instance, activating mutations in NRAS, KRAS and hematopoietic tyrosine kinases such as FLT3, or c-KIT can trigger hyper-activation of RAS-RAF, JAK-STAT or PI3K-AKT signaling pathways, leading to a proliferative advantage (Dohner, Weisdorf et al., 2015, Ley et al., 2013). Mutations in hematopoietic transcription factors such as RUNX1, CEBPA, GATA2, MYB or PU.1 can lead to transcriptional deregulation and faulty terminal differentiation of hematopoietic cells (Dohner et al., 2015, Ley et al., 2013, Tenen, 2003). Nucleophosmin (NPM1) shuttles between cytoplasm and nucleus. Frequent mutations of NPM1 can lead to atypical cytoplasmic localization of this protein and its interactors (for example ARF) (Dohner et al., 2015, Ley et al., 2013).

Furthermore, mutations in genes involved in epigenetic regulation, such as ASXL1, EZH2 or chromosomal translocations of MLL can deregulate tightly controlled deposition of various histone marks. Similarly, mutations of DNMT3A, TET2, IDH1/2 lead to alterations in global and focal DNA methylation patterns (Dohner et al., 2015, Ley et al., 2013). Moreover, mutations of tumor suppressor genes such as p53 contribute to deregulation of transcription (Dohner et al., 2015, Ley et al., 2013). Mutations of genes involved in splicing, such as SRSF2, SF3B1, U2AF1 or ZRSR2 lead to abnormal RNA processing. Finally, mutations of genes involved in the cohesin complex, including STAG2 or RAD21 can mediate aberrant chromosome segregation and defective gene expression (Dohner et al., 2015, Khwaja, Bjorkholm et al., 2016, Tenen, 2003).

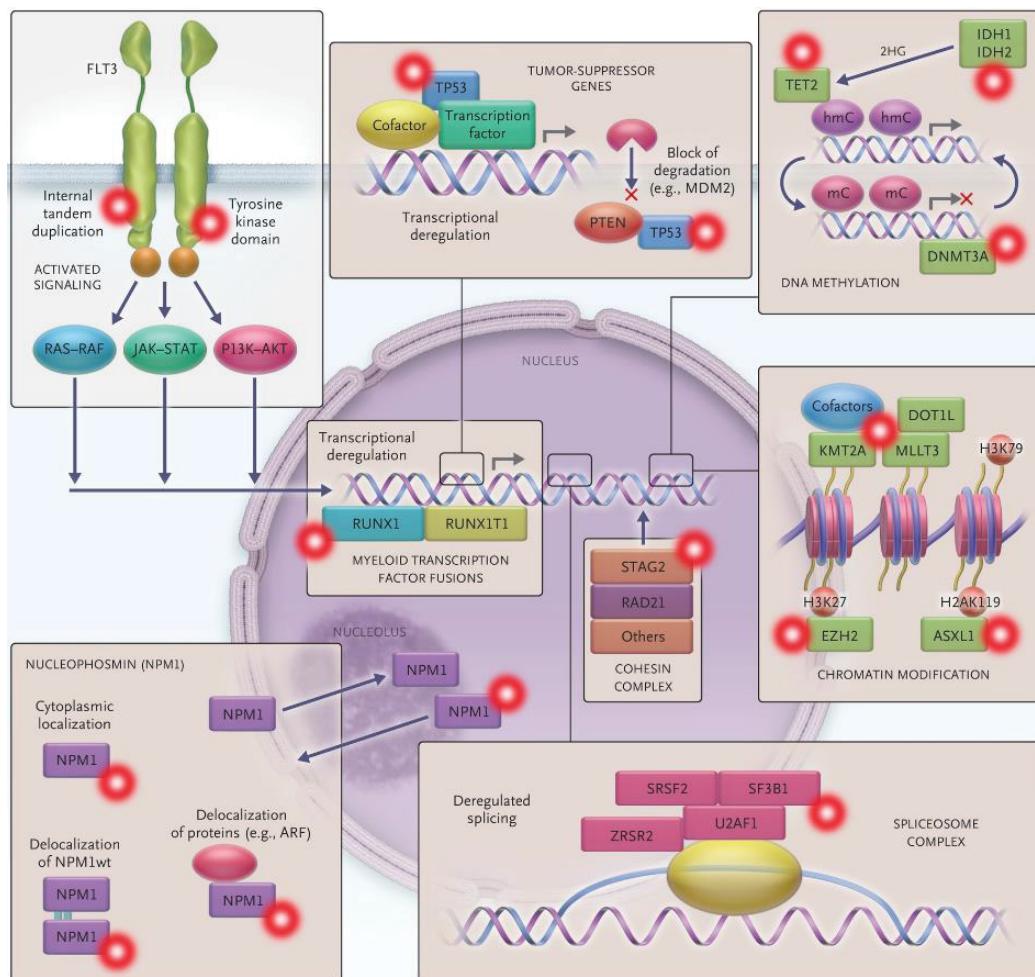


Figure 2. Mechanisms involved in the development and propagation of acute myeloid leukemia. Mutations of genes in acute myeloid leukemia result in defined mechanistic consequences (Reproduced with permission from: Dohner et al., 2015; Copyright Massachusetts Medical Society).

The current therapeutic strategy to treat AML patients is divided into induction, consolidation and maintenance stages of chemotherapy. Typically, therapy begins with an assessment to classify patients eligible for intensive induction chemotherapy. In the usual setup, induction chemotherapy involves infusions of cytarabine and anthracycline and aims at achieving complete remission. Response rates between 60-85% are observed for patients under 60 years of age and drops down to 40-60% in older patients. In consolidation therapy, lower doses of chemotherapeutic agents than during induction therapy are used. Post-remission strategies involve chemotherapy and transplantation of hematopoietic stem cells. In patients under 60 years, 2-4 cycles of cytarabine are administered. Allogeneic hematopoietic stem cell transplantation showed very high therapeutic efficacy in AML patients. Current improvements of standard therapy involve administration of higher doses of cytarabine, anthracyclines and the addition of other drugs, such as idarubicine and fludarabine. A combination of tyrosine III kinase inhibitors such as quizartinib or midostaurin with standard chemotherapy resulted in significantly prolonged overall and event free survival of AML patients harboring *FLT3* mutations (Altman, Foran et al., 2018, Gallogly, Lazarus et al., 2017, Stone, Larson et al., 2017).

Furthermore, new approaches involve monoclonal antibody-based immunotherapy, chimeric antigen receptor-modified T cells (CAR-T cells), bi-specific antibodies and drug-antibody conjugates (Dohner et al., 2015, Khwaja et al., 2016).

1.3 Multi-Partner Translocation (MPT) families in Acute Myeloid Leukemia

Leukemia-associated fusion proteins became a paradigm for cancer research and contributed fundamental insights into disease mechanisms. The first chromosomal translocation specific to a hematopoietic disease was reported in 1973. Improved cytogenetic techniques led to the discovery of a balanced translocation between chromosomes 9 and 22 in chronic myeloid leukemia (CML) patients, resulting in the formation of the Philadelphia chromosome. The fusion leads to the production of the hybrid BCR-ABL1 protein (Mertens, Johansson et al., 2015, Rowley, 1973). Detailed molecular characterization of the BCR-ABL1 fusion protein shed new light on the significance of chromosomal aberrations in hematological malignances and enormously increased our understanding of molecular mechanisms of oncogenesis, resulting in the generation of the targeted BCR-ABL1 inhibitor Imatinib (Mitelman, Johansson et al., 2007).

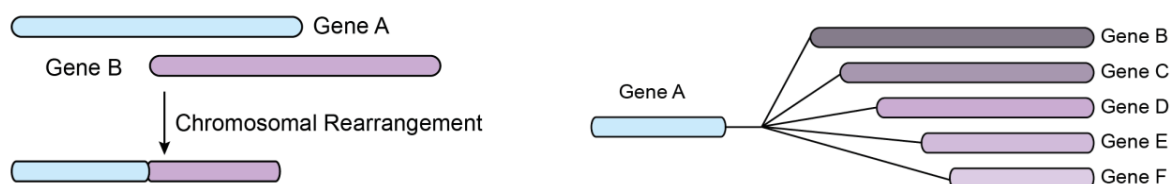


Figure 3. Multi-Partner Translocation (MPT) families. Schematic representation of the mechanisms leading to formation of fusion proteins (left) and classification of MPTs (right).

AML features the highest frequency of balanced chromosomal aberrations and number of gene fusions in cancer. To date, over 250 chromosomal translocations have been identified, leading to the recurrent expression of over 100 distinct fusion proteins (Mertens et al., 2015, Mitelman F, 2018). Many of these fusion genes are associated with a defined disease phenotype. Their recurrent expression within distinct AML subtypes implies a strong cancer driver function for many gene fusions, such as the oncoproteins resulting from the PML-RARA and RUNX1-RUNXT1 translocations (Table 1).

Interestingly, a significant number of these genes was found to be involved in cytogenetically distinct translocation events involving more than one fusion partner gene. This observation allowed to distinguish “Multi-Partner Translocation (MPT) Families”, in which one promiscuous gene is fused to several recipient loci (Figure 3). Fusion proteins belonging to MPT families share the biochemical properties of a common partner moiety but can also be characterized by very different functions of the unique fusion partner gene. Due to their modular character, these families represent unique opportunities to study how the combination of different functional protein modules can drive oncogenic transformation.

The largest MPT family in AML represents translocations of the *MLL* (Mixed Lineage Leukemia) gene with over 135 different *MLL* rearrangements identified to date (Meyer, Burmeister et al., 2017). Other MPT families associated with AML involve *RUNX1* (Runt related Transcription Factor 1) and *NUP98* (Nucleoporin 98) genes. These two genes were found to fuse to over 20 various fusion partner genes. Particularly high prevalence of these fusions in human acute leukemias (around 10% for *MLL*-fusions, 5-8% form *MLL*-ETO and 2-30% for *RUNX1*- and *NUP98*-rearrangements depending on patient age and leukemia subtype) indicates their clinical relevance (Gough, Slape et al., 2011, Martens & Stunnenberg, 2010, Muntean & Hess, 2012). The importance of fusion proteins in hematological malignances is highlighted by the fact that certain fusion protein were proposed to constitute disease defining entities within the WHO classification of genetic abnormalities in AML (Fritz, Percy et al., 2013). Globally, it is estimated

that all gene fusions identified so far are responsible for 20% of human morbidity associated with cancer (Mertens et al., 2015).

1.4 MLL rearranged leukemia

First research activities focusing on the *MLL* gene were initiated almost 30 years ago, and continued interest led to the discovery of its involvement in balanced translocations with almost 100 different fusion partner genes at the end of 2017 (Meyer et al., 2017, Thirman, Gill et al., 1993, Tkachuk, Kohler et al., 1992). The *MLL* gene consists of 37 exons and is located on the long arm of chromosome 11 (11q23), giving rise to a protein of ~450 kDa (3969 amino acids). *MLL* is ubiquitously expressed in human tissues, including hematopoietic stem and progenitor cells and localizes to the nucleus (Bagger, Sasivarevic et al., 2016).

The *MLL* gene encodes the human homolog of the *Drosophila melanogaster* trithorax (TrxG) histone methyltransferase (also known as KMT2A, ALL1, HRX, HTRX). In the fruit fly, the TrxG gene family was shown to positively regulate the expression of a set of genes through DNA binding and methylation of histone tails. Initially, the TrxG family of genes was identified in mutant flies with altered body segmentation phenotypes, predominantly showing three breast segments in the abdomen. This so-called “homeotic transformation” was caused by altered expression of homeobox (Hox) genes (Brock & van Lohuizen, 2001). Hox genes belong to transcriptional regulators that are responsible for correct body formation during embryogenesis. In addition, they are important regulators of tissue development, including the hematopoietic system. Expression of the Hox cluster genes is activated by MLL-induced methylation of lysine 4 in histone 3 (H3K4me) (Argiropoulos & Humphries, 2007). The importance of MLL in embryonic development of higher organisms was substantiated by experiments performed in MLL-knockout mice. MLL was critical for the maintenance of Hox gene expression and MLL-deficiency resulted in embryonic lethality at ~E10.5-16.5, depending on the knockout allele (McMahon, Hiew et al., 2007, Yagi, Deguchi et al., 1998, Yu, Hess et al., 1995). In analogy to the phenotype observed in *Drosophila*, MLL-knockout animals showed significant growth retardation, homeotic transformation, including altered expression patterns of Hox genes and skeletal malformations. Analysis of the hematopoietic system identified impaired proliferation of hematopoietic stem (HSC)- and progenitor cells (Yu et al., 1995). As in *Drosophila*, MLL was found to be involved in a number of protein complexes and its function likely reaches beyond the regulation of Hox gene expression, as will be discussed below.

The term Mixed Lineage Leukemia originated from the observation that blast cells from patients diagnosed with particular forms of acute leukemia expressed surface markers characteristic for both myeloid and lymphoid lineages (Mirro, Zipf et al., 1985, Van den Berghe, David et al., 1979). Later, cytogenetic studies revealed that these blasts often contained a translocation involving the 11q23 locus. Patients diagnosed with this type of translocation usually exhibited a significantly worse prognosis than other leukemia patients. Chromosomal rearrangements of the *MLL* gene are inherently associated with the phenotype of MLL-rearranged leukemia, which is diagnosed predominantly in pediatric patients, but also in adults (Krivtsov & Armstrong, 2007). In addition, MLL-rearrangements are found in therapy-induced acute leukemia patients (t-AML or t-ALL). t-AML can develop in patients that have been previously treated with topoisomerase II- inhibitor like etoposide for other, usually unrelated types of cancer (Pui, Chessells et al., 2003, Pui, Gaynon et al., 2002, Pui & Relling, 2000). Overall, around 10% of human acute leukemias of all types harbor translocations of the *MLL* gene (Huret, Dessen et al., 2001, Krivtsov & Armstrong, 2007). Even though translocations involving the *MLL* gene are the most frequent and very well characterized abnormalities of this gene, other aberrations have also been described. For instance, partial tandem duplications (PTD) of exons encoding the N-terminal part of MLL, originating from a short repeat within the MLL coding sequence are found in 5-10% of patients with AML (Basecke, Whelan et al., 2006, Schichman, Caligiuri et al., 1994). Moreover, some AML patients were discovered to exhibit an increased copy number of the *MLL* gene, resulting from an additional copy of chromosome 11 or amplification of 11q23 (Poppe, Vandesompele et al., 2004). Rearrangements of the *MLL* gene as identified in human acute leukemias are schematically depicted in Figure 4.

1.5 Structure of the MLL protein

Methylation of lysine 4 of histone 3 (H3K4) by complex of proteins associated with Set1 (COMPASS) family of proteins is highly conserved from yeast to humans and is always associated with positive regulation of transcription (Rickels, Hu et al., 2016, Shilatifard, 2012). Human MLL was discovered as a homologue to Set1, the first histone methyltransferase responsible for methylation of H3K4 ever described in *Saccharomyces cerevisiae*. It was shown that Set1 is part of a large multiprotein complex, termed COMPASS (Krogan, Dover et al., 2002, Miller, Krogan et al., 2001, Roguev, Schaft et al., 2001). In *Drosophila melanogaster*, the trithorax (trx) protein is a homologue of the mammalian MLL protein. Also in this case, trx is embedded in a COMPASS-like protein assembly (Mohan, Lin et al., 2010). The first mammalian H3K4

methyltransferase was discovered in 2002 (Milne, Briggs et al., 2002, Nakamura, Mori et al., 2002). The importance of COMPASS components in MLL-rearranged leukemia has been established and will be discussed in the later part of this thesis (Hughes, Rozenblatt-Rosen et al., 2004, Yokoyama, Wang et al., 2004).

The wild type MLL protein is proteolytically cleaved by the threonine endopeptidase Taspase1, giving rise to an N-terminal portion (300 kDa) and a C-terminal portion (180 kDa). Two cleavage sites have been identified: one after amino acid 2666 and another one after amino acid 2718 (Daser & Rabbitts, 2004). The resulting subunits of the protein are held together by two FYRN and FYRC interaction motifs, which are located in the N- and C-terminal cleavage products, respectively (Hsieh, Ernst et al., 2003, Nakamura et al., 2002, Yokoyama, 2002). Taspase1^{-/-} mice phenocopied MLL-deficiency, as they showed homeotic transformation and reduced proliferation of mouse embryonic fibroblasts, indicating that post-translational processing of MLL is required for its function (Takeda, Chen et al., 2006).

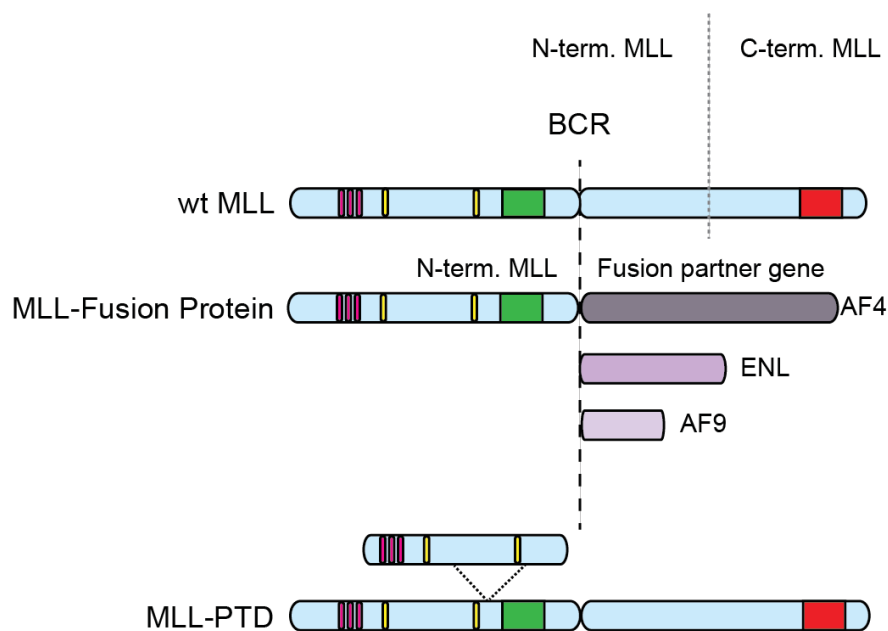


Figure 4. MLL rearrangements associated with human acute leukemias. Figure depicts known alterations of the *MLL* gene in cancer: translocations and partial tandem duplications (PTDs).

The wild type MLL protein harbors several functional domains (Figure 5). The N-terminal part of the protein contains three AT-hook domains which bind to the minor groove of DNA in regions rich in thymidine (T) and adenine (A) (Zelevnik-Le, Harden et al., 1994). This part is followed by two nuclear localization motifs (SNL1 and SNL2) conserved from *Drosophila melanogaster* to humans. The transcription repression domain (TRD) or methyl-CpG-binding domain (MBD) includes a CxxC motif, responsible for the recruitment of proteins associated with transcriptional repression, such as CtBP or the histone deacetylases HDAC1/2 (Xia, Anderson et al., 2003). The middle part of the protein contains four plant homology domains (PHD).

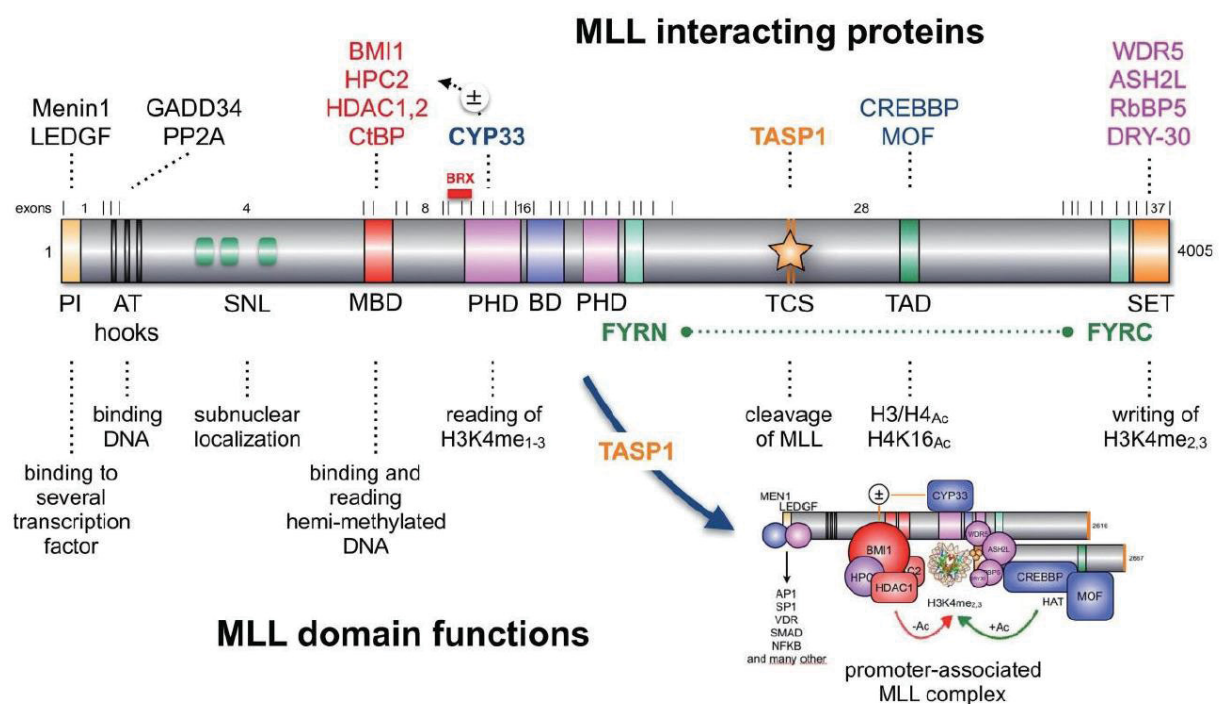


Figure 5. Detailed structure of the wild type MLL protein. Functional domains of MLL and their role in biology are depicted below the structure. The exon-structure and known MLL binding proteins are depicted above the structure. The assembled MLL complex is depicted in bottom/right. (Reprinted with permission from: Ann Lab Med. 2016;36(2):85-100).

One of the PHD fingers has the ability to bind methylated H3K4, important for the recruitment of MLL to its target genes (Fair, Anderson et al., 2001). The middle part of the protein also contains a bromodomain (BD). Bromodomains bind to acetylated lysine residues, thus further supporting the interaction of the MLL complex with active chromatin (Rao & Dou, 2015). A transcriptional activation domain is located in the C-terminal part of the MLL and can recruit CREB-binding protein (CBP) (Ernst, Wang et al., 2001). The SET domain (Suppressor of variegation; Su(var)3-

9, enhancer of zeste, trithorax) located at the extreme C-terminus of MLL has methyltransferase activity (Milne et al., 2002).

1.6 Composition of the MLL protein complex

Since the MLL protein complex was purified from mammalian nuclei around 15 years ago, it became evident that the wild type MLL protein never acts in isolation but is embedded in a large complex involving a multitude of proteins with different functions. The composition of the MLL complex is determined by a set of MLL-interaction partners that recognize distinct domains within the structure of MLL protein. Overall, the N-terminal portion of the protein was described to be essential for chromatin targeting of the MLL complex, while the C-terminal part was described to be responsible for activation of gene expression (Slany, 2016).

The very N-terminal residues of the MLL bind to Menin (multiple endocrine neoplasia I). Menin acts as a scaffold protein, providing an interaction interface for LEDGF (lens epithelium derived growth factor, also known as p75) (Caslini, Yang et al., 2007). The CxxC and AT-hooks domains in MLL are important for direct DNA-binding. However, it was also shown that targeting of the MLL complex to specific genes can be mediated by its interaction with other proteins (Risner, Kuntimaddi et al., 2013). For instance, LEDGF contains a PWWP domain, which recognizes and binds H3K36me3, a hallmark of actively transcribed chromatin (Vermeulen, Eberl et al., 2010). Moreover, the CxxC region of MLL can interact with repressive factors such as Polycomb proteins and histone deacetylases. Recruitment of these proteins is facilitated by conformational changes related to binding of Cyp33 to one of MLL PHD fingers (Xia et al., 2003).

The C-terminal part of MLL features two regions that have the capability to modify chromatin. The transactivation domain (TAD) is important for the interaction of MLL with histone acetyltransferases such as MOZ, p300/CBP or MOF, which are responsible for acetylation of H3K9, H3K27 and H4K16 respectively (Dou, Milne et al., 2005, Ernst et al., 2001, Paggetti, Largeot et al., 2010). The second region represents the SET domain in the very distal C-terminus, which has methyltransferase activity and can catalyze H3K4 di- and tri-methylation (Milne et al., 2002).

In general, methylation of H3K4 is catalyzed by methyltransferases belonging to the SET1 family, comprising MLL1, MLL2, MLL3, MLL4, SET1A and SET1B. Enzymes of the SET1 family require interaction with several core complex members for full activation. As these proteins are WDR5 (WD repeat-containing protein 5), RbBP5 (retinoblastoma-binding protein 5), ASH2L (Set1/Ash2 histone methyltransferase complex subunit ASH2) and DPY30, the MLL/SET core complex was

termed WRAD (Ernst & Vakoc, 2012). As depicted in Figure 5, the elements of the WRAD complex bind in close proximity to the SET domain of SET1 family proteins and invigorate their methyltransferase activity. The same components were identified in complex purifications of Set1 complexes from yeast, fly and humans, indicating that these factors are highly conserved (Ardehali, Mei et al., 2011, Dou, Milne et al., 2006, Goo, Sohn et al., 2003, Mohan, Herz et al., 2011, Wu, Wang et al., 2008). Moreover, loss of any WRAD component leads to aberrant gene expression that phenocopies deletion of SET1, indicating that the interaction of these proteins is functionally relevant *in vivo* (Miller et al., 2001, Nagy, Griesenbeck et al., 2002). Consistently, it was described that the methyltransferase activity of free MLL1 is weak and specific to H3K4 mono-methylation (Dou et al., 2006). Biochemical experiments with individual WRAD components and MLL1 showed that the methyltransferase activity of MLL is only moderately increased after binding to WDR5 and RbBP5, while it further increases up to 300 fold and acquires the ability to catalyze H3K4 tri-methylation only when all WRAD components are assembled (Dou et al., 2006, Patel, Dharmarajan et al., 2009). Interestingly, apart from its regulatory role, WRAD can mediate other protein:protein, protein:DNA and protein:RNA interactions, which facilitates the recruitment of SET1 complexes to their target genes. For instance, the winged-helix domain of ASH2L was shown to bind DNA (Chen, Wan et al., 2011, Sarvan, Avdic et al., 2011). WDR5 can interact with histone H3 by engaging its WD40 beta propeller domain (Wysocka, Swigut et al., 2005). Moreover, WDR5 was described to bind the transcription factor Oct4 and is responsible for the recruitment of SET1 proteins to co-activate gene regulatory networks to sustain the self-renewal of embryonic stem cells (Ang, Tsai et al., 2011).

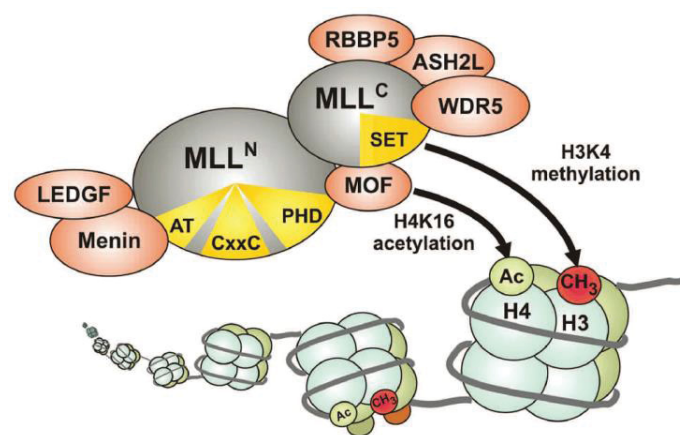


Figure 6. The protein complex around wild type MLL. The core interactors of N- and C-terminal part of MLL and their most relevant domains are shown. (Taken from: Slany, 2009; Reprinted with permission from: Haematologica).

1.7 Translocations of the *MLL* gene in leukemia

Chromosomal translocations originate when two fragments of non-homologous chromosomes are being exchanged during non-homologous end joining (NHEJ)-mediated DNA repair (Gillert, Leis et al., 1999, Richardson & Jasin, 2000). The origin of translocations involving the *MLL* gene was initially assigned to impaired repair of DNA double strand breaks during HSC development. In patients with therapy-related leukemia, translocations of the *MLL* gene also originate from aberrantly repaired double strand breaks of DNA, which develop after inhibition of Topoisomerase II, an enzyme that can cut both strands of DNA to amend DNA supercoils and tangles (Libura, Slater et al., 2005).

The major breakpoint cluster region (BCR) is located between introns 9 and 11 of the *MLL* gene. This region of ~8kb represents a target for most *MLL* rearrangements. Thus, the majority of *MLL*-fusion proteins retain the first 9-13 exons of *MLL* and a variable number of exons of a translocation partner gene (Meyer, Burmeister et al., 2017).

Translocations of the *MLL* gene are found in 5-10% of all human acute leukemias. The most recent study describes a total number of 135 distinct translocations of the *MLL* gene, of which 94 direct translocation partner genes have been characterized at the molecular level (Meyer et al., 2017).

Even though a large number of translocation partner genes has been identified, some fusion partners recombine with *MLL* particularly often. These fusion proteins include *MLL*-AF4 (t(4;11)(q21;q23), *MLL*-AF6 (t(6;11)(q27;q23)), *MLL*-AF9 (t(9;11)(p22;q23)), *MLL*-AF10 (t(10;11)(p12;q23)) and *MLL*-ENL (t(11;19)(q23;p13.3)), making up for >75 % of all *MLL*-fusions. Interestingly, the abundance of different *MLL*-fusion proteins distribute differently between childhood- and adult leukemias. In addition, there is a significant difference between the recombinomes of *MLL*-rearranged ALL vs. AML. Analysis based on disease phenotype revealed a prevalence of *MLL*-AF4, *MLL*-AF9 and *MLL*-AF10 in ALL, where these fusions constitute over 90% of *MLL*-rearranged cases. *MLL*-rearranged AML predominantly involves *MLL*-AF9, *MLL*-ELL and *MLL*-AF10 fusions, although the frequencies are different between infant, pediatric and adult patients (Figure 7) (Meyer et al., 2017). Moreover, it has been described that *MLL*-rearranged leukemias have significantly fewer cooperating mutations as compared to other mutational drivers in leukemia. On average, pediatric *MLL*-rearranged leukemia carries 1.3 non-silent mutations. In most of the cases, however, translocations of the *MLL* gene are the only detected aberrations (Papaemmanuil et al., 2016).

MLL-fusion proteins can be functionally categorized based on the fusion partner gene and its putative function. The most frequently identified MLL-fusion partners AF4, AF9, AF10, EEN and ENL all encode nuclear proteins. This indicates that in most cases, the process of translocation may not be stochastic and is associated with particular functional consequences (Meyer, Schneider et al., 2006). Another group of MLL-fusion partners represents cytoplasmic proteins, including AF1p, EEN, GAS7, AFX and AF10. These fusion partners contain coiled-coil oligomerization domains, which are necessary for oncogenic transformation induced by the respective MLL-fusion protein (So & Cleary, 2003, So, Lin et al., 2003b). In 1% of cases, MLL recombines with septins (SEPT2/5/6/9/11), which have a known function in cell cycle control or mitosis (Hall & Russell, 2004). Fusions of the *MLL* gene with histone acetyltransferases p300 and CBP constitute another small functional group of fusion partner genes. Interestingly, the enzymatic activity of these proteins remains unaffected after recombination (Ida, Kitabayashi et al., 1997, Taki, Sako et al., 1997).

These examples account for fusion proteins in which the 5'-portion of the *MLL* gene is fused to numerous loci. However, several publications described "reciprocal MLL-fusions". These fusion proteins involve the C-terminal portion of the *MLL* gene. While AF4-MLL and NEBL-MLL fusions were functionally characterized, their relevance in leukemia development remains unclear (Bursen, Schwabe et al., 2010, Emerenciano, Kowarz et al., 2013a, Meyer, Hofmann et al., 2013).

1.8 Clinical consequences of MLL-fusion proteins expression in leukemia

In affected patients, the frequency of MLL-rearranged ALL is highest in the first two years of life, decreases during childhood and adulthood and then increases with age. Such a characteristic postnatal peak has not been described for MLL-rearranged AML, however a similar distribution has been observed (Meyer et al., 2017, Meyer, Hofmann et al., 2013). Several studies have described that the prognosis of leukemias harboring rearrangements of the *MLL* gene varies according to the nature of the fusion partner gene, cell lineage and patients age. Moreover, it was shown that the rearrangements of infant patients occur already in utero (Greaves, Maia et al., 2003). Overall, MLL-AF10, MLL-ELL and MLL-AF6 fusion proteins are mainly found in AML but not in ALL, while MLL-AF4 is predominantly associated with ALL (Meyer et al., 2013).

In general, MLL-rearranged AMLs present with similar prognosis as other, non-rearranged AMLs (Felix, Hosler et al., 1995, Mrozek, Heinonen et al., 1997). In contrast, MLL-rearranged ALL is associated with a significantly worse clinical outcome than AML, and the prognosis is particularly bad for pediatric patients (Winters & Bernt, 2017). For instance, infants with the most commonly

found translocation in ALL, t(4;11) (MLL-AF4), present with particularly poor event free survival (EFS) rates of around 19%, which raises up to 40% in older pediatric patients (Huret, Dessen et al., 2000). In comparison, the EFS of infants with non-MLL-rearranged ALL was estimated for 80-96% (Mohan et al., 2010). Similarly, MLL-AF9 ALL presents with poor clinical outcome, with event free survival (EFS) rate around 38% in infants. Interestingly, AML patients expressing the same, most common translocation for this lineage, t(9;11), is associated with significantly improved outcome (EFS 63%). However, the median survival for all *de novo* identified cases of MLL-AF9 AML was estimated to be only 4 years (Huret, Dessen et al., 2003, Mrozek et al., 1997, Rubnitz, Raimondi et al., 2002). MLL-ENL is commonly found in infant ALL and is associated with poor clinical outcome. A median survival rate of less than 12 months was estimated for these cases (Huret, Senon et al., 2004). Rearrangements of MLL with AF6 and AF10 are predictors of significantly worse survival in AML (Balgobind, Raimondi et al., 2009, Grimwade, Hills et al., 2010, Pui, Carroll et al., 2011). This is in contrast to MLL-ELL and MLL-ENL, which are associated with a favorable prognosis in this lineage (Balgobind et al., 2009, Grimwade et al., 2010, Pui et al., 2011). Finally, therapy related MLL-rearranged leukemias also present with very poor clinical outcome (Krivtsov & Armstrong, 2007).

The location of the breakpoint cluster region varies between patient groups of different age. Childhood MLL-rearranged ALL is often associated with a BCR located in intron 11 of the *MLL* gene (Emerenciano, Meyer et al., 2013). Conversely, late disease onset, corresponding to adult leukemia patients is frequently associated with a BCR located in MLL introns 9 and 10 (Reichel, Gillert et al., 2001). Moreover, the location of the BCR results in various expression levels of different fusions. This observation could be explained by the domain structure of the *MLL* gene (Figure 5). MLL exons 11-16 encode the PHD domains. If the breakpoint lies within this region, it likely disrupts the structure of the first PHD finger, leading to alternative cysteine-histidine pairing and eventually aberrant folding of the neighboring domains (Emerenciano et al., 2013b). A BCR within MLL exon 11 intron could therefore result in impaired binding of ASB2, a component of the ECS(ASB) E3 ubiquitin ligase and CYP33. Expression of the ASB2 complex was shown to mediate degradation of MLL (Wang, Muntean et al., 2012). As ASB2 expression levels increase, this leads to lower levels of MLL protein and significant downregulation of MLL target genes during hematopoietic differentiation. Interestingly, due to a different location of the BCR, fusion proteins such as MLL-AF9 or MLL-AF4 do not interact with ASB2 and remain insensitive to ASB2-mediated degradation. In turn, this results in continuous expression of MLL-fusion proteins target genes and leukemia development (Marschalek, 2011a, Wang et al., 2012).

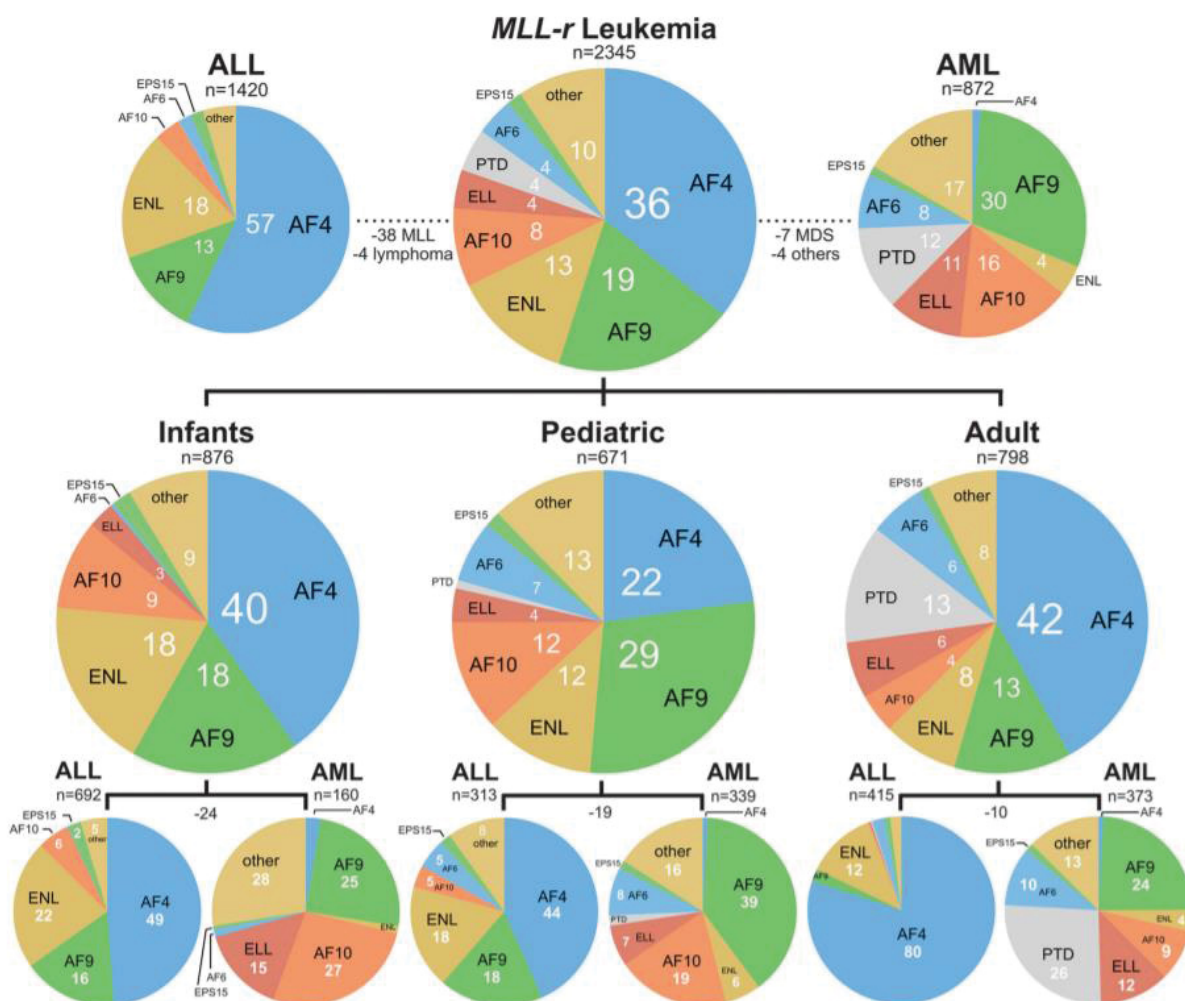


Figure 7. Classification of patients with MLL-rearranged leukemia. Classification of MLL-fusion partner genes based on age and disease phenotype is shown (Taken from: Meyer et al., 2017; reprinted under the license: <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>).

Furthermore, the poly(ADP-ribose) polymerase CYP33 associates with CtCB, HDAC1/2 and Polycomb proteins, which constitute transcriptional repressors. Thus, in physiological situation, CYP33 binding can convert MLL into a transcriptional repressor. As the CYP33 binding site is downstream of the BCR and therefore absent in MLL-translocations, this repressive mechanism does not exist in MLL-rearranged leukemia and might result in activation of transcription of MLL target genes (Hom, Chang et al., 2010, Wang, Song et al., 2010). BCRs affecting other introns will not have the same consequences (Emerenciano et al., 2013b, Fair et al., 2001, Xia et al., 2003).

An additional prognostic factor is associated with the expression levels of master regulatory factors, such as genes belonging to the HOXA cluster. It was shown that a subset of early onset ALL patients harboring MLL-AF4 show particularly low expression levels of HOXA genes and are associated with a particularly high risk of relapse (Stam, Schneider et al., 2010, Trentin, Giordan et al., 2009).

Overall, as the expression of various MLL-translocations is associated with defined prognostic risks, the majority of patients with MLL-rearrangements are treated according to high-risk protocols which include high-dose chemotherapy and stem cell transplantation.

Based on this and other criteria, several studies focused on further classification of MLL-fusion proteins and associated clinical outcomes which is expected to improve patient stratification in the future (Balgobind et al., 2009, Emerenciano et al., 2013b). A better understanding of how various MLL-fusion proteins affect gene expression and thus contributing to oncogenesis might result in more effective therapies.

1.9 Model systems establishing MLL-fusions as driver mutations

Different experimental strategies have been established to investigate the molecular properties associated with the oncogenic activity of different MLL-fusions. In particular, animal models recapitulating the human disease phenotype became fundamental for comprehensive characterization of disease mechanisms and testing novel therapeutic approaches (Milne, 2017). Yet, despite a lot of similarities, the process of hematopoiesis is physiologically different in mice and humans. Human peripheral blood is significantly richer in neutrophils, while murine blood has more lymphocytes, which might influence the behavior of leukemic cells between these organisms (Mestas & Hughes, 2004).

One very basic approach aims to stably introduce a potential oncogene into a cell culture model via viral transduction. This frequently used strategy is usually utilized to characterize changes in proliferation rates, alterations of gene expression or changes in cell cycle or apoptotic phenotypes caused by the exogenous expression of the gene of interest (Krivtsov & Armstrong, 2007, Marschalek, 2011b).

Another method investigates phenotypic consequences associated with an expression of an oncoprotein in a mouse model. For instance, non-obese diabetic severe combined immunodeficient (NOD/SCID) mice can be directly transplanted with a population of primary human cells or human cell lines. Through this procedure, transformed cells can be expanded *in*

vivo and used for detailed downstream analyses (Krivtsov & Armstrong, 2007, Marschalek, 2011b).

Alternatively, C57Bl/6 mice might be utilized in transplantation experiments of murine cell populations that were generated *ex vivo* via viral transduction of an oncogene. Bone marrow is isolated from mice, transduced with a viral vector encoding the MLL-fusion protein and the resulting cell population is injected into irradiated recipient mice to assess leukemia development *in vivo*. Depending on the characteristics of the retro- or lentiviral backbone used in this approach, transgene expression can have constitutive or inducible character. Although this method can have potential technical issues, such as transcriptional silencing of the transgene or unphysiological expression levels, it is frequently used by researches as a validation strategy for *in vitro* experiments (Krivtsov & Armstrong, 2007, Marschalek, 2011b, Zuber, McJunkin et al., 2011a). This approach was employed to study the oncogenic potential of MLL-GAS7 or MLL-ENL (Lavau, Szilvassy et al., 1997, So, Karsunky et al., 2003a). Mouse models were also used to demonstrate that the cell of origin of the developing leukemia has an impact on gene expression and epigenetic programs in a fully developed disease (Krivtsov, Figueroa et al., 2013).

Alternatively, an oncogene can be expressed from an endogenous promotor using a transgenic approach. However, generation of these models is often associated with extended time and financial resources (Krivtsov & Armstrong, 2007, Marschalek, 2011b). The generation of knock-in mice for studies of MLL-AF9 induced leukemia *in vivo* was published over two decades ago (Corral, Lavenir et al., 1996). A conditional knock-in model generating the MLL-AF9 fusion by interchromosomal translocation was described soon afterwards (Collins, Pannell et al., 2000). Other studies established conditional knock-in models of MLL-ENL and MLL-CBP, resulting in improved characterization of phenotypes associated with rearrangements of the *MLL* gene (Forster, Pannell et al., 2003, Milne, 2017, Wang, Iwasaki et al., 2005).

1.10 Transcriptional targets of MLL-fusion proteins

Because the N-terminus of MLL is retained in MLL-fusion proteins, the ability of MLL to bind DNA remains unchanged in the context of MLL-fusions. This indicates that wild type MLL and MLL-fusion proteins share common target genes. For instance, the HOXA cluster, especially *HOXA9*, as well as the genes encoding the transcription factors *MEIS1* and *PBX3* were shown to be major target genes of both wild type MLL and MLL-fusion proteins (Garcia-Cuellar, Steger et al., 2015, Li, Zhang et al., 2013b, Zeisig, Milne et al., 2004). Apart from the function of HOX genes in embryonic development, HOXA cluster genes, PBX3 and MEIS1 were shown to be critical for the

self-renewal potential of hematopoietic progenitor and stem cells. HOX gene expression is particularly high in hematopoietic stem- and progenitor cells and is significantly down-regulated during terminal maturation of blood cells (Ernst, Mabon et al., 2004). High, ectopic expression of HOX genes results in a differentiation block and maintains an abnormal population of progenitor cells that are able to self-renew (Bach, Buhl et al., 2010).

Even though the HOX cluster is the most prominent example of MLL-target genes, several other loci were shown to be regulated by MLL-fusion proteins. Wild type MLL resides at promoter regions of almost all actively transcribed genes (Scacheri, Davis et al., 2006). Furthermore, chromatin immunoprecipitation experiments revealed many binding sites that are specific to MLL-fusion proteins. For instance, chromatin immunoprecipitation of MLL-AF4 followed by deep sequencing revealed over 1000 genomic regions that are bound by the fusion protein in ALL cells (Krivtsov, Feng et al., 2008). Expression profiling experiments revealed that MLL-rearranged ALL exhibits a distinct pattern of gene expression than non-MLL-rearranged ALL or AML. That is consistent with an immature hematopoietic cell type (Armstrong, Staunton et al., 2002). A more recent study has performed a detailed chemo-genomic characterization of MLL-rearranged AML, highlighting a link between MLL-fusion proteins and MLL-PTD. This study found that the expression of a small set of adjacent genes that are duplicated on chromosome 15 is driven by both types of rearrangements of the *MLL* gene (Lavalley, Baccelli et al., 2015).

Ecotropic viral integration site-1 (EVI-1) is a nuclear transcription factor and a target gene of MLL-fusions. EVI-1 plays a pivotal role in hematopoiesis and in the regulation of hematopoietic stem cells. Expression of EVI-1 is significantly up-regulated in MLL-rearranged AML (Arai, Jagasia et al., 2011). Cyclins and Cyclin Dependent Kinase Inhibitors (CDKIs), in particular p27 and p18, constitute another group of genes that are regulated by MLL-fusion proteins. Interestingly, inactivating mutations of Menin and not MLL itself lead to aberrant induction of CDKIs (Milne, Martin et al., 2005, Xia, Popovic et al., 2005). Moreover, HCF-1, a protein involved in cell cycle regulation and a known interactor of MLL-fusion proteins was described to target MLL to S-phase promoters of cyclins and activate the transcription factor E2F, thus implicating MLL in the regulation of the cell cycle and DNA replication (Tyagi, Chabes et al., 2007). Furthermore, MLL can associate with telomeric regions of chromosomes, where it contributes to H3K4 methylation and positively regulates transcription of telomere repeat containing RNAs (Caslini, Connelly et al., 2009). In line with this, up-regulation of *HOXA7* by MLL-fusions was found to influence the transcription of subunits of the telomerase enzyme TERT (Gessner, Thomas et al., 2010). Moreover, MLL-fusion proteins were shown to induce expression of *Frat1* and *Frat2*, which are necessary for the oncogenicity of MLL-fusions (Walf-Vorderwulbecke, de Boer et al., 2012). *Frat2*

was also found to be associated with MLL-fusion-mediated activation of the small Rho GTPases Rac1 and Rac2 (van Amerongen, Nawijn et al., 2010). This established an interesting mechanistic link between MLL-fusion mediated transcriptional regulation and Rac activation to be necessary for MLL-rearranged leukemia (Mizukawa, Wei et al., 2011). Finally, another study found that MLL not only binds to gene promoters but also to enhancer regions of many genes, which links signal-dependent enhancer transcription with H3K4 methylation (Kaikkonen, Spann et al., 2013). Interestingly, a recent hypothesis states that MLL-fusion target genes other than HOXA cluster genes, *MEIS1* and *PBX3* do not play any important roles in leukemic transformation (Bernt, Zhu et al., 2011, Guenther, Lawton et al., 2008, Wilkinson, Ballabio et al., 2013). In support of this statement, wild type MLL was found to reside at every active promoter, while ChIP-seq studies confirmed binding sites of MLL-fusion proteins at only around 400 loci (Bernt et al., 2011, Guenther et al., 2008, Wilkinson et al., 2013).

1.11 Protein complexes around MLL-fusion proteins and mechanisms of MLL-fusion-driven leukemic transformation

As the N-terminal part of MLL containing the CxxC domain and the AT-hooks remains unchanged in MLL-translocations, MLL chimaeras retain the ability to bind to MLL target genes. At the same time, MLL-fusion proteins lack intrinsic methyltransferase activity, as the SET domain and additional protein modules, such as WRAD, are absent. It was proposed that MLL translocation partners exist in various biochemically and molecularly distinct complexes (Mohan et al., 2010, Slany, 2009). This might explain how MLL-fusion proteins mediate the de-regulation of gene expression and contribute to leukemogenesis. As not all MLL-fusion proteins share functional similarities, it is unclear whether any common molecular mechanisms, via which various unrelated MLL-fusion proteins contribute to leukemia development, exist.

Several clinically relevant MLL-fusion proteins were proposed to be associated with transcriptional elongation, as they incorporate into multi-protein complexes with factors involved in the control of this biological process (Figure 8a). ELL was the first translocation partner of MLL described to exhibit this functional link (Shilatifard, Lane et al., 1996). Because of its function as a RNA polymerase II (Pol II) elongation factor, it was proposed that the regulation of transcriptional elongation by Pol II might play a role in disease development driven by MLL-fusion proteins. Indeed, few years later it was shown that the transcription elongation factors ENL (ENL1, ENL2 and ENL3) form the super elongation complex (SEC), which acts as a scaffold for positive transcription elongation factor b (P-TEFb, containing cyclin T1, cyclin T2 and CDK9), ELL

associated factors (EAF1 and EAF2) and several frequent MLL translocation partners such as AFF1, AFF3, AF9, ENL or EEN (Lin, Smith et al., 2010, Luo, Lin et al., 2012, Martin, Milne et al., 2003). All three proteins of the ENL family share an R4 domain located at the C-terminus, which is necessary in the context of MLL-ENL to induce leukemia (DiMartino, Miller et al., 2000). In addition, AF9 and ENL contain highly conserved YEATS domains, which have the ability to recognize acetylation and crotonylation of histone lysine residues (Schulze, Wang et al., 2009, Zhao, Li et al., 2017). Furthermore, the YEATS domain was shown to be essential for MLL-ENL-mediated leukemia maintenance, which highlights ENL as a critical effector of MLL-fusion proteins and establishes its chromatin-reader domain as a potential target in MLL-rearranged leukemia (Erb, Scott et al., 2017). MLL-translocation partners AFF1 and AFF4 interact with various components of SEC. They are able to dimerize, directly bind P-TEFb and thus can regulate its activity *in vivo* (Bitoun, Oliver et al., 2007, Mueller, Bach et al., 2007). RNA Pol II was shown to pause at the promoter regions of genes involved in development, implying that its pause release might be a critical step involved in expression of these genes (Muse, Gilchrist et al., 2007). P-TEFb-mediated phosphorylation of the carboxy-terminal domain (CTD) of Pol II and negative elongation factor (NELF) are critical for the transition from transcriptional initiation to productive transcriptional elongation by RNA-Pol II (Peterlin & Price, 2006).

The DotCom complex represents another protein complex that links MLL-fusions with transcriptional elongation. Similarly to SEC, it involves several proteins that are frequently found in patients harboring MLL-rearrangements, such as AF9, AF10, ENL and AF17 (Okada, Feng et al., 2005). A critical component of this complex is the histone methyltransferase DOT1L, which catalyzes mono-, di- and tri-methylation of H3K79 (Steger, Lefterova et al., 2008, van Leeuwen, Gafken et al., 2002). Interestingly, DOT1L does not contain a SET domain, a catalytic domain present in all other known histone lysine methyltransferases. Structural analysis of DOT1L identified the SAM binding pocket as well as other domains responsible for its enzymatic activity (Feng, Wang et al., 2002, Min, Feng et al., 2003). While many protein-protein interactions within the DotCom complex were shown to require the same domain, and thus being mutually exclusive, the exact architecture of the complex remains uncharacterized (Muntean, Tan et al., 2010). However, known functions assigned to individual components strongly links DotCom to transcriptional elongation. In addition to the role of pTEFb in regulating transcriptional elongation, also DOT1L is involved in this process (Peterlin & Price, 2006, Steger et al., 2008). Furthermore, ENL was shown to bind histone H3, which might further influence chromatin modification activities of this complex (Mohan et al., 2010).

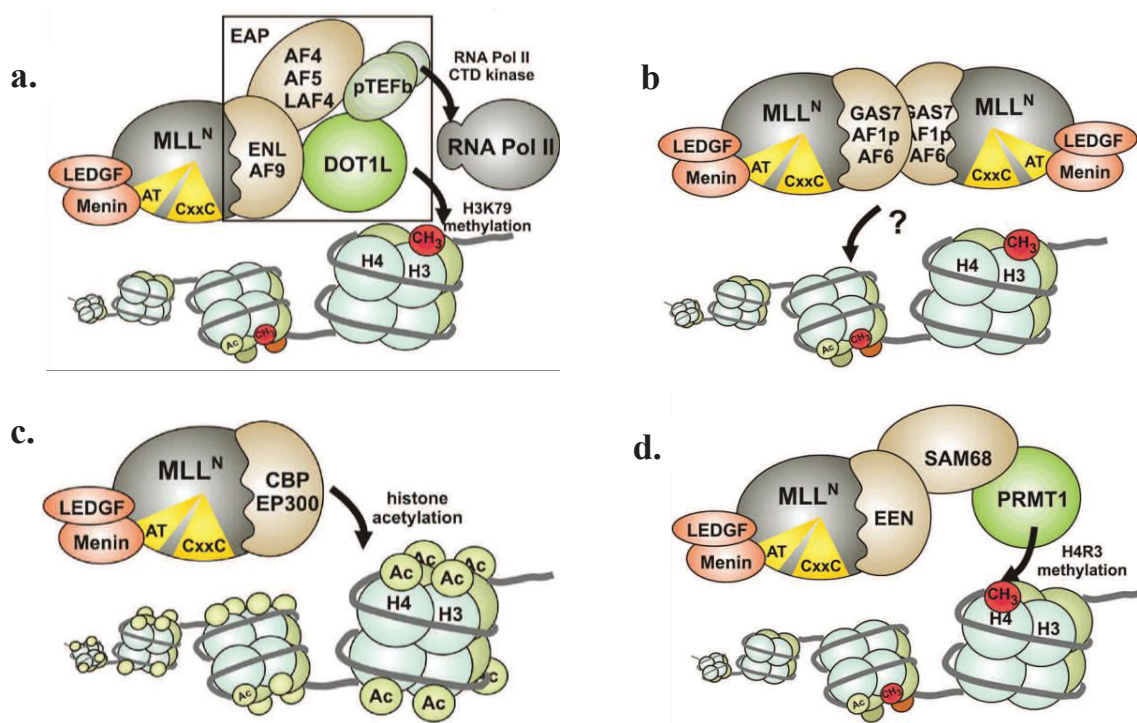


Figure 8. Proposed molecular pathways involved in the development of MLL-rearranged leukemia. Core interactors of MLL-fusion complexes and their proposed activities are shown. (Taken from: Slany, 2009; Reprinted with permission from: Haematologica).

In line with the interaction of MLL-fusions with the DotCom complex, particularly high levels of H3K79 di-methylation were found on MLL-fusion target genes, such as *HOXA9* or *MEIS1* in hematopoietic progenitor cells that express MLL-translocations (Bernt et al., 2011, Daigle, Olhava et al., 2013, Deshpande, Chen et al., 2013, Guenther et al., 2008, Milne et al., 2005). shRNA-mediated downregulation of *DOT1L* resulted in loss of H3K79 methylation. Moreover, *DOT1L* silencing impaired the activity of MLL-fusions, MLL-fusion-mediated leukemic transformation and leukemia maintenance. As DOT1L is the only known writer of H3K79 methylation and H3K79me2 deposition is specifically associated with MLL-fusion-target genes, DOT1L classifies as an ideal target for inhibition with a small molecule with high relevance for MLL-rearranged leukemia patients. As of today, Pinometostat (EPZ-5676) is the most potent, selective inhibitor of DOT1L currently in clinical development (Daigle et al., 2013). Other therapeutic strategies in MLL-rearranged leukemia are discussed a separate paragraph below.

In general, MLL-fusion proteins often take part in related biological processes and are frequently involved in the same protein complexes controlling transcriptional elongation. However, the large number of MLL-fusion proteins indicates that alternative molecular pathways exist which could

contribute to their oncogenic activity (Figure 8b-d). For instance, several MLL-fusion partner genes such as AF1p, AF6 or GAS7 belong to a group of cytoplasmic proteins, which contain a coiled-coiled domain that mediates protein-protein interactions. In fact, artificially induced dimerization of MLL-fusions was necessary and sufficient for leukemic transformation by MLL-fusion proteins (Figure 8b) (So et al., 2003b). While the exact mechanism of oncogenic transformation driven by dimerization events remains poorly understood, it has been hypothesized that cytoplasmic MLL translocation partners must be first imported to the nucleus in response to strong nuclear import signals that are located in the N-terminal portion of MLL to activate MLL-fusion target gene expression (Slany, 2009). This assumption was supported by a study demonstrating that a fusion of MLL with the cytoplasmic protein ABI1 is imported to the nucleus where it interacts with ENL (Garcia-Cuellar, Schreiner et al., 2000). Similar mechanisms might apply to MLL-fusions with other cytoplasmic partner proteins.

MLL-fusions with histone acetyltransferases CREB binding protein (CBP) and p300 were identified in patients with therapy-related leukemia. These fusions represent another possible mechanism of oncogenic transformation, where MLL is fused to a histone acetyltransferase (Figure 8c). Both the bromodomain and histone acetyltransferase domain of the CBP were required for oncogenic transformation *in vitro* and sufficient to promote leukemia *in vivo* (Lavau, Du et al., 2000, Rowley, Reshmi et al., 1997, Sobulo, Borrow et al., 1997).

Furthermore, MLL-EEN was proposed to interact with the arginine methyltransferase PRMT1, a protein that can shuttle between cytoplasm and nucleus (Figure 8d) (Cheung, Chan et al., 2007). Several substrates of PRMT1 have been described, including arginine 3 of histone 4 (H4R3). Interestingly, histone arginine methylation correlates with increased acetylation, which could indicate that MLL-EEN mediated PRMT recruitment shares similar molecular mechanisms of oncogenic transformation with MLL-CBP and MLL-p300 fusion proteins (Pal & Sif, 2007). Last but not least, certain subunits of the SWI/SNF chromatin-remodeling complex were described to contribute to MLL-rearranged leukemia development (Cruickshank, Sroczynska et al., 2015). Thus, MLL-fusion proteins are operating in the context of large multi-protein complexes whose composition determines their exact function.

1.12 Signaling pathways involved in MLL-fusion-mediated leukemogenesis

Proliferation and differentiation are tightly regulated by numerous signaling pathways. 70% of AML patients are estimated to carry genetic or epigenetic alterations that lead to constitutively active signaling of the PI3K/AKT/mTOR pathway, which is critical for the regulation of cell cycle

progression (Grandage, Gale et al., 2005, Kornblau, Tibes et al., 2009, Min, Cheong et al., 2004, Min, Eom et al., 2003, Xu, Simpson et al., 2003). Given the importance of PI3K/AKT/mTOR signaling pathway in hematopoiesis, targeting this axis has become an attractive target in AML. Pharmacological inhibition of AKT with MK-2260, mTORC1 with Rapamycin and PI3K/mTORC1/2 with BEZ-235 resulted in anti-leukemic effects that were particularly pronounced in cells harboring rearrangements of the *MLL* gene (Martelli, Tazzari et al., 2007, Sandhöfer, Metzeler et al., 2014). Some substrates of AKT are inactivated upon phosphorylation, such as the FOXO family of transcription factors, which in turn negatively regulate cell proliferation and survival. Active FOXOs shuttle to the nucleus and regulate the transcription of genes that induce cell cycle arrest and apoptosis. Elevated levels of FOXOs were shown to be necessary for the maintenance of leukemia initiating cells (LICs) *in vivo*. Moreover, FOXO deficiency resulted in differentiation and apoptosis of human MLL-AF9 leukemia cells (Sykes, Lane et al., 2011).

Furthermore, high expression of the FMS-like tyrosine kinase FLT3 is often found together with translocations of the *MLL* gene in AML and ALL (Armstrong, Golub et al., 2003, Armstrong, Kung et al., 2003, Libura, Asnafi et al., 2003, Stam, den Boer et al., 2005). Under physiologic conditions, binding of FLT3 ligand leads to receptor dimerization and phosphorylation and subsequent activation of downstream signaling pathways including PI3K/AKT, RAS/MAPK and Stat5 (Annesley & Brown, 2014). Internal tandem duplications (ITD) or point mutations within the tyrosine kinase domain (TKD) FLT3 are common in hematopoietic malignancies with poor prognosis and result in constitutive FLT3 activation (Annesley & Brown, 2014, Meshinchi, Woods et al., 2001). Transplantation experiments showed that FLT3-ITD cooperated with MLL-AF9 in inducing a lethal AML with a very short latency (Stubbs, Kim et al., 2008).

The Wnt signaling pathway is involved in the development of various types of cancer, including leukemia (Muller-Tidow, Steffen et al., 2004). β -catenin is the major mediator of the Wnt signaling pathway. Binding of Wnt ligands to the transmembrane receptor Frizzled leads to stabilization of β -catenin, its translocation from the cytoplasm to nucleus, interaction with transcription factors and transcription of Wnt-responsive target genes. MLL as well as other H3K4 methyltransferases were shown to be able to associate with β -catenin (Sierra, Yoshida et al., 2006). β -catenin is required for MLL-AF9-mediated transformation and down-regulation of β -catenin impaired the development of MLL-rearranged leukemia (Wang, Krivtsov et al., 2010, Yeung, Esposito et al., 2010). Moreover, activation of canonical Wnt signaling in MLL-rearranged cancer stem cells was implicated in development of resistance to pharmacological inhibition of BET bromodomain proteins, which is an emerging therapeutic strategy for patients harboring translocations of the *MLL* gene (Rathert, Roth et al., 2015). Interestingly, the activity of β -catenin has been shown to

be critical for the proliferation and survival of hematopoietic stem cells during fetal development while it is dispensable for adult HSCs (Cobas, Wilson et al., 2004, Jeannet, Scheller et al., 2008, Koch, Wilson et al., 2008, Reya, Duncan et al., 2003). Thus, targeting this signaling pathway might be an attractive and safe strategy of treatment for patients suffering from MLL-rearranged leukemia.

