# Decoding Epigenomic Evolution of Acute Myeloid Leukemia

Sabrina Andrea Weser




**Dissertation an der Fakultät für Biologie der**

**Ludwig-Maximilians-Universität München**

München 2022

1. Gutachter: Prof. Dr. Wolfgang Enard

2. Gutachter: Prof. Dr. Karsten Spiekermann

Tag der Abgabe: 11.07.22

Tag der mündlichen Prüfung: 09.10.23

# Decoding Epigenomic Evolution of Acute Myeloid Leukemia

Sabrina Andrea Weser

# Statutory Declaration and Statement
(Eidestattliche Versicherung und Erklärung)

## Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 12.10.23 .....

Sabrina Weser

## Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist und dass ich mich anderweitig einer Doktorprüfung ohne Erfolg **nicht** unterzogen habe.

München, den 12.10.23 .....

Sabrina Weser

# Table of contents

# Abbreviations

| | |
|---|---|
| 5hmC | 5-Hydroxymethylcytosine |
| 5mC | 5-Methylcytosine |
| 6-TG | Thioguanine |
| AML | Acute Myeloid Leukemia |
| APL | Promyelocytic Leukemia |
| APL | Acute Promyelocytic Leukemia |
| AraC | Cytarabine |
| ARS | Aminoacyl-tRNA synthetase |
| ATAC-Seq | Assay for Transposase-Accessible Chromatin Sequencing |
| BER | Base Excision Repair |
| cDNA | complementary DNA |
| CGI | CpG-Island |
| CHIP | Clonal Hematopoiesis of Indeterminate Potential |
| ChIP-Seq | Chromatin Immunoprecipitation Sequencing |
| CML | Chronic Myeloid Leukemia |
| CN | Cytogenetically normal |
| DE | Differential Expression |
| DNA | Deoxyribonucleic Acid |
| DNMT | DNA Methyltransferase |
| DNMT3A-Mut | DNMT3A mutated |
| DNMT3A-WT | DNMT3A wild type |
| ELN | European LeukemiaNet |
| ESC | Embryonic Stem Cell |
| ESI | Electrospray Ionization |
| FAB | French American British |
| H3K27 | Histone H3 Lysine 27 |
| HMA | Hypomethylating Agents |
| HPLC | High-performance Liquid Chromatography |
| HSC | Hematopoietic Stem Cell |
| KD | Knockdown |

| | |
|---|---|
| KO | Knockout |
| LDAC | Low Dose Cytarabine |
| LSC | Leukemia Stem Cell |
| LSD | Lysine-Specific Demethylase |
| MALDI | Matrix-Assisted Laser Desorption Ionization |
| MBD | Methyl-CpG-Binding Domain |
| MM | Multiple Myeloma |
| MRC | Medical Research Council |
| MRD | Measurable Residual Disease |
| mRNA | Messenger RNA |
| MS | Mass-Spectronomy |
| NGS | Next generation sequencing |
| OS | Overall Survival |
| PDX | Patient Derived Xenograft |
| PE | Paired-End |
| PRC2 | polycomb repressive complex 2 |
| PRMT | Protein Arginine N-Methyltransferase |
| QC | Quality Control |
| RNA | Ribonucleic Acid |
| RNA-Seq | RNA Sequencing |
| RRBS | Reduced Representation Bisulfite Sequencing |
| rRNA | ribosomal RNA |
| SBS | Sequencing By Synthesis |
| scRNA-Seq | single cell RNA-Seq |
| SCT | Stem Cell Transplantation |
| SE | Single-End |
| TF | Transcription factor |
| TOF | Time-Of-Flight |
| TSG | Tumor Suppressor Gene |
| UMI | Unique Molecular Identifier |
| WBC | White Blood Cell |
| WES | Whole Exome Sequencing |
| WGBS | Whole Genome Bisulfite Sequencing |
| WGS | Whole Genome Sequencing |

# List of publications

1. Stief SM, Hanneforth AL, **Weser S,** Mattes R, Carlet M, Liu WH, Bartoschek MD, Domínguez Moreno H, Oettle M, Kempf J, Vick B, Ksienzyk B, Tizazu B, Rothenberg-Thurley M, Quentmeier H, Hiddemann W, Vosberg S, Greif PA, Metzeler KH, Schotta G, Bultmann S, Jeremias I, Leonhardt H, Spiekermann, K. *Loss of KDM6A confers drug resistance in acute myeloid leukemia*. **Leukemia (2019**).

2. Kempf JM*, **Weser S***, Bartoschek MD, Metzeler KH, Vick B, Herold T, Völse K, Mattes R, Scholz M, Wange LE, Festini M, Ugur E, Roas M, Weigert O, Bultmann S, Leonhardt H, Schotta G, Hiddemann W, Jeremias I, Spiekermann K. *Loss-of-function mutations in the histone methyltransferase EZH2 promote chemotherapy resistance in AML*. **Scientific Reports (2021)**.

3. **Weser S**, Parekh S, Hartmann L, Spiekermann K. Disease progression in AML is driven by loss-of-function mutations in DNMT3A. **Manuscript in preparation**

# Declaration of Contribution as Co-Author

**Loss of KDM6A confers drug resistance in acute myeloid leukemia**

Sophie Stief and Karsten Spiekermann conceived the study. I designed, performed, and analyzed the RNA-Seq experiments. Anna-Li Hanneforth, Raphael Mattes, Michael D. Bartoschek, Belay Tizazu, Julia Kempf, Helena Domínguez Moreno, Michela Carlet and Wen-Hsin Liu helped in data collection and analysis. Binje Vick and Irmela Jeremias established, performed, and analyzed PDX experiments. Maja Rothenberg-Thurley, Klaus H. Metzeler, Sebastian Vosberg, and Philipp A. Greif helped with analysis. Heinrich Leonhardt and Karsten Spiekermann supervised the study. The manuscript was written by Sophie Stief.

**Disease progression in AML is driven by loss-of-function mutations in DNMT3A**

Karsten Spiekermann and I conceived the study. I analyzed and interpreted the data. Swati Parekh performed preprocessing of RNA-Seq data. Luise Hartmann designed the experiments. Karsten Spiekermann supervised the study. The manuscript was written by me.

According to the regulations for the Cumulative Doctoral Thesis at the Faculty of Biology, LMU München, I confirm the above contributions of Sabrina Weser to these publications.

| | |
|---|---|
| Prof. Dr. Karsten Spiekermann | Prof. Dr. Wolfgang Enard |

**Loss-of-function mutations in the histone methyltransferase EZH2 promote chemotherapy resistance in AML**

Karsten Spiekermann, Julia M. Kempf and I conceived the study. I designed and analyzed the proteomics and RNA-Seq experiments. Lucas E. Wange, Enes Ugur, Klaus H. Metzeler, Manuela Scholz and Tobias Herold helped with data collection and analysis. Michael D. Bartoschek, Binje Vick, Sebastian Bultmann, Raphael Mattes, Kerstin Völse, Maike Roas, Moreno Festini and Heinrich Leonhardt supported the experiments. Irmela Jeremias, Gunnar Schotta, Wolfgang Hiddemann and Oliver Weigert interpreted the data. Karsten Spiekermann supervised the study. The manuscript was written by Julia M. Kempf and me.

According to the regulations for the Cumulative Doctoral Thesis at the Faculty of Biology, LMU München, we confirm the above contributions to this publication.

.....

Sabrina Weser

.....

Julia Kempf

.....

Prof. Dr. Karsten Spiekermann

.....

Prof. Dr. Wolfgang Enard

# Aim of the Thesis

Acute myeloid leukemia (AML) is a heterogenous hematologic malignancy, characterized by very high relapse rates. Although recurrent gene mutations have been identified, they are rarely associated with relapse and relapse-specific genes have not been found. It can therefore be assumed that epigenetic regulation of gene expression has a major influence on disease progression and resistance development. The aim of this work is to study the impact of mutations and expression changes of the epigenetic modifier genes KDM6A, EZH2 and DNMT3A.

# Summary

The epigenetic landscape occupies a major role in the regulation of normal hematopoiesis. Since AML relapse was found to be largely independent of genetic factors, epigenetic regulation of gene expression is a likely cause of disease progression and chemoresistance. Studying the effect of mutations and aberrant expression of epigenetic modifiers is therefore imperative for the development of new treatment strategies and the improvement of patient survival.

In the first study of this thesis, we examined the consequence of loss-of-function in the H3K27 demethylase KDM6A. To this end, we established a lentiviral shRNA-mediated knockdown, as well as a CRISPR/Cas9-mediated knockout of KDM6A in the leukemia cell line K562. We found that cells with a reduction or total loss of KDM6A exhibited higher resistance to the chemotherapeutic agents cytarabine (AraC) and daunorubicin, but not to thioguanine (6-TG). This effect was found to be reversible by re-expression of KDM6A in those cells. Comparing RNA-Seq and ChIP-Seq data, we identified the reason for chemoresistance to be downregulation of the drug influx transporter gene ENT1, conferred by reduced H3K27 acetylation. To quantify the impact of KDM6A loss in AML relapse, we screened AML patient and patient-derived xenograft (PDX) samples for KDM6A protein and/or mRNA expression and observed decreased expression levels in either measure in 44-50% of relapsed cases.

We next studied the impact of loss-of-function in the H3K27 methyltransferase EZH2. Accordingly, we established a CRISPR/Cas9-mediated knockout of EZH2 in the cell lines HEK293T and K562. We found that EZH2 knockout cells as well were more resistant to cytarabine. However, in contrast to KDM6A, loss of EZH2 did not induce resistance against daunorubicin. Analyzing RNA-Seq and proteomics data of EZH2 knockout cells, we

identified transcriptional as well as translational upregulation of EZH2 target genes due to the loss of H3K27 trimethylation. Those genes included FHL1, recently described to confer cytarabine resistance through its involvement in drug efflux transport. Similar to KDM6A, we found a high relevance of EZH2 loss in AML relapse, observing decreased EZH2 expression in 50% of relapsed patients.

The third epigenetic modifier we studied was DNMT3A. In a cohort of matched diagnosis and relapse samples from cytogenetically normal (CN) AML patients, we examined the consequence of DNMT3A loss on disease progression. To this end we screened patients for DNMT3A mutations and analyzed RNA-Seq as well as DNA methylation data comparing disease progression in DNMT3A mutated (DNMT3A-Mut) and DNMT3A wild type (DNMT3A-WT) patients. We found DNMT3A-Mut patients to exhibit overexpression of several poor prognosis marker genes involved in chemotherapy resistance including the antiapoptotic genes BCL6 and BCL2A1 as well as the membrane receptor CLEC7A. In the relapse of DNMT3A-Mut patients, we furthermore observed downregulation of genes involved in myeloid differentiation, as a response of elevated DNA methylation, leading to increased stemness. Moreover, we identified DNMT3B upregulation as a possible cause of this observed DNA hypermethylation.

All in all, we demonstrated that mutations and aberrant expression leading to loss of function in epigenetic modifiers play a key role in AML disease progression and the development of resistance against chemotherapy.

# Introduction

## Acute Myeloid Leukemia

Acute myeloid leukemia (AML) is a hematological cancer affecting the myeloid line of blood cells. Arising from the acquisition of chromosomal rearrangements and gene mutations, altering normal hematopoietic growth and differentiation, AML is a very heterogeneous disease with patient-specific mutation patterns (Deneberg et al. 2011). Chromosomal rearrangements can only be found in about half of all cases, leaving the other half with a cytogenetically normal (CN) karyotype (Marchesi et al. 2011). Acute leukemias are characterized by a rapid onset as well as fast disease progression. Patients usually suffer from weakness, shortness of breath, easy bruising, and excess bleeding as a result of anemia and thrombocytopenia (De Kouchkovsky and Abdul-Hay 2016; Albrecht 2014). White blood cell (WBC) counts can be increased (>11x$10^3$/µl), normal or even decreased (<4.5x$10^3$/µl), but usually contain a high amount of immature cells, impairing normal immune system function (White blood cell count 2021; Santos et al. 2009).

### Classification of AML Subtypes

The current AML classification is the revised 4th edition of the WHO classification published in 2016 (Swerdlow et al. 2017; Arber et al. 2016), superseding the FAB classification from 1985 (Bennett et al. 1985). It distinguishes six AML subtypes, considering cytogenetic and genetic abnormalities, immunophenotypic and morphological traits as well as the patient's prior medical history. For classification, gene panels screening for mutations in NPM1, CEBPA and RUNX1 are currently necessary. In most cases diagnosis of AML requires a myeloblast count of ≥ 20% in either bone marrow or peripheral

blood. Exceptions hereof are AML subtypes with certain genetic abnormalities and erythroid leukemia (Arber et al. 2016). Furthermore, AML can be classified into three prognostic risk groups (favorable, intermediate, adverse) according to the European LeukemiaNet (ELN) recommendation of 2017 (Table 2). Depending on cytogenetic and genetic abnormalities, those three categories predict responses to standard AML therapy independent of age and depend on the determination of the TP53, ASXL1, RUNX1, NPM1, CEBPA and FLT3-ITD mutation status (Döhner et al. 2017). The risk stratification of the Medical Research Council (MRC) additionally takes KMT2A-PTD and DNMT3A into consideration (Grimwade et al. 2016).

**Table 1: WHO classification.**

| **Acute myeloid leukemia (AML) and related neoplasms** |
|---|
| **AML with recurrent genetic abnormalities**<br>• AML with t(8;21)(q22;q22.1);RUNX1-RUNX1T1<br>• AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22);CBFB-MYH11<br>• APL with PML-RARA<br>• AML with t(9;11)(p21.3;q23.3);MLLT3-KMT2A<br>• AML with t(6;9)(p23;q34.1);DEK-NUP214<br>• AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); GATA2, MECOM<br>• AML (megakaryoblastic) with t(1;22)(p13.3;q13.3);RBM15-MKL1<br>• Provisional entity: AML with BCR-ABL1<br>• AML with mutated NPM1<br>• AML with biallelic mutations of CEBPA<br>• Provisional entity: AML with mutated RUNX1 |
| **AML with myelodysplasia-related changes** |
| **Therapy-related myeloid neoplasms** |
| **AML, NOS**<br>• AML with minimal differentiation<br>• AML without maturation<br>• AML with maturation<br>• Acute myelomonocytic leukemia<br>• Acute monoblastic/monocytic leukemia<br>• Pure erythroid leukemia<br>• Acute megakaryoblastic leukemia<br>• Acute basophilic leukemia<br>• Acute panmyelosis with myelofibrosis |
| **Myeloid sarcoma** |
| **Myeloid proliferations related to Down syndrome**<br>• Transient abnormal myelopoiesis (TAM)<br>• Myeloid leukemia associated with Down syndrome |

Adapted from (Arber et al. 2016)

**Table 2: ELN classification.**

| **Risk category** | **Genetic abnormality** |
|---|---|
| **Favorable** | t(8;21)(q22;q22.1); RUNX1-RUNX1T1 |
| | inv(16)(p13.1q22) or t(16;16)(p13.1;q22); CBFB-MYH11 |
| | Mutated NPM1 without FLT3-ITD or with FLT3-ITD$^{low}$ |
| | Biallelic mutated CEBPA |
| **Intermediate** | Mutated NPM1 and FLT3-ITD$^{high}$ |
| | Wild-type NPM1 without FLT3-ITD or with FLT3-ITD$^{low}$ |
| | t(9;11)(p21.3;q23.3); MLLT3-KMT2A |
| | Cytogenetic abnormalities not classified as favorable or adverse |
| **Adverse** | t(6;9)(p23;q34.1); DEK-NUP214 |
| | t(v;11q23.3); KMT2A rearranged |
| | t(9;22)(q34.1;q11.2); BCR-ABL1 |
| | inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); GATA2,MECOM(EVI1) |
| | −5 or del(5q); −7; −17/abn(17p) |
| | Complex karyotype, monosomal karyotype |
| | Wild-type NPM1 and FLT3-ITD$^{high}$ |
| | Mutated RUNX1 |
| | Mutated ASXL1 |
| | Mutated TP53 |

Adapted from (Döhner et al. 2017)

## AML Treatment Options

Developed in the 1970s, the '7+3' induction therapy, consisting of cytarabine treatment for seven days followed by an anthracycline (i.e daunorubicin) for another three days in combination with or without a purine analog, is to date still the most widely used regimen (Kantarjian et al. 2021). Depending on patients' fitness and relapse risk, induction therapy is followed by either consolidation chemotherapy or allogeneic stem cell transplantation (SCT) (Chen et al. 2019). Although 60-80% of younger patients (<60 years) and 40-60% of older patients (≥ 60 years) achieve complete remission, the 5 year overall survival rate (OS) is very low (30-35% in younger and 10-15% in older patients) (Kantarjian et al. 2021; Tang et al. 2021). Chemotherapy resistance due to persistent leukemia stem cells is the major cause of treatment failure (Senft and Jeremias 2019). To overcome and prevent chemotherapy resistance and to provide better treatment options for patients unfit for intensive chemotherapy, eleven new drugs have been approved by the EMA and FDA since 2012, nine within the last four years (Daver et al. 2020). Depending on the patient's age, fitness, leukemia subgroups and risk classification, patients eligible for intensive chemotherapy will generally receive standard therapy with or without targeted treatments. Patients that are not eligible for intensive chemotherapy will either receive low dose cytarabine (LDAC), or one of the hypomethylating agents (HMA) decitabine or azacitidine in combination with venetoclax or glasdegib and targeted therapies (Figure 1) (Daver et al. 2020; Tang et al. 2021). Even though targeted therapies lead to a substantial increase of survival in affected patients, the heterogeneity of AML between and within patients is still a major challenge in the development of new treatment approaches. In total, more than 60 genes have been found to be recurrently mutated in AML, however none of them are recurrently gained and only very few are lost at relapse, as seen across multiple studies (Vosberg and Greif 2019).

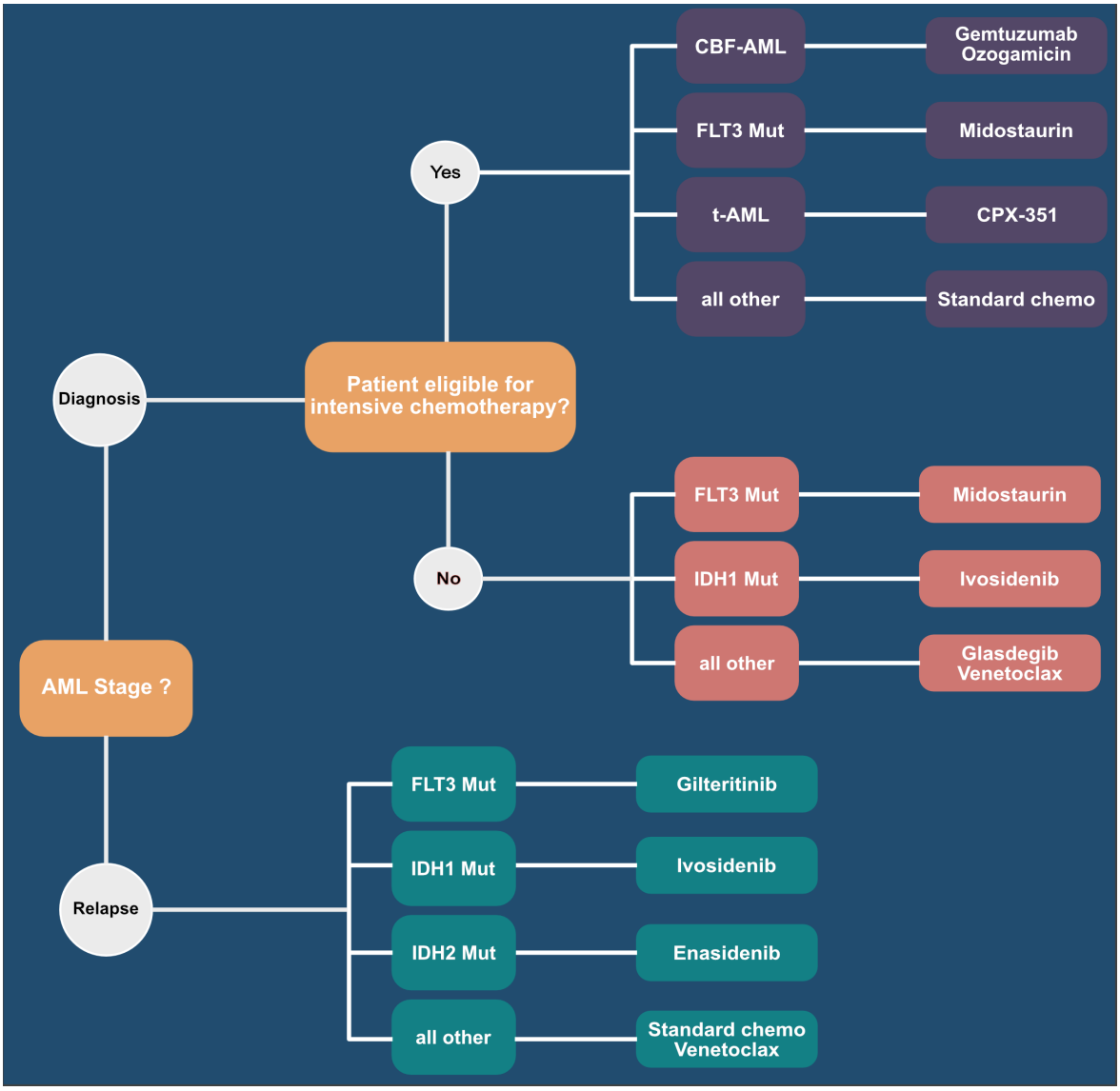



**Figure 1: Options for AML treatment**. Chemotherapeutics approved since 2017 for the treatment of AML. **Purple:** Chemotherapeutics approved for the treatment of patients eligible for intensive chemotherapy, usually in combination with standard therapy. **Orange:** Chemotherapeutics approved for the treatment of patients not eligible for intensive chemotherapy, usually in combination with LDAC or HMA. **Blue:** Chemotherapeutics approved for the treatment of relapsed AML.

## Clonal Evolution of AML

The change in the molecular profile during disease progression from diagnosis to relapse has been repeatedly described as genetic and epigenetic evolution of AML. It is commonly believed that AML arises from the sequential acquisition of mutations starting from one

single hematopoietic stem cell (HSC) or myeloid precursor cell with the potential of self-renewal. Acquired mutations leading to a block in myeloid differentiation and the preservation of the self-renewal capacity can create the basis for clonal evolution in AML (Vosberg and Greif 2019). The first mutations to be gained in the initiation of de-novo AML were found to arise in genes encoding epigenetic modifiers (e.g., DNMT3A, ASXL1 or TET2). Insufficient to induce leukemia but providing a crucial growth advantage for the cell, they are commonly termed preleukemic (Corces-Zimmerman et al. 2014; Jan et al. 2012). Those mutations can be detected years ahead of leukemia onset but are also detected in healthy people although associated with a higher risk of disease development and therefore pertain to a condition called clonal hematopoiesis of indeterminate potential (CHIP) (Jaiswal et al. 2014). CHIP mutations are usually followed by a mutation in NPM1 or in one of the transcription factors RUNX1, CEBPA or GATA2, which are again followed by mutations in signaling pathways like FLT3, RAS or KIT (Martignoles et al. 2018). Other types of clonal hierarchies in de-novo AML can follow paths without CHIP mutations or chromosomal aberrations as the initiating event (Martignoles et al. 2018). The model of clonal evolution can follow a linear path (linear evolution), in which mutations are acquired stepwise, or mutations can be acquired in different daughter clones giving rise to mutationally diverse populations called subclones (branching evolution) (Romer-Seibert and Meyer 2021). Subclonal diversity has been commonly described as the origin of relapse, as small subpopulations of cells are thought to survive chemotherapy and persist during remission to eventually reinstate the disease (Figure 2)(Ding et al. 2012; Shlush et al. 2017). Those cells are frequently termed leukemia stem cells (LSCs) as they possess stem cell properties including self-renewal, chemoresistance and repopulation potential. The measurable residual disease (MRD), comprising all surviving cells after chemotherapy, including LSCs, can be used as a measure of treatment efficacy and has recently been described as a predictor of AML relapse (Short et al. 2020).

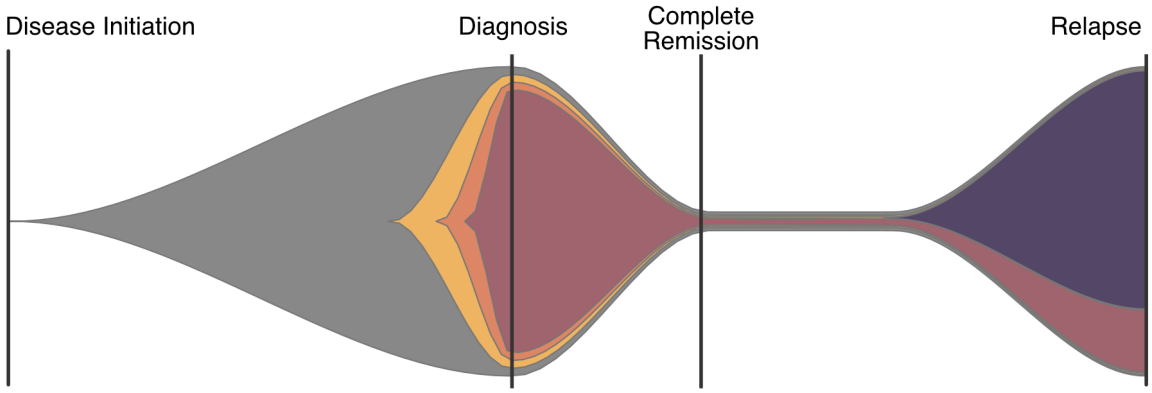



**Figure 2: Clonal Evolution**. Schematic representation of disease progression from leukemia initiation to relapse. **Colors** indicate different subsequent subclonal populations. **Gray:** Subclone with CHIP mutations occurring years before leukemia onset. **White** areas represent normal hematopoiesis.

## Epigenetic Modifications

Epigenetic modifications are reversible alterations in histone or DNA structure that regulate gene expression without changing the underlying DNA sequence. Involved in processes affecting development and differentiation of distinct cell types, they are crucial for regular cellular function including hematopoiesis. In AML, mutations in epigenetic modifiers play an important role in the onset of the disease (CHIP mutations) as well as in the progression to relapse and chemoresistance (Tulstrup et al. 2021; Li et al. 2016; Stief et al. 2019). According to their functional affiliation, epigenetic modifications can be assorted into the three major categories: DNA methylation, histone modification and nucleosome positioning (Portela and Esteller 2010).