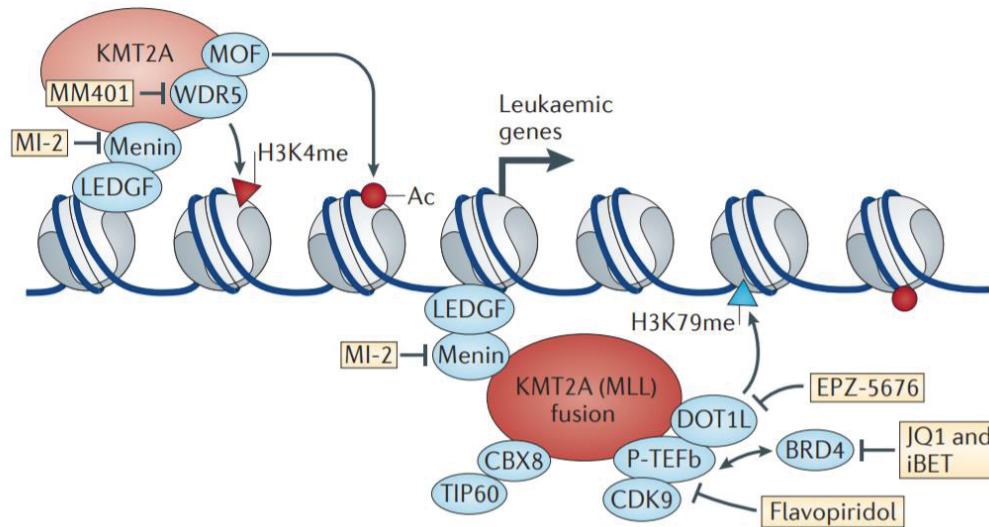
Finally, glycogen synthase kinase 3 (GSK3), which functions in a number of signaling pathways involved in human diseases, could endorse proliferation and leukemic transformation of cells harboring MLL-rearrangements via a mechanism that involved destabilization of the cyclin-dependent kinase inhibitor p27Kip1. Furthermore, inhibition of GSK3 with lithium chloride in a mouse model of MLL-rearranged leukemia induced anti-leukemic effects (Wang, Smith et al., 2008).

In summary, MLL-fusion proteins cooperate with cancer-associated deregulation of multiple signaling pathways, thus providing new targeting strategies for patients harboring MLL rearrangements.

1.13 Therapeutic strategies in MLL-rearranged leukemia

A number of proteins that associate with MLL-fusions have emerged as potential therapeutic targets. Recent efforts were invested in the identification of small molecule inhibitors of molecular mechanisms associated with MLL-fusions, but also in targeting cellular pathways whose inhibition might be synthetic lethal with the expression of MLL-fusion proteins (Figure 9).

Figure 9. Therapeutic targeting of MLL-fusion proteins. Small molecule inhibitors targeting different interactors within MLL-fusion complexes are shown (Taken from: Rao & Dou, 2015; Reprinted with permission from Springer Nature).



A number of publications showed that the enzymatic activity of DOT1L is critical for the pathogenesis of MLL-rearranged leukemia, endorsing DOT1L as a candidate target for therapeutic inhibition. DOT1L is able to catalyze methylation of ϵ -amino group of H3K79 using S-adenosyl-methionine (SAM) as a cofactor, producing methylated lysine and S-adenosyl-L-homocysteine (SAH). SAH has been shown to be a competitive inhibitor of DOT1L as well as of other SAM-dependent methyltransferases (Spurr, Bayle et al., 2016). Several small molecules resembling the structure of SAH have been tested for DOT1L inhibition, establishing EPZ-5676 (Pinometostat) as the most potent and selective DOT1L inhibitor (Daigle et al., 2013, Daigle, Olhava et al., 2011). Inhibition of DOT1L activity with EPZ-5676 leads to reduced levels of H3K79me₂ and concomitant reduction of MLL-target gene expression, including *HOXA9* and *MEIS1*. This ultimately releases the MLL-fusion dependent differentiation block and causes inhibition of cell growth *in vitro* and *in vivo* within ~14 days of treatment (Daigle et al., 2013). The effect is selective to cell lines harboring translocations of the *MLL* gene (Basavapathruni, Jin et al., 2012, Daigle et al., 2013, Deshpande et al., 2013). Promising results in pre-clinical studies have led to the initiation and successful completion of phase I clinical trials with EPZ-5676 (ClinicalTrials.gov identifiers: NCT02141828, NCT01684150) (Chen & Armstrong, 2015). However, the exact molecular mechanisms of DOT1L-mediated maintenance of leukemic gene expression in MLL-leukemia remains unclear. A recent genome-wide RNAi screen in MLL-AF9-expressing cells identified sirtuin1 (SIRT1) as an agonist of DOT1L in human leukemia. Inhibition of DOT1L was shown to lead to SIRT1-dependent decrease of acetylation of lysine 9 of histone 3 (H3K9) and up-regulation of methylation at the histone residue, resulting in decreased chromatin

accessibility (Chen, Koche et al., 2015). This discovery implicated that SIRT1 activation of SIRT1 together with DOT1L inhibition might constitute a novel therapeutic strategy for patients suffering from MLL-rearranged leukemia.

An alternative strategy of targeting MLL-rearranged leukemia is the inhibition of critical co-factors of MLL-fusion proteins. Among those, bromodomain containing protein 4 (BRD4) represents a promising target. BRD4 was discovered as a non-oncogene addiction target through an RNAi screen in an MLL-AF9; NrasG12D AML mouse model. Another research group identified BRD4 as a target in MLL-rearranged leukemia through a global proteomic approach (Dawson, Prinjha et al., 2011, Zuber, Shi et al., 2011c). BRD4 was shown to activate transcription through recruitment of P-TEFb within the super elongation complex and to be essential for the expression of critical hematopoietic transcription factors (Jang, Mochizuki et al., 2005, Roe, Mercan et al., 2015, Yang, Yik et al., 2005). Several potent small-molecule inhibitors of BRD4 have been developed. These inhibitors bind to the bromodomain of BRD4, and block its interactions with acetylated chromatin. This leads to inhibition of P-TEFb recruitment decreased expression of BRD4 target genes such as MYC, BCL2, and PIM1. Small-molecule inhibition of BRD4 phenocopies shRNA-mediated BRD4 downregulation, resulting in differentiation of leukemia cells and suppressed leukemia development *in vivo* (Dawson et al., 2011, Filippakopoulos, Qi et al., 2010, Zuber et al., 2011c). Inhibitors of BRD4 have been included in clinical trials for a variety of cancers and phase I clinical trials in AML patients have been completed (ClinicalTrials.gov identifiers: NCT02308761, NCT01587703). Given the success of early clinical trials, a careful evaluation of possible resistance mechanisms became necessary to optimize the clinical efficacy of these drugs. RNAi screening in MLL-AF9;G12D AML mouse model implicated increased WNT/ β -catenin signaling as a driver of resistance to BRD4 inhibition events (Fong, Gilan et al., 2015, Rathert et al., 2015).

Alternatively, several inhibitors have been designed to target CDK9, which represents the enzymatically active subunit of P-TEFb. Flavopiridol (also known as alvocidib) inhibits phosphorylation of the RNA Pol II C-terminal domain and leads to downregulation of CDK9 target genes, such as MYC and HOXA9, highlighting its potential as novel treatment alternative for patients with mixed-lineage leukemia (Garcia-Cuellar, Fuller et al., 2014).

Alternative strategies aim to disrupt critical protein-protein interactions within MLL-fusion protein complexes. The N-terminal portion of MLL involving the LEDGF-Menin binding interface and the CxxC regions are present in all MLL-fusion proteins and are necessary for the transforming capacity of MLL-fusion proteins. Both Menin and LEDGF are necessary for MLL-fusion mediated leukemic transformation and for the expression of MLL target genes such as *HOXA9* and *MEIS1*

(El Ashkar, Schwaller et al., 2018, Yokoyama & Cleary, 2008, Yokoyama, Somervaille et al., 2005). Therefore, small molecules targeting the interaction of MLL with its binding partners have been developed. MLL-fusion mediated oncogenic transformation could be blocked by outcompeting LEDGF from the ternary complex without inhibition of the MLL-Menin interaction, highlighting the LEDGF-MLL interface as a potential target in MLL-rearranged leukemias (Cermakova, Tesina et al., 2014, Mereau, De Rijck et al., 2013).

Inhibition of the Menin-MLL interaction with small molecules such as MI-463 and MI-503 or peptidomimetics such as MCP-1 led to growth arrest, differentiation and downregulation of MLL-fusion target genes in cells harboring MLL-rearrangements (Cierpicki & Grembecka, 2014, Grembecka, He et al., 2012, He, Malik et al., 2015). Similarly, targeting of the MLL-WDR5 interaction with peptidomimetics MM401 and MM589 were shown to impair the assembly of the MLL complex, leading to terminal myeloid differentiation and apoptosis of MLL-rearranged cells, without toxicity to normal, non-rearranged hematopoietic cells (Cao, Townsend et al., 2014, Karatas, Townsend et al., 2013).

Even though the presented strategies show robust effects in experimental and pre-clinical setups, further advancements as well as combinatorial approaches may lead to more efficient targeting of MLL-leukemia. For instance, combined inhibition of the Menin-MLL-fusion interaction and inhibition of DOT1L resulted in enhanced differentiation of leukemic cells (He, Malik et al., 2016). Similarly, simultaneous small-molecule inhibition of DOT1L and BRD4 resulted in synergistic antiproliferative effects, providing a rationale for future combination epigenetic therapies in AML (Gilan, Lam et al., 2016).

1.14 Experimental tools for functional characterization of MLL-fusions interactors

1.14.1 Biochemical characterization

In the last decades, mass spectrometry (MS)-based proteomics provided insights into numerous biological processes and signal transduction pathways, allowing deeper understanding of molecular biology. Advancements in the steps of data acquisition and overall reduction of time necessary to process the samples by MS machines enabled the design of complex experiments and resulted in high quality results (Aebersold & Mann, 2003, Aebersold & Mann, 2016). Different proteomics workflows have been established, allowing for the identification and quantification of protein expression, post-translational modifications (PTMs) and protein-protein interactions.

Expression proteomics aims to characterize a complete, expressed proteome of a cell at a given time point and biological state. Initially, only highly expressed proteins could be detected by this method. Recently, improvements resulted in higher sensitivity and resolution, allowing for the characterization of entire human proteomes (Baker, Ahn et al., 2017, Elguoshy, Magdeldin et al., 2016, Kim, Pinto et al., 2014, Omenn, Lane et al., 2017, Wilhelm, Schlegl et al., 2014). Changes in protein abundance can be assessed using several quantification strategies. Label-free quantification associates spectral counts with protein abundance without the need for additional protein labelling. Stable isotope labeling with amino acids (SILAC) introduces heavy variants of amino acids into cell culture media, while isobaric tags for relative and absolute quantitation (iTRAQ) or tandem mass tags (TMT) approaches are used to chemically label digested peptides (Ross, Huang et al., 2004, Thompson, Schafer et al., 2003). Such large datasets can be integrated and depicted as protein networks that provide novel insights about the function and evolution of biological systems in a comprehensive and quantitative fashion (Kocher & Superti-Furga, 2007). Moreover, unbiased and global profiling of various post-translational modifications became possible. For instance, phosphoproteomics allows the global investigation of phosphorylation states of proteins in cell populations *in vitro* and *in vivo* via enrichment of peptides with phosphorylated residues of serine, threonine and tyrosine using ion metal chromatography (IMAC) (Karisch, Fernandez et al., 2011, Olsen, Blagoev et al., 2006). Alternatively, peptides containing phosphotyrosine (pY) can be enriched by immune-precipitation prior to MS (Abe, Nagano et al., 2017). Moreover, titanium dioxide (TiO₂) became an attractive method for phosphopeptide enrichment due to its high affinity for phosphopeptides (Thingholm & Larsen, 2016).

Interaction proteomics assumes that intracellular proteins do not act in isolation but are embedded in a network of protein-protein interactions which contribute to defined cellular functions. Affinity purification coupled to mass spectrometry (AP-MS) emerged as a powerful tool for the characterization of protein-protein interaction networks of any protein of interest in a global fashion. It involves affinity tagging of key players in a complex assembly and the use of biochemical methods to purify their entire protein interactome. An affinity tag can be introduced using standard cloning techniques and the tagged protein can be expressed in target cells in a transient or stable manner. Available tagging systems offer a wide variety of tags and their combinations were used in different configurations (Li, 2010). For instance, tandem affinity purifications (TAP) strategy utilizes two consecutive enrichment steps employing affinity purifications performed consecutively with two molecular tags, such as STREP- and HA-tags, FLAG-HA-tags or protein A-CBP tags (Burckstummer, Bennett et al., 2006, Giambruno, Grebien

et al., 2013, Glatter, Ahrne et al., 2015). As the experiments are performed in mild buffer conditions, the composition of eluted complexes often remains intact. The obtained protein samples are subjected to proteolytic digestion using trypsin and resulting peptides are analyzed by MS (Kocher & Superti-Furga, 2007, Li, 2011). However, each protocol presents with some drawbacks. Usually, results obtained by tandem AP-MS are clean, but due to extensive experimental procedures weak interactions can be lost. In contrast, single-step purifications allow the identification of a higher number of contaminating proteins. Both approaches have been compared in the context of hematopoietic transcription factors (Giamb Bruno et al., 2013). The study shows that the exact protocol for sample processing depends on the affinity tag used in the experiments and highlights the importance of combining a minimum of two independent tag purifications for sufficient robustness of a complex analysis.

Several bioinformatic tools have been optimized to precisely assess the significance and reproducibility of the data within an AP-MS dataset, including CompPASS and SAINT (Choi, Larsen et al., 2011, Jackson, 2009). The contaminant repository for AP-MS experiments (CRAPome), represents a database that can be used for better discrimination between bona fide interactors and background contaminants (Mellacheruvu, Wright et al., 2013). Finally, different software packages such as Gephi and Cytoscape enable easy visualization and manipulation of protein-protein interaction networks (Bastian M., 2009, Shannon, Markiel et al., 2003).

Several strategies have been employed to characterize and validate the architecture of protein complexes engaged by wild type MLL and various MLL-fusion proteins. Valuable mechanistic insights into the role of MLL-fusion proteins as well as MLL-fusion mediated disease were obtained from purifications of MLL-associated complexes from cell lines overexpressing FLAG-HA-tagged MLL-AF4, MLL-AF9 and MLL-ELL (Biswas, Milne et al., 2011). A pipeline combining different proteomic approaches followed by functional validation led to the identification of BRD4 as a critical target of MLL-fusion proteins (Dawson et al., 2011). A quantitative approach was used to investigate the composition of human SET1/MLL histone methyltransferase complexes contributing valuable knowledge to the H3K4 methylation landscape in human cells (van Nuland, Smits et al., 2013). A MS-based approach contributed to characterize the interaction between MLL and the polymerase associated factor complex (PAFc) (Muntean et al., 2010). In conclusion, a large number of studies has employed MS for better characterization of proteins involved in MLL-rearranged leukemia, highlighting the importance of this technology in understanding of this disease.

1.14.2 Genetic characterization

The past decades established RNA interference (RNAi) as a robust and powerful genetic tool to survey gene function. The first references addressing RNAi date back to the last years of the 20th century when several experiments conducted in model organisms showed that homologous RNA sequences can lead to down-regulation of mRNA levels of the endogenous gene (Napoli, Lemieux et al., 1990). Extended studies in *C. elegans* led to the discovery of the RNAi pathway that is used for gene silencing by double stranded RNA and were awarded with the Nobel Prize for Physiology or Medicine in 2006 (Nobel Media AB, 2014).

Over the last two decades, our understanding of RNA biology established short RNA molecules as important regulatory units of eukaryotic genomes. While short interfering RNAs (siRNAs) are able to process exo- and endogenous double strand RNAs (dsRNAs) of viral origin, mammalian expression of over 1000 distinct micro-RNAs (miRNAs) is responsible for a large variety of regulatory functions (Carthew & Sontheimer, 2009).

The endogenous micro-RNA (miRNA) pathway starts in the nucleus, where Pol II- miRNA precursors (termed pri-miRNAs) are generated by Pol II and cleaved by the nuclease Drosha to produce ~100 bp long pre-miRNAs. Exportin 5 (XPO5) mediates export of the pre-miRNAs to the cytoplasm where another nuclease enzyme, Dicer, further processes them to mature duplexes of miRNAs of ~21-25 bp in length. These small fragments can recruit argonaute proteins (Ago) and assemble into the RNA-induced silencing complex (RISC). Mature siRNA duplexes contain the passenger strand, which is degraded in later stages of the miRNA pathway, while the guide strand targets the RISC complex to homologous regions within cellular mRNAs, leading to their cleavage, destabilization or translational repression (Fellmann & Lowe, 2014).

RNAi has become a fundamental tool for loss-of-function studies. Short hairpin RNA (shRNA) sequences, designed to be fully complementary to any gene of interest, are embedded into retro- or lentiviral vector backbones. Their stable or inducible expression fuels endogenous miRNA processing pathways to generate mature siRNAs and to suppress the expression of virtually any gene. Initially, shRNA expression was driven from RNA Polymerase III (Pol III) promoters and led to the production of short stem-loop structures. These products would enter the miRNA pathway at the stage of pre-miRNAs (Brummelkamp, Bernards et al., 2002). This strategy turned out to be sub-optimal due to cellular toxicity of stem-loop structures as well as due to difficulties in their processing by enzymatic components of the miRNA maturation pathway (Grimm, Streetz et al., 2006, Gu, Jin et al., 2012). These setbacks have been overcome by using endogenous miRNA scaffolds such as miR-30 or miR-155 to embed shRNA sequences. Further optimization resulted

in the generation of improved miR-E scaffolds, which were designed to enhance pre-miRNA processing, boosting mature miRNA levels up to 30 fold over the original miR-30 backbone (Fellmann, Hoffmann et al., 2013).

Large shRNA libraries can be designed to target all genes in any genome of interest and can be delivered to target cells via viral transduction, allowing for stable integration of the shRNA into the genome. A large number of shRNA screens have been performed, employing various modifications of the method. For instance, expression of shRNAs can be coupled to fluorescent reporters, thus allowing to track shRNA expression through FACS based assays. Moreover, available systems work robustly not only *in vitro*, but also *in vivo*. In this case, shRNAs are introduced into cultured cells prior to transplantation into recipient mice. This approach was particularly successful in the hematopoietic system (Zuber, Radtke et al., 2009, Zuber, Rappaport et al., 2011b).

Overall, shRNA screening approaches have contributed essential information about cancer vulnerabilities as well as insights into the regulation of numerous signaling pathways. Multiple targets critical to the biology of MLL-rearranged leukemia have been identified by this strategy, including BRD4, RNF20, ITGB3, CDK6, JMJD1C, ZNF521 or protein kinase Msk1 (Germano, Morello et al., 2017, Miller, Al-Shahrour et al., 2013, Placke, Faber et al., 2014, Sroczynska, Cruickshank et al., 2014, Wang, Kawaoka et al., 2013, Wiersma, Bussiere et al., 2016, Zuber et al., 2011c) Furthermore, modifications of the RNAi technology have been employed to decipher dependencies and synthetic lethal relationships in nearly 400 different cancer cell lines (McDonald, de Weck et al., 2017).

More recently development of the CRISPR/Cas9 technology has revolutionized the approach of genome engineering. Initially described as a bacterial immune mechanism for defense against pathogens, the system was repurposed for genome editing and targeted mutagenesis (Cong, Ran et al., 2013, Mali, Yang et al., 2013). The CRISPR/Cas9 system utilizes the Cas9 nuclease, an enzyme initially discovered in *Streptococcus pyogenes*, and a single guide RNA (sgRNA). The Cas9 nuclease binds the sgRNA, which directs the complex to the complementary regions within the genome. Double stranded target DNA is recognized by a three-nucleotide length protospacer adjacent motive (PAM) NGG sequence located at the 3' end of the recognition site. Subsequently, the Cas9 nuclease introduces the formation of a double strand break at this site (Shalem, Sanjana et al., 2014). CRISPR/Cas9 mediated mutagenesis became a versatile tool to introduce genomic mutations, as it can easily lead to homo- or heterozygous loss-of-function of the targeted gene. Similar to the RNAi technology, large, genome-wide libraries of sgRNAs have been generated

and are being applied in complex multiplexed screening approaches (Sanjana, Shalem et al., 2014, Shalem et al., 2014).

Given the primary application of the CRISPR/Cas9 system in loss-of-function experiments *in vitro* and *in vivo*, the method contributed to a number of interesting findings, also in the field of hematology. For instance, a CRISPR dropout screen was employed to identify genetic vulnerabilities and therapeutic targets in AML (Tzelepis, Koike-Yusa et al., 2016, Wang, Yu et al., 2017). The same method was used to successfully introduce translocations of the *MLL* gene in human hematopoietic progenitor cells *in vivo* (Buechele, Breese et al., 2015, Reimer, Knoss et al., 2017). Furthermore, a CRISPR/Cas9 screen lead to identification of novel molecular targets that sensitize MLL-rearranged cell population to DOT1L inhibition (Chen, Delaney et al., 2016).

1.15 Aims of this thesis

Acute Myeloid Leukemia frequently harbors chromosomal rearrangements involving the Mixed Lineage Leukemia (*MLL*) gene. Up to date, more than 75 different MLL-fusion genes have been identified and many of them have been described to act as strong cancer drivers. All MLL-fusion proteins were proposed to be embedded into large protein complexes comprising modules with various functions. While critical effectors of several distinct MLL-fusion proteins were identified, it is not clear whether transforming mechanisms are conserved across the entire family of MLL-fusions. We hypothesized that common oncogenic mechanisms are encoded in stable physical and genetic MLL-fusion-specific interaction networks. We aimed to identify common critical effectors of different MLL-fusion proteins that are presumed to employ different mechanisms of oncogenic transformation. Thus, work during this thesis was centered around the following aims:

1. Characterization of the interactome of 7 selected MLL-fusion proteins through AP-MS.
2. Bioinformatic annotation of conserved proteins associated with at least five out of seven selected MLL-fusion proteins among the MLL-fusion-interactome
3. Functional validation of conserved interactors to identify high confidence effectors of distinct MLL-fusion proteins through RNAi screening.
4. Obtaining new insights into molecular mechanisms leading to development of MLL-rearranged leukemia

2. Results

2.1 Prologue

MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity

Anna Skucha, Jessica Ebner, Johannes Schmöllerl, Mareike Roth, Thomas Eder, Adrián César-Razquin, Alexey Stukalov, Sarah Vittori, Matthias Muhar, Bin Lu, Martin Aichinger, Julian Jude, André C. Müller, Balázs Györfy, Christopher R. Vakoc, Peter Valent, Keiryn L. Bennett, Johannes Zuber*, Giulio Superti-Furga*, Florian Grebien*

* equal contribution

lead contact: Florian Grebien

Here we describe a combined proteomic-genomic cellular screening approach which was established to identify critical conserved effectors of MLL-fusion proteins.

Characterization of protein complexes around seven selected MLL-fusion proteins by AP-MS revealed a densely interconnected protein-protein interaction network of 963 proteins, of which 128 proteins were found to interact with at least five of all seven MLL-fusions. Systematic functional investigation of the conserved MLL-fusion interactome using subtractive shRNA screens identified the methyltransferase SETD2 as a critical effector of MLL-fusion proteins. shRNA- and CRISPR/Cas9-mediated downregulation of *SETD2* led to induction of DNA-damage, myeloid differentiation and apoptosis of leukemic blasts *in vitro* and *in vivo*. In summary, this study establishes a novel role for SETD2 in the maintenance of genomic integrity during initiation and progression of MLL-rearranged AML.

Nature Communications

Accepted for publication: 16.03.2018

Detailed contribution of authors is listed below:

A.Skucha planned, performed and analyzed most experiments; J.E. conducted interaction studies, performed apoptosis assays and cell cycle analysis, performed CRISPR/Cas9-mediated mutagenesis experiments in human MLL-rearranged cell lines and data analysis; J.S. performed *in vivo* experiments; M.R. contributed constructs and cell lines used throughout the manuscript, and helped with *in vivo* experiments; T.E. performed bioinformatic analysis of NGS datasets; A.C-

R. and A. St. performed bioinformatic analysis of the protein-protein interaction network; S.V. prepared cell lines used for the AP-MS approach; M.M. contributed constructs and cell lines used throughout the manuscript and helped to establish FACS-based assays used in the study; B.L. and C.R.V. contributed stable Cas9-expressing cell lines; M.A. and J.J. contributed constructs and cell lines used throughout the manuscript; A.C.M., helped to plan experiments, processed samples, and performed MS analysis; B.G. performed bioinformatic analysis of publicly available datasets; P.V. provided access to patient materials and contributed to writing of the manuscript; K.L.B. provided access to MS technology, helped to plan experiments and contributed to writing of the manuscript; F.G. G.S-F. and J.Z. designed the study and planned experiments, F.G. wrote the manuscript. All co-authors read and commented on the manuscript.

2.2 MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity

ARTICLE

DOI: 10.1038/s41467-018-04329-y

OPEN

MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity

Anna Skucha^{1,2}, Jessica Ebner², Johannes Schmöller^{1,2}, Mareike Roth³, Thomas Eder², Adrián César-Razquin¹, Alexey Stukalov¹, Sarah Vittori¹, Matthias Muhar³, Bin Lu⁴, Martin Aichinger³, Julian Jude³, André C. Müller¹, Balázs Györffy⁵, Christopher R. Vakoc⁴, Peter Valent⁶, Keiryn L. Bennett¹, Johannes Zuber³, Giulio Superti-Furga^{1,7} & Florian Grebien^{2,8}

MLL-fusions represent a large group of leukemia drivers, whose diversity originates from the vast molecular heterogeneity of C-terminal fusion partners of MLL. While studies of selected MLL-fusions have revealed critical molecular pathways, unifying mechanisms across all MLL-fusions remain poorly understood. We present the first comprehensive survey of protein-protein interactions of seven distantly related MLL-fusion proteins. Functional investigation of 128 conserved MLL-fusion-interactors identifies a specific role for the lysine methyltransferase SETD2 in MLL-leukemia. SETD2 loss causes growth arrest and differentiation of AML cells, and leads to increased DNA damage. In addition to its role in H3K36 tri-methylation, SETD2 is required to maintain high H3K79 di-methylation and MLL-AF9-binding to critical target genes, such as *Hoxa9*. SETD2 loss synergizes with pharmacologic inhibition of the H3K79 methyltransferase DOT1L to induce DNA damage, growth arrest, differentiation, and apoptosis. These results uncover a dependency for SETD2 during MLL-leukemogenesis, revealing a novel actionable vulnerability in this disease.

¹CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna 1090, Austria. ²Ludwig Boltzmann Institute for Cancer Research, Vienna 1090, Austria. ³Research Institute of Molecular Pathology, Vienna 1030, Austria. ⁴Cold Spring Harbor Laboratory, Cold Spring Harbor, 11724 NY, USA. ⁵MTA TTK Lendület Cancer Biomarker Research Group, Institute of Enzymology, Second Department of Pediatrics, Semmelweis University, Budapest 1094, Hungary. ⁶Department of Internal Medicine I. Division of Hematology and Hemostaseology, Ludwig Boltzmann Cluster Oncology, Medical University of Vienna, Vienna 1090, Austria. ⁷Center for Physiology and Pharmacology, Medical University of Vienna, Vienna 1090, Austria. ⁸Institute for Medical Biochemistry, University of Veterinary Medicine, Vienna 1210, Austria. These authors contributed equally: Johannes Zuber, Giulio Superti-Furga, Florian Grebien. Correspondence and requests for materials should be addressed to F.G. (email: florian.grebien@vetmeduni.ac.at)

Leukemia-associated fusion proteins serve as a paradigm for modern cancer research, as the molecular machineries that provide their functional cellular context have emerged as amenable to targeted molecular approaches^{1,2}. Families of related leukemia fusion proteins that share genomic and biological properties represent unique opportunities to study how the combination of distinct functional protein modules can drive oncogenic transformation. The largest family of “multi-partner translocations” in acute leukemia comprises fusions involving the product of the *KMT2A* (*MLL*) gene. *MLL*-fusion proteins are found in acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) and are often associated with adverse prognosis, particularly in pediatric patients³. Expression of *MLL*-fusions enhances proliferation and blocks myeloid differentiation of hematopoietic progenitor cells, leading to their pathological accumulation. In line, many *MLL*-fusions can act as potent oncogenes in cell line models and animal models of leukemia⁴.

In leukemia, the *MLL* N-terminus takes part in >120 different translocations, resulting in the generation of *MLL*-fusion proteins encompassing more than 75 different partner genes⁵. It has therefore been proposed that the oncogenic activity of *MLL*-fusion proteins depends on chromatin targeting functions exerted by the *MLL* N-terminus in combination with other functional properties encoded by the fusion partners⁶. Several regions in the *MLL* N-terminus are critical for the activity of *MLL*-fusions. For instance, the CxxC-domain is essential for DNA binding of *MLL*-fusion proteins⁷. Furthermore, the *MLL*-interacting protein Menin links *MLL*-fusion proteins with LEDGF, and the H3K36me3-binding PWWP domain of LEDGF is critical for the function of *MLL*-fusions⁸. In fact, a direct fusion of the LEDGF PWWP domain to *MLL* was able to replace Menin altogether⁹.

Numerous studies have established strong links between the molecular function of the fusion partner and the mechanistic basis of oncogenic transformation in *MLL*-fusion-induced leukemogenesis⁴. Pioneering biochemical experiments have shown that several fusion partners of *MLL*, such as AF4, AF9, and ENL are members of the DOT1L complex (DotCom) and the super-elongation complex (SEC)^{10–13}, which are both involved in transcriptional control. As the SEC can regulate the transcriptional activity of RNA polymerase II, it was hypothesized that these *MLL*-fusions induce aberrant regulation of transcriptional elongation on *MLL*-target genes¹⁴.

A large number of factors was shown to influence the oncogenic properties of *MLL*-fusions, including signaling proteins^{15–17}, epigenetic modulators^{18–21}, and transcription factors^{22–24}, as well as the wild-type *MLL* protein²⁵. However, it is unclear whether these molecular mechanisms pertain to the entire family of *MLL*-fusions or if they specifically affect the leukemogenicity of isolated *MLL*-fusion proteins. In fact, specific molecular mechanisms of oncogenic transformation were postulated to prevail for selected *MLL*-fusions. For instance, inhibition of the arginine methyltransferase PRMT1 was shown to reduce the leukemic potential of several oncogenic fusion proteins, including *MLL*-EEN and *MLL*-GAS7, but not *MLL*-AF9, *MLL*-AF10, or *MLL*-ENL^{26,27}. Furthermore, the enzymatic activity of CBP was shown to be required for leukemogenic activity of fusions of *MLL* with the histone acetyltransferase CREBBP^{28,29}. Finally, dimerization might play an important role in nuclear translocation and oncogenic transformation in fusions of *MLL* to the cytoplasmic partner proteins GAS7 and AF1p, yet the underlying molecular mechanism is unclear^{30,31}.

Here, we set out to survey the molecular composition of a diverse subset of distantly related *MLL*-fusion protein complexes to characterize their unique and common properties, and to reveal possible actionable vulnerabilities that are based on specific molecular mechanisms shared by *MLL*-fusions. We identify the

methyltransferase SETD2 as an interactor of all *MLL*-fusion proteins. shRNA-mediated and CRISPR/Cas9-mediated loss of SETD2 leads to growth arrest and differentiation of *MLL*-fusion-expressing cells in vitro and in vivo. Moreover, we show that loss of SETD2 is associated with increased DNA damage. SETD2 loss disrupts a H3K36me3-H3K79me2 signature on *MLL*-target genes and sensitizes *MLL*-AML cells to pharmacologic inhibition of the known *MLL*-fusion protein effector DOT1L. In summary, we describe a novel dependency for SETD2 in the initiation and maintenance of *MLL*-rearranged leukemia, highlighting a novel vulnerability in this disease.

Results

Functional proteomic survey of *MLL*-fusion proteins. Reasoning that critical effectors might be enriched among the physical interaction partners of distantly related *MLL*-fusion proteins, we undertook an unbiased survey of the protein–protein interactions of *MLL*-fusion proteins in leukemia cells. Using FRT/Flp-mediated locus-specific cassette exchange, we generated isogenic Jurkat leukemia cell lines allowing for Doxycycline (Dox)-inducible, single-copy expression of affinity-tagged variants of seven *MLL*-fusions that were previously proposed to employ different molecular mechanisms of oncogenic transformation (*MLL*-AF1p, *MLL*-AF4, *MLL*-AF9, *MLL*-CBP, *MLL*-EEN, *MLL*-ENL, *MLL*-GAS7, Fig. 1a, b and Supplementary Fig. 1a–c). Subcellular fractionation revealed that all selected *MLL*-fusion proteins localized to the nucleus (Supplementary Fig. 1d) and were capable of inducing expression of the *MLL*-fusion-target genes *HOXA5*, *HOXA9*, *HOXA10*, and *MEIS1* (Fig. 1c).

Protein complexes around *MLL*-fusion proteins were purified from nuclear lysates of cell lines expressing seven distinct *MLL*-fusions (Fig. 1d and Supplementary Fig. 2a). Purifications were analyzed by LC-MS/MS using both one-dimensional and two-dimensional gel-free proteomic approaches, recovering 4600 proteins in total, engaging in 15,094 putative interactions (Fig. 1e and Supplementary Fig. 2b)^{32–34}. *p*-value-based filtering for the 300 most significant interactions per *MLL*-fusion resulted in a network of 960 high-confidence cellular proteins (Supplementary Fig. 2b). Validation of the network confirmed previously reported interactions of *MLL*-fusions with protein complexes important for transcriptional control and epigenetic regulation, including the PAF complex, the SWI/SNF complex, and Polycomb Repressor Complex 1 (Supplementary Fig. 2c)^{12,35–37}. The network also revealed abundant unique interaction partners of all *MLL*-fusion proteins, indicating that distinct *MLL*-fusions can engage specific molecular pathways. 406 proteins in the network (42.3%) co-purified with more than one *MLL*-fusion protein, while 128 proteins (13.3%) interacted with at least five of the seven *MLL*-fusions (Supplementary Fig. 2b), indicating a strong degree of topological conservation within the *MLL*-interaction network. Further analysis of the 128 conserved partners of *MLL*-fusion proteins revealed six distinct protein communities (*p* < 0.01; Supplementary Table 1), whose annotation retrieved molecular functions that are highly relevant to the biology of *MLL*-fusion proteins, including chromatin remodeling, transcriptional elongation, and hematopoiesis (Fig. 1e). Interestingly, protein families that had not been reported to interact with *MLL*-fusion proteins before were also identified, such as factors involved in DNA-repair, RNA splicing, and RNA transport.

In summary, our comprehensive analysis and validation of the cellular interaction networks shows that distinct *MLL*-fusion proteins engage in a high number of direct, as well as indirect protein–protein interactions. Structurally different *MLL*-fusion proteins share 128 conserved interaction partners, which are

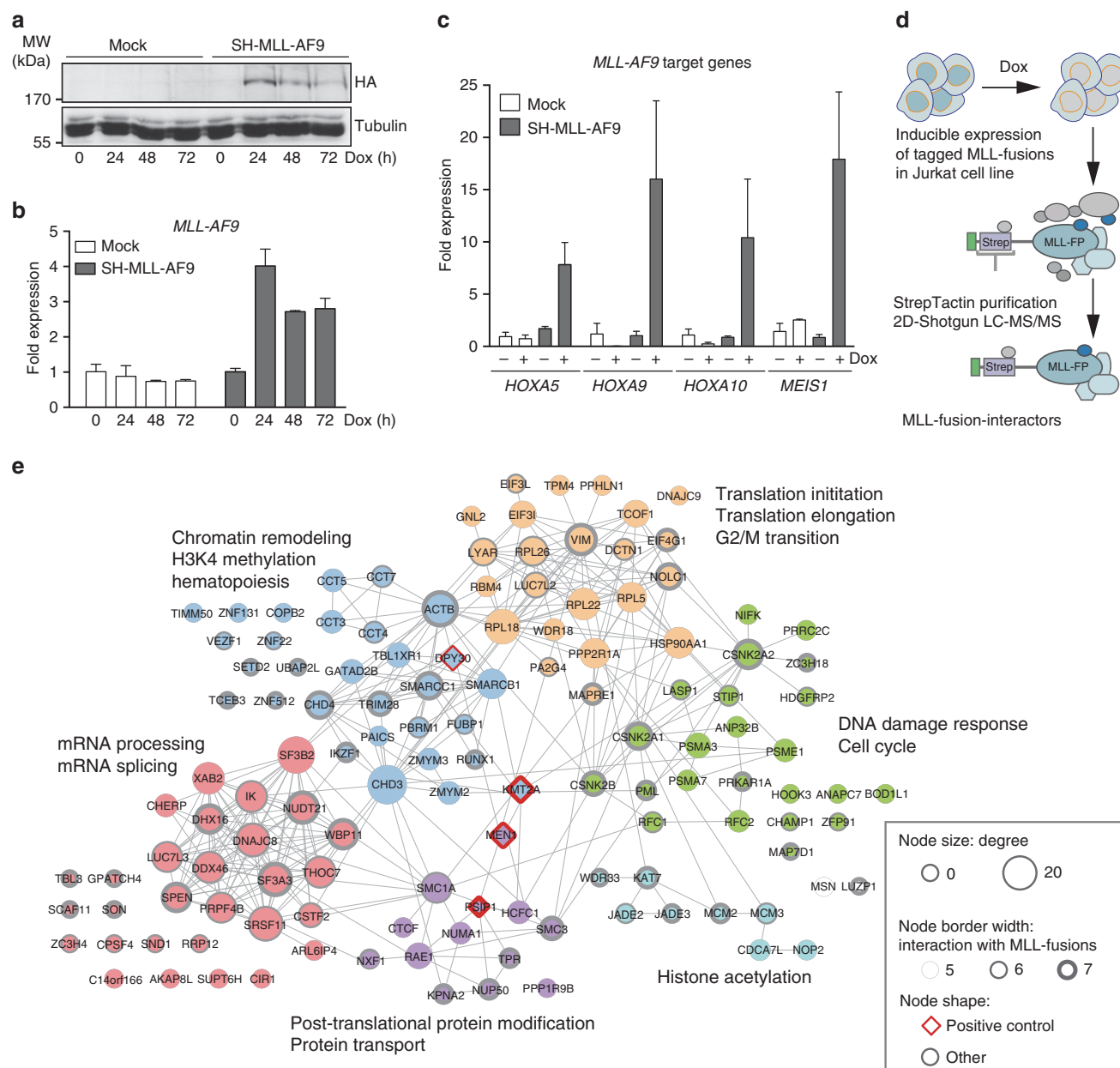


Fig. 1 Functional proteomic survey of the MLL-fusion interactome. Cells expressing Strep-HA (SH)-tagged MLL-AF9 or mock-transfected cells were treated with Dox for the indicated time points and transgene expression was monitored by immunoblotting (**a**) and qPCR (**b**) (means \pm s.d. $n = 3$). **c** SH-MLL-AF9-expressing cells were treated with Dox for 24 h and the expression of indicated MLL-target genes was measured by qPCR (mean \pm s.d. $n = 3$). **d** Schematic illustration of the strategy of affinity purification of protein complexes associated with MLL-fusion proteins from nuclear lysates of cell lines expressing affinity-tagged MLL-fusion proteins. **e** Gene ontology (GO) enrichment of six distinct protein communities among the core 128 interactors shared by at least 5 of 7 MLL-fusion proteins

enriched in six functional communities that are highly relevant for the biology of MLL-fusion proteins.

shRNA screen identifies SETD2 as an effector of MLL-fusions.

As our primary validation reduced the number of potential critical effectors in the network of MLL-fusion protein-interactors from 960 to 128, we next aimed to further narrow down the circle of candidate proteins using sequential functional genomic approaches (Fig. 2a). To systematically investigate the functional contribution of the conserved 128 MLL-interaction partners to MLL-fusion-dependent leukemia, we devised a shRNA screen in the human MLL-AF9-expressing AML cell line MOLM-13. In the

system used by us, transcriptional coupling of fluorescent reporter proteins to shRNA expression upon Dox-induction allows for dynamic monitoring of competing growth kinetics in mixed cell populations expressing experimental shRNAs (GFP) vs. non-targeting control shRNAs (shRen.713, dsRed, Fig. 2b). While expression of a control shRNA did not differentially affect cell proliferation in mixed populations over time, strong shRNA-induced negative selection of GFP-positive cells was observed upon targeting of *MEN1*, an interaction partner of all seven investigated MLL-fusion proteins with a well-known function in MLL-fusion-induced leukemogenesis³⁸ (Fig. 2b, bottom). We used this setup to systematically test the effects of 128 shRNA-pools targeting conserved MLL-fusion interaction partners on

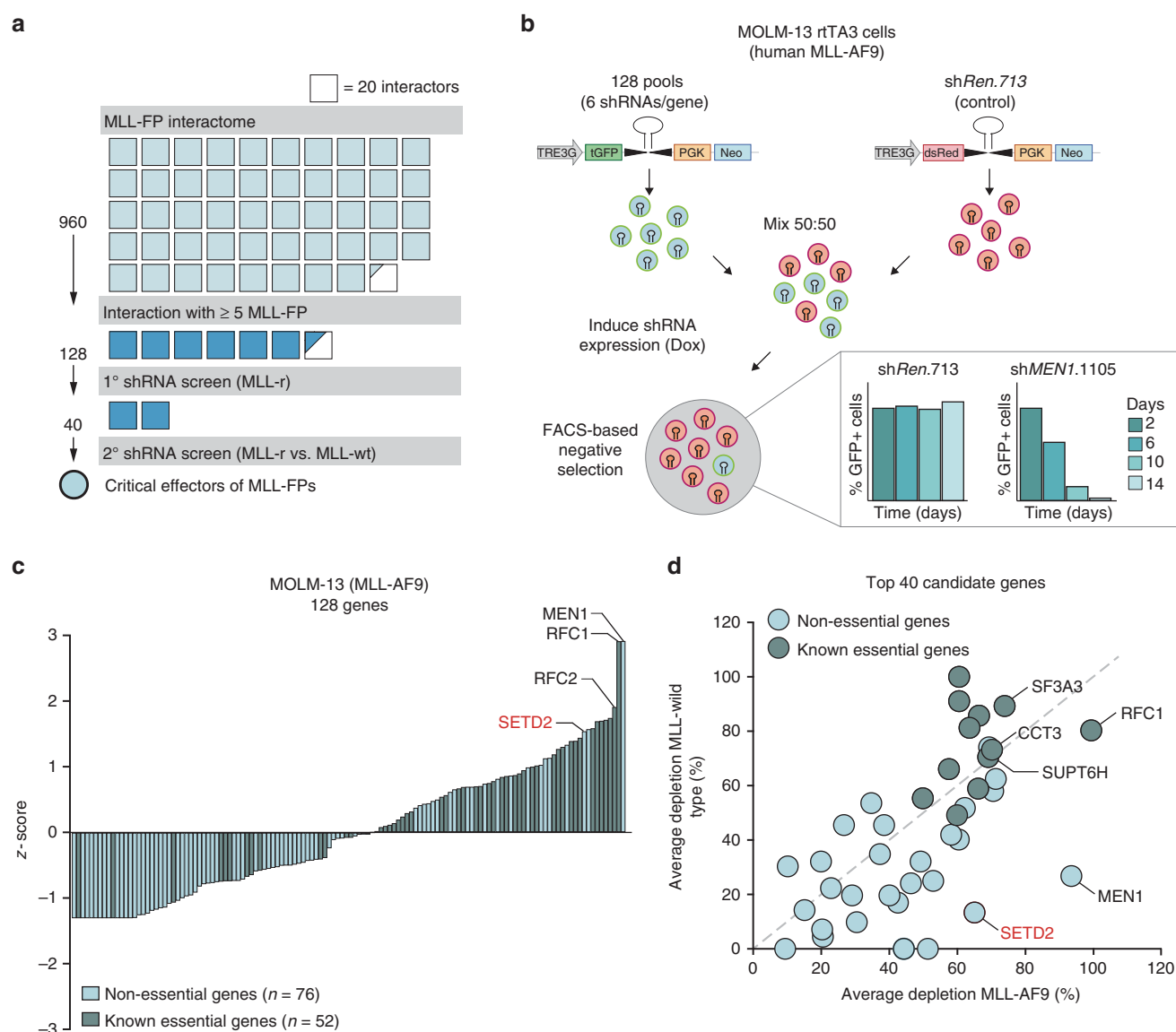


Fig. 2 shRNA screen identifies SETD2 as a critical effector of MLL-fusions. **a** Schematic representation of the filtering strategy. Affinity purification coupled to mass spectrometry identified 960 candidate genes (top 300 interactors per bait, ranked by *p*-value) to interact with at least one of seven selected MLL-fusion proteins. 128 proteins interacted with ≥ 5 of seven MLL-fusions. 40 candidate genes were screened in MLL-rearranged vs. MLL-wild-type cell lines. Each square corresponds to 20 interactors. **b** Schematic outline of retroviral vectors and experimental design of the FACS-based negative selection RNAi screen. Competitive proliferation assays were set up by mixing cells in a 50:50 ratio (experimental-GFP vs. control-dsRed) and cultivation in the presence of Dox. The relative ratio of GFP-positive vs. dsRed-positive cells was monitored by flow cytometry over 14 days. Bar graphs (bottom) represent the performance of positive (shMEN1.1105) and negative (shRen.713) controls shown as percentage of GFP+ cells over time. **c** Summary of RNAi screening data in the MOLM-13 cell line. Positive and negative *z*-scores correspond to candidate genes with stronger and weaker depletion phenotypes. Gene essentiality was assigned based on published datasets. **d** Average depletion values obtained from two subsequent RNAi screens performed in the MLL-AF9-expressing MOLM-13 cell line are plotted against mean depletion values from counterscreens performed in two MLL-wild-type cell lines (K562 and HL-60)

AML cell growth. Relative depletion of all shRNA-pools was normalized to a negative-control shRNA (shRen.713) and to a strong growth inhibitory positive-control shRNA (shMyb.670)²². As the read-out of this screen is inhibition of proliferation, we would expect that essential genes would be enriched among the strongest hits. Indeed, scoring of shRNA-induced effects upon knockdown of all 128 MLL-fusion interactors revealed a strong positive correlation between growth inhibition and reported gene essentiality (Fig. 2c)^{39–42}. However, as we intended to identify proteins with MLL-fusion-specific roles in the network, we reasoned that their loss-of-function might preferably affect the viability of MLL-fusion-expressing leukemia cells. Thus, we re-

screened the 40 candidate genes with the highest confidence in MLL-AF9-expressing MOLM-13 cells and in the MLL-wild-type leukemia cell lines K562 and HL-60. As expected, knockdown of MLL-interactors with known essential functions, such as RFC1, SF3A3, or CCT3, led to growth inhibition in both MLL-fusion cells and MLL-wild-type leukemia cells (Fig. 2d). In contrast, depletion of the known MLL-interaction partner MEN1 selectively inhibited growth of MLL-fusion cells while sparing MLL-wild-type cells, proving the validity of the chosen strategy.

Knockdown of the gene encoding the H3K36me3-specific methyltransferase SETD2 showed a strong bias towards inhibition of proliferation of MLL-fusion-expressing cells, while causing

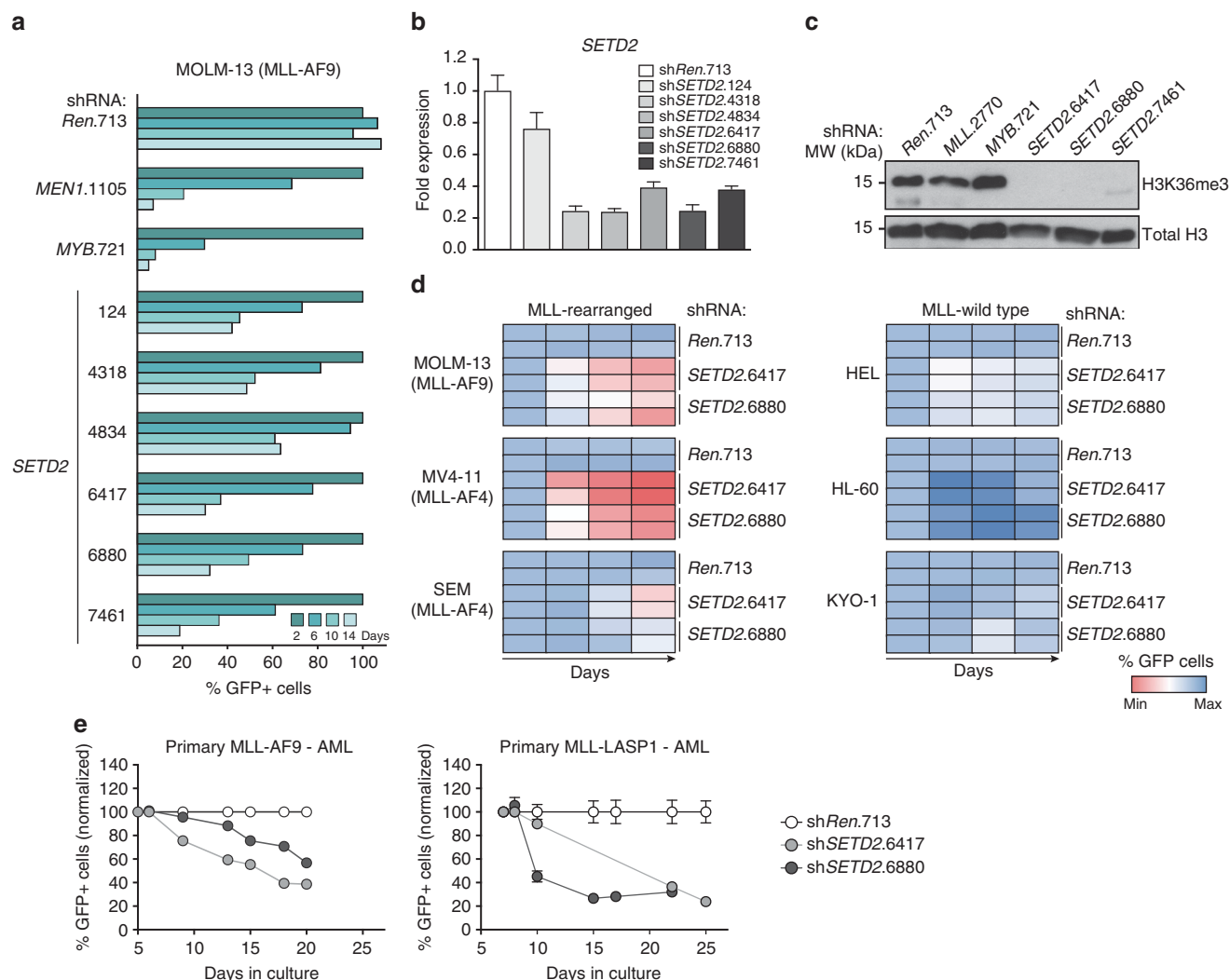


Fig. 3 SETD2 is required for proliferation of MLL-leukemia cells. **a** Results of FACS-based competitive proliferation assay shown as the percentage of GFP-positive MOLM-13 cells expressing individual SETD2-targeting shRNAs in the presence of Dox over 14 days. One representative experiment of four is shown. **b** qPCR analysis of SETD2 mRNA levels in MOLM-13 cells expressing indicated shRNAs after 48 h of Dox treatment (mean \pm s.d. $n = 3$). **c** Western blot analysis of H3K36me3 levels in MOLM-13 cells expressing indicated shRNAs after 72 h of Dox treatment. **d** Heatmap representation of competitive proliferation assays performed in human cell lines harboring MLL rearrangements (left) vs. MLL-wild-type cells (right) expressing indicated shRNAs targeting SETD2 as described in **a**. Representative results of two out of three experiments are shown. **e** Time course of GFP expression of primary human AML cells from patients expressing MLL-AF9 and MLL-LASP1 fusion genes expressing indicated shRNAs (mean \pm s.d. $n = 3$)

negligible cell depletion in K562 and HL-60 cells, suggesting an MLL-fusion-specific function (Fig. 2d, Supplementary Fig. 3a). SETD2 was one of 42 core proteins that interacted with all seven MLL-fusions, as it co-purified with MLL-fusion proteins in all affinity-purification experiments with significant peptide coverage (Supplementary Fig. 3b). Consistently, co-immunoprecipitation experiments showed that this interaction involved the N-terminal part of MLL, which is conserved in all MLL-fusion proteins studied, and the C-terminus of SETD2, which encompasses all annotated functional domains of the SETD2 protein (Supplementary Fig. 3c). SETD2 expression was higher in AML samples than in normal hematopoietic stem and progenitor cell types and mature myeloid cells⁴³ (Supplementary Fig. 3d). SETD2 expression was highest in patients with 11q23 aberrations featuring MLL-translocations, as compared to samples with normal karyotype AML or myelodysplastic syndrome (Supplementary Fig. 3e).

Thus, we identified the methyltransferase SETD2 as a selective effector of MLL-AF9 AML cells through functional genomic investigation of conserved interaction partners of MLL-fusion proteins.

SETD2 is essential for MLL-fusion-expressing cells. We next aimed at validating the shRNA screen results at the level of individual shRNAs. Expression of all six SETD2-targeting shRNAs induced strong growth inhibition in MOLM-13 cells, in line with significant downregulation of SETD2 mRNA (Fig. 3a, b). As SETD2 is the only protein known to mediate tri-methylation of H3K36⁴⁴, we investigated the effect of SETD2 downregulation on total cellular H3K36me3 levels. The three strongest SETD2-targeting shRNAs caused near-complete clearance of global H3K36me3 signals (Fig. 3c). Importantly, growth inhibition was not generally associated with H3K36me3 loss, as we did not observe changes in global H3K36me3 levels upon downregulation of MLL and MYB, which strongly affected proliferation of MLL-AF9 AML cells (Fig. 3c).

As our screening data indicate that SETD2 knockdown selectively inhibits the proliferation of MLL-fusion-expressing cells, we sought to extend this observation to a larger panel of human leukemia cell lines. In addition to MOLM-13 cells also the MLL-AF4-expressing cell lines MV4-11 and SEM showed

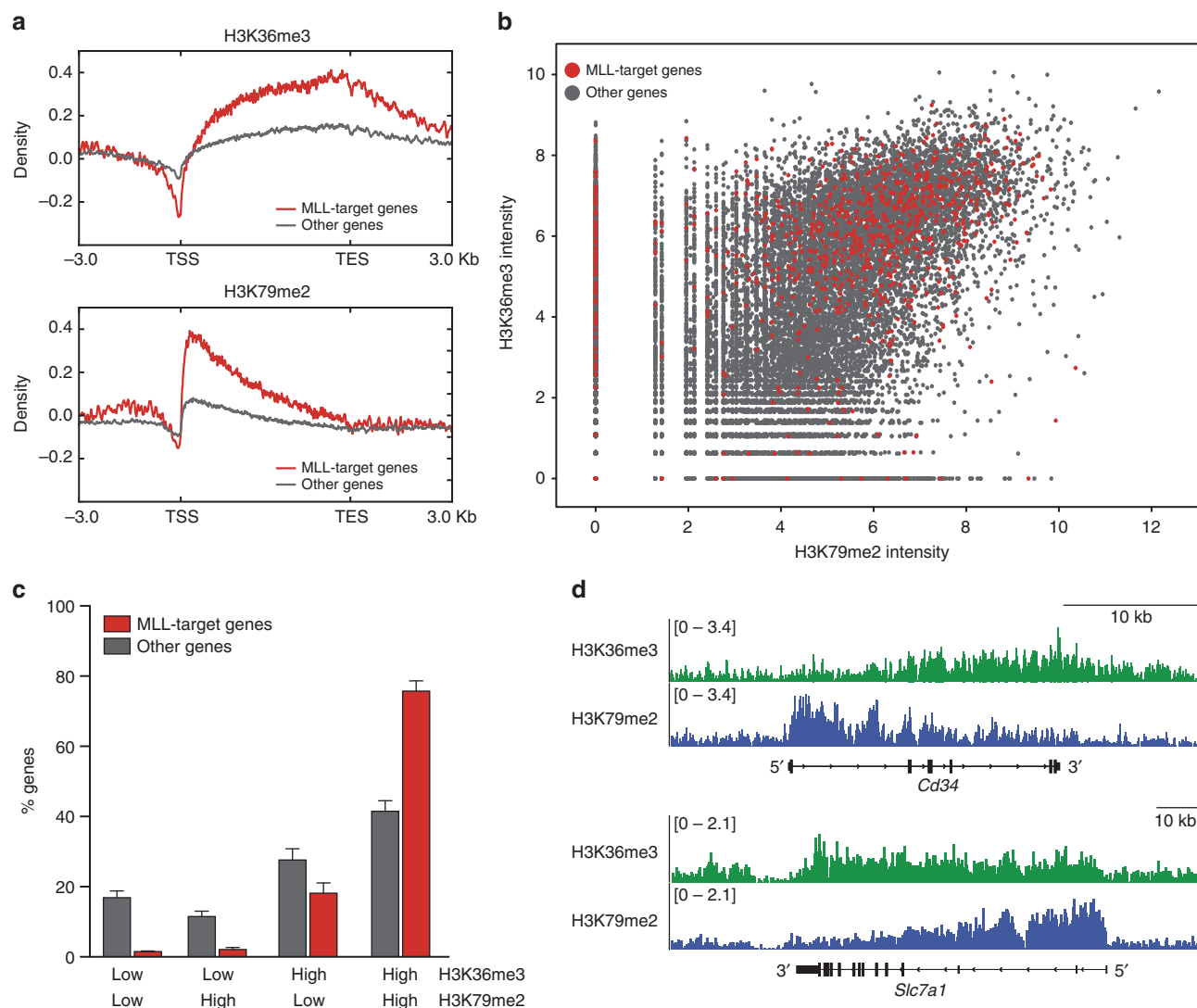


Fig. 4 MLL-target genes are marked by a H3K36me3-H3K79me2 signature. **a** Metagen plots of ChIP-Rx data for H3K36me3 (top) and H3K79me2 (bottom) for MLL-target genes (red) or non-MLL-target genes (gray) in mouse *MLL-AF9/NrasG12D* AML cells. **b** Dot plot of normalized ChIP-Rx signal intensities for H3K36me3 vs. H3K79me2 marks on MLL-target genes (red) or non-MLL-target genes (gray) in the mouse genome (Pearson $R = 0.44$, $p < 2.2 \times 10^{-16}$). **c** Bar graph showing percentages of genes among MLL-target genes and non-MLL-target genes associated with the indicated histone marks; low: not exceeding input counts per gene, high: exceeding input counts per gene (mean \pm s.d. $n = 2$). **d** H3K36me3 (green) and H3K79me2 profiles (blue) of selected MLL-AF9-target genes

significant anti-proliferative responses and induction of apoptosis upon *SETD2* knockdown (Fig. 3d, left, and Supplementary Fig. 4a-c). In contrast, *SETD2* downregulation in the MLL-wild-type cell lines HEL, HL-60, and KYO-1 only marginally affected proliferation (Fig. 3d, right, Supplementary Fig. 4c). *SETD2* knockdown resulted in a strong proliferative disadvantage in primary human AML cells from patients expressing *MLL-AF9* and *MLL-LASP1* fusion genes (Fig. 3e).

Taken together, downregulation of *SETD2* caused a strong anti-proliferative response in primary human AML cells and cell lines expressing various MLL-fusion genes, suggesting a requirement for *SETD2* in the oncogenic context of MLL-fusion proteins.

MLL-target genes exhibit high H3K36me3 levels. To investigate the relationship between *SETD2* and MLL-fusions we profiled the global distribution of the *SETD2*-dependent H3K36me3 mark in a

mouse AML cell line expressing *MLL-AF9* and activated *Nras* (G12D)²² using ChIP-Rx⁴⁵. As expected, H3K36me3 was present on gene bodies of expressed genes. We found that MLL-AF9 target genes²² displayed significantly higher H3K36me3 levels than non-MLL-target genes (Fig. 4a, top). In line with previous data, MLL-fusion target genes were also highly positive for the DOT1L-dependent H3K79me2 mark²¹ (Fig. 4a, bottom), and the global levels of H3K36me3 and H3K79me2 modifications showed a strong positive correlation in mouse *MLL-AF9/NrasG12D* cells (Fig. 4b). However, while only 42% of non-MLL-target genes were highly positive for both marks, 76% of MLL-target genes displayed a combined H3K36me3/H3K79me2-high signature (Fig. 4c, d, and Supplementary Fig. 5). As MLL-fusion-binding was shown to correlate with H3K79me2 on MLL-target genes²¹ and depend on recognition of H3K36me3 marks⁹, these data suggest that the *SETD2*-dependent H3K36me3 modification is part of an epigenetic signature that marks target genes of MLL-fusion proteins together with H3K79me2.