### DNA methylation

DNA methylation can occur on either adenine or cytosine nucleotides, building N6-methyladenine, 5-methylcytosine, 5-hydroxymethylcytosine or N4-methylcytosine. Of those, 5-methylcytosine is the most common, occurring mainly on CpG dinucleotides (Faulk 2019). CpG and non-CpG (CpA, CpT and CpC) methylation have been described in the

context of gene regulation, however non-CpG methylation is mostly restricted to certain cell types such as stem cells, oocytes, neurons, skeletal muscle or hematopoietic cells (Lister et al. 2009; Jang et al. 2017; Pinney 2014). CpG methylation is predominantly linked to gene silencing and commonly located on CpG-Islands in promoter regions (Weber et al. 2007). The inhibition of transcription is thereby either conferred directly by blocking transcription factor (TF) binding or indirectly by the recruitment of methyl-CpG-binding domain (MBD) proteins inducing chromatin remodeling or likewise preventing transcription factor binding (Esteller 2007; Domcke et al. 2015). Transcriptional activation conferred by CpG methylation can occur, although quite rarely, due to TFs with the ability to bind methylated CpGs (Héberlé and Bardet 2019). Furthermore, gene body methylation has been associated with gene activation through the repression of intragenic transcription start sites (Jjingo et al. 2012). The transfer of methyl groups is mediated by a family of DNA methyltransferases (DNMTs) that comprises five members, the catalytically active DNMT1, DNMT3A and DNMT3B, as well as the regulatory factor DNMT3L and the RNA methyltransferase DNMT2 (TRDMT1). Responsible for de-novo DNA methylation, DNMT3A and DNMT3B are believed to construct the methylation pattern during embryogenesis and have been found to obtain essential roles in the differentiation of hematopoietic stem cells (Watanabe et al. 2002; Challen et al. 2011). While DNMT3A and DNMT3B display equal binding preferences for unmethylated and hemimethylated DNA, DNMT1 preferentially methylates hemimethylated DNA and is therefore seen as a maintenance methyltransferase (Robert et al. 2003; Rhee et al. 2000). In AML, aberrant DNA methylation occurs frequently and has been proposed as a clinically relevant classification for AML subtypes as well as a predictor for overall and event-free survival (Figueroa et al. 2010). One of the most prevalently mutated genes in AML is DNMT3A. Occurring in ~25% of all AML cases, DNMT3A mutations are a marker for poor prognosis resulting in reduced overall survival and shorter time to relapse (Spencer et al. 2014;

Lauber et al. 2020). In most cases, DNMT3A mutations are acquired early in disease onset at a preleukemic state (CHIP mutations) and are stable during disease progression to relapse (Greif et al. 2018). In contrast to DNMT3A, mutation events in DNMT3B and DNMT1 are rare, however elevated expression levels have been linked to poor disease outcome and chemoresistance (Cancer Genome Atlas Research Network ...: Niederwieser et al. 2015). DNMT3A variants, identified in AML, have been predominantly linked to a reduction in methyltransferase activity. These variants, including the highly abundant R882H variant, therefore cause hypomethylation of specific DNMT3A target sites (Russler-Germain et al. 2014). Responsible for DNA demethylation, TET2 holds an antagonistic role opposed to DNMT3A, and TET2 mutations have been associated with DNA hypermethylation. However, mutations in either of those genes provide HSCs with a severe growth advantage resulting in similar disease phenotypes (Ostrander et al. 2020).

## Histone modifications

Histone proteins assembled to nucleosomes participate in the regulation of transcription by impeding DNA accessibility. However, all histones undergo post-translational modifications, providing multiple mechanisms to alter this state. The most common histone modifications are acetylation, methylation, phosphorylation, ubiquitylation and sumoylation, appearing on both histone tails and core domains (Kouzarides 2007). Histone methylation, characterized by the addition of one, two or three methyl groups to histone proteins, was found at lysine, arginine or histidine residues (Greer and Shi 2012). Depending on the specific location of the methylated residue, transcription can either be promoted or inhibited. In either case, methyl groups are transferred from S-adenosylmethionine to their specific location by three enzyme families: the SET-domain-containing proteins and the DOT1-like proteins methylating lysines and the protein arginine N-methyltransferases (PRMTs) methylating arginines (Rea et al. 2000;

Feng et al. 2002; Bannister and Kouzarides 2011). So far only one methyltransferase (SETD3) has been identified capable of methylating histidine residues (Wilkinson et al. 2019). Trimethylation of histone H3 lysine 27 (H3K27), typically responsible for transcriptional repression, is conferred by the two SET-domain-containing proteins EZH2 and EZH1 (Lavarone et al. 2019). Responsible for mono- di- and trimethylation, either of them can be the central core protein of the PRC2 complex, comprising furthermore of the subunits EED, SUZ12 and RBBP7 (RbAp46) (Margueron and Reinberg 2011). Histone methylation has in general a slow turnover, however, has been found to be reversible. While two families of histone demethylases exist (The amine-oxidase type lysine-specific demethylases (LSDs) and the JumonjiC (JMJC) domain-containing demethylases), responsible for the removal of lysine methylation, so far only two arginine demethylases (PAD4 and JMJD6) have been identified (Zhang et al. 2019). KDM6A (UTX), belongs to the group of JMJC domain containing proteins, and specifically demethylates H3K27, making it an antagonist of EZH2 (Agger et al. 2007). Mutations in EZH2 or KDM6A are rare in AML and have been found to be mutually exclusive. Despite their antagonistic effects, loss-of-function of either protein was associated with poor prognosis and chemoresistance (Greif et al. 2018; Stief et al. 2019; Kempf et al. 2021). In addition to its H3K27 demethylase activity, KDM6A was found to confer H3K27 acetylation as well as H3K4 monomethylation and furthermore possess the ability to alter chromatin accessibility (Gozdecka et al. 2018). Histone acetylation was observed on lysine residues and is usually associated with transcriptional activation. Besides KDM6A, about 30 more histone acetyltransferases have been identified. Grouped into two classes, Type A histone acetyltransferases act in the nucleus in a transcription-related manner, while type B proteins acetylate newly synthesized histones in the cytoplasm (Gujral et al. 2020). Histones have been observed to harbor several modifications simultaneously, generating a combined effect to regulate transcription (Portela and Esteller 2010). Furthermore,

cross-talk between DNA-methylation and histone modifications, like the inhibition of de-novo DNA methylation through H3K4 methylation, adds to the complexity of epigenetic modifications (Ooi et al. 2007).

## Chromosome remodeling

Nucleosomes represent the basic packaging unit of eukaryotic DNA and consist of DNA wrapped around eight histone core molecules, two copies of each of H2A, H2B, H3, and H4, around which chromosomal DNA is wrapped. They inhibit transcription by blocking transcription factors from accessing the DNA and hinder polymerases from further elongation. The exact position of nucleosomes is therefore of great importance. Rearrangements of nucleosomes of as few as 30 bp in TSS regions have been described to perturb RNA polymerase activity (Schones et al. 2008). Histone variants are capable of regulating nucleosome position by recoiling more or less DNA than the original core histone (Talbert and Henikoff 2021). They can be exchanged with the original core histone through large chromatin-remodeling complexes, capable of unwrapping the DNA in an ATP-dependent process. Chromatin-remodeling complexes additionally regulate transcription by assisting in the establishment of accurate density and spacing of nucleosomes and in the ejection or relocation of histones to enable transcription factor binding (Clapier et al. 2017). Interactions between chromatin-remodelers and other epigenetic modifications occur frequently and have been found to regulate each other. For example, the position of nucleosomes has a crucial impact on DNA methylation, which preferentially targets nucleosome-bound DNA (Chodavarapu et al. 2010). However some histone variants like H2A.Z also have the potential to prevent DNA-methylation at CpG sites (Conerly et al. 2010).

## Omics Data and Analysis

Omics is the umbrella term for several fields of biological studies, including genomics, transcriptomics, epigenomics, proteomics and metabolomics, engaging with the analysis of complete genetic or molecular profiles (Hasin et al. 2017). For the characterization or quantification of the biomolecules of interest in the respective omics field, several high throughput techniques have been developed in recent years. Next generation sequencing (NGS) based techniques, determining the exact nucleic acid sequence of DNA or cDNA molecules, are frequently used in genomics, transcriptomics and epigenomics studies and include RNA-Seq (RNA Sequencing), ChIP-Seq (Chromatin Immunoprecipitation Sequencing), ATAC-Seq (Assay for Transposase-Accessible Chromatin Sequencing), WGBS (whole genome bisulfite sequencing), RRBS (Reduced Representation Bisulfite Sequencing), WGS (Whole Genome Sequencing) or WES (Whole Exome Sequencing) (Mortazavi et al. 2008; Barski et al. 2007; Buenrostro et al. 2013; Lister et al. 2009; Meissner et al. 2008; Bentley et al. 2008; van Dijk et al. 2014). Furthermore, mass-spectronomy (MS) based techniques, measuring the mass-to-charge ratio of ionized molecules, are commonly applied to characterize proteins and metabolites (Chakraborty et al. 2018). Although array-based techniques are widely replaced by NGS in transcriptomics and epigenomics studies, Methylation BeadChip Arrays, relying on the hybridization of molecules to predesigned probes, are still commonly used in epigenomic experiments (Hong et al. 2020; Xu et al. 2021). The rapidly decreasing costs of NGS and mass spectrometry permitted their application in clinical research and diagnostics and opened the door for personalized medicine (Morganti et al. 2020). NGS already plays a major role in AML diagnostics, as targeted panel sequencing is routinely used for the identification of marker genes, employed for AML classification and risk stratification (Leisch et al. 2019). It is likely to become even more important in the future for cancer prevention and screening in

patients with CHIP. NGS-based prediction models for this matter have been described, predicting AML up to 10 years ahead of its clinical diagnosis (Desai et al. 2018; Abelson et al. 2018). In this work we focus on RNA-Seq, Methylation BeadChip Arrays and mass-spectrometry to analyze the influence of loss of epigenetic modifiers in AML.

## RNA Sequencing

RNA Sequencing (RNA-Seq) is used to analyze the whole transcriptome through the detection and quantification of RNA molecules present in any given system. In cancer research, RNA-Seq has been employed to address numerous questions and aspects (Ozsolak and Milos 2011). Differential gene expression, one of the most common RNA-Seq applications, as well as transcript fusion detection have helped in the identification of several cancer biomarkers and potential therapeutic targets (Hong et al. 2020). Furthermore, single cell RNA-Seq (scRNA-Seq) can provide more insight into cancer evolution by the discovery and characterization of cancer heterogeneity and identification of resistant cell types (Hong et al. 2020). Various RNA-Seq library preparation methods for NGS have been established, each with its specific benefits and distinct scope of application. However, almost all of them follow five major steps: RNA extraction, reverse transcription, fragmentation, adapter ligation and sequencing (Figure 3) (Levin et al. 2010). Since total RNA mainly consists of ribosomal RNA (rRNA) and only about 1-2% is composed of messenger RNA (mRNA), mRNA needs to be purified either using rRNA depletion or mRNA capture with oligo-dT particles binding to their polyadenylated tails (Conesa et al. 2016). Illumina sequencers have a molecule size limitation due to reduced cluster efficiency of longer fragments, therefore fragmentation of mRNA or reverse transcribed mRNA (cDNA) is necessary. This can be achieved by either heat, chemical hydrolysis, enzymatic digestion or sonication (Hrdlickova et al. 2017). Another approach is the transposon-based tagmentation, combining fragmentation and adapter ligation into a single step, significantly

reducing library preparation time (Picelli et al. 2014). Before sequencing, adapters need to be ligated to the cDNA libraries, ensuring fragment binding to the flow cell (Fedurco et al. 2006). Furthermore, sample and/or molecule-specific identifiers (UMIs) can be added to multiplex samples, reduce batch effects and correctly identify duplicated reads (Marioni et al. 2008; Kivioja et al. 2011; Parekh et al. 2016). The final libraries will then undergo amplification by PCR to enrich the adapter-containing fragments and to produce the required amount of molecules for detection (van Dijk et al. 2014).

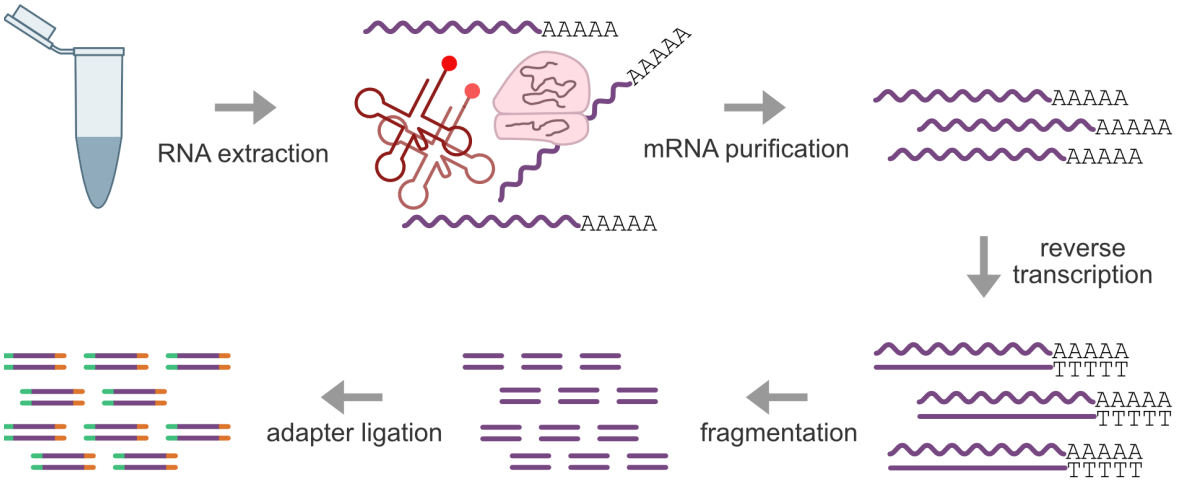



**Figure 3: RNA-Seq library preparation.** Illustration of the basic steps of RNA-Seq library preparation. **Purple**: nucleotide sequences of mRNA (wavy line) or cDNA (straight line). **Orange and green:** sequencing adapter for hybridization on the flow cell.

Introduced in 2003, Illumina's sequencing approach, sequencing by synthesis (SBS), has slowly outperformed all other techniques and became the current market leader (Giani et al. 2020). Illumina sequencing is based on bridge amplification, in which cluster generation takes place through repeated cycles of amplification on DNA strands forming a bridge by bending over and hybridizing to the next oligo (Kircher and Kelso 2010). The resulting clusters consisting of several thousand copies are then sequenced by the sequential incorporation of fluorescently labeled nucleotides. Each of the four bases is attached to a

unique fluorophore that simultaneously acts as a blocking group, permitting only one base to be detected per cycle (Slatko et al. 2018). Depending on the required application, several options for the sequencing layout exist. Single-end (SE) runs, sufficient for most quantitative gene expression studies, sequence a fragment from only one end, while paired-end (PE) runs, required for applications like de-novo transcript discovery, generate reads from both ends (Conesa et al. 2016). Read lengths starting from 50bp provide adequate mappability, but longer read lengths can be required for example for splice junction detection. With the current technology read lengths of up to 600bp (300bp paired-end) can be generated (Chhangawala et al. 2015). Another important factor to consider is the read depth, while more than 2 mio reads can be sufficient for some experiments others might need up to 200 mio reads depending on the required application. Adequate read depth as well as the number of replicates is essential to provide sufficient statistical power (Baccarella et al. 2018).

After sequencing, the Illumina instrument software converts the light intensity signals to bases, in a process called base-calling and predicts an accuracy score for those calls (Ledergerber and Dessimoz 2011). The resulting massive number of short reads then need to undergo basic data processing involving demultiplexing, mapping and counting.


**Figure 4: Processing of RNA-Seq data.** Illustration of RNA-Seq data processing from mapping of reads to gene counts. Reads of two samples (shades of **green**) are depicted as mapping to the genome (**gray** line), with exons of two genes indicated as colored boxes (shades of **purple**). Dashed lines indicate reads spanning over intron-exon junctions.

Multiplexing as many samples as possible to be sequenced in the same run prevents batch effects and ensures accurate analysis. Therefore, files need to be demultiplexed after sequencing, generating separate FASTQ files per sample (Renaud et al. 2015). To obtain information on the genomic location of each read, reads are then mapped (aligned) to a reference genome or transcriptome. This process requires the use of a splice-aware aligner, since transcripts are constructed from non-adjacent genomic regions (Dobin and Gingeras 2015). In contrast to other sequencing-based omics data, RNA-Seq reads do not require read trimming, a process that removes low quality bases or adapter sequences. Instead of improving mappability, trimmed reads were found to significantly change estimated expression levels, introducing bias in differential expression analysis. (Liao and Shi 2020; Williams et al. 2016). After mapping, expression levels are estimated by counting the aligned reads per gene or any genomic feature of interest, using specific read summarization programs such as featureCounts (Liao et al. 2014). During data analysis, several quality control (QC) steps are implemented to ensure high data quality and suitability for subsequent analyses. The per base sequence quality, GC content, presence of adapter or primer sequences and overrepresentation of k-mers as well as duplicated reads are evaluated on raw read data and present a measure of library preparation and sequencing quality. After mapping, the fraction of uniquely assigned reads to exonic regions is calculated and will range from 50 to 80% in a successful experiment (Sheng et al. 2017). One of the most common applications of RNA-Seq is the differential expression (DE) analysis. Used to quantify changes in normalized read counts between different states or groups, I applied DE analysis in this study to assess gene expression shifts between diagnosis and relapse samples as well as to to study the mechanism of expression regulation by EZH2 and KDM6A.

## Methylation BeadChip Arrays

Methylation BeadChip Arrays are used to analyze the DNA methylome by assessing the methylation status of cytosine nucleotides throughout the genome. This is achieved by sodium bisulfite treatment, converting unmodified cytosines to uracil leaving 5-Methylcytosine (5mC) and 5-Hydroxymethylcytosine (5hmC) intact (Booth et al. 2013; Dedeurwaerder et al. 2014). Illumina BeadChip arrays are widely used in comparative epigenetic population studies including cancer research (McEwen et al. 2018). Both the Infinium HumanMethylation450 (450k) array and its successor, the Infinium Methylation EPIC array (EPIC), use bead technology to measure methylation levels of >480,000 and >850,000 probes, respectively (Nakabayashi 2017). Each silica microbead is coated with several hundred thousand copies of one unique oligonucleotide probe and multiple beads for each probe are randomly distributed on the array surface (Morris and Beck 2015). Comprising 50 bases, probe sequences are designed to harbor a CpG (or non-CpG site) at its 3'end, enabling genotyping of the C/T conversion by single-base extension. EPIC and 450k arrays feature two different probe chemistries, the Infinium I assay, employing two different beads for the methylated and unmethylated state of each probe, and the Infinium II assay, managing with only one bead for both states (Pidsley et al. 2016). For Infinium I probes, methylation levels are measured in the same color channel for both methylated and unmethylated bead types, through the incorporation of a fluorescently labeled ddNTP depending on the CpG adjacent base (Bibikova et al. 2009). In contrast, for Infinium II probes methylation is measured through the incorporation of either a fluorescently labeled G or A base depending on a methylated (C) or unmethylated (T) state of the query site (Figure 5) (Bibikova et al. 2011). The proportion of methylation per site (β-values) is then calculated as the ratio of the methylated signal over the total signal (methylated +

unmethylated), where a β-value of 0 represents 0% methylation and a β-value of 1 represents 100% methylation (Dedeurwaerder et al. 2014).

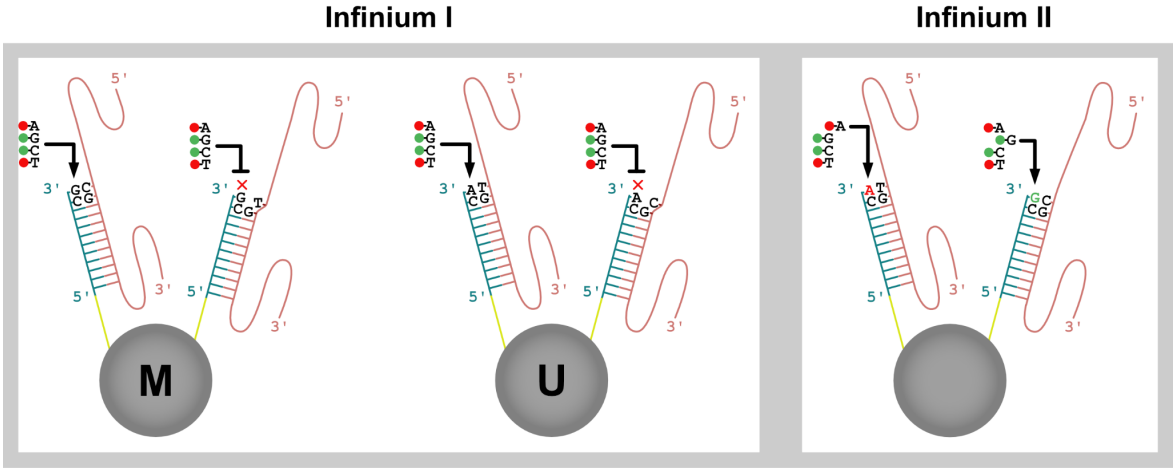



**Figure 5: Illumina Infinium Probe Design.** Illustration of the probe design for the Illumina Methylation BeadChip Array. The **infinium I** assay is based on two bead types **M** and **U** for the methylated and unmethylated loci respectively. Binding of the respective loci enables incorporation of a labeled nucleotide. The **Infinium II** assay supports binding of both loci on one bead, determining methylation status by single-base extension.

The combination of Infinium I and Infinium II probes allowed for the incorporation of additional cytosine loci, however posts challenges on data analysis due to differing dynamic ranges of both assays (Dedeurwaerder et al. 2014). Data preprocessing is therefore critical and includes the following steps: quality control, probe filtering and normalization. BeadChip arrays feature several hundred control probes to infer sample quality by measuring metrics like bisulfite conversion, or hybridization. In addition, the detection p-value, determining the detection of the total signal (methylated + unmethylated) over the background signal (negative control probes), can be calculated to identify failed positions and samples (Aryee et al. 2014). Probes depicting a high detection p-value pose an issue for data reliability and should be filtered out prior to normalization. Depending on the intended study it can also be beneficial to filter probes that hybridize to multiple genomic locations (cross-reactive probes), probes containing a single nucleotide polymorphism

(SNP) on the target CpG site or probes on sex chromosomes (Wang et al. 2018). To reduce technical noise introduced by the array design, performing within array normalization including background correction, dye bias adjustment and Infinium I/II- type bias correction is applicable (Wu et al. 2014).

## Mass-spectrometry

Mass-spectrometry (MS) based proteomics aims to analyze the whole proteome by investigating protein abundances, varieties of isoforms and protein-protein interactions. Of the two major approaches, top-down (analyzing the intact protein) and bottom-up (analyzing peptides), the latter is the most widely used. The bottom-up approach comprises an enzymatic digestion step, generating smaller peptides by the use of proteases, usually trypsin (Figure 6). To decrease complexity, this is commonly followed by fractionation of the peptides through gel electrophoresis or liquid chromatography (Dupree et al. 2020). High-performance liquid chromatography (HPLC), provides slow flow rates and facilitates the identification of peptides with equal mass, by employing hydrophobic separation of peptides in a capillary column (Karpievitch et al. 2010). The peptides are then eluted into a mass spectrometer, composed of three major components: ion source, mass analyzer and ion detector (Yates 2000). Relying on the analyzation of gaseous ions, mass spectrometers will first convert the peptides from a liquid state to gaseous ions, through either electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI) methods (Timp and Timp 2020). In a first run, ions are analyzed by measuring their mass-to-charge ratios (m/z) and are then further fragmented by employing either data-dependent analysis (DDA) or data-independent analysis (DIA) mode. While DDA selects the most intense ions only, DAI can analyze all ions by sequential fragmentation of selected m/z ranges (Zhang et al. 2013). Quadrupole mass analyzers, separating ions based on their trajectory stabilities in an oscillating electrical field, are the most used analyzers in proteomics. They are usually

combined with time-of-flight (TOF) or Orbitrap analyzers, distinguishing ions by velocity differences or oscillation frequencies respectively (Sinha and Mann 2020). The resulting ion current is then detected by electrodes, amplifying the response in an electron cascade to produce a measurable signal (Parker et al. 2011). From the combined data of all obtained measures, analyzed peptides can be identified through specific search algorithms in a predefined database (Schmidt et al. 2014). Subsequently intensity values representing protein expression are generated and can be used in quantitative proteomics analysis.

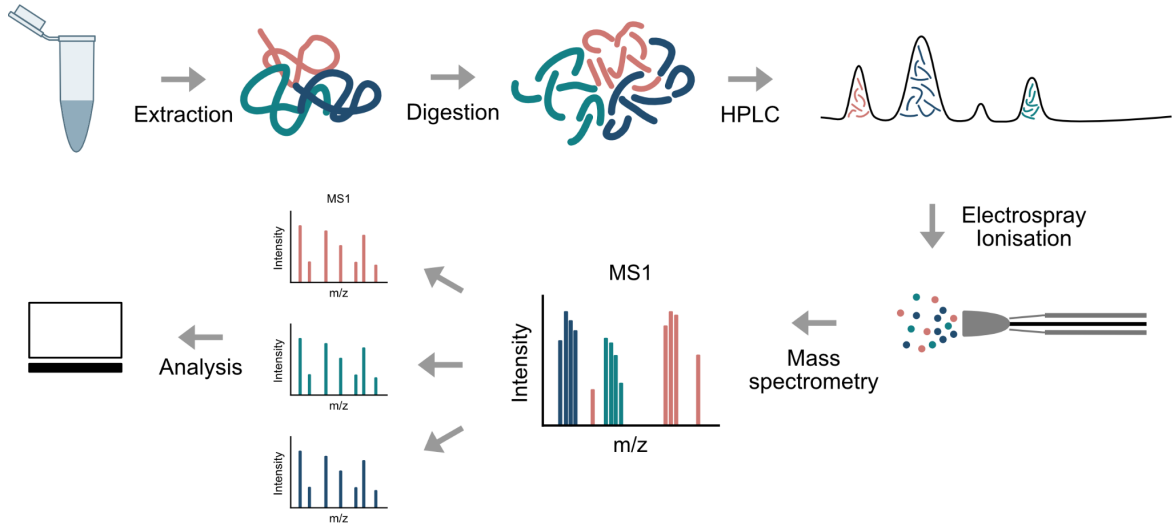



**Figure 6: Proteomics workflow.** Illustration of the major steps of the bottom-up proteomics workflow. Purified proteins are proteolytically digested in a first step and resulting peptides are then separated by liquid chromatography and charged by electrospray ionization. Mass spectrometry then reconstructs information about the primary proteins.

## Results

### Assessing the Impact of KDM6A loss in AML

---

# Loss of KDM6A confers drug resistance in acute myeloid leukemia

Leukemia 2019

---

Leukemia (2020) 34:50–62
https://doi.org/10.1038/s41375-019-0497-6

ARTICLE

Acute myeloid leukemia

# Loss of KDM6A confers drug resistance in acute myeloid leukemia

Sophie M. Stief[1,2,3] · Anna-Li Hanneforth[1] · Sabrina Weser[1] · Raphael Mattes[1] · Michela Carlet[4] · Wen-Hsin Liu[4] · Michael D. Bartoschek[5] · Helena Domínguez Moreno[6] · Matthias Oettle[1] · Julia Kempf[1] · Binje Vick[2,3,4] · Bianka Ksienzyk[1] · Belay Tizazu[1] · Maja Rothenberg-Thurley[1] · Hilmar Quentmeier[7] · Wolfgang Hiddemann[1,2] · Sebastian Vosberg[1,2,3] · Philipp A. Greif[1,2,3] · Klaus H. Metzeler[1] · Gunnar Schotta[6] · Sebastian Bultmann[5] · Irmela Jeremias[2,4,8] · Heinrich Leonhardt[5] · Karsten Spiekermann[1,2,3]

Received: 18 November 2018 / Revised: 25 March 2019 / Accepted: 18 April 2019 / Published online: 14 June 2019


## Abstract

Acute myeloid leukemia (AML) is an aggressive hematologic neoplasm resulting from the malignant transformation of myeloid progenitors. Despite intensive chemotherapy leading to initial treatment responses, relapse caused by intrinsic or acquired drug resistance represents a major challenge. Here, we report that histone 3 lysine 27 demethylase *KDM6A* (*UTX*) is targeted by inactivating mutations and mutation-independent regulation in relapsed AML. Analyses of matched diagnosis and relapse specimens from individuals with *KDM6A* mutations showed an outgrowth of the *KDM6A* mutated tumor population at relapse. KDM6A expression is heterogeneously regulated and relapse-specific loss of KDM6A was observed in 45.7% of CN-AML patients. *KDM6A*-null myeloid leukemia cells were more resistant to treatment with the chemotherapeutic agents cytarabine (AraC) and daunorubicin. Inducible re-expression of KDM6A in *KDM6A*-null cell lines suppressed proliferation and sensitized cells again to AraC treatment. RNA expression analysis and functional studies revealed that resistance to AraC was conferred by downregulation of the nucleoside membrane transporter *ENT1* (*SLC29A1*) by reduced H3K27 acetylation at the ENT1 locus. Our results show that loss of KDM6A provides cells with a selective advantage during chemotherapy, which ultimately leads to the observed outgrowth of clones with *KDM6A* mutations or reduced KDM6A expression at relapse.