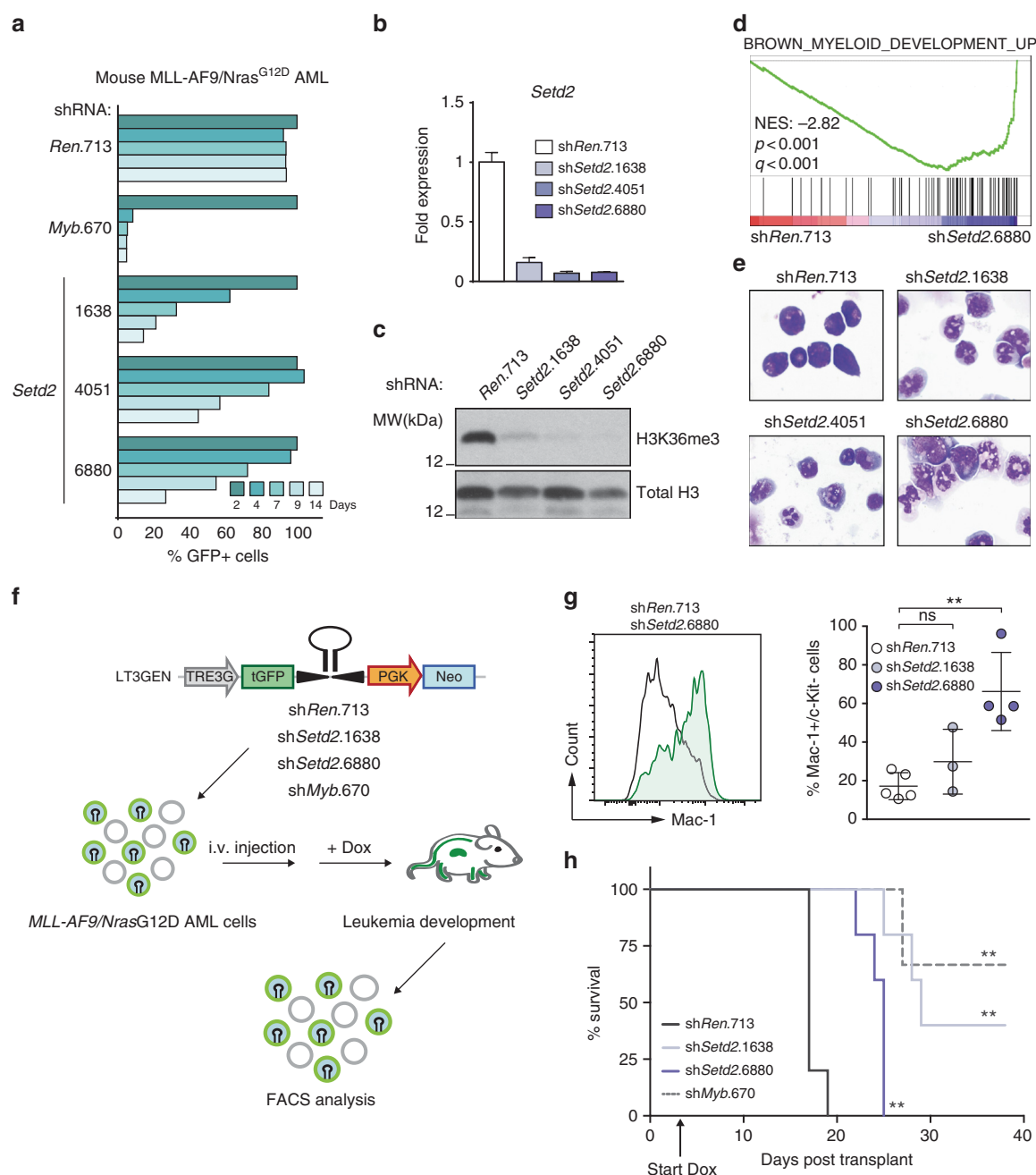


Fig. 5 SETD2 loss induces myeloid differentiation of MLL-leukemia cells. **a** Results of competitive proliferation assay shown as the percentage of GFP-positive mouse MLL-AF9/*Nras*G12D AML cells expressing indicated shRNAs in the presence of Dox over 14 days. One representative experiment out of three is shown. **b** qPCR analysis of *Setd2* mRNA levels in MLL-AF9/*Nras*G12D cells expressing indicated shRNAs after 48 h of Dox treatment (mean \pm s.d. $n = 3$). **c** Western blot analysis of H3K36me3 levels in MLL-AF9/*Nras*G12D cells expressing indicated shRNAs after 72 h of Dox treatment. **d** Gene Set Enrichment Analysis indicating myeloid differentiation of MLL-AF9/*Nras*G12D AML cell line upon knockdown of *Setd2*. NES, Normalized Enrichment Score. **e** Micrographs of cytospin preparations of MLL-AF9/*Nras*G12D AML cells after expression of indicated shRNAs. **f** Schematic representation of the in vivo transplantation assay. MLL-AF9/*Nras*G12D AML cells expressing indicated shRNAs were transplanted into sub-lethally irradiated C57BL/6 Ly5.1 mice. Dox was administered to the drinking water starting from day 3 (arrow in **h**) and disease progression was monitored by bioluminescence imaging. Terminally sick mice were sacrificed and analyzed. **g** Flow cytometric analysis of Mac-1 on MLL-AF9/*Nras*G12D AML cells upon shRNA-mediated *Setd2* knockdown in vivo (mean \pm s.d. $n \geq 3$). **h** Kaplan-Meier survival curves of C57BL/6 Ly5.1 mice transplanted with MLL-AF9/*Nras*G12D AML cells expressing *Setd2*-targeting shRNAs. Survival curves of mice transplanted with cells expressing *Setd2*-targeting shRNAs were compared to cells expressing control shRNAs using a Log-rank test. ns, not significant, ** $p < 0.01$ (t-test)

Loss of SETD2 induces myeloid differentiation and DNA damage. Next we sought to characterize global changes in gene expression upon SETD2 ablation. Dox-inducible knockdown of *Setd2* caused a strong growth disadvantage in mouse MLL-AF9/

*Nras*G12D cells (Fig. 5a, b, Supplementary Fig. 6a). *Setd2* downregulation led to almost complete loss of cellular H3K36me3 signals (Fig. 5c). RNA-seq analysis showed that 868 genes were differentially expressed upon *Setd2* knockdown in

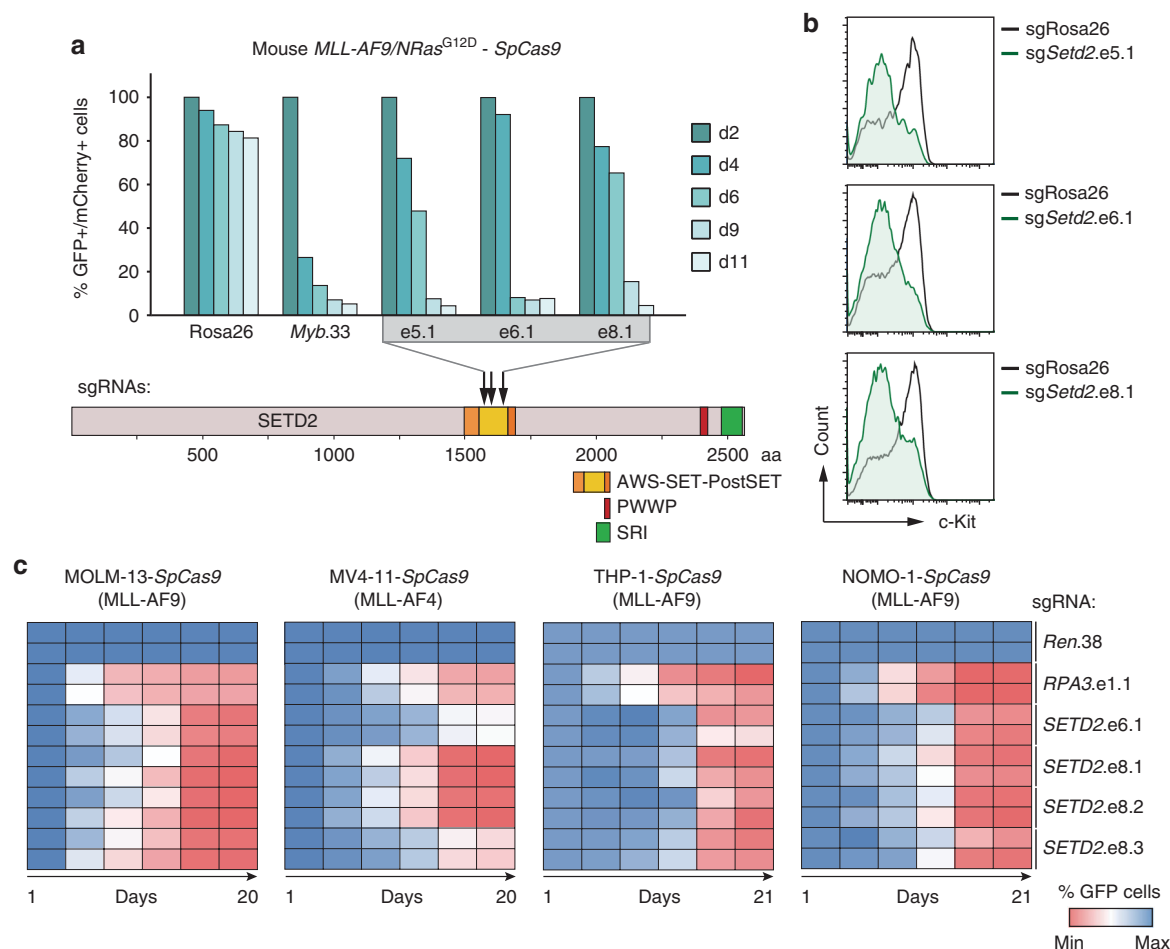


Fig. 6 The SETD2 SET domain is required for proliferation of MLL-leukemia cells. **a** Results of competitive proliferation assays shown as percentages of mCherry+/GFP+ mouse *MLL-AF9/NrasG12D-SpCas9* AML cells expressing indicated sgRNAs (top). Data from one representative experiment of two are shown. Schematic representation of the domain structure of *SETD2* (bottom). **b** Flow cytometric analysis of surface expression of c-Kit in *MLL-AF9/NrasG12D* AML-*SpCas9* cells upon CRISPR/Cas9-mediated mutagenesis of *SETD2*. **c** Heatmap representation of competitive proliferation assays shown as the percentage of GFP-positive human AML cell lines stably expressing *SpCas9* cells upon CRISPR/Cas9-mediated mutagenesis of the *SETD2* SET domain

MLL-AF9/NrasG12D AML cells. While 458 genes were upregulated, 410 genes were downregulated in *shSetd2*-cells, ($p_{adj} < 0.01$, Supplementary Fig. 6b). Consistent with a role of SETD2 in the DNA damage response⁴⁶, *MLL-AF9/NrasG12D* AML cells expressing two different *Setd2*-targeting shRNAs showed upregulation of DNA damage-associated gene expression (Supplementary Fig. 6c). Indeed, *Setd2* downregulation resulted in significantly higher levels of DNA damage in the absence of genotoxic agents, as measured by alkaline comet assay and phosphorylated histone H2AX (γ -H2AX, Supplementary Fig. 7a–c). Knockdown of *Setd2* led to induction of p21, reduced cell cycle progression, and induction of apoptosis of *MLL-AF9/NrasG12D* AML cells (Supplementary Fig. 7c–e).

Gene Set Enrichment Analysis revealed that *Setd2* downregulation induced gene expression programs associated with myeloid differentiation (Fig. 5d). Indeed, *Setd2*-deficient cells displayed clear signs of terminal myeloid maturation, including nuclear segmentation and increased granularity (Fig. 5e), as well as downregulation of the progenitor marker c-Kit and upregulation of the mature myeloid marker Mac-1 (Supplementary Fig. 8a). Similarly, *SETD2* downregulation in the human *MLL-AF4*-expressing cell line MV4-11 and in *MLL-AF9*-expressing MOLM-13 cells induced increased cell surface levels of the differentiation marker CD36 together with macroscopic changes characteristic of myeloid maturation (Supplementary Fig. 8b, c).

To test whether loss of SETD2 could overcome the *MLL-AF9*-dependent differentiation block in myeloid progenitors in vivo, we transplanted *MLL-AF9/NrasG12D* AML cells expressing *Setd2*-targeting or control shRNAs into recipient mice. shRNA expression was induced by Dox-administration and the immuno-phenotype of the developing leukemia was analyzed by flow cytometry (Fig. 5f). Knockdown of *Setd2* induced strong downregulation of c-Kit concomitant with upregulation of Mac-1 in leukemic cells in vivo, resulting in a significant increase in disease latency (Fig. 5g, h and Supplementary Fig. 8d). This is consistent with recent results showing that knockout of *Setd2* greatly increased the latency of *MLL-AF9*-induced AML⁴⁷. While most leukemia cells isolated from moribund recipients of control AML cells showed robust shRNA expression (as measured by GFP levels), shRNA-expressing cells were strongly outcompeted by shRNA-negative cells in case of *Setd2* knockdown in vivo (Supplementary Fig. 8e).

In summary, shRNA-mediated downregulation of SETD2 caused growth arrest, induction of apoptosis, and increased DNA damage. Furthermore, SETD2 loss induced terminal myeloid differentiation in *MLL*-fusion-expressing mouse and human AML cells in vitro and in vivo, indicating that the *MLL*-fusion-induced differentiation block is SETD2-dependent.

The SETD2 SET domain is required for AML growth. To interrogate the translational potential of our findings, we next

wanted to establish whether the methyltransferase activity of SETD2 is necessary for the observed effects. Direction of *SpCas9*-cleavage to functional protein domains was shown to greatly increase the read-out in competitive proliferation assays⁴⁸. We employed CRISPR/Cas9-mediated mutagenesis of the enzymatic SET domain to investigate whether catalytic activity of SETD2 was required for the oncogenicity of MLL-fusion proteins. Introduction of three sgRNAs targeting the *Setd2* SET domain in *SpCas9*-expressing MLL-AF9/*Nras*G12D AML cells led to a strong depletion of transduced cells over time, as shown before⁴⁸ (Fig. 6a). Notably, CRISPR/Cas9-mediated mutagenesis of the *Setd2* SET domain was sufficient to induce myeloid differentiation of MLL-AF9/*Nras*G12D cells, as measured by down-regulation of c-Kit together with upregulation of Mac-1 (Fig. 6b, Supplementary Fig. 9a). In line, we found strong anti-proliferative effects, induction of myeloid differentiation, and apoptosis upon mutagenesis of the SETD2 SET domain in the human MLL-rearranged AML cell lines MOLM-13 and MV4-11, THP-1, and NOMO-1 (Fig. 6c; Supplementary Fig. 9b–f, Supplementary Fig. 10).

These data show that the SET domain of SETD2 is required to sustain the proliferative capacity and differentiation block of MLL-fusion protein-expressing AML cells. In addition, these results imply a functional involvement of the H3K36me3 mark in the maintenance of MLL-fusion-dependent leukemia and offer a plausible route for future pharmacological intervention.

Efficient MLL-fusion-mediated transformation requires SETD2. All our data show a strong functional requirement for the expression and activity of SETD2 in the progression of MLL-leukemia. As it is possible that alternative molecular mechanisms pertain during initiation of MLL-rearranged leukemia, we tested the involvement of SETD2 in this process. *Setd2* knockdown resulted in a significant reduction in MLL-AF9-induced serial replating capacity of mouse hematopoietic stem/progenitor cells (HSPC), indicating that *Setd2* expression is required to unleash the full oncogenic potential of MLL-AF9 (Fig. 7a). *Setd2* ablation induced loss of compact colony morphology characteristic of blast-like cells and induced the formation of large, dispersed colonies reminiscent of mature myeloid clusters (Fig. 7b). Flow cytometry confirmed that *Setd2*-deficient colonies expressed high levels of the mature myeloid marker Mac-1 (Fig. 7c). To investigate the effect of SETD2 on oncogenic transformation in vivo, we co-transduced fetal-liver-derived HSPC expressing a *SpCas9* transgene⁴⁹ with retroviral vectors encoding MLL-ENL and *Setd2*-targeting or control sgRNAs. The contribution of cells carrying sgRNA-induced mutations in the *Setd2* SET domain to leukemia development was investigated by flow cytometric analysis of mCherry expression upon transplantation (Fig. 7d). While cells expressing a control sgRNA showed robust contribution to MLL-ENL-induced leukemia in vivo (56%), cells carrying *Setd2*-mutations induced by two different sgRNAs were clearly under-represented in the leukemic population (5–25%, Fig. 7e, f).

Thus, both downregulation and mutagenesis of SETD2 was incompatible with efficient oncogenic transformation by MLL-fusion oncoproteins in vitro and in vivo. These results indicate that SETD2 expression is required for leukemogenesis and establish SETD2 as a novel actionable target in MLL-rearranged leukemia.

SETD2 loss sensitizes MLL-AML cells to DOT1L inhibition. Finally, we aimed to obtain insight into the molecular mechanism that functionally connects SETD2 activity with MLL-fusion-induced leukemia. ChIP-Rx showed that *Setd2* downregulation led to a concomitant reduction of both H3K36me3 and

H3K79me2 levels on MLL-target genes (Fig. 8a, b and Supplementary Fig. 11a), while it did not alter H3K4me3 density (Supplementary Fig. 11b). As chromatin binding of MLL-fusion proteins was shown to depend on H3K36me3 recognition via the conserved interaction partner LEDGF⁹, we hypothesized that reduction of H3K36me3 levels upon *Setd2* loss might impair chromatin binding of MLL-fusions. Indeed, knockdown of *Setd2* caused reduced binding of MLL-AF9 to the promoters of the canonical MLL-target genes *Hoxa9* and *Meis1* (Fig. 8c), leading to reduced *Hoxa9* expression (Fig. 8d).

Given the dependence of the dual H3K36me3–H3K79me2 signature across MLL-target genes on SETD2 and the strong requirement of MLL-leukemia for the H3K79 methyltransferase DOT1L²¹, we reasoned that SETD2 loss might cooperate with pharmacologic inhibition of DOT1L. Treatment of mouse MLL-AF9/*Nras*G12D and human MLL-AF4-expressing MV4-11 cells with the clinical DOT1L inhibitor EPZ5676⁵⁰ potentiated the effects of SETD2 downregulation, including growth inhibition, induction of apoptosis, and onset of myeloid differentiation. Importantly, none of these parameters were altered in SETD2-proficient cells in the presence of the same concentrations of the inhibitor (Fig. 8e, f, Supplementary Fig. 11c–f). Finally, and consistent with a role of DOT1L in DNA repair⁵¹, we found that combination of SETD2 loss and DOT1L inhibition synergized in the induction of DNA damage (Fig. 8g and Supplementary Fig. 11g).

These data show that loss of SETD2 expression in MLL-fusion AML cells interferes with the H3K36me3–H3K79me2-signature on MLL-target genes and impairs chromatin binding of MLL-fusion proteins. In consequence, SETD2 loss led to hypersensitization of MLL-leukemia cells to small-molecule-mediated DOT1L inhibition, which provides a rationale for potential future combination therapies in AML.

Discussion

Here, we provide the first comprehensive protein–protein interaction network of MLL-fusion proteins in leukemia cells. We show that functional annotation of conserved MLL-interaction partners by loss-of-function screening enables the identification of conserved actionable nodes among the molecular network of MLL-fusions. As exemplified by our discovery of the histone methyltransferase SETD2 as an essential factor in MLL-rearranged leukemia, this approach can reveal novel genetic dependencies and yield new entry points for targeting of the entire group of MLL-rearranged leukemia, comprising over 75 different MLL-fusion partners.

Our results show that MLL-fusion proteins engage a large number of distinct protein–protein interactions. This could be explained by the modular nature of wild-type core MLL complexes^{52,53} and by the specific architecture of their leukemic counterparts⁴. Our analysis of protein–protein interactions of selected, molecularly distinct MLL-fusion proteins greatly expands the cellular catalog of MLL-interacting proteins. In addition, it also provides novel insights into the topologies of MLL-fusion proteins that transform cells via unknown mechanisms. For instance, interactome analysis of the MLL-GAS7-fusion protein showed that it specifically interacts with components of the CTLH complex, which is involved in microtubule dynamics and chromosome segregation⁵⁴.

A core set of 128 proteins constitutes the conserved interactome of MLL-fusion proteins. In addition to known interaction partners of the wild-type MLL protein, such as MEN1, DPY30, and LEDGF, it also contains several proteins whose link to AML biology have only recently been established. For instance, the protein SON interacts with MEN1 to regulate the expression of leukemia-specific genes in a MLL-dependent

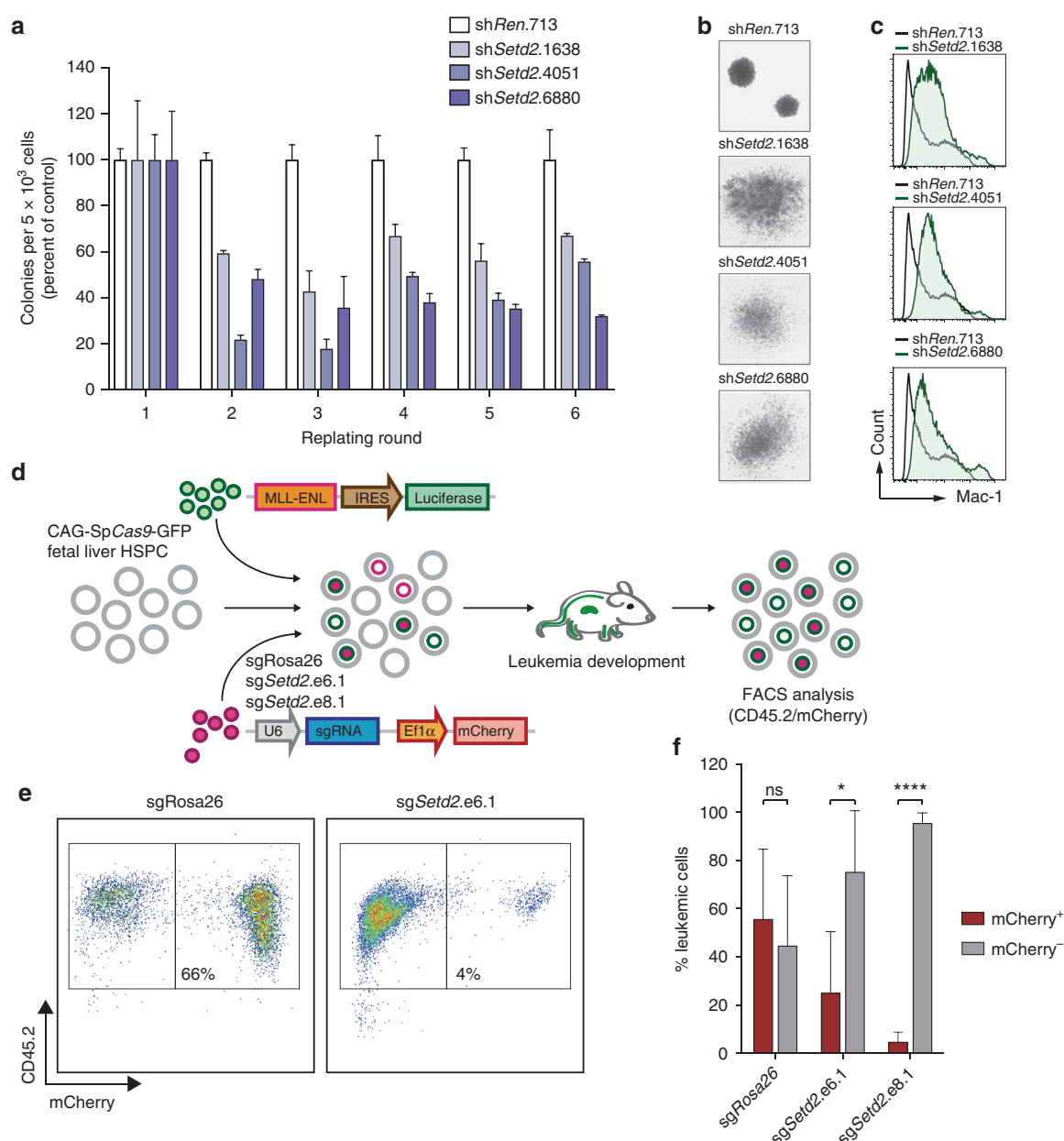


Fig. 7 SETD2 is required for oncogenic transformation by MLL-fusions. **a** Serial replating assay of primary MLL-AF9-transformed fetal liver cells upon shRNA-mediated knockdown of *Setd2*. Colony numbers were normalized to cells expressing shRen.713 (mean \pm s.d. $n = 3$). **b** Morphology of colonies of MLL-AF9-transformed fetal liver cells upon shRNA-mediated knockdown of *Setd2*. **c** Flow cytometric analysis of Mac-1 expression of MLL-AF9-transformed fetal liver cells upon shRNA-mediated knockdown of *Setd2*. **d** Schematic representation of the in vivo transformation assay. Fetal liver cells from SpCas9-transgenic mice were co-transduced with retroviral vectors expressing MLL-ENL and *Setd2*-targeting or control sgRNAs. Cells were transplanted into lethally irradiated C57BL/6 recipient mice. Terminally sick mice were sacrificed and bone marrow cells were analyzed. **e** Representative flow cytometry plots of donor-derived bone marrow cells from mice transplanted with MLL-ENL and a control sgRNA (sgRosa.26, left) or a sgRNA targeting the SET domain of *Setd2* (sgSetd2.e6.1, right). Live cells were gated. **f** Quantification of flow cytometry analysis of donor-derived bone marrow cells as shown in **e** (mean \pm s.d. $n \geq 4$). ns, not significant, * $p < 0.05$, **** $p < 0.0001$ (t-test)

manner⁵⁵. Thus, functional annotation of the core network of MLL-fusion interactors will contribute to establish novel links between MLL-fusion proteins and important cellular processes that had previously not been associated with the biology of MLL-fusion proteins, such as mRNA splicing or protein transport. Given the involvement of these molecular pathways in basic cellular physiology, it is not surprising that more than one third of proteins in the network of conserved MLL-fusion protein interaction partners were identified as essential in recent genome-wide screens^{39–42}.

To discern MLL-fusion-associated genetic dependencies from essential genes we employed a subtractive shRNA screening approach. Strikingly, the gene encoding the methyltransferase SETD2 was identified as an MLL-fusion-specific hit from this screen with high confidence. shRNA-mediated knockdown as well as CRISPR/Cas9-induced mutagenesis of SETD2 caused proliferation arrest and myeloid differentiation of MLL-fusion-expressing primary and transformed human, and mouse AML cells in vitro and in vivo. This is surprising, because SETD2 has been implied to have tumor suppressor activity in various

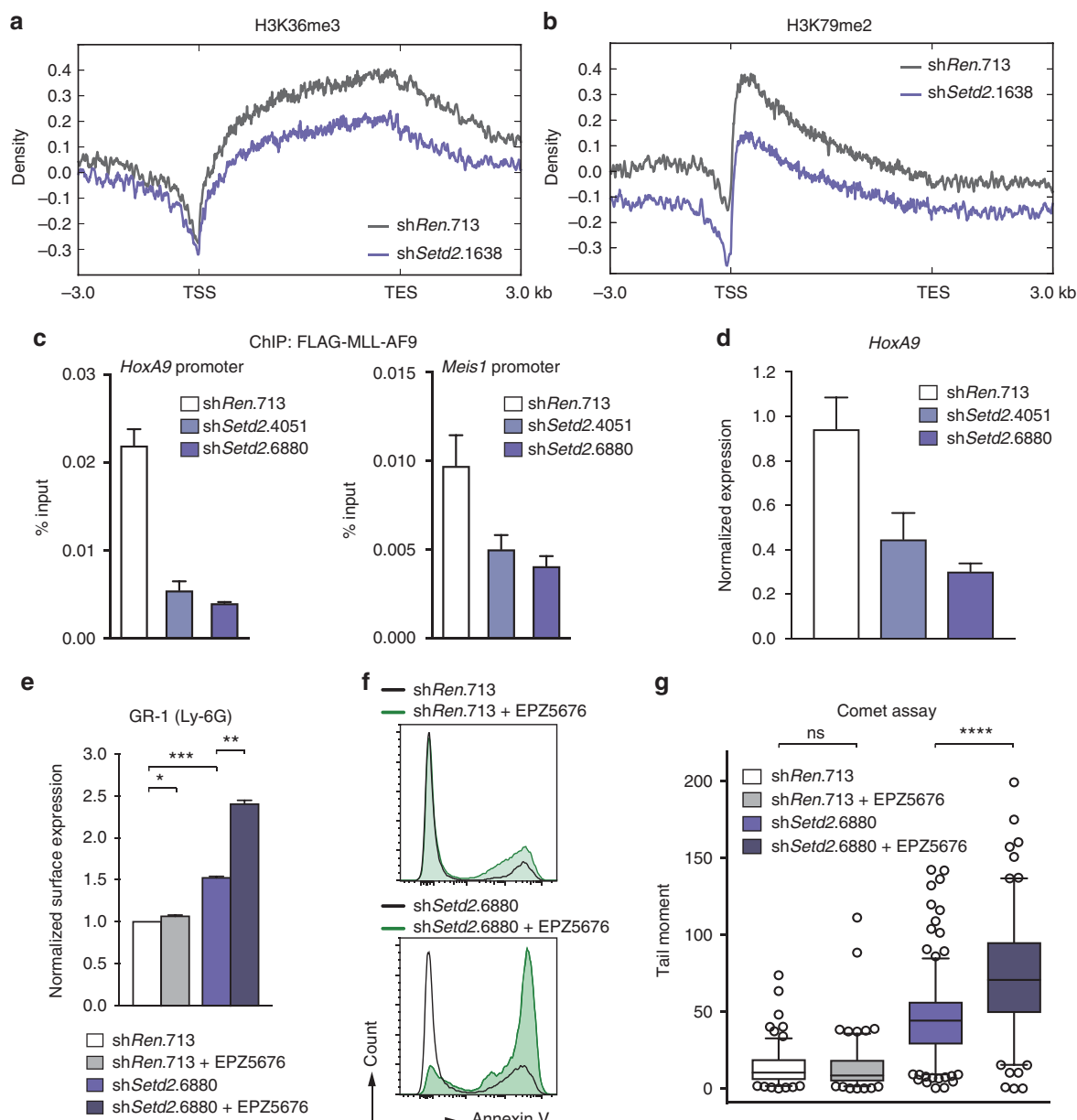


Fig. 8 SETD2 loss sensitizes AML cells to DOT1L inhibition. Metagenome plots of ChIP-Rx data for H3K36me3 (**a**) and H3K79me2 (**b**) after shRNA-mediated knockdown of *Setd2* in mouse *MLL-AF9/NrasG12D* AML cells. **c** qPCR analysis of enrichment on *HoxA9* and *Meis1* promoter regions after MLL-AF9-FLAG-ChIP upon shRNA-mediated knockdown of *Setd2* (mean \pm s.d. $n = 3$). **d** qPCR analysis of *HoxA9* mRNA levels in MLL-AF9-FLAG cells expressing indicated shRNAs (mean \pm s.d. $n = 3$). **e** Quantification of surface expression of Gr-1 (Ly-6G) after shRNA-mediated knockdown of *Setd2* in mouse *MLL-AF9/NrasG12D* AML cells treated with EPZ5676 (500 nM) (mean \pm s.d. $n = 2$). **f** Flow cytometric analysis of Annexin V-positive cells in MV4-11 cells treated with EPZ5676 (50 nM) after shRNA-mediated knockdown of *SETD2*. **g** Quantification of tail moments in an alkaline comet assay performed after shRNA-mediated knockdown of *Setd2* in mouse *MLL-AF9/NrasG12D* AML cells treated with EPZ5676 (500 nM). Quantification of >150 cells is shown. ns, not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p > 0.0001$ (t-test)

malignancies, including leukemia^{56–58}. SETD2 knockdown was reported to cause a driver-independent proliferative advantage of leukemia cells in vitro and in vivo⁵⁸. In contrast, another report showed that mutational disruption of the *SETD2* SET domain was incompatible with MLL-AF9 AML cell growth⁴⁸. Furthermore, several genome-wide CRISPR/Cas9 screens identified *SETD2* as an essential gene in leukemia cell lines^{39–42}. Finally, a recent report showed that while homozygous *Setd2* deletion in the mouse strongly delayed leukemogenesis, heterozygous *Setd2* deletion accelerated MLL-AF9-induced leukemia and caused chemoresistance⁴⁷. This is consistent with increased frequencies of *SETD2* mutations in high-risk leukemia patients that show

increased genomic complexity and chromothripsis⁵⁷, and often exhibit therapy resistance and relapse⁵⁶. Therefore, as the majority of cancer patients carry heterozygous *SETD2* mutations, *SETD2* might act as a haplo-insufficient tumor suppressor. In contrast complete loss of *SETD2* strongly impedes leukemia development.

The most prominent cellular function of SETD2 is its non-redundant H3K36 tri-methylation activity⁴⁴. The H3K36me3 mark is enriched on gene bodies of actively transcribed genes⁵⁹. We found that MLL-target genes displayed high H3K36me3 levels, validating our proteomic identification of SETD2 as an interactor of MLL-fusion proteins at the genomic level. Given the

large number of proteins harboring a H3K36me3-recognizing PWWP motif⁶⁰, several key cellular processes were postulated to be influenced by this epigenetic mark, including transcriptional elongation, splicing, and epigenetic control of gene expression⁶¹. Consistent with a role for the PWWP motif of the MLL-interactor LEDGF in chromatin binding of MLL-fusion proteins⁹, SETD2 loss caused reduced binding of MLL-AF9 to target promoters. Thus, the interaction between SETD2 and MLL-fusion proteins could be required to ensure efficient chromatin binding of MLL-fusions through the maintenance of high H3K36me3 levels on MLL-target genes.

Our results clearly show that SETD2 is involved in the control of the DNA damage response in MLL-fusion-expressing cells. SETD2-deficient cells exhibited high amounts of DNA damage and increased γ -H2AX levels, even in the absence of exogenous genotoxic stress. This is in line with a recent study showing that SETD2 mutations in leukemia impair the DNA damage response, thereby leading to chemotherapy resistance⁴⁷. This defect is attributed to loss of H3K36me3-dependent recruitment of repair proteins to sites of DNA damage. It was shown that MLL-AF9-transformed cells require an intact DNA damage response for full oncogenicity, as experimental induction of DNA damage led to differentiation of leukemia cells⁶². In line with this, down-regulation of SETD2 was sufficient to induce myeloid differentiation of AML blasts. Therefore, a physical and functional interaction between MLL-fusion proteins and SETD2 could be required to guarantee efficient, H3K36me3-dependent repair of DNA lesions that continuously occur during MLL-fusion-induced oncogenic transcription. However, given the accumulation of DNA damage upon SETD2 loss, continuous SETD2 inhibition might result in increased formation of chemoresistant AML subclones.

We found that SETD2 was required for the maintenance of a specific dual H3K36me3-H3K79me2 signature on target genes of MLL-fusions. The H3K79me2 mark is catalyzed by the histone methyltransferase DOT1L, which is critical for the establishment and maintenance of MLL-rearranged leukemia²¹. SETD2 down-regulation rendered MLL-fusion-expressing AML cells hypersensitive to the pharmacological DOT1L-inhibitor EPZ5676 (Pinometostat), which is currently in clinical development. It will be interesting to test whether this synergy can be exploited to efficiently target chemoresistant AML cells that are carrying SETD2 mutations.

In summary, our combined proteomic-functional genomic analysis of MLL-fusion protein interactors enabled us to reveal the molecular logic of how modular protein–protein interactions can influence the oncogenicity of MLL-fusion proteins. Our studies provide novel insights into the biology of MLL-fusion proteins and identify an unexpected dependency of MLL-fusion-expressing leukemia cells on the methyltransferase SETD2 during leukemia initiation and maintenance, validating SETD2 as an actionable target MLL-rearranged leukemia.

Methods

Constructs. MLL-fusion genes were assembled by fusing the cDNA of the MLL N-terminus (amino acids 1–1396) to C-terminal parts of AF1p (EPS15), AF4 (AFF1), AF9 (MLLT3), CBP (CREBBP), EEN (SH3GL1), ENL (MLLT1), and GAS7 (GAS7) and cloned into pcDNA5/FRT/TO/SH/GW. Generation of the miR-E shRNA vectors RT3GEN and RT3REN was previously described^{63,64}. The SEM cell line was infected with SGEN⁶⁴. A pMSCV-MLL-AF9-IRES-Venus construct was used for the in vitro re-plating assay, while a pMSCV-MLL-ENL-IRES-Luc2 construct was used for the in vivo transformation assay. A V5-tagged version of the N-terminal part of MLL (amino acids 1–1396) was cloned into a vector containing a Doxycycline-inducible promoter. The C-terminal fragment of SETD2 (amino acids 950–2570) was cloned with a N-terminal 6 \times -Myc tag. The library of 768 shRNAs was designed to target 128 conserved interaction partners of ≥ 5 selected MLL-fusions with six shRNAs per gene. 97-mer oligomers (Integrated DNA Technologies) were reconstituted in H₂O and stored at -80°C . Mini-pools of six shRNAs

targeting the same candidate gene were amplified in parallel PCR reactions using Pfx DNA polymerase (Invitrogen) as described⁶⁴. Reactions were pooled and purified using PCR Clean-up kit (Qiagen). PCR products were digested with EcoRI and XhoI (New England Biolabs) and ligated with retro- or lentiviral vectors allowing for inducible or constitutive shRNA expression together with selection markers. After dialysis, ligations were introduced into Mega X DH10 β T1 electro-competent cells (Invitrogen) by electroporation (2 kV, 200 Ω , 25 μF) using a MicroPulser Electroporator (Bio-Rad). The library was purified using Midi Prep Kit (Qiagen). The presence of shRNA cassettes was verified by Sanger sequencing. For CRISPR/Cas9-mediated mutagenesis, sgRNAs were cloned into lentiviral vectors allowing for constitutive sgRNA expression together with GFP or mCherry as previously described⁶⁵. Sequences of sgRNAs used in the study are listed in Supplementary Table 2.

Cell culture. All standard human leukemia cell lines such as: MOLM13, MV4-11, HEL, HL-60, KYO-1, were obtained from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, www.dsmz.de)) or the American Type Culture Collection (ATCC, www.atcc.org) and modified to express the ecotropic receptor and rtTA3. The murine Tet-On MLL-AF9/NrasG12D AML cell line (RN2) was previously described⁶⁶. All cell lines were cultured in RPMI 1640 (Gibco) supplemented with 10%FBS, 100 U/ml penicillin, and 100 $\mu\text{g}/\text{ml}$ streptomycin. Platinum-E cells were maintained in DMEM (Gibco) supplemented with 10% FBS, 100 U/ml penicillin, and 100 $\mu\text{g}/\text{ml}$ streptomycin. SpCas9-expressing variants of MOLM-13 and MLL-AF9/NrasG12D cells were generated by lentiviral transduction followed by selection with Blasticidin (10 $\mu\text{g}/\text{ml}$). The SpCas9-expressing subclone of MV4-11 was a gift from G. Winter (Dana Farber Cancer Institute, Harvard University). The SpCas9-expressing THP-1 and NOMO-1 cell lines were previously described⁶⁷. MLL-AF9-FLAG cells were previously described²². For proliferation curves, cells were seeded at low densities in triplicates and cell numbers were determined using a multi-channel electronic cell counter (CASY-I; Omni Life Science) in regular intervals. The DOT1L inhibitor EPZ5676 was obtained from BPS Bioscience. Human leukemic blast cells from heparinized samples of AML patients ($n = 3$) were isolated on Ficoll-Hypaque gradients and stored in liquid nitrogen. After thawing, cells were cultured in RPMI 1640 medium containing 10% BIT 9500 Serum Substitute, 100 ng/ml SCF, 50 ng/ml Flt3L, 20 ng/ml IL-3, 20 ng/ml G-CSF (all PeproTech), 10^{-4}M β -mercaptoethanol, 50 $\mu\text{g}/\text{ml}$ gentamicin, and 10 $\mu\text{g}/\text{ml}$ ciprofloxacin plus 500 nM SR1 and 1 μM UM729⁶⁸. This protocol typically leads to sustained proliferation of primary human AML cells over 20 days, yielding a >10 -fold expansion in vitro. All patients gave written informed consent before blood or bone marrow was obtained. The study was approved by the Institutional Review Board of the Medical University of Vienna. Personal data from AML patients were used according to ethics approvals of clinical partners for collection of clinical and genetic data upon informed consent. All cell lines have been tested for mycoplasma contamination. Cell lines used in this study were not listed in the database of commonly misidentified cell lines maintained by ICLAC.

Viral transduction. For retroviral transductions, Platinum-E cells were transiently transfected with pGAG-POL and retroviral expression vectors using the calcium-phosphate method in the presence of Chloroquine (25 μM , Sigma-Aldrich). Virus-containing supernatant was harvested, filtered (0.45 μm), and supplemented with polybrene (5 $\mu\text{g}/\text{ml}$). Target cells were spinoculated at 1300 \times g for 90 min. For lentiviral transductions, HEK293T cells were transiently transfected with psPAX2, pMD2.G, and lentiviral expression vectors. Virus-containing supernatant was harvested, filtered (0.45 μm), and supplemented with polybrene (5 $\mu\text{g}/\text{ml}$). Target cells were spinoculated at 1300 \times g for 90 min. Human primary AML cells were transduced with concentrated lentiviral supernatants via centrifugation (1200 \times g, 90 min) at a multiplicity of infection of 20.

Generation of Flp-In cell lines. Jurkat Flp-In cells (Invitrogen) were transduced with pLenti6/TR (Thermo) and a clone expressing high levels of the tetracycline repressor (TR) was isolated. Cells were transfected with targeting constructs (in pcDNA5/FRT/TO) together with pCAAGS-Flp-E by nucleofection using program X-001 (Amaxa). Targeted cells were selected in Clonacell TCS medium (Stem Cell Technologies) supplemented with 600 $\mu\text{g}/\text{ml}$ Hygromycin B. Clones were isolated and expanded in liquid medium in the presence of Hygromycin B. Expression of MLL-fusions was tested by qRT-PCR after induction of transgene expression by addition of 1 $\mu\text{g}/\text{ml}$ Doxycycline for 24 h.

Affinity purification of protein complexes. Nuclear extracts from transgene-expressing Jurkat cells were prepared and single-step STREP-Tactin purifications of MLL-fusion proteins were performed as described⁵³. All purifications of MLL-fusion proteins were performed from 1×10^9 freshly harvested cells. After being washed with PBS, cells were incubated in buffer N (300 mM sucrose, 10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 0.1 mM DTT, 0.75 mM spermidine, 0.15 mM spermine, 0.1% Nonidet P-40, 50 mM NaF, 1 mM Na₃VO₄, protease inhibitors) for 5 min on ice. Nuclei were collected by centrifugation (500 \times g for 5 min), and the supernatant was removed. The nuclear pellet was washed with buffer N. For the extraction of nuclear proteins, nuclei were

resuspended in buffer C420 (20 mM HEPES pH 7.9, 420 mM NaCl, 25% glycerol, 1 mM EDTA, 1 mM EGTA, 0.1 mM DTT, 50 mM NaF, 1 mM Na_3VO_4 , protease inhibitors), vortexed briefly, and shaken vigorously for 30 min. After centrifugation for 1 h at 100,000 \times g, the protein concentration of the soluble nuclear fraction was measured by Bradford assay. Prior to purification, all nuclear extracts were adjusted to 2 mg/ml and 150 mM NaCl with HEPES buffer (20 mM HEPES, 50 mM NaF, 1 mM Na_3VO_4 , protease inhibitors). 15 mg of nuclear extract were pre-treated with benzonase (20 U/ml) and RNase A (50 ng/ml) for 15 min at 4 °C. Nonspecific binding to the affinity resin was blocked by the addition of avidin (1 μ g/ml). 150 μ l StrepTactin sepharose (IBA) was added and lysates were incubated for 2 h at 4 °C with agitation. Beads were washed 3 times with TNN-HS buffer (50 mM HEPES pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 50 mM NaF, 1 mM Na_3VO_4 , and protease inhibitors). Bound proteins were eluted by the addition of 100 μ l 2.5 mM Biotin (Alfa Aesar) in TNN-HS buffer. Samples were digested with trypsin and processed for LC-MS/MS analysis.

Mass spectrometry. Analysis of affinity purification samples was performed as described previously^{33,34}. All affinity purifications were analyzed on a hybrid linear trap quadrupole (LTQ) Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) coupled to a 1200 series high-performance liquid chromatography system (Agilent Technologies) via a nano-electrospray ion source using liquid junction (Proxeon). Solvents for HPLC separation of peptides were as follows: solvent A consisted of 0.4% formic acid in water, and solvent B consisted of 0.4% formic acid in 70% methanol and 20% isopropanol. 8 μ l of the tryptic peptide mixture were automatically loaded onto a trap column (Zorbax 300SB-C18 5 μ m, 5 \times 0.3 mm, Agilent Biotechnologies). After washing, peptides were eluted by back-flushing onto a 16-cm-fused silica analytical column with an inner diameter of 50 μ m packed with C18-reversed phase material (ReproSil-Pur 120 C18-AQ, 3 μ m, Dr. Maisch) with a 27-min gradient ranging from 3 to 30% solvent B, followed by a 25-min gradient from 30 to 70% solvent B and, finally, a 7-min gradient from 70 to 100% solvent B at a constant flow rate of 100 nl/min. Analyses were performed in a data-dependent acquisition mode, and dynamic exclusion for selected ions was 60 s. A top 15 collision-induced dissociation (CID) method was used, and a single lock mass at m/z 445.120024 ($\text{Si}(\text{CH}_3)_2\text{O}$)₆ was employed. Maximal ion accumulation time allowed in CID mode was 50 ms for MSⁿ in the LTQ and 500 ms in the C-trap. Automatic gain control was used to prevent overfilling of the ion traps and was set to 5000 in MSⁿ mode for the LTQ and 10⁶ ions for a full FTMS scan. Intact peptides were detected in the Orbitrap Velos at 60,000 resolution at m/z 400.

Protein identification and network analysis. For protein identification, raw MS data files were converted into Mascot generic format (.mgf) files and searched against the human SwissProt protein database (v. 2013.01) using the two search engines Mascot (v2.3.02, MatrixScience, London, UK) and Phenyx (v2.6, GeneBio, Geneva, Switzerland). Carbamidomethyl cysteine and oxidized methionine were set as fixed and variable modifications, respectively; one missed tryptic cleavage site per peptide was permitted. The Mascot and Phenyx identifications were combined and filtered as described³² to provide <1% protein false discovery rate (FDR). Known MS contaminants, such as trypsin and keratin were removed from the results, and further analysis of proteins specifically binding to the baits was achieved by fitting the MS data to the generalized linear statistical model: $\log(\text{data}) \sim A_{0,j} + A_{i,j} + \alpha_i + \beta_{i,k}$, where $A_{0,j}$ is the logarithm of the baseline abundance of the j -th prey protein (estimated from the control AP-MS experiments), $A_{i,j}$ is the specific enrichment of the j -th prey in the pull-downs of i -th bait, and α_i and $\beta_{i,k}$ are the normalization terms that model the abundance of background proteome in the k -th replicate pull-down of i -th bait (to estimate α_i , DDX5 and DDX17 proteins were used, as these are known components of the nuclear proteome background and were ubiquitously present in all AP-MS experiments). To improve the accuracy, the model was independently applied to three different types of MS data: Protein spectral counts (the Poisson distribution was used to model the data) and the sum of peptide scores from either Mascot or Phenyx search results, assuming the log-normal distribution. Only peptides unique to the protein groups were used. The inference of the model parameters was achieved using JAGS v3.0. For each type of MS data, the p -value for the hypothesis that $A_{i,j} > 0$ (i.e., that the j -th prey binds specifically to the i -th bait) was calculated and then the three p -values were combined into a single p -value using the Fisher method. All the identified bait-prey pairs were ranked by the combined p -value. The 300 most significant interactions per bait were retained. This cutoff represents a compromise between ensuring high statistical significance of the included interactors, but also capturing sufficient diversity in the interactomes of the selected MLL-fusion proteins. The proteins shared by at least five baits were selected for further analysis. Seven proteins were manually removed as these were either frequently observed contaminants or were not detected in human hematopoietic cell lines. The final set was comprised of 128 proteins. The resulting network was extended by the known protein-protein interactions, which were retrieved from three different datasets: (i) the set of non-redundant complexes in CORUM⁶⁹, from which binary protein-protein interactions were extracted using the matrix model. (ii) The set of interactions described in ref. ⁷⁰, which combines data from several public repositories. (iii) The set of interactions reported in ref. ⁷¹, integrating different data sets. After removing self-interactions, the final network consisted of 365 PPIs between 101 core MLL-fusion interactors, while 27 other interactors identified by AP-MS remained connected

only to the MLL-fusion baits. The network was partitioned into distinct protein communities by maximizing the modularity score of the network over all possible partitions using the “cluster_optimal” function of the “igraph” package in R. Gene ontology (GO) term enrichment analysis for each separate network community was performed. Enrichment was computed with the topGO package from R, using the default algorithm and the annotation file from geneontology.org (18 November 2015). All human proteins in UniProtKB/Swiss-Prot were used as the background population. p -values were corrected for multiple testing using the Benjamini-Hochberg procedure (FDR). Based on functional annotation similarity, unconnected nodes were assigned to the most enriched GO terms in each community. Enrichment of protein complexes within the network of 960 MLL-fusion-interactors was estimated by Fisher’s exact test. Before enrichment, CORUM core complexes sharing >70% of the proteins were iteratively merged to reduce the redundancy. All proteins present in at least one complex were used as the background population. p -values were corrected for multiple testing as explained above. Protein interaction networks were visualized using Cytoscape and Gephi. Detailed information about all 960 identified interactors of MLL-fusion proteins is provided in Supplementary Data 1.

Negative selection RNAi screening. MOLM-13 cells transduced with mini-pools of retroviral vectors for shRNA-mediated targeting of conserved MLL-interactors (coupled to GFP) were mixed in a 50:50 ratio with cells expressing control shRNAs (coupled to dsRed) and cultured in the presence of Doxycycline (1 μ g/ml). Changes in GFP/dsRed ratios were examined by flow cytometry over time. Percentages of GFP-positive cells were measured at each time point during the experiment and normalized to initial measurement after 2 days of Dox treatment. Gene essentiality was assessed based on recent large-scale datasets from genome-wide screens^{39–42}. Based on individual scores from single screens, we assigned scores of 1 (essential) vs. 0 (non-essential) to each gene in our data set. Our combined essentiality score reflects the sum of all essentiality information per gene from 18 different experiments. Thus, a gene that is ubiquitously essential will obtain a score of 18, while a ubiquitously non-essential gene will obtain a score of 0. A gene was called essential if it scored in ≥ 10 of 18 cell lines. Sequences of shRNAs used for the RNAi screen are listed in Supplementary Data 2 and Supplementary Table 3.

Chromatin immunoprecipitation (ChIP-Rx) and sequencing. MLL-AF9/*NrasG12D* AML cells and *Drosophila melanogaster* S2 cells were separately cross-linked with 10% formaldehyde and quenched with glycine (2.5 M). Pellets were washed, pooled, and resuspended in SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0). Chromatin was sonicated to obtain fragments of 150 bp using a Bioruptor sonicator (Diagenode). 0.5% Triton X-100 was added to the samples to allow solubilization of the sheared DNA. Chromatin was incubated with antibodies overnight (5 μ g each). Antibody-bound material was enriched using protein-G-coupled magnetic beads (Invitrogen), washed (50 mM Hepes-KOH, pH 7.4; 500 mM LiCl; 1 mM EDTA; 1% NP40 and 0.5% Na-Deoxycholate), and released using elution buffer (50 mM Tris-HCl, pH 8.0; 10 mM EDTA and 1% SDS) at 65 °C. DNA-protein crosslinks were reverted by incubating the samples overnight at 65 °C in the presence of 0.2 M NaCl. The DNA was treated with RNaseA (0.2 mg/ml) and proteinase K (0.2 mg/ml) and purified using PCR clean-up kit (Qiagen). Chromatin immunoprecipitation of FLAG-tagged MLL-AF9 was performed using the High Sensitivity ChIP Kit (Abcam, 185913) according to the manufacturer’s instructions. Antibodies used were: anti-H3K4me3 (Abcam, 8580) anti-H3K36me3 (Abcam, 9050), anti-H3K79me2 (Abcam, 3594), anti-Flag (Sigma, F1804). Sequencing libraries were prepared using NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs) and sequenced on Illumina HiSeq 4000 using 50 bp single-read chemistry.

Raw ChIP-seq reads were evaluated with FastQC (version 0.11.4). Quality-filtering and trimming was done with PRINSEQ-lite (version 0.20.4). Resulting high-quality reads were simultaneously mapped against the *Mus musculus* (GRCm38) and *Drosophila melanogaster* (dm6) reference genomes via BWA (version 0.7.15). SAMtools (version 1.4) was used to split the alignments into mouse and *Drosophila* reads. Read normalization via the *Drosophila melanogaster* spike-in material was carried out with Deeptools (version 2.5.0.1) for each sample. Profile plots of histone marks were also generated with Deeptools (version 2.5.0.1). For the comparison of H3K79me2 vs. H3K36me3 signal intensities on MLL-target genes vs. non-MLL-targets, an equally sized set of randomly selected non-MLL-target genes was chosen. MLL-target genes represent genes that were downregulated upon MLL-AF9 withdrawal as measured by microarray analysis²². IGV was used for manual inspection and visualization of data. For the analysis of histone mark intensities in genes, mapped reads per gene were counted with featureCounts (1.5.0), respective input counts subtracted, and normalized via TMM using the edgeR package. The Pearson correlation coefficient between changes in respective histone marks over gene bodies after *Setd2* knockdown was calculated with the functions bigwigCompare, multiBigwigSummary, and plotCorrelation of Deeptools.

RNA sequencing. RNA was isolated using RNeasy kit (Qiagen). The amount of total RNA was quantified using the Qubit 2.0 Fluorometric Quantitation system (Life Technologies) and the RNA integrity number was determined using the

Experion Automated Electrophoresis System (Bio-Rad). RNA-seq libraries were prepared with TruSeq Stranded mRNA LT sample preparation kit (Illumina) using Sciclone and Zephyr liquid handling robotics (PerkinElmer). Sequencing libraries were pooled, diluted, and sequenced on an Illumina HiSeq 3000 using 50 bp single-read chemistry. Base calls provided by the Illumina Realtime Analysis software were converted into BAM format using Illumina2bam and demultiplexed using BamIndexDecoder (<https://github.com/wtsi-npg/illumina2bam>). Initial quality control of raw sequencing reads was done with FastQC (version 0.11.4) followed by pre-processing with PRINSEQ-lite (version 0.20.4). Resulting high-quality reads were mapped via STAR⁷² (version 2.5.0b) against the mouse (GRCm38) reference genome. After processing of the alignment results with SAMtools (0.1.19) counts per gene were obtained by HTSeq⁷³ (version 0.6.0). Normalization and differential expression analysis between two samples was carried out with DESeq2⁷⁴. For the visualization of gene expression and unsupervised hierarchical clustering of the samples the rlog normalization in DESeq2 was applied. We used the R library pheatmap for sample clustering (euclidean distance, complete linkage clustering) and heatmap.2 from the gplots package to visualize differentially expressed genes (Pearson correlation and ward.D clustering).

Real-time PCR analysis. Total RNA was isolated using RNeasy mini kit (Qiagen). Reverse transcription was performed with RevertAid RT Kit (Thermo Scientific) using 300 ng RNA. Quantitative PCR was performed using SensiMix SYBR Hi-ROX kit (Bioline) on a RotorGene Q PCR machine (RG-600, Qiagen). Results were analyzed using the $2^{-ddC(t)}$ method. Sequences of primers used for qPCR are listed in Supplementary Table 4.

FACS analysis. Cells were incubated in Fc block reagent (murine: Biolegend, 14-0161-85, clone 93; human: Biolegend 422301) prior to incubation with the following antibodies: anti-human CD36 (Biolegend, 336207, clone 5-271), Brilliant Violet 421 anti-mouse/human CD11b (Biolegend, 101235, clone M1/70), APC Rat anti-Mouse CD117 (BD Pharmingen, 553356 = cell, clone 2B8), anti-Mouse Ly-6G (Gr-1) (Biolegend, 108411, clone RB6-8C5). Samples were measured on LSR Fortessa or Canto II flow cytometers (BD Biosciences) and analyzed using FlowJo software (Tree Star).

Co-immunoprecipitation. HEK293 cells stably expressing pMSCV-rtTA3-IRES-EcoR-PGK-Puro were transiently transfected with the indicated constructs using the PEI transfection method. The expression of the N-terminal MLL-fragment was induced with Doxycycline for 24 h (1 µg/ml, Sigma Aldrich). The Proteasome inhibitor MG-132 (Sigma-Aldrich, 10 mM in DMSO), was added to the medium 2 h before cell harvest. Cells were harvested in IP-lysis buffer (50 mM Tris/HCl pH 7.5, 150 mM NaCl, 1% NP-40, 5 mM EDTA, 5 mM EGTA) supplemented with protease and phosphatase inhibitors. Protein concentrations were determined with Bradford protein assay (Biorad) using γ-globulin (Biorad) as a standard. Subsequently, lysates were incubated with Anti-Myc-tag mAb-Magnetic Beads (Biomedica GmbH) for 1.5 h with continuous rotation at 4 °C. Beads were recovered by centrifugation and washed in IP-buffer. Bound proteins were released by addition of Laemmli-sample-buffer (Biorad) and boiling for 10 min at 95 °C before SDS-PAGE analysis and immunoblotting.

Western blotting. Western blotting was performed according to standard laboratory protocols. Antibodies used were: anti-H3K36me3 (Abcam, 9050; 1:1000), anti-H3 (Abcam, 1791; 1:5000), anti-HA (Covance, MMS-101P; 1:2000), anti-H2AX (Millipore 05-636; 1:5000) anti-Tubulin (Abcam, 7291; 1:5000), anti-RCC-1 (Santa Cruz, sc-55559; 1:2000), anti-p21 (Santa Cruz, sc-6246; 1:1000), anti-GAPDH (Santa Cruz, sc-365062; 1:1000), anti-V5-tag (Cell Signaling, 13202; 1:2000), anti-Myc (Abcam, 9106; 1:10000). Secondary antibodies used were: goat anti-mouse HRP (Jackson ImmunoResearch, 115-035-03 or Thermo Fisher Scientific, 31430; 1:5000), goat anti-rabbit HRP (Jackson ImmunoResearch, 111-035-003 or Thermo Fisher Scientific 31460; 1:5000). Uncropped scans of all blots are shown in Supplementary Fig. 12.

Cytospin analysis. Cells were cytocentrifuged onto glass slides and stained with Giemsa staining solution before microscopic analysis. Images were processed using Adobe Photoshop (Adobe).

Comet assay. Cells were treated with Doxycycline (1 µg/ml) to induce shRNA expression. shRen.713-expressing cells were treated with 150 µM H₂O₂ for 10 min. 4×10^4 cells were washed in PBS and mixed with 100 µl, 0.5% low melting agarose. The cell suspension was deposited on pre-chilled frosted glass slides pre-coated with 1% agarose. Slides were immersed in pre-chilled lysis buffer (2.5 M NaCl, 10 mM Tris-HCl, 100 mM Na₂EDTA, 10% DMSO, and 1% Triton X; pH 10) for 1–2 h and washed with cold H₂O (3 times for 10 min). Slides were incubated in electrophoresis buffer (55 mM NaOH, 1 mM EDTA, 1% DMSO; pH 12.8) for 45 min followed by electrophoresis at 35 V for 40 min. Samples were neutralized in 400 mM Tris-HCl buffer pH 7.0 for 1 h and washed once with pre-chilled H₂O before staining with SYBR Gold. Comet tail moments, defined as the average distance

migrated by the DNA multiplied by the fraction of DNA in the comet tail, were scored using the CASP image-analysis software.

Transplantation experiments. 1×10^6 Murine MLL-AF9/NrasG12D cells were injected into the tail-vein of sub-lethally (5.5 Gy) irradiated C57BL/6 Ly5.1 recipient ($n = 5$). Disease progression was monitored by bioluminescence imaging. Doxycycline (4 mg/ml) was supplied to the drinking water of mice to activate the expression of shRNAs. E14.5 fetal liver cells from C57BL/6 Ly5.2 embryos with heterozygous expression of the SpCas9 transgene⁴⁹ were co-transduced with retroviral vectors allowing for constitutive expression of MLL-ENL and Luciferase, and sgRNAs coupled to mCherry. The efficiency of infection ranged between 5–11%. Transduced fetal liver cells were injected into the tail-vein of lethally (2×5.5 Gy) irradiated C57BL/6 Ly5.1 recipient mice. Terminally sick animals were sacrificed after 50–60 days, and bone marrow was isolated from femurs and tibia. Animals suffering from obvious other symptoms than leukemia were excluded from the analyses. During all animal experiments we adhered to the 3 R principles (reduction, replacement, and refinement). Animal numbers were determined by the investigator using previous experience and based on judgement of pilot experiments. In general, animal numbers were chosen to be as small as possible but large enough to provide needed estimates for statistical tests, based on previous experience. All animal experiments were performed according to ethical animal license protocols approved by the authorities of the Austrian government. No randomization was used in transplantation experiments. The investigator was not blinded to the group allocation during the experiment and while assessing the outcome.

Hematopoietic progenitor re-plating assay. Fetal liver cells were retrovirally co-transduced with MLL-AF9 coupled to Venus and vectors allowing for constitutive expression of shRNAs coupled to mCherry. Venus/mCherry double-positive cells were isolated by FACS sorting and seeded in complete methylcellulose medium (MethoCult M3434). Colonies were scored in 7-day intervals and 5×10^3 cells were re-plated.

Apoptosis assays. Annexin V staining was performed according to the manufacturer's protocol (Annexin V Apoptosis Detection Kit PE, Affymetrix, eBioscience). The TUNEL assay was performed according to the manufacturer's instructions (ApoBrdU Red DNA Fragmentation Kit; BioVision (K404-60)). Cells were analyzed by flow cytometry.

Genotyping of cells with CRISPR/Cas9-induced mutations. Targeted regions were amplified in a PCR reaction using LA Taq[®] DNA Polymerase (TaKaRa RR002A). PCR products were purified (Qiagen) and analyzed by Sanger sequencing. Chromatograms were analyzed with the TIDE tool (Tracking of Indels by Decomposition, <https://tide-calculator.nki.nl>)⁷⁵ to quantify nature and frequency of generated indels.

Cell cycle analysis. Murine MLL-AF9/NrasG12D AML cells were cultured in the presence of Doxycycline (1 µg/ml), harvested, fixed in 70% Ethanol and stored at –20 °C until further analysis. Cells were stained with PI staining solution and examined by flow cytometry.

Statistical analysis. Two-tailed Student's *t*-tests were used for statistical analysis if not stated otherwise.

Data availability. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD009338. RNA-seq and ChIP-seq data was deposited into the Gene Expression Omnibus (GEO). GEO accession GSE110521.

Received: 31 October 2017 Accepted: 23 April 2018

Published online: 18 May 2018

References

- Mitelman, F., Johansson, B. & Mertens, F. The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **7**, 233–245 (2007).
- Shortt, J., Ott, C. J., Johnstone, R. W. & Bradner, J. E. A chemical probe toolbox for dissecting the cancer epigenome. *Nat. Rev. Cancer* **17**, 160–183 (2017).
- Grimwade, D., Ivey, A. & Huntly, B. J. P. Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. *Blood* **127**, 29–42 (2015).
- Slany, R. K. The molecular mechanics of mixed lineage leukemia. *Oncogene* **35**, 5215–5223 (2016).

5. Meyer, C. et al. The MLL recombinome of acute leukemias in 2013. *Leukemia* **27**, 2165–2176 (2013).
6. Milne, T. A. et al. Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol. Cell* **38**, 853–863 (2010).
7. Cierpicki, T. et al. Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. *Nat. Struct. Mol. Biol.* **17**, 62–68 (2010).
8. Yokoyama, A. & Cleary, M. L. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* **14**, 36–46 (2008).
9. Okuda, H. et al. MLL fusion proteins link transcriptional coactivators to previously active CpG-rich promoters. *Nucleic Acids Res.* **42**, 4241–4256 (2014).
10. Mohan, M. et al. Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev.* **24**, 574–589 (2010).
11. Biswas, D. et al. Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc. Natl Acad. Sci. USA* **108**, 15751–15756 (2011).
12. Muntean, A. G. et al. The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell* **17**, 609–621 (2010).
13. Yokoyama, A., Lin, M., Naresh, A., Kitabayashi, I. & Cleary, M. L. A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell* **17**, 198–212 (2010).
14. Lin, C. et al. AFF4, a component of the ELL / P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell* **37**, 429–437 (2010).
15. Wang, Z. et al. Glycogen synthase kinase 3 in MLL leukaemia maintenance and targeted therapy. *Nature* **455**, 1205–1209 (2008).
16. Wang, Y. et al. The Wnt/ β -catenin pathway is required for the development of leukemia stem cells in AML. *Science* **327**, 1650–1653 (2010).
17. Miller, P. G. et al. In vivo RNAi screening identifies a leukemia-specific dependence on integrin β 3 signaling. *Cancer Cell* **3**, 1–14 (2013).
18. Zuber, J. et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).
19. Harris, W. J. et al. The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell* **21**, 473–487 (2012).
20. Shen, C. et al. NSD3-short is an adaptor protein that couples BRD4 to the CHD8 chromatin remodeler. *Mol. Cell* **60**, 847–859 (2015).
21. Bernt, K. M. et al. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).
22. Zuber, J. et al. An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. *Genes Dev.* **25**, 1628–1640 (2011).
23. Ohlsson, E. et al. Initiation of MLL-rearranged AML is dependent on C/EBP α . *J. Exp. Med.* **211**, 5–13 (2013).
24. Ye, M. et al. Hematopoietic differentiation is required for initiation of acute myeloid leukemia. *Cell Stem Cell* **17**, 611–623 (2015).
25. Thiel, A. T. et al. MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. *Cancer Cell* **17**, 148–159 (2010).
26. Cheung, N. et al. Targeting aberrant epigenetic networks mediated by PRMT1 and KDM4C in acute myeloid leukemia. *Cancer Cell* **29**, 32–48 (2016).
27. Cheung, N., Chan, L. C., Thompson, A., Cleary, M. L. & So, C. W. Protein arginine-methyltransferase-dependent oncogenesis. *Nat. Cell Biol.* **9**, 1208–1215 (2007).
28. Sobulo, O. M. et al. MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). *Proc. Natl Acad. Sci. USA* **94**, 8732–8737 (1997).
29. Lavau, C., Du, C., Thirman, M. & Zeleznik-Le, N. Chromatin-related properties of CBP fused to MLL generate a myelodysplastic-like syndrome that evolves into myeloid leukemia. *EMBO J.* **19**, 4655–4664 (2000).
30. So, C. W., Lin, M., Ayton, P. M., Chen, E. H. & Cleary, M. L. Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. *Cancer Cell* **4**, 99–110 (2003).
31. Martin, M. E. et al. Dimerization of MLL fusion proteins immortalizes hematopoietic cells. *Cancer Cell* **4**, 197–207 (2003).
32. Pichlmair, A. et al. Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* **487**, 486–490 (2012).
33. Giambruno, R. et al. Affinity purification strategies for proteomic analysis of transcription factor complexes. *J. Proteome Res.* **12**, 4018–4027 (2013).
34. Grebier, F. et al. Pharmacological targeting of the Wdr5-MLL interaction in C/EBP α N-terminal leukemia. *Nat. Chem. Biol.* **11**, 571–578 (2015).
35. Yokoyama, A. et al. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell Biol.* **24**, 5639–5649 (2004).
36. Nie, Z. et al. Novel SWI/SNF chromatin-remodeling complexes contain a mixed-lineage leukemia chromosomal translocation partner. *Mol. Cell Biol.* **23**, 2942–2952 (2003).
37. Maethner, E. et al. MLL-ENL inhibits polycomb repressive complex 1 to achieve efficient transformation of hematopoietic cells. *Cell Rep.* **3**, 1553–1566 (2013).
38. Yokoyama, A. et al. The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123**, 207–218 (2005).
39. Blomen, V. A. et al. Gene essentiality and synthetic lethality in haploid human cells. *Science* **350**, 1092–1096 (2015).
40. Tzelepis, K. et al. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep.* **17**, 1193–1205 (2016).
41. Hart, T. et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities screens reveal fitness genes. *Cell* **163**, 1515–1526 (2015).
42. Wang, T. et al. Identification and characterization of essential genes in the human genome. *Science* **350**, 1092–1096 (2015).
43. Bagger, F. O. et al. BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic Acids Res.* **44**, D917–D924 (2015).
44. Edmunds, J. W., Mahadevan, L. C. & Clayton, A. L. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.* **27**, 406–420 (2008).
45. Orlando, D. A. et al. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome resource quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* **9**, 1163–1170 (2014).
46. Pfister, S. X. et al. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep.* **7**, 2006–2018 (2014).
47. Mar, B. G. et al. SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood* **130**, 2631–2641 (2017).
48. Shi, J. et al. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661–667 (2015).
49. Platt, R. J. et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).
50. Daigle, S. R. et al. Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* **122**, 1017–1025 (2013).
51. Huyen, Y. et al. Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **432**, 406–411 (2004).
52. Dou, Y. et al. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).
53. van Nuland, R. et al. Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol. Cell Biol.* **33**, 2067–2077 (2013).
54. Kobayashi, N. et al. RanBPM, Muskelein, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8 α and ARMC8 β are components of the CTLH complex. *Gene* **396**, 236–247 (2007).
55. Kim, J.-H. et al. SON and its alternatively spliced isoforms control MLL complex-mediated H3K4me3 and transcription of leukemia-associated genes. *Mol. Cell* **61**, 859–873 (2016).
56. Mar, B. G. et al. Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. *Nat. Commun.* **5**, 1–6 (2014).
57. Parker, H. et al. Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. *Leukemia* **30**, 2179–2186 (2016).
58. Zhu, X. et al. Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat. Genet.* **46**, 287–293 (2014).
59. Krogan, N. J. et al. Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell Biol.* **23**, 4207–4218 (2003).
60. Qin, S. & Min, J. Structure and function of the nucleosome-binding PWWP domain. *Trends Biochem. Sci.* **39**, 536–547 (2017).
61. Li, J. et al. SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget* **7**, 50719–50734 (2015).
62. Santos, M. A. et al. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. *Nature* **514**, 107–111 (2014).
63. Dow, L. E. et al. A pipeline for the generation of shRNA transgenic mice. *Nat. Protoc.* **7**, 374–393 (2012).
64. Fellmann, C. et al. An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1–10 (2013).
65. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
66. Zuber, J. et al. Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat. Biotechnol.* **29**, 79–83 (2011).
67. Tarumoto, Y. et al. LKB1, salt-inducible kinases, and MEF2C are linked dependencies in acute myeloid leukemia. *Mol. Cell* **69**, 1017–1027 (2018).
68. Pabst, C. et al. Identification of small molecules that support human leukemia stem cell activity ex vivo. *Nat. Methods* **11**, 4–6 (2014).

69. Ruepp, A. et al. CORUM: the comprehensive resource of mammalian protein complexes—2009. *Nucleic Acids Res.* **38**, D497–D501 (2010).
70. Colinge, J. et al. Building and exploring an integrated human kinase network: global organization and medical entry points. *J. Proteom.* **107**, 113–127 (2014).
71. Menche, J. et al. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science* **347**, 1257601 (2015).
72. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
73. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
74. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
75. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).

Acknowledgements

We thank G. Winter for SpCas9-expressing MV4-11 cells and J. Loizou, V. Sexl, and A. Villunger for critical advice on the manuscript. We also thank all members of the Grebien and Superti-Furga laboratories for discussions. We are grateful for the expert technical assistance of A. Mazouzi with comet assays. We thank G. Stefanl for storing and preparing primary leukemia cells, A. Spittler and G. Hofbauer (Core Facility Flow Cytometry, Medical University of Vienna) for cell sorting, the IMP animal facility staff for animal caretaking and the Biomedical Sequencing Facility at CeMM for assistance with next-generation sequencing. N. Krall is acknowledged for preparing ethics application 1676/2016 for work with biobanked samples. A.S. was supported by the European FP7 Marie-Curie Initial Training Network HEM_ID. B.G. was supported by the NVKP_16-1-2016-0037 grant of the National Research, Development and Innovation Office, Hungary. C.R.V. is supported by the National Institutes of Health grant NCI RO1 CA174793, NCI 5P01CA013106-Project 4, and a Leukemia & Lymphoma Society Scholar Award. Research in the Valent laboratory is supported by the Austrian Science Fund (FWF) F4704. Research in the Zuber laboratory is supported by a SFB grant (F4710) of the Austrian Science Fund (FWF), and generous institutional funding by Boehringer Ingelheim. Research in the Superti-Furga Laboratory is supported by ERA-NET Grant I 2192-B22, the FWF Stand-Alone Grant P 29250-B30, and FWF SFB Grant F 4711-B20. Work in the Grebien laboratory is supported by grant no. 857935 from the Austrian Research Promotion Agency (FFG) to J.E. This project has received funding from the European Research Council under the European Union's Horizon 2020 research and

innovation programme (grant agreement n° 336860/StG to J.Z., grant agreements n° 695214/AdG, and n° 727416/PoC to G.S.F. and grant agreement n° 636855/StG to F.G.).

Author contributions

A.S., J.E., J.S., M.R., S.V., and F.G. planned and performed experiments and analyzed results. T.E., A.C.R., A.St., and B.G. performed bioinformatic analyses. M.R., M.M., J.J., M.A., B.L., C.V., and J.Z. provided protocols, cell lines, and vectors used in the study. A. C.M. and K.L.B. performed MS analysis, planned, and performed experiments and analyzed results. P.V. provided access to patient materials. A.S., J.Z., G.S.F., and F.G. designed the study, planned experiments, analyzed results, and wrote the paper.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-04329-y>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

Supplementary Information to the publication

MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity

Anna Skucha^{1,2}, Jessica Ebner², Johannes Schmöller², Mareike Roth³, Thomas Eder², Adrián César-Razquin¹, Alexey Stukalov¹, Sarah Vittori¹, Matthias Muhar³, Bin Lu⁴, Martin Aichinger³, Julian Jude³, André C. Müller¹, Balázs Györfy⁵, Christopher R. Vakoc⁴, Peter Valent⁶, Keiryn L. Bennett¹, Johannes Zuber^{3*}, Giulio Superti-Furga^{1,7*}, Florian Grebien^{2,8*#}

¹ CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria

² Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria

³ Research Institute of Molecular Pathology, Vienna, Austria

⁴ Cold Spring Harbor Laboratory, Cold Spring Harbor, USA

⁵ MTA TTK Lendület Cancer Biomarker Research Group, Institute of Enzymology, Second Department of Pediatrics, Semmelweis University, Budapest, Hungary

⁶ Department of Internal Medicine I. Division of Hematology & Hemostaseology, Ludwig Boltzmann Cluster Oncology, Medical University of Vienna; Austria

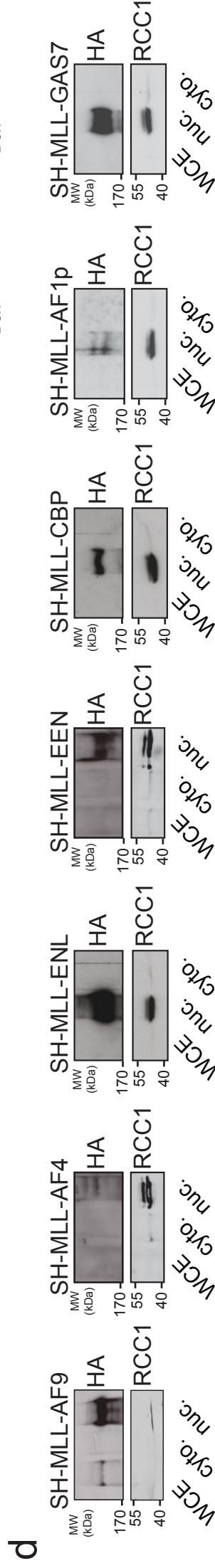
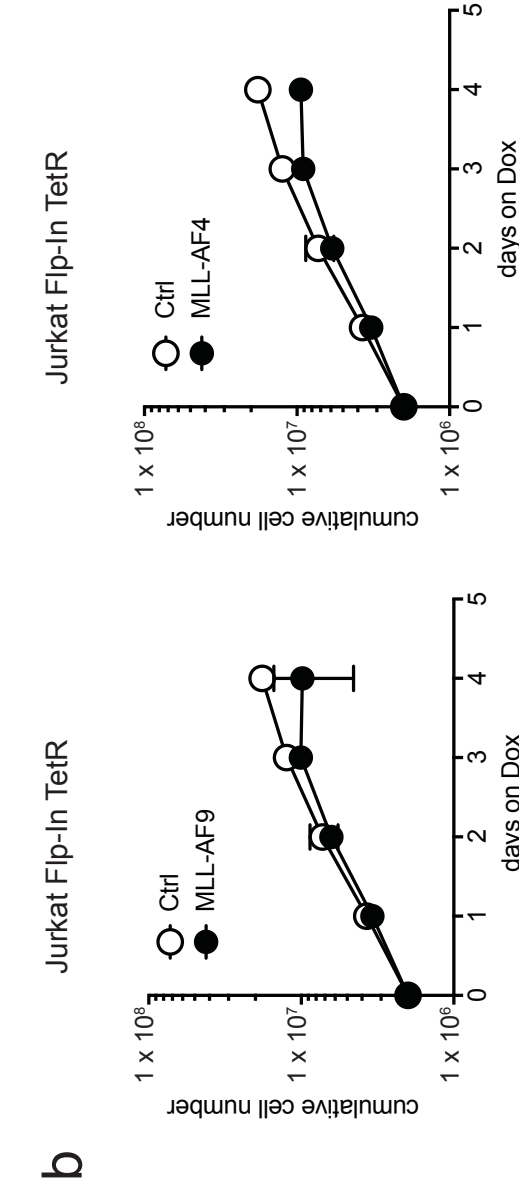
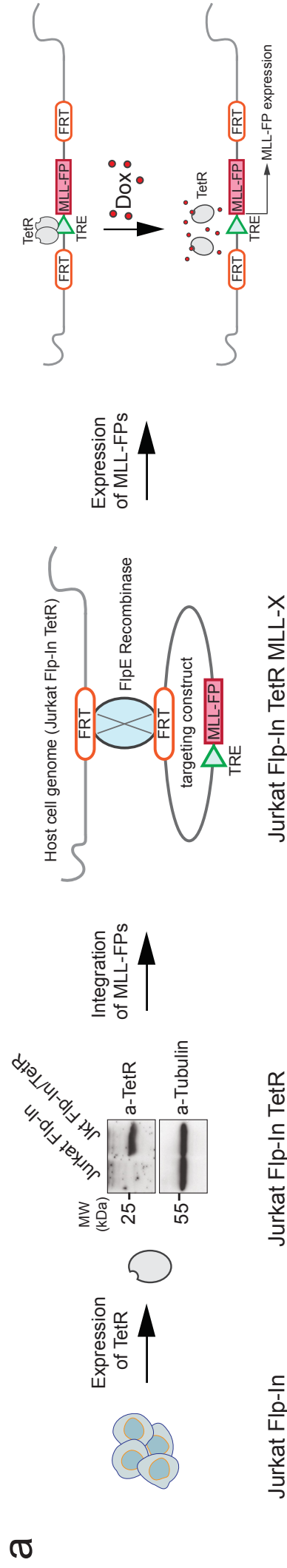
⁷ Center for Physiology and Pharmacology, Medical University of Vienna, Austria

⁸ Institute for Medical Biochemistry, University of Veterinary Medicine, Vienna, Austria

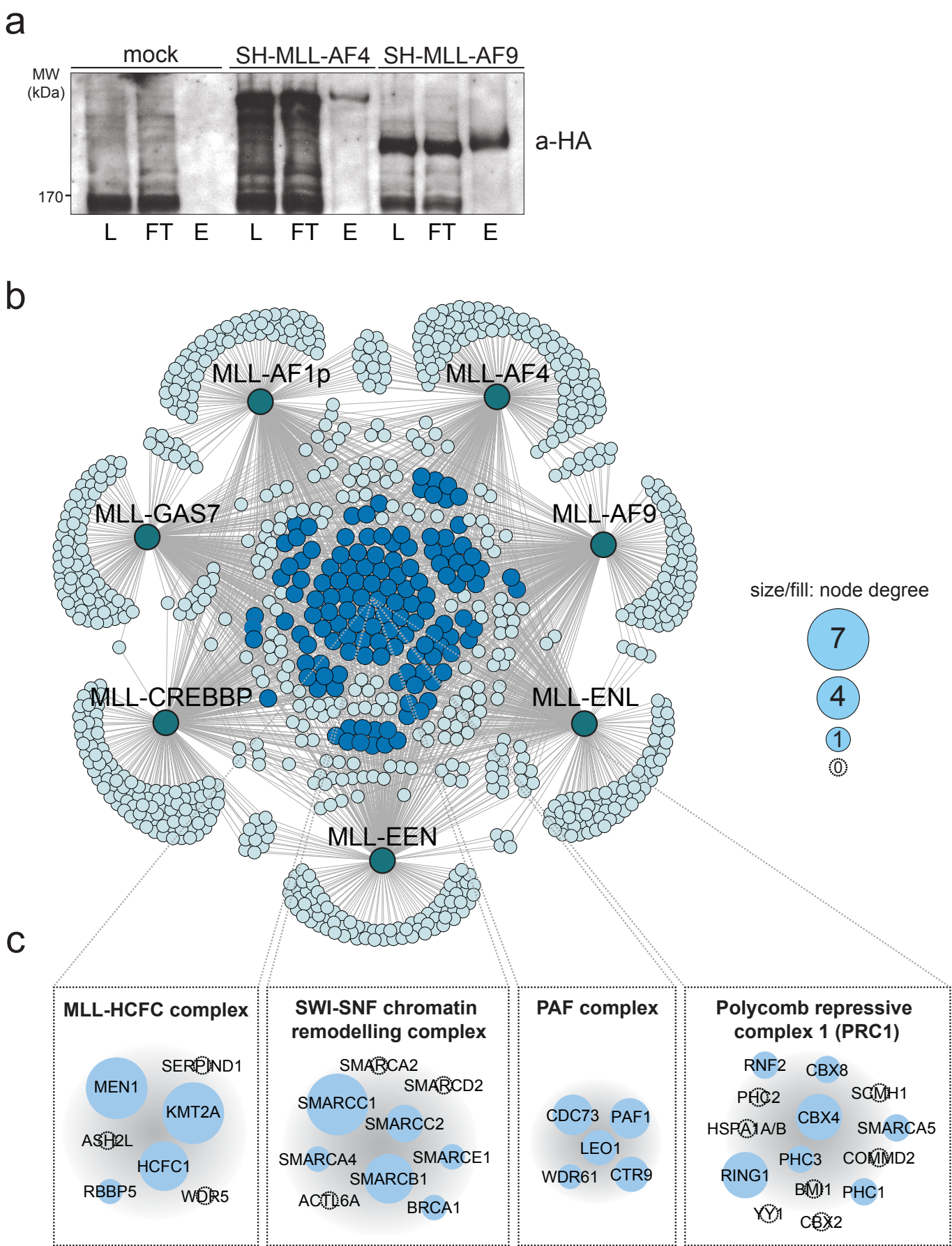
* these authors contributed equally to this work

lead contact: Florian Grebien:

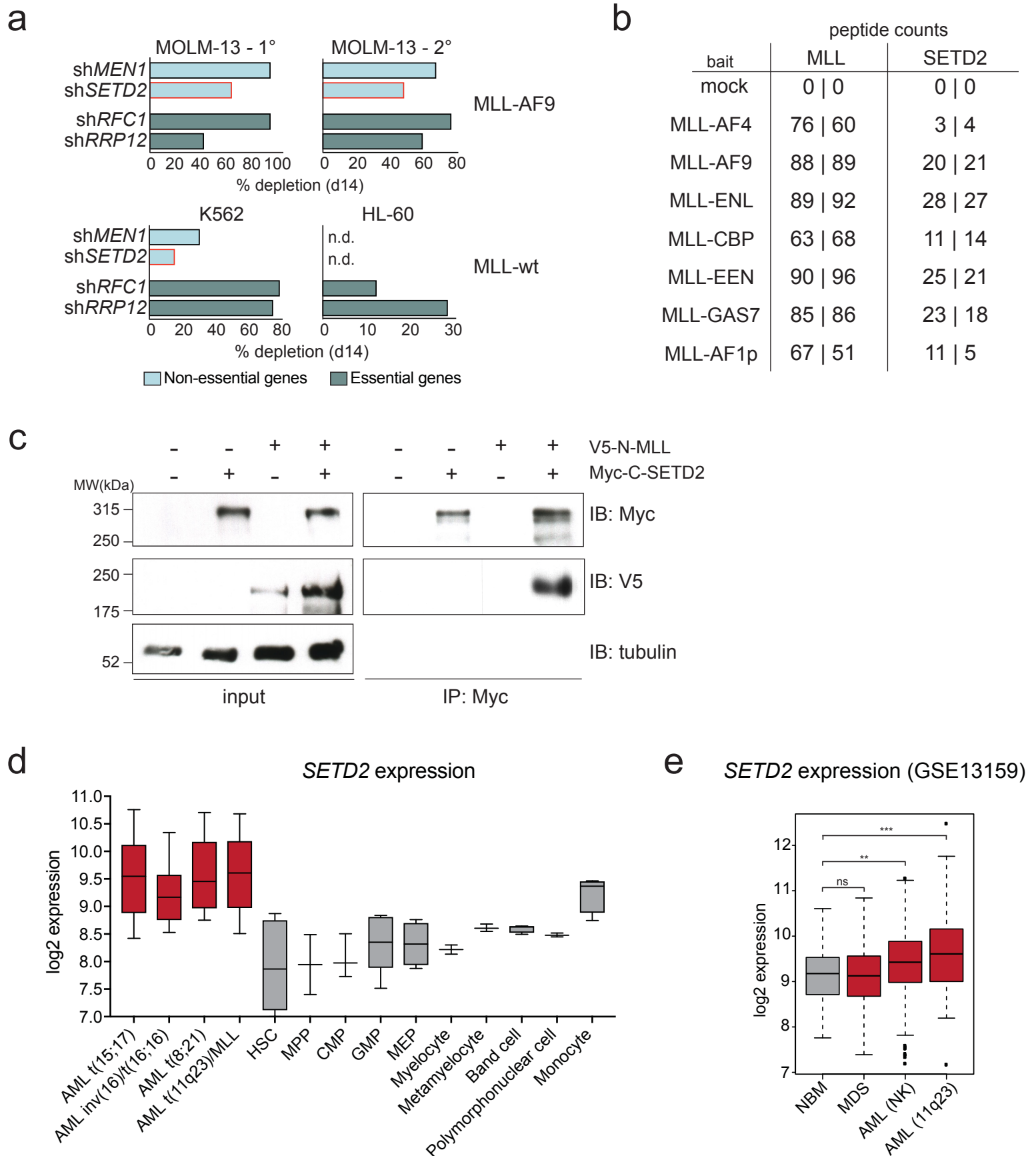
florian.grebien@lbcr.lbg.ac.at, florian.grebien@vetmeduni.ac.at



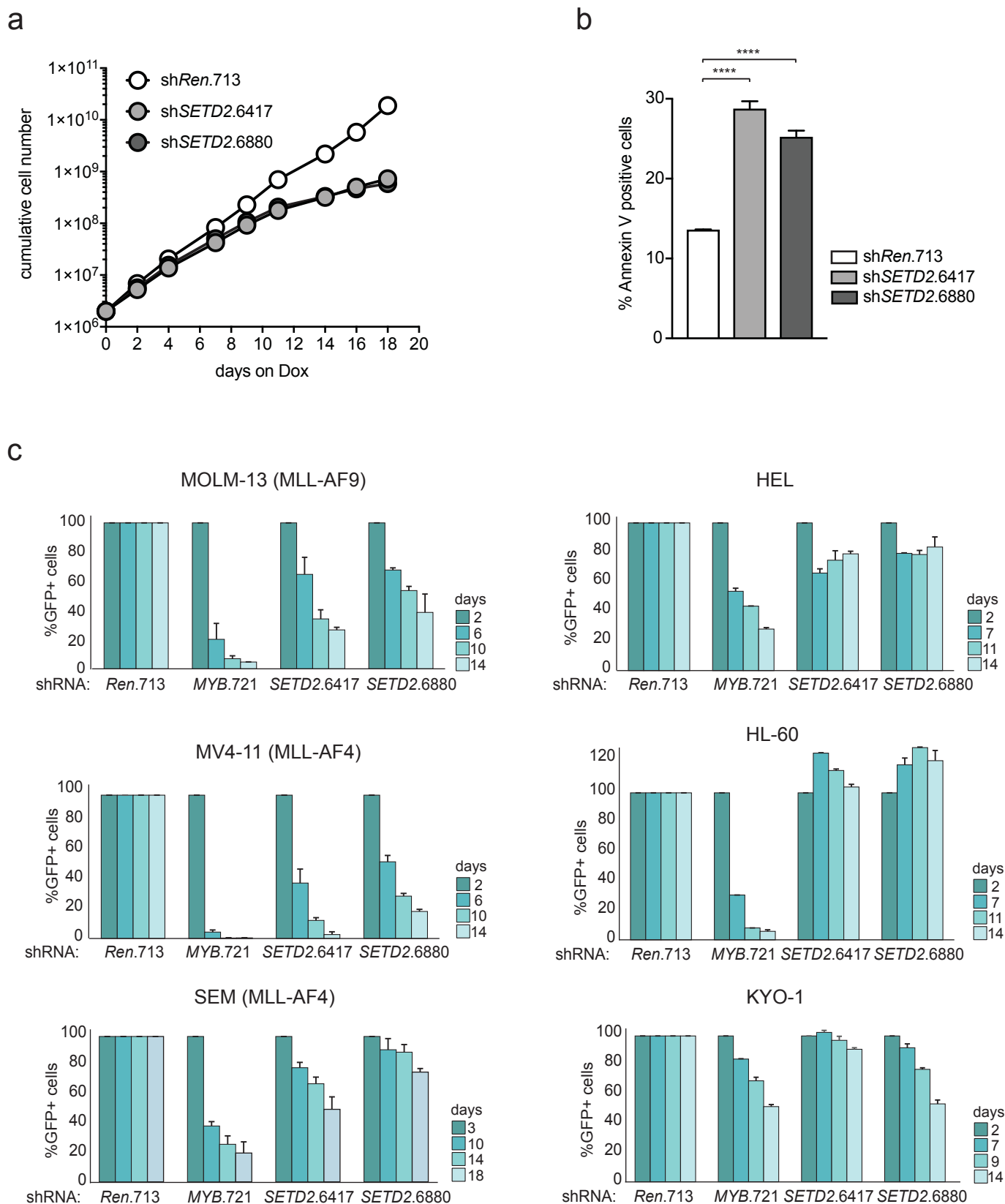
Supplementary Figure 1. Generation and characterization of isogenic cell lines for expression of affinity-tagged variants of MLL-fusion proteins. (a) Schematic illustration of the workflow for establishing the cell line models for affinity purification of MLL fusion proteins. The Flp-In system involves introduction of a Flp Recombination Target (FRT) site into the genome of the Jurkat cell line. An expression vector encoding for MLL-fusion protein is integrated into the genome via Flp recombinase-mediated DNA recombination at the FRT site. **(b)** Growth curves of Jurkat Flp-In cell lines expressing indicated MLL-fusion proteins (mean \pm s.d. n=3). **(c)** Cells expressing Strep-HA (SH)-tagged MLL fusion proteins were treated with Dox for 24h and transgene expression was monitored by qPCR (mean \pm s.d. n=3). **(d)** Extracts from Jurkat cells expressing indicated MLL-fusion proteins were fractionated and analyzed for expression of hemagglutinin (HA). RCC-1 was used as loading control (WCE: whole cell extract; nuc: nuclear fraction; cyto: cytosolic fraction).



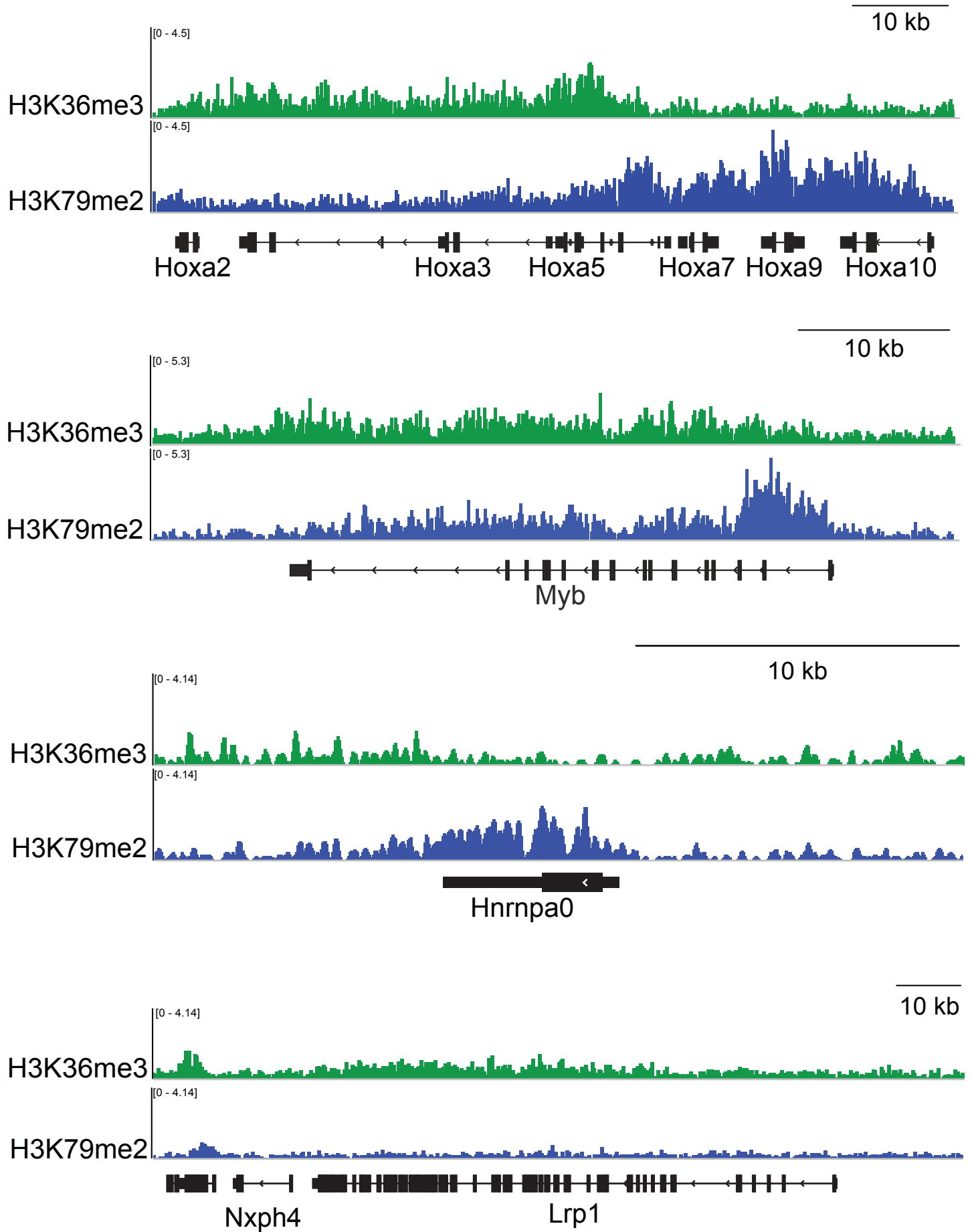
Supplementary Figure 2. Identification of protein complexes among the MLL-fusion protein-interactome. (a) Western blot (WB) analysis of lysate (L), flow-through (FT) and eluate (E) from StrepTactin-purifications of indicated affinity-tagged MLL-fusion proteins. **(b)** Network representation of the AP-MS-derived protein-protein interaction network of seven MLL-fusion proteins (dark green nodes), showing the top 300 interactors per bait ranked by p-value (light blue nodes). Nodes labelled in dark blue represent 128 proteins that interact with ≥ 5 of the 7 MLL-fusion proteins. **(c)** Representation of enriched protein complexes previously reported to be associated with MLL. The size and the color of the nodes correspond to the number of interaction partners.



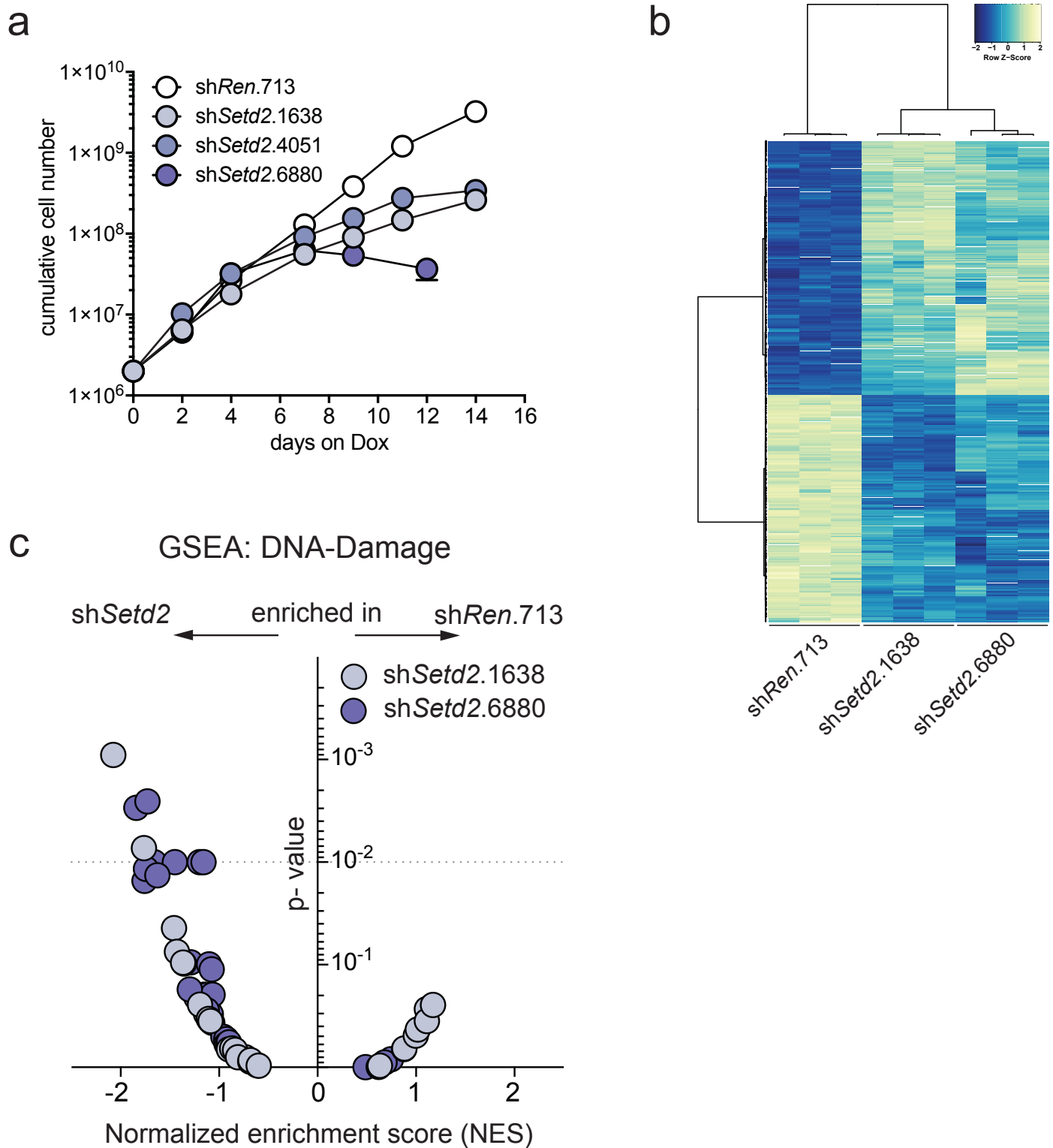
Supplementary Figure 3. shRNA screening of conserved MLL-interactors identifies SETD2 as a critical effector of MLL-fusion proteins. (a) Examples of depletion phenotypes of four selected genes in MLL-rearranged vs. MLL-wild type leukemia cell lines. Gene essentiality was assigned based on published datasets. (b) Peptide counts for MLL and SETD2 from AP-MS experiments of MLL-fusion proteins. (c) HEK293 rTA3 cells were transiently transfected with indicated constructs. Following immunoprecipitation using anti-Myc beads, eluates (IP) and whole cell extracts (input) were analyzed by immunoblotting with the indicated antibodies. (d) Expression levels of SETD2 in leukemia and normal hematopoiesis (bloodspot.binf.ku.dk). (e) Expression levels of SETD2 in different AML subtypes (NBM: Normal Bone marrow; MDS: Myelodysplastic Syndrome; AML (NK): Normal Karyotype; AML (11q23): AML harboring rearrangements of the MLL gene). ns, not significant, ** $p < 0.01$, *** $p < 0.001$ (t-test).



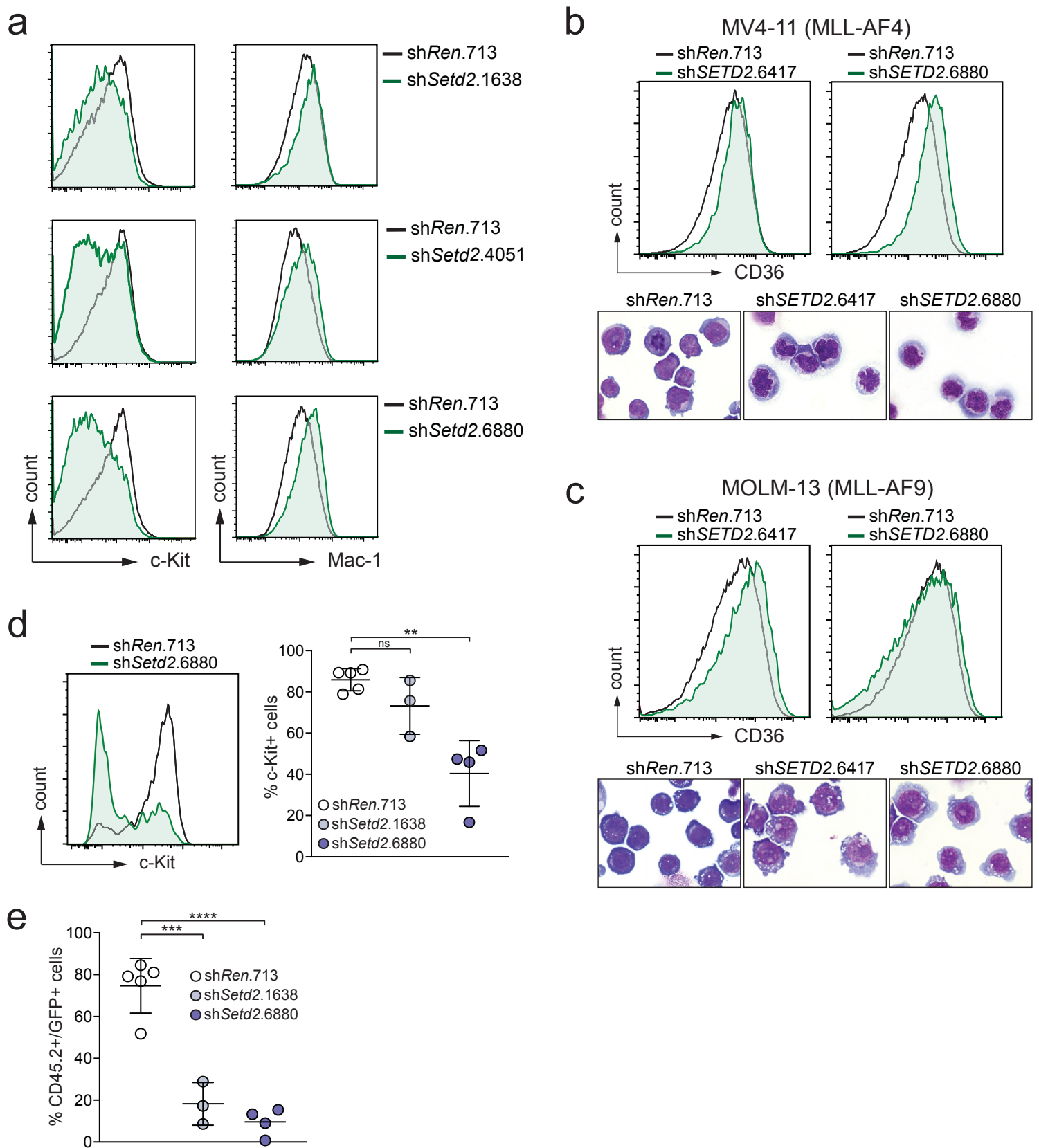
Supplementary Figure 4. shRNA-mediated knockdown of SETD2 impairs proliferation of human MLL-rearranged cell lines. (a) Growth curves of MV4-11 cells expressing indicated shRNAs (mean \pm s.d. n=3). **(b)** Quantification of flow cytometric analysis of apoptosis as measured by Annexin V-staining in MV4-11 cells. Cells were treated with Dox for 11 days and stained according to manufacturer's protocol. **** p<0.0001 (t-test) (mean \pm s.d. n=3). **(c)** Results of FACS-based competitive proliferation assay shown as the percentage of GFP-positive cells expressing individual SETD2-targeting shRNAs in the presence of Dox over 14-18 days in indicated human cell lines (mean \pm s.d. n=2).



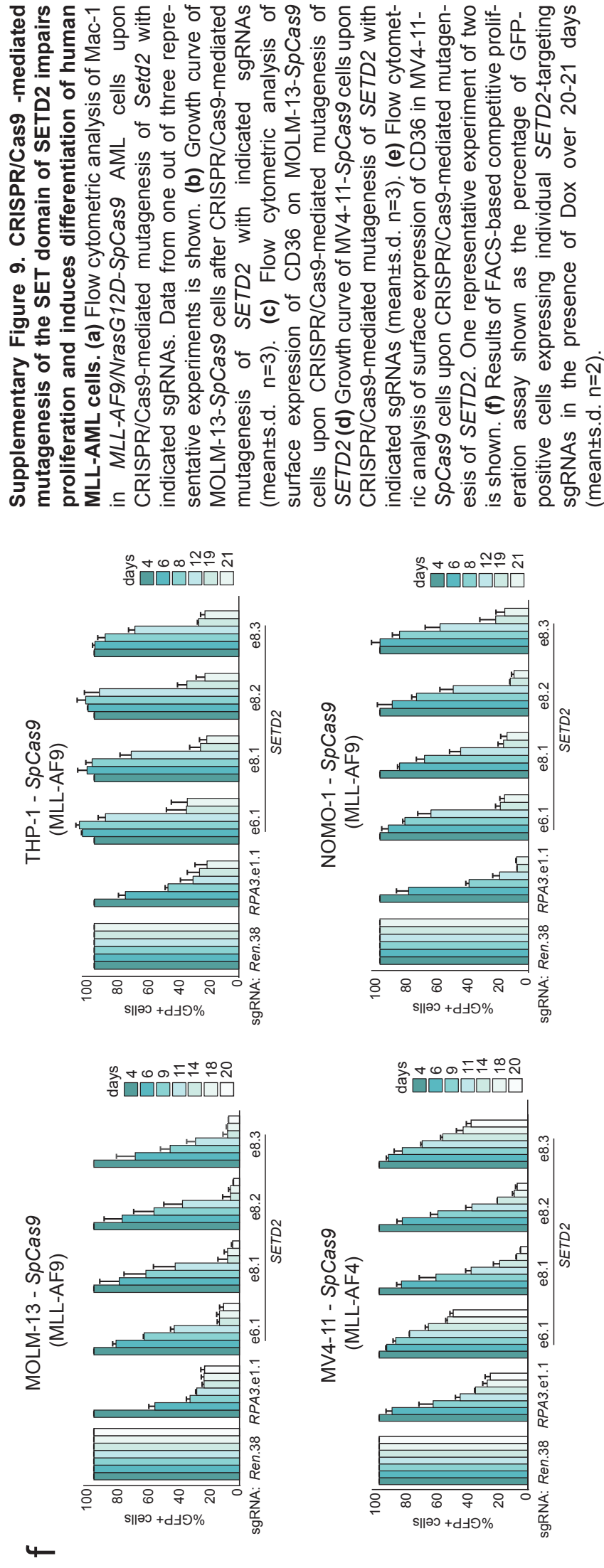
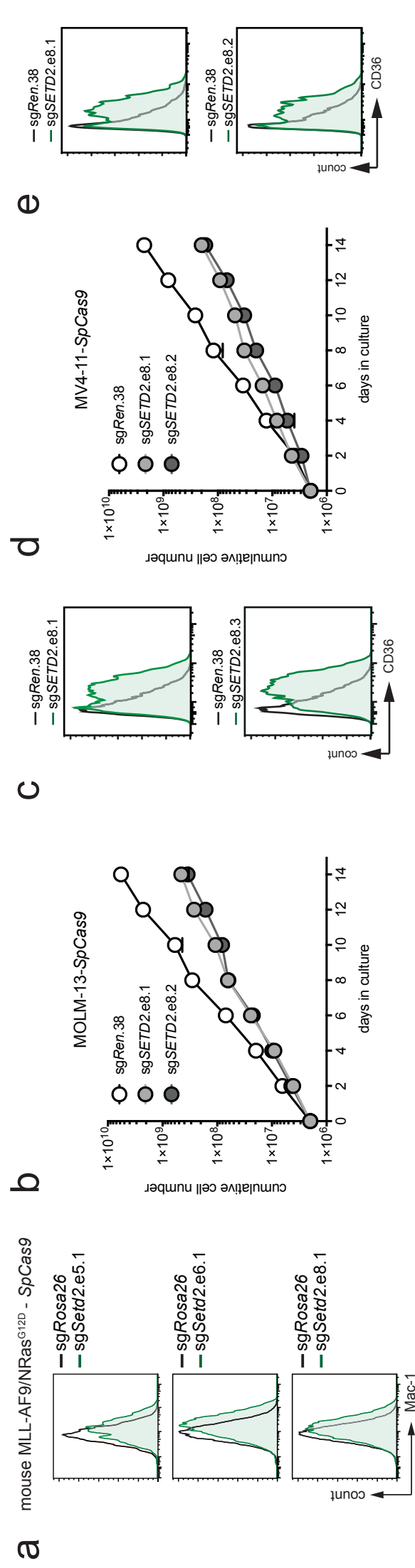
Supplementary Figure 5. MLL-target genes show high levels of H3K36me3 and H3K79me2. H3K36me3 (green) and H3K79me2 profiles (blue) of selected MLL-AF9 target genes Hoxa-cluster genes and Myb vs. non-MLL target genes (Hnrnpa0 and Lrp1).



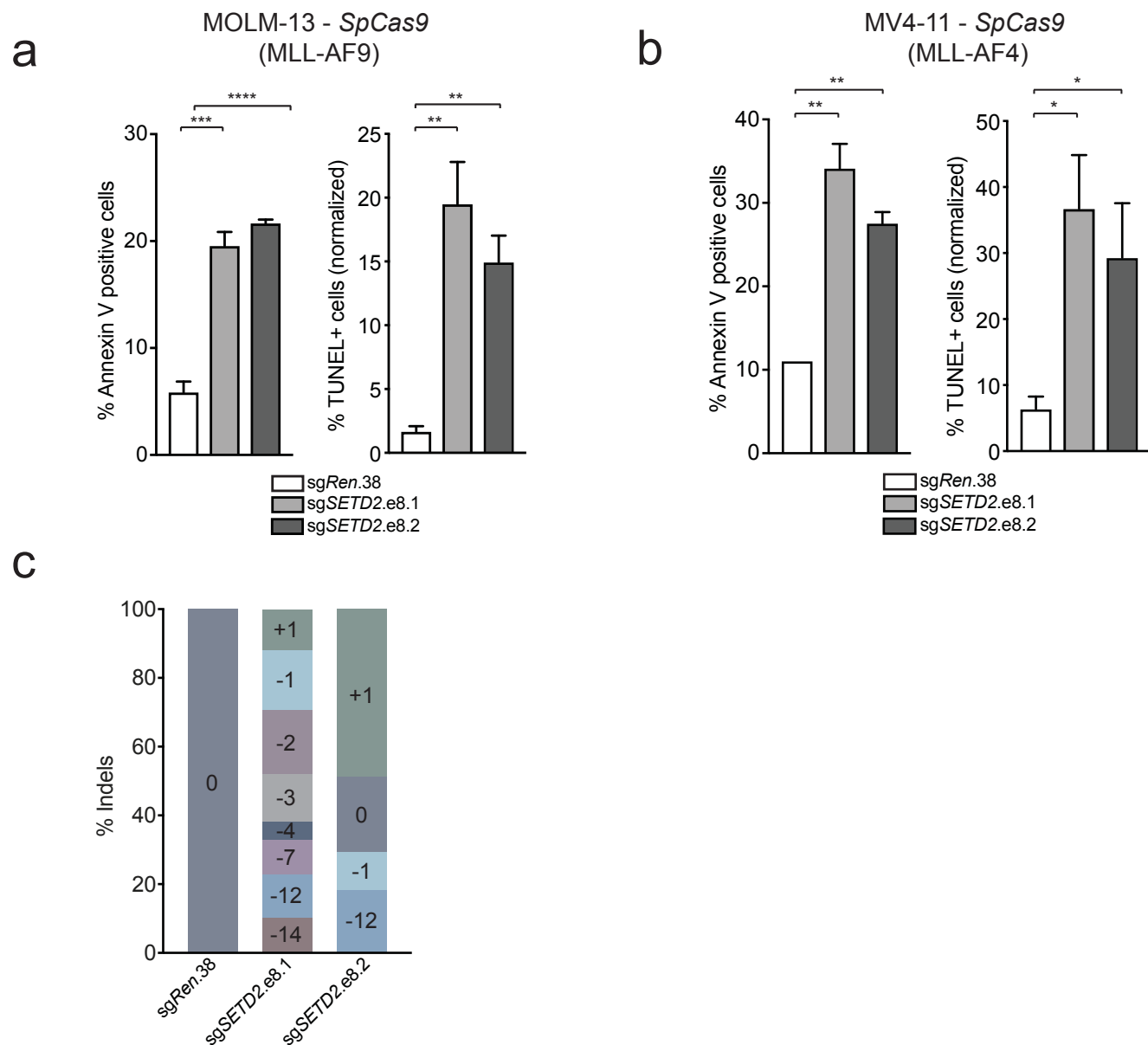
Supplementary Figure 6. Loss of SETD2 induces DNA damage-associated transcriptional changes in MLL-AF9 expressing cells. (a) Growth curve of mouse *MLL-AF9/NrasG12D* AML cells expressing indicated shRNAs. **(b)** Heatmap representation of differentially expressed genes in *MLL-AF9/NrasG12D* AML cells upon *Setd2* knockdown (mean \pm s.d. n=3). **(c)** Diagram showing results from GSEA of gene sets associated with DNA damage (from MSigDB) in *MLL-AF9/NrasG12D* AML cells upon shRNA-mediated knockdown of *Setd2*.



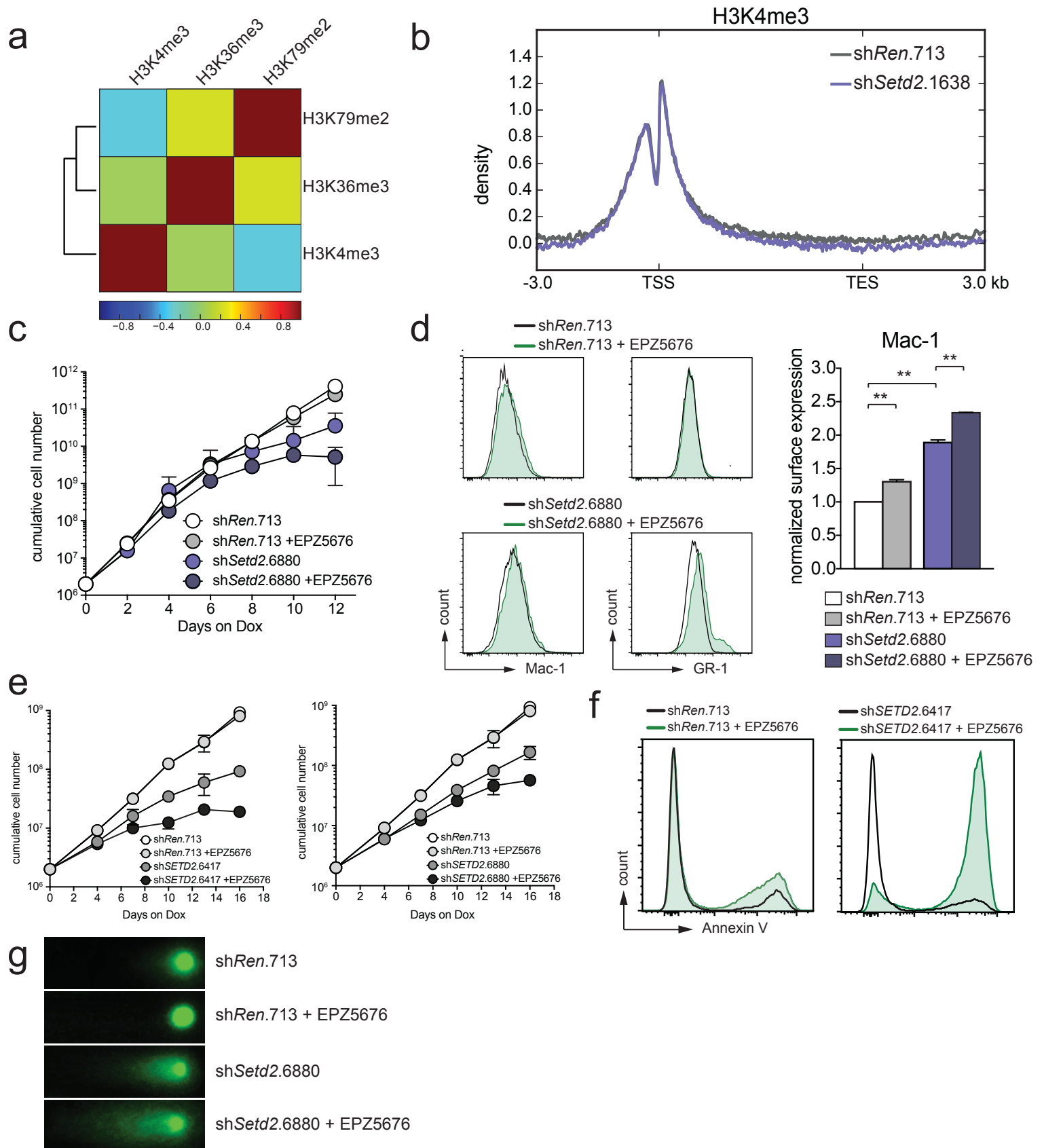
Supplementary Figure 8. shRNA-mediated SETD2 downregulation leads to myeloid differentiation of leukemia cells *in vitro* and *in vivo*. (a) Flow cytometric analysis of c-Kit and Mac-1 in *MLL-AF9/NrasG12D* AML cells upon *Setd2* knockdown. Data from one out of three representative experiments are shown. (b) Flow cytometric analysis of CD36 in MV4-11 cells upon *SETD2* knockdown (top). Data from one out of three representative experiments are shown. Micrographs of cytospin preparations of MV4-11 cells after expression of indicated shRNAs (bottom). (c) Flow cytometric analysis of CD36 on MOLM-13 cells upon *SETD2* knockdown (top). Data from one out of three representative experiments are shown. Micrographs of cytospin preparations of MV4-11 cells after expression of indicated shRNAs (bottom). (d) Flow cytometric analysis of c-Kit on *MLL-AF9/NrasG12D* AML cells upon *Setd2* knockdown *in vivo* (left). Quantification of flow cytometric analysis of c-Kit on *MLL-AF9/NrasG12D* AML cells upon *Setd2* knockdown with indicated shRNAs *in vivo* (right) (mean±s.d. n≥3) (e) Quantification of flow cytometric analysis of CD45.2/GFP on *MLL-AF9/NrasG12D* AML cells upon *Setd2* knockdown with indicated shRNAs *in vivo*. ns, not significant, ** p<0.01, *** p<0.001, **** p<0.0001 (t-test) (mean±s.d. n≥3).



Supplementary Figure 9. CRISPR/Cas9 -mediated mutagenesis of the SET domain of SETD2 impairs proliferation and induces differentiation of human MLL-AML cells. (a) Flow cytometric analysis of Mac-1 in MLL-AF9/NrasG12D-SpCas9 AML cells upon CRISPR/Cas9-mediated mutagenesis of Setd2 with indicated sgRNAs. Data from one out of three representative experiments is shown. (b) Growth curve of MOLM-13-SpCas9 cells after CRISPR/Cas9-mediated mutagenesis of SETD2 with indicated sgRNAs (mean \pm s.d. n=3). (c) Flow cytometric analysis of surface expression of CD36 on MOLM-13-SpCas9 cells upon CRISPR/Cas9-mediated mutagenesis of SETD2 (d) Growth curve of MV4-11-SpCas9 cells upon CRISPR/Cas9-mediated mutagenesis of SETD2 with indicated sgRNAs (mean \pm s.d. n=3). (e) Flow cytometric analysis of surface expression of CD36 in MV4-11-SpCas9 cells upon CRISPR/Cas9-mediated mutagenesis of SETD2. One representative experiment of two is shown. (f) Results of FACS-based competitive proliferation assay shown as the percentage of GFP-positive cells expressing individual SETD2-targeting sgRNAs in the presence of Dox over 20-21 days (mean \pm s.d. n=2).



Supplementary Figure 10. Mutagenesis of the SETD2 SET domain leads to apoptosis of MLL-fusion expressing cells. (a) Flow cytometric apoptosis detection as measured by Annexin V- (left) and TUNEL staining (right) in MOLM-13-*SpCas9* cells upon CRISPR/Cas9-mediated mutagenesis of *SETD2* with indicated sgRNAs (mean±s.d. n=3). **(b)** Flow cytometric apoptosis detection as measured by Annexin V- (left) and TUNEL staining (right) in MV4-11-*SpCas9* cells upon CRISPR/Cas9-mediated mutagenesis of *SETD2* with indicated sgRNAs (mean±s.d. n=3). **(c)** Analysis of indel formation by *SETD2*-targeting sgRNAs in MV4-11-*SpCas9* cells using Sanger sequencing and TIDE analysis. * $p > 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ (t-test).



Supplementary Figure 11. SETD2 loss disrupts the H3K36me3-H3K79me2 signature on MLL-target genes and sensitizes AML cells to DOT1L inhibition. (a) Heatmap representation of the Pearson correlation coefficients between changes in the respective histone marks over gene bodies from ChIP-Rx experiments after shRNA-mediated knockdown of *Setd2* in *MLL-AF9/NrasG12D* cells. (b) Metagene plots of ChIP-Rx data for H3K4me3 after *Setd2* knockdown. (c) Growth curves of *MLL-AF9/NrasG12D* AML cells treated with EPZ5676 (500 nM) upon shRNA-mediated mutagenesis of *Setd2* with indicated shRNAs (mean±s.d. n=3). (d) Flow cytometric analysis of Mac-1 and Gr-1 expression in *MLL-AF9/NrasG12D* cells treated with EPZ5676 (500nM) upon shRNA-mediated knockdown of *Setd2* (left). Quantification of surface expression of Mac-1 on *MLL-AF9/NrasG12D* AML cells treated with EPZ5676 (500nM) after shRNA-mediated knockdown of *Setd2* (mean±s.d. n=2) (right). (e) Growth curves of MV4-11 cells upon shRNA-mediated knockdown of *SETD2* with indicated shRNAs treated with EPZ5676 (50nM) (mean±s.d. n=3). (f) Flow cytometric analysis of apoptosis induction as measured by Annexin V-staining in MV4-11 cells treated with EPZ5676 (50nM) after shRNA-mediated knockdown of *SETD2*. (g) Representative micrographs of *MLL-AF9/NrasG12D* AML cells showing DNA breaks after treated with EPZ5676 (500nM) upon shRNA-mediated knockdown of *Setd2*. ** p<0.01 (t-test).

2.3 Interlude

Pharmacological targeting of the Wdr5-MLL interactions in C/EBP α N-terminal leukemia.

Florian Grebien*, Masoud Vedadi*, Matthäus Getlik*, Roberto Giambruno, Amit Grover, Roberto Avellino, Anna Skucha, Sarah Vittori, Ekatarina Kuznetsova, David Smil, Dalia Barsyte-Lovejoy, Fengling Li, Gennadiy Poda, Matthieu Schapira, Hong Wu, Aiping Dong, Guillermo Senisterra, Alexey Stukalov, Kilian V. M. Huber, Andreas Schönegger, Richard Marcellus, Martin Bilban, Christoph Bock, Peter J. Brown, Johannes Zuber, Keiryn L. Bennett, Rima Al-awar, Ruud Delwel, Claus Nerlov, Cheryl H. Arrowsmith*, Giulio Superti-Furga*

* equal contribution

lead contact: Giulio Superti-Furga and Cheryl H. Arrowsmith

Nature Chemical Biology

Published: 13 July 2015

DOI: 10.1038/NCHEMBIO.1859

Mutations in the CEBPA gene, encoding a master regulator of myeloid gene expression, are found in about 9% of patients presenting with AML. Predominantly, frameshift mutations were identified of the N-terminal part of C/EBP α , which result in loss of expression of the full-length p42 C/EBP α isoform but retain expression of a shorter p30 variant from an ATG codon downstream of the mutated region. Although large number of interaction partners of C/EBP α have been identified, it is not clear whether these proteins' effector functions differ between p42 vs. p30 C/EBP α isoforms. Here we hypothesize that different C/EBP α variants associate with different molecular machineries, which might reflect different cellular processes. Interaction proteomics identified the essential MLL complex component WDR5 to preferentially interact with the p30 isoform of C/EBP α . Wdr5 was responsible for the transforming properties of mutated C/EBP α . Furthermore, we identified and characterized OICR-9429 as a novel small molecule, which disrupts the interaction of Wdr5 and MLL. OICR-9429 was able to release the p30-dependent differentiation and induced to apoptosis in CEBPA-mutant AML cells. Thus, targeting of the MLL complex might represent an alternative strategy in AML patients with CEBPA-mutations.

F.G., M.V., R.G., A.G., R.A., **A. Skucha**, S.V., E.K., D.B.-L., F.L, G.S., K.V.M.H. and R.M. planned, performed and analyzed biochemical, biophysical, cellular and *in vivo* experiments. M.G., D.S., G.P., M.S., P.J.B. and R.A. contributed to chemical design and

synthesis of OICR-9429 and OICR-0547. H.W., A.D. and M.S. solved and analyzed the X-ray crystal structure of WDR5 in complex with OICR-9429. A. Stukalov, A. Schönegger, M.B. and C.B. performed bioinformatic analyses. J.Z., K.L.B., R.D. and C.N. provided access to vital tools and technologies, planned experiments and analyzed results. F.G., M.V., P.J.B., R.A., C.H.A. and G.S.-F. designed the study, planned experiments, analyzed results and wrote the paper.

2.4 Pharmacological targeting of the Wdr5-MLL interactions in C/EBP α N-terminal leukemia

Pharmacological targeting of the Wdr5-MLL interaction in C/EBP α N-terminal leukemia

Florian Grebien^{1,2,14*}, Masoud Vedadi^{3,4,14}, Matthäus Getlik^{5,14}, Roberto Giambruno¹, Amit Grover⁶, Roberto Avellino⁷, Anna Skucha¹, Sarah Vittori¹, Ekaterina Kuznetsova³, David Smil³, Dalia Barsyte-Lovejoy³, Fengling Li³, Gennadiy Poda^{5,8}, Matthieu Schapira^{3,4}, Hong Wu³, Aiping Dong³, Guillermo Senisterra³, Alexey Stukalov¹, Kilian V M Huber¹, Andreas Schönegger¹, Richard Marcellus⁵, Martin Bilban⁹, Christoph Bock¹, Peter J Brown³, Johannes Zuber¹⁰, Keiryn L Bennett¹, Rima Al-awar^{4,5}, Ruud Delwel⁷, Claus Nerlov⁶, Cheryl H Arrowsmith^{11,12,14} & Giulio Superti-Furga^{1,13,14*}

The *CEBPA* gene is mutated in 9% of patients with acute myeloid leukemia (AML). Selective expression of a short (30-kDa) CCAAT-enhancer binding protein- α (C/EBP α) translational isoform, termed p30, represents the most common type of *CEBPA* mutation in AML. The molecular mechanisms underlying p30-mediated transformation remain incompletely understood. We show that C/EBP α p30, but not the normal p42 isoform, preferentially interacts with Wdr5, a key component of SET/MLL (SET-domain/mixed-lineage leukemia) histone-methyltransferase complexes. Accordingly, p30-bound genomic regions were enriched for MLL-dependent H3K4me3 marks. The p30-dependent increase in self-renewal and inhibition of myeloid differentiation required Wdr5, as downregulation of the latter inhibited proliferation and restored differentiation in p30-dependent AML models. OICR-9429 is a new small-molecule antagonist of the Wdr5-MLL interaction. This compound selectively inhibited proliferation and induced differentiation in p30-expressing human AML cells. Our data reveal the mechanism of p30-dependent transformation and establish the essential p30 cofactor Wdr5 as a therapeutic target in *CEBPA*-mutant AML.

Leukemia is characterized by deregulation of hematopoietic progenitor cell function, leading to ectopic self-renewal and disruption of normal differentiation properties. As lineage-specific transcription factors are key regulators of hematopoietic homeostasis, their dysregulated activity is frequently associated with leukemia.

The basic-region leucine zipper transcription factor C/EBP α is a master regulator of myeloid gene expression programs in the hematopoietic system, linking terminal differentiation to growth arrest¹. C/EBP α deficiency leads to a complete block of terminal myeloid differentiation at the pre-granulocyte/monocyte-cell stage².

CEBPA mutations are present in 9% of patients presenting with AML³. The most prevalent type of mutation involves frameshifts in the N-terminal part of the C/EBP α coding sequence⁴. These mutations ablate expression of the full-length p42 isoform but still allow a shorter protein (termed p30) to be expressed from an AUG codon downstream of the mutated region. Under physiological circumstances, nutrient and growth factor availability decreases the p42/p30 ratio by increasing bypass of the upstream initiation codon, promoting the maintenance of an undifferentiated cellular state⁵. Mice engineered to express only the p30 variant of C/EBP α (*Cebpa*^{p30/p30} genotype) develop AML with complete penetrance⁶. *Cebpa*^{p30/p30} hematopoietic progenitors hyperproliferate *in vitro*⁶, and overexpression of p30 in hematopoietic progenitor cell lines blocks myeloid differentiation⁵.

Various hypotheses have been put forward to explain the molecular basis for the transforming function of C/EBP α p30. N-terminal deletions strongly reduce transcriptional activation by C/EBP α ^{7,8}. Yet certain properties that are exerted by the p42 isoform are still maintained by the p30 variant. These include the *trans*-activating potential and the recruitment of the SWI/SNF (Switch/Sucrose Non-Fermentable) complex to regulate the expression of downstream target genes^{8,9}. In addition, C/EBP α p42 can inhibit the activity of E2F proteins through direct, physical interaction¹⁰. This is critically required for C/EBP α -dependent induction of terminal myeloid differentiation¹¹, with c-Myc as a critical E2F-regulated C/EBP α target^{12,13}. The ability of C/EBP α to repress E2F target genes is lost upon deletion of its N terminus^{11,12}, indicating that derepression of E2F target genes is key to leukemogenesis upon p42 loss. Additionally, C/EBP α p42 and p30 isoforms might have variable affinities for consensus C/EBP binding sites in the genome¹⁴.

Finally, although a plethora of interaction partners has been described for C/EBP α ^{15–19}, it is not clear to what extent these cofactors differ between p42 and p30 isoforms and whether any differences in the protein interactomes contribute to the differential effects of C/EBP α p42 versus p30.