**Supplementary information** The online version of this article (https://doi.org/10.1038/s41375-019-0497-6) contains supplementary material, which is available to authorized users.

✉ Karsten Spiekermann
karsten.spiekermann@med.uni-muenchen.de

[1] Department of Medicine III, University Hospital, LMU Munich, Munich, Germany

[2] German Cancer Consortium (DKTK), Heidelberg, Germany

[3] German Cancer Research Centre (DKFZ), Heidelberg, Germany

[4] Department of Apoptosis in Hematopoietic Stem Cells (AHS), Helmholtz Zentrum München, Munich, Germany

[5] Department of Biology II and Center for Integrated Protein Science Munich (CIPSM), Human Biology and BioImaging, LMU Munich, Planegg-Martinsried, Germany

[6] Biomedical Center and Center for Integrated Protein Science Munich, LMU Munich, Martinsried, Germany

[7] Department of Human and Animal Cell Lines, Leibniz-Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

[8] Department of Pediatrics, Dr. von Hauner Children's Hospital, LMU Munich, Munich, Germany

## Introduction

Acute myeloid leukemia (AML) is characterized by expansion of abnormal myeloid precursor cells in the bone marrow and blood. If not treated, AML can progress quickly and become fatal within a few months. Standard strategies for initial therapy, which have not changed substantially during the last 30 years [1], include cytarabine (AraC) in combination with anthracyclines like daunorubicin (DNR). Although the majority of AML patients achieve complete remission following induction chemotherapy, the disease reoccurs in 60–65% of younger adult patients (≤60 years) within 3 years after diagnosis [2]. Relapse represents the major cause for treatment failure and drug resistance is likely to play a role in its development. Recently, we have analyzed paired diagnosis and relapse samples of 50 cytogenetically normal (CN) AML patients and found *KDM6A* as a novel relapse-associated gene in AML [3]. KDM6A (or UTX) is a JmjC domain containing histone H3 lysine 27 (H3K27)-specific demethylase [4, 5] and belongs to the KDM6 family that include *KDM6B* and *UTY* [4, 6]. KDM6A can facilitate gene activation through

the catalytic JmjC domain and is also a component of the COMPASS-like complex, which is important for chromatin enhancer activation [7–11]. *KDM6A* is frequently targeted by somatic loss-of-function mutations in cancer [12–15] including leukemia [16–18]. Dependent on the cancer type, KDM6A appears to possess distinct tumor-suppressive functions. In T-cell acute lymphoblastic leukemia (T-ALL), *KDM6A* mutations are located almost exclusively in the JmjC domain [16, 17] and inactivation of the single *KDM6A* copy in males is sufficient to contribute to T-ALL pathogenesis [17]. In contrast, hematopoietic-specific loss of *Kdm6a* induces leukemogenesis through demethylase-independent alterations in H3K27 acetylation, H3K4 monomethylation and chromatin accessibility [19].

Using diagnosis and relapse samples from AML patients, patient-derived xenografts (PDX), and leukemia cell lines, we investigated the status of KDM6A during disease progression and the impact of KDM6A loss on chemotherapy resistance. We found three AML patients with enrichment of *KDM6A* loss-of-function mutations at relapse and relapse-specific loss of KDM6A mRNA and protein expression in 45.7% of CN-AML patients and 44.4% of AML patients, respectively. Reduction or loss of KDM6A expression in myeloid cell lines leads to increased resistance towards AraC and DNR treatment. Whereas re-expression of KDM6A in *KDM6A*-null cell lines sensitizes cells to AraC treatment. AraC resistance is achieved by reduction of the drug influx transporter ENT1. Taken together, our findings highlight KDM6A as a novel mediator of drug resistance in AML.

## Materials and methods

### Cell culture

Cell lines (Supplementary Table 1) were obtained from DSMZ (Braunschweig, Germany) and cultured according to the supplier's recommendation. PDX AML samples were recovered from mice and cultured as previously described [20]. Authentication of cell lines was performed using short tandem repeat typing. Testing for Mycoplasma contamination was done using the MycoAlert Mycoplasma detection kit (Lonza).

### Patients

Our analysis was based on samples from AML patients from the AMLCG-99 trial (NCT00266136), AMLCG-2008 trial (NCT01382147), and the Department of Medicine III, University Hospital, LMU. Informed consent for scientific use of sample material was received from all study participants in accordance with the Declaration of Helsinki.

### Proliferation assay

Cells were treated with cytarabine (Selleck Chemicals, Houston, TX, USA), DNR (in-house), 6-thioguanine (Sigma Aldrich), or NBMPR (Sigma Aldrich). After 72–96 h, viable cells were counted on Vi-Cell Cell Viability Analyzer (Beckman Coulter, Krefeld, Germany). For long-term proliferation, cells were treated three times (d0, d4, d8) and viable cells were counted every second day. Unpaired, two-tailed Student's *t*-test and calculation of $IC_{50}$ values were performed using GraphPad Prism version 6.07 (GraphPad Software, La Jolla, CA, USA). PiggyBac KDM6A cells were cultured +/− doxycycline (0.5 µg/mL) for 48 h followed by treatment with AraC +/− doxycycline for 72 h. For longer proliferation, cells were cultured +/− doxycycline (0.5 µg/mL) for 8 days. Every 2 days, cells were counted and cultured in fresh medium +/− doxycycline.

Additional methods are provided in supplementary methods.

## Results

### Gain of *KDM6A* mutations at relapse

Despite their initial response to chemotherapy, the majority of AML patients will develop chemotherapy resistance and relapse. Acquired *KDM6A* mutations were reported at relapse [3] pointing towards a novel mechanism of resistance in AML. To get insight into the biological relevance of *KDM6A* mutations, we first analyzed their locations in 20 AML patients at diagnosis. Patients with *KDM6A* mutations were from the AMLCG-99 trial ($n = 6$), AMLCG-2008 trial ($n = 9/664$), a diagnosis-relapse cohort ($n = 2/50$) [3] and this work ($n = 3$). We found a broad location pattern with the majority of mutations either located at or within the proximity of the tetratricopeptide repeat (TRP) or the JmjC domain (Fig. 1a). 65% of patients harbor either frameshift insertions/deletions or nonsense mutations suggesting a loss-of-function phenotype. In the majority of patients (12/20), the mutation occurred only in a subpopulation of AML cells, with a variant allele frequency (VAF) below 15% (Supplementary Fig. 1a). In addition to two previously described CN-AML patients [3], we identified three patients with recurrent *KDM6A* mutations using matched diagnosis and relapse samples, which were available for 3/18 patients (Fig. 1b; Supplementary Fig. 1b–d). In all patients we observed an increase in VAF of *KDM6A* mutations at relapse (Fig. 1b). The mutant clone E1325X showed the most striking increase at relapse (68.2% VAF), as it was barely detectable at diagnosis (0.58% VAF). Transplantation of relapsed tumor

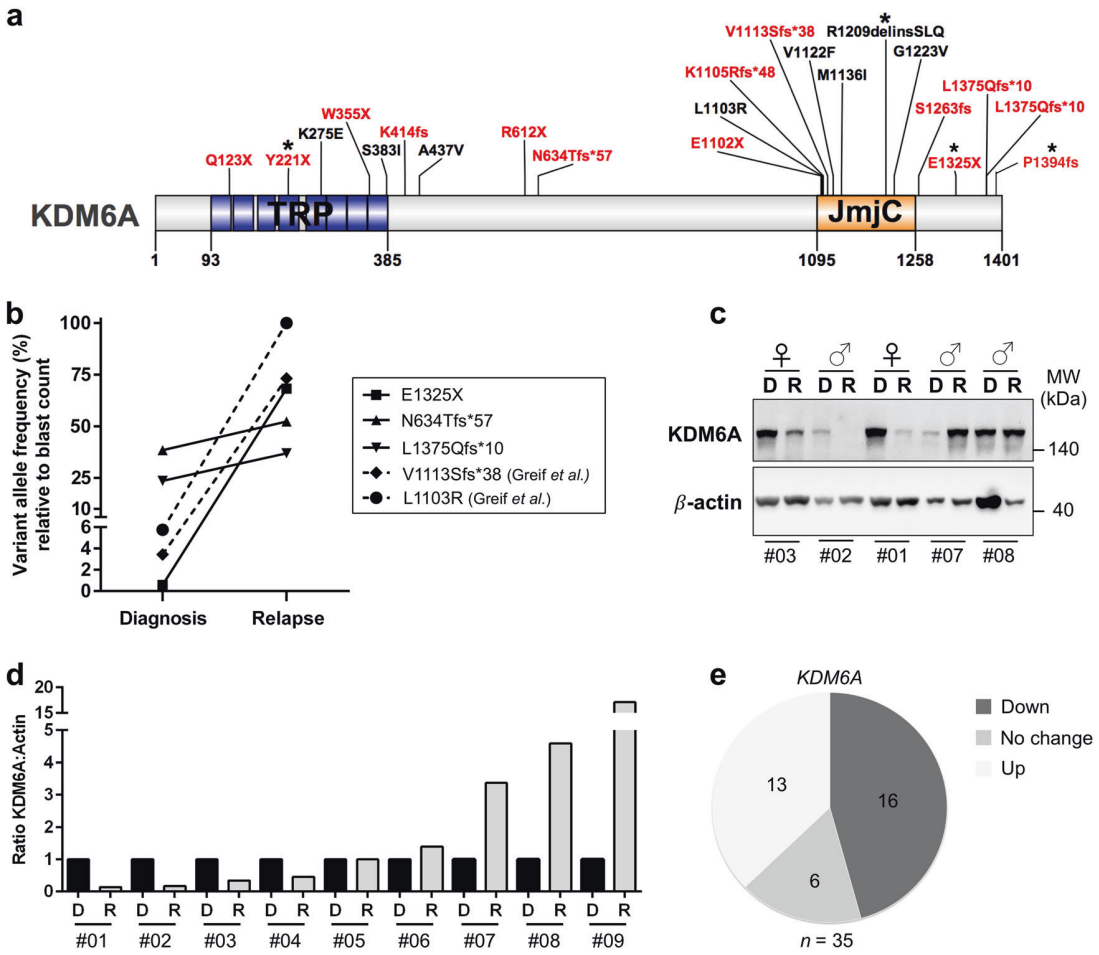



**Fig. 1** Gain of recurrent *KDM6A* mutations at relapse and change in KDM6A RNA and protein expression at relapse. **a** Schematic overview of KDM6A protein structure (NP_066963.2) and mutations (red = truncating; black = missense) identified at diagnosis in 20 AML patients, illustrated using IBS software [40]. Location of mutations is displayed and amino-acid positions are indicated below the graph. Asterisk (*) signifies two patients harboring two mutations each. Presented *KDM6A* mutations are from AMLCG-99 trial (NCT00266136), AMLCG-2008 trial (NCT01382147), a CN-AML diagnosis-relapse cohort [3] and this work. TRP tetratricopeptide repeat, JmjC Jumonji C. **b** Comparison of variant allele frequency (VAF) between diagnosis and relapse in 5 AML patients with *KDM6A* mutations. Due to variations in blast count, VAF was calculated relative to the respective blast count. Raw data for mutation L1130R and V1113Sfs*38 originate from our previous study [3]. **c**, Immunoblotting for KDM6A expression in five AML patients at diagnosis (D) and relapse (R). Their respective gender is shown on top and the UPN is displayed below. MW, molecular weight; β-actin, loading control. **d** Comparison of KDM6A protein expression in nine AML patients without *KDM6A* mutations at diagnosis and relapse. The ratio of KDM6A to β-actin expression is displayed. Respective values at relapse were normalized to the corresponding diagnosis sample. **e** Pie chart illustrating the regulation of *KDM6A* mRNA expression in 35 CN-AML patients. The three groups, *KDM6A*-up, *KDM6A*-down and *KDM6A*-no change were defined as a change in expression between diagnosis and relapse of above or below 20%, respectively

cells from this patient into immunodeficient mice (PDX model [20]) resulted in stable regeneration of *KDM6A* E1325X mutant clone (PDX AML-393; Supplementary Fig. 1b), which was verified by Sanger sequencing (Supplementary Fig. 1e). A second *KDM6A* mutation, P1394fs, was present in the same diagnosed patient with a 12.8-fold greater VAF (8.1%) than E1325X, but was lost at relapse (Supplementary Fig. 1b).

Next, we compared KDM6A expression in matched diagnosis and relapse samples of 9 AML patients (listed in Supplementary Table 2) with no detectable *KDM6A* mutation (Fig. 1c, d; Supplementary Fig. 1f). A strong decrease in KDM6A protein expression at relapse was observed in four patients whereas three patients showed increased expression at relapse. Additional analysis of *KDM6A* mRNA regulation in 35 CN-AML patients revealed a downregulation of *KDM6A* in 45.7% of patients ($n = 16/35$) and an upregulation in 37.1% of patients ($n = 13/35$; Fig. 1e and Supplementary Fig. 2). Interestingly, DNA methylation levels of KDM6A varied between AML patients at diagnosis and AML patients with high DNA methylation levels of KDM6A (top 25%) showed

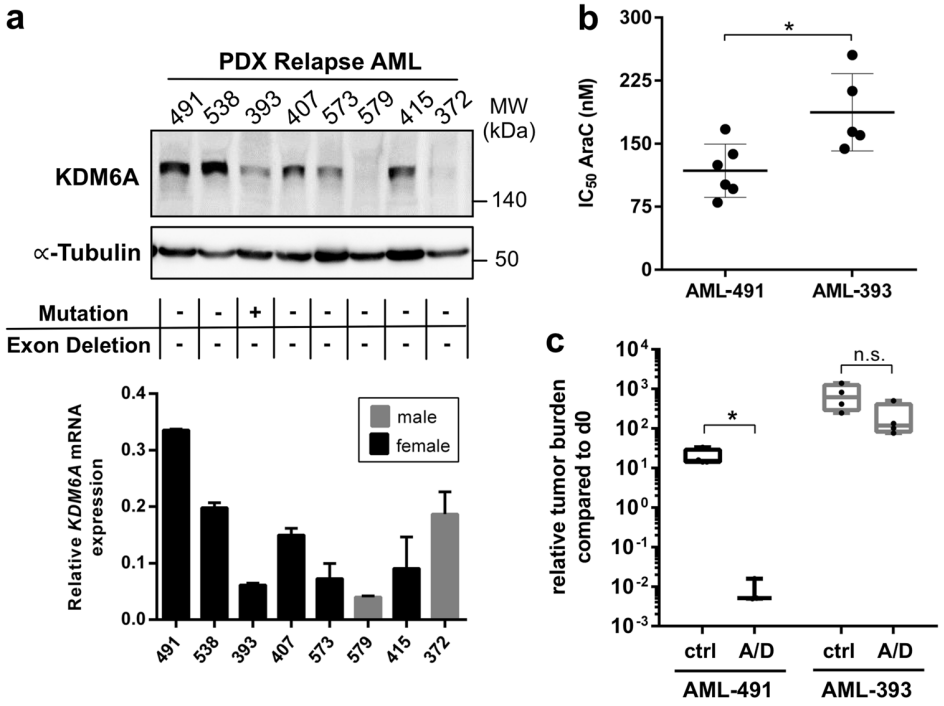



**Fig. 2** Heterogeneous KDM6A expression in relapsed PDX AML cells and decreased drug sensitivity in *KDM6A* mutant PDX AML-393 cells. **a** KDM6A mRNA and protein expression was analyzed in eight PDX samples from relapsed AML patients. For Western Blot, α-Tubulin was used as a loading control. *KDM6A* mRNA expression is shown relative to *GAPDH* expression. Male samples are highlighted as gray bar plots. The occurrence of *KDM6A* mutations or exon deletions is indicated with the "+" symbol. **b** Comparison of AraC $IC_{50}$ values between *KDM6A* WT PDX AML-491 and *KDM6A* mutant PDX AML-393 in vitro. Mean ± s.d. are given for at least five independent experiments performed in duplicates. Unpaired, two-tailed Student's *t*-test; $*P = 0.016$. **c** Comparison of tumor load reduction under in vivo chemotherapy in AML-491 and AML-393 bearing animals. Tumor burden was quantified by bioluminescence before (d0) and after (d28) two cycles of treatment with AraC (days 1–4, 14–17) and DNX (days 1, 4, 14, and 17) (A/D) or control treated animals (ctrl). Relative tumor burden at day 28 compared to d0 was calculated. Unpaired, two-tailed Student's *t*-test; $*P = 0.0157$; n.s., not significant

significantly shorter overall survival in a publicly available dataset [21] (Supplementary Fig. 3).

To further investigate the relevance of KDM6A at relapse, we analyzed KDM6A mRNA and protein expression in 8 PDX AML samples, established from primary patients′ cells at relapse. 50% of PDX AML samples showed low or undetectable protein expression, which, except for AML-372, correlated with mRNA expression (Fig. 2a). Mutational analysis by MLPA and targeted sequencing revealed only the above described *KDM6A* mutation (E1325X) in PDX AML-393 (Supplementary Fig. 1b). No additional *KDM6A* exon deletion mutations were detected (Supplementary Fig. 4). mRNA expression of the histone demethylase *KDM6B* and the histone methyltransferase *EZH2* were slightly increased in AML-579 cells, whereas AML-538 showed low *KDM6B* and AML-491 low *EZH2* mRNA expression (Supplementary Fig. 5a, b). Analysis of the mRNA expression of *UTY* in PDX AML samples showed normal *UTY* levels (Supplementary Fig. 6d). Since we were unable to detect a low molecular weight band corresponding to the premature stop mutation E1325X (estimated protein weight: 145 kDa) in the female PDX AML-393 cells, we treated these cells in vitro with the proteasomal inhibitor MG132. No increase in overall KDM6A expression, but also no appearance of an additional band corresponding to E1325X was observed (Supplementary Fig. 5c), which might point towards a nonsense-mediated mRNA decay. When we overexpressed E1325X, L1103R, and V1113Sfs*38 *KDM6A* mutants in HEK293T cells, proteasomal inhibition resulted in a significantly elevated expression of these mutants whereas wildtype (WT) and the demethylase-dead mutant H1146A did not change (Supplementary Fig. 5d, e). These results suggest that reduction of KDM6A expression might be either facilitated by nonsense-mediated mRNA decay or by proteasomal degradation.

Next, we asked whether drug resistance might have resulted in the outgrowth of the *KDM6A* mutated population at relapse. Since AraC was a component of the induction therapy, we investigated whether *KDM6A* mutant cells are less sensitive to AraC treatment. Therefore, we used PDX AML models of the same adverse ELN classification [20] with (i) *KDM6A* WT (AML-491) and (ii) *KDM6A* mutant (AML-393) and compared the half-inhibitory concentration

($IC_{50}$) of AraC. In vitro treatment for 72 h revealed a 0.63-fold decreased sensitivity towards AraC treatment for *KDM6A* mutant AML-393 compared to *KDM6A* WT AML-491 cells (187.3 nM vs. 117.9 nM; Fig. 2b). We observed the same effect in vivo after treating mice bearing AML-491 or AML-393 with two cycles of AraC and DaunoXome (DNX; liposomal DNR; Fig. 2c). Treatment dramatically decreased the tumor burden in *KDM6A* WT AML-491 bearing mice compared to control ($P = 0.0157$; Fig. 2c), whereas only a modest drop in tumor burden was observed in treated AML-393 bearing mice. Overall, these results indicate that a reduced KDM6A expression is associated with a decreased AraC sensitivity.

## Decreased AraC sensitivity in *KDM6A* mutant cells

*KDM6A* exon deletion mutations have been observed in AML cell lines, MONO-MAC-6 (MM-6) and THP-1 [12]. To identify the frequency of *KDM6A* deletions in leukemia, we performed MLPA analysis for the *KDM6A* gene in 40 myeloid leukemia cell lines (Supplementary Table 1). We identified two additional AML cell lines, OCI-AML3 and HL-60, with in-frame deletions in exon 3–4 and 5–6, respectively (Supplementary Fig. 4). These deletions were confirmed by independent CytoScan HD Array hybridization analysis (Fig. 3a). Both in-frame deletions lead to a truncated protein of approximately 147 kDa (WT: 154 kDa). Although low to intermediate mRNA expression was detectable in the mutant cell lines (Supplementary Fig. 6a), KDM6A protein expression was completely absent (Fig. 3b). *KDM6A* mutant cells showed increased H3K27 tri-methylation, whereas H3K27 di- and mono-methylation levels were similar between WT and mutant cells (Fig. 3b). Global H3K27me3 was inversely correlated with KDM6A levels ($r = -0.76$; $P = 0.0042$; Fig. 3c) indicating a KDM6A dependent epigenetically altered phenotype. Additionally, we analyzed expression of *KDM6B* and *EZH2* via qPCR as these genes might compensate for KDM6A loss. For both genes, mRNA expression was similar between mutant and WT cells (Supplementary Fig. 6b, c). Analysis of *UTY* mRNA expression showed loss of *UTY* mRNA expression in two *KDM6A* mutant and three *KDM6A* WT cell lines (Supplementary Fig. 6d).

Next, we investigated whether KDM6A loss leads to increased AraC resistance. *KDM6A* mutant cells showed a trend towards higher AraC $IC_{50}$ values compared to WT (Supplementary Fig. 7a). To eliminate gender-specific effects of *KDM6A* WT cells (higher expression in females as *KDM6A* escapes X inactivation [3, 17]), we compared the $IC_{50}$ values of male AML cell lines. Male *KDM6A* mutant AML cell lines had significantly increased $IC_{50}$ values compared to WT cells ($P = 0.0441$; Fig. 3d). We demonstrated previously [3] that MM-1 cells (*KDM6A* WT) are more sensitive to AraC treatment than the *KDM6A* mutant sister cells MM-6. To investigate if MM-6 also has a competitive growth advantage compared to MM-1 under AraC therapy, we performed a competitive assay mixing MM-1 with MM-6 cells in a 9:1 ratio. Native conditions as well as treatment with AraC significantly increased the number of MM-6 cells to 32.3% (native conditions) or 52.6% (270 nM AraC) after 9 days (Supplementary Fig. 7b). Treatment of MM-1 cells with AraC, DNR or 6-TG for 72 h, applying concentrations in the range of their respective $IC_{50}$ values, induced an upregulation of KDM6A protein expression (Supplementary Fig. 8).

## Knockdown of *KDM6A* confers decreased AraC and DNR sensitivity in K562 cells

Although AraC was part of the induction regimen in the five investigated patients with *KDM6A* mutations, the composition of induction therapy varied between patients and in certain cases also included the drugs DNR and/or 6-thioguanine (6-TG). To investigate if reduced expression of KDM6A leads to increased resistance towards multiple drugs, we performed lentiviral shRNA-mediated knockdown (KD) of *KDM6A* in the myeloid cell line K562. Of several tested shRNAs, sh*KDM6A* #3, #4, and #7 decreased KDM6A expression by 70% (#3, #4) or 90% (#7; Fig. 4a). Next, *KDM6A* KD or control cells were treated for 72 h with AraC, DNR, or 6-TG. *KDM6A* KD cells displayed decreased sensitivity towards AraC treatment (Fig. 4b) applying doses within the range of reported AraC plasma concentrations in patients [22] (Supplementary Fig. 9a). Only KD with the most potent sh*KDM6A* #7 resulted in a significantly increased resistance to AraC (Fig. 4b). However, the effect of *KDM6A* KD on response towards DNR or 6-thioguanine (6-TG) was not as prominent or even absent: only KD cells sh*KDM6A* #7 were slightly more resistant to DNR treatment (Fig. 4c), and no change in $IC_{50}$ values was observed after 6-TG treatment (Fig. 4d). The applied DNR concentrations (5–75 nM) are within the lower range of the reported concentrations in AML patients (403.8 ± 349 nM) [23]. Since induction therapy typically involves continuous treatment for seven ("7 + 3") or 10 days (TAD regime), we next applied a prolonged time course with multiple treatments. Prolonged treatment with 6-TG showed no difference in the amount of viable cells between control and KD after 14 days (Fig. 4g). Differences in growth under AraC treatment started at day 4, and resulted in a significant proliferative advantage for *KDM6A* KD cells compared to control (Fig. 4e). Growth of both control groups was completely arrested under DNR treatment after day 8, whereas *KDM6A* KD cells were strongly proliferating (Fig. 4f). *KDM6A* KD efficiency and proliferative advantage under DNR were positively correlated. To compare the

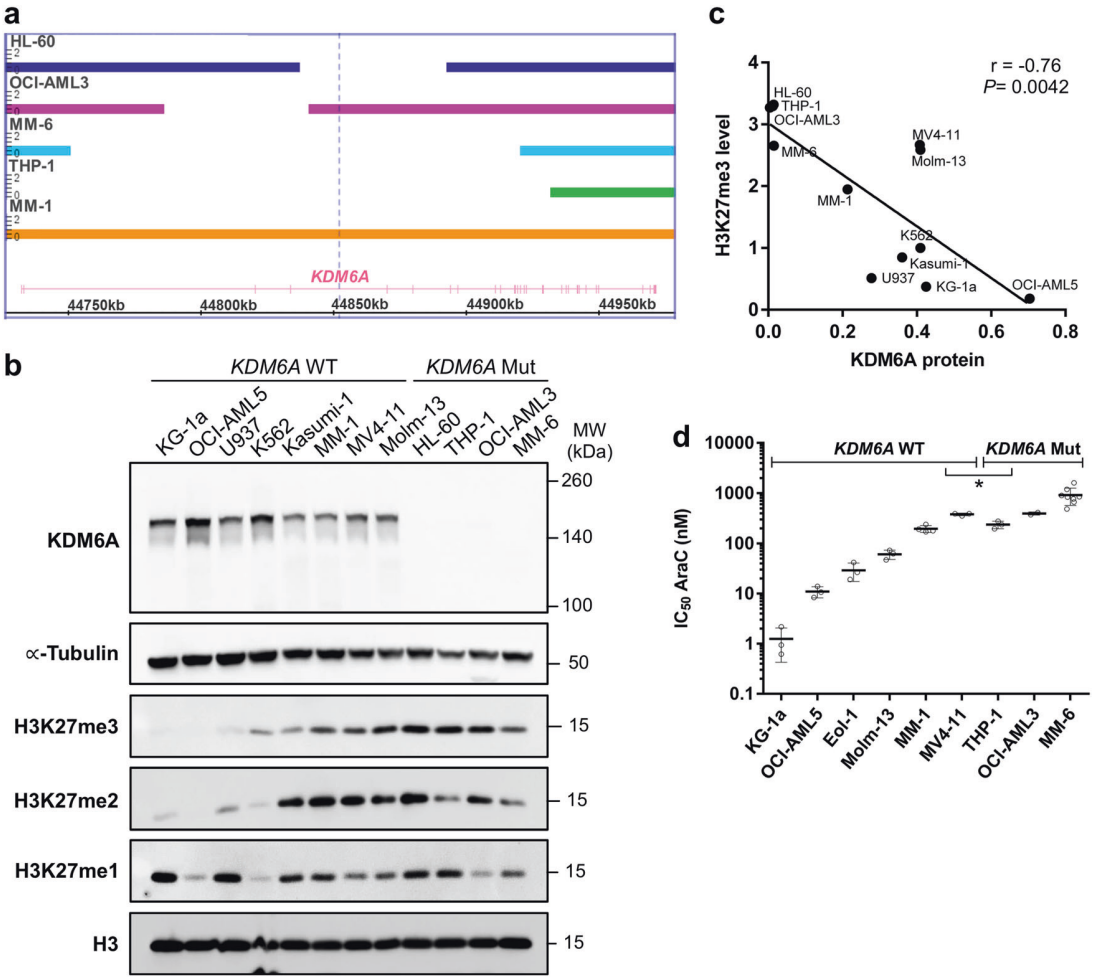



**Fig. 3** AML cell lines with KDM6A loss show high H3K27 trimethylation and increased AraC resistance. **a** Identification of *KDM6A* aberrations in AML cell lines HL60, OCI-AML3, MM-6, and THP-1. MM-1 serves as a WT control. Bar thickness ranging from 0 to 2 indicates the copy number (CN) status. Haploidy (CN = 1) in cell lines from male patients is due to the X-chromosomal localization of *KDM6A*. HL-60, the only cell line derived from a woman has lost one of its X chromosomes. **b** Western blot for KDM6A expression and global H3K27 mono-, di- and tri-methylation levels in *KDM6A* WT and mutant human leukemia cell lines. α-Tubulin and total H3 were used as loading controls. Blots are representative of three independent experiments. MW, molecular weight. **c** Negative correlation between KDM6A protein expression and global H3K27me3 level in *KDM6A* WT and mutant human leukemia cell lines (Pearson's correlation; $r = -0.76$, $^*P = 0.0042$). Mean values of three independent analyzes are shown. **d** Comparison of AraC $IC_{50}$ values between *KDM6A* WT and mutant male AML cell lines. To determine their respective $IC_{50}$ values, AML cell lines were treated with increasing concentrations of AraC for 72 h. Mean of $IC_{50}$ values from at least three independent experiments ± standard deviation (s.d.) are shown. Unpaired, two-tailed Student's *t*-test; $^*P = 0.044$

impact of *KDM6A* KD vs. knockout (KO), we applied CRISPR/Cas9 genome editing to ablate KDM6A expression in K562 cells. We established single cell *KDM6A* KO clones (Supplementary Fig. 9c), leading to complete loss of KDM6A expression (Fig. 4h). After 72 h AraC treatment, $IC_{50}$ values were significantly increased for both *KDM6A* KO clones compared to controls (Fig. 4i). We observed a trend towards higher $IC_{50}$ values or no difference between KO and control cells after DNR or 6-TG treatment, respectively (Supplementary Fig. 9d, e). These data indicate that reduction or loss of KDM6A expression in K562 cells increases resistance to AraC and DNR but not 6-TG.

## Loss of KDM6A in MM-1 recapitulates the drug phenotype of the *KDM6A* mutant sister cells MM-6

The sister cell lines MM-1 and MM-6 have originally been established in culture from the same male AML patient at relapse [24]. Whereas MM-1 cells express KDM6A, MM-6 cells harbor a *KDM6A* exon deletion, rendering them a good model to examine the implications of KDM6A loss within a similar genetic background. To investigate whether *KDM6A* KO in MM-1 cells results in the same drug resistance phenotype observed in MM-6 cells (Fig. 5d, e), we applied CRISPR/Cas9 targeting the *KDM6A* locus in MM-1 cells

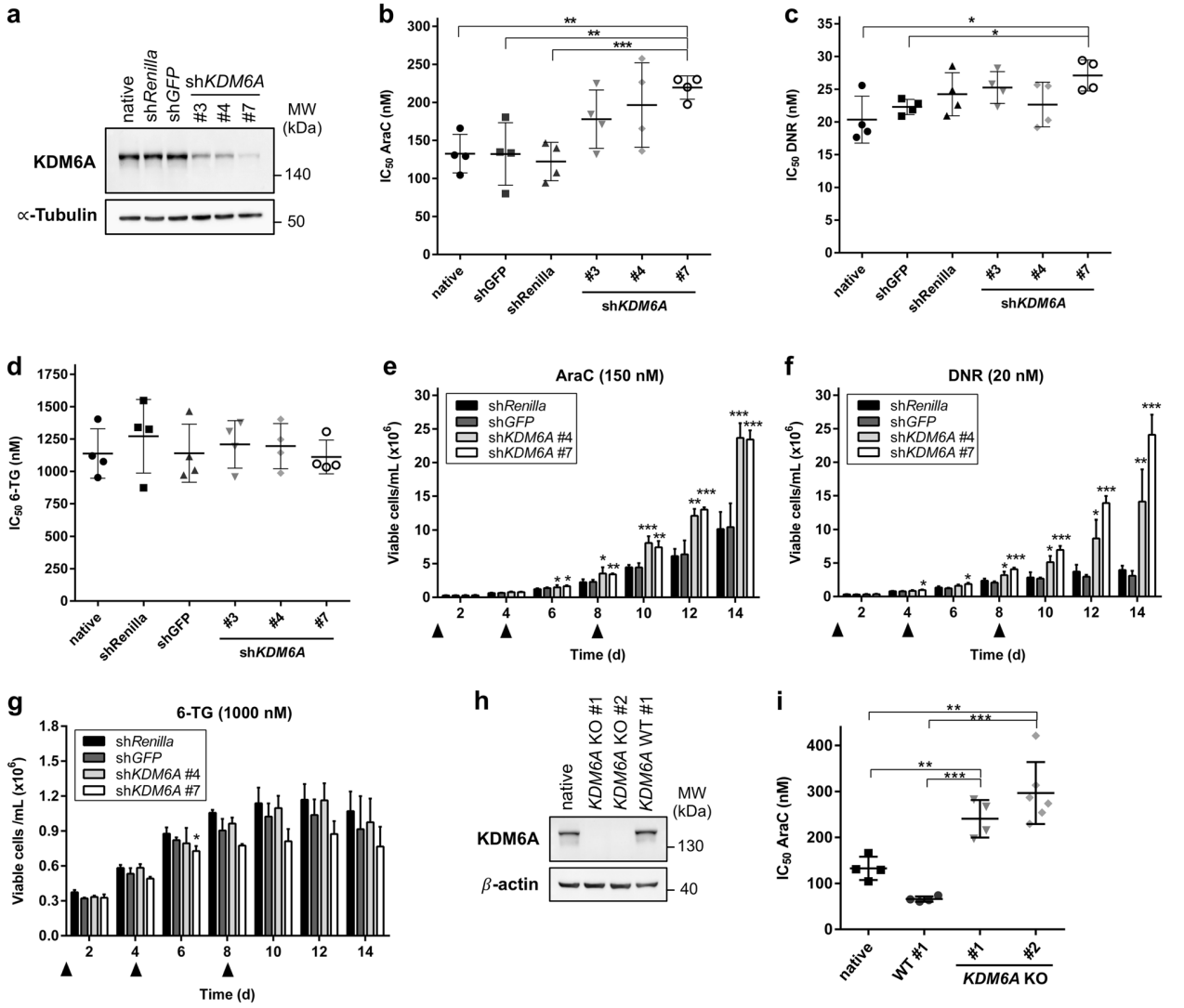



**Fig. 4** Reduction or loss of KDM6A expression confers decreased AraC and DNR sensitivity in K562 cells. **a** Immunoblot showing knockdown (KD) of KDM6A expression in K562 cells lentiviral transduced with three different shRNAs against *KDM6A*. sh*Renilla* and sh*GFP* serve as controls. Blot is representative of three independent experiments. MW molecular weight, α-Tubulin loading control. **b**–**d** Comparison of $IC_{50}$ values for AraC (**b**), DNR (**c**), and 6-TG (**d**) between control and *KDM6A* KD in K562 cells. Cells were treated for 72 h with the respective drug. Mean of $IC_{50}$ values from four independent experiments ± s.d. are shown. Unpaired, two-tailed Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. **e**–**g** Long-term proliferation assay measuring the amount of viable K562 cells, sh*Control* and sh*KDM6A*, every 2 days for 14 days in total. Cells were treated with 150 nM AraC (**e**), 20 nM DNR (**f**), or 1000 nM 6-TG (**g**) on Day 0, 4, and 8 as indicated with the triangle. Mean ± s.d. are given for three independent experiments. Unpaired, two-tailed Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. **h**, Immunoblot showing loss of KDM6A protein in *KDM6A* knockout (KO) K562 cells. Results of one representative experiment are shown ($n = 3$ experiments). MW molecular weight, β-actin loading control. **i** Increase in AraC $IC_{50}$ values in both *KDM6A* KO K562 clones compared to WT #1 clone or native cells. Mean of $IC_{50}$ values ± s.d. from at least four independent experiments are shown. Unpaired, two-tailed Student's *t*-test; $**P < 0.01$; $***P < 0.001$

(Fig. 5a). Compared to MM-1 parental and WT single cell clones, KDM6A expression was lost in one clone (Fig. 5b). In agreement with our observation that MM-6 cells are 4.3-fold more resistant to AraC after 72 h than MM-1 (Supplementary Fig. 7a), *KDM6A* KO in MM-1 increased the AraC $IC_{50}$ compared to both WT clones (3.4 to 8.8-fold increase after 96 h treatment; Fig. 5c). Comparison of the $IC_{50}$ values after DNR (Figs. 5d) and 6-TG (Fig. 5e) treatment indicated that *KDM6A* KO MM-1 are significantly less sensitive to DNR and 6-TG than *KDM6A* WT MM-1 cells. Strikingly, KO of *KDM6A* in MM-1 conferred a similar resistance to DNR (MM-1: KO vs. WT#1/#2: 2.5 to 2.6-fold; native MM-1 vs. MM-6: 2.5-fold increase in $IC_{50}$) and 6-TG (MM-1: KO vs. WT#1/#2: 1.6 to 2.1-fold; native MM-1 vs. MM-6: 1.9-fold) as in MM-6 cells. Together these data suggest that loss of KDM6A in MM-1

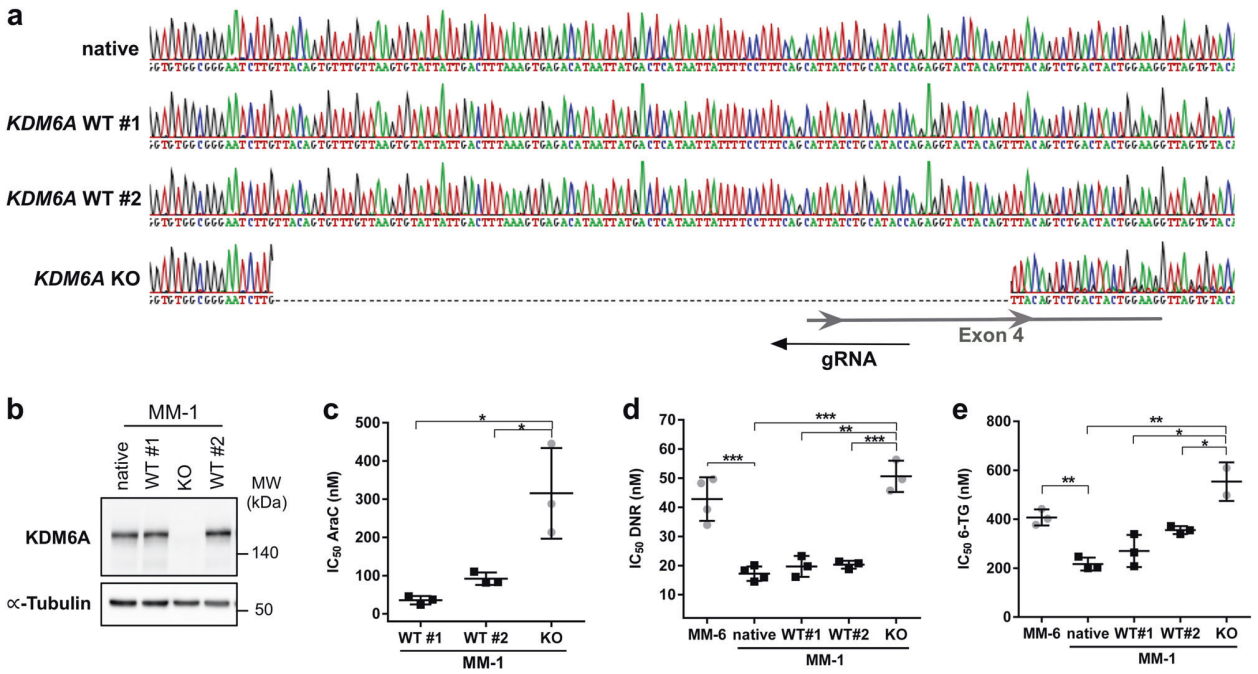



**Fig. 5** KDM6A loss in MM-1 recapitulates the drug phenotype of the *KDM6A* mutant sister cell line MM-6. **a** DNA sequencing chromatogram showing *KDM6A* frameshift mutation A112Vfs*3 of *KDM6A* knockout (KO) MM-1 clone which is absent in native MM-1 cells and two *KDM6A* WT clones. Last 74 bp of Intron 3 and 29 bp of exon 4 are deleted. WT #1 and #2 clones were tested negative for *KDM6A* KO after CRISPR/Cas9 targeting. **b** Immunoblotting for KDM6A expression in KDM6A WT and KO cells. Blot is representative of three independent experiments. MW molecular weight, α-Tubulin loading control. **c**–**e** Comparison of $IC_{50}$ values for AraC (**c**), DNR (**d**), and 6-TG (**e**) between MM-6 and MM-1 cells including MM-1 native, two *KDM6A* WT and one *KDM6A* KO clone. Cells were treated for 96 h with the respective drug. Mean ± s.d are given for three independent experiments. Unpaired, two-tailed Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$

is responsible for a decreased sensitivity towards AraC, DNR, and 6-TG.

## Re-expression of KDM6A sensitizes cells to AraC treatment

To investigate if re-introduction of KDM6A in *KDM6A*-null cell lines sensitizes them to AraC treatment, we generated stable cell lines with a doxycycline inducible PiggyBac (PB) KDM6A expression system (Fig. 6a–c). Re-expression of KDM6A significantly suppressed proliferation in THP-1 and K562 *KDM6A* KO cells (Fig. 6d, e). Furthermore, re-expression of KDM6A significantly decreased the amount of viable cells after treatment with AraC in both cell lines (Fig. 6f, g). *KDM6A* KO K562 cells, which were more resistant to AraC than *KDM6A* WT cells, were sensitized again to AraC treatment by re-expressing KDM6A (Fig. 6f). To investigate if the demethylase activity of KDM6A is essential for re-sensitizing cells to AraC treatment, we expressed a catalytically dead *KDM6A* mutant, H1146A [25], in the K562 *KDM6A* KO #1 and #2 cells. Expression of KDM6A H1146A showed only a trend in decreasing the amount of viable cells after treatment with AraC (Fig. 6h, i).

## Decreased ENT1 expression by *KDM6A* loss mediates increased AraC resistance

To identify genes involved in KDM6A-mediated drug resistance, we performed RNA-Seq analysis in K562 cells treated with siRNA (Supplementary Fig. 10a, b) or shRNA against *KDM6A* (Fig. 4a). Transient KDM6A KD with siRNA, which showed a trend towards higher $IC_{50}$ for AraC compared to control (Supplementary Fig. 10c), resulted in transcriptional downregulation of 39 genes and upregulation of 7 genes (Supplementary Fig. 10d). For the most potent sh*KDM6A* #7 we detected transcriptional deregulation of 295 genes compared to 7 or 54 deregulated genes during sh*KDM6A* #3 or #4 mediated KD, respectively (Supplementary Fig. 10e). Whereas the majority of differentially expressed genes (39/46) was downregulated in the siRNA-mediated KD (Supplementary Fig. 10d), sh*KDM6A* #7 KD resulted in similar transcriptional down- (150, 50.8%) and upregulation (145, 49.2%; Fig. 7a). Treatment with AraC (150 nM) during shRNA-mediated KD led to increased transcriptional deregulation (sh*KDM6A* #7: 2,357; sh*Renilla*: 2,272) in comparison to the native state with 40.3% of genes exclusively being deregulated in sh*KDM6A* #7 (Supplementary Fig. 10f). Next, we compared KDM6A

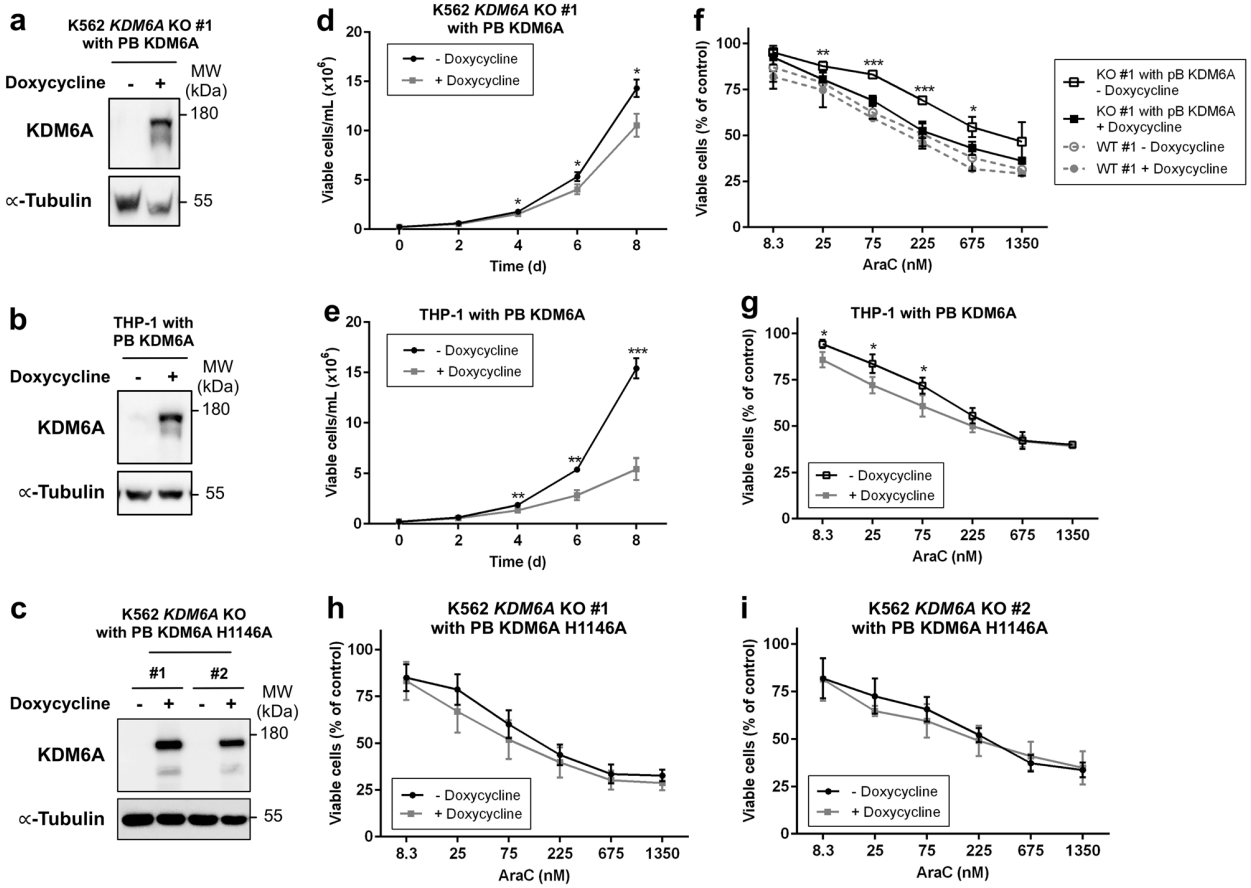



**Fig. 6** Re-expression of KDM6A suppresses cell growth and sensitizes cells to AraC therapy. **a**–**c** Immunoblot showing doxycycline inducible re-expression of KDM6A in K562 *KDM6A* KO #1 PB KDM6A (**a**), *KDM6A* mutant THP-1 PB KDM6A (**b**) and K562 *KDM6A* KO #1 and #2 PB KDM6A H1146A cells (**c**) after treatment with 0.5 µg/mL doxycycline for 48 h. Blot is representative of two independent experiments. MW molecular weight, α-Tubulin loading control. **d**, **e** Proliferation assay measuring the amount of viable K562 *KDM6A* KO #1 PB KDM6A (**d**) and *KDM6A* mutant THP-1 PB KDM6A cells (**e**) in the absence or presence of doxycycline (0.5 µg/mL) every 2 days for 8 days. Mean ± s.d. are given for three independent experiments. Unpaired, two-tailed Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. **f**–**i** AraC dose-response analysis after treatment for 72 h in K562 *KDM6A* KO #1 PB KDM6A and K562 *KDM6A* WT #1 (**f**), *KDM6A* mutant THP-1 PB KDM6A (**g**) and K562 *KDM6A* KO PB KDM6A H1146A cells (**h**,**i**) in the absence or presence of doxycycline (0.5 µg/mL). Mean ± s.d. are given for at least three independent experiments. Unpaired, two-tailed Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$

regulated genes with known candidate genes in drug metabolic pathways and found that *ENT1* was consistently downregulated in KDM6A KD cells in both RNA-Seq screenings (Fig. 7a, Supplementary Fig. 10d). ENT1, or SLC29A1, is a membrane transporter important for the cellular uptake of nucleosides and its analogues [26]. Sh*KDM6A* K562 cells showed significantly reduced *ENT1* mRNA compared to controls (Fig. 7b). AraC treatment slightly increased *ENT1* mRNA expression in sh*Renilla* cells. In sh*KDM6A* #7 cells *ENT1* expression was reduced even after AraC administration (Fig. 7b). Additionally, decreased ENT1 expression was detected in *KDM6A* KO K562 (Supplementary Fig. 11a) and MM-6 cells (Fig. 7c). Treatment of K562, MM-1, and MM-6 cells with a selective ENT1 inhibitor, NBMPR, in combination with AraC resulted in increased cell survival compared to AraC alone (Fig. 7d, Supplementary Fig. 11b, c). By contrast, no change was observed when we combined ENT1 inhibition with DNR or 6-TG in K562 cells (Supplementary Fig. 11d, e). To further elucidate the mechanism of ENT1 regulation by KDM6A, we performed ChIP-seq analysis for H3K27me3 and H3K27ac in MM-1 and MM-6 cells as recent studies have reported that the tumor suppressor effect is largely demethylase independent [19, 27]. ChIP-seq for H3K27me3 showed no enrichment on the *ENT1* locus, however, we detected differential H3K27ac peaks in the promoter and a putative enhancer region of ENT1 in MM-1 compared to MM-6 (Supplementary Fig. 12). Additionally, K562 and THP-1 cells with loss of KDM6A showed low H3K27ac peaks on the *ENT1* locus, which were increased

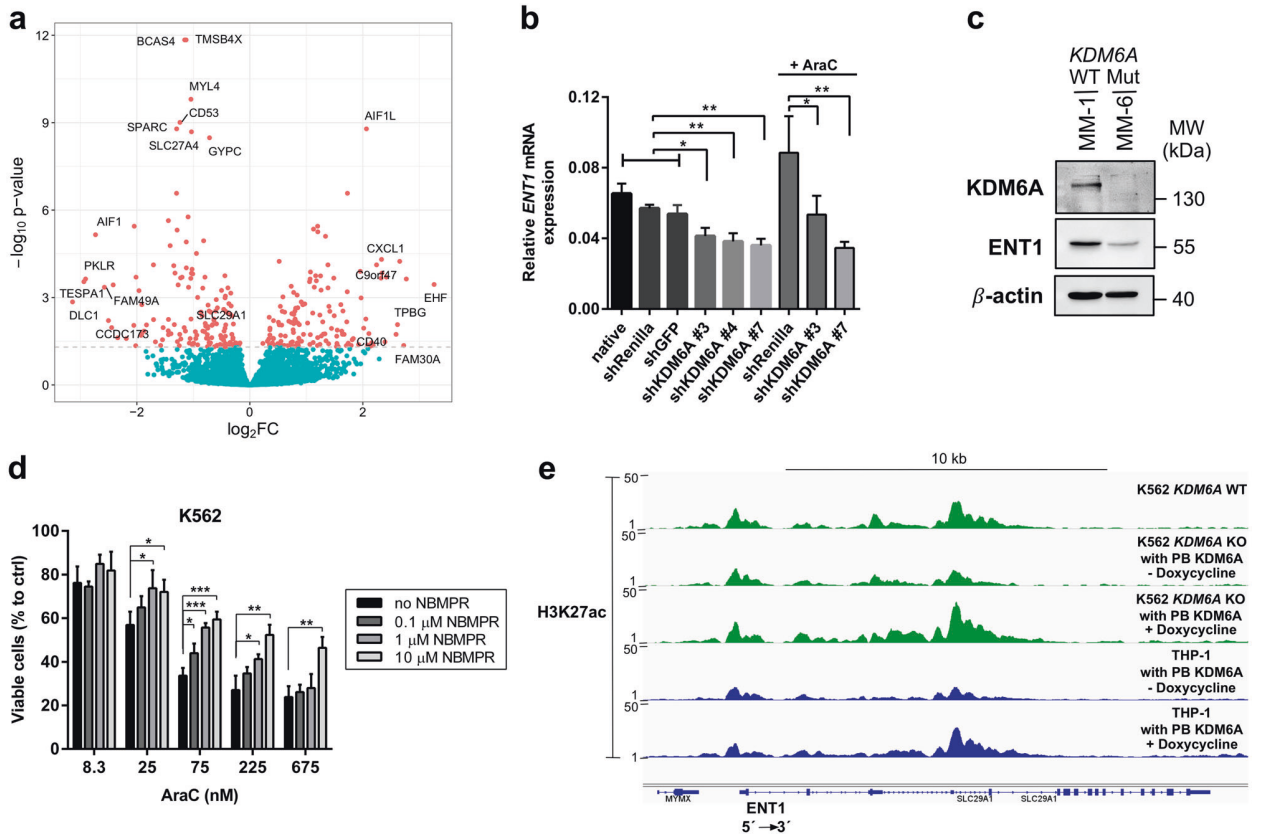



**Fig. 7** Downregulation of ENT1 by KDM6A loss is regulated by H3K27 acetylation and increases AraC resistance. **a** Volcano plot showing $\log_2$ fold change on the x-axis and adjusted *P* value on the y-axis for the differential gene expression between shRNA-mediated KD of *KDM6A* (sh*KDM6A* #7) and control (sh*GFP*) in K562 cells ($n = 6$). Genes with adjusted *P* value < 0.05 are highlighted in red and those with a $\log_2$ FC >2.5 or <−2.5 are labeled with the gene name. In addition, genes with adjusted *P* value < 1e-8 and the gene *SLC29A1* are labeled. **b** qRT-PCR for *ENT1* in K562 cells: native, sh*Renilla*, sh*GFP* and three different sh*KDM6A* KDs. *ENT1* mRNA for sh*Renilla*, sh*KDM6A* #3 and sh*KDM6A* #7 K562 cells is also shown after treatment with 150 nM AraC for 72 h. The mean ± s.d. for *ENT1* mRNA relative to *GAPDH* for three independent experiments is shown. **c** Immunoblotting showing strong reduction of ENT1 protein in *KDM6A* mutant MM-6 cells compared to the *KDM6A* WT sister cell line MM-1. Immunoblots are representative of three independent experiments. MW, molecular weight; β-actin, loading control. **d** Inhibition of ENT1 by NBMPR increases the amount of viable cells during AraC treatment. K562 cells were treated with different AraC concentrations in combination with 0, 0.1, 1 and 10 μM of NBMPR for 72 h. Mean ± s.d are given for three independent experiments. Unpaired, two-tailed Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. **e** Genomic snapshot of H3K27ac ChIP-seq in K562 *KDM6A* WT, K562 *KDM6A* KO PB KDM6A and *KDM6A* mutant THP-1 cells in the absence or presence of doxycycline at the *ENT1* locus. Cells were treated with media +/− doxycycline (0.5 μg/mL) every 24 h for 72 h

upon doxycycline inducible re-expression of KDM6A in both cell lines (Fig. 7e). In summary, our data demonstrate that strong reduction or complete loss of KDM6A decreases ENT1 expression, probably through direct or indirect effects of KDM6A on enhancer regions, which then promotes increased AraC resistance.