We hypothesized that differences in the biochemical properties of C/EBP α isoforms could be reflected in a differential ability to engage the molecular machinery of the cell. The protein complexes associated with p42 or p30 should be able to account for differences

¹CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. ²Ludwig Boltzmann Institute for Cancer Research, Vienna, Austria. ³Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada. ⁴Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada. ⁵Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, Ontario, Canada. ⁶MRC Molecular Hematology Unit, Weatherall Institute of Molecular Medicine, Oxford, UK. ⁷Department of Hematology, Erasmus University Medical Center, Rotterdam, the Netherlands. ⁸Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada. ⁹Department of Laboratory Medicine and Core Facility Genomics, Core Facilities, Medical University Vienna, Vienna, Austria. ¹⁰Research Institute of Molecular Pathology, Vienna, Austria. ¹¹Princess Margaret Cancer Centre, Toronto, Ontario, Canada. ¹²Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada. ¹³Center for Physiology and Pharmacology, Medical University Vienna, Vienna, Austria.

¹⁴These authors contributed equally to this work. *e-mail: florian.grebien@lbcir.lbg.ac.at, carrow@uhnresearch.ca or gusuperti@cemm.oew.ac.at

in their biological activities, in particular with respect to altered regulation of self-renewal, proliferation and differentiation. Thus, the identification of any factor conferring differential and druggable vulnerability to p30-expressing cells would represent an attractive target for therapy. Here we identified Wdr5 as such a target and demonstrate that a new, first-in-class small-molecule antagonist of the Wdr5-MLL interaction can selectively target AML cells with N-terminal CEBPA mutations.

RESULTS

p30 interacts with Wdr5 and colocalizes with H3K4me3

We expressed affinity-tagged variants of C/EBP α p42 and p30 in the mouse myeloid progenitor cell line FDCP-1 (Supplementary Results, Supplementary Fig. 1a). While p42 overexpression in this system induced the downregulation of the progenitor surface marker c-Kit as well as the acquisition of morphological signs of myeloid maturation, these features were blocked upon p30 overexpression (Supplementary Fig. 1b,c). Gene expression profiling revealed that p30 had profound effects on global gene expression patterns, including the upregulation of genes important for cell growth and division and the induction of E2F target genes (Supplementary Fig. 1d,e). This confirms that our isogenic cell system could recapitulate principal aspects of C/EBP α -mediated terminal myeloid differentiation as well as p30-associated changes in gene expression. We affinity-purified protein complexes containing p42 and p30 from nuclear extracts of these cell lines and characterized them by LC/MS/MS. Using stringent statistical evaluation, we found 64 proteins that reproducibly associated with p42, whereas 52 proteins purified together with p30 (Fig. 1a and Supplementary Tables 1 and 2). We recapitulated several previously published interactions of C/EBP α p42 (for example, C/EBP ζ , Ddit3 and Hdac1-Hdac2) under the conditions used, supporting the relevance of our interactome data. Only 14 proteins interacted with both p42 and p30 (Fig. 1a), suggesting that p30-selective protein-protein interactions could be instrumental to leukemogenesis. Further, we hypothesized that any p30-specific characteristics would mostly manifest on the level of transcription, chromatin regulation or both. Among all of the p30-specific interactors, three proteins were annotated with the Gene Ontology (GO) term 'transcription' and one protein with 'histone modification/chromatin remodeling'. Two proteins were positive for both selected GO terms (Fig. 1a). Of those, Smarcd2, a component of the SWI/SNF complex, has been previously shown to be associated with and regulate the activity of C/EBP α . The other p30-specific interactor was Wdr5, a component of SET/MLL histone methyltransferase (HMT) complexes. MLL belongs to the Trithorax (TrxG) family of proteins, which positively regulate gene expression through the catalysis of H3K4 methylation. The H3K4 methyltransferase activity of MLL critically depends on the presence of Wdr5 (ref. 20).

As Wdr5 is part of large, multisubunit protein complexes, we next tested whether p30 also displays enhanced association with reported Wdr5-MLL interaction partners. Out of 29 proteins previously reported to participate in Wdr5-MLL-containing complexes, 23 proteins (79%) were present at greater or equal levels in purifications of p30, whereas only 6 proteins (21%) showed enrichment in p42 analyses (Supplementary Fig. 2a). This indicates that, globally, p30 has a stronger tendency to incorporate into Wdr5-containing protein complexes than does p42, despite the fact that Wdr5 expression is equal in p42- and p30-expressing cells (Fig. 1b and Supplementary Fig. 2b). Notably, the selective interaction of p30 with Wdr5 could be confirmed by coimmunoprecipitation experiments in FDCP-1 cells (Supplementary Fig. 2b) and was conserved in a pair of isogenic mouse 32D cell lines engineered to express affinity-tagged p42 or p30 (Fig. 1b). Although the specificity of the Wdr5-p30 interaction was not retained upon overexpression in HEK293 cells, we observed a strong interaction between

endogenous Wdr5 and C/EBP α p30 proteins in a mouse *Cebpa*^{p30/p30} myeloid progenitor cell line (Supplementary Fig. 2c,d).

As the association of p30 with Wdr5 and components of the SET/MLL HMT complex suggested a specific role in the regulation of gene expression, we investigated the genome-wide distribution of p42 and p30. Both C/EBP α isoforms showed very similar distributions across genomic loci and displayed enriched chromatin binding around annotated transcription start sites, consistent with functional roles in transcriptional regulation (Supplementary Figs. 2e and 3a,b). Furthermore, 66% (5,170) of all sites occupied by p30 also overlapped with p42 binding, whereas p30 was exclusively bound to 2,663 sites (Fig. 1c and Supplementary Fig. 3c).

As Wdr5 is required for efficient H3K4 methylation by SET/MLL²⁰ and the H3K4me3 mark is strongly associated with actively transcribed gene promoters, we performed H3K4me3 chromatin immunoprecipitation (ChIP-seq) experiments in cells expressing p42 or p30. Despite the smaller overall number of binding sites, p30 ChIP-seq peaks showed greater overlap with H3K4me3-marked genomic regions than p42-bound sites (47% versus 32%, $P < 10^{-10}$; Fig. 1d,e). Only 18% of regions with only p42 binding but no p30 binding were H3K4me3 positive, whereas 53% of regions exclusively associated with p30 were marked with H3K4me3 (Fig. 1d). This strong positive correlation of p30 binding and H3K4me3 marks on chromatin is in line with increased interaction of p30 with Wdr5 and the SET/MLL HMT complex.

Assignment of ChIP-seq peaks with overlapping p30-H3K4me3 occupancy to gene targets based on their proximity to the nearest annotated transcription start site yielded 2,606 actively transcribed p30 target genes. 268 of these genes showed significant p30-dependent changes in mRNA expression levels (FDR 0.01, $P = 3.3 \times 10^{-7}$; Supplementary Fig. 4a). p30-bound and upregulated genes were enriched for functions in cell cycle control and mitosis, whereas p30-dependent downregulated genes were involved in the regulation of endoplasmic reticulum stress and transcriptional suppression (Supplementary Fig. 4b). Together, these data indicate that p30 preferably associates with Wdr5 and SET/MLL HMT complex components, leading to increased p30-H3K4me3 colocalization on chromatin.

The p30-mediated differentiation block is Wdr5 dependent

Next we tested the functional consequences of inducible Wdr5 loss on p30-dependent cellular features. In the system we used, shRNA expression is coupled to the induction of a fluorescent reporter after doxycycline (Dox) administration (Fig. 2a and Supplementary Fig. 5a). Two independent *Wdr5*-targeting shRNAs efficiently silenced *Wdr5* mRNA and protein expression in different cell systems (Supplementary Fig. 5b-e). Given the prominent p30-H3K4me3 colocalization on chromatin and the described role of Wdr5 in the SET/MLL HMT complex, we hypothesized that Wdr5 downregulation might have a specific effect on H3K4me3 marks at promoters of p30 target genes. Indeed, H3K4me3 levels on promoters of the p30-target genes *Cdk6*, *Etv6* and *Kit* were higher in FDCP-1 cells expressing p30 than in those expressing p42 (Supplementary Fig. 5f). Furthermore, we observed a clear reduction of H3K4me3 marks on these promoters upon shRNA-mediated depletion of Wdr5 in p30-expressing cells, whereas they remained unchanged in p42-expressing cells (Supplementary Fig. 5f).

Gene expression profiling in p30-expressing FDCP-1 cells revealed that knockdown of *Wdr5* led to increased expression of genes associated with terminal myeloid differentiation (Supplementary Fig. 5g). Notably, the same upregulation of myeloid genes was observed in p30-expressing 32D cells after *Wdr5* knockdown (Fig. 2b), pointing to a conserved role of Wdr5 in preventing the expression of differentiation-associated genes in p30-mutated leukemia. To test this hypothesis in a mechanistically defined model, we turned to granulocyte colony-stimulating factor (G-CSF)-induced

granulocytic maturation of 32D cells, which is characterized by upregulation of myeloid surface markers and induction of myeloid gene expression. Expression of p30 in 32D cells completely blocked all signs of G-CSF-induced maturation, validating this model for the monitoring of p30-dependent effects on myeloid differentiation (**Supplementary Fig. 6a,b**). If *Wdr5* was indeed a key effector of p30 action, then its loss of function should overcome the p30-dependent differentiation block. As expected, expression of a control shRNA (shRen.713) did not alter the expression of myeloid surface markers or genes associated with terminal granulocytic maturation (**Fig. 2c,d**). In contrast, suppression of *Wdr5* using two independent shRNAs triggered robust expression of Gr-1 and Mac-1 surface markers in response to G-CSF, resulting in a fourfold increase in the abundance of Mac-1-positive cells (**Fig. 2c** and **Supplementary Fig. 7a**). Furthermore, knockdown of *Wdr5* allowed high expression of the granulocyte marker genes *Lcn2* and *Lyz2* in G-CSF-treated 32D cells expressing p30 (**Fig. 2d**). Notably, the G-CSF-mediated differentiation-promoting effect was restricted to p30-expressing cells. Knockdown of *Wdr5* in mock- or C/EBP α p42-transduced cells only slightly increased the abundance of Mac-1-positive cells and myeloid marker gene expression, as the basal levels of both parameters were elevated in these cell types (**Supplementary Fig. 7a–e**). Also, *Wdr5* downregulation did not have any effect in culture conditions that do not permit myeloid differentiation (**Supplementary Fig. 7f**). Finally, knockdown of *WDR5* did not result in altered proliferation rates in human K562 and U-937 cells, further supporting the specificity of the observed effects in p30-expressing cells (**Supplementary Fig. 8**).

As *Wdr5* was shown to be a critical mediator of MLL activity, we next tested the involvement of the SET domain-containing MLL enzymes MLL1, MLL2, MLL3 and MLL4 (also known as Wbp7) in maintaining the p30-induced differentiation block of 32D cells. Only shRNA-mediated knockdown of *Mill3*, and to a lesser extent *Mill2*, could phenocopy the differentiation-inducing effect of *Wdr5* downregulation in p30-expressing 32D cells (**Supplementary Fig. 9a–c**). This suggests differential preferences of *Wdr5* in its functional interaction with different MLL family members.

Hence, *Wdr5* is a critical effector of C/EBP α p30 action, as inactivation of a *Wdr5*-MLL1 axis reduces H3K4me3 levels on p30 target genes and leads to the upregulation of myeloid marker genes, culminating in a release of the p30-dependent block of myeloid differentiation.

p30-induced self-renewal and leukemogenesis require *Wdr5*

We next asked whether *Wdr5* was required for the development of C/EBP α p30-dependent leukemia. As cells from *Cebpa*^{p30/p30} animals show enhanced self-renewal capacity consistent with a preleukemic state, we tested whether *Wdr5* was required for this

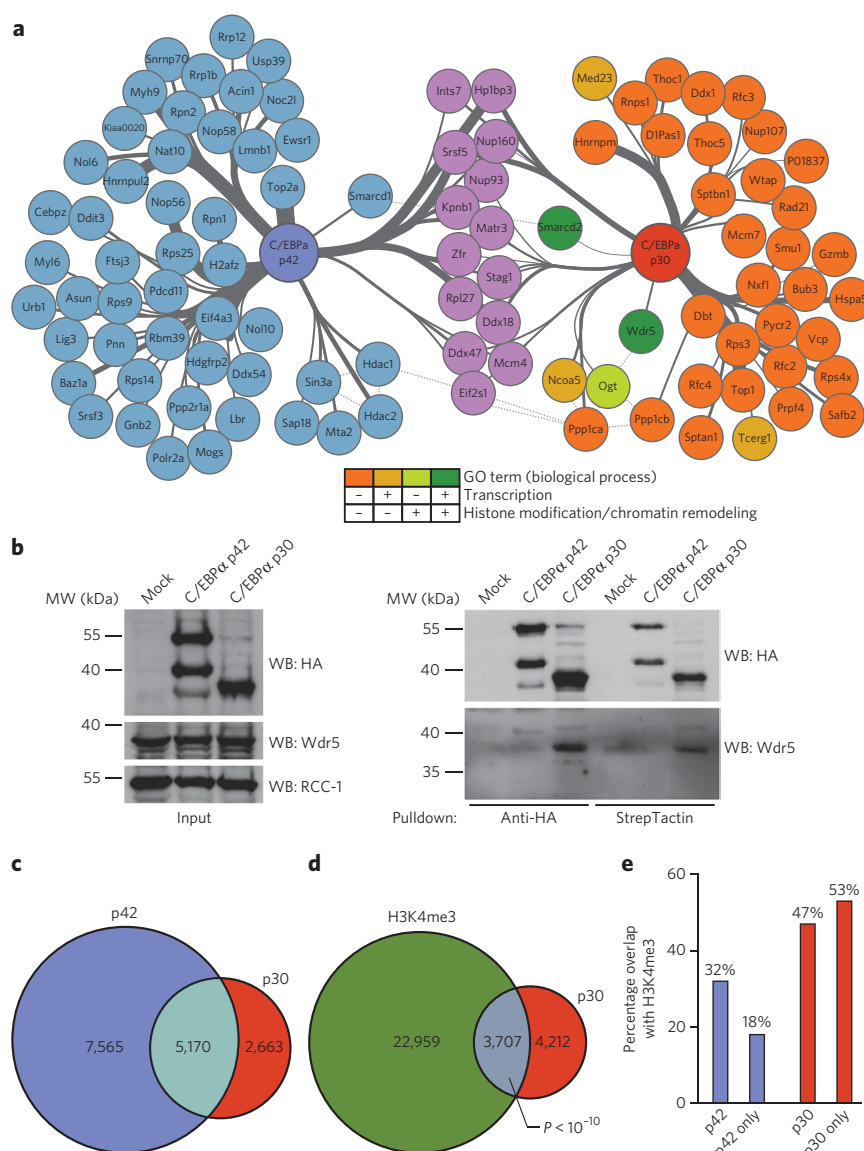


Figure 1 | C/EBP α p30 preferably interacts with *Wdr5* and the SET/MLL HMT complex, leading to increased colocalization of p30 and H3K4me3 on chromatin.

(a) Network representation of proteins reproducibly identified in AP-MS experiments of tagged C/EBP α variants from FDCP-1 after background normalization. Proteins identified in the purifications of both C/EBP α p42 and p30 variants are shown in purple. p42-specific interactors are shown in blue, and proteins specifically bound to p30 are color-coded according to the scheme below the network. The thickness of edges represents the strength of the interaction (based on spectral counts). (b) Left, nuclear extracts from 32D cells expressing tagged variants of p42 or p30 were analyzed by western blotting (WB) for expression of hemagglutinin (HA) and *Wdr5*. RCC-1 was used as loading control. Right, western blotting analysis of HA and *Wdr5* from anti-HA or StrepTactin purifications of tagged C/EBP α variants from 32D cell lines stably expressing p42 and p30. Representative images of at least two replicate experiments are shown. MW, molecular weight. (c,d) Venn diagrams showing the overlap of ChIP-seq peaks between C/EBP α p42 and p30 (c) and between C/EBP α p30 and H3K4me3 in p30-expressing cells (hypergeometric t-test, duplicate experiments) (d). (e) Bar diagram showing percentages of overlap between p42 and p42-only and between p30 and p30-only peaks and H3K4me3 ChIP-seq data in the respective cellular background.

phenotype. Although hematopoietic stem and progenitor cells from *Cebpa*^{p30/p30} mice expressing a control shRNA (shCtrl) showed sustained proliferation in a serial replating assay, the proliferative capacity of sh*Wdr5*-expressing cells gradually dropped until no immature cells could be isolated after the third plating round (**Fig. 3a** and **Supplementary Fig. 10**). Consistently, immunophenotypic

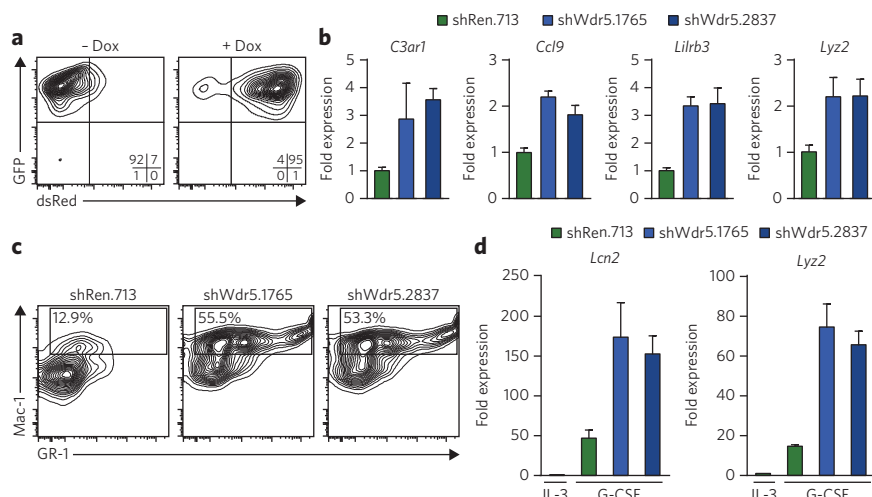


Figure 2 | Loss of Wdr5 restores granulocytic differentiation potential in C/EBPα p30-expressing cells. (a) Flow cytometric analysis of dsRed-reporter induction in 32D reverse Tet activator (rtTA3) p30-expressing cells upon Dox administration. (b) qRT-PCR analysis of the indicated genes in 32D rtTA3 p30-expressing cells expressing indicated shRNA constructs after 48 h of Dox treatment. (c) Flow cytometry analysis for Mac-1 and GR-1 surface markers of 32D rtTA3 p30-expressing cells expressing the indicated shRNA constructs 96 h after Dox treatment and 48 h after exposure to G-CSF (10 ng ml⁻¹). Presented events are gated on the GFP⁺ dsRed⁺ population. (d) qRT-PCR analysis of *Lcn2* and *Lyz2* expression in 32D rtTA3 p30-expressing cells transduced with the indicated shRNA constructs after 96 h of Dox treatment followed by 48 h exposure to G-CSF. Data are presented as mean ± s.d. of triplicate experiments.

analysis of these cells showed a high proportion of mature myeloid cells with strong Mac-1 staining in *Cebpa*^{p30/p30} cultures with *Wdr5* knockdown, whereas Mac-1 expression was mostly absent on shCtrl-expressing *Cebpa*^{p30/p30} cells (Fig. 3b). We next determined the

influence of *Wdr5* on the maintenance of p30-dependent leukemia *in vivo*. We knocked down *Wdr5* in freshly isolated *Cebpa*^{p30/p30} leukemia cells using three different lentiviral constructs expressing GFP-coupled *Wdr5*-targeting shRNAs and transplanted the transduced cells into irradiated recipient mice (Fig. 3c). Using the same titer for all viral stocks, we observed robust GFP expression in the donor-derived fraction of bone marrow in mice transplanted with control shRNA-expressing lentivirus. In contrast, two shRNAs that efficiently downregulate *Wdr5* expression led to a strong reduction in the percentage of GFP-positive bone marrow cells (Fig. 3d,e). The same trend was observed when the Mac-1^{lo} c-Kit⁺ p30 leukemia-initiating cell (LIC) fraction was analyzed (Fig. 3d,e). Notably, the abundance of GFP-positive p30 LICs was correlated to *Wdr5* expression levels as a dysfunctional *Wdr5*-targeting shRNA (*Wdr5*-1) had no effect in this system (Fig. 3e,f). Taken together, this indicates that the maintenance of transformed *Cebpa*^{p30/p30} leukemic cells *in vitro* and *in vivo* requires *Wdr5* expression.

OICR-9429 occupies the MLL-binding pocket of Wdr5

As our loss-of-function experiments suggest that *Wdr5* is indeed an obligate and specific effector of p30 action, we reasoned that pharmacological antagonism of *Wdr5* could selectively target p30-dependent cellular functions. We previously reported a series of compounds that bound human WDR5 and antagonized its interaction with MLL complex

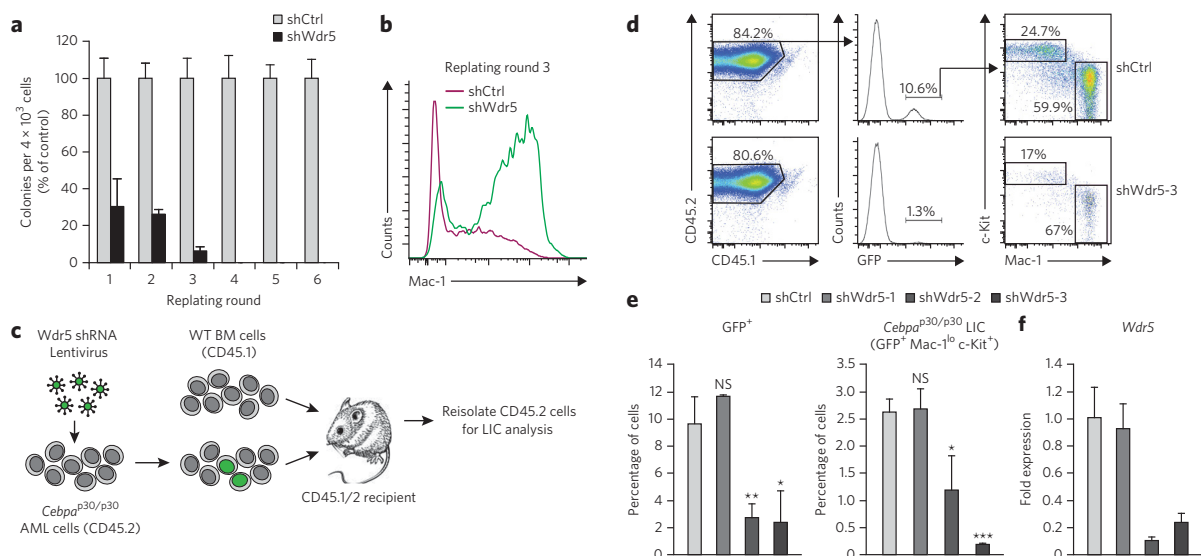


Figure 3 | Wdr5 is required to maintain C/EBPα p30-dependent self-renewal *in vitro* and *in vivo*. (a) Serial replating assay of primary *Cebpa*^{p30/p30} fetal liver cells after knockdown of *Wdr5*. To prevent outgrowth of cells escaping retroviral shRNA expression after several rounds of replating, we used a lentiviral shRNA construct allowing for continuous selection of shRNA-expressing cells through puromycin addition. Colony numbers were normalized to cells expressing a control shRNA construct (shCtrl) in every round of replating. (b) Flow cytometry analysis of Mac-1 surface marker expression of cells from the experiment shown in a after round 3 of replating. (c) Schematic outline of the experimental setup of the transplantation experiment. WT, wild type; BM, bone marrow. (d) Flow cytometric analysis of *Cebpa*^{p30/p30} LIC from the bone marrow of transplanted mice. Presented events are gated on the CD45.2⁺ lineage singlet population. *Cebpa*^{p30/p30} LICs display a c-Kit⁺ Mac-1^{lo} immunophenotype. Top, representative example of a mouse transplanted with *Cebpa*^{p30/p30} AML cells transduced with a control shRNA-expressing lentivirus. Bottom, representative example of a mouse transplanted with *Cebpa*^{p30/p30} AML cells transduced with sh*Wdr5*-3-expressing lentivirus. (e) Statistical representation of percentages of GFP⁺ cells (left) and LICs (GFP⁺ Mac-1^{lo} c-Kit⁺ cells, right) within the CD45.2⁺ population from the experiment described in b. (f) qRT-PCR analysis of *Wdr5* expression in wild-type fetal liver cells transduced with the indicated shRNA constructs. Data throughout are presented as mean ± s.d. of triplicate experiments. NS, nonsignificant; **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001 (Student's *t*-test).

components in biochemical assays^{21,22}. Optimization of this series using structure-guided medicinal chemistry resulted in OICR-9429 (**1**; Fig. 4a), which binds WDR5 with high affinity ($K_d = 93 \pm 28$ nM (mean \pm s.d.); Supplementary Fig. 11) and competitively disrupts its interaction with a high-affinity Wdr5-interacting (WIN) peptide of MLL ($K_{\text{dis}} = 64 \pm 4$ nM; Fig. 4b). OICR-9429 is highly selective for WDR5 as it showed no apparent binding to or inhibition of 22 human methyltransferases; 9 different WD40 and histone reader domains; and a panel of over 250 human kinase, G protein-coupled receptor, ion channel and transporter drug targets (Supplementary Tables 3 and 4).

To obtain insight into the detailed mode of interaction, we determined the crystal structure of WDR5 in complex with OICR-9429 at 1.5-Å resolution (Supplementary Table 5). As expected from our molecular design, OICR-9429 bound in the MLL WIN peptide-binding pocket of WDR5, thereby preventing its interaction with MLL via the same protein surface (Fig. 4c,d). Thus, OICR-9429 occupies the arginine-binding pocket exploited by the WDR5-binding peptides of SET1 methyltransferases²³ and KANSL1 (ref. 24). Conformational rearrangements critical for OICR-9429 binding were observed (i) at the bottom of the cavity (Phe263) to accommodate the *N*-methylpiperazine moiety in lieu of the WIN peptide guanidinium group and (ii) at the rim (Phe133 and Phe149) to accommodate the 'northern' substituent of the OICR-9429 compound (Supplementary Fig. 12). The piperazine, amide linker and pyridone groups of OICR-9429 form direct or water-mediated hydrogen bonds with surrounding residues (Cys261, Ser91 and Asp107; Fig. 4e), whereas the northern fragment interacts with a hydrophobic patch in WDR5 (Phe133, Tyr191, Phe149 and Pro173; Supplementary Fig. 12). Ser91 in WDR5 is critical for compound binding, as a WDR5^{S91K} mutant did not bind OICR-9429 (Fig. 4f).

OICR-9429 is a potent and selective Wdr5 antagonist

First we tested whether OICR-9429 was able to disrupt the C/EBP α p30-Wdr5 interaction. Wdr5 was readily detected in C/EBP α immunoprecipitates from lysates of *Cebpa*^{p30/p30} cells in the presence of OICR-9429, suggesting that the interaction of Wdr5 with MLL did not influence p30 binding (Supplementary Fig. 13). We next tested the effect of OICR-9429 on Wdr5-dependent protein-protein interactions in cells using a biotinylated variant of the compound in a chemical proteomics experiment. Although we were able to efficiently isolate Wdr5 using the biotinylated variant of OICR-9429, this enrichment was lost upon competition with excess unmodified OICR-9429 (Fig. 5a). Bioinformatic analysis of LC/MS/MS data revealed that Wdr5 was the primary target protein of OICR-9429 in cells (Fig. 5b). Notably, our analysis did not identify any other components of SET/MLL HMT complexes, indicating that OICR-9429 disrupts integral protein-protein interactions between Wdr5 and its binding partners. Indeed, OICR-9429 reduced the amount of endogenous MLL and RBBP5 that coimmunoprecipitated with exogenously expressed Flag-tagged WDR5 in a dose-dependent manner (Fig. 5c).

Gene expression profiling of OICR-9429-treated *Cebpa*^{p30/p30} cells showed that Wdr5 antagonism led to the upregulation of myeloid-specific transcripts (Supplementary Fig. 14a). Notably, gene set enrichment analysis²⁵ showed a close correlation between OICR-9429-induced genes and genes that were upregulated after Wdr5 knockdown. Furthermore, the gene signature of *Cebpa*^{p30/p30} LICs⁶ was downregulated upon Wdr5 antagonism by OICR-9429 (Supplementary Fig. 14b).

Taken together, these data demonstrate that OICR-9429 is a potent, selective and cell-active antagonist of the Wdr5-MLL interaction that is able to elicit a profound disruption of the protein-protein interaction network around Wdr5 and the SET/MLL HMT complex.

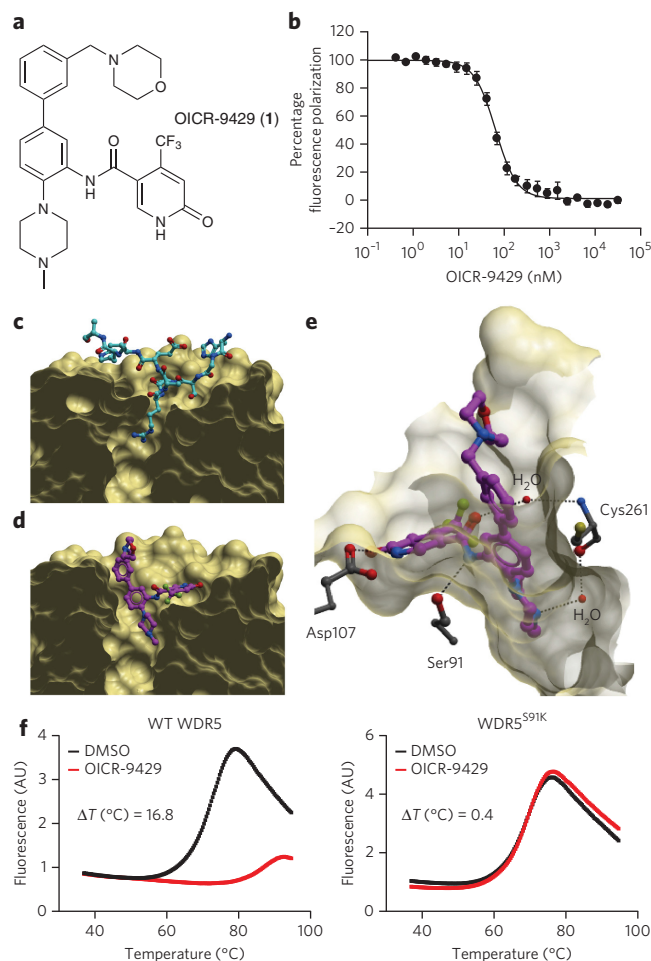


Figure 4 | OICR-9429 binds WDR5 in the MLL WIN motif-binding pocket.

(a) Chemical structure of OICR-9429 (*N*-(4-(4-methylpiperazin-1-yl)-3'-(morpholinomethyl)-[1,1'-biphenyl]-3-yl)-6-oxo-4-(trifluoromethyl)-1,6-dihydropyridine-3-carboxamide, **1**). (b) Peptide displacement assay monitoring the decrease in fluorescence polarization (FP) signal of a fluorescently labeled MLL peptide upon OICR-9429-induced dissociation from WDR5. (c) Crystal structure of WDR5 in complex with a WIN peptide (cyan; PDB 4ESG). (d) Structure of WDR5 bound to OICR-9429 (purple, PDB code 4QL1). (e) Key direct and water-mediated hydrogen bonds between OICR-9429 and WDR5 are shown as dashed lines. Extensive hydrophobic contacts and other interactions are shown in Supplementary Figure 12. (f) Differential scanning fluorimetry of wild-type (WT) WDR5 (left) and the WDR5^{S91K} mutant (right) incubated with DMSO or OICR-9429 (50 μ M). AU, arbitrary units.

p30-expressing cells are sensitive to Wdr5 antagonism

Next we tested whether OICR-9429 would have a selective inhibitory effect on cells expressing C/EBP α p30. Primary fetal liver cells from *Cebpa*^{p30/p30} animals showed a much greater sensitivity toward OICR-9429 treatment than did wild-type fetal liver cells in colony-formation assays (Fig. 6a). Similarly, the sensitivity to OICR-9429 was markedly greater in a *Cebpa*^{p30/p30} cell line than in mouse cells transformed to overexpress HoxA9 and Meis1 and different human leukemia cell lines (Supplementary Fig. 15a,b). Also, OICR-9429 did not affect the number and distribution of mouse bone marrow colony-formation units (Supplementary Fig. 15c). The specificity of OICR-9429 is further illustrated by the fact that the closely related derivative OICR-0547 (**2**), an inactive control compound that no longer binds WDR5, did not show any effect on the viability of cells sensitive or resistant to OICR-9429

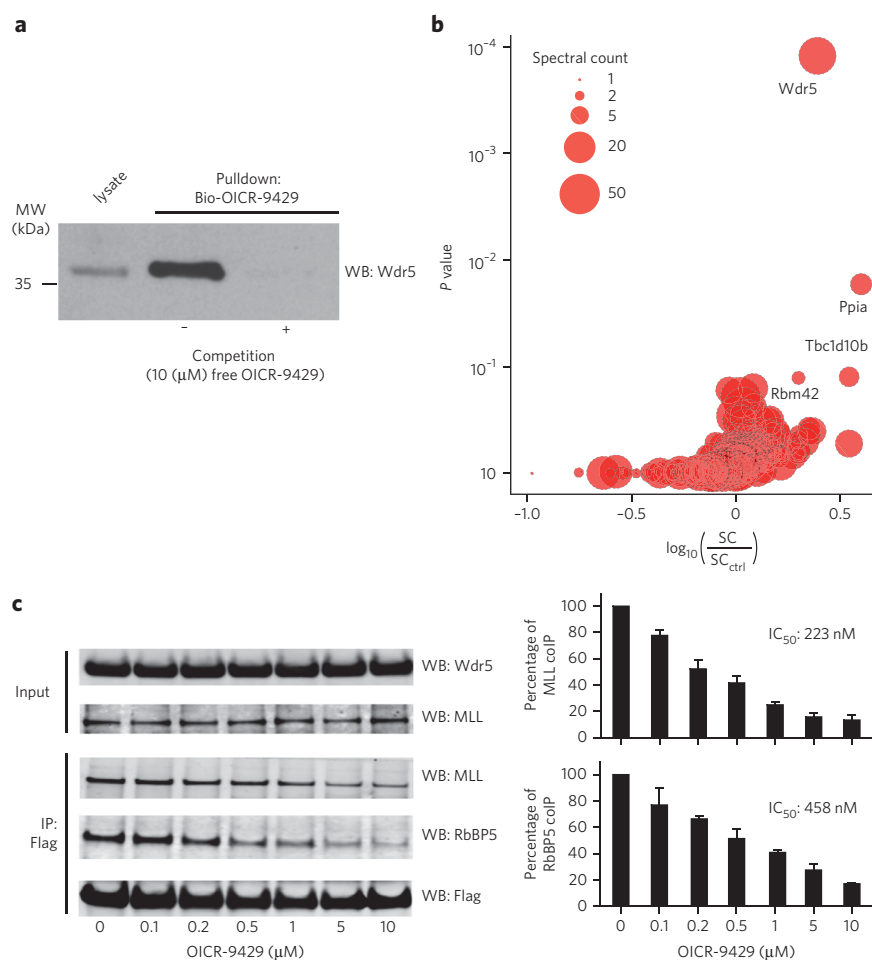


Figure 5 | OICR-9429 is a small-molecule antagonist of the Wdr5-MLL interaction.

(a) Western blot (WB) analysis of drug-affinity purification experiments using a biotinylated variant of OICR-9429 in *Cebpa*^{p30/p30} cells. MW, molecular weight. (b) Cellular target profile of OICR-9429. Log ratio of total spectral counts (SCs) in normal drug pull-down experiments versus competition experiments (x axis) and corresponding binding specificity *P* value, as calculated by Decontaminator³⁶ (y axis). Ctrl, control. (c) Left, HEK293 cells were transfected with Flag-tagged WDR5 and treated with vehicle alone (DMSO) or the indicated doses of OICR-9429. Flag immunoprecipitates were blotted for MLL and RBBP5. Right, histogram representation of quantified data from immunoprecipitation (IP) experiments. Data are presented as mean \pm s.d. of triplicate experiments. IC₅₀, half-maximum inhibitory concentration.

at relevant concentrations (Supplementary Fig. 16). Prolonged exposure of *Cebpa*^{p30/p30} cells to OICR-9429 caused upregulation of surface markers associated with myeloid differentiation and loss of progenitor morphology (Fig. 6b,c). In fact, OICR-9429 induced the upregulation of myeloid marker genes in *Cebpa*^{p30/p30} cells, such as *C3ar1*, *Ccl9*, *Lcn2*, *Lilrb3* and *Lyz2* (Fig. 6d), indicating that Wdr5 antagonism is able to overcome the differentiation block of *Cebpa*^{p30/p30} cells.

Finally, we tested whether primary human AML cells with N-terminal C/EBP α mutations would be sensitive to WDR5 antagonism. At 5 μ M, OICR-9429 caused a significant ($P < 0.001$) decrease in viability in the majority of AML patient-derived cells with mutations in the N-terminal part of the *CEBPA* gene (mean viability 53%, $n = 8$; Fig. 6e). Notably, the same concentration of OICR-9429 had little effect on viability in AML patient cells with other mutations (mean viability 86%, $n = 5$).

In summary, our data indicate that p30-recruited WDR5-containing protein complexes represent a functional vulnerability that can be pharmacologically exploited to trigger differentiation and growth arrest in C/EBP α -mutant AML.

DISCUSSION

Here we demonstrate that major differences in the protein-protein interactions as well as protein-DNA interactions of p42 versus p30 isoforms of C/EBP α underlie their distinct biological functions. Comparative affinity purification (AP)-MS experiments showed that the leukemogenic p30 variant of C/EBP α , but not the wild-type p42 isoform, preferably interacted with Wdr5, a critical component of mammalian SET/MLL HMT complexes. These complexes are composed of the SET/MLL HMT and three catalytically inactive core subunits (Wdr5, Ash2l and Rbbp5)²⁶. Among those, Wdr5 mediates MLL-dependent H3K4 methylation by physically interacting with both histone H3 and the MLL protein^{27–30}. Thus, the maintenance of MLL-dependent global H3K4me3 levels and gene expression requires Wdr5 (refs. 20,26). Wdr5 expression was required for embryonic stem cell pluripotency³¹, suggesting that Wdr5 is a key factor in linking chromatin remodeling and differentiation potential.

As deposition of the H3K4me3 mark is mediated by the SET/MLL HMT complex, we used the distribution of H3K4me3 ChIP-seq as a proxy for the presence and activity of Wdr5 on chromatin. Colocalization of H3K4me3 with p30 was enriched over that with p42 and was predominantly found in proximity to genes that showed p30-dependent expression changes. Formally, our data do not exclude the possibility that the p30-Wdr5 interaction occurs outside of the SET/MLL complex. However, Wdr5 downregulation led to a selective reduction of H3K4me3 levels on promoters of p30 target genes, indicating a functional consequence of the p30-Wdr5 interaction in the SET/MLL-dependent regulation of gene expression. These data not only cross-validate our AP-MS results of the specific association of p30 with Wdr5 but also extend the functional aspect of the p30-Wdr5 interaction to the chromatin level.

Wdr5 was not the only p30-specific interactor in our AP-MS survey. Indeed, C/EBP α p42 and p30 were associated with an unexpectedly different subset of the molecular machinery of the cell. The observation that a smaller isoform of a protein can have specific binding partners that are not shared with the longer variant seems counterintuitive. It is possible that different biophysical properties of C/EBP α variants may account for changes in subnuclear localization and thus also manifest in differential protein interaction patterns. This could be reflected by the differences in chromatin association between p42 and p30, which are likely to be influenced by specific binding proteins, as was previously shown to be important for transcription factor function³².

As Wdr5 could also interact with C/EBP α p42 upon overexpression in HEK293 cells, we speculate that the preferred interaction of p30 and Wdr5 in myeloid cells is likely to be indirect and mediated through other proteins whose expression is restricted to the hematopoietic system and which are differentially regulated by p30.

Wdr5 and MLL1 were critical to sustain p30-dependent functions, including inhibition of myeloid differentiation, enhanced

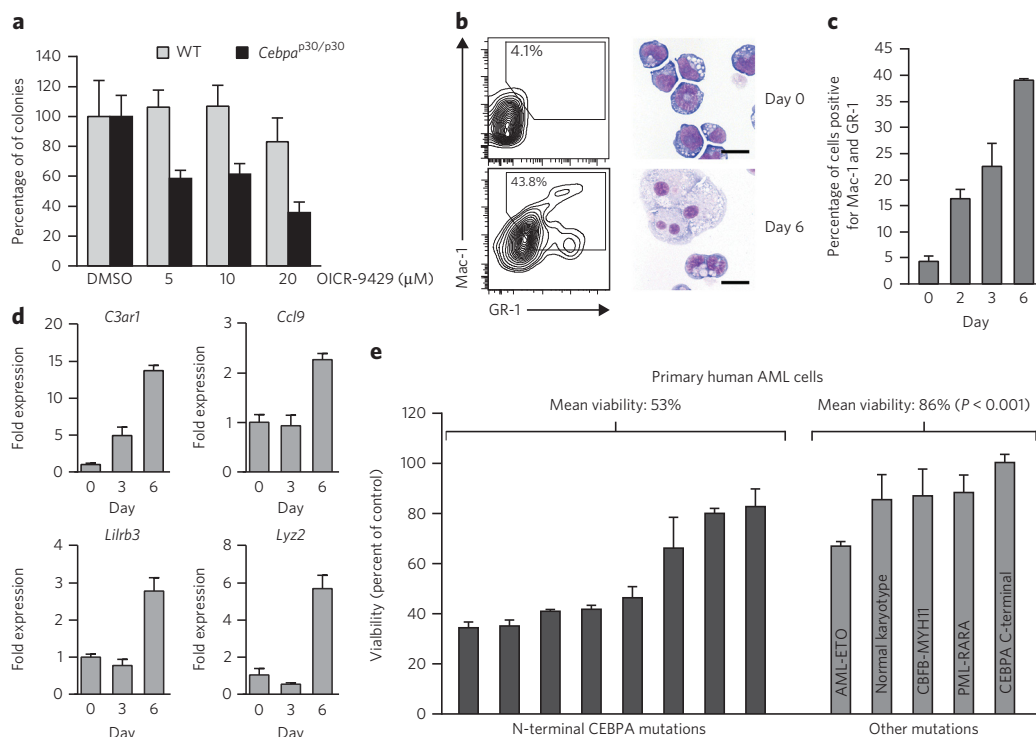


Figure 6 | Pharmacological antagonism of the WDR5-MLL interaction by OICR-9429 selectively affects p30-expressing cells. (a) Colony-formation assay of primary mouse *Cebpa*^{p30/p30} and wild-type (WT) fetal liver cells in response to OICR-9429 at indicated concentrations. Numbers were normalized to colonies generated in response to DMSO for each genotype. **(b)** *Cebpa*^{p30/p30} cells were left untreated (top) or treated with 20 μM OICR-9429 (bottom) for 6 d. Left, flow cytometry analysis of Mac-1 and GR-1 surface expression. Right, representative micrographs of stained cytopins. Scale bar, 10 μm. **(c)** Quantification of Mac-1 GR-1 double-positive cells at the indicated time points after exposure to OICR-9429 (20 μM). **(d)** qRT-PCR analysis of the indicated genes in *Cebpa*^{p30/p30} cells treated with 20 μM OICR-9429 for the indicated periods. **(e)** Cell viability of primary human AML cells with the indicated mutational status after 3 d of exposure to 5 μM of OICR-9429. In samples with N-terminal CEBPA mutations, these are either isolated or occur in conjunction with a C-terminal mutation on the other allele. Numbers were normalized values obtained in response to DMSO for each patient sample. Data throughout are presented as mean ± s.d. of triplicate experiments. *P* values were calculated using the Student's *t*-test.

self-renewal and leukemogenesis. Notably, this effect was restricted to the p30 isoform of C/EBPα, as depletion of Wdr5 had little effect on wild-type cells or on cells overexpressing p42. In line with this, shRNA-mediated downregulation of *Wdr5* had no effects in fibroblasts or myoblasts³¹. We were able to phenocopy the effect of *Wdr5* downregulation by using a new pharmacological antagonist of the Wdr5-MLL interaction. OICR-9429 has all of the essential properties of a high-quality chemical probe. Its on-target, dose-dependent activity in cells directly reflects its *in vitro* biochemical activity. Furthermore, OICR-9429 directly engages the endogenous cellular target at pharmacologically relevant concentrations and displays a disease-related activity that is related to the target hypothesis. OICR-9429 displays exquisite cellular selectivity and specificity in disrupting critical protein-protein interactions between WDR5 and other MLL complex members, thereby compromising MLL activity.

C/EBPα p42 was shown to sequester E2F proteins through direct protein-protein interactions, leading to repression of E2F transcriptional activity, cell cycle arrest and differentiation⁵. Reduction of the E2F-C/EBPα interaction in the p30 variant has been postulated to contribute to leukemogenic effects of p30. Our identification of Wdr5 as a specific interactor of p30 allows speculation over another mechanism: Wdr5 has been reported to interact with E2F6 (ref. 33), and another reported binding partner of Wdr5, Hcf-1, was shown to mediate recruitment of the SET/MLL HMT complex to E2F target genes by physical interaction with E2F family members³⁴. Thus, p30 could actively induce the expression of E2F target genes via the specific recruitment of the SET/MLL HMT complex.

It has been reported that LICs from *Cebpa*^{p30/p30} mice show gene expression patterns that are related to AML with MLL rearrangements⁶. Consistent with these findings, a peptide-based antagonist of the Wdr5-MLL interaction was recently shown to selectively target MLL leukemia cells³⁵.

In summary, our comparative integrated analyses of C/EBPα variants provide evidence for the existence of gain-of-function features of the C/EBPα p30 protein. We reveal a new and specific functional connection between the p30 variant of the C/EBPα transcription factor and Wdr5, which controls the transforming properties of p30. We describe OICR-9429 as a new compound that disrupts the interaction of Wdr5 with MLL in cells, thereby selectively triggering a differentiation program in p30-expressing leukemia cells. Thus, our data provide an attractive alternative by targeting the MLL complex in C/EBPα-mutated AML. As direct pharmacological targeting of transcription factors remains difficult, targeting of critical cofactors such as the MLL complex could represent an appealing strategy for targeting cancer entities that depend on oncogenic transcription factor mutations.

Received 20 October 2014; accepted 28 May 2015; published online 13 July 2015

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Protein Data Bank. Crystallographic data are deposited under accession code 4QL1.

References

- Koschmieder, S., Halmos, B., Levantini, E. & Tenen, D.G. Dysregulation of the C/EBP α differentiation pathway in human cancer. *J. Clin. Oncol.* **27**, 619–628 (2009).
- Zhang, P. *et al.* Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP α . *Immunity* **21**, 853–863 (2004).
- Leroy, H. *et al.* CEBPA point mutations in hematological malignancies. *Leukemia* **19**, 329–334 (2005).
- Fasan, A. *et al.* The role of different genetic subtypes of CEBPA mutated AML. *Leukemia* **28**, 794–803 (2014).
- Nerlov, C. C/EBP α mutations in acute myeloid leukaemias. *Nat. Rev. Cancer* **4**, 394–400 (2004).
- Kirstetter, P. *et al.* Modeling of C/EBP α mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. *Cancer Cell* **13**, 299–310 (2008).
- Friedman, A.D. & McKnight, S. Identification of two polypeptide segments of CCAAT/enhancer-binding protein required for transcriptional activation of the serum albumin gene. *Genes Dev.* **4**, 1416–1426 (1990).
- Nerlov, C. & Ziff, E.B. Three levels of functional interaction determine the activity of CCAAT/enhancer binding protein- α on the serum albumin promoter. *Genes Dev.* **8**, 350–362 (1994).
- Pedersen, T.A., Kowenz-Leutz, E., Leutz, A. & Nerlov, C. Cooperation between C/EBP α TBP/TFIIB and SWI/SNF recruiting domains is required for adipocyte differentiation. *Genes Dev.* **15**, 3208–3216 (2001).
- Slomiany, B.A., Arigo, K.L.D., Kelly, M.M. & Kurtz, D.T. C/EBP α inhibits cell growth via direct repression of E2F-DP-mediated transcription. *Mol. Cell. Biol.* **20**, 5986–5997 (2000).
- Porse, B.T. *et al.* E2F repression by C/EBP α is required for adipogenesis and granulopoiesis *in vivo*. *Cell* **107**, 247–258 (2001).
- D'Alo, F. *et al.* The amino terminal and E2F interaction domains are critical for C/EBP α -mediated induction of granulopoietic development of hematopoietic cells. *Blood* **102**, 3163–3171 (2003).
- Wang, Q.-F., Cleaves, R., Kummalue, T., Nerlov, C. & Friedman, A.D. Cell cycle inhibition mediated by the outer surface of the C/EBP α basic region is required but not sufficient for granulopoiesis. *Oncogene* **22**, 2548–2557 (2003).
- Cleaves, R., Wang, Q. & Friedman, A.D. C/EBP α p30, a myeloid leukemia oncoprotein, limits G-CSF receptor expression but not terminal granulopoiesis via site-selective inhibition of C/EBP DNA binding. *Oncogene* **23**, 716–725 (2004).
- Zada, A.A. *et al.* Proteomic discovery of Max as a novel interacting partner of C/EBP α : a Myc/Max/Mad link. *Leukemia* **20**, 2137–2146 (2006).
- Trivedi, A.K. *et al.* Proteomic identification of C/EBP-DBD multiprotein complex: JNK1 activates stem cell regulator C/EBP α by inhibiting its ubiquitination. *Oncogene* **26**, 1789–1801 (2007).
- Bararia, D. *et al.* Proteomic identification of the MYST domain histone acetyltransferase TIP60 (HTATIP) as a co-activator of the myeloid transcription factor C/EBP α . *Leukemia* **22**, 800–807 (2008).
- Koleva, R.I. *et al.* C/EBP α and DEK coordinately regulate myeloid differentiation. *Blood* **119**, 4878–4888 (2012).
- Fujimoto, T., Anderson, K., Jacobsen, S.E.W., Nishikawa, S.-I. & Nerlov, C. Cdk6 blocks myeloid differentiation by interfering with Runx1 DNA binding and Runx1-C/EBP α interaction. *EMBO J.* **26**, 2361–2370 (2007).
- Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3K4 methylation and vertebrate development. *Cell* **121**, 859–872 (2005).
- Bolshan, Y. *et al.* Synthesis, optimization, and evaluation of novel small molecules as antagonists of WDR5-MLL interaction. *ACS Med. Chem. Lett.* **4**, 353–357 (2013).
- Senisterra, G. *et al.* Small-molecule inhibition of MLL activity by disruption of its interaction with WDR5. *Biochem. J.* **449**, 151–159 (2013).
- Migliori, V. *et al.* Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.* **19**, 136–144 (2012).
- Dias, J. *et al.* Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. *Genes Dev.* **28**, 929–942 (2014).
- Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
- Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).
- Couture, J.-F., Collazo, E. & Trievel, R.C. Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat. Struct. Mol. Biol.* **13**, 698–703 (2006).
- Song, J.-J. & Kingston, R.E. WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J. Biol. Chem.* **283**, 35258–35264 (2008).
- Patel, A., Dharmarajan, V. & Cosgrove, M.S. Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J. Biol. Chem.* **283**, 32158–32161 (2008).
- Ruthenburg, A.J. *et al.* Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. *Nat. Struct. Mol. Biol.* **13**, 704–712 (2006).
- Ang, Y.-S. *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183–197 (2011).
- Yu, M. *et al.* Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol. Cell* **36**, 682–695 (2009).
- Dou, Y. *et al.* Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873–885 (2005).
- Tyagi, S., Chabes, A.L., Wysocka, J. & Herr, W. E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Mol. Cell* **27**, 107–119 (2007).
- Cao, F. *et al.* Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol. Cell* **53**, 247–261 (2014).
- Lavallée-Adam, M., Cloutier, P., Coulombe, B. & Blanchette, M. Modeling contaminants in AP-MS/MS experiments. *J. Proteome Res.* **10**, 886–895 (2011).

Acknowledgments

We thank M. Gridling, M. Planavsky, D. Printz and A. Spittler for experimental help and K. Kandasamy and M. Schuster for bioinformatic help. Next-generation sequencing was performed at the Campus Science Support Facilities Next-Generation Sequencing Unit (<http://www.csf.ac.at/>). F.G. and R.G. were funded by the Austrian Science Fund (FWF grant P22282-B11). A. Skucha is supported by FP7-PEOPLE-2011-ITN Project HemID (289611). The Superti-Furga laboratory is supported by the Austrian Academy of Sciences and by European Research Council (ERC) grant ERC-2009-AdG-250179-i-FIVE. The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funds from AbbVie, Bayer, Boehringer Ingelheim, Genome Canada through the Ontario Genomics Institute (OGI-055), GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda and the Wellcome Trust (092809/Z/10/Z). The Ontario Institute for Cancer Research is funded by the Government of Ontario. Funding was also provided by the Leukemia and Lymphoma Society of Canada.

Author contributions

F.G., M.V., R.G., A.G., R.A., A. Skucha, S.V., E.K., D.B.-L., F.L., G.S., K.V.M.H. and R.M. planned, performed and analyzed biochemical, biophysical, cellular and *in vivo* experiments. M.G., D.S., G.P., M.S., P.J.B. and R.A. contributed to chemical design and synthesis of OICR-9429 and OICR-0547. H.W., A.D. and M.S. solved and analyzed the X-ray crystal structure of WDR5 in complex with OICR-9429. A. Stukalov, A. Schönegger, M.B. and C.B. performed bioinformatic analyses. J.Z., K.L.B., R.D. and C.N. provided access to vital tools and technologies, planned experiments and analyzed results. F.G., M.V., P.J.B., R.A., C.H.A. and G.S.-F. designed the study, planned experiments, analyzed results and wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information, chemical compound information and chemical probe table is available in the [online version of the paper](http://www.nature.com/reprints/index.html). Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Correspondence and requests for materials should be addressed to F.G., C.H.A. or G.S.-F. OICR-9429 is available commercially.

ONLINE METHODS

Constructs. The coding regions of C/EBP α p42 and p30 (ref. 8) were cloned into a modified pMSCV-IRES-GFP retroviral vector with a C-terminal Strep-HA tag. A two-component miR-30-based retroviral expression system was used for inducible shRNA studies³⁷. A DNA fragment encoding human WDR5 (residues 24–334) was amplified by PCR and subcloned into the pET28-MHL vector, downstream of the polyhistidine coding region. shRNA sequences are listed in **Supplementary Table 6**.

Cell culture. Primary mouse fetal liver cells were cultivated in DMEM plus 10% FCS plus SCF, IL-6 and IL-3. c-Kit-positive *Cebpa*^{p30/p30} mutant bone marrow cells were enriched using magnetic beads (Miltenyi). Cells were cultured in IMDM medium supplemented with IL-6, IL-3, SCF, polybrene and β -mercaptoethanol. FDCP-1 and 32D cell lines and immortalized *Cebpa*^{p30/p30} fetal liver cells were maintained in RPMI 1640 medium supplemented with 10% FCS and IL-3. Plat-E and HEK293T cells were maintained in DMEM medium with 10% FCS. For colony assays, 1×10^4 cells were plated in complete methylcellulose medium (MethoCult M3434) and colonies were scored 7 d later. For replating assays, colonies were scored in 7-d intervals and 1×10^4 cells were replated. The *Cebpa*^{p30/p30} cell line was established by picking single cell clones after the sixth round of replating and continuous liquid culture in the presence of SCF, IL-3 and IL-6 for 4 weeks. The HoxA9/Meis1-transformed cell line was a kind gift from R. Slany (University Erlangen, Germany). Ecotropic retroviral particles were generated using the Plat-E packaging cell line. Lentiviruses were produced by transient transfection of HEK293T cells. Human leukemic blast cells were derived from diagnosed AML patients and isolated on Ficoll-Hypaque gradients. Cells were cultured in RPMI 1640 supplemented with 20% FCS plus human IL-3, IL-6, G-CSF, GM-CSF and SCF. Patient recruitment and sample processing were performed according to protocols from the Dutch-Belgian Hematology/Oncology Cooperative Group (HOVON trials). All studies were approved through the institutional human ethics review board of the Erasmus Medical Center Rotterdam, and all patients provided written informed consent in accordance with the Declaration of Helsinki.

Bone marrow transplantation. Cells were transplanted in lethally irradiated CD45.1/2 8- to 10-week-old recipients along with 1×10^6 whole BM competitor cells (CD45.1) via tail vein injection. Six mice were transplanted for every candidate shRNA in two independent experiments. Mice were killed 28 d after transplantation and subjected to LIC analysis. All mice were bred and maintained at University of Oxford in accordance with Institutional and UK Home Office guidelines.

Affinity purification of C/EBP α protein complexes. All steps described in the protocol were carried out at 4 °C. Purifications were performed from 1×10^9 freshly harvested cells. After washing with PBS, cells were incubated in buffer N (300 mM sucrose, 10 mM HEPES, pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 0.1 mM DTT, 0.75 mM spermidine, 0.15 mM spermine, 0.1% Nonidet P-40, 50 mM NaF, 1 mM Na₃VO₄, protease inhibitors) for 5 min on ice to lyse the cytoplasm. Nuclei were collected by centrifugation (500g for 5 min), and the supernatant was removed. The nuclear pellet was washed with buffer N. Nuclei were resuspended in buffer C420 (20 mM HEPES, pH 7.9, 420 mM NaCl, 25% glycerol, 1 mM EDTA, 1 mM EGTA, 0.1 mM DTT, 50 mM NaF, 1 mM Na₃VO₄, protease inhibitors) and shaken vigorously for 30 min. Nuclear extracts were cleared by centrifugation for 1 h at 100,000g. Prior to purification, 15 mg of extracts were adjusted to 2 mg/ml and 150 mM NaCl with 20 mM HEPES, 50 mM NaF and 1 mM Na₃VO₄, protease inhibitors. After 20 min of preclearing with RNase A, benzonase and avidin, nuclear extracts were incubated with 200 μ l StrepTactin Sepharose beads for 2 h on a rotating wheel at 4 °C. Beads were washed 3 \times with TNN-HS buffer and 2 \times using TNN-HS buffer without detergent and inhibitors. Bound proteins were eluted by incubation with 100 μ l 2.5 mM biotin in TNN-HS buffer 10 min at 4 °C, followed by centrifugation for 3 min at 300g. Samples were alkylated with iodoacetamide and separated by 1D SDS-PAGE on a 4–12% bis-Tris gel (NuPAGE, Invitrogen). Proteins were visualized by silver staining, the entire gel lane was excised and 20 slices were digested *in situ* with modified porcine trypsin (Promega) as previously described³⁸. Peptides were pooled into 10 samples and analyzed by online LC/MS/MS (1D-gel-MS).

Chemical proteomics. Per experiment, 100 μ l UltraLink streptavidin bead slurry (Pierce) was centrifuged, and the supernatant was removed. After washing

with lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 0.2% NP-40, 5% glycerol, 1.5 mM MgCl₂, 25 mM NaF, 1 mM Na₃VO₄, 1 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol (DTT), 10 μ g/ml TLCK, 1 μ g/ml leupeptin, 1 μ g/ml aprotinin and 10 μ g/ml soybean trypsin inhibitor (Sigma), pH 7.5), biotin-conjugated OICR-9429 derivative (0.05 μ mol) was added and incubated on a rotoshaker for 30 min at 4 °C. After centrifugation and one additional washing step, beads were resuspended in cell lysates (10 mg total protein per pull-down) and incubated on a rotoshaker for 2 h at 4 °C. For competition experiments, lysates were preincubated with 10 μ M of unmodified OICR-9429 for 20 min at 4 °C. After centrifugation, beads were transferred to spin columns (MoBioTec) and washed with lysis buffer and HEPES, followed by elution of bound proteins with elution buffer (50% 6 M urea, 50% 100 mM formic acid).

MS. All affinity purifications were analyzed on a hybrid linear trap quadrupole (LTQ) Orbitrap Velos mass spectrometer (ThermoFisher Scientific) coupled to a 1200 series high-performance liquid chromatography system (Agilent Technologies) via a nanoelectrospray ion source using liquid junction (Proxeon). Solvents for HPLC separation of peptides were as follows: solvent A consisted of 0.4% formic acid in water, and solvent B consisted of 0.4% formic acid in 70% methanol and 20% isopropanol. 8 μ l of the tryptic peptide mixture were automatically loaded onto a trap column (Zorbax 300SB-C18 5 μ m, 5 \times 0.3 mm, Agilent Biotechnologies) with a binary pump at a flow rate of 45 μ l/min. 0.1% trifluoroacetic acid was used for loading and washing the pre-column. After washing, the peptides were eluted by back-flushing onto a 16-cm fused silica analytical column with an inner diameter of 50 μ m packed with C18 reversed-phase material (ReproSil-Pur 120 C18-AQ, 3 μ m, Dr. Maisch GmbH) with a 27-min gradient ranging from 3% to 30% solvent B, followed by a 25-min gradient from 30% to 70% solvent B and, finally, a 7-min gradient from 70% to 100% solvent B at a constant flow rate of 100 nl/min³⁹. Analyses were performed in a data-dependent acquisition mode, and dynamic exclusion for selected ions was 60 s. A top 15 collision-induced dissociation (CID) method was used, and a single lock mass at m/z 445.120024 (Si(CH₃)₂O)₆ was employed⁴⁰. Maximal ion accumulation time allowed in CID mode was 50 ms for MSⁿ in the LTQ and 500 ms in the C-trap. Automatic gain control was used to prevent overfilling of the ion traps and was set to 5,000 in MSⁿ mode for the LTQ and 10⁶ ions for a full FTMS scan. Intact peptides were detected in the Orbitrap Velos at 60,000 resolution at m/z 400.

Protein identification. Peak list information was extracted from the RAW MS files and converted into an MGF format with the msconvert tool (ProteoWizard Library v2.1.2708). MGF files were searched against the mouse component of the UniProtKB/SwissProt database (<http://www.uniprot.org>), including all protein isoforms plus rat C/EBP α and known contaminant sequences. An initial search was performed with Mascot (<http://www.matrixscience.com/>, version 2.3.02). Mass error tolerances on the precursor and fragment ions were ± 10 p.p.m. and ± 0.6 Da, respectively. Only fully tryptic peptides were considered with a maximum of one missed cleavage site, and carbamidomethyl cysteine and methionine oxidation were set as fixed and variable modifications, respectively. The Mascot peptide ion score threshold was equal to 30, and at least three peptide identifications per protein were required.

For both the precursor and fragment ion data, linear recalibration transformations that minimize the mean square deviation of the measured from theoretical values were deduced from initial identifications. Recalibrated files were searched against the same protein database with Mascot and Phenix (GeneBio) using narrower mass tolerances (± 4 p.p.m. and ± 0.3 Da)⁴¹. All other search parameters were identical to the initial first pass search. Mascot and Phenix output files were processed by internally developed parsers to filter and integrate protein identifications. The following peptide score thresholds were used: $T_1 = 14$, $T_2 = 40$ and $T_3 = 10$; and $T_1 = 4.2$, $T_2 = 4.75$ and $T_3 = 3.5$, respectively ($P < 10^{-3}$). Proteins with at least two unique peptides above score T_1 or with a single peptide above T_2 were selected as unambiguous identifications. Additional peptides from these validated proteins with a score $>T_3$ were appended to the final result. The validated identifications from both algorithms were merged, spectral conflicts were discarded and protein groups were defined according to shared peptides. A false discovery rate (FDR) of <0.01 for protein identifications and <0.001 for peptides (including peptides exported with lower scores) was estimated by applying the same filtering procedure against a database of reversed protein sequences. All affinity purifications were analyzed on the basis of protein spectral counts. For each purification strategy, proteins identified in the control cells were subtracted from proteins



identified from the corresponding C/EBP α affinity purifications. Moreover, contaminants such as keratin, spectrin and plectin were removed from the list as known nonspecific binders frequently observed in AP-MS.

To estimate confident drug targets in chemical proteomics experiments, the drug pull-downs were compared with the competition experiments using the Decontaminator algorithm modified to take advantage of using multiple protein search engines (*P* values were calculated separately for spectral counts, and both Mascot and Phenix protein scores were calculated as described in ref. 42 and then combined into single *P* value using the Fisher method). The *P*-value cutoff for the confident hits was set to 0.001.

Expression and purification of human WDR5. WDR5 was overexpressed in *E. coli* BL21 by addition of 1 mM IPTG overnight at 15 °C. Cells were resuspended in 50 mM HEPES buffer, pH 7.4, containing 250 mM NaCl, 5 mM imidazole, 2 mM β -mercaptoethanol and 5% glycerol and lysed using a microfluidizer (Microfluidics Corporation, 20,000 psi). The clarified lysate was loaded onto a Ni²⁺-charged HiTrap Chelating column (GE Healthcare). After washing with 10 column volumes of 20 mM HEPES, pH 7.4, 250 mM NaCl, 50 mM imidazole and 5% glycerol, the protein was eluted with elution buffer (20 mM HEPES, pH 7.4, 250 mM NaCl, 250 mM imidazole, 5% glycerol) and loaded on a Superdex200 column (GE Healthcare) equilibrated with 20 mM PIPES buffer, pH 6.5, and 250 mM NaCl. TEV protease was added to combined WDR5-containing fractions to remove the His tag. The protein was further purified to homogeneity by ion-exchange chromatography.

Crystallization of the WDR5–OICR-9429 complex. Purified WDR5 protein (10 mg/ml) was mixed with OICR-9429 at 1:5 molar ratio of protein/compound and crystallized using the sitting-drop vapor diffusion method by mixing 1 μ l of protein solution with 1 μ l of the reservoir solution containing 25% PEG 3350, 0.2 M ammonium acetate and 0.1 M Bis-Tris, pH 6.5. Crystals were soaked in the corresponding mother liquor supplemented with 20% ethylene glycol as cryoprotectant before freezing in liquid nitrogen.

Crystallographic data collection and structure determination. X-ray diffraction data for WDR5 structure in complex with OICR-9429 was collected at 100K at beam-line 08ID-1 of CLS, Canadian Light Source. Data sets were processed using the HKL-3000 suite⁴³. The structures of WDR5 in complex with OICR-9429 inhibitor was solved by molecular replacement using PHASER⁴⁴ with PDB entry 4IA9 as search template. PRODRG was used to generate geometry restraints for the compound OICR-9429 refinement⁴⁵. REFMAC was used for structure refinement⁴⁶. The graphics program COOT was used for model building and visualization⁴⁷. MOLPROBITY was used for structure validation⁴⁸.

Fluorescence polarization assays. The 9-Ala-FAM ((Ac)-ARA-EVHLRK-(Ahh-Ahh)-K(5,6-FAM)) peptide for WDR5 was synthesized, C-terminally labeled with FAM and purified by Peptide 2.0 (Chantilly). Peptide displacement assays were performed in 125- μ l reactions, at a constant labeled peptide concentration of 5 nM and WDR5 concentration of 50 nM in 80 mM sodium phosphate, pH 6.5, 20 mM KCl and 0.008% Triton X-100 using 96-well Microfluor 2 plates (Thermo Scientific). Fluorescence polarization was measured using a ViewLux imager (PerkinElmer) at an excitation wavelength of 480 nm and an emission wavelength of 540 nm.

OICR-9429 selectivity assays. OICR-9429 was evaluated at concentrations up to 50 μ M for potential inhibition of the following human protein methyltransferases using a radioactivity-based catalytic assay, as previously reported^{22,49,50}: SMYD2, G9a, EHMT1, SUV39H2, SETDB1, SETD7, SETD8, SUV420H1, SUV420H2, PRMT1, PRMT3, PRMT5-MEP50 complex, PRMT6, PRMT8, PRMD9, EZH1, EZH2, NSD1, NSD2, NSD3, SETD2, DOT1L and DNMT1. The following chromatin histone binding 'reader domains' were tested for binding to OICR-9429 (200 μ M) using differential scanning fluorimetry (DSF) and differential static light scattering (DSLS) as previously described⁵¹: TDRD3, SND1, L3MBTL1, L3MBTL3, UHRF1, 53BP1, STRAP and EED. Additional selectivity screening of >200 enzymes, receptors and transporters was performed by Cerep.

Transient transfection and immunoprecipitation of WDR5. 293 cells were transfected with 1 μ g of Flag-tagged human WDR5 plasmid (a kind gift

from J. Wysocka, Stanford University) using GeneJuice (EMD), and 24 h later the medium was replaced with fresh medium containing OICR-9429, and the cells were incubated for 5 h. Cells were collected and lysed in 0.5 ml of cold 0.3 M KCl, 20 mM Tris-HCl, pH 8, 1 mM EDTA, 0.1% NP40, 10% glycerol and protease inhibitors (Roche) for 10 min on ice. The lysate was centrifuged for 5 min at 14,000 r.p.m., and the supernatant was collected. Supernatants were used for immunoprecipitation with 1 μ g Flag M2 (Sigma) and 30 μ l protein G Dynabeads (Invitrogen) overnight at 4 °C. Beads were washed 3 times with lysis buffer containing 0.15 M KCl before SDS-PAGE loading buffer was added.

Chromatin immunoprecipitation. 1×10^8 cells were cross-linked with 10% formaldehyde for 10 min, quenched with glycine for 5 min and then harvested. Cells were resuspended in LB1 buffer (50 mM Hepes, pH 7.6, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40 and 0.25% Triton X-100) to lyse the cytoplasm. Nuclei were washed once in LB2 buffer (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) before lysis in LB3 buffer (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 0.5% *N*-lauroylsarcosine, EDTA, 1 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM Na₃VO₄ and protease inhibitors). The released chromatin was sonicated to obtain fragments of 150 bp using a COVARIS sonicator. 0.5% Triton X-100 was added to the samples immediately after sonication to aid solubilization of the sheared DNA. Samples were spun at 10,000g for 10 min. Half of the supernatant was incubated with 5 μ g anti-HA (Abcam) or anti-H3K4me3 (Millipore) at 4 °C overnight. Antibody-bound material was pulled down using Dynal protein G magnetic beads (Invitrogen), washed 5 times and released using elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA and 1% SDS) at 65 °C. DNA-protein crosslinks were reverted by incubating the samples overnight at 65 °C. The DNA was treated with RNaseA and proteinase K and purified through phenol-chloroform extraction. 10 ng of precipitated material was used for sequencing on a Genome Analyzer IIx (Illumina).

Deep sequencing and data analysis. Short read sequences were aligned to the Genome Reference Consortium Mouse Build 38 (GCA_000001635.2, GRChm38, UCSC mm10) using Bowtie2, version 2.1.0 (ref. 52 in end-to-end, gapped alignment mode). Peaks were called with the Model-based Analysis of ChIP-Seq (MACS) algorithm, version 1.4.2 (ref. 53 with parameter settings suggested by the authors⁵⁴). The resulting peak lists were annotated with regard to transcription start sites using the Bioconductor package ChIPpeakAnno (<http://www.bioconductor.org/packages/release/bioc/html/ChIPpeakAnno.html>) and PeakAnnotator⁵⁵ and post-processed with custom R scripts using the ggplot2 plotting system (<http://ggplot2.org/>).

Microarray analysis. Total RNA (200 ng) was used for GeneChip analysis. Preparation of terminal-labeled cDNA, hybridization to genome-wide mouse Gene Level 1.0 ST GeneChips (Affymetrix) and scanning of the arrays were performed as described previously⁵⁶ and according to the manufacturer's protocols. Signal extraction, normalization using the RMA algorithm and probe filtering was performed using R/bioconductor, as described previously^{57,58}.

Western blotting. Western blotting was performed according to standard laboratory protocols. Antibodies used were: peroxidase-labeled anti-HA (HA-7, Sigma, 1:2,000), anti-HA (HA-11, Covance, 1:2,000), anti-Wdr5 (Abcam, 1:5,000), anti-C/EBP α (14AA, Santa Cruz, 1:1,000), anti-tubulin (Abcam, 1:5,000), anti-RCC-1 (E-6, Santa Cruz, 1:500), anti-MLL (N-terminal, Santa Cruz, 1:200), anti-Flag (Sigma, 1:5,000) and anti-RbBP5 (Abcam, 1:1,000). For **Figure 5c**, secondary IR600- and IR700-conjugated antibodies (LiCor, 1:5,000) were used to visualize the bands on an Odyssey scanner (LiCor). The band intensity was quantified using LiCor software, with data normalized to the WDR5-IP signal and plotted as a percentage of control with standard errors representing s.e.m. Experiments were performed 3–4 independent times with the compound incubated on cells for 5–6 h.

Real-time PCR analysis. Total RNA was isolated using RNeasy Mini Kit (Qiagen). 300 ng RNA was reverse transcribed using oligo(dT) primers using RevertAid Reverse Transcriptase (Fermentas). Quantitative PCR was carried out on a Qiagen RotorGene RG-600 PCR machine using the SensiMix SYBR kit (Bioline). Results were quantified using the 2^{- $\Delta\Delta C_t$} method⁵⁹. qPCR primer sequences are listed in **Supplementary Table 7**.

FACS. Cells were stained with fluorescence-labeled antibodies against c-Kit (BD Pharmingen 553142, 1:200), Mac-1 (BD Pharmingen 51-01712J, 1:200) and GR-1 (Biolegend 108412, 1:200) after incubation with anti-mouse CD16/CD32 (BD Pharmingen 553142, 1:200) to block Fc receptors. Data were collected on a FACS Fortessa Instrument (BD Biosciences). For the analysis of the LIC fraction in the bone marrow of transplanted mice, bone marrow cells were stained with a mix of antibodies (lineage cocktail) against CD4 (Biolegend 100514, 1:800), CD5 (Biolegend 100610, 1:600), CD8 (Biolegend 100710, 1:800) and Ter119 (Biolegend 116210, 1:300) and antibodies against c-Kit (eBioscience 17-1171-82, 1:200), Sca-1 (Biolegend 122520, 1:200), Mac1 (eBioscience 12-0112-83, 1:200), CD45.2 (eBioscience 47-0454-82, 1:50) and CD45.1 (eBioscience 25-0453-82, 1:100) following red blood cell lysis. 7-Amino-actinomycin D (7-AAD, Sigma) was added to exclude dead cells. Cells were analyzed on a LSR Fortessa SORP analyzer (BD). Data were analyzed with FlowJo software (Treestar).

Cytospin analysis. Cells were cytocentrifuged onto glass slides and stained with Rapid-Chrome Kwik-Diff Staining System (Thermo Scientific) before microscopic analysis.

Cell viability measurements. 20,000 viable, actively proliferating primary human AML cells per well were seeded in 96-well plates in triplicates and treated with 0.05% DMSO or OICR-9429. Cell viability was measured using the Cell Titer-Glo luminescent cell viability assay (Promega) on a VICTOR X4 luminometer (PerkinElmer) after 72 h.

Statistical analysis. Two-tailed Student's *t*-tests were used for statistical analysis if not stated otherwise.

37. Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat. Biotechnol.* **29**, 79–83 (2011).
38. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858 (1996).
39. Bennett, K.L. *et al.* Proteomic analysis of human cataract aqueous humour: Comparison of one-dimensional gel LCMS with two-dimensional LCMS of unlabelled and iTRAQ-labelled specimens. *J. Proteomics* **74**, 151–166 (2011).
40. Olsen, J.V. *et al.* Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021 (2005).
41. Colinge, J., Masselot, A., Giron, M., Dessigny, T. & Magnin, J. OLAV: towards high-throughput tandem mass spectrometry data identification. *Proteomics* **3**, 1454–1463 (2003).
42. Huber, K.V. *et al.* Stereospecific targeting of MTH1 by (S)-crizotinib as an anticancer strategy. *Nature* **508**, 222–227 (2014).
43. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in Os. *Meth. Enzymol.* **276**, 307–326 (1997).
44. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
45. Schüttelkopf, A.W. & van Aalten, D.M.F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).
46. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
48. Davis, I.W., Murray, L.W., Richardson, J.S. & Richardson, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619 (2004).
49. Siarheyeva, A. *et al.* An allosteric inhibitor of protein arginine methyltransferase 3. *Structure* **20**, 1425–1435 (2012).
50. Yu, W. *et al.* Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. *Nat. Commun.* **3**, 1288 (2012).
51. Niesen, F.H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
52. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
53. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
54. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X.S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
55. Salmon-Divon, M., Dvinge, H., Tammoja, K. & Bertone, P. PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics* **11**, 415 (2010).
56. Astapova, I. *et al.* The nuclear corepressor, NCoR, regulates thyroid hormone action in vivo. *Proc. Natl. Acad. Sci. USA* **105**, 19544–19549 (2008).
57. Irizarry, R.A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
58. Bilban, M. *et al.* Deregulated expression of fat and muscle genes in B-cell chronic lymphocytic leukemia with high lipoprotein lipase expression. *Leukemia* **20**, 1080–1088 (2006).
59. Schmittgen, T.D. & Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3**, 1101–1108 (2008).

Supplementary information to the paper

Pharmacological targeting of the Wdr5-MLL interaction in C/EBP α N-terminal leukemia

Florian Grebien^{1,2*§}, Masoud Vedadi^{3,4*}, Matthäus Getlik^{5*}, Roberto Giambruno¹, Amit Grover⁶, Roberto Avellino⁷, Anna Skucha¹, Sarah Vittori¹, Ekaterina Kuznetsova³, David Smil³, Dalia Barsyte-Lovejoy³, Fengling Li³, Gennadiy Poda^{5,8}, Matthieu Schapira^{3,4}, Hong Wu³, Aiping Dong³, Guillermo Senisterra³, Alexey Stukalov¹, Kilian V. M. Huber¹, Andreas Schönegger¹, Richard Marcellus⁵, Martin Bilban⁹, Christoph Bock¹, Peter J. Brown³, Johannes Zuber¹⁰, Keiryn L. Bennett¹, Rima Al-awar^{4,5}, Ruud Delwel⁷, Claus Nerlov⁶, Cheryl H. Arrowsmith^{3,11*§} and Giulio Superti-Furga^{1,12*§}

¹ CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna 1090, Austria

² Ludwig Boltzmann Institute for Cancer Research, Vienna 1090, Austria

³ Structural Genomics Consortium, University of Toronto, Toronto, ON, M5G 1L7, Canada

⁴ Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, M5S 1A8, Canada

⁵ Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, ON, M5G 0A3, Canada

⁶ MRC Molecular Hematology Unit, Weatherall Institute of Molecular Medicine, Oxford OX3 9DS, United Kingdom

⁷ Department of Hematology, Erasmus University Medical Center, Rotterdam 3015 GE, The Netherlands

⁸ Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, M5S 3M2, Canada

⁹ Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University Vienna, Vienna 1090, Austria

¹⁰ Research Institute of Molecular Pathology (IMP), Vienna 1030, Austria

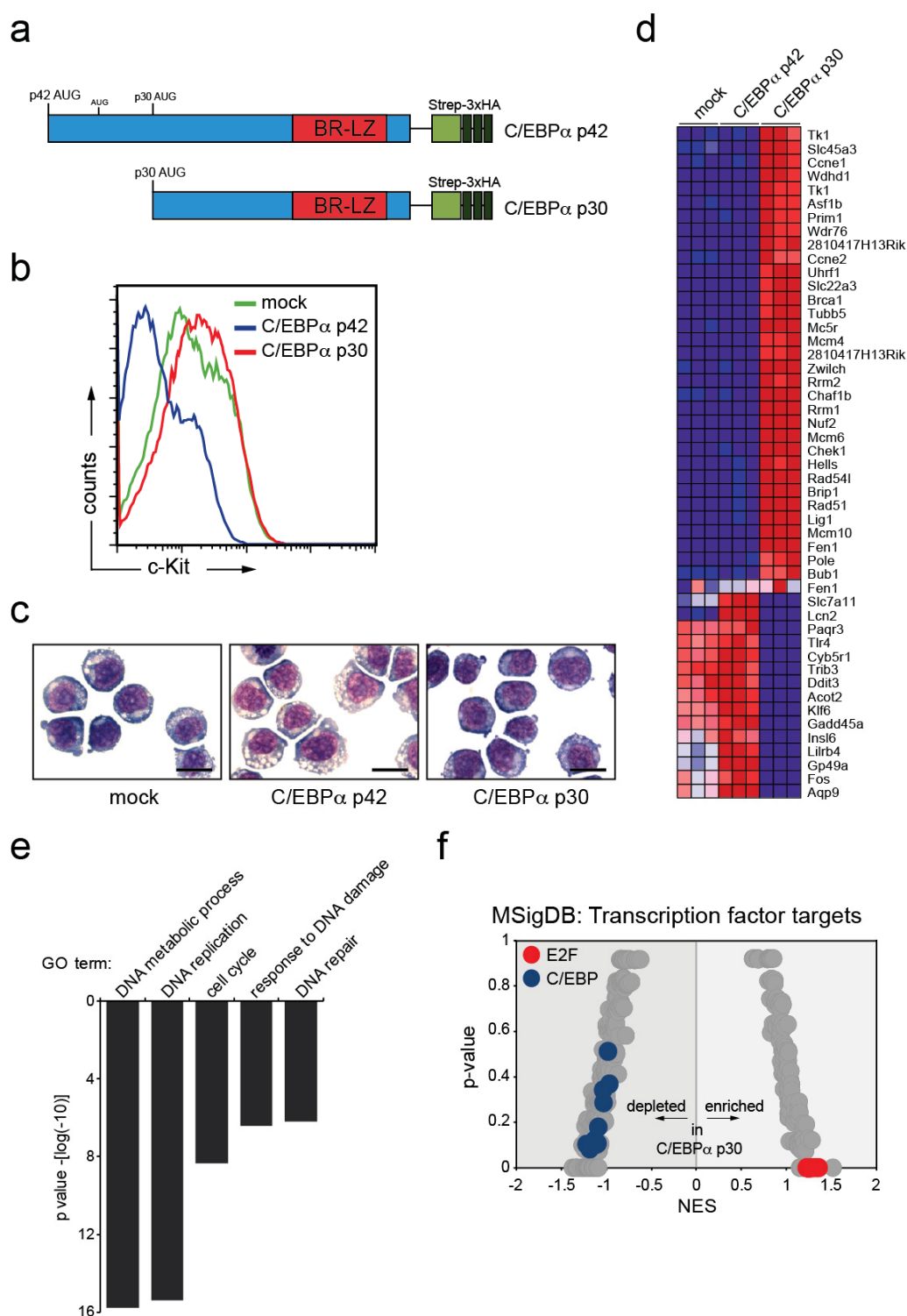
¹¹ Princess Margaret Cancer Centre and Department of Medical Biophysics, University of Toronto, Toronto, ON, M5G 2M9, Canada

¹² Center for Physiology and Pharmacology, Medical University of Vienna, 1090 Vienna

¹³ Center for Physiology and Pharmacology, Medical University Vienna, Vienna, Austria

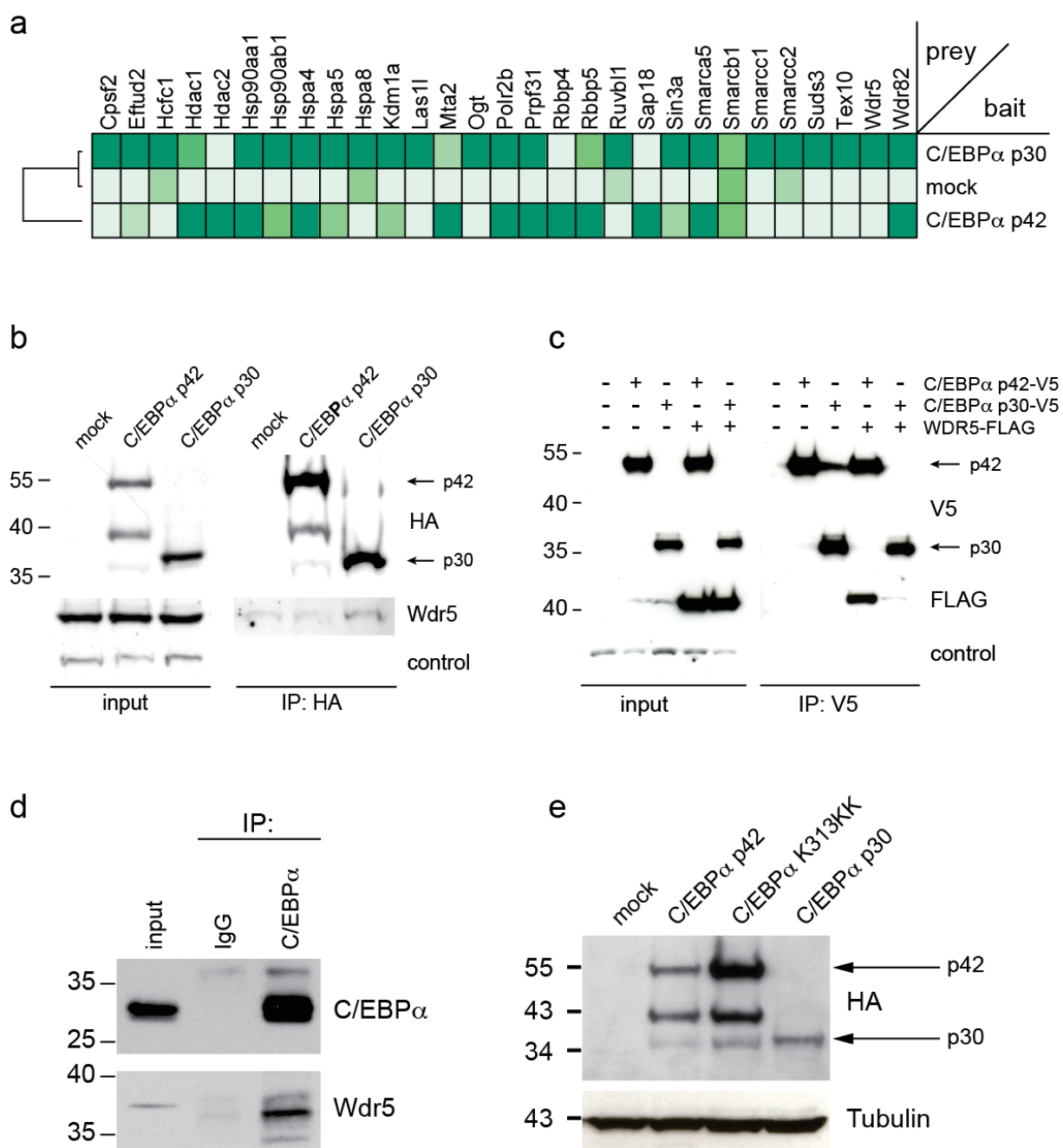
* equal contribution

§ correspondence to



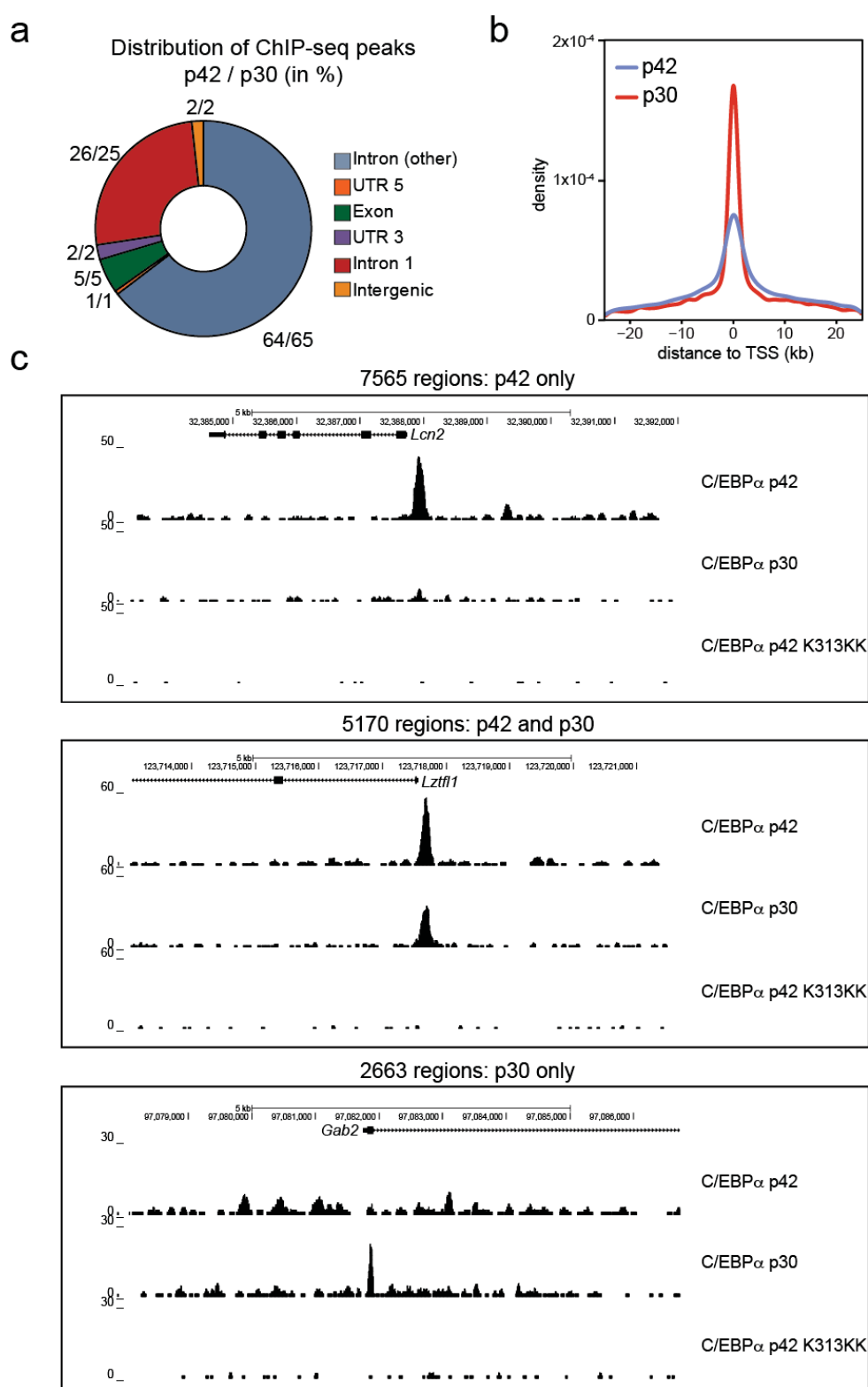
Supplementary Figure 1. An isogenic cell system to study cellular effects of C/EBP α variants. (a) Schematic representation of constructs used in this study. The Strep-3xHA tag was fused to the C-terminus of C/EBP α p42 or p30. (b) Flow cytometric analysis of c-Kit expression of FDCP-1 cells expressing C/EBP α p42 or p30. (c) Histological staining of cytopsin preparations of FDCP-1 cells expressing C/EBP α p42 or p30. Scale bar, 10 μ m (d) Heatmap representations of the top 47 genes differentially regulated upon p30 expression in FDCP-1 cells (FDR 0.01, fold

change >6) **(e)** Gene Ontology (GO)-analysis of genes shown in (e) **(f)** Gene Set enrichment (GSEA) for all transcription factor target gene sets available from the MSigDB. p-values are plotted versus normalized enrichment score (NES) for each gene set. Gene sets for E2F and C/EBP transcription factors are highlighted in red and blue, respectively.



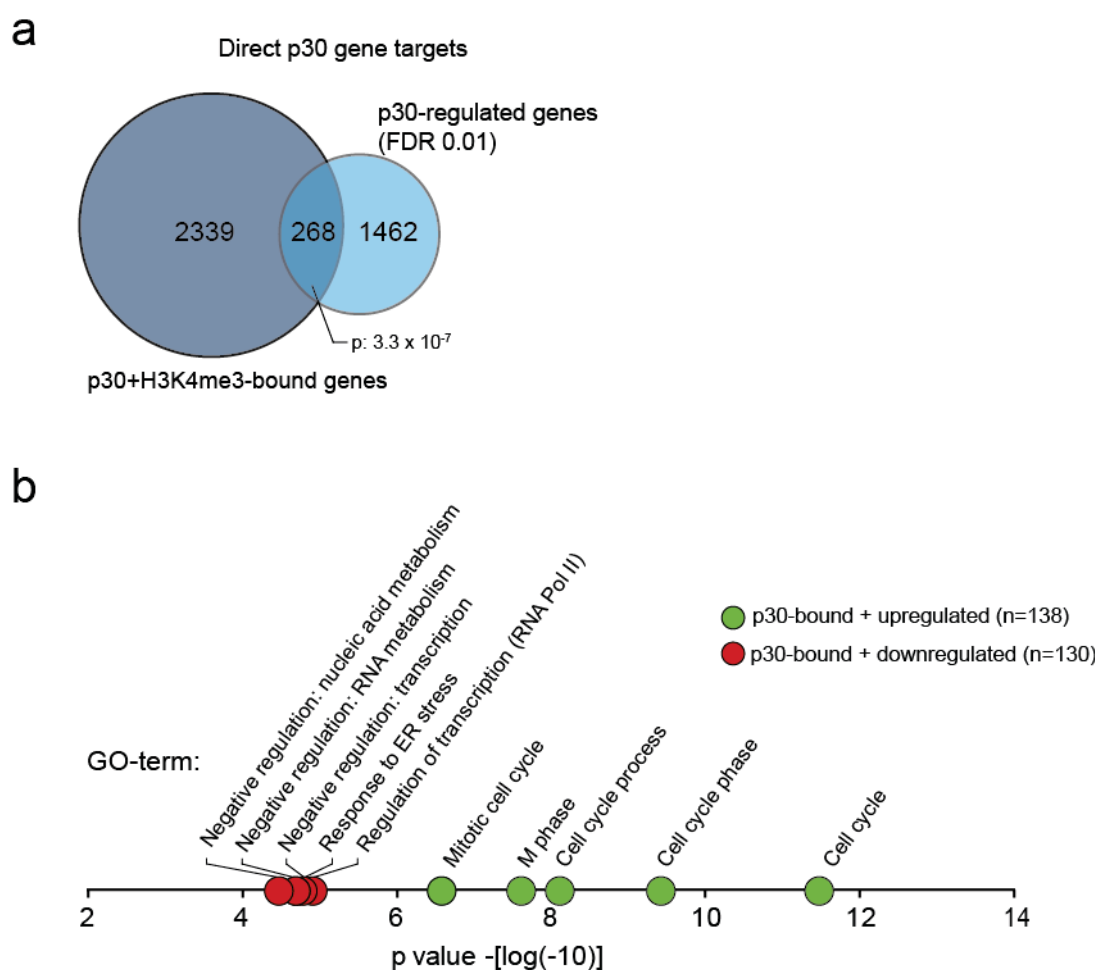
Supplementary Figure 2. Wdr5 preferably interacts with C/EBP α p30 in myeloid cells **(a)** Proteins present in the p42 and/or p30 AP-MS datasets were filtered for proteins that were previously found associated with Wdr5 and MLL (as published in the CORUM database, <http://mips.helmholtz-muenchen.de/genre/proj/corum/>). Data are presented based on average spectral counts of each protein in the respective AP-MS datasets. **(b)** Left panel: extracts from FDCP-1 cells expressing tagged variants of p42 or p30 were analyzed by Western blot for expression of HA and Wdr5. Right panel: Western blot analysis of HA and Wdr5 from anti-HA-purifications of tagged C/EBP α variants from FDCP-1 cell lines stably expressing p42 and p30. **(c)** Left panel: extracts from HEK293 cells transfected with indicated tagged variants of p42, p30 and Wdr5 were analyzed by Western blot for expression of V5 and FLAG. Right panel: Western blot analysis of FLAG and V5 from anti-V5-immunoprecipitates

of the same samples. **(d)** Western blot analysis of C/EBP α and Wdr5 from IgG- and anti-C/EBP α immunoprecipitates from lysates of *Cebpa*^{p30/p30} cells. **(e)** Extracts from FDCP-1 cells expressing tagged variants of p42, p42 K313KK or p30 were analyzed by Western blot for expression of HA and tubulin. Representative images of at least 2 replicate experiments are shown. Note that tagged variants of p42 and p30 migrate at higher molecular weight than 42 and 30 kDa, respectively. The additional HA-reactive band at 40 kDa in p42-expressing samples arises from translation initiation at an ATG codon that lies between the initiation codons of p42 and p30 (see Supplementary Fig. 1a).

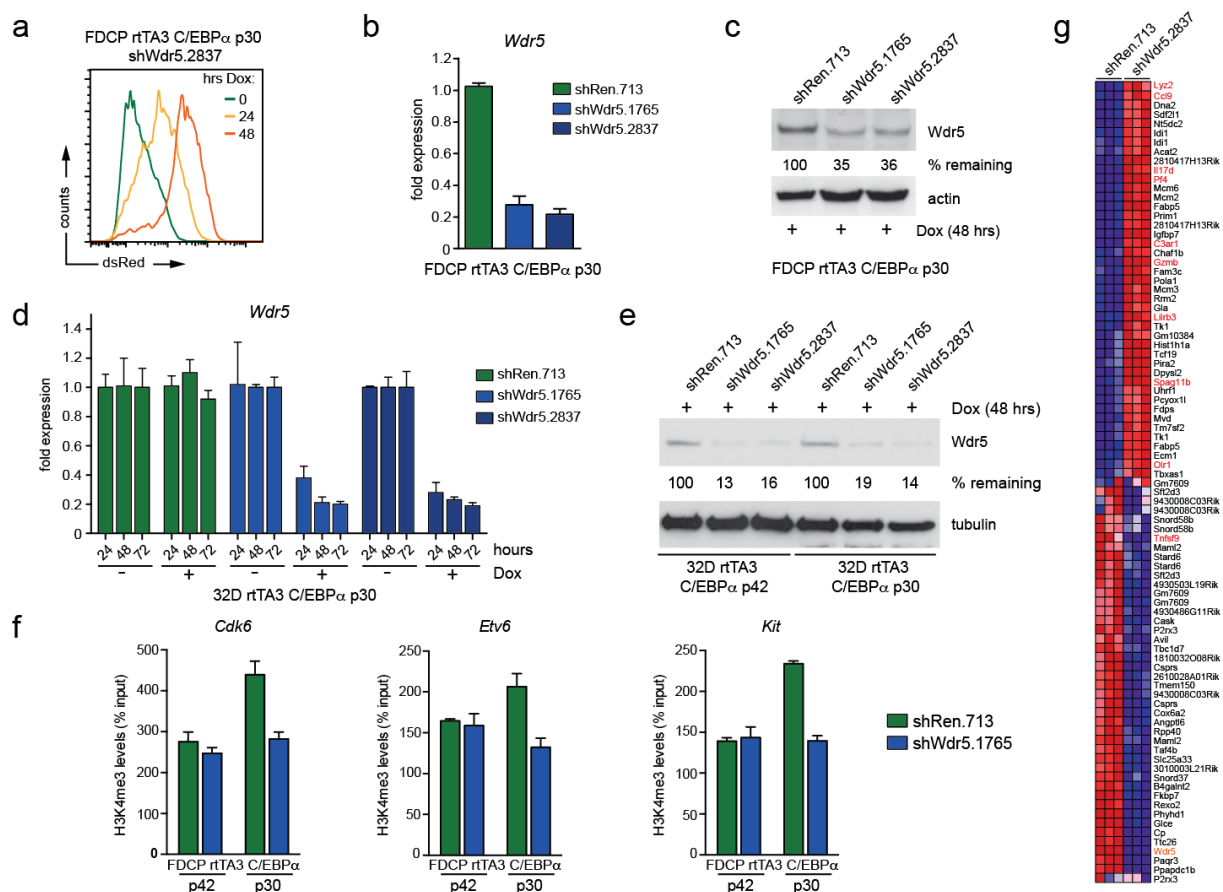


Supplementary Figure 3. C/EBP α variants display differential chromatin binding. (a) Distribution of p42/p30 binding sites at various genomic loci. Numbers represent percentages for p42 and p30, respectively. (b) Relative distances of p42- and p30-ChIP-seq peaks to annotated Transcription Start Sites (TSS) (c) Representative UCSC genome browser tracks for promoters bound by C/EBP α p42 and p30 (upper panel), promoters bound by C/EBP α p42, but not p30 (middle panel),

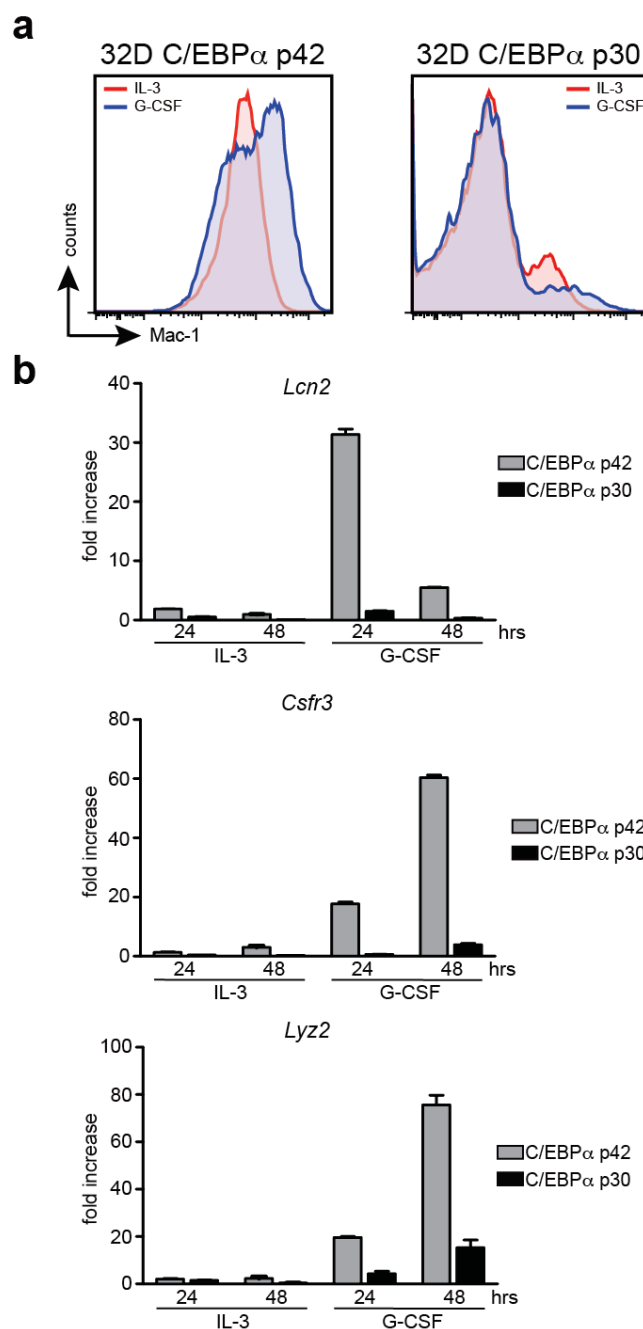
or promoters bound by C/EBP α p30, but not p42 (lower panel). The specificity of our experimental approach is validated by a control ChIP-seq experiment for a DNA-binding deficient C/EBP α mutant (K313KK), which did not show any significant chromatin association.



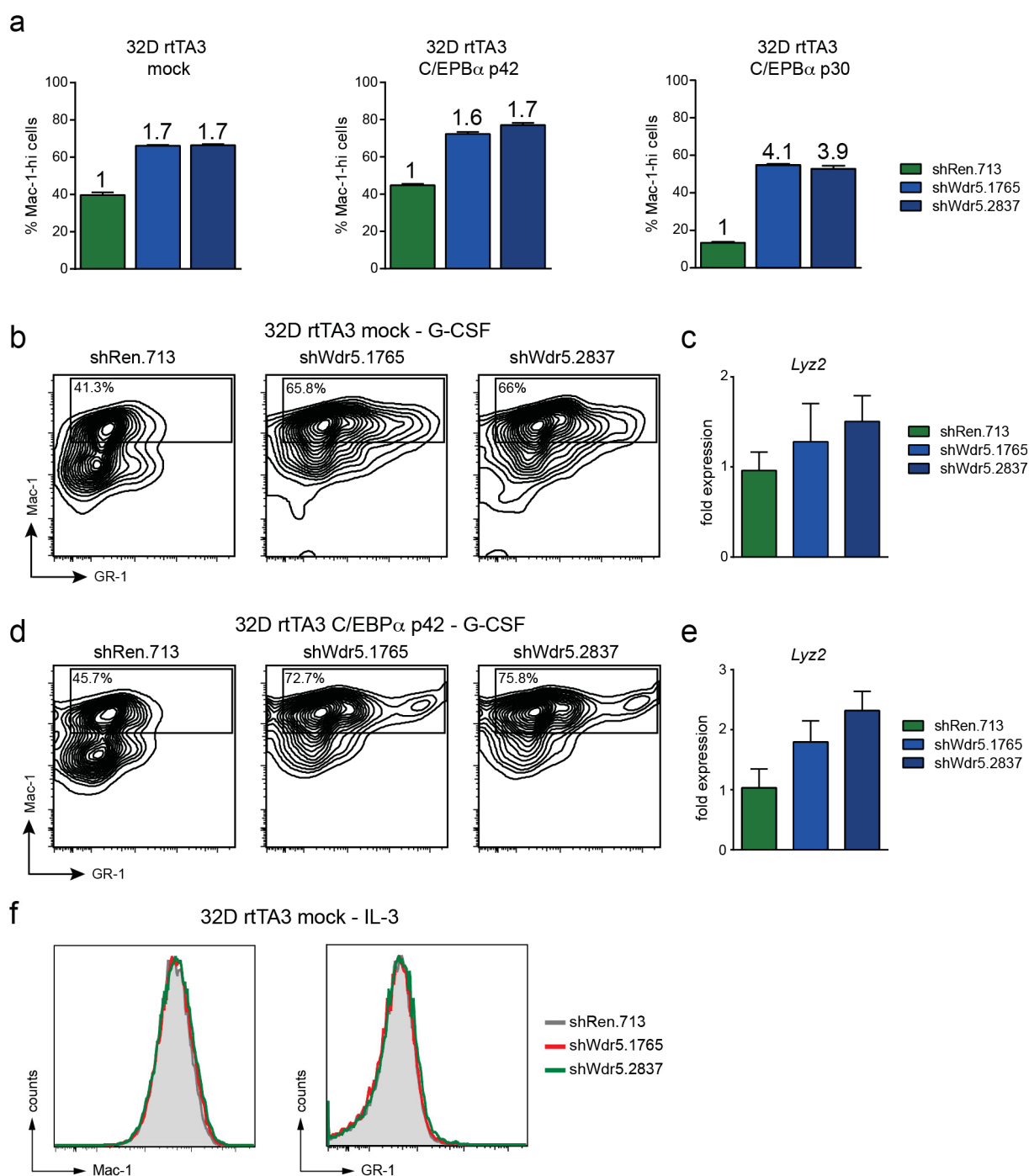
Supplementary Figure 4. C/EBP α p30 regulates the expression of distinct gene sets. **(a)** The Venn diagram represents an intersection of genes whose promoters were bound by p30 and were marked with H3K4me3 to indicate active transcription (left circle, dark blue, 2607 genes) and genes whose expression is de-regulated by p30 expression (right circle, light blue, 1730 genes, FDR 0.01). The 268 genes in the intersection are present in both datasets and are therefore designated “direct p30 gene targets”. **(b)** Functional annotation of direct p30 gene targets. The list of genes derived in (A) was further subdivided in up-regulated (green) and down-regulated genes (red). Functional annotation was done based on GO terms describing Biological Processes using DAVID. p-values for the top 5 biological processes are plotted for each gene set.



Supplementary Figure 5. Inducible shRNA-mediated knockdown of Wdr5 in p30-expressing cells reduces H3K4me3 levels on promoters of p30 target genes and induces upregulation of myeloid gene expression. (a) Flow cytometric analysis of dsRed-reporter induction in FDCP-1 rtTA3 p30 cells upon Dox administration at indicated timepoints. (b) qRT-PCR- and (c) Western Blot analysis of Wdr5 expression in FDCP-1 rtTA3 p30 cells expressing indicated shRNA constructs after 48 h of Dox treatment. Representative images of at least 2 replicate experiments are shown. (d) qRT-PCR- and (e) Western Blot analysis of Wdr5 expression in 32D rtTA3 p30 cells expressing indicated shRNA constructs after 48 hrs of Dox treatment. Representative images of at least 2 replicate experiments are shown. (f) qPCR analysis of the indicated cell types for specific enrichment on promoters of the *Cdk6* (left), *Etv6* (middle) and *Kit* (right) genes after H3K4me3 ChIP. (g) Heatmap representation of differentially regulated genes in FDCP-1 rtTA3 p30 cells expressing indicated shRNA constructs after 48 h of Dox treatment (FDR 0.05). Myeloid specific genes and are highlighted in red. Wdr5 is highlighted in orange. Data are presented as mean \pm standard deviation (SD) of triplicate experiments

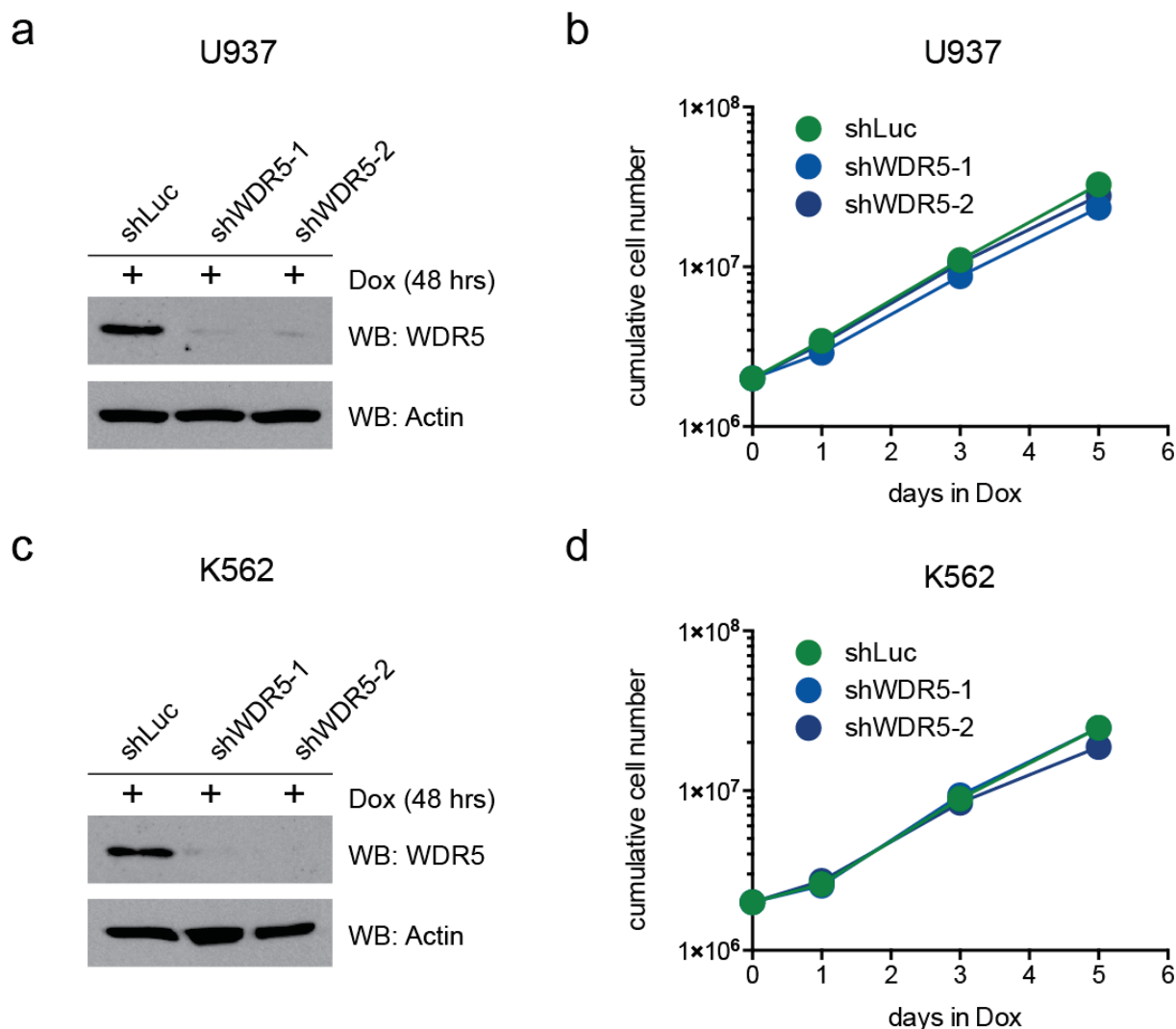


Supplementary Figure 6. Effects of C/EBP α p42 and p30 isoforms and Wdr5 on the granulocytic differentiation potential of 32D cells **(a)** Flow cytometric analysis of Mac-1 surface expression in 32D cells expressing C/EBP α p42 (left panel) or C/EBP α p30 (right panel) after culture in the presence of IL-3 (5 ng/mL, red curves) or G-CSF (10 ng/mL, blue curves) **(a)** qRT-PCR-analysis of the indicated genes in 32D cell expressing C/EBP α p42 (white bars) or C/EBP α p30 (black bars) after culture in the presence of IL-3 or G-CSF for the indicated time. Data are presented as mean \pm standard deviation (SD) of triplicate experiments

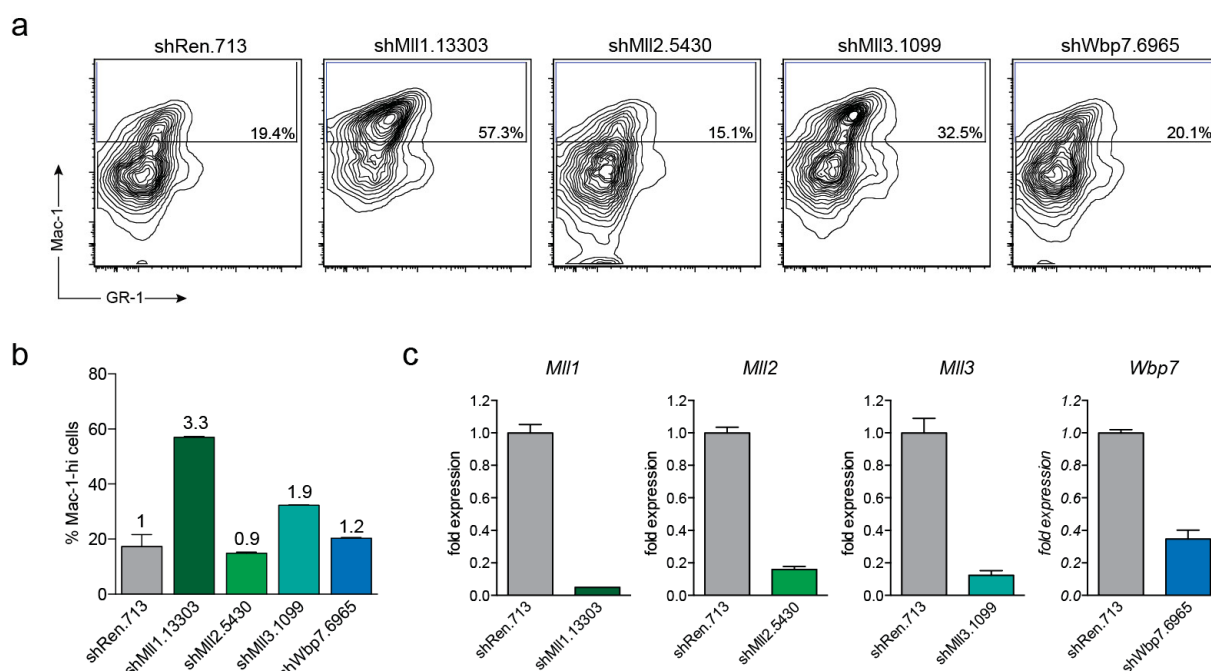


Supplementary Figure 7. Knockdown of Wdr5 has marginal effects on the granulocytic differentiation potential of 32D cells (a) statistical representation of percentages of Mac-1^{hi} cells from the G-CSF experiment described in Fig. 2c and Supplementary Fig. 7b-e. Numbers above bars represent the fold induction in Mac-1^{hi} cells over cells expressing a control shRNA- (b) Flow cytometry analysis for Mac-1 and GR-1 surface markers of 32D rtTA3 cells expressing indicated shRNA constructs 96 h after Dox treatment and 48 h after exposure to G-CSF (10 ng/mL). Presented events are gated on the dsRed⁺ population. (c) qRT-PCR analysis of *Lyz2* expression in 32D rtTA3 cells transduced with indicated shRNA constructs after 96 h of Dox treatment followed by 48 h exposure to G-CSF (10 ng/mL). Data are presented as mean \pm standard deviation (SD) of triplicate experiments. (d) Flow cytometry analysis for Mac-1 and GR-1 surface markers of 32D rtTA3 p42 cells

expressing indicated shRNA constructs 96 h after Dox treatment and 48 h after exposure to G-CSF (10 ng/mL). Presented events are gated on the GFP⁺ dsRed⁺ population. **(e)** qRT-PCR analysis of *Lyz2* expression in 32D rtTA3 p42 cells transduced with indicated shRNA constructs after 96 h of Dox treatment followed by 48 h exposure to G-CSF. Data are presented as mean \pm standard deviation (SD) of triplicate experiments. **(f)** Flow cytometry analysis for Mac-1 and GR-1 surface markers of 32D rtTA3 cells expressing indicated shRNA constructs 96 h after Dox in the presence of IL-3 (5 ng/mL).

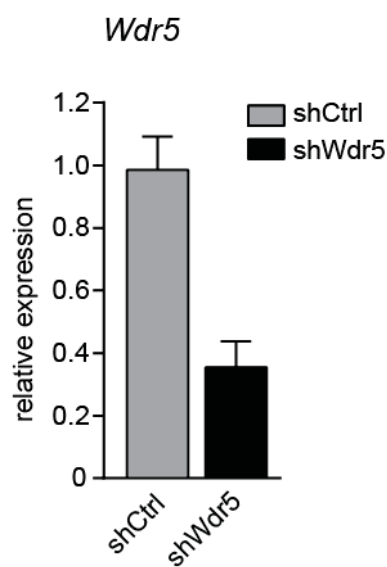


Supplementary Figure 8. Knockdown of WDR5 does not impair proliferation of U937 and K562 cells. (a) Extracts from U937 cells expressing indicated shRNA constructs were analyzed by Western blot for expression of WDR5 and Actin. (b) Proliferation curves of U937 cells expressing indicated shRNA constructs after Dox treatment. (c) Extracts from K562 cells expressing indicated shRNA constructs were analyzed by Western blot for expression of WDR5 and Actin. (d) Proliferation curves of K562 cells expressing indicated shRNA constructs after Dox treatment. Data are presented as mean \pm standard deviation (SD) of triplicate experiments. Representative images of at least 2 replicate experiments are shown.

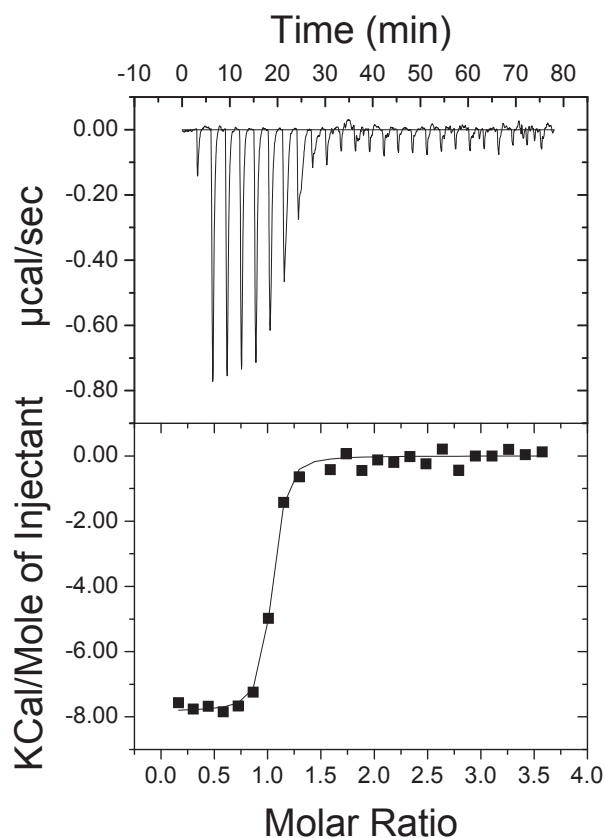


Supplementary Figure 9. Loss of MII1 or MII3, but not MII2 or MII4/Wbp7 restores granulocytic differentiation potential in C/EBP α p30-expressing cells.

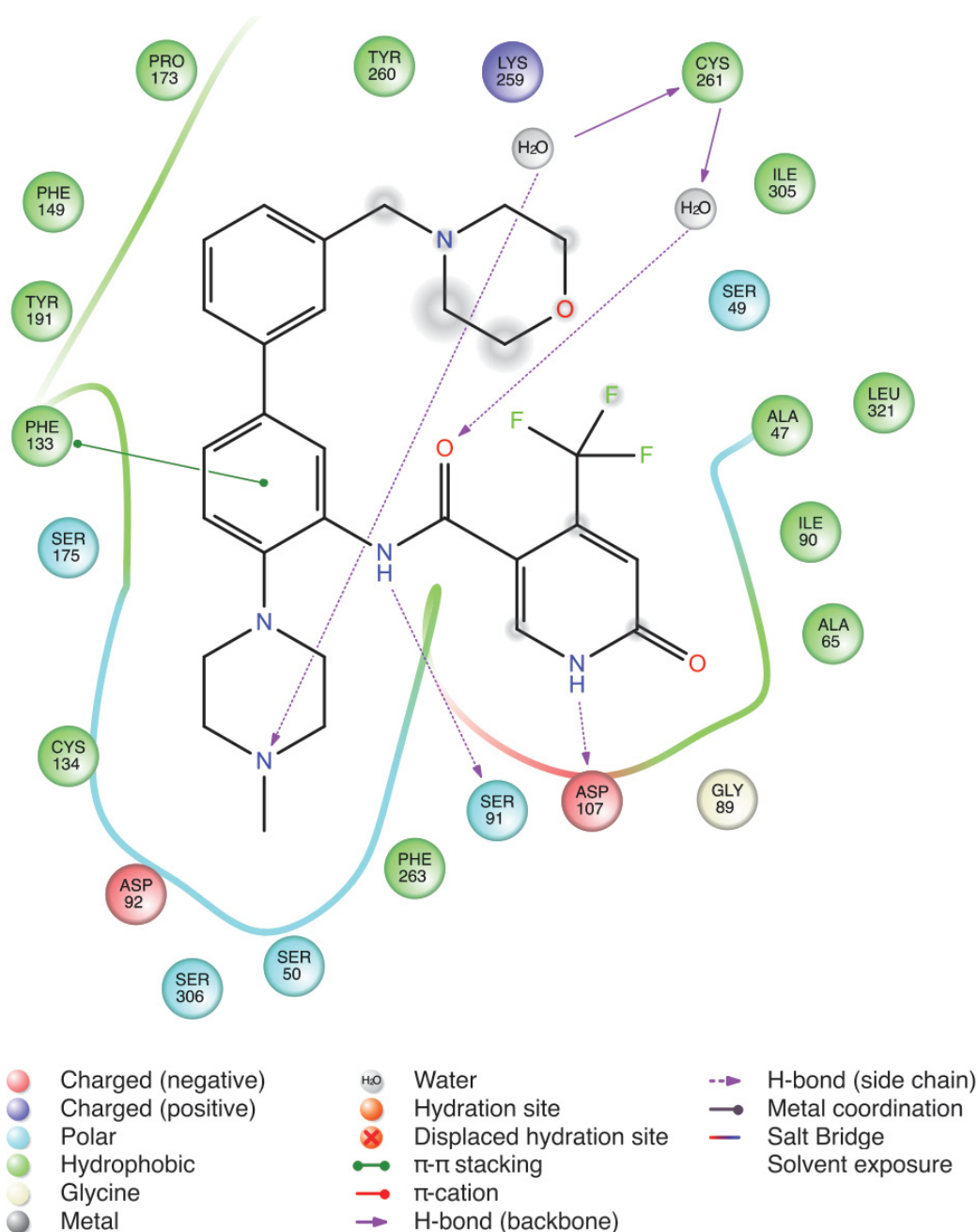
(a) Flow cytometry analysis for Mac-1 and GR-1 surface markers of 32D rtTA3 p30 cells expressing indicated shRNA constructs 96 h after Dox treatment and 48 h after exposure to G-CSF (10 ng/mL). Presented events are gated on the GFP⁺ dsRed⁺ population. **(b)** Statistical representation of percentages of Mac-1^{hi} cells from the experiment described in (a) **(c)** qRT-PCR analysis of the indicated genes *in* 32D rtTA3 p30 cells expressing indicated shRNA constructs after 48 h of Dox treatment. Data are presented as mean \pm standard deviation (SD) of triplicate experiments.



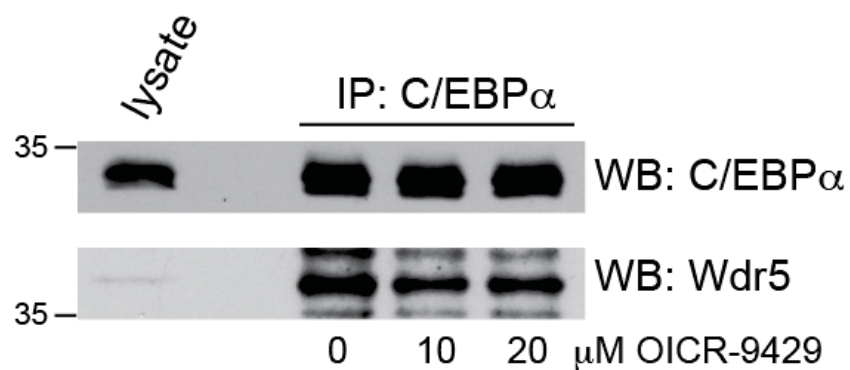
Supplementary Figure 10. Wdr5 knockdown in primary fetal liver cells from *Cebpa*^{p30/p30} mice. qRT-PCR-analysis of Wdr5 expression levels in *Cebpa*^{p30/p30} primary fetal liver cells expressing a control (grey bars) or Wdr5-targeting lentiviral shRNA construct used in Fig. 3a, b. Data are presented as mean +/- standard deviation (SD) of triplicate experiments.



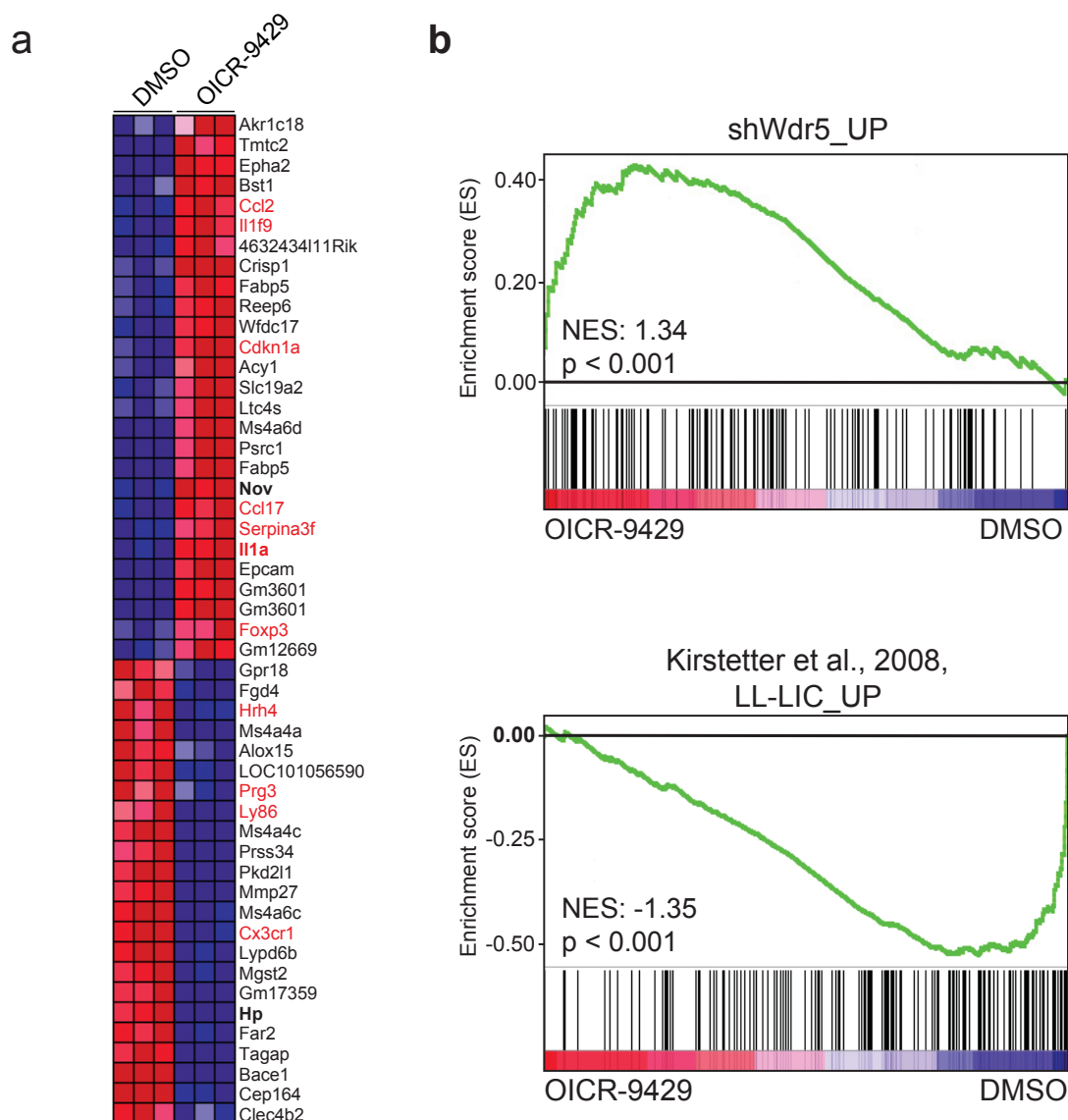
Supplementary Figure 11. OICR-9429 binds to WDR5. Binding of OICR-9429 to WDR5 was assessed by ITC. 10 μ L aliquots of 300 μ M compound solution were titrated over a 15 μ M solution of WDR5. DMSO concentration was kept constant at 3% throughout the experiment. The titrations were performed at 25 $^{\circ}$ C in 100 mM Hepes, 150 mM NaCl, pH 7.4 using a General Electric VP MicroCal calorimeter. A K_D value of 93 ± 28 nM was calculated from 5 independent experiments.