## Discussion

In cancer, except for bladder cancer [28], the frequency of mutation in the *KDM6A* gene is rather low [29, 30]. In AML, frequency ranges from 0.7% [31] to 4.0% [3, 32, 33] and the majority of mutations are missense mutations. In our study, *KDM6A* mutations were exclusively missense and truncating mutations. One of these mutations, E1325X, has been previously described in an AML patient at diagnosis and was present in a subclone only [31]. Reduced expression or mutations/deletions of KDM6A correlate with poor overall survival in patients with CN-AML [3] or myeloma [29], respectively. We identified three AML patients harboring *KDM6A* mutations at diagnosis and observed an outgrowth of the *KDM6A* mutated population at relapse. These results are in agreement with our previous study [3] and suggest that *KDM6A* loss may contribute to increased chemo-resistance in AML.

As *KDM6A* is not X-inactivated [34], females with T-ALL benefit from two functional copies [17], and compared

to female patients shortened relapse-free survival is seen in male CN-AML patients [3]. Dependent on expression dosage, KDM6A deficiency was found to accelerate and promote cancer progression in a mouse lymphoma model [35]. We observed that KDM6A mRNA and protein expression is reduced in 45.7% and 44.4% of AML patients at relapse, respectively. As most of these samples showed no *KDM6A* mutation, another mechanism contributing to KDM6A regulation might be DNA methylation. We found that AML patients with high DNA methylation levels of KDM6A have a significantly shorter overall survival. These results are in line with our previous study [3] showing a correlation of low KDM6A expression and poor overall survival in CN-AML patients.

In agreement with our data showing higher AraC resistance in *KDM6A* mutant AML cell lines, we observed that a *KDM6A* mutant PDX sample is less sensitive towards in vitro AraC treatment compared to a *KDM6A* WT PDX sample. Prominent differences in treatment response during in vivo therapy indicate that even small differences in drug sensitivity observed in vitro can significantly impact long-term in vivo therapy.

We have recently demonstrated that *KDM6A* mutant MM-6 are less sensitive to AraC than MM-1 cells [3]. Under competitive growth conditions, we observed a selective growth advantage of MM-6 during AraC treatment. Furthermore, inducible re-expression of KDM6A in *KDM6A*-null cells sensitized to AraC treatment. Very recently, Gozdecka et al. [19] could show that lentiviral expression of KDM6A in MM-6 cells suppressed spontaneous cell proliferation. The data presented here extend these findings and show that KDM6A acts as a tumor suppressor and mediates drug resistance in AML.

UTY, a catalytically inactive KDM6A homolog that is encoded on the Y chromosome, was recently shown to suppress myeloid leukemogenesis in *KDM6A*-deficient male mice [19]. It was also reported to be lost or reduced in conjunction with *KDM6A* mutations in pancreatic cancers with squamous differentiation in male patients [27]. We found that UTY mRNA expression was lost or barely detectable in MM-6 and MM-1 cells suggesting that the drug resistant phenotype of MM-6 is not affected by UTY.

Various mechanisms of drug resistance in AML have been proposed in the last decades. AraC for instance can only exert its cytotoxic effect after cellular uptake and conversion into its active form. Among the key candidate genes in AraC metabolism, we consistently found differential expression of the drug influx transporter gene *ENT1*. We demonstrate that KD/KO of *KDM6A* leads to decreased expression of ENT1 linking decreased KDM6A levels to increased AraC resistance. Previous studies showed that KD or inhibition of ENT1 in AML cell lines confers AraC resistance [36, 37]. In AraC based therapy, AML patients with low ENT1 levels were reported to have shorter disease-free or overall survival [38]. In agreement with previous results, inhibition of ENT1 by NBMPR increased AraC resistance. Furthermore, we demonstrate that transport of DNR and 6-TG across the cell membrane is ENT1 independent. Recently, KDM6A was reported to regulate gene expression during myeloid leukemogenesis mainly by modifying levels of H3K27 acetylation, H3K4 monomethylation and chromatin accessibility [19]. Our ChIP-seq analysis suggests that ENT1 expression is regulated by H3K27 acetylation whereas H3K27 demethylase activity is dispensable for ENT1 expression. As re-expression of a catalytically dead mutant in K562 *KDM6A* KO cells had not the same effect as *KDM6A* WT in re-sensitizing cells to AraC, it remains to be determined if disruption of the catalytic domain impacts the regulation of ENT1 expression.

Prolonging the treatment time for DNR, resulted in a significant increase in DNR resistance in MM-6 cells compared to a shorter treatment as previously described [3]. Additionally, we demonstrate that deletion of *KDM6A* in MM-1 recapitulates the same drug resistant phenotype observed in MM-6. The mechanisms leading to DNR resistance need further investigation, but loss of KDM6A-mediated upregulation of the DNR metabolizing enzymes AKR1C1 and AKR1C2 (data not shown) might be involved in DNR resistance. A study demonstrated that upregulation of AKR1C1/3 facilitated reduction of DNR efficacy in leukemic U937 cells [39].

Taken together, our results show that KDM6A inactivation either by loss-of-function mutations or protein downregulation mediates drug resistance in AML.

**Acknowledgements** The authors thank all study participants. We thank M. Fritschle for animal handling. KHM, PAG, GS, HL, IJ, and KS were supported by the German Research Council (DFG) within the SFB 1243 "Cancer Evolution". SMS is a fellow of the Life Science Munich (LSM) graduate school. MDB is a fellow of IMPRS-LS. SMS, RM, ALH, SW, HDM and JK are members of the IRTG-1243 within the SFB 1243.

**Author contributions** SMS designed and performed the experiments, collected, analyzed and interpreted the data, and wrote the manuscript; ALH, RM, MDB, BT, JK, and HDM performed the experiments, collected, and analyzed the data. SW performed and analyzed the RNA-seq experiments. HDM performed and analyzed the numerical aberration analysis. MC and WHL designed and established the knockdown system. BV and IJ established and provided serially transplantable PDX AML samples, and performed and analyzed in vivo therapy trial. MRT, KM, SV, and PAG were involved in the analysis of patient data. HL supervised the study. KS designed, interpreted the data and supervised the study. All authors discussed the results and commented on the manuscript.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

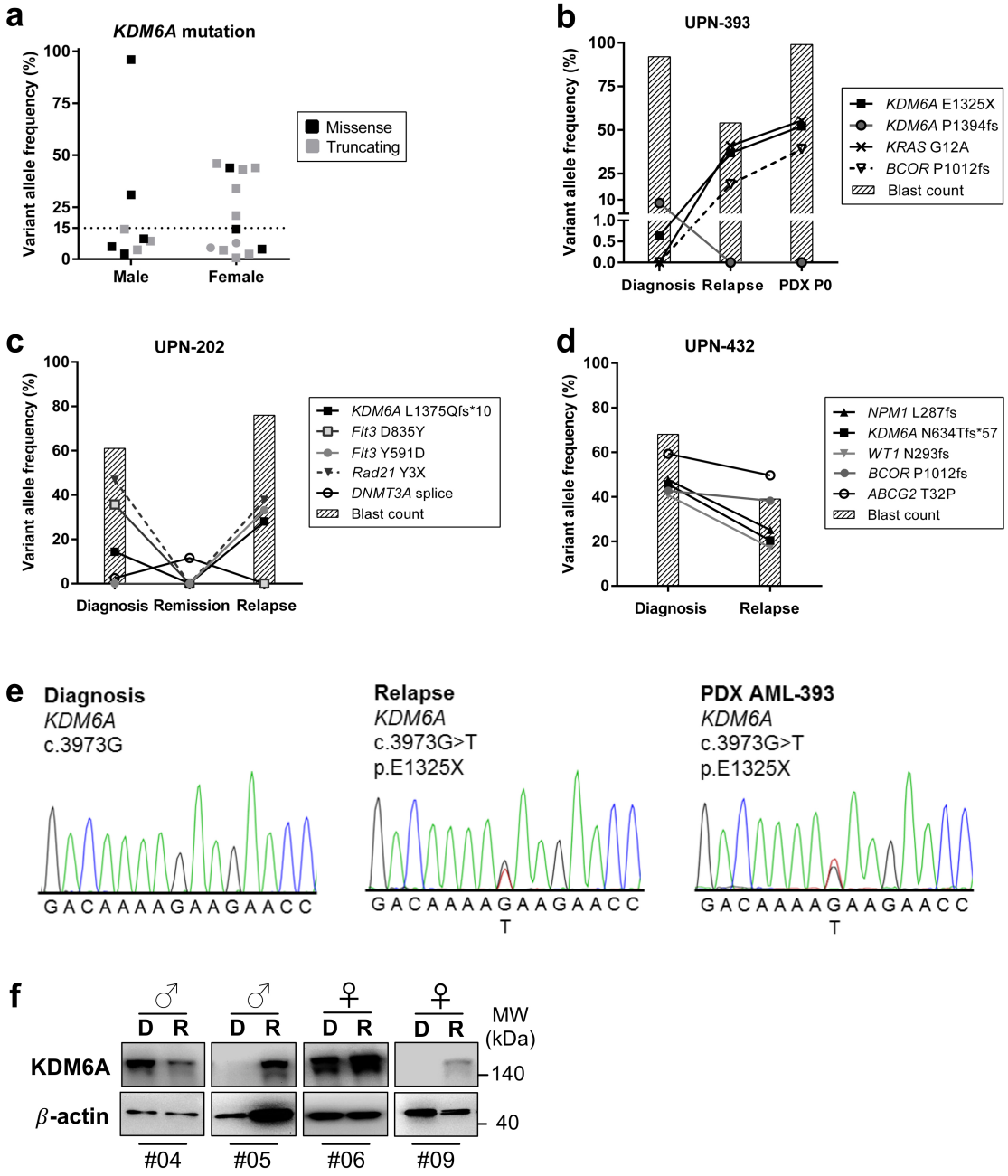
1. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129:424–47. https://doi.org/10.1182/blood-2016-08-733196.
2. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. N Engl J Med. 2015;373:1136–52. https://doi.org/10.1056/NEJMra1406184.
3. Greif PA, Hartmann L, Vosberg S, Stief SM, Mattes R, Hellmann I, et al. Evolution of cytogenetically normal acute myeloid leukemia during therapy and relapse: an exome sequencing study of 50 patients. Clin Cancer Res. 2018;24:1716–26. https://doi.org/10.1158/1078-0432.CCR-17-2344.
4. Hong S, Cho Y-W, Yu L-R, Yu H, Veenstra TD, Ge K. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. Proc Natl Acad Sci USA. 2007;104:18439–44. https://doi.org/10.1073/pnas.0707292104.
5. Agger K, Cloos PAC, Christensen J, Pasini D, Rose S, Rappsilber J, et al. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature. 2007;449:731–4. https://doi.org/10.1038/nature06145.
6. Min GL, Villa R, Trojer P, Norman J, Yan KP, Reinberg D, et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science. 2007;318:447–50. https://doi.org/10.1126/science.1149042.
7. Hu D, Gao X, Morgan MA, Herz H-M, Smith ER, Shilatifard A. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol Cell Biol. 2013;33:4745–54. https://doi.org/10.1128/MCB.01181-13.
8. Dhar SS, Zhao D, Lin T, Gu B, Pal K, Wu SJ, et al. MLL4 is required to maintain broad H3K4me3 peaks and super-enhancers at tumor suppressor genes. Mol Cell. 2018;70:825–41.e6. https://doi.org/10.1016/j.molcel.2018.04.028.
9. Froimchuk E, Jang Y, Ge K. Histone H3 lysine 4 methyltransferase KMT2D. Gene. 2017;627:337–42. https://doi.org/10.1016/j.gene.2017.06.056.
10. Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res. 2010;38:4958–69. https://doi.org/10.1093/nar/gkq244.
11. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA. 2010;107:21931–6. https://doi.org/10.1073/pnas.1016071107.
12. van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nat Genet. 2009;41:521–3. https://doi.org/10.1038/ng.349.
13. Ross JS, Badve S, Wang K, Sheehan CE, Boguniewicz AB, Otto GA, et al. Genomic profiling of advanced-stage, metaplastic breast carcinoma by next-generation sequencing reveals frequent, targetable genomic abnormalities and potential new treatment options. Arch Pathol Lab Med. 2015;139:642–9. https://doi.org/10.5858/arpa.2014-0200-OA.
14. Nickerson ML, Dancik GM, Im KM, Edwards MG, Turan S, Brown J, et al. Concurrent alterations in TERT, KDM6A, and the BRCA pathway in bladder cancer. Clin Cancer Res. 2014;20:4935–48. https://doi.org/10.1158/1078-0432.CCR-14-0330.
15. Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun. 2014;5:1–7. https://doi.org/10.1038/ncomms4630.
16. Ntziachristos P, Tsirigos A, Welstead GG, Trimarchi T, Bakogianni S, Xu L, et al. Contrasting roles of histone 3 lysine 27 demethylases in acute lymphoblastic leukaemia. Nature. 2014;514:513–7. https://doi.org/10.1038/nature13605.
17. Van der Meulen J, Sanghvi V, Mavrakis K, Durinck K, Fang F, Matthijssens F, et al. The H3K27me3 demethylase UTX is a gender-specific tumor suppressor in T-cell acute lymphoblastic leukemia. Blood. 2015;125:13–22. https://doi.org/10.1182/blood-2014-05-577270.H.-G.W.
18. Bolli N, Manes N, McKerrell T, Chi J, Park N, Gundem G, et al. Characterization of gene mutations and copy number changes in acute myeloid leukemia using a rapid target enrichment protocol. Haematologica. 2015;100:214–22. https://doi.org/10.3324/haematol.2014.113381.
19. Gozdecka M, Meduri E, Mazan M, Tzelepis K, Dudek M, Knights AJ, et al. UTX-mediated enhancer and chromatin remodeling suppresses myeloid leukemogenesis through noncatalytic inverse regulation of ETS and GATA programs, Nat Genet2018;1–12. https://doi.org/10.1038/s41588-018-0114-z.
20. Vick B, Rothenberg M, Sandhöfer N, Carlet M, Finkenzeller C, Krupka C, et al. An advanced preclinical mouse model for acute myeloid leukemia using patients' cells of various genetic subgroups and in vivo bioluminescence imaging. PLoS ONE. 2015;10:e0120925. https://doi.org/10.1371/journal.pone.0120925.
21. Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. Cancer Cell. 2010;17:13–27. https://doi.org/10.1016/j.ccr.2009.11.020.
22. Early AP, Preisler HD, Slocum H, Rustum YM. A pilot study of high-dose 1-beta-D-arabinofuranosylcytosine for acute leukemia and refractory lymphoma: clinical response and pharmacology. Cancer Res. 1982;42:1587–94. http://www.ncbi.nlm.nih.gov/pubmed/6949642. Accessed 6 June 2018.
23. Bogason A, Quartino AL, Lafolie P, Masquelier M, Karlsson MO, Paul C, et al. Inverse relationship between leukaemic cell burden and plasma concentrations of daunorubicin in patients with acute myeloid leukaemia. Br J Clin Pharmacol. 2011;71:514–21. https://doi.org/10.1111/j.1365-2125.2010.03894.x.
24. MacLeod RA, Voges M, Drexler HG. Mono Mac 6: a mature monoblastic leukemia cell line with t(9;11)(p21; q23). Blood. 1993;82:3221–2. http://www.ncbi.nlm.nih.gov/pubmed/7695659. Accessed 6 June 2018.
25. Wang C, Lee J-E, Cho Y-W, Xiao Y, Jin Q, Liu C, et al. UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. Proc Natl Acad Sci USA. 2012;109:15324–9. https://doi.org/10.1073/pnas.1204166109.

26. M Pastor-Anglada, S Perez-Torras, Nucleoside transporter proteins as biomarkers of drug responsiveness and drug targets, Front. Pharmacol. 2015; 6. https://doi.org/10.3389/fphar.2015.00013.
27. Andricovich J, Perkail S, Kai Y, Casasanta N, Peng W, Tzatsos A. Loss of KDM6A activates super-enhancers to induce gender-specific squamous-like pancreatic cancer and confers sensitivity to BET inhibitors. Cancer Cell. 2018;33:512–26.e8. https://doi.org/10.1016/j.ccell.2018.02.003.
28. Ler LD, Ghosh S, Chai X, Thike AA, Heng HL, Siew EY, et al. Loss of tumor suppressor KDM6A amplifies PRC2-regulated transcriptional repression in bladder cancer and can be targeted through inhibition of EZH2. Sci Transl Med. 2017;9:1–14. https://doi.org/10.1126/scitranslmed.aai8312.
29. Pawlyn C, Kaiser MF, Heuck C, Melchor L, Wardell CP, Murison A, et al. The spectrum and clinical impact of epigenetic modifier mutations in Myeloma. Clin Cancer Res. 2016;22:5783–94. https://doi.org/10.1158/1078-0432.CCR-15-1790.
30. J-F Spinella, P Cassart, C Richer, V Saillour, M Ouimet, S Langlois, et al. Genomic characterization of pediatric T-cell acute lymphoblastic leukemia reveals novel recurrent driver mutations, Oncotarget. 2016;7. https://doi.org/10.18632/oncotarget.11796.
31. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med. 2016;374:2209–21. https://doi.org/10.1056/NEJMoa1516192.
32. Metzeler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Dennis G, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood. 2016;128:686–99. https://doi.org/10.1182/blood-2016-01-693879.Presented.
33. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368:2059–74. https://doi.org/10.1056/NEJMoa1301689.
34. Greenfield A, Carrel L, Pennisi D, Philippe C, Quaderi N, Siggers P, et al. The UTX gene escapes X inactivation in mice and humans. Hum Mol Genet. 1998;7:737–42. https://doi.org/10.1093/hmg/7.4.737.
35. Li X, Zhang Y, Zheng L, Liu M, Chen CD, Jiang H. UTX is an escape from X-inactivation tumor-suppressor in B cell lymphoma. Nat Commun. 2018;9:2720. https://doi.org/10.1038/s41467-018-05084-w.
36. Kim JH, Lee C, Cheong HS, Koh Y, Ahn KS, Kim HL, et al. SLC29A1 (ENT1) polymorphisms and outcome of complete remission in acute myeloid leukemia. Cancer Chemother Pharmacol. 2016;78:533–40. https://doi.org/10.1007/s00280-016-3103-x.
37. Macanas-Pirard P, Broekhuizen R, González A, Oyanadel C, Ernst D, García P, et al. Resistance of leukemia cells to cytarabine chemotherapy is mediated by bone marrow stroma, involves cell-surface equilibrative nucleoside transporter-1 removal and correlates with patient outcome. Oncotarget. 2017;8:23073–86. https://doi.org/10.18632/oncotarget.14981.
38. Wan H, Zhu J, Chen F, Xiao F, Huang H, Han X, et al. SLC29A1 single nucleotide polymorphisms as independent prognostic predictors for survival of patients with acute myeloid leukemia: an in vitro study. J Exp Clin Cancer Res. 2014;33:90 https://doi.org/10.1186/s13046-014-0090-9.
39. Matsunaga T, Yamaguchi A, Morikawa Y, Kezuka C, Takazawa H, Endo S, et al. Induction of aldo-keto reductases (AKR1C1 and AKR1C3) abolishes the efficacy of daunorubicin chemotherapy for leukemic U937 cells. Anticancer Drugs. 2014;25:868–77. https://doi.org/10.1097/CAD.0000000000000112.
40. Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z, et al. IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics. 2015;31:3359–61. https://doi.org/10.1093/bioinformatics/btv362.

**Supplemental Figures, Table and Methods**

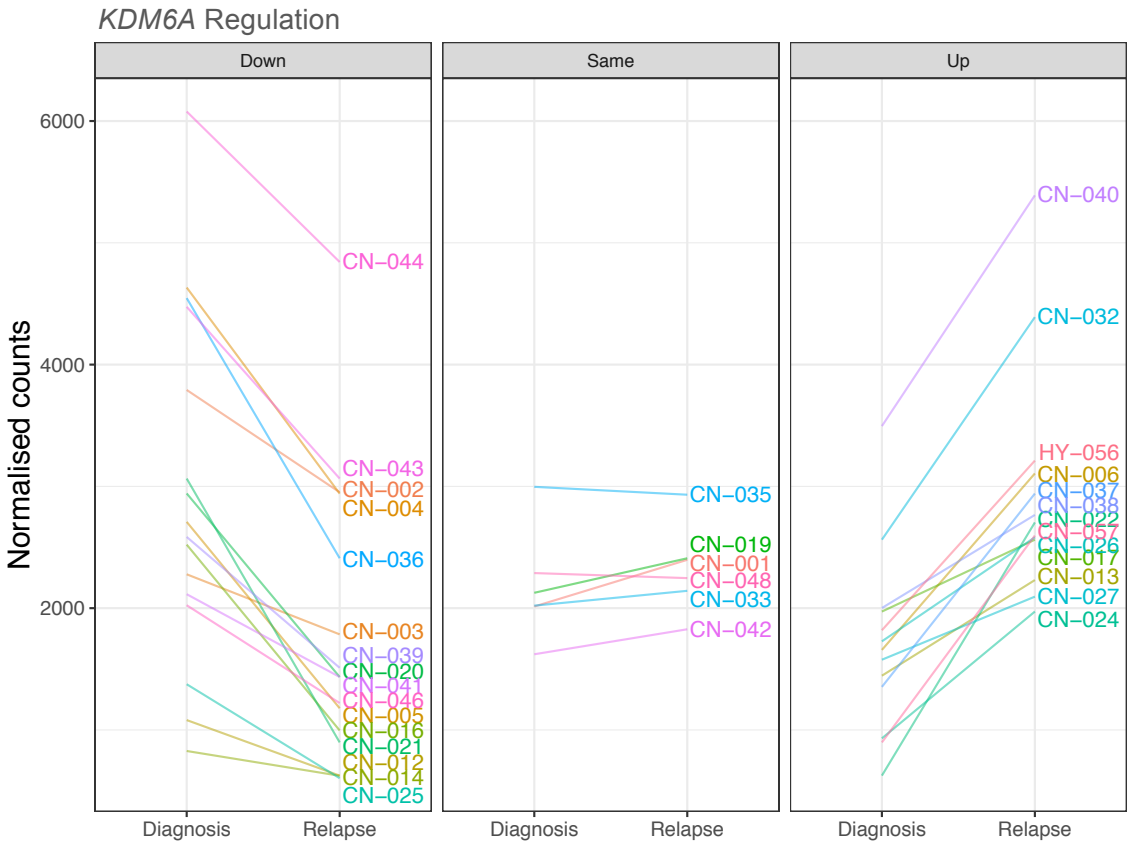
**Loss of KDM6A confers drug resistance in acute myeloid leukemia** Stief *et al*.

**Table of Contents**

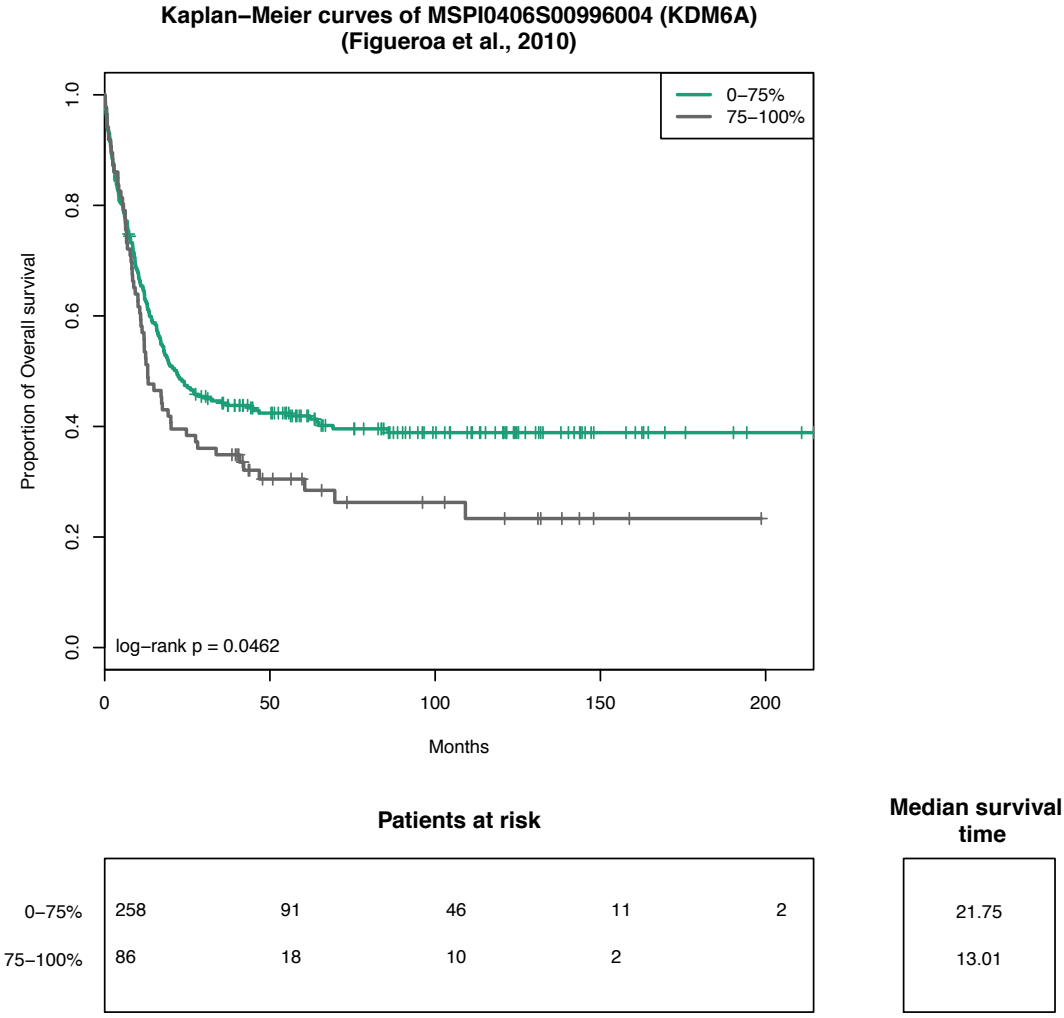



**Supplementary Fig. 1, related to Fig. 1. Outgrowth of *KDM6A* mutations at relapse. a,** Overview of variant allele frequency (VAF) of *KDM6A* mutations at diagnosis in 20 AML patients. The types of mutation including missense and truncating mutations are highlighted with their respective color. VAFs are shown separately for male and female patients and a VAF of 15% is pointed out by a dotted line. Presented *KDM6A* mutations are from AMLCG-99 trial (NCT00266136), AMLCG-2008 trial (NCT01382147), a CN-AML diagnosis-relapse cohort[1] and this work. One female patient harboring two *KDM6A* mutations is shown with a circle. **b,c,d,** VAF plots for different evolutionary patterns observed from diagnosis to relapse in three AML patients. Bars represent the blast count and each line represents one mutation. In addition, for patient UPN-393 (**b**) and UPN-202 (**c**), VAFs of a PDX sample from passage 0 (P0) or a remission sample are shown, respectively. **e,** DNA sequencing chromatogram showing a *KDM6A* mutation in the

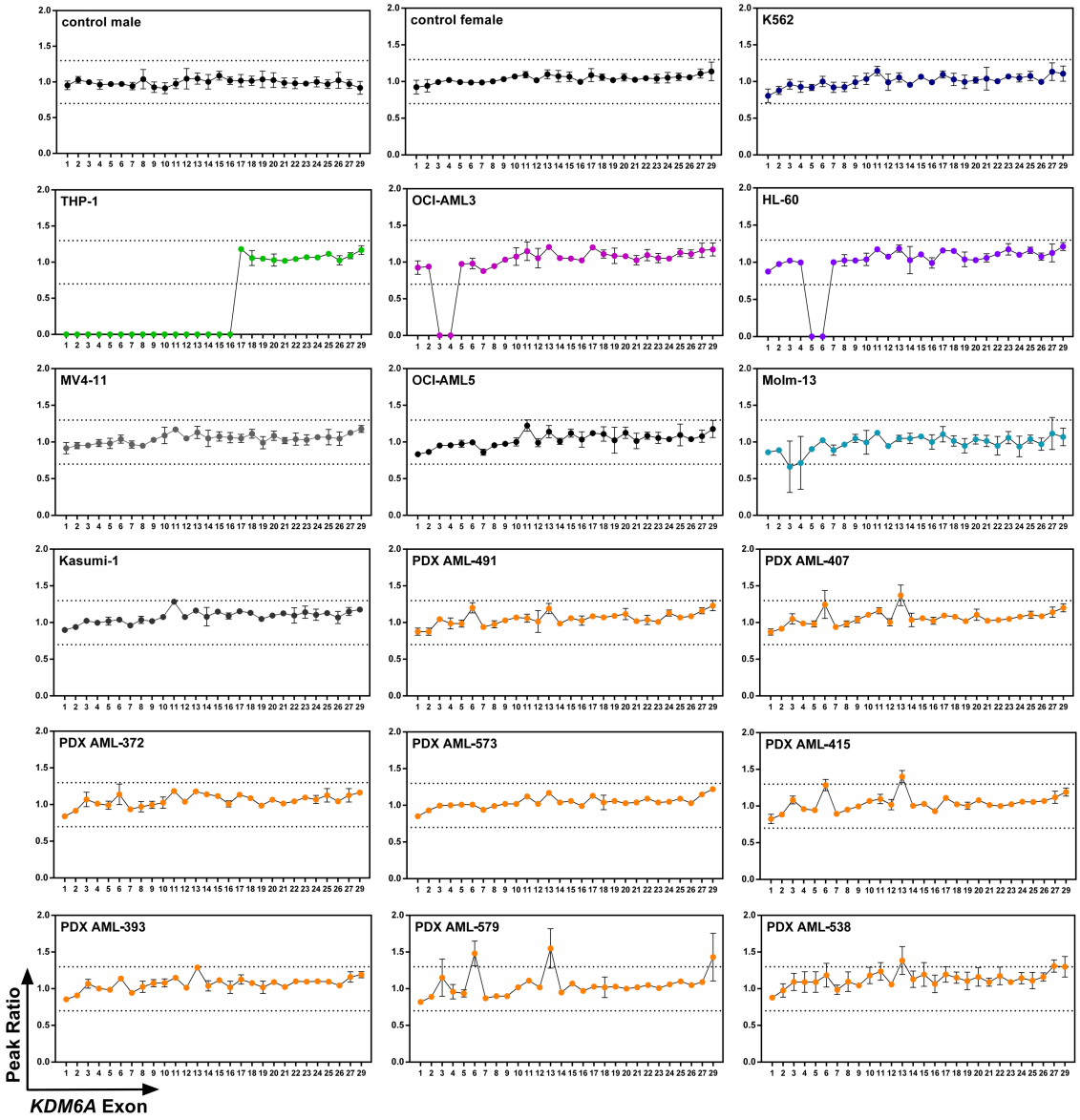
gDNA of a female primary AML patient sample at relapse and in PDX AML-393 cells, established from the primary relapse sample. The mutation is not detectable in diagnosis material of the same patient. **f**, Immunoblotting for KDM6A protein expression in 4 AML patients at diagnosis (D) and relapse (R). Their respective gender is shown on top. The ratio of KDM6A to β-actin expression is shown in Figure 1d. MW, molecular weight; β-actin, loading control.



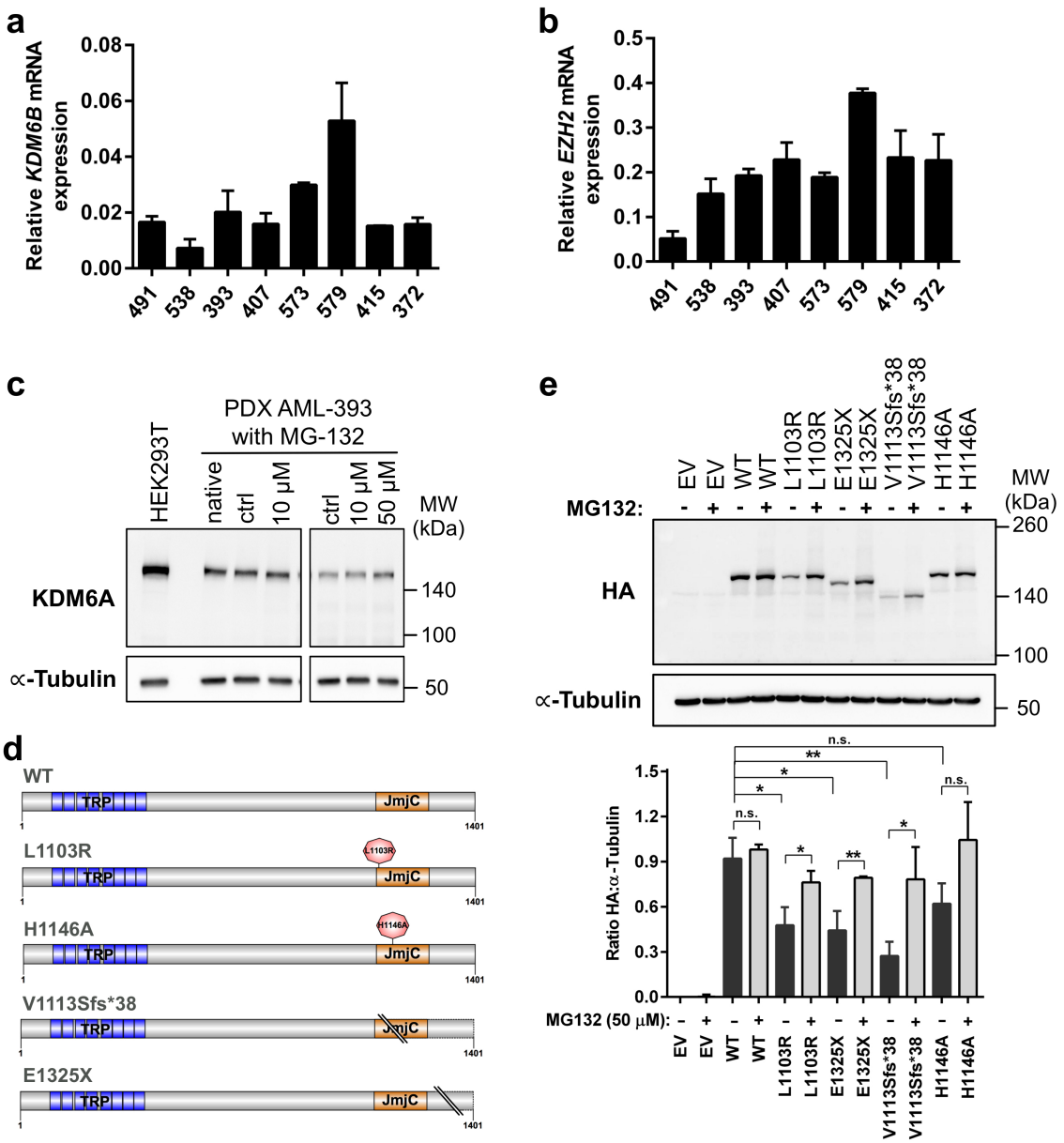

**Supplementary Fig. 2, related to Fig. 1. Regulation of *KDM6A* mRNA expression in 35 CN-AML patients.** The three groups, *KDM6A*-up, *KDM6A*-down and *KDM6A*-same (no change) were defined as a change in expression between diagnosis and relapse of above or below 20% respectively. Regulation of *KDM6A* was determined based on 35 previously described CN-AML patients[1].



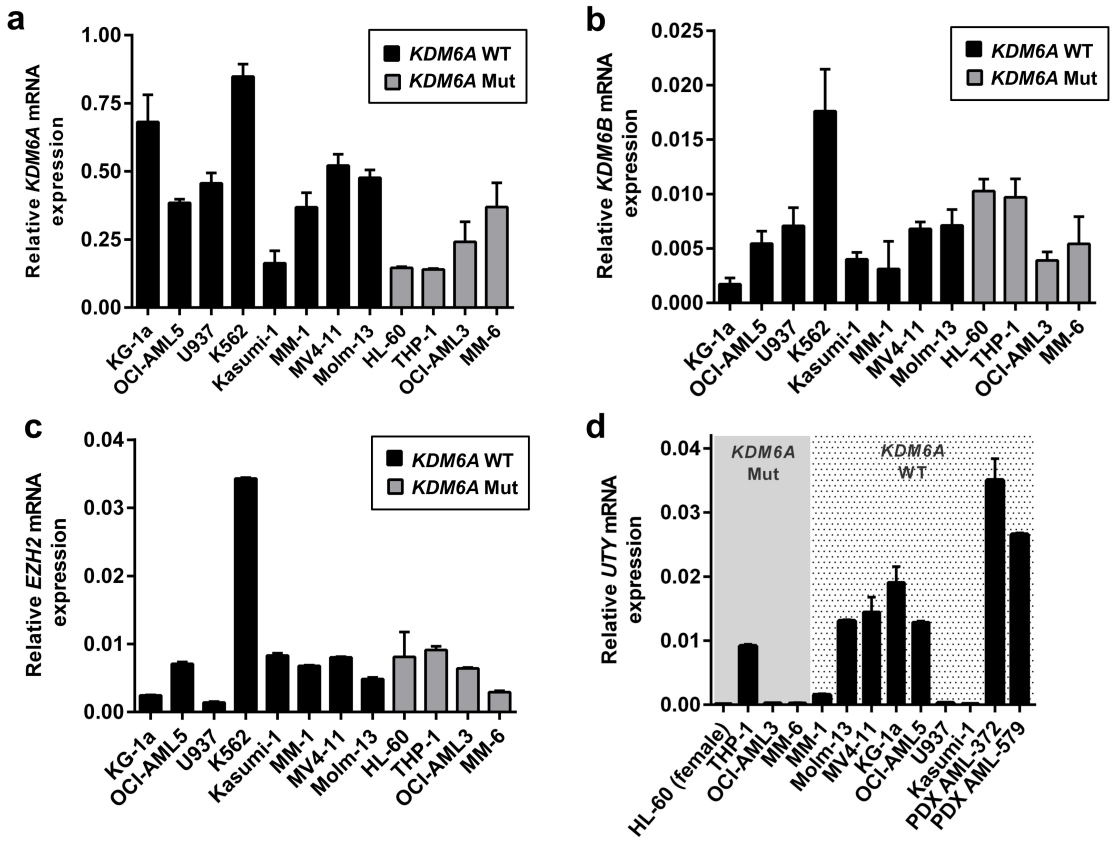

**Supplementary Fig. 3, related to Fig. 1. AML patients with high DNA methylation levels of KDM6A show a shorter overall survival.** The clinical correlation of KDM6A methylation status in 344 previously described AML patients[2] (top 25% versus 0-75%) and overall survival is shown.



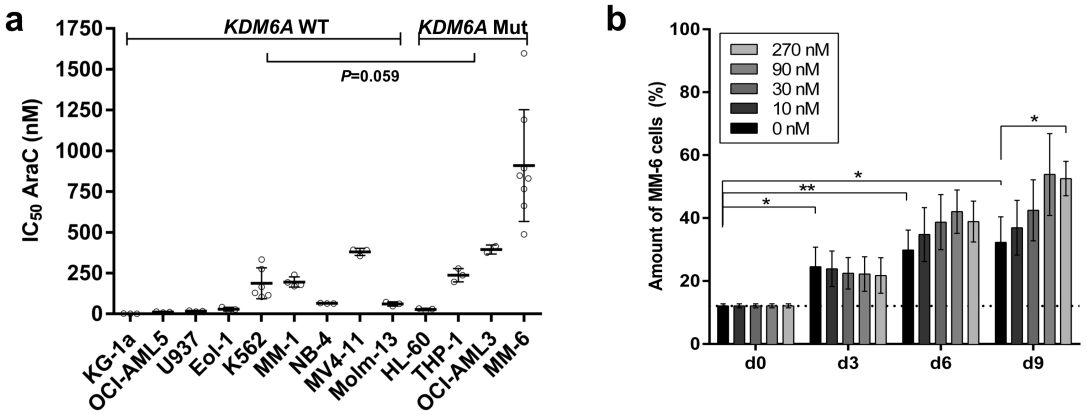

**Supplementary Fig. 4, related to Fig. 2 and Fig. 3. *KDM6A* exon deletions in AML cell lines and AML PDX samples identified.** The peak ratio for each *KDM6A* exon specific probe, detected by quantitative MLPA analysis, is shown. Results for 8/40 of the investigated myeloid cell lines (summarized in **Supplementary Table 1**) and 8/8 PDX AML cells are shown. The area of a normal peak ratio lies within the two dotted lines and ranges from 0.7 to 1.3. Mean ± s.d. are given for at least two independent experiments.



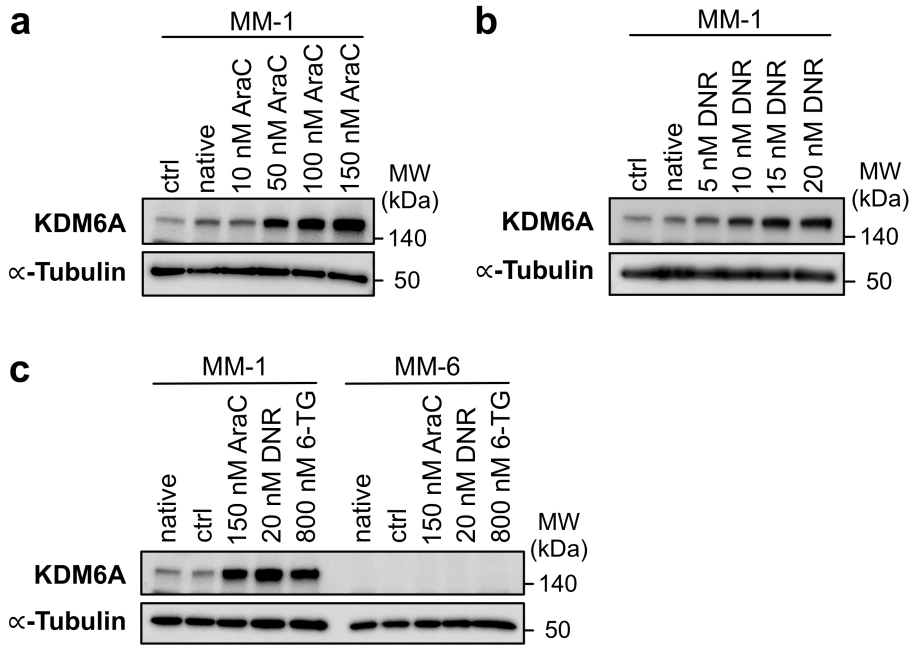

**Supplementary Fig. 5, related to Fig. 2. *KDM6A* mutations are prone to degradation.** **a,b,** qRT-PCR for *KDM6B* (**a**) and *EZH2* (**b**) in PDX relapsed AML cells. The mean ± s.d. relative to the endogenous control *GAPDH* for three experiments is shown. **c,** Immunoblot showing KDM6A expression in PDX AML-393 cells after treatment with the proteasomal inhibitor MG-132 for 6 hours. MW, molecular weight; α-Tubulin, loading control. **d,** Schematic overview of *KDM6A* WT and mutant structures illustrated with IBS software[3]. Point mutations are highlighted with a red dot. TRP, tetratricopeptide repeat; JmjC, Jumonji C. **e,** Effect of MG132 treatment (50 µM, 6h) on KDM6A expression in HEK293T cells transfected with KDM6A WT or mutants (N-terminal HA tag) is shown. Immunoblots are representative of three independent experiments. The mean ratio of HA: α-Tubulin expression ± s.d is given for three independent experiments. Unpaired, two-tailed Student's *t*-test; **P*<0.05; ***P*<0.01; n.s., not significant. EV, empty vector; MW, molecular weight; α-Tubulin, loading control.



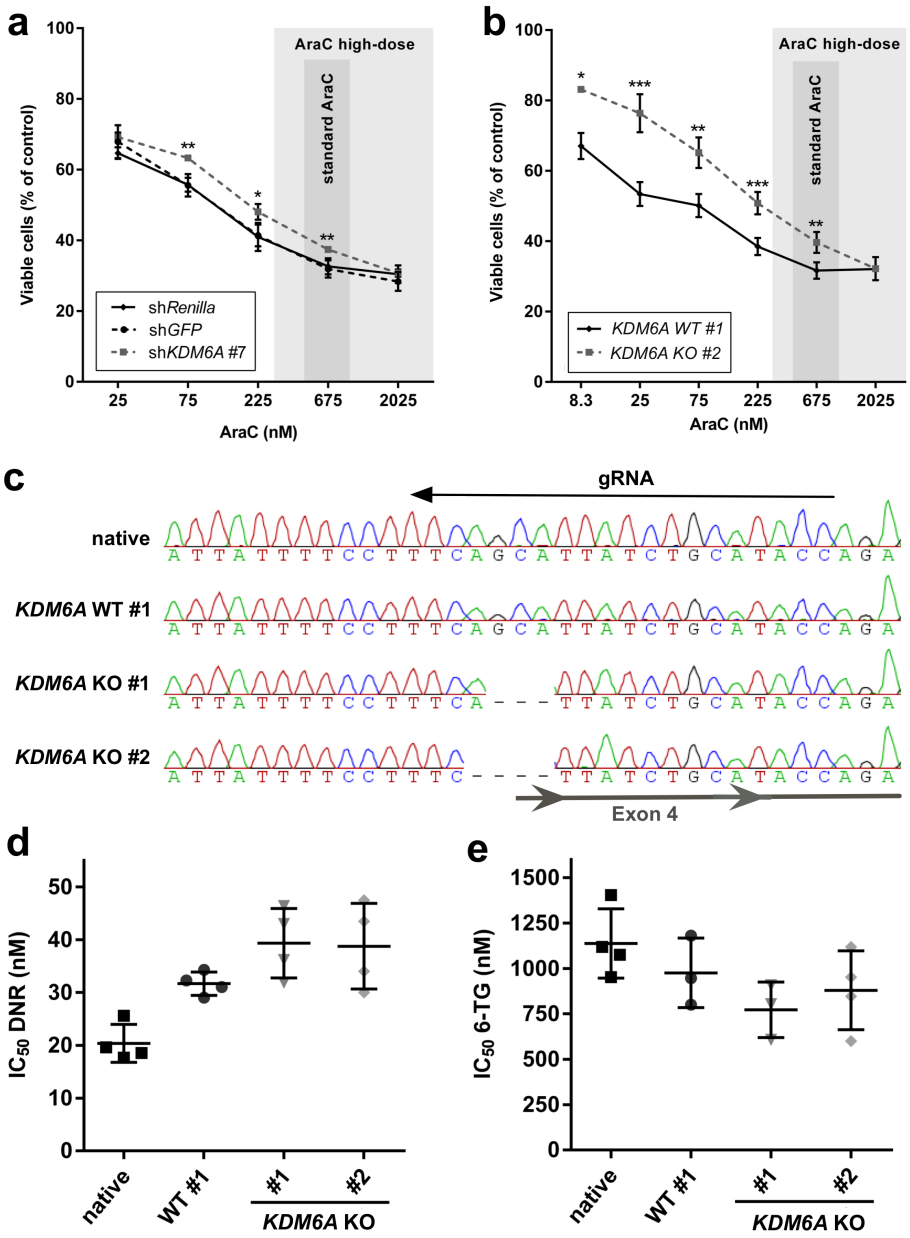

**Supplementary Fig. 6, related to Fig. 3. No changes in *KDM6B* and *EZH2* mRNA expression, but loss of UTY in KDM6A mutant AML cells. a,b,c,d,** qRT-PCR for *KDM6A* (**a**), *KDM6B* (**b**), *EZH2* (**c**), and *UTY* **(d)** in *KDM6A* WT and mutant AML cell lines. Two male PDX relapsed AML samples are additionally shown for *UTY* **(d)**. The mean ± s.d. relative to *GAPDH* for three experiments is shown.



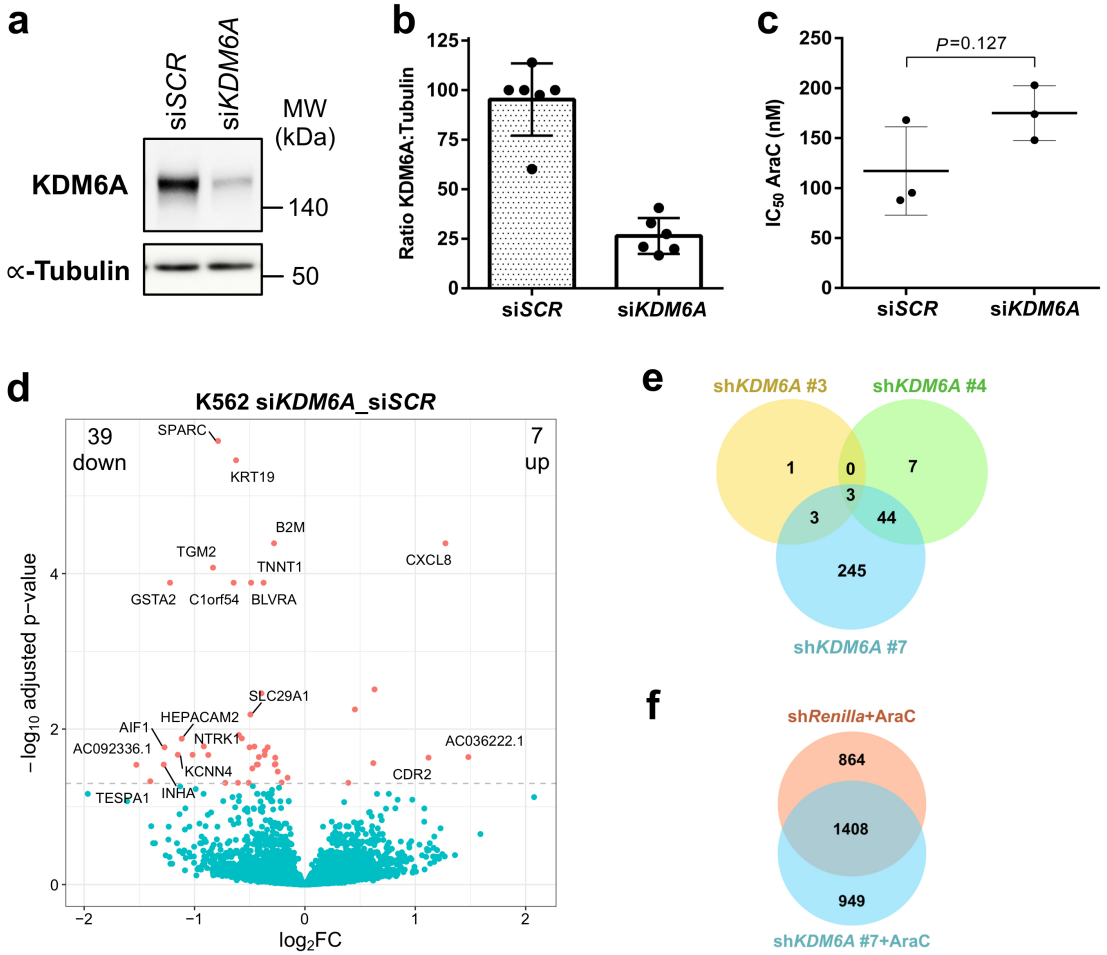

**Supplementary Fig. 7, related to Fig 3. Competitive growth advantage of KDM6A mutant MM-6 cells under native conditions and AraC therapy. a,** Comparison of AraC $IC_{50}$ values between *KDM6A* WT and mutant AML cell lines. Mean ± s.d. from at least three independent experiments are shown. Unpaired, two-tailed Student's *t*-test; *P*=0.059. **b,** Changes in the amount of *KDM6A* mutant MM-6 cells relative to MM-1 cells for a time period of 9 days. MM-6 cells, labeled with Cell Violet, were mixed 1:9 with MM-1 (labeled with Cell Trace CFSE) cells on Day 0 and treated with increasing concentrations of AraC. Mean ± s.d. is given for three independent experiments. Unpaired, two-tailed Student's *t*-test; **P*<0.05; ***P*<0.01.




**Supplementary Fig. 8. Upregulation of KDM6A protein expression after treatment with AraC, DNR and 6-TG. a, b,** Immunoblot showing upregulation of KDM6A protein expression in MM-1 cells with increasing concentrations of AraC (**a**) and DNR (**b**) after 72h. **c,** Immunoblot showing upregulation of KDM6A protein expression in MM-1 cells after treatment with AraC, DNR or 6-TG for 72h. Applied concentrations were approximately the $IC_{50}$ of each drug. *KDM6A* mutant MM-6 were used as a control. ctrl, DMSO control; MW, molecular weight; α-Tubulin, loading control.