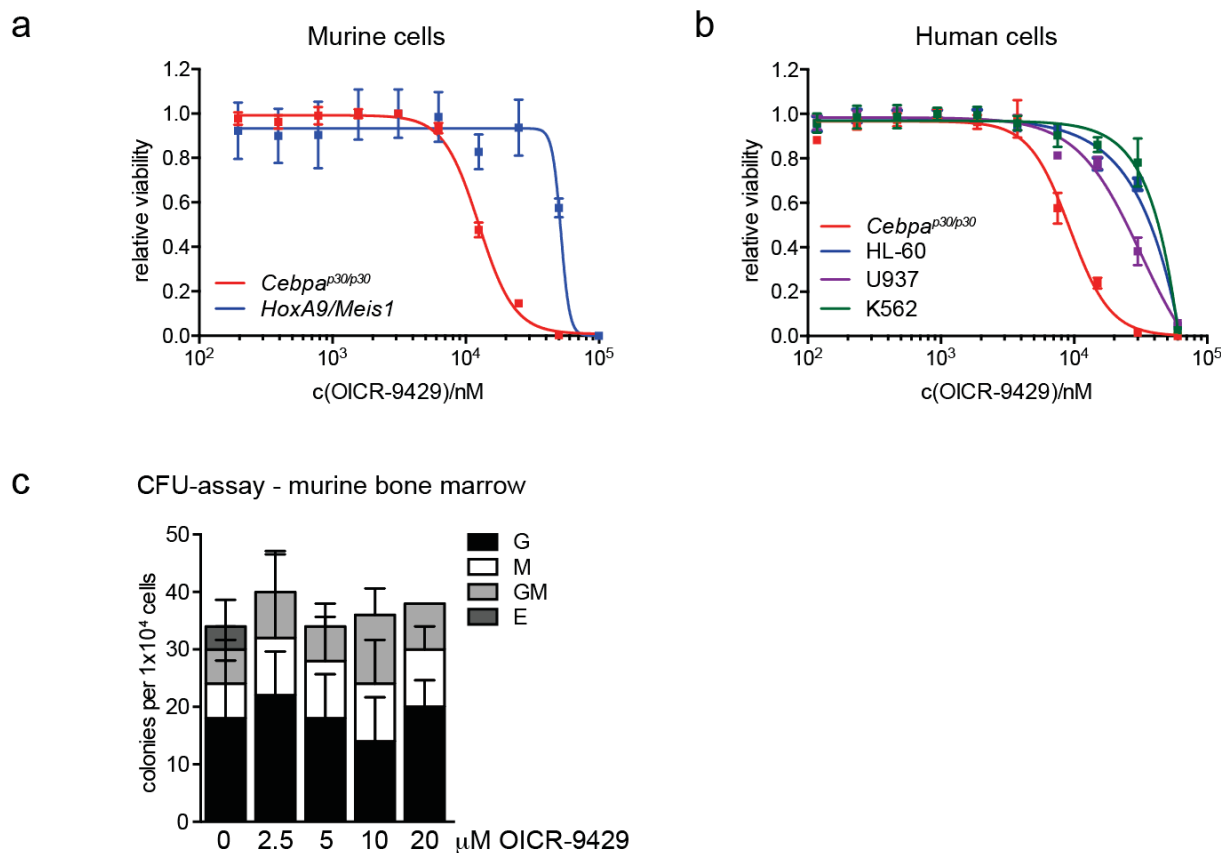
Supplementary Figure 12. Schematic illustration of key atomic interactions between OICR-9429 and WDR5. OICR-9429's south-end (methyl-piperazine), east-end (pyridinone) and amide linker form hydrogen bonds with surrounding residues of WDR5 (colored circles), while the north moiety is surrounded with hydrophobic side-chains (interaction map was generated with Maestro).



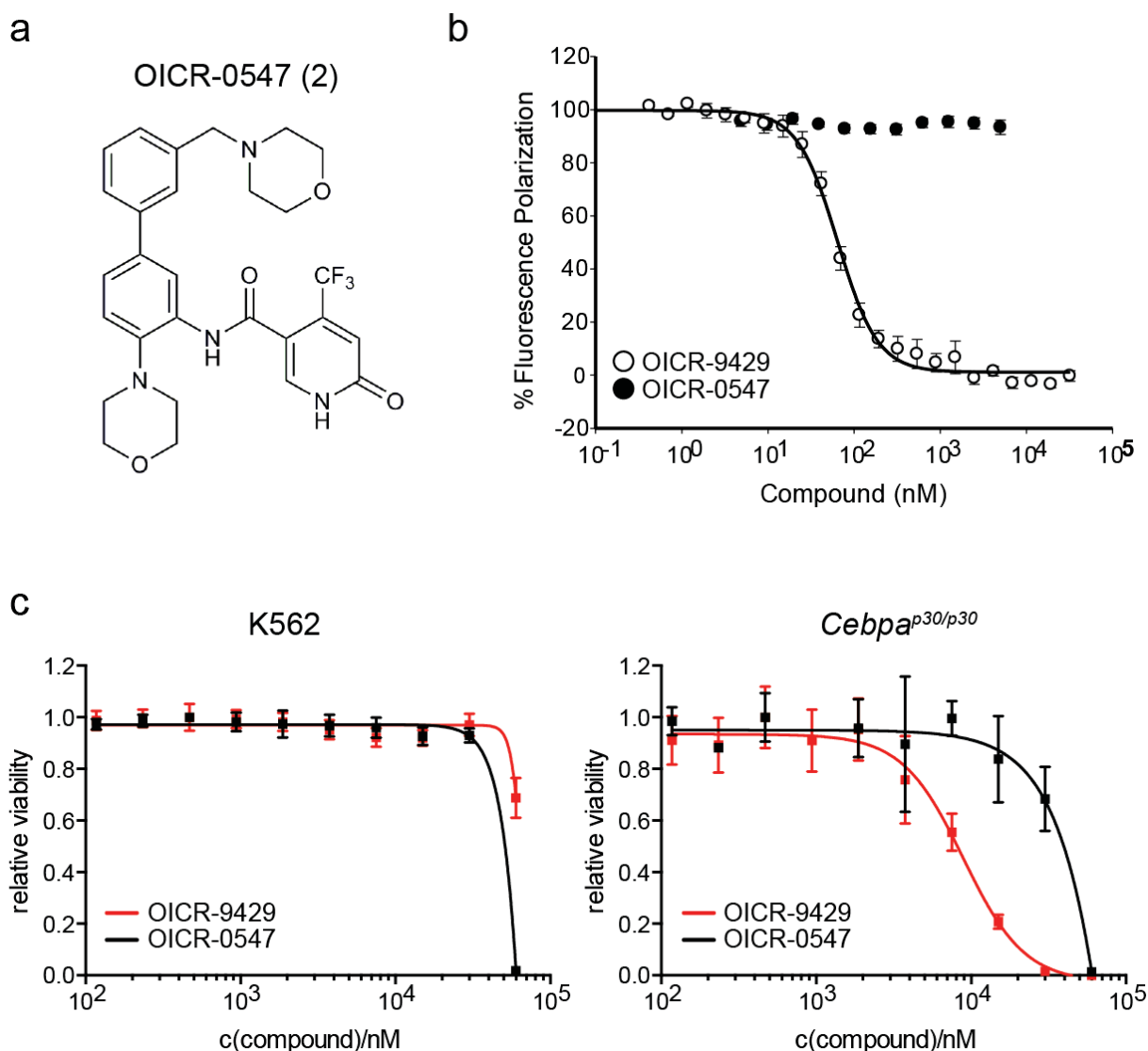
Supplementary Figure 13. OICR-9429 does not inhibit the C/EBP α p30-Wdr5 interaction. Western blot analysis of C/EBP α -immunoprecipitates from *Cepba*^{p30/p30} cells treated with the indicated concentrations of OICR-9429. Representative images of at least 2 replicate experiments are shown.



Supplementary Figure 14. Gene expression analysis of *Cebpa*^{p30/p30} cells upon OICR-9429 treatment. (a) Heatmap representation of differentially regulated genes in *Cebpa*^{p30/p30} cells after 72 h of OICR-9429 treatment (20 μ M, FDR 0.01, fold change 1.75). Myeloid specific genes and are highlighted in red. (b) Gene Set Enrichment Analysis showing global up-regulation of Genes induced by Wdr5 knockdown (top) and global down-regulation of the *Cepba*^{p30/p30} LIC signature upon OICR-9429 treatment of *Cebpa*^{p30/p30} cells (bottom), NES, Normalized Enrichment Score.



Supplementary Figure 15. *Cebpa*^{p30/p30}-cells are specifically sensitive to Wdr5 antagonism. (a) Dose response curves for OICR-9429 in *Cebpa*^{p30/p30} cells (red) and HoxA9/Meis1-transformed wild-type cells (blue). (b) Dose response curves for OICR-9429 in *Cebpa*^{p30/p30} cells (red), HL-60 (blue), U937 (purple) and K562 cells (green). (c) CFU-assay of primary murine bone marrow cells in the presence of indicated concentrations of OICR-9429. Data are presented as mean \pm standard deviation (SD) of triplicate experiments.



Supplementary Figure 16. OICR-0547, an inactive control compound, does not bind to WDR5 and has no effect on cellular viability. (a) Chemical structure of OICR-0547. OICR-0547 was designed as an inactive control compound where the methylpiperazine moiety was replaced with a morpholine ring, which should lead to the inability of the compound to bind in the arginine-binding pocket of WDR5. **(b)** Peptide displacement assay monitoring the decrease in fluorescence polarization (FP) signal of a modified fluorescein labeled MLL peptide upon compound-induced dissociation from WDR5. **(c)** Dose response curves for OICR-9429 (red) and OICR-0547 (black) in K562 cells (left panel) and *Cebpa*^{p30/p30} cells (right panel). Note that OICR-0547 induces non-selective toxicity at high doses in both cell types, which might results from off-target activity. Data are presented as mean \pm standard deviation (SD) of triplicate experiments.

3. Discussion

3.1 The role of chromosomal translocations in leukemia

Discovery of oncogenes contributing to leukemia development and even more importantly, deep understanding of oncogenic mechanisms and their translation to diagnostic and therapeutic approaches represent an ultimate goal in cancer research. Even though the last decade brought a plethora of novel innovative technological approaches, which resulted in an enormous number of scientific studies focusing on cancer research, a complete understanding of the molecular mechanisms driving distinct types of hematological malignances is missing.

The study of oncogenes is intimately linked to the discovery of chromosomal aberrations. Fusion proteins resulting from chromosomal translocations have yielded fundamental insights into a number of cellular events that orchestrate the initiation and progression of neoplasms. The unique features of translocation-associated fusion proteins made them attractive candidates to study the evolution of proteins with novel functions and the molecular machines that provide their functional cellular context have emerged as relevant targets for therapeutic approaches. For instance, expression of BCR-ABL and PML-RARA is associated with acute promyelocytic leukemia (APL) and chronic myelogenous leukemia (CML), respectively. Both diseases are paradigms for the success of targeted cancer therapy, which has turned these fatal diseases into curable conditions. AML displays the highest frequency of balanced chromosomal aberrations and number of gene fusions among all types of cancer. The high prevalence of this phenomenon in AML allows to distinguish families of multi partner translations (MPTs), in which one gene is fused to multiple different recipient loci. The largest MPT family in AML represents translocations of the MLL (Mixed Lineage Leukemia) gene, featuring over 135 translocation partners (Meyer et al., 2017). Furthermore, fusion proteins were described to act in the context of large multi-protein complexes whose composition dictates their molecular functions. A large number of studies confirmed that various interaction partners of fusion proteins are required for the initiation and maintenance of oncogenic programs and that inhibition of their activity can result in the loss of fusion protein oncogenicity.

However, as the majority of insights result from detailed investigation of isolated fusion proteins, it is not clear whether many of the reported molecular mechanisms apply to all members of the respective MPT family. For instance, in the MLL MPT family some critical interactors, such as DOT1L were described to be associated with more than one MLL-fusion protein. Yet, an

underlying molecular mechanism of transformation that is common to the entire group of MLL-fusions has not been proposed. Here, we present a novel experimental pipeline aiming to elucidate conserved molecular mechanisms of action of different MLL-fusion proteins.

In our study, a diverse group of MLL-fusion proteins has been selected: MLL-AF1p, MLL-AF4, MLL-AF9, MLL-CBP, MLL-ENL, MLL-EEN and MLL-GAS7. This subgroup includes translocations that are commonly found in MLL-rearranged leukemias of different lineages, such as MLL-AF4 in MLL-rearranged ALL or MLL-AF9 and MLL-ENL associated with MLL-rearranged AML. The prevalence of other MLL-fusions selected by us is significantly lower, such as MLL-GAS7 or MLL-CBP. Additionally, molecularly distinct mechanisms of action have been proposed for all selected MLL-fusions protein. Thus, the selected subset of MLL-fusions accounts for a comprehensive collection to represent the large repertoire of known MLL-translocations.

3.2 Functional proteomics identifies SETD2 as a critical effector of MLL-fusion proteins

Reasoning that critical effectors of MLL-rearranged leukemia development might be enriched among physical interactors of MLL-fusion proteins, an unbiased approach aiming at identification of protein-protein interaction partners of selected MLL-fusions was employed. Flp-FRT mediated generation of human cell lines allowing for Doxycycline-inducible overexpression of affinity-tagged variants of seven selected MLL-fusion proteins served as an excellent tool to survey the molecular composition of distantly related MLL-fusion proteins in an isogenic background. Functional proteomics identified over 4600 proteins in 140 MS runs. p-value based filtering for the 300 most significant interactors per bait resulted in a protein-protein interaction network of 960 proteins. Many known proteins and protein complexes previously associated with MLL-rearrangements were identified. For instance, multiple members of the MLL-HCFC complex, the SWI/SNF chromatin remodeling complex, the PAF and Polycomb repressive complex 1 were present in the dataset with high coverage. Moreover, each selected MLL-fusion was associated with a large group of unique interaction partners. While these exclusive interactomes have not been further validated in our study, they might serve as an interesting resource for future investigation of specific molecular pathways that are engaged by selected MLL-fusions.

A core subset of 128 conserved binding partners was shared between at least five of the seven selected, structurally different MLL-fusions. This large number indicates that the vast majority of identified proteins in our dataset is engaged in indirect, rather than direct interactions with MLL-fusion proteins. More experimental work would be required to clarify the exact architecture and

position of these proteins within the MLL-fusion-associated complexes. Interestingly, this conserved dataset included known interactors of MLL-fusion complexes, such as wild-type MLL, MEN1, DPY30 or LEDGF, which further validated the chosen approach. Furthermore, bioinformatic analysis of the 128 conserved interactors revealed their affiliation with six distinct protein communities associated with various molecular functions. This analysis already pointed towards DNA damage repair and RNA splicing, molecular functions associated with the methyltransferase SETD2, as will be discussed later.

Various approaches could have been employed for investigating the functional contribution of the conserved set of 128 interactors to MLL-fusion-dependent leukemogenesis. Here, pools of six shRNAs targeting the same candidate gene were cloned into retroviral vectors allowing for transcriptional coupling of shRNA- and GFP expression and introduced into an MLL-AF9-expressing human AML cell line via retroviral transduction. Flow cytometry-based quantitative analysis of GFP expression in a mixed population of cells expressing control- and experimental shRNAs over time enabled the clear assessment of hairpin-mediated effects on cell viability and differentiation status. Alternatively, a pooled approach could have been employed. In fact, a strategy using a pooled NGS-based RNAi approach has been tested, but failed to identify high-confidence candidate genes for further validation steps.

The approach described above led to the identification of SETD2 as a critical effector among the interactors of MLL-fusion proteins. A number of functional experiments proved the importance of SETD2 for the initiation and maintenance of leukemia and its selectivity to the MLL-rearranged genotype.

3.3 The function of SETD2-mediated deposition of H3K36me3

Even though Set2, the *Saccharomyces cerevisiae* homologue of human SETD2 is responsible for all three steps of H3K36 methylation, the human SETD2 protein is a non-redundant *in vivo* writer of H3K36 tri-methylation. RNAi-mediated down-regulation of SETD2 leads to complete loss of H3K36 tri-methylation, while the levels of mono- and di-methylation remain unaffected (Edmunds, Mahadevan et al., 2008). The SET (Suppressor of Variegation, Enhancer of zeste and Thithorax) domain of SETD2 is evolutionarily conserved from yeasts to mammals (Rea, Eisenhaber et al., 2000). The H3K36me3 mark is recognized by a wide range of reader proteins, which bind the mark through chromo- and chromo-like domains, such as PWWP or Tudor. Several studies of Set2 in yeast demonstrated that H3K36me3 is deposited mostly within gene bodies and is strongly associated with active transcription (Bannister, Schneider et al., 2005, Barski,

Cuddapah et al., 2007, Krogan, Kim et al., 2003). This is in line with our results, as we also found strong H3K36me3 occupancy within gene bodies of MLL-fusion-target genes. High H3K36me3 levels were strongly correlated with the presence of the DOT1L-dependent H3K79me2 mark across the same set of genes. Furthermore, we showed that knockdown of SETD2 not only led to downregulation of H3K36me3 but also of H3K79me2. The functional role of this interdependency remains unknown.

Constitutive *Setd2* knockout results in embryonic lethality at E10.5-E11.5. This phenotype was associated with severe vascular defects (Hu, Sun et al., 2010). Interestingly, a conditional *Setd2* knockout mouse models revealed the importance of SETD2 in self-renewal and differentiation programs of multiple hematopoietic compartments. *Setd2*-deficient hematopoietic stem and progenitor cells had activated signaling pathways related to the erythroid specific transcription factor Klf-1, which is often associated with malignant transformation. Moreover, *Setd2* deficiency induced additional replicative stress in hematopoietic stem cells, which affected their proliferation and cell cycle status. These events result in genomic instability, accumulation of secondary mutations and further acceleration of malignant transformation (Zhang, Sun et al., 2018). Independently, another group demonstrated that *Setd2*-knockout mice present with reduced counts of myeloid, lymphoid and megakaryocyte progenitor cells, and a complete depletion of phenotypic and functional hematopoietic stem cells (Zhou, Yan et al., 2018). In line with our observations, hematopoietic stem cells from *Setd2*-conditional knockout mice showed increased apoptosis and differentiation towards mature cell types.

In higher eukaryotes, histone methylation is often associated with the regulation of alternative splicing. Genes containing constitutively included exons feature significantly higher levels of H3K36me3 than alternatively spliced exons (Kolasinska-Zwierz, Down et al., 2009). Furthermore, intron-containing genes were shown to have higher levels of H3K36me3 than intron-less genes in mouse and humans (de Almeida, Grosso et al., 2011). These observations led to the hypothesis that the H3K36me3 signature correlates with the presence of RNA-splicing machineries. In line with this, down-regulation of SETD2 resulted in a global switch of splice sites and lowered the levels of inclusion of exons in actively transcribed genes (Luco, Pan et al., 2010). Proteins containing a Tudor domain, such as MRG15, which is involved in transcriptional regulation, were shown to bind H3K36me2 and H3K36me3 and recruit the RNA splicing machinery (Luco et al., 2010, Zhang, Du et al., 2006). Similarly, the PWWP-domain containing protein ZMYND11 is a reader of H3K36me3, associating with RNA splicing regulatory factors (Wen, Li et al., 2014). Also, bromodomain and PHD finger-containing protein 1 (BRPF1) was found to co-localize with H3K36me3 in gene bodies in complex with the histone acetyltransferase MOZ. BRPF1 also

contains a PWWP domain and its binding to H3K36me3 was linked to transcriptional regulation (Ullah, Pelletier et al., 2008, Vezzoli, Bonadies et al., 2010).

RNA polymerase II constitutes a critical part of the general transcription machinery and interacts with a plethora of transcription factors and accessory proteins. The C-terminal domain (CTD) of RNA Pol II functions as a scaffold for factors involved in the regulation of transcriptional initiation, elongation, termination and mRNA processing. In particular, serine residues in the CTD of RNA Pol II can undergo regulatory phosphorylation events, which are initially catalyzed by cyclin dependent kinases CDK7 and CDK8 during transcriptional initiation and later by CDK9 and CTK1 during RNA processing (Hahn, 2004). A link between SETD2 and transcriptional elongation has been initially described in yeast, as an interaction between the Set2 Rpb1 interacting (SRI) domain of Set2 and the phosphorylated CTD of RNA Pol II was reported (Li, Moazed et al., 2002). Soon after, an analogous interaction was identified in mammalian cells (Li, Phatnani et al., 2005). It was shown that Rpb1, the largest subunit of RNA Pol II, interacts with the SRI domain of SETD2 when it is in the hyper-phosphorylated state. This interaction was proven to be dependent on CTK1 kinase, which phosphorylates Ser2 in the CTD of RNA Pol II (Krogan et al., 2003). Deletion mutants of SETD2 lacking the SRI domain resulted in abrogated interaction with RNA Pol II and reduced H3K36me3 levels (Kizer, Phatnani et al., 2005, Li, Howe et al., 2003).

In addition, SETD2 has been mechanistically linked to transcriptional regulation. Intragenic methylation of DNA by DNMT3B was shown to protect gene bodies from spurious entry of RNA Pol II and cryptic transcriptional initiation. Interestingly, the function of DNMT3B is dependent on its selective recruitment to SETD2-dependent H3K36me3 within gene bodies and H3K36me3 recognition was associated with a functional PWWP domain of DNMT3B (Baubec, Colombo et al., 2015, Neri, Rapelli et al., 2017). This demonstrates that deposition of the H3K36me3 histone mark and the processes of transcriptional elongation might be closely related and could be linked via DNA methylation.

Furthermore, the SETD2-mediated H3K36me3 mark was also implicated in DNA repair processes. For instance, loss of SETD2 was shown to abolish double strand break repair by homologous recombination (HR) (Aymard, Bugler et al., 2014). Down-regulation of SETD2 caused impaired chromatin recruitment of proteins involved in HR, such as 53BP1, RPA and RAD51 (Aymard et al., 2014, Kanu, Gronroos et al., 2015, Pfister, Ahrabi et al., 2014). PHF1, a member of the PRC2 complex is another Tudor domain-containing protein recruited to sites of DNA double strand breaks marked with H3K36me3 and involved in early DNA damage response processes. Thus, loss of SETD2-mediated H3K36me3 resulted in reduced retention of PHF1 at double-strand break sites and impaired DNA damage repair. Furthermore, loss of H3K36me3

precludes the ability of the PRC2 complex to mark transcriptionally active genes with the repressive H3K27me3 histone mark *in vitro* and *in vivo* (Musselman, Avvakumov et al., 2012). Moreover, SETD2 was shown to play an important role in DNA mismatch repair (MMR). H3K36me3 is required for chromatin recruitment of the mismatch recognition complex MutS α *in vitro* and *in vivo* (Li, Mao et al., 2013a). Similar to the recognition mechanisms described above, MSH6, a known component of the MutS α complex, contains a PWWP domain, which interacts with the SETD2-dependent H3K36me3 histone mark. Loss of SETD2-mediated H3K36me3 deposition as well as mutations in the H3K36me3-binding PWWP domain of MSH6 led to elevated frequencies of mutations, establishing the function of the SETD2-mediated H3K36me3 histone mark and its binding domain in the regulation of mismatch repair in human cells (Li et al., 2013a). Moreover, the known interactor of MLL and MLL-fusion proteins LEDGF has been linked to all molecular events mentioned above. LEDGF was initially described as a transcriptional co-activator as well as a component of the transcriptional machinery associated with RNA Pol II (Ge, Si et al., 1998). Through its PWWP domain-mediated binding of H3K36me3, LEDGF was shown to facilitate DNA-end resection in the recruitment step of the homologous recombination process and to recruit the DNA damage repair machinery including C-terminal binding protein interacting protein (CtIP), RPA and Rad51 to sites of DNA damage (Daugaard, Baude et al., 2012, Pfister et al., 2014). Furthermore, it was proposed that LEDGF might modulate splicing through its binding of H3K36me3 and spliceosome factors such as Srsf1 (Pradeepa, Sutherland et al., 2012). Interestingly, SETD2 was not only shown to be essential for the integrity of the histone code, but also to have targets other than lysine 36 in histone 3. It was shown that SETD2 specifically binds and methylates lysine 40 of α -tubulin (α -TubK40me3) (Park, Powell et al., 2016). This process was found to occur on mitotic microtubules, and loss of SETD2 was associated with abnormal mitoses, cytokinesis defects and overall genomic instability. This study provided the first evidence of a potential dual function of SETD2 as a remodeler of chromatin and the cytoskeleton (Park et al., 2016).

Interferon α (INF α)-based therapy is an effective therapeutic strategy against hepatitis B, where INF α -triggered immune signaling induces the production of antiviral proteins. Recently, SETD2 was found to mono-methylate lysine 525 of the transcription factor STAT1 downstream of INF α -activated JAK-STAT signaling (Chen, Liu et al., 2017). This posttranslational modification was necessary for phosphorylation and transcriptional activity of STAT1 and thus to critical to enhance antiviral immunity via strengthening INF α signaling in hepatocytes (Chen et al., 2017).

In summary, the functional repertoire of SETD2 likely includes alternative splicing, transcriptional initiation and elongation and DNA damage repair.

3.4 Mutations of SETD2 in cancer

Mutations in *SETD2* have been discovered in a variety of human malignances. Inactivating mutations of *SETD2* were initially described in clear cell renal cell carcinoma (ccRCC) cell lines and primary cells (Dalglish, Furge et al., 2010, Duns, Hofstra et al., 2012). Studies conducted by The Cancer Genome Atlas Research Network (TCGA) consortium estimated the prevalence of *SETD2* mutations in ccRCC to be about 15.5%, with the majority constituting nonsense and inactivating frameshift events (The Cancer Genome Atlas Research Network, 2013, Gerlinger, Rowan et al., 2012, Gossage, Murtaza et al., 2014). Furthermore, mutations of *SETD2* were frequently found in pediatric high-grade gliomas (HGGs), where almost 30% of patients carried alterations of this gene (Fontebasso, Schwartzentruber et al., 2013). Low-frequency mutations of *SETD2* have been observed in other types of cancer, such as melanoma, high-risk gastrointestinal stromal cancer, breast cancer and bladder cancer, indicating a tumor suppressive role of SETD2 in these malignances (Al Sarakbi, Sasi et al., 2009, Berger, Hodis et al., 2012, The Cancer Genome Atlas Research Network, 2014, Huang, McPherson et al., 2016). SETD2 was also established as a putative tumor suppressor in colorectal cancer, where its deficiency was associated with augmented Wnt/ β -catenin signaling, affecting self-renewal and differentiation of intestinal cells (Yuan, Li et al., 2017). Interestingly, bi-allelic loss of *SETD2* was also associated with a tumor suppressor phenotype in ccRCC (Duns et al., 2012).

Interestingly, somatic mutations in the most important substrate of SETD2 were also identified. Missense mutations in Lys 36 of histone H3 (H3K36I) in chondroblastoma resulted in aberrant expression of genes associated with differentiation processes, reduced ability of cells to differentiate into chondrocytes, osteocytes and adipocytes, increased activity of transcription factors involved with mesenchymal multipotency and eventually enhanced proliferation of immature chondroblast-like cell population (Lu, Jain et al., 2016).

Mutations of *SETD2* were also identified in hematological malignances. For instance, whole genome sequencing studies found that 10% of patients suffering from early T-cell precursor acute lymphoblastic leukemia (ETP-ALL) presented with focal deletions of *SETD2* (Zhang, Ding et al., 2012). Exome sequencing identified bi-allelic loss of function mutations of *SETD2* in patients with mast cell leukemia (MCL) (Martinelli, Mancini et al., 2018). *SETD2* was also found to be the most frequently silenced gene in Enteropathy-associated T cell lymphoma. *SETD2* expression was down-regulated in the developmental stages of $\gamma\delta$ T-cell precursors, thus implicating a role for SETD2 in T-cell maturation (Moffitt, Ondrejka et al., 2017). Mutations of *SETD2* were also found

in 3% of chronic lymphoblastic leukemias (CLL) and were associated with chromothripsis and poor overall survival of these patients (Parker, Rose-Zerilli et al., 2016). Interestingly, mutations in *SETD2* and other epigenetic regulators were enriched in relapsed pediatric acute lymphoblastic leukemia (ALL), indicating a potential role of these events in chemotherapy resistance (Mar, Bullinger et al., 2014). Indeed, the same group recently showed that heterozygous loss of *SETD2* in leukemia resulted in resistance to DNA-damaging agents, such as cytarabine, 6-thioguanine, etoposide and doxorubicin (Mar, Chu et al., 2017).

Furthermore, nonsense and frameshift mutations in *SETD2* were also found in pediatric MLL-rearranged leukemia. Interestingly, shRNA-mediated knockdown of *SETD2* caused a proliferative advantage, increased colony formation and significantly accelerated leukemia development after transplantation of MLL-rearranged leukemia cells *in vitro* and *in vivo* (Zhu, He et al., 2014). These findings established a tumor suppressor role of *SETD2* in cancer.

On the contrary, we have demonstrated that CRISPR/Cas9- and shRNA-mediated loss of *SETD2* leads to enhanced differentiation, DNA damage and apoptosis in AML cells *in vitro* and *in vivo*. Thus, our results establish *SETD2* as a strong dependence rather than a tumor suppressor in MLL-rearranged leukemia. Importantly, in our studies, both shRNA- as well as CRISPR/Cas9-mediated targeting of *SETD2* resulted in near-complete loss of *SETD2* expression. A recent study found that heterozygous loss of *Setd2* accelerates MLL-AF9-induced leukemogenesis and contributes to chemotherapy resistance. However, homozygous *Setd2* knockout strongly delayed MLL-AF9-induced leukemia formation, which is in line with our data. Several other publications describe a similar phenotype. For instance, CRISPR/Cas9-mediated mutagenesis of the *Setd2* SET domain in an MLL-AF9-cell line model caused strong depletion of targeted cells (Shi, Wang et al., 2015). Moreover, several genome-scale CRISPR/Cas9 screens identified *SETD2* as an essential gene in leukemia cells with an MLL-rearranged background (Hart, Chandrashekhara et al., 2015, Tzelepis et al., 2016, Wang, Birsoy et al., 2015). As the vast majority of cancer patients present with heterozygous mutations of *SETD2*, it is likely that *SETD2* acts as a haplo-insufficient tumor suppressor while homozygous loss of *SETD2* efficiently hampers leukemia development. In summary, even though the importance of *SETD2* has been demonstrated in various types of cancer models, the exact role of *SETD2* in cancer remains unclear and more studies are required to clarify the particular function of *SETD2* in MLL-rearranged leukemia.

3.5 Mechanisms of action of *SETD2* in cancer

The mechanism of action of SETD2 in MLL-rearranged leukemia has not been fully established. Several avenues have been examined and the current knowledge suggests a complex interplay between different molecular mechanisms.

It was previously shown that H3K36me3 loss leads to impaired recruitment of DNA-damage response proteins. Thus, SETD2-deficient cells may accumulate a higher mutational burden caused by environmental stress or oncogene-induced DNA damage. This hypothesis is supported by our results. We observed increased DNA damage upon SETD2 knockdown as measured by comet assays and western blotting. We also found activation of p53 and its target gene, p21 in SETD2-depleted cells. Transcriptomic analysis revealed that DNA damage-associated gene sets were significantly enriched in *SETD2* knockdown cells. A different phenotype has been observed in ccRCC cells, where loss of SETD2 caused inefficient activation of p53 and further accumulation of DNA damage (Carvalho, Vitor et al., 2014). This indicates that the observed phenotype might be tissue specific and dependent on the p53 status. Additionally, the architecture of the protein complexes that are engaged by SETD2 might be particularly different. Our data indicate that in MLL-rearranged leukemia SETD2 interacts with the N-terminal portion of the MLL protein, which is retained after chromosomal translocation. Identification of protein complexes containing SETD2 in ccRCC might help to characterize the existence of other mechanisms resulting in different phenotypes.

It has been shown that the presence of H3K36me3 is necessary to restore a normal chromatin state in bodies of transcribed genes after the passage of the RNA Pol II machinery. Many factors involved in the control of this process have been identified, including the SPT6 or FACT (Facilitates Chromatin Transcription) complex. These complexes serve as histone chaperons and aid the elongation machinery in its passage through nucleosome re-positioning (Cheung, Chua et al., 2008). Furthermore, these factors are responsible for the repression of cryptic transcription on intragenic transcription sites. In line with this, it was reported that downregulation of SETD2 leads to induction of cryptic transcription from intragenic sites within over 1000 genes (Carvalho, Raposo et al., 2013). At the same time, the occupancy of RNA Pol II at intragenic regions of active genes was significantly increased (Carvalho et al., 2013). Moreover, histone exchange after the passage of RNA Pol II is SETD2-dependent. Lower levels of H3K36me3 after SETD2 downregulation resulted in reduced binding of FACT complex subunits on active genes. This indicates that chromatin binding of protein complexes involved in transcriptional control might be dependent on SETD2-mediated histone marks (Carvalho et al., 2013).

In agreement with this, the role of SETD2 in MLL-rearranged leukemia could be linked to chromatin targeting of MLL-fusion complexes by the PWWP-domain containing protein LEDGF,

a known interactor of MLL chimeras. Thus, loss of H3K36me3 may result in compromised chromatin binding of MLL-complexes and deregulation of MLL-fusion target genes expression. This hypothesis is supported by our experiments, as we found that *SETD2* knockdown resulted in reduced binding of MLL-AF9 to promoter regions of the known MLL target genes *Meis1* and *HoxA9*, leading to reduced expression of these genes.

Another study demonstrated that the ASH1L-dependent H3K36me2 histone mark is preferentially bound by LEDGF *in vivo* and recruits MLL to promoter regions, implying that H3K36me3 was dispensable for MLL occupancy. In fact, reduction of H3K36me3 was associated with upregulation of H3K36me2, thus enhancing the recruitment of LEDGF containing MLL complex to chromatin (Zhu, Li et al., 2016).

Furthermore, a cross-talk between H3K36me3 and other epigenetic marks has been proposed. A recent report showed that loss of H3K36me3 leads to increased global tri-methylation of H3K27 in sarcomas (Lu et al., 2016). Conversely, binding of the Tudor domain containing protein PHF1 to H3K36me3 sterically precludes the ability of the Polycomb complex PRC2 to methylate lysine 27 of histone 3 *in vitro* and *in vivo* (Musselman et al., 2012). Moreover, a mechanistic link between H3K36me3 and acetylation of H4K16 was shown. H3K36me3 can trigger deposition of the H4K16ac mark upon induction of DNA damage. Both histone marks were elevated upon treatment with double strand break-inducing agents (Li & Wang, 2017). Loss of H3K36me3 led to increased levels of ASH1L-mediated deposition of H3K36me2 in MLL-AF9 rearranged cells. It was also observed that the global chromatin occupancy of MLL-fusion complexes was shifted to the gene bodies marked with H3K36me2 (Zhu et al., 2016). We identified a dependence between SETD2-dependent deposition of H3K36me3 and DOT1L-dependent deposition of H3K79me2. A dual H3K36me3-H3K27me2 signature was strongly associated with MLL-target genes, and loss of SETD2 not only reduced the density of H3K36me3 but also caused a reduction in H3K79me2 levels. Epigenetic cross-talk between H3K36me3 and H3K79me2 was also described in a recent study, which found aberrant regulation of tumor suppressor genes and oncogenes upon SETD2 loss. In particular, loss of SETD2 was linked to downregulated expression of the tumor suppressor *ASXL1* and upregulation of *ERG*, which mediated acceleration of leukemia development (Bu, Chen et al., 2018). Interestingly, the authors reported that SETD2 wild type cells expressing MLL-AF9 presented with increased H3K36me3 as compared to non-translocated cells and that loss of SETD2-mediated histone mark was associated with increased rates of H3K79me2.

Overall, an unambiguous mechanism of action of SETD2 in cancer and particularly in MLL-rearranged leukemia has not been revealed. However, recent reports have established the role

of SETD2 in DNA damage repair as well as a number of epigenetic dependencies associated to aberrant expression of genes affecting leukemia development.

3.6 Potential therapeutic opportunities

To further clarify the activity and role of SETD2 in malignant diseases, the availability of a small molecule to probe the functions of SETD2 in cancer would be of high interest. A screen of N-alkyl sinefungin derivatives against a panel of methyltransferases identified a potent inhibitor of SETD2 (Zheng, Ibanez et al., 2012). Unfortunately, sinefungin is not cell permeable, which precludes its applications in the majority of cell-based functional experiments.

The concept of gene-drug interaction is of particular interest in cancer therapy, where small molecules might be particularly active against cancer cells that have specific genomic vulnerabilities or resistance mechanisms. In line with this, mutations of *SETD2* sensitized ccRCC tumor cells to the PI3K β inhibitors TGX221 or AZD6482 and this treatment significantly decreased the invasive potential of tumor cells (Feng, Sun et al., 2015). Moreover, a synthetic lethality screening approach revealed that H3K36me3-deficient cancer cells were significantly more sensitive to small-molecule-mediated inhibition of WEE1 kinase using AZD1775. Deficiency of H3K36me3 was sufficient to induce down-regulated expression of RRM2. RRM2 is a ribonucleotide reductase subunit responsible for catalyzing the biosynthesis of deoxyribonucleotides, thus contributing to DNA replication stress in S phase of the cell cycle. In addition, WEE1 inhibition activates CDK1/2, which leads to further reduction of RRM2 expression. This combinatorial effect resulted in strong reduction of RRM2 expression, decrease of the dNTP pool and caused S-phase arrest and apoptosis (Pfister, Markkanen et al., 2015).

In our manuscript we found a combinatorial effect of SETD2 downregulation and pharmacological inhibition of DOT1L using EPZ5676 (Pinometostat). Combined SETD2/DOT1L targeting caused growth inhibition, induction of myeloid differentiation and apoptosis of MLL-rearranged cell lines. We also showed that loss of SETD2 and DOT1L inhibition synergize in the induction of DNA damage.

These examples emphasize the need for efficient inhibitory agents of SETD2. Chemical probes as well as approaches based on combinatorial therapies might be essential for further characterization of SETD2 role in hematological malignances but also later, as a therapeutic strategy for patients suffering from MLL-rearranged leukemia.

3.7 Interaction proteomics identifies Wdr5 as an interactor of the p30 isoform of C/EBP α in AML

A similar experimental pipeline was employed to identify critical effectors associated with the short isoform of C/EBP α , a transcription factor that is frequently mutated in acute myeloid leukemia. Reasoning that interactors of the p30 variant of C/EBP α might be critical for leukemia development, affinity-purification experiments of both wild type- (p42) and the short, leukemia associated p30 isoforms of C/EBP α were performed and protein complexes were characterized by LC-MS/MS. Wdr5, a known component of SET1/MLL histone methyltransferase complexes was preferably associated with p30, which was validated by co-immunoprecipitation experiments. Furthermore, we showed that Wdr5 was responsible for maintenance of the p30-mediated differentiation block, as shRNA-mediated downregulation of *Wdr5* resulted in granulocytic differentiation of C/EBP α p30-expressing cells *in vitro* and *in vivo*. Finally, p30-expressing cells were particularly sensitive to OICR-9429, a small-molecule selective antagonist of the Wdr5-MLL interaction.

This manuscript represents another link between leukemia development and epigenetic regulation. Wdr5 is an integral component of the WRAD complex that is associated with MLL and mediates deposition of the H3K4me3 histone mark. We have described Wdr5 to be a key factor controlling the differentiation potential of cells. Similar to SETD2, critical effectors in C/EBP α mutant AML appear to be necessary for the maintenance of the differentiation block in the leukemic cell population and their loss restores terminal myeloid maturation. Interestingly, this effect could be phenocopied by inactivation of Wdr5 with the small molecule OICR-9429.

Overall, the presented experimental pipeline characterized the gain of function properties of C/EBP α p30 and revealed a functional link between p30 and Wdr5 as responsible for the transforming properties of C/EBP α p30. Targeting the MLL complex in C/EBP α mutated AML might thus constitute an attractive alternative for treatment of malignances caused by mutated transcription factors.

3.8 Conclusions and future prospects

Overall, we have established an efficient experimental pipeline that enables the identification of critical effectors of oncogenic mechanisms in AML. This approach was applied to the MLL multi-partner translocation family, as represented by seven MLL translocation partners in a systematic fashion. Three distinct experimental parts complemented our comprehensive approach. Initially, a comparative biochemical characterization of MLL-fusion family was performed, resulting in a

detailed charting of the interactome of seven MLL-fusion proteins. Functional analysis of conserved candidate gene lists that were assembled based on interactome data through RNAi screening identified SETD2 as an essential interactor of MLL-fusion proteins. Finally, in-depth validation of SETD2 as a critical effector in MLL-rearranged leukemia was performed *in vitro* and *in vivo*, identifying a role for SETD2 in maintaining genomic integrity in MLL-AML.

Our combined proteomic-functional genomic analysis of MLL-fusion-interactors shed new light on the role of SETD2 in AML with MLL -fusions. Our study provides interesting insights into the biology of MLL-rearranged leukemia and highlights the importance of a dependency of MLL-fusions on SETD2 during leukemia development and maintenance. Taken together, the results presented in this thesis establish SETD2 as a potential target for patients suffering from MLL-rearranged acute myeloid leukemia.

Although SETD2 has attracted a lot of attention in recent years, a lot of controversy and open questions exist regarding its specific role in hematopoietic malignances.

For instance, it remains unknown whether SETD2 has specific target genes in MLL-rearranged cells. Thus, chromatin immunoprecipitation followed by sequencing (ChIP-seq) experiments would be instrumental to reveal global patterns of SETD2 chromatin binding. However, the lack of a suitable ChIP-grade antibody with a strong affinity to endogenously expressed SETD2 represents a major hurdle towards reaching these goals. CRISPR/Cas9-mediated tagging of endogenous proteins might turn out to be of help to overcome this technical drawback. In line with this, ChIP-seq of MLL-fusion proteins after the loss of SETD2-mediated H3K36me3 would contribute additional insight about the molecular basis of the strong dependence of MLL-leukemia on SETD2.

A large body of evidence suggests a role for SETD2 in the DNA-damage response in various types of cancer. It was shown that low levels of H3K36me3 impair DNA damage response and result in higher mutational rates (Mar et al., 2017). ChIP-seq analysis of proteins involved in the DNA-damage response might reveal distinct sets of genes that are involved in this process. For instance, given our observation of elevated levels of phosphorylation of histone H2AX, γ -H2AX ChIP-seq in SETD2 deficient cells might identify sites of DNA damage that are associated with the loss of SETD2.

Alternatively, as SETD2 was implicated in transcriptional regulation, it would be intriguing to investigate the global transcriptional status in SETD2-deficient versus wild type MLL-rearranged cells. A number of specific NGS-based assays focusing on this molecular process have been developed in recent years, including nascent RNA-seq, Pro-seq, GRO-seq or TT-seq (Core,

Waterfall et al., 2008, Garcia-Cuellar, Buttner et al., 2016, Gardini, 2017, Kwak, Fuda et al., 2013, Schwalb, Michel et al., 2016). Last but not least, a potent SETD2 inhibitor would contribute to a better understanding of the role of this protein in various types of cancer.

In summary, we have identified a novel dependency for the lysine methyltransferase SETD2 in the initiation and maintenance of MLL-rearranged leukemia, which highlights an unexpected vulnerability in this disease. Given the controversy about the role of SETD2 in cancer, including leukemia harboring MLL-translocations, additional experimental work is required to validate SETD2 as an actionable target in MLL-rearranged leukemia.

4. References

- Abe Y, Nagano M, Tada A, Adachi J, Tomonaga T (2017) Deep phosphotyrosine proteomics by optimization of phosphotyrosine enrichment and MS/MS parameters. *J Proteome Res* 16: 1077-1086
- Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422: 198-207
- Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* 537: 347-55
- Al Sarakbi W, Sasi W, Jiang WG, Roberts T, Newbold RF, Mokbel K (2009) The mRNA expression of SETD2 in human breast cancer: correlation with clinico-pathological parameters. *BMC cancer* 9: 290
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjord JE, Foekens JA, Greaves M et al. (2013) Signatures of mutational processes in human cancer. *Nature* 500: 415-21
- Altman JK, Foran JM, Pratz KW, Trone D, Cortes JE, Tallman MS (2018) Phase 1 study of quizartinib in combination with induction and consolidation chemotherapy in patients with newly diagnosed acute myeloid leukemia. *American journal of hematology* 93: 213-221
- American Cancer Society. Cancer Facts and Figures 2017. Atlanta: American Cancer Society; 2017
- Ang YS, Tsai SY, Lee DF, Monk J, Su J, Ratnakumar K, Ding J, Ge Y, Darr H, Chang B, Wang J, Rendl M, Bernstein E, Schaniel C, Lemischka IR (2011) Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* 145: 183-97
- Annesley CE, Brown P (2014) The biology and targeting of FLT3 in pediatric leukemia. *Front Oncol* 4: 263
- Arai S, Jagasia M, Storer B, Chai X, Pidala J, Cutler C, Arora M, Weisdorf DJ, Flowers ME, Martin PJ, Palmer J, Jacobsohn D, Pavletic SZ, Vogelsang GB, Lee SJ (2011) Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. *Blood* 118: 4242-9
- Ardehali MB, Mei A, Zobeck KL, Caron M, Lis JT, Kusch T (2011) Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *Embo j* 30: 2817-28
- Argiropoulos B, Humphries RK (2007) Hox genes in hematopoiesis and leukemogenesis. *Oncogene* 26: 6766-76
- Armstrong SA, Golub TR, Korsmeyer SJ (2003) MLL-rearranged leukemias: insights from gene expression profiling. *Seminars in hematology* 40: 268-73
- Armstrong SA, Kung AL, Mabon ME, Silverman LB, Stam RW, Den Boer ML, Pieters R, Kersey JH, Sallan SE, Fletcher JA, Golub TR, Griffin JD, Korsmeyer SJ (2003) Inhibition of FLT3 in

- MLL. Validation of a therapeutic target identified by gene expression based classification. *Cancer Cell* 3: 173-83
- Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, Sallan SE, Lander ES, Golub TR, Korsmeyer SJ (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 30: 41-7
- Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S, Iacovoni JS, Daburon V, Miller KM, Jackson SP, Legube G (2014) Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* 21: 366-74
- Bach C, Buhl S, Mueller D, Garcia-Cuellar MP, Maethner E, Slany RK (2010) Leukemogenic transformation by HOXA cluster genes. *Blood* 115: 2910-8
- Bagger FO, Sasivarevic D, Sohi SH, Laursen LG, Pundhir S, S nderby CK, Winther O, Rapin N, Porse BT (2016) BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic Acids Research* 44: D917-D924
- Baker MS, Ahn SB, Mohamedali A, Islam MT, Cantor D, Verhaert PD, Fanayan S, Sharma S, Nice EC, Connor M, Ranganathan S (2017) Accelerating the search for the missing proteins in the human proteome. *Nature communications* 8: 14271
- Balgobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A, Beverloo HB, Chang M, Creutzig U, Dworzak MN, Forestier E, Gibson B, Hasle H, Harrison CJ, Heerema NA, Kaspers GJ, Leszl A, Litvinko N, Nigro LL, Morimoto A et al. (2009) Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. *Blood* 114: 2489-96
- Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T (2005) Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *The Journal of biological chemistry* 280: 17732-6
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-37
- Basavapathruni A, Jin L, Daigle SR, Majer CR, Therkelsen CA, Wigle TJ, Kuntz KW, Chesworth R, Pollock RM, Scott MP, Moyer MP, Richon VM, Copeland RA, Olhava EJ (2012) Conformational adaptation drives potent, selective and durable inhibition of the human protein methyltransferase DOT1L. *Chemical biology & drug design* 80: 971-80
- Basecke J, Whelan JT, Griesinger F, Bertrand FE (2006) The MLL partial tandem duplication in acute myeloid leukaemia. *Br J Haematol* 135: 438-49
- Bastian M. HS, Jacomy M. (2009) Gephi: an open source software for exploring and manipulating networks.
- Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schubeler D (2015) Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* 520: 243-7

- Berger MF, Hodis E, Heffernan TP, Deribe YL, Lawrence MS, Protopopov A, Ivanova E, Watson IR, Nickerson E, Ghosh P, Zhang H, Zeid R, Ren X, Cibulskis K, Sivachenko AY, Wagle N, Sucker A, Sougnez C, Onofrio R, Ambrogio L et al. (2012) Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature* 485: 502-6
- Bernt KM, Zhu N, Sinha AU, Vempati S, Faber J, Krivtsov AV, Feng Z, Punt N, Daigle A, Bullinger L, Pollock RM, Richon VM, Kung AL, Armstrong SA (2011) MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* 20: 66-78
- Biswas D, Milne TA, Basrur V, Kim J, Elenitoba-Johnson KS, Allis CD, Roeder RG (2011) Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc Natl Acad Sci U S A* 108: 15751-6
- Bitoun E, Oliver PL, Davies KE (2007) The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Human molecular genetics* 16: 92-106
- Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, Ma Y, Moore R, Mungall AJ, Marra MA, Zhang J, Ma X, Liu Y, Liu Y, Auviel JMG, Davidsen TM, Gesuwan P, Hermida LC, Salhia B, Capone S et al. (2018) The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nature medicine* 24: 103-112
- Brock HW, van Lohuizen M (2001) The polycomb group--no longer an exclusive club? *Curr Opin Genet Dev* 11: 175-81
- Brummelkamp TR, Bernards R, Agami R (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296: 550-3
- Bu J, Chen A, Yan X, He F, Dong Y, Zhou Y, He J, Zhan D, Lin P, Hayashi Y, Sun Y, Zhang Y, Xiao Z, Grimes HL, Wang QF, Huang G (2018) SETD2-mediated crosstalk between H3K36me3 and H3K79me2 in MLL-rearranged leukemia. *Leukemia* 32: 890-899
- Buechele C, Breese EH, Schneidawind D, Lin CH, Jeong J, Duque-Afonso J, Wong SH, Smith KS, Negrin RS, Porteus M, Cleary ML (2015) MLL leukemia induction by genome editing of human CD34+ hematopoietic cells. *Blood* 126: 1683-94
- Burckstummer T, Bennett KL, Preradovic A, Schutze G, Hantschel O, Superti-Furga G, Bauch A (2006) An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nature methods* 3: 1013-9
- Bursen A, Schwabe K, Ruster B, Henschler R, Ruthardt M, Dingermann T, Marschalek R (2010) The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL.AF4. *Blood* 115: 3570-9
- Calkins GN, Boveri T (1914) Zur Frage der Entstehung maligner Tumoren. *Science* 40: 857-859
- The Cancer Genome Atlas Research Network (2013) Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 499: 43-9

- The Cancer Genome Atlas Research Network (2014) Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* 507: 315-22
- The Cancer Genome Atlas Research Network (2015) Genomic classification of cutaneous melanoma. *Cell* 161: 1681-96
- Cao F, Townsend EC, Karatas H, Xu J, Li L, Lee S, Liu L, Chen Y, Ouillet P, Zhu J, Hess JL, Atadja P, Lei M, Qin ZS, Malek S, Wang S, Dou Y (2014) Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol Cell* 53: 247-61
- Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell* 136: 642-55
- Carvalho S, Raposo AC, Martins FB, Grosso AR, Sridhara SC, Rino J, Carmo-Fonseca M, de Almeida SF (2013) Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. *Nucleic Acids Res* 41: 2881-93
- Carvalho S, Vitor AC, Sridhara SC, Martins FB, Raposo AC, Desterro JM, Ferreira J, de Almeida SF (2014) SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *eLife* 3: e02482
- Caslini C, Connelly JA, Serna A, Broccoli D, Hess JL (2009) MLL associates with telomeres and regulates telomeric repeat-containing RNA transcription. *Mol Cell Biol* 29: 4519-26
- Caslini C, Yang Z, El-Osta M, Milne TA, Slany RK, Hess JL (2007) Interaction of MLL amino terminal sequences with menin is required for transformation. *Cancer Res* 67: 7275-83
- Cermakova K, Tesina P, Demeulemeester J, El Ashkar S, Mereau H, Schwaller J, Rezacova P, Veverka V, De Rijck J (2014) Validation and structural characterization of the LEDGF/p75-MLL interface as a new target for the treatment of MLL-dependent leukemia. *Cancer Res* 74: 5139-51
- Chen C-WD, Delaney C, Xu H, Koche R, Lujambio Goizueta A, Hoshii T, Cusan M, Chabon J, Valerio DG, Kuehn M, Wang X, Cai S, Chu SH, Chen YK, He J, Huang C-H, Feng Z, Krivtsov A, Qi J, Bradner JE et al. (2016) An epigenetic regulator screen identifies novel targets that tensitize MLL-rearranged leukemia to DOT1L inhibition. *Blood* 128: 571-571
- Chen CW, Armstrong SA (2015) Targeting DOT1L and HOX gene expression in MLL-rearranged leukemia and beyond. *Exp Hematol* 43: 673-84
- Chen CW, Koche RP, Sinha AU, Deshpande AJ, Zhu N, Eng R, Doench JG, Xu H, Chu SH, Qi J, Wang X, Delaney C, Bernt KM, Root DE, Hahn WC, Bradner JE, Armstrong SA (2015) DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia. *Nature medicine* 21: 335-43
- Chen K, Liu J, Liu S, Xia M, Zhang X, Han D, Jiang Y, Wang C, Cao X (2017) Methyltransferase SETD2-mediated methylation of STAT1 is critical for interferon antiviral activity. *Cell* 170: 492-506.e14
- Chen SJ, Shen Y, Chen Z (2013) A panoramic view of acute myeloid leukemia. *Nat Genet* 45: 586-7

- Chen Y, Wan B, Wang KC, Cao F, Yang Y, Protacio A, Dou Y, Chang HY, Lei M (2011) Crystal structure of the N-terminal region of human Ash2L shows a winged-helix motif involved in DNA binding. *EMBO reports* 12: 797-803
- Cheung N, Chan LC, Thompson A, Cleary ML, So CWE (2007) Protein arginine-methyltransferase-dependent oncogenesis. *Nature Cell Biology* 9: 1208-1215
- Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. *PLoS Biol* 6: e277
- Choi H, Larsen B, Lin ZY, Breitkreutz A, Mellacheruvu D, Fermin D, Qin ZS, Tyers M, Gingras AC, Nesvizhskii AI (2011) SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nature methods* 8: 70-3
- Cierpicki T, Grembecka J (2014) Challenges and opportunities in targeting the menin-MLL interaction. *Future medicinal chemistry* 6: 447-62
- Cobas M, Wilson A, Ernst B, Mancini SJ, MacDonald HR, Kemler R, Radtke F (2004) Beta-catenin is dispensable for hematopoiesis and lymphopoiesis. *J Exp Med* 199: 221-9
- Collins EC, Pannell R, Simpson EM, Forster A, Rabbitts TH (2000) Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development. *EMBO reports* 1: 127-32
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339: 819-23
- Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322: 1845-8
- Corral J, Lavenir I, Impey H, Warren AJ, Forster A, Larson TA, Bell S, McKenzie AN, King G, Rabbitts TH (1996) An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell* 85: 853-61
- Cotton RG (1993) Current methods of mutation detection. *Mutation research* 285: 125-44
- Cruickshank VA, Sroczynska P, Sankar A, Miyagi S, Rundsten CF, Johansen JV, Helin K (2015) SWI/SNF subunits SMARCA4, SMARCD2 and DPF2 collaborate in MLL-rearranged leukaemia maintenance. *PLoS One* 10: e0142806
- Daigle SR, Olhava EJ, Therkelsen CA, Basavapathruni A, Jin L, Boriack-Sjodin PA, Allain CJ, Klaus CR, Raimondi A, Scott MP, Waters NJ, Chesworth R, Moyer MP, Copeland RA, Richon VM, Pollock RM (2013) Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood* 122: 1017-25
- Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, Johnston LD, Scott MP, Smith JJ, Xiao Y, Jin L, Kuntz KW, Chesworth R, Moyer MP, Bernt KM, Tseng JC, Kung AL, Armstrong SA, Copeland RA, Richon VM et al. (2011) Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* 20: 53-65

- Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, Davies H, Edkins S, Hardy C, Latimer C, Teague J, Andrews J, Barthorpe S, Beare D, Buck G, Campbell PJ, Forbes S, Jia M, Jones D, Knott H et al. (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* 463: 360-3
- Daser A, Rabbitts TH (2004) Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis. *Genes Dev* 18: 965-74
- Daugaard M, Baude A, Fugger K, Povlsen LK, Beck H, Sorensen CS, Petersen NH, Sorensen PH, Lukas C, Bartek J, Lukas J, Rohde M, Jaattela M (2012) LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nat Struct Mol Biol* 19: 803-10
- Dawson MA, Prinjha RK, Dittmann A, Giotopoulos G, Bantscheff M, Chan WI, Robson SC, Chung CW, Hopf C, Savitski MM, Huthmacher C, Gudgin E, Lugo D, Beinke S, Chapman TD, Roberts EJ, Soden PE, Auger KR, Mirguet O, Doehner K et al. (2011) Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* 478: 529-33
- de Almeida SF, Grosso AR, Koch F, Fenouil R, Carvalho S, Andrade J, Levezinho H, Gut M, Eick D, Gut I, Andrau JC, Ferrier P, Carmo-Fonseca M (2011) Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat Struct Mol Biol* 18: 977-83
- Deshpande AJ, Chen L, Fazio M, Sinha AU, Bernt KM, Banka D, Dias S, Chang J, Olhava EJ, Daigle SR, Richon VM, Pollock RM, Armstrong SA (2013) Leukemic transformation by the MLL-AF6 fusion oncogene requires the H3K79 methyltransferase Dot1L. *Blood* 121: 2533-41
- DiMartino JF, Miller T, Ayton PM, Landewe T, Hess JL, Cleary ML, Shilatifard A (2000) A carboxy-terminal domain of ELL is required and sufficient for immortalization of myeloid progenitors by MLL-ELL. *Blood* 96: 3887-93
- Dohner H, Weisdorf DJ, Bloomfield CD (2015) Acute myeloid leukemia. *N Engl J Med* 373: 1136-52
- Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD, Roeder RG (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 13: 713-9
- Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, Allis CD, Chait BT, Hess JL, Roeder RG (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* 121: 873-85
- Dulbecco R (1986) A turning point in cancer research: sequencing the human genome. *Science* 231: 1055-6
- Duns G, Hofstra RM, Sietzema JG, Hollema H, van Duivenbode I, Kuik A, Giezen C, Jan O, Bergsma JJ, Bijnen H, van der Vlies P, van den Berg E, Kok K (2012) Targeted exome sequencing in clear cell renal cell carcinoma tumors suggests aberrant chromatin regulation as a crucial step in ccRCC development. *Human mutation* 33: 1059-62
- Edmunds JW, Mahadevan LC, Clayton AL (2008) Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *Embo j* 27: 406-20

- El Ashkar S, Schwaller J, Pieters T, Goossens S, Demeulemeester J, Christ F, Van Belle S, Juge S, Boeckx N, Engelman A, Van Vlierberghe P, Debyser Z, De Rijck J (2018) LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. *Blood* 131: 95-107
- Elguoshy A, Magdeldin S, Xu B, Hirao Y, Zhang Y, Kinoshita N, Takisawa Y, Nameta M, Yamamoto K, El-Refy A, El-Fiky F, Yamamoto T (2016) Why are they missing? : Bioinformatics characterization of missing human proteins. *Journal of proteomics* 149: 7-14
- Emerenciano M, Kowarz E, Karl K, de Almeida Lopes B, Scholz B, Bracharz S, Meyer C, Pombo-de-Oliveira MS, Marschalek R (2013a) Functional analysis of the two reciprocal fusion genes MLL-NEBL and NEBL-MLL reveal their oncogenic potential. *Cancer letters* 332: 30-4
- Emerenciano M, Meyer C, Mansur MB, Marschalek R, Pombo-de-Oliveira MS (2013b) The distribution of MLL breakpoints correlates with outcome in infant acute leukaemia. *Br J Haematol* 161: 224-36
- Erb MA, Scott TG, Li BE, Xie H, Paulk J, Seo HS, Souza A, Roberts JM, Dastjerdi S, Buckley DL, Sanjana NE, Shalem O, Nabet B, Zeid R, Offei-Addo NK, Dhe-Paganon S, Zhang F, Orkin SH, Winter GE, Bradner JE (2017) Transcription control by the ENL YEATS domain in acute leukaemia. *Nature* 543: 270-274
- Ernst P, Mabon M, Davidson AJ, Zon LI, Korsmeyer SJ (2004) An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Current biology : CB* 14: 2063-9
- Ernst P, Vakoc CR (2012) WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Briefings in functional genomics* 11: 217-26
- Ernst P, Wang J, Huang M, Goodman RH, Korsmeyer SJ (2001) MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Mol Cell Biol* 21: 2249-58
- Fair K, Anderson M, Bulanova E, Mi H, Tropschug M, Diaz MO (2001) Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Mol Cell Biol* 21: 3589-97
- Farrar JE, Schuback HL, Ries RE, Wai D, Hampton OA, Trevino LR, Alonzo TA, Guidry Auvil JM, Davidsen TM, Gesuwan P, Hermida L, Muzny DM, Dewal N, Rustagi N, Lewis LR, Gamis AS, Wheeler DA, Smith MA, Gerhard DS, Meshinchi S (2016) Genomic profiling of pediatric acute myeloid leukemia reveals a changing mutational landscape from disease diagnosis to relapse. *Cancer Res* 76: 2197-205
- Felix CA, Hosler MR, Winick NJ, Masterson M, Wilson AE, Lange BJ (1995) ALL-1 gene rearrangements in DNA topoisomerase II inhibitor-related leukemia in children. *Blood* 85: 3250-6
- Fellmann C, Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, Lai DY, Barbosa IA, Kwon JS, Guan Y, Sinha N, Zuber J (2013) An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep* 5: 1704-13
- Fellmann C, Lowe SW (2014) Stable RNA interference rules for silencing. *Nat Cell Biol* 16: 10-8