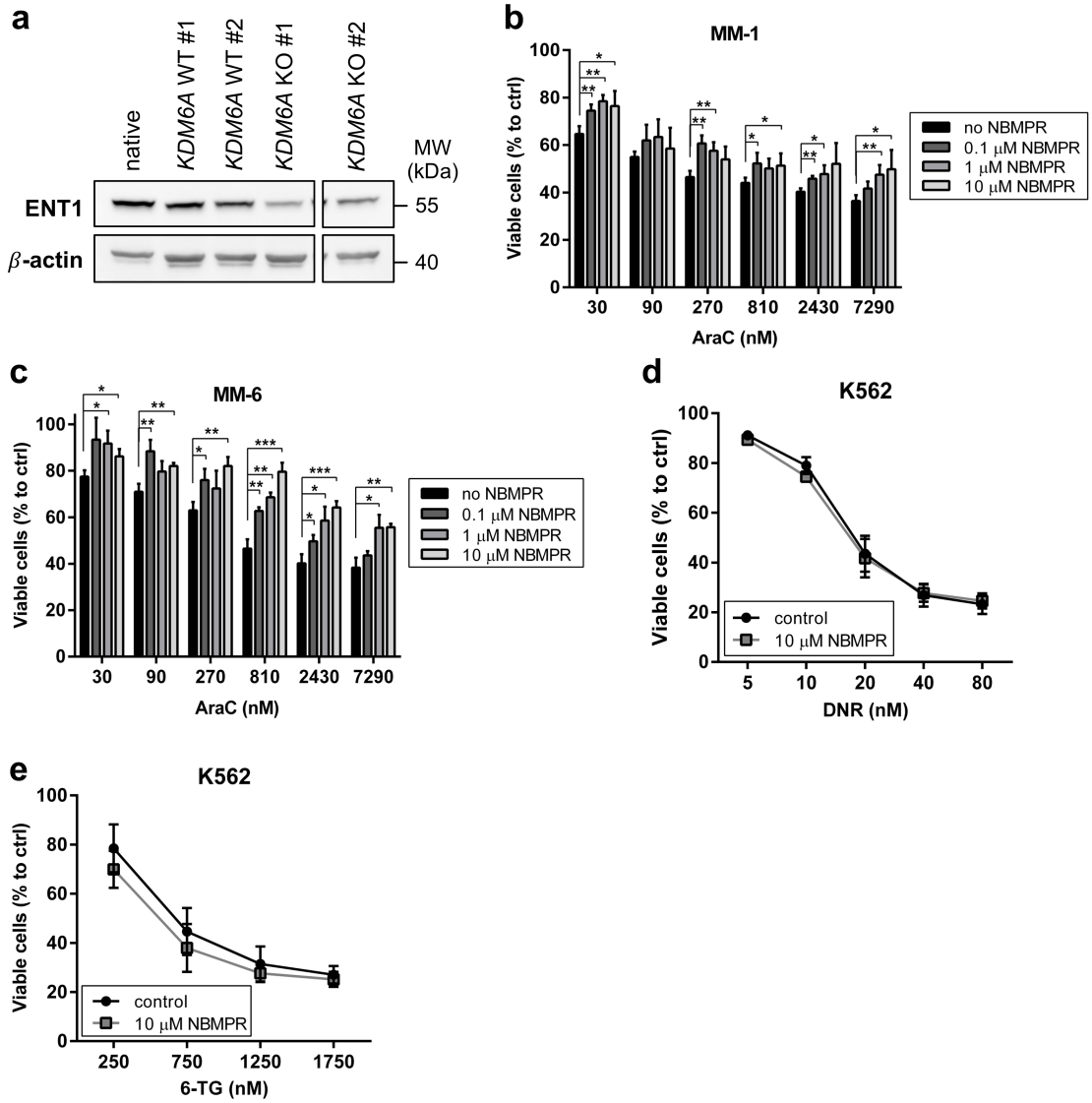



**Supplementary Fig. 9, related to Fig. 4. Loss of *KDM6A* decreases sensitivity to AraC and DNR but not 6-TG. a,b,** AraC dose-response analysis in K562 cells with modified KDM6A expression. Amount of viable cells after treatment with different AraC concentrations for 72h was compared between sh*Control* and sh*KDM6A* #7 cells (**a**) and between WT and *KDM6A* KO cells (**b**). The mean ± s.d. is given for technical duplicates of at least three independent experiments. The area shaded in dark grey (514 nM to 1028 nM) and light grey (226 nM-144 µM) indicates the range of steady-state plasma concentrations measured in patients during standard AraC (100-200mg/m$^2$) and after high-dose AraC (3000mg/m$^2$) treatment, respectively[4]. **c,** DNA sequencing chromatogram showing *KDM6A* frameshift mutations A112Vfs*12 in *KDM6A* KO K562 clone #1 and #2 compared with parental cells and a WT #1 clone. WT #1 clone was tested negative for *KDM6A* KO after CRISPR/Cas9 targeting. **d,e,** Comparison of $IC_{50}$ values for DNR (**d**) and 6-TG (**e**) between K562 control cells (native and WT #1) and *KDM6A* KO clones #1 and #2 (72h treatment). Mean of $IC_{50}$ values ± s.d. ($n$=3-4) are shown.

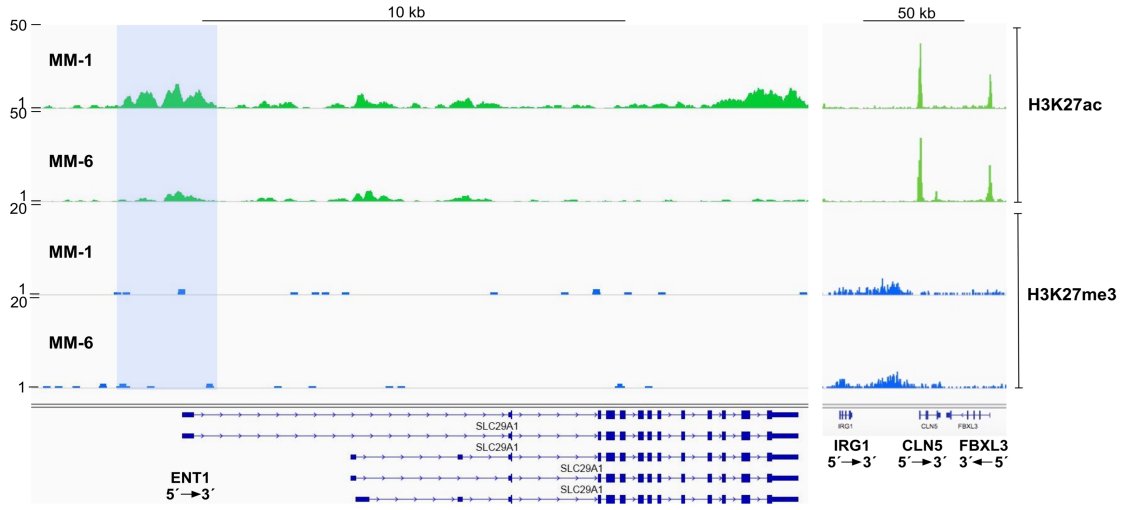



**Supplementary Fig. 10, related to Fig. 7. Effect of *KDM6A* silencing regarding AraC resistance and target gene expression. a,** Immunoblot showing KD of KDM6A expression in K562 cells after two rounds of transfection (2x72h) with siRNA against control (si*SCR*) and *KDM6A* (si*KDM6A*). MW, molecular weight; α-Tubulin, loading control. **b,** Knockdown efficiency of si*KDM6A* compared with control (si*SCR*) is shown for six independent experiments (mean ± s.d.). The ratio of KDM6A to α-Tubulin protein expression was calculated for each experiment and values were normalized to a control. **c,** Pro-proliferative effect of AraC treatment after *KDM6A* KD in K562 cells. 72h after second transfection, K562 cells (si*SCR* and si*KDM6A*) were treated with different concentrations of AraC for 72h. Mean of $IC_{50}$ values from three independent experiments ± s.d. are shown. Unpaired, two-tailed Student's *t*-test; *P*=0.127. **d,** Volcano plot showing $log_2$ fold change on the x-axis and adjusted *P* value on the y-axis for the differential gene expression between siRNA-mediated knockdown of *KDM6A* (si*KDM6A*) and control (si*SCR*) in K562 cells (*n*=6). Genes with adjusted *P* value <0.05 are highlighted in red and those with a $log_2$ FC >1 or <-1 are labeled with the gene name. In addition, genes with adjusted *P* value < 0.01 and the gene *SLC29A1* are labeled. **e,** Overlap between differentially expressed genes (*P*<0.05) in three different sh*KDM6A* KD K562 cells. Differentially expressed genes of each sh*KDM6A* (#3, #4, and #7) compared to sh*GFP* control K562 cells are shown (*n*=6). **f,** AraC specific gene expression changes (*P*<0.05) in sh*KDM6A* #7 K562 cells compared with sh*Renilla* control K562 cells (*n*=6). Differentially

expressed genes in AraC (150 nM, 72h) treated samples were acquired by comparison with the respective untreated samples.




**Supplementary Fig. 11, related to Fig. 7. Inhibition of ENT1 promotes resistance to AraC but not DNR and 6-TG. a,** Immunoblot showing reduction of ENT1 expression in *KDM6A* KO K562 clones #1 and #2 compared with parental and WT clones. Both images, which have been separated, are from the same gel. MW, molecular weight; β-actin, loading control **b,c,** Inhibition of ENT1 by NBMPR increases the amount of viable cells during AraC treatment. MM-1 (**b**) and MM-6 (**c**) cells were treated with different AraC concentrations in combination with 0, 0.1, 1 and 10 μM of NBMPR for 72h and viable cells relative to untreated cells are shown. Mean ± s.d. are given for three independent experiments. Unpaired, two-tailed Student's *t*-test; **P*<0.05; ***P*<0.01; ****P*<0.001. **d,e,** Proliferative effect of ENT1 inhibition by NBMPR (10μM) compared with control (no inhibitor) in combination with different concentrations of DNR (**d**) and 6-TG (**e**) in K562 cells. The mean ± s.d. for three independent experiments is given.




**Supplementary Fig. 12, related to Fig. 7. *KDM6A* mutant MM-6 cells show decreased H3K27ac peaks at the *ENT1* locus.** Genomic snapshot of H3K27ac and H3K27me3 ChIP-seq in *KDM6A* WT MM-1 and *KDM6A* mutant MM-6 cells at the *ENT1* locus. As control regions with no obvious changes, the *IRG1*, *CLN5*, and *FBXL3* loci are shown. The light blue tinted box highlights the differential peaks of interest.

**Supplementary Table 1. Summary of analyzed cell lines.** *KDM6A* exon deletions, detected by quantitative MLPA and CytoScan HD Array hybridization analysis, are listed together with name, cell type, and gender.

| Name | Cell Type | Gender | *KDM6A* exon deletion |
|---|---|---|---|
| AP-1060 | AML | male | - |
| Eol-1 | AML | male | - |
| F-36P | AML | male | - |
| FKH-1 | AML | male | - |
| GF-D8 | AML | male | - |
| HEK293T | embryonic kidney | female | n.a. |
| HL-60 | AML | female | del exon 5-6 |
| HNT-34 | AML | female | - |
| HT-93 | AML | male | - |
| K562 | CML in blast crisis | female | - |
| Kasumi-1 | AML | male | - |
| Kasumi-3 | AML | male | - |
| KG-1a | AML | male | - |
| M-07e | acute megakaryoblastic leukemia | female | - |
| ME-1 | AML | male | - |
| MEGAL | acute megakaryoblastic leukemia | n.s. | - |
| MKPL-1 | acute megakaryoblastic leukemia | male | - |
| MONO-MAC-1 (MM-1) | acute monocytic leukemia | male | - |
| MONO-MAC-6 (MM-6) | acute monocytic leukemia | male | del exon 3-10 |
| Molm-13 | AML | male | - |
| Molm14 (sister Molm-13) | AML | male | - |
| Molm16 | AML | female | - |
| MUTZ-2 | AML | male | - |
| MUTZ-3 | acute myelomonocytic leukemia | male | - |
| MV4-11 | acute monocytic leukemia | male | - |
| NB-4 | acute promyelocytic leukemia | female | n.a. |
| OCI-AML1 | AML | female | - |
| OCI-AML3 | AML | male | del exon 3-4 |
| OCI-AML4 | AML | female | - |
| OCI-AML5 | AML | male | - |
| OCI-AML6 | sAML | female | - |
| OCI-M1 | AML | n.s. | - |
| PLB-985 (sister HL-60) | AML | female | - |
| SKM-1 | AML | male | - |
| SKNO-1 | AML | male | - |
| TF-1 | erythroleukemia | male | - |
| THP-1 | acute monocytic leukemia | male | del exon 1-16 |
| U-937 | histiocytic lymphoma | male | - |
| UCSD-AML1 | AML | female | - |
| UOC-M1 | AML | male | - |
| UT-7 | AML | male | - |
| YNH-1 | AML | male | - |

n.a., not analyzed; n.s., not specified.

## Supplementary methods

### CFSE staining

MM-1 and MM-6 cells were labeled with the respective Cell Trace CFSE and Violet Cell Proliferation Kit (Thermo Fisher Scientific). Cell staining was performed as previously described[5]. Cells were then mixed in a 9:1 ratio and incubated with different AraC concentrations. Every 72h, a small portion was analyzed by flow cytometry (FACSCanto II, BD Bioscience, Franklin Lakes, NJ, USA) using 1 µg/mL propidium iodide (Sigma Aldrich, St. Louis, MO, USA) to exclude dead cells.

### DNA constructs

Human KDM6A with N-terminal HA was amplified from pCMV-HA-UTX (Addgene #24168), and cloned into the pcDNA6/V5-His A vector (Thermo Fisher Scientific) using the In-Fusion HD Cloning Plus Kit (Takara Bio, Saint-Germain-en-Laye, France). KDM6A mutants were generated using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA, USA) and correct sequence was confirmed by Sanger sequencing.

### Immunoblotting

Immunoblotting was performed as previously described[1]. Following antibodies were used: anti-KDM6A (#33510, Cell Signaling Technology, Danvers, MA, USA), anti-HA-Tag (#2367**,** Cell Signaling), anti-ENT1 (STJ96396, St John's Laboratory, London, UK), anti-β-actin (A5441, Sigma Aldrich), anti-α-Tubulin (T6199, Sigma Aldrich), anti-H3 (ab1791, Abcam, Cambridge, UK), anti-H3K27me1 (07-448, Merck Millipore, Billerica, MA, USA), anti-H3K27me2 (07-452, Merck Millipore), anti-H3K27me3 (07-449, Merck Millipore)**.** Western blot signals were quantified using ImageJ version 1.50d and relative levels were normalized to loading control.

### qRT-PCR analysis

Total RNA was isolated using the RNeasy Mini and RNase-Free DNase Kit (Qiagen). cDNA was synthesized by reverse transcription with the ThermoScript RT-PCR System (Thermo Fisher Scientific, Waltham, MA, USA) using 1 µg of total RNA as input. qRT-PCR assays were performed with QuantiTect SYBR Green PCR Kit (Qiagen) using 500 ng of cDNA and the following primers: *KDM6A*[6]; *KDM6B*[6]; *GAPDH*[6]; *EZH2*: fwd 5′-CCCTGACCTCTGTCTTACTTGTGGA-3′, rev 5′-ACGTCAGATGGTGCCAGCAATA-3′; *ENT1*[7], and *UTY*: fwd 5′-TTAGCCTGACAGTCGAGGAAA-3′, rev 5′-GTAGGGTCTTCGTTCTGGCG-3′. Reactions were run on a LightCycler480 II (Roche, Basel, CH). Fold changes were calculated using the $\Delta\Delta C_t$ method and normalized against *GAPDH* expression.

### CRISPR/Cas9 gene editing

*KDM6A* specific gRNA (5´-GGTATGCAGATAATGCTGAA-3`) was cloned into pSpCas9(BB)-2A-GFP (PX458), a gift from Feng Zhang (Addgene #48138). The gRNA was designed using Benchling (Biology Software, 2018). 48h after nucleofection with 1-2 µg plasmid, GFP positive cells were enriched by single-cell sorting into 96-well V-bottom

plates with FACSVantage SE (BD Bioscience). Cells were cultured until colonies were readily visible. Cell lysis, PCR on lysates and restriction digest were performed as previously described with minor modifications[8]. Briefly, for gDNA isolation cells in 96-well plate were washed with PBS (2x), resuspended in 50 µL/well lysis buffer (50 mM TRIS/HCl pH 7.5, 10 mM CaCl2, 1.7 µM SDS, 50 µg/ml Proteinase K), frozen at -80°C for 30 min, incubated at 56°C for 3 h, and 85°C for 30 min. 2.5 µL/well of the resulting lysate were subjected to PCR using forward 5´-GGGGTTAGCCTAGATGCTGTTC-3`, and reverse primer 5`-ATTGGCAATAATCTGCCCAAAACA-3´. KO clones were identified by restriction-fragment length polymorphism (RFLP) analysis of PCR products using HpyF10VI (Thermo Fisher Scientific). Sanger sequences were analyzed and aligned with Geneious 8.1.7 (Biomatters Ltd, Auckland, New Zealand) software and Benchling.

**DNA sequencing**

gDNA was isolated with the QIAamp DNA Blood Mini Kit (Qiagen) manually or using a Qiacube instrument (Qiagen). Mutation E1325X was verified by Sanger sequencing both DNA strands of PCR-amplified gDNA using 3500/3500xL Genetic Analyzer (Applied Biosystems). First, PCR was performed with forward primer 5`-CACGGATGAGGAAATTGACTCC-3` and reverse primer 5`-GGCATCTGTGTACATCTAGATTGTTCTTAG-3` followed by a second PCR and sequence analysis with forward primer 5`-CAGGCCTGCTGAGCATTG-3`and reverse primer 5`-GAAACCAACAGTGGAGAGGG-3. Targeted, multiplexed amplicon resequencing covering the entire open-reading frame of *KDM6A* and mutational hotspots/entire open-reading frame of genes known to be recurrently mutated in myeloid malignancies was performed as described previously[9].

**Transfection**

HEK293T cells were transiently transfected using the calcium-phosphate precipitation method. Briefly, 13µg plasmid DNA in 450µL $H_2O$ was mixed with 50µL 2M $CaCl_2$. The resulting mixture was slowly added to 500µL 2x HBS buffer (Sigma Aldrich), incubated for 4 min and added drop wise to a 10cm dish containing 90% confluent cells. For MG-132 treatment, 42h after transfection, cells were incubated with fresh media containing 50 µM MG-132 (Sigma Aldrich) for 6h. Suspension cells were transfected using the Amaxa nucleofection system with Nucleofector Kit V and preprogrammed settings (K562: T-016; MM-1: T-030/T-036; THP:1 U-001) on the Amaxa Nucleofector II device (Lonza Group AG, Basel, CH).

***In* vivo therapy trial**

Patient-derived xenograft (PDX) cells expressing firefly luciferase were established as previously described[9]. $5*10^5$ AML 491 or AML 393 cells were injected i.v. into groups (n=4; n=3 for AML 491 treated with AraC) of NSG mice (NOD scid gamma, The Jackson Laboratory, Bar Harbour, ME, USA), and tumor growth was regularly monitored by bioluminescence imaging (BLI) as previously described[9]. At defined imaging signals, mice were treated with a combination of cytarabine (100 mg/kg, i.p., days 1-4 and 14-17) and DaunoXome (1 mg/kg, i.v., days 1, 4, 14, 17). 28 days after start of therapy, BLI was performed and increase in BLI signals relative to day 0 were calculated. No randomization was used. All animal trials were performed in accordance with the current ethical standards

of the official committee on animal experimentation (Regierung von Oberbayern, number 55.2-1-54-2531-95-2010).

**siRNA**

K562 cells were nucleofected with Stealth siRNA (Thermo Fisher Scientific) against *KDM6A* (HS111232; 5'-GCAAAUGUUCCAGUGUAUAGGUUUA-3`) or negative control siRNA (si*SCR*; low GC; 12935200). After 72h, cells were nucleofected for a second time and incubated for 72h, after which cells were used for subsequent analysis.

**Lentiviral transduction**

Lentiviruses expressing *KDM6A*-targeting shRNA #3 (5′-TACTTGAATAGCACCTTCCGA-3′), #4 (5′-TTTAATGGCATCCTGAGGCTG-3′), #7 (5′-TTTATCAATAGACTGCCTGTA-3′) or control shRNA targeting Renilla (5′-TAGATAAGCATTATAATTCCT-3′) or eGFP (5′-CAGCCACAACGTCTATATCAT-3′) were generated by cloning into the pCDH-EF1α-MCS-T2A-copGFP vector (SBI, CD521A-1). dsRED-miR30 fragment (TRMPVIR vector; Addgene #27994) was amplified using primers carrying MfeI and SalI as 5' and 3' restriction sites. The 22mer shRNA target sequences have been synthetized as part of 110bps ssDNA oligos (Eurofins Scientific, Luxembourg), annealed and cloned into the vector. To enhance the shRNA expression, the EF1α promoter was replaced by the viral promoter SFFV. Production of lentiviral particles and transduction was performed as previously described[10,11]. After a few days, transgene positive cells were enriched in two consecutive rounds by flow cytometry using FACSVantage SE (BD Bioscience).

**MLPA and numerical aberrations**

MLPA analysis for *KDM6A* exon deletions was carried out as previously described[1]. To identify numerical *KDM6A* aberrations in AML cell lines, CytoScan HD Array (Affymetrix, Santa Clara, CA, USA) hybridization analysis were performed. The DNA was prepared using the Qiagen Gentra Puregene Kit (Qiagen). Labeling, hybridization and scanning were performed according to the manufacturer´s protocol. Data were analyzed using the Chromosome Analysis Suite software version 2.0.1.2 (Affymetrix).

**PiggyBac Construct**

To generate the PiggyBac (PB) donor vector harboring the doxycycline inducible expression cassette, the tet-3xFLAG-AsiSI-NotI-IRES-DsRed-Express-M2rtTA-P2A-PuroR cassette was amplified from pSBtet-3xFLAG-IRES-DsRed-Express-PuroR[12] using primers PB.fwd (5`-acggggaaaaggcctccaGGTCCGCTATCTAGACGA-3) and PB.rev (5`-aagccataccaatgggccGCTAGTAGGCCCAGCT-3) with overhangs for subsequent DNA assembly. The insert Cuo-MCS-IRES-GFP-EF1α-CymR-PuroR was excised from the PiggyBac expression vector PB-Cuo-MCS-IRES-GFP-EF1α-CymR-Puro (System Biosciences, #PBQM812A-1) using ApaI and SfiI (Thermo Fisher Scientific). The 3xFLAG-AsiSI-NotI-IRES-DsRed-Express-M2rtTA-P2A-PuroR fragment was then cloned into the ApaI and SfiI linearized PiggyBac backbone using the NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs) according to the manufacturer's

instructions generating the vector pPBtet-3xFLAG-IRES-DsRed-Express-PuroR. Human full length KDM6A without the N-terminal HA Tag was amplified from pcDNA6 HA KDM6A with overhangs for subsequent DNA assembly. The KDM6A fragment was cloned into the BmtI and NotI linearized pPBtet-3xFLAG-IRES-DsRed-Express-PuroR vector (3xFLAG was removed by linearization) using the In-Fusion HD Cloning Plus Kit (Takara Bio, Saint-Germain-en-Laye, France).

**Generation of inducible PiggyBac KDM6A cell lines**

To generate stable cell lines carrying doxycycline inducible KDM6A, cells were first nucleofected with equimolar amounts of pPBtet-KDM6A-IRES-DsRed-Express-PuroR and PiggyBac transposase (Biocat, PB200PA-1). Two days after transfection, cells were subjected to puromycin selection (2 µg/mL) for 3 days. Viable cells were enriched by single-cell sorting into 96-well V-bottom plates with FACSVantage SE (BD Bioscience). Cells were cultured with puromycin (2 µg/mL) until colonies were readily visible. To screen for successful KDM6A re-expression, clones were treated with doxycycline (0.5 µg/mL) for 48h and inducible KDM6A expression was analyzed by Western Blot.

**Library Preparation and Sequencing of Cell Lines**

To construct bulk libraries from mRNA, a protocol adapted from the SCRB-seq method was used[13]. Briefly, 50 ng mRNA was reverse transcribed using Maxima H Minus Reverse Transcriptase and tagged with sample-specific barcodes and unique molecular identifiers (UMIs). Only 2 µM of the E3V6NEXT was used. Samples were pooled and purified by SPRI beads, followed by an Exonuclease I treatment. Full-length cDNA was pre-amplified by single primer PCR for 10 cycles with the modification of using KAPA Hifi 2x ready mix (KAPA Biosystems). For the Nextera XT kit, 4 ng of cDNA was used as input and library preparation performed according to the manufacturer's protocol, with the exception of using a custom i5 primer. Sequencing was performed on an Illumina HiSeq 1500, flow cell with single end layout utilizing the standard Illumina sequencing and index primers. Sample reads were sequenced using 50 cycles and the UMI sequence using 16 cycles. To obtain expression data the raw fastq files were processed by the zUMIs pipeline using default parameters[14]. Mapping to the human reference genome hg38 was performed by STAR[15] (version 2.5.2b) and the gene annotation GRCh38.84 was taken from Ensembl. Differential expression analysis was performed using limma[16]. Genes with a read count below 10 in all samples were filtered out and library sizes scaled using the package edgeR[17]. Count data was transformed to log2-counts per million and the mean variance calculated to compute the precision weights. In order to increase statistical power empirical Bayes moderation was applied and false discovery rate (FDR) calculated by the Benjamini-Hochberg procedure.

**Sequencing of Primary Patient Data Analysis**

Sequencing of mRNA was performed as previously described[18] using the TruSeq RNA Sample Preparation protocol, followed by sequencing on a HiSeq 2000 Instrument (Illumina). Primary patient data was mapped to the human reference genome hg19 using 2-pass mapping with the RNA-Seq aligner STAR (version 2.4.0.1)[15], allowing only one alignment per read. Otherwise default parameters were used and the gene annotation GRCh37.75 taken from Ensembl. Mapped reads were then assigned to genes (Ensembl annotation version GRCh37.75) using the function featureCounts from the R package

Rsubread[19]. Subsequently, count tables for reads per genes were generated and normalized using the functions estimateSizeFactors and estimateDispersions of the DESeq2 package[20]. The three groups, *KDM6A*-up, *KDM6A*-down and *KDM6A*-same were defined as a change in expression between Diagnosis and Relapse of above or below 20% respectively.

**DNA Methylation Analysis**

KDM6A methylation status was determined based on 344 previously described AML patients[2]. This data is accessible via the Gene Expression Omnibus database repository (GSE18700). Clinical correlation of KDM6A methylation status and overall survival was determined using the Leukemia Gene Atlas version 2.1.0[21].

**ChIP-seq of Histone Modifications**

ChIP-seq reads were aligned to the human reference genome (hg38) using bowtie[22] with "-q-n2--best --chunkmbs2000- p 32 -m1" options. bigWig files were generated from homer tag directories using makeBigWig.pl. 1-2 million cross-linked cells (1% formaldehyde, 10min at RT) were lysed in 100 µL Buffer-B (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1%SDS, 1x protease inhibitors) and sonicated until most of the DNA fragments were 200-500 base pairs long (Covaris S220; 4°C, 30min, duty cycle 2%, 105 Watts). Lysates were then centrifuged (10min, 4°C, 12000g) and supernatant was diluted with 900 µL Buffer-A (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM EGTA,1% Triton X-100, 0.1% SDS, 0.1% Sodiumdeoxycholate, 140 mM NaCl, 1x protease inhibitors). 150 µL of sonicated chromatin was then incubated on a rotating wheel (3h, 4°C) with 2 µg of antibody conjugated to 10 µL Dynabeads. Antibodies used: H3K27ac (Diagenode; Pab-174-050), H3K27me3 (Pab-069-050). Beads were washed 4x with Buffer-A and once with Buffer-C (10 mM Tris-HCl, pH 8.0, 10 mM EDTA). Beads were incubated with 70 µL elution buffer (0.5% SDS, 300 mM NaCl, 5 mM EDTA, 10 mM Tris HCl pH 8.0) containing 2 µL of RNase (10mg/ml) for 30 min (37°C, 900rpm) and then with 2 µL of Proteinase K (20mg/ml) for 1h (55°C) and 8h (65°C) and supernatant was transferred to a new tube. Another 30 µL of elution buffer was added to the beads for 1 minute and eluates were combined and incubated with another 1 µL of Proteinase K for 1h at 55°C. Finally, DNA was purified with SPRI AMPure XP beads (Beckman Coulter), sample-to-beads ratio 1:2. Purified DNA was used as input for library preparation with NEBNext Ultra II DNA Library Prep Kit for Illumina (Biolabs, E7645S) and processed according to the manufacturer´s instructions. Libraries were quality controlled by Qubit and Agilent DNA Bioanalyzer analysis. Deep sequencing was performed on HiSeq1500/2500 according to the standard Illumina protocol for 50bp single-end reads. The dataset is publicly available under GSE128262.