- Feng C, Sun Y, Ding G, Wu Z, Jiang H, Wang L, Ding Q, Wen H (2015) PI3Kbeta inhibitor TGX221 selectively inhibits renal cell carcinoma cells with both VHL and SETD2 mutations and links multiple pathways. *Scientific reports* 5: 9465
- Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Current biology* : CB 12: 1052-8
- Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, Philpott M, Munro S, McKeown MR, Wang Y, Christie AL, West N, Cameron MJ, Schwartz B, Heightman TD, La Thangue N et al. (2010) Selective inhibition of BET bromodomains. *Nature* 468: 1067-73
- Fong CY, Gilan O, Lam EY, Rubin AF, Ftouni S, Tyler D, Stanley K, Sinha D, Yeh P, Morison J, Giotopoulos G, Lugo D, Jeffrey P, Lee SC, Carpenter C, Gregory R, Ramsay RG, Lane SW, Abdel-Wahab O, Kouzarides T et al. (2015) BET inhibitor resistance emerges from leukaemia stem cells. *Nature* 525: 538-42
- Fontebasso AM, Schwartzentruber J, Khuong-Quang DA, Liu XY, Sturm D, Korshunov A, Jones DT, Witt H, Kool M, Albrecht S, Fleming A, Hadjadj D, Busche S, Lepage P, Montpetit A, Staffa A, Gerges N, Zakrzewska M, Zakrzewski K, Liberski PP et al. (2013) Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. *Acta neuropathologica* 125: 659-69
- Forster A, Pannell R, Drynan LF, McCormack M, Collins EC, Daser A, Rabbitts TH (2003) Engineering de novo reciprocal chromosomal translocations associated with MLL to replicate primary events of human cancer. *Cancer Cell* 3: 449-58
- Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM, Whelan S, Organisation WH (2013) International Classification of Diseases for Oncology.
- Gallooly MM, Lazarus HM, Cooper BW (2017) Midostaurin: a novel therapeutic agent for patients with FLT3-mutated acute myeloid leukemia and systemic mastocytosis. *Therapeutic advances in hematology* 8: 245-261
- Garcia-Cuellar MP, Buttner C, Bartenhagen C, Dugas M, Slany RK (2016) Leukemogenic MLL-ENL fusions induce alternative chromatin states to drive a functionally dichotomous group of target genes. *Cell Rep* 15: 310-22
- Garcia-Cuellar MP, Fuller E, Mathner E, Breitingner C, Hetzner K, Zeitlmann L, Borkhardt A, Slany RK (2014) Efficacy of cyclin-dependent-kinase 9 inhibitors in a murine model of mixed-lineage leukemia. *Leukemia* 28: 1427-35
- Garcia-Cuellar MP, Schreiner SA, Birke M, Hamacher M, Fey GH, Slany RK (2000) ENL, the MLL fusion partner in t(11;19), binds to the c-Abl interactor protein 1 (ABI1) that is fused to MLL in t(10;11)+. *Oncogene* 19: 1744-51
- Garcia-Cuellar MP, Steger J, Fuller E, Hetzner K, Slany RK (2015) Pbx3 and Meis1 cooperate through multiple mechanisms to support Hox-induced murine leukemia. *Haematologica* 100: 905-13

- Gardini A (2017) Global Run-On Sequencing (GRO-Seq). *Methods Mol Biol* 1468: 111-20
- Ge H, Si Y, Roeder RG (1998) Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *Embo j* 17: 6723-9
- Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald NQ, Butler A, Jones D, Raine K, Latimer C, Santos CR et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366: 883-892
- Germano G, Morello G, Aveic S, Pinazza M, Minuzzo S, Frasson C, Persano L, Bonvini P, Viola G, Bresolin S, Tregnago C, Paganin M, Pigazzi M, Indraccolo S, Basso G (2017) ZNF521 sustains the differentiation block in MLL-rearranged acute myeloid leukemia. *Oncotarget* 8: 26129-26141
- Gessner A, Thomas M, Castro PG, Buchler L, Scholz A, Brummendorf TH, Soria NM, Vormoor J, Greil J, Heidenreich O (2010) Leukemic fusion genes MLL/AF4 and AML1/MTG8 support leukemic self-renewal by controlling expression of the telomerase subunit TERT. *Leukemia* 24: 1751-9
- Giambruno R, Grebien F, Stukalov A, Knoll C, Planyavsky M, Rudashevskaya EL, Colinge J, Superti-Furga G, Bennett KL (2013) Affinity purification strategies for proteomic analysis of transcription factor complexes. *J Proteome Res* 12: 4018-27
- Gilan O, Lam EY, Becher I, Lugo D, Cannizzaro E, Joberty G, Ward A, Wiese M, Fong CY, Ftouni S, Tyler D, Stanley K, MacPherson L, Weng CF, Chan YC, Ghisi M, Smil D, Carpenter C, Brown P, Garton N et al. (2016) Functional interdependence of BRD4 and DOT1L in MLL leukemia. *Nat Struct Mol Biol* 23: 673-81
- Gillert E, Leis T, Repp R, Reichel M, Hosch A, Breitenlohner I, Angermuller S, Borkhardt A, Harbott J, Lampert F, Griesinger F, Greil J, Fey GH, Marschalek R (1999) A DNA damage repair mechanism is involved in the origin of chromosomal translocations t(4;11) in primary leukemic cells. *Oncogene* 18: 4663-71
- Glatter T, Ahrne E, Schmidt A (2015) Comparison of different sample preparation protocols reveals lysis buffer-specific extraction biases in gram-negative bacteria and human Cells. *J Proteome Res* 14: 4472-85
- Goo YH, Sohn YC, Kim DH, Kim SW, Kang MJ, Jung DJ, Kwak E, Barlev NA, Berger SL, Chow VT, Roeder RG, Azorsa DO, Meltzer PS, Suh PG, Song EJ, Lee KJ, Lee YC, Lee JW (2003) Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins. *Mol Cell Biol* 23: 140-9
- Gossage L, Murtaza M, Slatter AF, Lichtenstein CP, Warren A, Haynes B, Marass F, Roberts I, Shanahan SJ, Claas A, Dunham A, May AP, Rosenfeld N, Forshew T, Eisen T (2014) Clinical and pathological impact of VHL, PBRM1, BAP1, SETD2, KDM6A, and JARID1c in clear cell renal cell carcinoma. *Genes, chromosomes & cancer* 53: 38-51
- Gough SM, Slape CI, Aplan PD (2011) NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. *Blood* 118: 6247-57

- Grandage VL, Gale RE, Linch DC, Khwaja A (2005) PI3-kinase/Akt is constitutively active in primary acute myeloid leukaemia cells and regulates survival and chemoresistance via NF-kappaB, Mapkinase and p53 pathways. *Leukemia* 19: 586-94
- Greaves MF, Maia AT, Wiemels JL, Ford AM (2003) Leukemia in twins: lessons in natural history. *Blood* 102: 2321-33
- Grembecka J, He S, Shi A, Purohit T, Muntean AG, Sorenson RJ, Showalter HD, Murai MJ, Belcher AM, Hartley T, Hess JL, Cierpicki T (2012) Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. *Nature chemical biology* 8: 277-84
- Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441: 537-41
- Grimwade D, Hills RK, Moorman AV, Walker H, Chatters S, Goldstone AH, Wheatley K, Harrison CJ, Burnett AK (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: 354-65
- Gu S, Jin L, Zhang Y, Huang Y, Zhang F, Valdmanis PN, Kay MA (2012) The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. *Cell* 151: 900-11
- Guenther MG, Lawton LN, Rozovskaia T, Frampton GM, Levine SS, Volkert TL, Croce CM, Nakamura T, Canaani E, Young RA (2008) Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. *Genes Dev* 22: 3403-8
- Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* 11: 394-403
- Hall PA, Russell SE (2004) The pathobiology of the septin gene family. *The Journal of pathology* 204: 489-505
- Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* 100: 57-70
- Hart T, Chandrashekhara M, Aregger M, Steinhart Z, Brown Kevin R, MacLeod G, Mis M, Zimmermann M, Fradet-Turcotte A, Sun S, Mero P, Dirks P, Sidhu S, Roth Frederick P, Rissland Olivia S, Durocher D, Angers S, Moffat J (2015) High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* 163: 1515-1526
- Hayward NK, Wilmott JS, Waddell N, Johansson PA, Field MA, Nones K, Patch AM, Kakavand H, Alexandrov LB, Burke H, Jakrot V, Kazakoff S, Holmes O, Leonard C, Sabarinathan R, Mularoni L, Wood S, Xu Q, Waddell N, Tembe V et al. (2017) Whole-genome landscapes of major melanoma subtypes. *Nature* 545: 175-180
- He S, Malik B, Borkin D, Miao H, Shukla S, Kempinska K, Purohit T, Wang J, Chen L, Parkin B, Malek SN, Danet-Desnoyers G, Muntean AG, Cierpicki T, Grembecka J (2015) Menin-MLL inhibitors block oncogenic transformation by MLL-fusion proteins in a fusion partner-independent manner. *Leukemia* 30: 508-513

- Hom RA, Chang PY, Roy S, Musselman CA, Glass KC, Selezneva AI, Gozani O, Ismagilov RF, Cleary ML, Kutateladze TG (2010) Molecular mechanism of MLL PHD3 and RNA recognition by the Cyp33 RRM domain. *Journal of molecular biology* 400: 145-54
- Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ (2003) Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol Cell Biol* 23: 186-94
- Hu M, Sun XJ, Zhang YL, Kuang Y, Hu CQ, Wu WL, Shen SH, Du TT, Li H, He F, Xiao HS, Wang ZG, Liu TX, Lu H, Huang QH, Chen SJ, Chen Z (2010) Histone H3 lysine 36 methyltransferase Hypb/Setd2 is required for embryonic vascular remodeling. *Proc Natl Acad Sci U S A* 107: 2956-61
- Huang KK, McPherson JR, Tay ST, Das K, Tan IB, Ng CC, Chia NY, Zhang SL, Myint SS, Hu L, Rajasegaran V, Huang D, Loh JL, Gan A, Sairi AN, Sam XX, Dominguez LT, Lee M, Soo KC, Ooi LL et al. (2016) SETD2 histone modifier loss in aggressive GI stromal tumours. *Gut* 65: 1960-1972
- Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC, Hayes DN, Shanmugam KS, Bhattacharjee A, Biondi CA, Kay GF, Hayward NK, Hess JL, Meyerson M (2004) Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. *Mol Cell* 13: 587-97
- Huret JL, Dessen P, Bernheim A (2001) An atlas of chromosomes in hematological malignancies. Example: 11q23 and MLL partners. *Leukemia* 15: 987-9
- Huret JL, Dessen P, Bernheim A (2003) Atlas of Genetics and Cytogenetics in Oncology and Haematology, year 2003. *Nucleic Acids Res* 31: 272-4
- Huret JL, Dessen P, Le Minor S, Bernheim A (2000) The "Atlas of genetics and cytogenetics in oncology and haematology" on the internet and a review on infant leukemias. *Cancer genetics and cytogenetics* 120: 155-9
- Huret JL, Senon S, Bernheim A, Dessen P (2004) An Atlas on genes and chromosomes in oncology and haematology. *Cellular and molecular biology (Noisy-le-Grand, France)* 50: 805-7
- Ida K, Kitabayashi I, Taki T, Taniwaki M, Noro K, Yamamoto M, Ohki M, Hayashi Y (1997) Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13). *Blood* 90: 4699-704
- Jackson PK (2009) Navigating the deubiquitinating proteome with a CompPASS. *Cell* 138: 222-4
- Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* 19: 523-34
- Jeannet G, Scheller M, Scarpellino L, Duboux S, Gardiol N, Back J, Kuttler F, Malanchi I, Birchmeier W, Leutz A, Huelsken J, Held W (2008) Long-term, multilineage hematopoiesis occurs in the combined absence of beta-catenin and gamma-catenin. *Blood* 111: 142-9

- Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD, Chun HB, Tough DF, Prinjha RK, Benner C, Glass CK (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell* 51: 310-25
- Kanu N, Gronroos E, Martinez P, Burrell RA, Yi Goh X, Bartkova J, Maya-Mendoza A, Mistrik M, Rowan AJ, Patel H, Rabinowitz A, East P, Wilson G, Santos CR, McGranahan N, Gulati S, Gerlinger M, Birnbak NJ, Joshi T, Alexandrov LB et al. (2015) SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. *Oncogene* 34: 5699-708
- Karatas H, Townsend EC, Cao F, Chen Y, Bernard D, Liu L, Lei M, Dou Y, Wang S (2013) High-affinity, small-molecule peptidomimetic inhibitors of MLL1/WDR5 protein-protein interaction. *Journal of the American Chemical Society* 135: 669-82
- Karisch R, Fernandez M, Taylor P, Virtanen C, St-Germain JR, Jin LL, Harris IS, Mori J, Mak TW, Senis YA, Ostman A, Moran MF, Neel BG (2011) Global proteomic assessment of the classical protein-tyrosine phosphatome and "Redoxome". *Cell* 146: 826-40
- Khwaja A, Bjorkholm M, Gale RE, Levine RL, Jordan CT, Ehninger G, Bloomfield CD, Estey E, Burnett A, Cornelissen JJ, Scheinberg DA, Bouscary D, Linch DC (2016) Acute myeloid leukaemia. *Nature reviews Disease primers* 2: 16010
- Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabudhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LD et al. (2014) A draft map of the human proteome. *Nature* 509: 575-81
- Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol* 25: 3305-16
- Koch U, Wilson A, Cobas M, Kemler R, Macdonald HR, Radtke F (2008) Simultaneous loss of beta- and gamma-catenin does not perturb hematopoiesis or lymphopoiesis. *Blood* 111: 160-4
- Kocher T, Superti-Furga G (2007) Mass spectrometry-based functional proteomics: from molecular machines to protein networks. *Nature methods* 4: 807-15
- Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* 41: 376-81
- Kornblau SM, Tibes R, Qiu YH, Chen W, Kantarjian HM, Andreeff M, Coombes KR, Mills GB (2009) Functional proteomic profiling of AML predicts response and survival. *Blood* 113: 154-64
- Krivtsov AV, Armstrong SA (2007) MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* 7: 823-33
- Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, Xia X, Jesneck J, Bracken AP, Silverman LB, Kutok JL, Kung AL, Armstrong SA (2008) H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* 14: 355-68

- Krivtsov AV, Figueroa ME, Sinha AU, Stubbs MC, Feng Z, Valk PJ, Delwel R, Dohner K, Bullinger L, Kung AL, Melnick AM, Armstrong SA (2013) Cell of origin determines clinically relevant subtypes of MLL-rearranged AML. *Leukemia* 27: 852-60
- Krogan NJ, Dover J, Khorrami S, Greenblatt JF, Schneider J, Johnston M, Shilatifard A (2002) COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *The Journal of biological chemistry* 277: 10753-5
- Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003) Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* 23: 4207-18
- Kwak H, Fuda NJ, Core LJ, Lis JT (2013) Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* 339: 950-3
- Lavallee VP, Baccelli I, Kros J, Wilhelm B, Barabe F, Gendron P, Boucher G, Lemieux S, Marinier A, Meloche S, Hebert J, Sauvageau G (2015) The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. *Nat Genet* 47: 1030-7
- Lavau C, Du C, Thirman M, Zeleznik-Le N (2000) Chromatin-related properties of CBP fused to MLL generate a myelodysplastic-like syndrome that evolves into myeloid leukemia. *Embo j* 19: 4655-64
- Lavau C, Szilvassy SJ, Slany R, Cleary ML (1997) immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *EMBO J* 16: 4226-37
- Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH, Pugh TJ, Stransky N, Helman E, Kim J et al. (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499: 214-218
- Lewin R (1987) National Academy looks at human genome project, sees progress. *Science* 235: 747-8
- Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, Robertson A, Hoadley K, Triche TJ, Jr., Laird PW, Baty JD, Fulton LL, Fulton R, Heath SE, Kalicki-Veizer J, Kandoth C, Klco JM, Koboldt DC, Kanchi KL, Kulkarni S, Lamprecht TL et al. (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 368: 2059-74
- Li B, Howe L, Anderson S, Yates JR, 3rd, Workman JL (2003) The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *The Journal of biological chemistry* 278: 8897-903
- Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li GM (2013a) The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell* 153: 590-600
- Li J, Moazed D, Gygi SP (2002) Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *The Journal of biological chemistry* 277: 49383-8

- Li L, Wang Y (2017) Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair. *The Journal of biological chemistry* 292: 11951-11959
- Li M, Phatnani HP, Guan Z, Sage H, Greenleaf AL, Zhou P (2005) Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc Natl Acad Sci U S A* 102: 17636-41
- Li Y (2010) Commonly used tag combinations for tandem affinity purification. *Biotechnology and applied biochemistry* 55: 73-83
- Li Y (2011) The tandem affinity purification technology: an overview. *Biotechnology letters* 33: 1487-99
- Li Z, Zhang Z, Li Y, Arnovitz S, Chen P, Huang H, Jiang X, Hong GM, Kunjamma RB, Ren H, He C, Wang CZ, Elkahoun AG, Valk PJ, Dohner K, Neilly MB, Bullinger L, Delwel R, Lowenberg B, Liu PP et al. (2013b) PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood* 121: 1422-31
- Libura J, Slater DJ, Felix CA, Richardson C (2005) Therapy-related acute myeloid leukemia-like MLL rearrangements are induced by etoposide in primary human CD34+ cells and remain stable after clonal expansion. *Blood* 105: 2124-31
- Libura M, Asnafi V, Tu A, Delabesse E, Tigaud I, Cymbalista F, Bennaceur-Griscelli A, Villarese P, Solbu G, Hagemeijer A, Beldjord K, Hermine O, Macintyre E (2003) FLT3 and MLL intragenic abnormalities in AML reflect a common category of genotoxic stress. *Blood* 102: 2198-204
- Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, Shilatifard A (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Molecular Cell* 37: 429-437
- Lu C, Jain SU, Hoelper D, Bechet D, Molden RC, Ran L, Murphy D, Venneti S, Hameed M, Pawel BR, Wunder JS, Dickson BC, Lundgren SM, Jani KS, De Jay N, Papillon-Cavanagh S, Andrulis IL, Sawyer SL, Grynszpan D, Turcotte RE et al. (2016) Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* 352: 844-9
- Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T (2010) Regulation of alternative splicing by histone modifications. *Science* 327: 996-1000
- Luo Z, Lin C, Shilatifard A (2012) The super elongation complex (SEC) family in transcriptional control. *Nature reviews Molecular cell biology* 13: 543-7
- Ma X, Liu Y, Liu Y, Alexandrov LB, Edmonson MN, Gawad C, Zhou X, Li Y, Rusch MC, Easton J, Huether R, Gonzalez-Pena V, Wilkinson MR, Hermida LC, Davis S, Sioson E, Pounds S, Cao X, Ries RE, Wang Z et al. (2018) Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* 555: 371-376
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339: 823-6

- Mar BG, Bullinger LB, McLean KM, Grauman PV, Harris MH, Stevenson K, Neuberg DS, Sinha AU, Sallan SE, Silverman LB, Kung AL, Lo Nigro L, Ebert BL, Armstrong SA (2014) Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. *Nature communications* 5: 3469
- Mar BG, Chu SH, Kahn JD, Krivtsov AV, Koche R, Castellano CA, Kotliar JL, Zon RL, McConkey ME, Chabon J, Chappell R, Grauman PV, Hsieh JJ, Armstrong SA, Ebert BL (2017) SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood* 130: 2631-2641
- Marschalek R (2011a) It takes two-to-leukemia: about addictions and requirements. *Leuk Res* 35: 424-5
- Marschalek R (2011b) Mechanisms of leukemogenesis by MLL fusion proteins. *Br J Haematol* 152: 141-54
- Marschalek R (2016) Systematic Classification of Mixed-Lineage Leukemia Fusion Partners Predicts Additional Cancer Pathways. *Ann Lab Med* 36(2):85-100
- Martelli AM, Tazzari PL, Evangelisti C, Chiarini F, Blalock WL, Billi AM, Manzoli L, McCubrey JA, Cocco L (2007) Targeting the phosphatidylinositol 3-kinase/Akt/mammalian target of rapamycin module for acute myelogenous leukemia therapy: from bench to bedside. *Current medicinal chemistry* 14: 2009-23
- Martens JH, Stunnenberg HG (2010) The molecular signature of oncofusion proteins in acute myeloid leukemia. *FEBS letters* 584: 2662-9
- Martin ME, Milne TA, Bloyer S, Galoian K, Shen W, Gibbs D, Brock HW, Slany R, Hess JL (2003) Dimerization of MLL fusion proteins immortalizes hematopoietic cells. *Cancer Cell* 4: 197-207
- Martinelli G, Mancini M, De Benedittis C, Rondoni M, Papayannidis C, Manfrini M, Meggendorfer M, Calogero R, Guadagnuolo V, Fontana MC, Bavaro L, Padella A, Zago E, Pagano L, Zanotti R, Scaffidi L, Specchia G, Albano F, Merante S, Elena C et al. (2018) SETD2 and histone H3 lysine 36 methylation deficiency in advanced systemic mastocytosis. *Leukemia* 32: 139-148
- McDonald ER, 3rd, de Weck A, Schlabach MR, Billy E, Mavrakis KJ, Hoffman GR, Belur D, Castelletti D, Frias E, Gampa K, Golji J, Kao I, Li L, Megel P, Perkins TA, Ramadan N, Ruddy DA, Silver SJ, Sovath S, Stump M et al. (2017) Project DRIVE: a compendium of cancer dependencies and synthetic lethal relationships uncovered by large-scale, deep RNAi screening. *Cell* 170: 577-592.e10
- McMahon KA, Hiew SY, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ, Kioussis D, Williams O, Brady HJ (2007) Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* 1: 338-45
- Mellacheruvu D, Wright Z, Couzens AL, Lambert JP, St-Denis NA, Li T, Miteva YV, Hauri S, Sardiou ME, Low TY, Halim VA, Bagshaw RD, Hubner NC, Al-Hakim A, Bouchard A, Faubert D, Fermin D, Dunham WH, Goudreault M, Lin ZY et al. (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nature methods* 10: 730-6

- Mereau H, De Rijck J, Cermakova K, Kutz A, Juge S, Demeulemeester J, Gijsbers R, Christ F, Debyser Z, Schwaller J (2013) Impairing MLL-fusion gene-mediated transformation by dissecting critical interactions with the lens epithelium-derived growth factor (LEDGF/p75). *Leukemia* 27: 1245-53
- Mertens F, Johansson B, Fioretos T, Mitelman F (2015) The emerging complexity of gene fusions in cancer. *Nat Rev Cancer* 15: 371-81
- Meshinchi S, Woods WG, Stirewalt DL, Sweetser DA, Buckley JD, Tjoa TK, Bernstein ID, Radich JP (2001) Prevalence and prognostic significance of Flt3 internal tandem duplication in pediatric acute myeloid leukemia. *Blood* 97: 89-94
- Mestas J, Hughes CC (2004) Of mice and not men: differences between mouse and human immunology. *Journal of immunology (Baltimore, Md : 1950)* 172: 2731-8
- Meyer C, Burmeister T, Gröger D, Tsaur G, Fechina L, Renneville A, Sutton R, Venn NC, Emerenciano M, Pombo-de-Oliveira MS, Barbieri Blunck C, Almeida Lopes B, Zuna J, Trka J, Ballerini P, Lapillonne H, De Braekeleer M, Cazzaniga G, Corral Abascal L, van der Velden VHJ et al. (2017) The MLL recombinome of acute leukemias in 2017. *Leukemia* 32: 273-284
- Meyer C, Hofmann J, Burmeister T, Gröger D, Park TS, Emerenciano M, Pombo de Oliveira M, Renneville A, Villarese P, Macintyre E, Cavé H, Clappier E, Mass-Malo K, Zuna J, Trka J, De Braekeleer E, De Braekeleer M, Oh SH, Tsaur G, Fechina L et al. (2013) The MLL recombinome of acute leukemias in 2013. *Leukemia* 27: 2165-2176
- Meyer C, Schneider B, Jakob S, Strehl S, Attarbaschi A, Schnittger S, Schoch C, Jansen MW, van Dongen JJ, den Boer ML, Pieters R, Ennas MG, Angelucci E, Koehl U, Greil J, Griesinger F, Zur Stadt U, Eckert C, Szczepanski T, Niggli FK et al. (2006) The MLL recombinome of acute leukemias. *Leukemia* 20: 777-84
- Miller PG, Al-Shahrour F, Hartwell KA, Chu LP, Jaras M, Puram RV, Puissant A, Callahan KP, Ashton J, McConkey ME, Poveromo LP, Cowley GS, Kharas MG, Labelle M, Shterental S, Fujisaki J, Silberstein L, Alexe G, Al-Hajj MA, Shelton CA et al. (2013) In Vivo RNAi screening identifies a leukemia-specific dependence on integrin beta 3 signaling. *Cancer Cell* 24: 45-58
- Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A (2001) COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* 98: 12902-7
- Milne TA (2017) Mouse models of MLL leukemia: recapitulating the human disease. *Blood* 129: 2217-2223
- Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* 10: 1107-17
- Milne TA, Martin ME, Brock HW, Slany RK, Hess JL (2005) Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res* 65: 11367-74
- Min J, Feng Q, Li Z, Zhang Y, Xu RM (2003) Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* 112: 711-23

- Min YH, Cheong JW, Kim JY, Eom JI, Lee ST, Hahn JS, Ko YW, Lee MH (2004) Cytoplasmic mislocalization of p27Kip1 protein is associated with constitutive phosphorylation of Akt or protein kinase B and poor prognosis in acute myelogenous leukemia. *Cancer Res* 64: 5225-31
- Min YH, Eom JI, Cheong JW, Maeng HO, Kim JY, Jeung HK, Lee ST, Lee MH, Hahn JS, Ko YW (2003) Constitutive phosphorylation of Akt/PKB protein in acute myeloid leukemia: its significance as a prognostic variable. *Leukemia* 17: 995-7
- Mirro J, Zipf TF, Pui CH, Kitchingman G, Williams D, Melvin S, Murphy SB, Stass S (1985) Acute mixed lineage leukemia: clinicopathologic correlations and prognostic significance. *Blood* 66: 1115-23
- Mitelman F JBaMF (2018) Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer.
- Mitelman F, Johansson B, Mertens F (2007) The impact of translocations and gene fusions on cancer causation. *Nature Reviews Cancer* 7: 233-245
- Mizukawa B, Wei J, Shrestha M, Wunderlich M, Chou FS, Griesinger A, Harris CE, Kumar AR, Zheng Y, Williams DA, Mulloy JC (2011) Inhibition of Rac GTPase signaling and downstream prosurvival Bcl-2 proteins as combination targeted therapy in MLL-AF9 leukemia. *Blood* 118: 5235-45
- Moffitt AB, Ondrejka SL, McKinney M, Rempel RE, Goodlad JR, Teh CH, Leppa S, Mannisto S, Kovanen PE, Tse E, Au-Yeung RKH, Kwong YL, Srivastava G, Iqbal J, Yu J, Naresh K, Villa D, Gascoyne RD, Said J, Czader MB et al. (2017) Enteropathy-associated T cell lymphoma subtypes are characterized by loss of function of SETD2. *J Exp Med* 214: 1371-1386
- Mohan M, Herz HM, Smith ER, Zhang Y, Jackson J, Washburn MP, Florens L, Eissenberg JC, Shilatifard A (2011) The COMPASS family of H3K4 methylases in Drosophila. *Mol Cell Biol* 31: 4310-8
- Mohan M, Lin C, Guest E, Shilatifard A (2010) Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. *Nat Rev Cancer* 10: 721-8
- Mrozek K, Heinonen K, Lawrence D, Carroll AJ, Koduru PR, Rao KW, Strout MP, Hutchison RE, Moore JO, Mayer RJ, Schiffer CA, Bloomfield CD (1997) Adult patients with de novo acute myeloid leukemia and t(9; 11)(p22; q23) have a superior outcome to patients with other translocations involving band 11q23: a cancer and leukemia group B study. *Blood* 90: 4532-8
- Mueller D, Bach C, Zeisig D, Garcia-Cuellar MP, Monroe S, Sreekumar A, Zhou R, Nesvizhskii A, Chinnaiyan A, Hess JL, Slany RK (2007) A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* 110: 4445-54
- Muller-Tidow C, Steffen B, Cauvet T, Tickenbrock L, Ji P, Diederichs S, Sargin B, Kohler G, Stelljes M, Puccetti E, Ruthardt M, deVos S, Hiebert SW, Koeffler HP, Berdel WE, Serve H (2004) Translocation products in acute myeloid leukemia activate the Wnt signaling pathway in hematopoietic cells. *Mol Cell Biol* 24: 2890-904

- Muntean AG, Hess JL (2012) The pathogenesis of mixed-lineage leukemia. *Annual Review of Pathology: Mechanisms of Disease* 7: 283-301
- Muntean AG, Tan J, Sitwala K, Huang Y, Bronstein J, Connelly JA, Basrur V, Elenitoba-Johnson KS, Hess JL (2010) The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell* 17: 609-21
- Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K (2007) RNA polymerase is poised for activation across the genome. *Nat Genet* 39: 1507-11
- Musselman CA, Avvakumov N, Watanabe R, Abraham CG, Lalonde ME, Hong Z, Allen C, Roy S, Nunez JK, Nickoloff J, Kulesza CA, Yasui A, Cote J, Kutateladze TG (2012) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat Struct Mol Biol* 19: 1266-72
- Nagy PL, Griesenbeck J, Kornberg RD, Cleary ML (2002) A trithorax-group complex purified from *Saccharomyces cerevisiae* is required for methylation of histone H3. *Proc Natl Acad Sci U S A* 99: 90-4
- Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell* 10: 1119-28
- Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *The Plant cell* 2: 279-289
- Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, Maldotti M, Anselmi F, Oliviero S (2017) Intragenic DNA methylation prevents spurious transcription initiation. *Nature* 543: 72-77
- Nobel Media AB (2014) www.nobelprize.org
- Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y (2005) hDOT1L links histone methylation to leukemogenesis. *Cell* 121: 167-78
- Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127: 635-48
- Omenn GS, Lane L, Lundberg EK, Overall CM, Deutsch EW (2017) Progress on the HUPO draft human proteome: 2017 metrics of the human proteome project. *J Proteome Res* 16: 4281-4287
- Paggetti J, Largeot A, Aucagne R, Jacquelin A, Lagrange B, Yang XJ, Solary E, Bastie JN, Delva L (2010) Crosstalk between leukemia-associated proteins MOZ and MLL regulates HOX gene expression in human cord blood CD34+ cells. *Oncogene* 29: 5019-31
- Pal S, Sif S (2007) Interplay between chromatin remodelers and protein arginine methyltransferases. *Journal of cellular physiology* 213: 306-15
- Palca J (1988) James Watson to head NIH human genome project. *Nature* 335: 193

- Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, Potter NE, Heuser M, Thol F, Bolli N, Gundem G, Van Loo P, Martincorena I, Ganly P, Mudie L, McLaren S, O'Meara S, Raine K, Jones DR, Teague JW et al. (2016) Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med* 374: 2209-2221
- Park IY, Powell RT, Tripathi DN, Dere R, Ho TH, Blasius TL, Chiang YC, Davis IJ, Fahey CC, Hacker KE, Verhey KJ, Bedford MT, Jonasch E, Rathmell WK, Walker CL (2016) Dual chromatin and cytoskeletal remodeling by SETD2. *Cell* 166: 950-962
- Parker H, Rose-Zerilli MJ, Larrayoz M, Clifford R, Edelmann J, Blakemore S, Gibson J, Wang J, Ljungström V, Wojdacz TK, Chaplin T, Roghanian A, Davis Z, Parker A, Tausch E, Ntoufa S, Ramos S, Robbe P, Alsolami R, Steele AJ et al. (2016) Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. *Leukemia* 30: 2179-2186
- Patel A, Dharmarajan V, Vought VE, Cosgrove MS (2009) On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *The Journal of biological chemistry* 284: 24242-56
- Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23: 297-305
- Pfister SX, Ahrabi S, Zalmas LP, Sarkar S, Aymard F, Bachrati CZ, Helleday T, Legube G, La Thangue NB, Porter AC, Humphrey TC (2014) SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep* 7: 2006-18
- Pfister SX, Markkanen E, Jiang Y, Sarkar S, Woodcock M, Orlando G, Mavrommati I, Pai CC, Zalmas LP, Drobinitzky N, Dianov GL, Verrill C, Macaulay VM, Ying S, La Thangue NB, D'Angiolella V, Ryan AJ, Humphrey TC (2015) Inhibiting WEE1 selectively kills histone H3K36me3-deficient cancers by dNTP starvation. *Cancer Cell* 28: 557-568
- Placke T, Faber K, Nonami A, Putwain SL, Salih HR, Heidel FH, Kramer A, Root DE, Barbie DA, Krivtsov AV, Armstrong SA, Hahn WC, Huntly BJ, Sykes SM, Milsom MD, Scholl C, Frohling S (2014) Requirement for CDK6 in MLL-rearranged acute myeloid leukemia. *Blood* 124: 13-23
- Poppe B, Vandesompele J, Schoch C, Lindvall C, Mrozek K, Bloomfield CD, Beverloo HB, Michaux L, Dastugue N, Herens C, Yigit N, De Paepe A, Hagemeijer A, Speleman F (2004) Expression analyses identify MLL as a prominent target of 11q23 amplification and support an etiologic role for MLL gain of function in myeloid malignancies. *Blood* 103: 229-35
- Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA (2012) Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS genetics* 8(5): e1002717
- Pui CH, Carroll WL, Meshinchi S, Arceci RJ (2011) Biology, risk stratification, and therapy of pediatric acute leukemias: an update. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 29: 551-65
- Pui CH, Chessells JM, Camitta B, Baruchel A, Biondi A, Boyett JM, Carroll A, Eden OB, Evans WE, Gadner H, Harbott J, Harms DO, Harrison CJ, Harrison PL, Heerema N, Janka-Schaub

- G, Kamps W, Masera G, Pullen J, Raimondi SC et al. (2003) Clinical heterogeneity in childhood acute lymphoblastic leukemia with 11q23 rearrangements. *Leukemia* 17: 700-6
- Pui CH, Gaynon PS, Boyett JM, Chessells JM, Baruchel A, Kamps W, Silverman LB, Biondi A, Harms DO, Vilmer E, Schrappe M, Camitta B (2002) Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet (London, England)* 359: 1909-15
- Pui CH, Relling MV (2000) Topoisomerase II inhibitor-related acute myeloid leukaemia. *Br J Haematol* 109: 13-23
- Rao RC, Dou Y (2015) Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* 15: 334-46
- Rathert P, Roth M, Neumann T, Muerdter F, Roe JS, Muhar M, Deswal S, Cerny-Reiterer S, Peter B, Jude J, Hoffmann T, Boryn LM, Axelsson E, Schweifer N, Tontsch-Grunt U, Dow LE, Gianni D, Pearson M, Valent P, Stark A et al. (2015) Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* 525: 543-547
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593-9
- Reichel M, Gillert E, Angermuller S, Hensel JP, Heidel F, Lode M, Leis T, Biondi A, Haas OA, Strehl S, Panzer-Grumayer ER, Griesinger F, Beck JD, Greil J, Fey GH, Uckun FM, Marschalek R (2001) Biased distribution of chromosomal breakpoints involving the MLL gene in infants versus children and adults with t(4;11) ALL. *Oncogene* 20: 2900-7
- Reimer J, Knoss S, Labuhn M, Charpentier EM, Gohring G, Schlegelberger B, Klusmann JH, Heckl D (2017) CRISPR-Cas9-induced t(11;19)/MLL-ENL translocations initiate leukemia in human hematopoietic progenitor cells in vivo. *Haematologica* 102: 1558-1566
- Reya T, Duncan AW, Ailles L, Domen J, Scherer DC, Willert K, Hintz L, Nusse R, Weissman IL (2003) A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* 423: 409-14
- Richardson C, Jasin M (2000) Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* 405: 697-700
- Rickels R, Hu D, Collings CK, Woodfin AR, Piunti A, Mohan M, Herz HM, Kvon E, Shilatfard A (2016) An evolutionary conserved epigenetic mark of polycomb response elements implemented by Trx/MLL/COMPASS. *Mol Cell* 63: 318-328
- Risner LE, Kuntimaddi A, Lokken AA, Achille NJ, Birch NW, Schoenfelt K, Bushweller JH, Zeleznik-Le NJ (2013) Functional specificity of CpG DNA-binding CXXC domains in mixed lineage leukemia. *The Journal of biological chemistry* 288: 29901-10
- Roe JS, Mercan F, Rivera K, Pappin DJ, Vakoc CR (2015) BET Bromodomain inhibition suppresses the function of hematopoietic transcription factors in acute myeloid leukemia. *Mol Cell* 58: 1028-39

- Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, Stewart AF (2001) The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *Embo j* 20: 7137-48
- Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics : MCP* 3: 1154-69
- Rowley JD (1973) Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 243: 290-3
- Rowley JD, Reshmi S, Sobulo O, Musvee T, Anastasi J, Raimondi S, Schneider NR, Barredo JC, Cantu ES, Schlegelberger B, Behm F, Doggett NA, Borrow J, Zeleznik-Le N (1997) All patients with the T(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders. *Blood* 90: 535-41
- Rubnitz JE, Raimondi SC, Tong X, Srivastava DK, Razzouk BI, Shurtleff SA, Downing JR, Pui CH, Ribeiro RC, Behm FG (2002) Favorable impact of the t(9;11) in childhood acute myeloid leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 20: 2302-9
- Sandhöfer N, Metzeler KH, Rothenberg M, Herold T, Tiedt S, Groiß V, Carlet M, Walter G, Hinrichsen T, Wachter O, Grunert M, Schneider S, Subklewe M, Dufour A, Fröhling S, Klein HG, Hiddemann W, Jeremias I, Spiekermann K (2014) Dual PI3K/mTOR inhibition shows antileukemic activity in MLL-rearranged acute myeloid leukemia. *Leukemia* 29: 828-838
- Sanjana NE, Shalem O, Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nature methods* 11: 783-784
- Sarvan S, Avdic V, Tremblay V, Chaturvedi CP, Zhang P, Lanouette S, Blais A, Brunzelle JS, Brand M, Couture JF (2011) Crystal structure of the trithorax group protein ASH2L reveals a forkhead-like DNA binding domain. *Nat Struct Mol Biol* 18: 857-9
- Scacheri PC, Davis S, Odom DT, Crawford GE, Perkins S, Halawi MJ, Agarwal SK, Marx SJ, Spiegel AM, Meltzer PS, Collins FS (2006) Genome-wide analysis of menin binding provides insights into MEN1 tumorigenesis. *PLoS genetics* 2: e51
- Schichman SA, Caligiuri MA, Gu Y, Strout MP, Canaani E, Bloomfield CD, Croce CM (1994) ALL-1 partial duplication in acute leukemia. *Proc Natl Acad Sci U S A* 91: 6236-9
- Schulze JM, Wang AY, Kobor MS (2009) YEATS domain proteins: a diverse family with many links to chromatin modification and transcription. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 87: 65-75
- Schwalb B, Michel M, Zacher B, Fruhauf K, Demel C, Tresch A, Gagneur J, Cramer P (2016) TT-seq maps the human transient transcriptome. *Science* 352: 1225-8
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343: 84-87

- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13: 2498-504
- Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR (2015) Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat Biotechnol* 33: 661-7
- Shilatifard A (2012) The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual review of biochemistry* 81: 65-95
- Shilatifard A, Lane WS, Jackson KW, Conaway RC, Conaway JW (1996) An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* 271: 1873-6
- Sierra J, Yoshida T, Joazeiro CA, Jones KA (2006) The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. *Genes Dev* 20: 586-600
- Slany RK (2009) The molecular biology of mixed lineage leukemia. *Haematologica* 94: 984-93
- Slany RK (2016) The molecular mechanics of mixed lineage leukemia. *Oncogene* 35: 5215-5223
- So CW, Cleary ML (2003) Common mechanism for oncogenic activation of MLL by forkhead family proteins. *Blood* 101: 633-9
- So CW, Karsunky H, Passegue E, Cozzio A, Weissman IL, Cleary ML (2003a) MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. *Cancer Cell* 3: 161-71
- So CW, Lin M, Ayton PM, Chen EH, Cleary ML (2003b) Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. *Cancer Cell* 4: 99-110
- Sobulo OM, Borrow J, Tomek R, Reshmi S, Harden A, Schlegelberger B, Housman D, Doggett NA, Rowley JD, Zeleznik-Le NJ (1997) MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). *Proc Natl Acad Sci U S A* 94: 8732-7
- Spurr SS, Bayle ED, Yu W, Li F, Tempel W, Vedadi M, Schapira M, Fish PV (2016) New small molecule inhibitors of histone methyl transferase DOT1L with a nitrile as a non-traditional replacement for heavy halogen atoms. *Bioorganic & medicinal chemistry letters* 26: 4518-4522
- Sroczynska P, Cruickshank VA, Bukowski JP, Miyagi S, Bagger FO, Walfridsson J, Schuster MB, Porse B, Helin K (2014) shRNA screening identifies JMJD1C as being required for leukemia maintenance. *Blood* 123: 1870-82
- Stam RW, den Boer ML, Schneider P, Nollau P, Horstmann M, Beverloo HB, van der Voort E, Valsecchi MG, de Lorenzo P, Sallan SE, Armstrong SA, Pieters R (2005) Targeting FLT3 in primary MLL-gene-rearranged infant acute lymphoblastic leukemia. *Blood* 106: 2484-90
- Stam RW, Schneider P, Hagelstein JA, van der Linden MH, Stumpel DJ, de Menezes RX, de Lorenzo P, Valsecchi MG, Pieters R (2010) Gene expression profiling-based dissection of MLL translocated and MLL germline acute lymphoblastic leukemia in infants. *Blood* 115: 2835-44

- Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, Vakoc AL, Kim JE, Chen J, Lazar MA, Blobel GA, Vakoc CR (2008) DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol Cell Biol* 28: 2825-39
- Steliarova-Foucher E, Colombet M, Ries LAG, Moreno F, Dolya A, Bray F, Hesselting P, Shin HY, Stiller CA (2017) International incidence of childhood cancer, 2001-10: a population-based registry study. *The Lancet Oncology* 18: 719-731
- Stone RM, Larson RA, Dohner H (2017) Midostaurin in FLT3-Mutated Acute Myeloid Leukemia. *N Engl J Med* 377: 1903
- Stubbs MC, Kim YM, Krivtsov AV, Wright RD, Feng Z, Agarwal J, Kung AL, Armstrong SA (2008) MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia: development of a model for rapid therapeutic assessment. *Leukemia* 22: 66-77
- Sykes SM, Lane SW, Bullinger L, Kalaitzidis D, Yusuf R, Saez B, Ferraro F, Mercier F, Singh H, Brumme KM, Acharya SS, Scholl C, Tothova Z, Attar EC, Frohling S, DePinho RA, Armstrong SA, Gilliland DG, Scadden DT (2011) AKT/FOXO signaling enforces reversible differentiation blockade in myeloid leukemias. *Cell* 146: 697-708
- Takeda S, Chen DY, Westergard TD, Fisher JK, Rubens JA, Sasagawa S, Kan JT, Korsmeyer SJ, Cheng EH, Hsieh JJ (2006) Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes Dev* 20: 2397-409
- Taki T, Sako M, Tsuchida M, Hayashi Y (1997) The t(11;16)(q23;p13) translocation in myelodysplastic syndrome fuses the MLL gene to the CBP gene. *Blood* 89: 3945-50
- Tenen DG (2003) Disruption of differentiation in human cancer: AML shows the way. *Nat Rev Cancer* 3: 89-101
- Thingholm TE, Larsen MR (2016) The use of titanium dioxide for selective enrichment of phosphorylated peptides. *Methods Mol Biol* 1355: 135-46
- Thirman MJ, Gill HJ, Burnett RC, Mbangkollo D, McCabe NR, Kobayashi H, Ziemann-van der Poel S, Kaneko Y, Morgan R, Sandberg AA, et al. (1993) Rearrangement of the MLL gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations. *N Engl J Med* 329: 909-14
- Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry* 75: 1895-904
- Tkachuk DC, Kohler S, Cleary ML (1992) Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. *Cell* 71: 691-700
- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians* 65: 87-108

- Trentin L, Giordan M, Dingermann T, Basso G, Te Kronnie G, Marschalek R (2009) Two independent gene signatures in pediatric t(4;11) acute lymphoblastic leukemia patients. *Eur J Haematol* 83: 406-19
- Tyagi S, Chabes AL, Wysocka J, Herr W (2007) E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Mol Cell* 27: 107-19
- Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey OM, Mupo A, Grinkevich V, Li M, Mazan M, Gozdecka M, Ohnishi S, Cooper J, Patel M, McKerrell T, Chen B, Domingues AF, Gallipoli P, Teichmann S, Ponstingl H et al. (2016) A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep* 17: 1193-1205
- Ullah M, Pelletier N, Xiao L, Zhao SP, Wang K, Degerny C, Tahmasebi S, Cayrou C, Doyon Y, Goh SL, Champagne N, Cote J, Yang XJ (2008) Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes. *Mol Cell Biol* 28: 6828-43
- van Amerongen R, Nawijn MC, Lambooj JP, Proost N, Jonkers J, Berns A (2010) Frat oncoproteins act at the crossroad of canonical and noncanonical Wnt-signaling pathways. *Oncogene* 29: 93-104
- Van den Berghe H, David G, Broeckeaert-Van Orshoven A, Louwagie A, Verwilghen R, Casteels-Van Daele M, Eggermont E, Eeckels R (1979) A new chromosome anomaly in acute lymphoblastic leukemia (ALL). *Human genetics* 46: 173-80
- van Leeuwen F, Gafken PR, Gottschling DE (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* 109: 745-56
- van Nuland R, Smits AH, Pallaki P, Jansen PW, Vermeulen M, Timmers HT (2013) Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol Cell Biol* 33: 2067-77
- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142: 967-80
- Vezzoli A, Bonadies N, Allen MD, Freund SM, Santiveri CM, Kvinlaug BT, Huntly BJ, Gottgens B, Bycroft M (2010) Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. *Nat Struct Mol Biol* 17: 617-9
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW (2013) Cancer genome landscapes. *Science* 339: 1546-58
- Walf-Vorderwulbecke V, de Boer J, Horton SJ, van Amerongen R, Proost N, Berns A, Williams O (2012) Frat2 mediates the oncogenic activation of Rac by MLL fusions. *Blood* 120: 4819-28
- Wang E, Kawaoka S, Yu M, Shi J, Ni T, Yang W, Zhu J, Roeder RG, Vakoc CR (2013) Histone H2B ubiquitin ligase RNF20 is required for MLL-rearranged leukemia. *Proc Natl Acad Sci U S A* 110: 3901-6

- Wang J, Iwasaki H, Krivtsov A, Febbo PG, Thorner AR, Ernst P, Anastasiadou E, Kutok JL, Kogan SC, Zinkel SS, Fisher JK, Hess JL, Golub TR, Armstrong SA, Akashi K, Korsmeyer SJ (2005) Conditional MLL-CBP targets GMP and models therapy-related myeloproliferative disease. *Embo j* 24: 368-81
- Wang J, Muntean AG, Hess JL (2012) ECSASB2 mediates MLL degradation during hematopoietic differentiation. *Blood* 119: 1151-61
- Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM (2015) Identification and characterization of essential genes in the human genome. *Science* 350: 1096-101
- Wang T, Yu H, Hughes NW, Liu B, Kendirli A, Klein K, Chen WW, Lander ES, Sabatini DM (2017) Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. *Cell* 168: 890-903.e15
- Wang Y, Krivtsov AV, Sinha AU, North TE, Goessling W, Feng Z, Zon LI, Armstrong SA (2010) The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science* 327: 1650-3
- Wang Z, Smith KS, Murphy M, Piloto O, Somervaille TC, Cleary ML (2008) Glycogen synthase kinase 3 in MLL leukaemia maintenance and targeted therapy. *Nature* 455: 1205-9
- Wang Z, Song J, Milne TA, Wang GG, Li H, Allis CD, Patel DJ (2010) Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell* 141: 1183-94
- Wen H, Li Y, Xi Y, Jiang S, Stratton S, Peng D, Tanaka K, Ren Y, Xia Z, Wu J, Li B, Barton MC, Li W, Li H, Shi X (2014) ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature* 508: 263-8
- Wiersma M, Bussiere M, Halsall JA, Turan N, Slany R, Turner BM, Nightingale KP (2016) Protein kinase Msk1 physically and functionally interacts with the KMT2A/MLL1 methyltransferase complex and contributes to the regulation of multiple target genes. *Epigenetics & chromatin* 9: 52
- Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmaier A et al. (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509: 582-7
- Wilkinson AC, Ballabio E, Geng H, North P, Tapia M, Kerry J, Biswas D, Roeder RG, Allis CD, Melnick A, de Bruijn MF, Milne TA (2013) RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. *Cell Rep* 3: 116-27
- Winters AC, Bernt KM (2017) MLL-rearranged leukemias - an update on science and clinical approaches. *Frontiers in Pediatrics* 5:4
- Wu M, Wang PF, Lee JS, Martin-Brown S, Florens L, Washburn M, Shilatifard A (2008) Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol Cell Biol* 28: 7337-44

- Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121: 859-72
- Xia ZB, Anderson M, Diaz MO, Zeleznik-Le NJ (2003) MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proc Natl Acad Sci U S A* 100: 8342-7
- Xia ZB, Popovic R, Chen J, Theisler C, Stuart T, Santillan DA, Erfurth F, Diaz MO, Zeleznik-Le NJ (2005) The MLL fusion gene, MLL-AF4, regulates cyclin-dependent kinase inhibitor CDKN1B (p27kip1) expression. *Proc Natl Acad Sci U S A* 102: 14028-33
- Xu Q, Simpson SE, Scialla TJ, Bagg A, Carroll M (2003) Survival of acute myeloid leukemia cells requires PI3 kinase activation. *Blood* 102: 972-80
- Yagi H, Deguchi K, Aono A, Tani Y, Kishimoto T, Komori T (1998) Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* 92: 108-17
- Yang Z, Yik JH, Chen R, He N, Jang MK, Ozato K, Zhou Q (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* 19: 535-45
- Yeung J, Esposito MT, Gandillet A, Zeisig BB, Griessinger E, Bonnet D, So CW (2010) beta-Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. *Cancer Cell* 18: 606-18
- Yokoyama A (2002) Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. *Blood* 100: 3710-3718
- Yokoyama A, Cleary ML (2008) Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* 14: 36-46
- Yokoyama A, Somervaille TC, Smith KS, Rozenblatt-Rosen O, Meyerson M, Cleary ML (2005) The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* 123: 207-18
- Yokoyama A, Wang Z, Wysocka J, Sanyal M, Aufiero DJ, Kitabayashi I, Herr W, Cleary ML (2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol Cell Biol* 24: 5639-49
- Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995) Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* 378: 505-8
- Yuan H, Li N, Fu D, Ren J, Hui J, Peng J, Liu Y, Qiu T, Jiang M, Pan Q, Han Y, Wang X, Li Q, Qin J (2017) Histone methyltransferase SETD2 modulates alternative splicing to inhibit intestinal tumorigenesis. *The Journal of clinical investigation* 127: 3375-3391
- Zeisig BB, Milne T, Garcia-Cuellar MP, Schreiner S, Martin ME, Fuchs U, Borkhardt A, Chanda SK, Walker J, Soden R, Hess JL, Slany RK (2004) Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol Cell Biol* 24: 617-28

- Zelevnik-Le NJ, Harden AM, Rowley JD (1994) 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proc Natl Acad Sci U S A* 91: 10610-4
- Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, Easton J, Chen X, Wang J, Rusch M, Lu C, Chen SC, Wei L, Collins-Underwood JR, Ma J, Roberts KG, Pounds SB, Ulyanov A, Becksfort J, Gupta P et al. (2012) The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 481: 157-63
- Zhang P, Du J, Sun B, Dong X, Xu G, Zhou J, Huang Q, Liu Q, Hao Q, Ding J (2006) Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Res* 34: 6621-8
- Zhang YL, Sun JW, Xie YY, Zhou Y, Liu P, Song JC, Xu CH, Wang L, Liu D, Xu AN, Chen Z, Chen SJ, Sun XJ, Huang QH (2018) Setd2 deficiency impairs hematopoietic stem cell self-renewal and causes malignant transformation. *Cell research* 28, pages476-90
- Zhao D, Li Y, Xiong X, Chen Z, Li H (2017) YEATS domain-a histone acylation reader in health and disease. *Journal of molecular biology* 429: 1994-2002
- Zheng W, Ibanez G, Wu H, Blum G, Zeng H, Dong A, Li F, Hajian T, Allali-Hassani A, Amaya MF, Sjarheyeva A, Yu W, Brown PJ, Schapira M, Vedadi M, Min J, Luo M (2012) Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2. *Journal of the American Chemical Society* 134: 18004-14
- Zhou Y, Yan X, Feng X, Bu J, Dong Y, Lin P, Hayashi Y, Huang R, Olsson A, Andreassen PR, Grimes HL, Wang QF, Cheng T, Xiao Z, Jin J, Huang G (2018) Setd2 regulates quiescence and differentiation of adult hematopoietic stem cells by restricting RNA polymerase II elongation. *Haematologica* 2018.187708
- Zhu L, Li Q, Wong SH, Huang M, Klein BJ, Shen J, Ikenouye L, Onishi M, Schneidawind D, Buechele C, Hansen L, Duque-Afonso J, Zhu F, Martin GM, Gozani O, Majeti R, Kutateladze TG, Cleary ML (2016) ASH1L links histone H3 lysine 36 dimethylation to MLL leukemia. *Cancer Discov* 6: 770-83
- Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, Yan X, Wei W, Pang Y, Cheng H, Hua C, Zhang Y, Yang X, Lu X, Cao L, Hao L, Dong L, Zou W, Wu J, Li X et al. (2014) Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat Genet* 46: 287-93
- Zuber J, McJunkin K, Fellmann C, Dow LE, Taylor MJ, Hannon GJ, Lowe SW (2011a) Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat Biotechnol* 29: 79-83
- Zuber J, Radtke I, Pardee TS, Zhao Z, Rappaport AR, Luo W, McCurrach ME, Yang MM, Dolan ME, Kogan SC, Downing JR, Lowe SW (2009) Mouse models of human AML accurately predict chemotherapy response. *Genes Dev* 23: 877-89
- Zuber J, Rappaport AR, Luo W, Wang E, Chen C, Vaseva AV, Shi J, Weissmueller S, Fellman C, Taylor MJ, Weissenboeck M, Graeber TG, Kogan SC, Vakoc CR, Lowe SW (2011b) An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-

renewal program as essential for leukemia maintenance. *Genes & Development* 25: 1628-1640

Zuber J, Shi J, Wang E, Rappaport AR, Herrmann H, Sison EA, Magoon D, Qi J, Blatt K, Wunderlich M, Taylor MJ, Johns C, Chicas A, Mulloy JC, Kogan SC, Brown P, Valent P, Bradner JE, Lowe SW, Vakoc CR (2011c) RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* 478: 524-8

5. Annex

5.1 Curriculum Vitae

Anna Skucha

Email Address: an.skucha@gmail.com **Contact Number:** +43 681 20132847
Date of Birth: September 18th, 1987 **Nationality:** Polish **Residency:** Vienna, Austria

Research experience

PhD Student

08/2012 present

Studies conducted under the shared supervision of
Prof. Dr. Giulio Superti-Furga, CeMM Research Center for Molecular Medicine of the OeAW
(cemm.at)

Dr. Florian Grebien, LBI-CR Ludwig Boltzmann Institute for Cancer Research (lbicr.lbg.ac.at),
including:

- Advanced molecular biology methods
- Involvement in an international consortium: regular reporting, project management, organisation of international meetings, active science communication during national and international conferences, international collaborations
- Supervision and mentoring of master student and trainees
- Contribution to scientific publications and grant applications

Main focus of the project: Cancer Research, Hematology, Epigenetics

Project supported by the highly competitive Maria Curie Action fellowship (EU FP7)

Part of European consortium Hem_ID (HEMatopoietic cell IDentity: genetic and epigenetic regulation in normal and malignant hematopoiesis)

Visiting Research Scholar (as part of the Maria Curie Action fellowship)

02/2015

The Biological Research Center (Centro de Investigaciones Biológicas), Madrid, Spain

Dr. Miguel Vidal's laboratory

Scope: *in vivo* models of human leukemia

Visiting Research Scholar (as part of the PhD training)

08/2013 - 02/2014

Research Institute of Molecular Pathology (IMP), Vienna, Austria; Dr. Johannes Zuber's laboratory

Scope: Setting up a NGS-based shRNA screening project

Research Technician

02/2012 - 08/2012

MFPL Vienna, Austria, Dr. Silke Dorner's group.

Scope: contribution to the finalization of a manuscript; advanced molecular biology methods, cell culture of insect cells, affinity purification, mentoring of students and trainees, writing activities: Standard Operating Procedures (SOPs)

Research Internship

08/2011 - 09/2011

MFPL Vienna, Austria; Dr. Silke Dorner's laboratory

Project: Co-translational mRNA degradation in *Drosophila* cells

Scope: basic molecular biology methods, cell culture of insect cells, cell fractionation, affinity purification

Summer Undergraduate Research Fellowship (SURF BioNET)

06/2010 - 09/2010

Part of UREKA Scholarship National University of Ireland, Galway, Ireland

Dr. Michael P. Carty's laboratory

Project: Investigation of DNA damage responses in human skin cells exposed to UVA radiation

Scope: advanced methods of molecular biology, DNA damage induction assays, cell culture of human primary fibroblasts, scientific writing (reports, press releases), oral and poster presentations

Research Internship

07/2008 - 09/2008

Cracow's Regional Blood Center, Cracow, Poland

Scope: Training in different departments of the Institution: patient qualification, blood testing, blood separation and processing, writing activities (reports)

Education

PhD in Molecular Medicine

08/2012 - Present

Center for Molecular Medicine of the Austrian Academy of Sciences (CeMM), Vienna, Austria

Prof. Dr. Giulio Superti-Furga's laboratory

Project: MLL fusion-driven leukemia requires SETD2 to safeguard genomic integrity

Master of Science in Biotechnology

10/2006 - 06/2011

Jagiellonian University, Cracow, Poland

Project: Regulation of antichymotrypsin transcript's half-life and ZFP36 gene expression

Skills and Certificates

- **Project Management:** Prince2® 2017 Foundation and Practitioner certified; issued by Axelos, UK
- Certified in **Good Laboratory and Manufacturing Practise**; issued by Whitehall Training, UK
- **Medicinal Chemistry:** the Molecular basis of drug discovery; issued by Davidson Collage, USA in collaboration with Novartis Institutes for Biomedical Research (NIBR)
- **Advanced laboratory skills:** advanced skills in molecular biology, flow cytometry, shRNA and CRISPR/Cas9 mediated mutagenesis, cell culture (cell lines and primary cells), Next Generation Sequencing, basic skills in animal work and mass spectrometry
- **Computer skills:** MS Office, Image processing tools, laboratory equipment software
- **Language skills:** Polish (native proficiency), English (full professional proficiency), German (professional working proficiency, B2 certificate), Italian (elementary proficiency)
- **Science communication:** strong publication record (annex I); 7 oral and 9 poster presentations at national and international conferences (annex II and III)

- **Networking skills:** PhD in a close collaboration of three laboratories in three top research institutes in Vienna (CeMM/LBI-CR/IMP); contributed to several national and international collaborations; organisation of several international meetings; supervised master student and trainees (annex IV);

Honours and awards

- Selected within the top 30 PhD candidates of Medical University of Vienna for the **NaturTalente** program promoted by key international pharma and biotech companies.
- PhD project supported by the competitive prestigious **FP7 Marie Curie Program** (HEM_ID ITN)

References Available upon request

List of oral presentations at international conferences

- EHA European Hematology Association meeting; Madrid, Spain; June 2017
- Epigenetic regulation of hematopoiesis and leukemogenesis meeting; Strasbourg, France; April 2015
- Invited Speaker: Centro de Investigaciones Biológicas; Madrid, Spain; February 2015
- Young Scientist Association Symposium; Vienna, Austria; June 2014
- Scientific Writing Workshop; Kos, Greece; May 2014
- Diagnostics of Childhood Acute Lymphoblastic Leukemia; Milan, Italy; September 2013
- Applications of proteomics in cell biology; Vienna, Austria; March 2013

List of poster presentations at international conferences

- Ludwig Boltzmann Gesellschaft Meeting for Health Sciences; Vienna, Austria; November 2016
- EHA European Hematology Association meeting; Copenhagen, Denmark; June 2016
- International Society for Experimental Hematology; Kyoto, Japan; September 2015
- EHA European Hematology Association meeting; Vienna, Austria; June 2015
- ISEH International Society for Experimental Hematology; Montreal, Canada; August 2014
- Interdisciplinary Signaling Workshop; Visegrad, Hungary; July 2014
- FEBS summer school, Molecular Mechanisms in Signal Transduction and Cancer; Spetses, Greece; August 2013
- Young Scientist Association Symposium; Vienna, Austria; June 2013
- PhD symposium; Cells don't play dice; Vienna, Austria; June 2013

Mentoring and supervision

- Master Student: Jessica Ebner, Ludwig Boltzmann Institute for Cancer Research; Laboratory of Dr. Florian Grebien; November 2015 - January 2017

- Trainee: Clara Gudenus, CeMM, Research Center for Molecular Medicine of the OeAW; Laboratory of Prof. Dr. Giulio Superti-Furga; August-September 2015
- Mentoring of trainees and students in the laboratory of Dr. Silke Dorner; MFPL; February – August 2012
- Mentoring of trainees in the laboratory of Dr. Aneta Kasza; UJ, Krakow, Poland; 2009-2011

Professional certificates

- **Project Management:** PRINCE2® Foundation and Practitioner Certificate; Axelos, UK
- **Good Manufacturing Practice** Certificate of Achievement
Whitehall Training, Accredited by the Faculty of Pharmaceutical Medicine of the Royal College of Physicians of the United Kingdom
- **Good Laboratory Practice** Certificate of Achievement
Whitehall Training, Accredited by the Faculty of Pharmaceutical Medicine of the Royal College of Physicians of the United Kingdom
- **German Language Certificate: ÖSD Zertifikat B2**
Österreichisches Sprachdiplom Deutsch (ÖSD), Vienna, Austria
- **Medicinal Chemistry:** the Molecular basis of drug discovery; issued by Davidson Collage, USA in collaboration with Novartis Institutes for Biomedical Research (NIBR)

5.2 Scientific Publications

- Skucha A, Ebner J, Schmöllerl J, Roth M, Eder T, César-Razquin A, Stukalov A, Vittori S, Muhar M, Lu B, Aichinger M, Jude J, Müller AC, Györfy B, Vakoc CR, Valent P, Bennett KL, Zuber J, Superti-Furga G, Grebien F. **MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity.** *Nature Commun*, in press.
- Pencik J, Schleder M, Gruber W, Unger C, Walker SM, Chalaris A, Marié IJ, Hassler MR, Javaheri T, Aksoy O, Blayney JK, Prutsch N, Skucha A, Herac M, Krämer OH, Mazal P, Grebien F, Egger G, Poli V, Mikulits W, Eferl R, Esterbauer H, Kennedy R, Fend F, Scharpf M, Braun M, Perner S, Levy DE, Malcolm T, Turner SD, Haitel A, Susani M, Moazzami A, Rose-John S, Aberger F, Merkel O, Moriggl R, Culig Z, Dolznig H, Kenner L. **STAT3 regulated ARF expression suppresses prostate cancer metastasis.** *Nature Commun*. 2015 Jul 22;6:7736.
- Grebien F, Vedadi M, Getlik M, Giambruno R, Grover A, Avellino R, Skucha A, Vittori S, Kuznetsova E, Smil D, Barsyte-Lovejoy D, Li F, Poda G, Schapira M, Wu H, Dong A, Senisterra G, Stukalov A, Huber KVM, Schönegger A, Marcellus R, Bilban M, Bock C, Brown PJ, Zuber J, Bennett KL, Al-Awar R, Delwel R, Nerlov C, Arrowsmith CH, Superti-Furga G. **Pharmacological targeting of the Wdr5-MLL interaction in C/EBPα N-terminal leukemia.** *Nature Chem Biol*. 2015 Aug;11(8):571-578.
- Antic S, Wolfinger MT, Skucha A, Hosiner S, Dorner S. **General and MicroRNA-Mediated mRNA Degradation Occurs on Ribosome Complexes in Drosophila Cells.** *Mol Cell Biol*. 2015 Jul;35(13):2309-20.
- Huber ML, Sacco R, Parapatics K, Skucha A, Khamina K, Müller AC, Rudashevskaya EL, Bennett KL. **abFASP-MS: affinity-based filter-aided sample preparation mass spectrometry for quantitative analysis of chemically labeled protein complexes.** *J Proteome Res*. 2014 Feb 7;13(2):1147-55.
- Florkowska M, Tymoszek P, Balwierz A, Skucha A, Kochan J, Wawro M, Stalinska K, Kasza A. **EGF activates TTP expression by activation of ELK-1 and EGR-1 transcription factors.** *BMC Mol Biol*. 2012 Mar 20;13:8.