[1] P.A. Greif, L. Hartmann, S. Vosberg, S.M. Stief, R. Mattes, I. Hellmann, et al., Evolution of Cytogenetically Normal Acute Myeloid Leukemia During Therapy and Relapse: An Exome Sequencing Study of 50 Patients., Clin. Cancer Res. 24 (2018) 1716–1726. doi:10.1158/1078-0432.CCR-17-2344.

[2] M.E. Figueroa, S. Lugthart, Y. Li, C. Erpelinck-Verschueren, X. Deng, P.J. Christos, et al., DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia, Cancer Cell. 17 (2010) 13–27. doi:10.1016/j.ccr.2009.11.020.

[3] W. Liu, Y. Xie, J. Ma, X. Luo, P. Nie, Z. Zuo, et al., IBS: an illustrator for the presentation and visualization of biological sequences., Bioinformatics. 31 (2015) 3359–3361. doi:10.1093/bioinformatics/btv362.

[4] A.P. Early, H.D. Preisler, H. Slocum, Y.M. Rustum, A pilot study of high-dose 1-beta-D-arabinofuranosylcytosine for acute leukemia and refractory lymphoma: clinical response and pharmacology., Cancer Res. 42 (1982) 1587–94. http://www.ncbi.nlm.nih.gov/pubmed/6949642 (accessed June 6, 2018).

[5] A.B. Lyons, K. V. Doherty, Flow Cytometric Analysis of Cell Division by Dye Dilution, Curr. Protoc. Cytom. (2004) 9.11.1-9.11.10.

[6] W. Jiang, D. Zhang, N. Bursac, Y. Zhang, WNT3 is a biomarker capable of predicting the definitive endoderm differentiation potential of hESCs., Stem Cell Reports. 1 (2013) 46–52. doi:10.1016/j.stemcr.2013.03.003.

[7] H. Kitao, Y. Morodomi, S. Niimi, M. Kiniwa, K. Shigeno, K. Matsuoka, et al., The antibodies against 5-bromo-2'-deoxyuridine specifically recognize trifluridine incorporated into DNA., Sci. Rep. 6 (2016) 25286. doi:10.1038/srep25286.

[8] C.B. Mulholland, M. Smets, E. Schmidtmann, S. Leidescher, Y. Markaki, M. Hofweber, et al., A modular open platform for systematic functional studies under physiological conditions, Nucleic Acids Res. 43 (2015). doi:10.1093/nar/gkv550.

[9] B. Vick, M. Rothenberg, N. Sandhöfer, M. Carlet, C. Finkenzeller, C. Krupka, et al., An Advanced Preclinical Mouse Model for Acute Myeloid Leukemia Using Patients' Cells of Various Genetic Subgroups and In Vivo Bioluminescence Imaging, PLoS One. 10 (2015) e0120925. doi:10.1371/journal.pone.0120925.

[10] N. Terziyska, C. Castro Alves, V. Groiss, K. Schneider, K. Farkasova, M. Ogris, et al., In vivo imaging enables high resolution preclinical trials on patients' leukemia cells growing in mice., PLoS One. 7 (2012) e52798. doi:10.1371/journal.pone.0052798.

[11] S. Ebinger, E.Z. Özdemir, C. Ziegenhain, S. Tiedt, C. Castro Alves, M. Grunert, et al., Characterization of Rare, Dormant, and Therapy-Resistant Cells in Acute Lymphoblastic Leukemia., Cancer Cell. 30 (2016) 849–862. doi:10.1016/j.ccell.2016.11.002.

[12] C.B. Mulholland, J. Ryan, W. Qin, M.D. Bartoschek, F.R. Traube, S. Bultmann, et al., TET1 drives global DNA demethylation via DPPA3-mediated inhibition of maintenance methylation, BioRxiv. 321604 (2018). doi:10.1101/321604.

[13] T. Soumillon, M. Cacchiarelli, D. Semrau, S. van Oudenaarden, A. Mikkelsen, Characterization of directed differentiation by highthroughput single-cell RNA-seq., BioRxiv. (2014).

[14] S. Parekh, C. Ziegenhain, B. Vieth, W. Enard, I. Hellmann, zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs, Gigascience. 7 (2018). doi:https://doi.org/10.1093/gigascience/giy059.

[15] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, et al., STAR: ultrafast universal RNA-seq aligner, Bioinformatics. 29 (2013) 15–21. doi:10.1093/bioinformatics/bts635.

[16] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., limma powers

differential expression analyses for RNA-sequencing and microarray studies., Nucleic Acids Res. 43 (2015) e47. doi:10.1093/nar/gkv007.

[17] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data., Bioinformatics. 26 (2010) 139–40. doi:10.1093/bioinformatics/btp616.

[18] L. Hartmann, S. Dutta, S. Opatz, S. Vosberg, K. Reiter, G. Leubolt, et al., ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation., Nat. Commun. 7 (2016) 11733. doi:10.1038/ncomms11733.

[19] Y. Liao, G.K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features., Bioinformatics. 30 (2014) 923–30. doi:10.1093/bioinformatics/btt656.

[20] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2., Genome Biol. 15 (2014) 550. doi:10.1186/s13059-014-0550-8.

[21] K. Hebestreit, S. Gröttrup, D. Emden, J. Veerkamp, C. Ruckert, H.-U. Klein, et al., Leukemia gene atlas--a public platform for integrative exploration of genome-wide molecular data, PLoS One. 7 (2012) e39148. doi:10.1371/journal.pone.0039148.

[22] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol. 10 (2009) R25. doi:10.1186/gb-2009-10-3-r25.

# Loss-of-function mutations in the histone methyltransferase EZH2 promote chemotherapy resistance in AML

Scientific Reports 2021

scientific reports

Check for updates

OPEN

# Loss-of-function mutations in the histone methyltransferase EZH2 promote chemotherapy resistance in AML

Julia M. Kempf[1,10], Sabrina Weser[1,10], Michael D. Bartoschek[2], Klaus H. Metzeler[1], Binje Vick[3], Tobias Herold[1], Kerstin Völse[3], Raphael Mattes[1], Manuela Scholz[4], Lucas E. Wange[8], Moreno Festini[1], Enes Ugur[2], Maike Roas[1], Oliver Weigert[1], Sebastian Bultmann[2], Heinrich Leonhardt[2], Gunnar Schotta[5], Wolfgang Hiddemann[6,9], Irmela Jeremias[3,6,7] & Karsten Spiekermann[1,6,9]✉

**Chemotherapy resistance is the main impediment in the treatment of acute myeloid leukaemia (AML). Despite rapid advances, the various mechanisms inducing resistance development remain to be defined in detail. Here we report that loss-of-function mutations (LOF) in the histone methyltransferase EZH2 have the potential to confer resistance against the chemotherapeutic agent cytarabine. We identify seven distinct EZH2 mutations leading to loss of H3K27 trimethylation via multiple mechanisms. Analysis of matched diagnosis and relapse samples reveal a heterogenous regulation of EZH2 and a loss of EZH2 in 50% of patients. We confirm that loss of EZH2 induces resistance against cytarabine in the cell lines HEK293T and K562 as well as in a patient-derived xenograft model. Proteomics and transcriptomics analysis reveal that resistance is conferred by upregulation of multiple direct and indirect EZH2 target genes that are involved in apoptosis evasion, augmentation of proliferation and alteration of transmembrane transporter function. Our data indicate that loss of EZH2 results in upregulation of its target genes, providing the cell with a selective growth advantage, which mediates chemotherapy resistance.**

Acute myeloid leukaemia (AML) is a heterogeneous haematological malignancy, characterised by clonal expansion of abnormal, undifferentiated myeloid precursor cells. Even though many patients with AML respond well to induction chemotherapy, relapse and refractory disease are common, representing the major cause of treatment failure. Treatment with cytarabine (AraC) and daunorubicin (DNR) remains the standard care for AML patients, although several new therapeutic strategies have been implemented within the last years[1–3]. Epigenetic dysregulation of DNA methylation or histone modifications has been identified in many malignant tumors[4,5] and can be considered as a cause of cancer development and progression[6,7]. Since considerable insight concerning those epigenetic changes has been gained in recent years, many therapy concepts targeting the involved regulatory factors have been proposed and hold promise for novel treatment approaches[8,9].

Enhancer of zeste homolog 2 (EZH2) is a lysine methyltransferase found as the central core protein of the polycomb repressive complex 2 (PRC2)[10]. Comprising four subunits (SUZ12, EED, EZH2/EZH1 and RbAp46), this complex mediates transcriptional repression by catalysing the trimethylation of histone H3 at lysine 27 (H3K27me3)[11]. EZH2 has been found to serve a dual purpose, as either tumour suppressor or oncogene,

[1]Department of Medicine III, University Hospital, LMU Munich, Munich, Germany. [2]Department of Biology II and Center for Integrated Protein Science Munich (CIPSM), Human Biology and BioImaging, LMU Munich, Planegg Martinsried, Germany. [3]Research unit Apoptosis in Haematopoietic Stem Cells (AHS), Helmholtz Zentrum München, Munich, Germany. [4]Center for Human Genetics and Laboratory Diagnostic (AHC), Martinsried, Germany. [5]Biomedical Center and Center for Integrated Protein Science Munich, LMU Munich, Martinsried, Germany. [6]German Cancer Consortium (DKTK), Heidelberg, Germany. [7]Department of Pediatrics, Dr. von Hauner Children's Hospital, LMU Munich, Munich, Germany. [8]Anthropology and Human Genomics, Department of Biology II, Ludwig-Maximilians-University, Martinsried, Germany. [9]German Cancer Research Center (DKFZ), Heidelberg, Germany. [10]These authors contributed equally: Julia M. Kempf and Sabrina Weser. ✉email: karsten.spiekermann@med.uni-muenchen.de

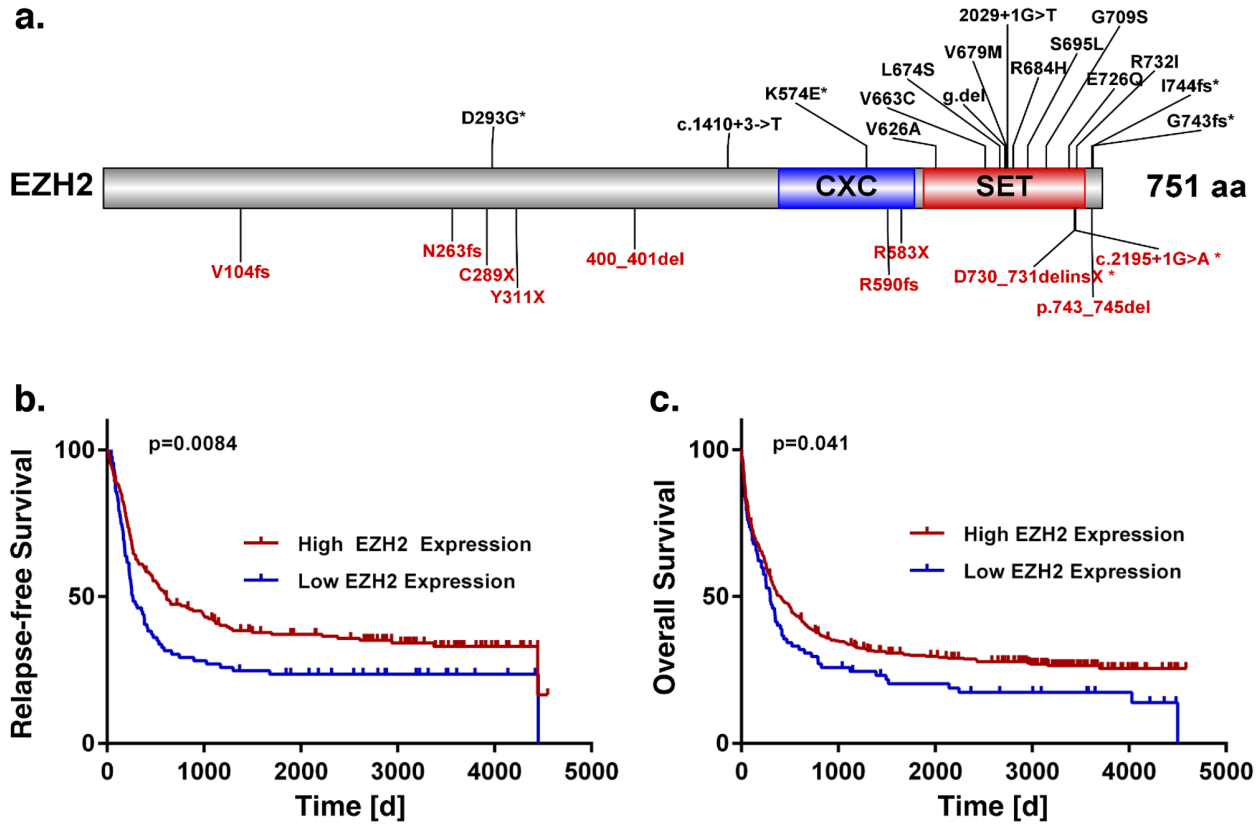



**Figure 1.** Recurrent *EZH2* mutations. **(a)** Schematic overview of EZH2 protein structure (NM_004456.4) and identified mutations (27 in total, c.2195+1G>A appeared twice) in a cohort of 664 AML patients at diagnosis. Functional domains are indicated at distinct locations and truncating mutations are displayed in red. Patients from Metzeler et al. 2016 (AMLCG-1999, AMLCG-2008). **(b–c)** Survival analysis of patients with low or high *EZH2* mRNA expression at the time point of diagnosis. *EZH2* high and low groups defined by the upper and lower quartile of *EZH2* mRNA expression, independent of mutation status. **(b)** Relapse-free survival (RFS). **(c)** Overall survival (OS). Patients from AMLCG 1999 (GSE37642), n = 517. 21 patients harboured an *EZH2* mutation. P-value calculated by log-rank test.

depending on the type of cancer[12–17]. In leukaemia, overexpression of EZH2 has been observed in CLL[18], paediatric T-ALL[19] and CML[20], while other studies reported EZH2 levels to be decreased in CMML[21] as well as ALL[18,19,22]. A recent study of Basheer et al. suggests opposing roles of EZH2 in initiation and maintenance of AML[23].

EZH1, an EZH2 homolog capable of partially compensating EZH2 function, holds an essential role in preserving pathological stem cells[24]. Therefore, it might contribute to the already complex role of EZH2 in hematopoietic malignancies[25–27]. Although EZH2 loss-of-function mutations seem to be rare in AML[28], loss of EZH2 by other mechanisms have been frequently reported and appear to play a major role in disease progression[29,30]. Absence of EZH2 in leukaemia cells was recently found to aberrantly activate BCAT1, resulting in enhanced mTOR signaling[27] and activation of the oncogene Hmga2 by causing an epigenetic switch from H3K27 trimethylation to H3K27 acetylation[31]. Furthermore, reduced disease-free survival was found to be associated with EZH2 mutations in myeloid malignancies[28,32,33] including AML[23]. In addition, chemoresistance was found in a recent study on AML patients with poor prognosis and downregulated EZH2[34].

In our previous study[35], examining diagnosis/relapse pairs of 50 cytogenetically normal (CN) AML patients, we found mutations in epigenetic modifiers, including EZH2, frequently gained at relapse, suggesting epigenetic mechanisms to be involved in disease progression in a subset of patients. The current study aims to evaluate the importance of chemotherapy resistance in AML. We investigated *EZH2* mutations and their functional loss of methyltransferase activity using patient samples, *in vivo* and *in vitro* patient-derived xenografts (PDX), and haematopoietic cell lines. We found *EZH2* loss-of-function mutations to be involved in the development of resistance against cytarabine and observed upregulation of EZH2 target genes due to loss of H3K27 trimethylation.

## Results

**Recurrent *EZH2* mutations at diagnosis.** In our previous work, we analysed 664 AML patients to study recurrently mutated genes, including *EZH2*[36]. In this cohort, 25 patients (4 %) carried an *EZH2* mutation at the time of diagnosis (27 mutations in total, Fig. 1a). Most of these mutations (n = 20, 74%) were located in the SET ([Su(var)3-9, Enhancer-of-zeste and Trithorax]) or CXC (cysteine-rich region, sometimes referred to as

pre-SET) domain at the C-terminus of the protein and are responsible for the catalytic activity of the methyltransferase. Furthermore, 41% (11) of mutations cause a stop-gain or frameshift, resulting in a truncated protein. An additional two frameshift mutations result in an elongated protein variant. Mutations most frequently co-occurring with mutated *EZH2* were found in *RUNX1*, *ASXL1*, *DNMT3A* and *TET2* (44%, 40%, 20% and 20%, Supplementary Fig. 1a). Additionally, *RUNX1* and *ASXL1* mutations were found to occur more often in *EZH2* mutated patients (44% and 40%) than in *EZH2* wild type patients (14% and 10%, $p$ = 4.6e−04, and $p$ = 9.3 e−05, Fisher's Exact Test). In contrast, NPM1, the most frequently mutated gene in our cohort, was found to be mutated less often in *EZH2* mutated (12%) than in *EZH2* wild type patients (34%) ($p$ = 2.8 e−02, Fisher's Exact Test). Interestingly, *KDM6A* and *EZH2* mutations were found to be mutually exclusive. Most patients with *EZH2* mutations (76%, n = 19) can be assigned to the adverse risk group (Supplementary Fig. 1a), according to the recent ELN classification[37].

In order to evaluate the prognostic importance of *EZH2,* we examined the survival of patients dependent on their *EZH2* mutation and expression status. The overall survival (OS) of patients harbouring *EZH2* mutations did not differ significantly from patients without mutation (Supplementary Fig. 1b). However, low *EZH2* mRNA expression was significantly associated with poor relapse-free survival (RFS) and OS in publicly available independent data sets of the AMLCG 1999 trial (GSE37642, Fig. 1b-c) and HOVON (GSE14468, Supplementary Fig. 1) study groups[38–40]. Additionally, monosomy 7, resulting in reduced *EZH2* expression, was associated with poor overall survival (Supplementary Fig. 1c).

**Relevance of EZH2 status in AML relapse.** To further investigate the poor survival in patients with low *EZH2* mRNA expression, we compared protein expression in a set of matched diagnosis and relapse pairs of ten AML patients without *EZH2* mutations (Supplementary Table 2). In 50% of patients, we observed decreased levels of EZH2 protein expression, whereas the other half revealed increased protein expression levels in relapse (Fig 2a). An increase of at least 2-fold in protein expression was found in four patients, whereas a strong decrease (2-fold or more) in protein expression was observed in three patients. An additional analysis of *EZH2* mRNA expression in 32 CN-AML patients revealed a similar heterogenous picture. Downregulation of *EZH2* was found in 22% of patients, while upregulation was found in 53% (Fig. 2b). Additionally, we identified two relapse-associated *EZH2* mutations. EZH2/p.A692G found in the second relapse of patient CN-021 from the Greif et al. cohort[35] and EZH2/Y733LfsX6 found in the first relapse of a patient from the AML-CG cohort. Both mutations revealed subclonal outgrowth during the course of treatment and increasing variant allele frequencies (VAFs) in relapsed patients (Fig. 2c). Additionally, we found an increase of VAFs in the relapse of three other EZH2 mutations found in the Greif et al. cohort[35] (Supplementary Fig. 2).

**Functional characterisation of *EZH2* mutations.** To evaluate the biochemical activity of the EZH2 variants, we measured global H3K27 trimethylation levels in a 293T/$EZH2^{-/-}$ model (Fig. 3a), which was established through CRISPR/Cas9 mediated genome editing, targeting exon 3 of *EZH2*. We found EZH2 protein expression levels to be strongly correlated with global H3K27me3 levels (Fig. 3b). In fact, global H3K27me3 was not detectable in any of the tested $EZH2^{-/-}$ clones, whereas $EZH2^{+/-}$ clones showed decreased EZH2 expression as well as reduced global H3K27me3 levels (Fig. 3a). Since EZH2 is only one part of the PRC2 complex, we additionally analysed protein expression of the remaining components SUZ12, RBAP46 and EED. We could not detect aberrant expression of these subunits in both 293T/$EZH2^{-/-}$ clones (Fig. 3c). Interestingly, both clones showed an increased resistance against AraC compared to the wild type clones (Fig. 3d, Supplementary Fig. 3a) and a slightly reduced colony count was observed in a colony formation assay (Supplementary Fig. 3c). Re-expression of seven different EZH2 variants, (Supplementary Fig. 3b) found in the AML-CG-1999 and AML-CG-2008 studies, could only partially rescue global H3K27me3 levels, indicating a LOF phenotype, while the re-expressed wildtype protein was able to restore complete activity (Fig. 3e).

In addition, co-expression of these variants with wild type *EZH2* led to a reduction of H3K27me3 levels in four mutations, suggesting a dominant-negative effect (Fig. 3f).

To validate the robustness of our 293T/$EZH2^{-/-}$ model, we performed the rescue experiment with two previously described EZH2 variants. EZH2/p.Y646N, a gain-of-function mutation found in lymphomas[16,17] and EZH2/Y731, a LOF mutation[41]. We were able to verify the functions of both mutations with our model (Supplementary Fig. 3d and e).

**EZH2 depletion promotes resistance in K562 cells.** In order to study the impact of *EZH2* mutations on chemoresistance in a hematopoietic context, we screened 12 AML cell lines with *EZH2* mutant or wild type background (Supplementary Table 1). We identified two *EZH2* mutated cell lines, SKM-1 and KG-1a, that seem to be more resistant against cytarabine and daunorubicin, respectively (Fig. 4a). Notably, three of the $EZH2^{wt}$ cell lines were harbouring *KDM6A* mutations, which can also affect drug resistance[42]. Furthermore, we found a strong positive correlation between H3K27me3 levels and EZH2 protein expression (Supplementary Fig. 3a). Next, we established seven $EZH2^{-/-}$ single cell (sc) knockout clones in the myeloid cell line K562, using CRISPR Cas9 genome editing (Fig. 4b, Supplementary Fig. 4e). Both knockout and control cells were treated for 72 h with either AraC or DNR. In K562/$EZH2^{-/-}$ clones, increased chemoresistance was found against AraC, while sensitivity against DNR was not affected (Fig. 4c, Supplementary Fig. 4c). Additionally, we observed reduced proliferation in K562/$EZH2^{-/-}$ clones compared to K562/$EZH2^{+/+}$ clones (Fig. 4d). Furthermore, the response towards AraC treatment was studied in a long-term proliferation assay, consequently treating single cell clones for 12 days with a low dose of AraC. In accordance with the short-term assay, also the K562/$EZH2^{-/-}$ sc clones of the long-term assay displayed higher resistance against AraC (Fig. 4e, Supplementary Fig. 4d).

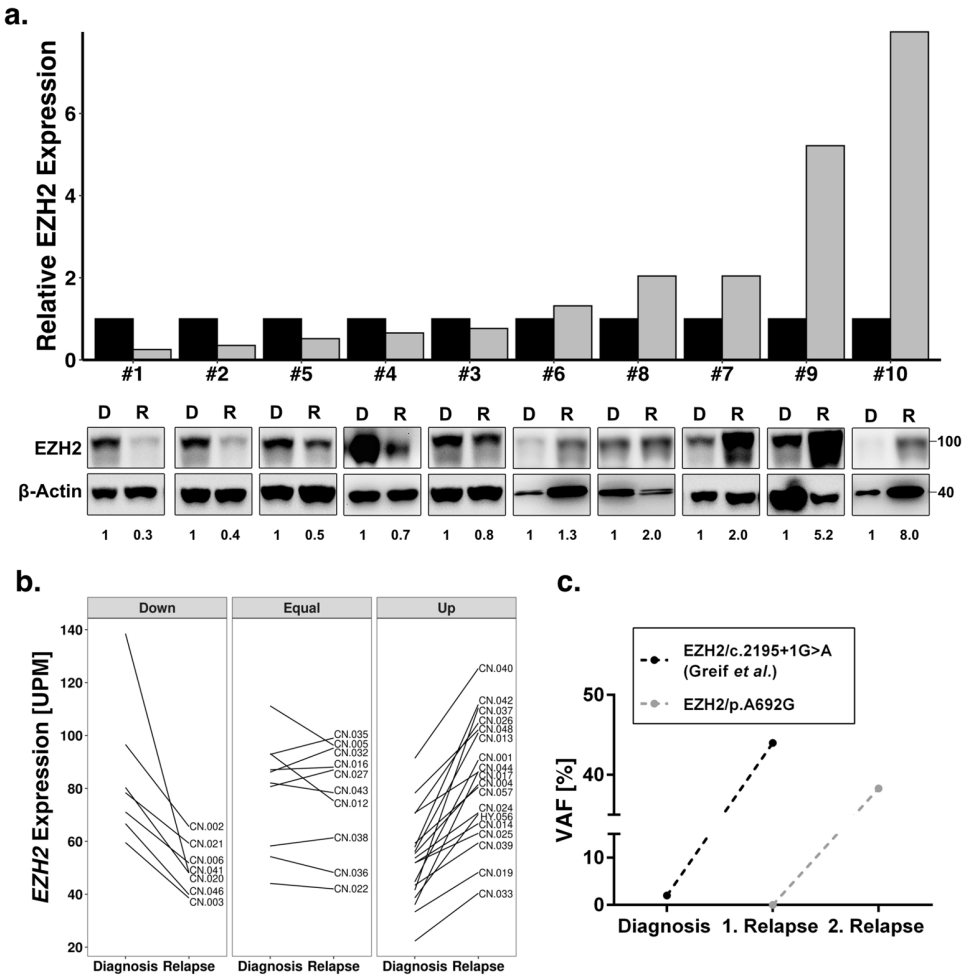



**Figure 2.** Relevance of EZH2 status in AML relapse. **(a)** Immunoblot for EZH2 protein expression in 10 AML patients at diagnosis and relapse. MW, molecular weight; β-actin, loading control. The ratio of EZH2 to β-actin expression is indicated below and presented in the histogram above. Each relapse value was normalized to the corresponding diagnosis sample. None of the patients carried an *EZH2* mutation. **(b)** *EZH2* mRNA expression between diagnosis and relapse of 32 CN-AML patients from Greif et al. cohort[35]. Up and down are defined as a change in mRNA expression of at least 20%. Three patients carried an *EZH2* mutation. **(c)** Variant allele frequency of the two relapse-associated *EZH2* mutations with outgrowth in first and second relapse.

**EZH2 re-expression sensitises to AraC treatment in K562 cells.** To investigate if re-expression of EZH2 can reconstitute baseline H3K27me3 levels and therefore sensitise cells to AraC treatment, we established a stable, doxycycline-inducible EZH2 expression system via the PiggyBac transposon system (Supplementary Fig. 5a–d). For this reason, DNA coding for *EZH2*$^{wt}$ (AA1-751) was introduced into K562/*EZH2*$^{-/-}$ cells (clone #7). Re-expression of wildtype *EZH2* was able to restore global H3K27me3 levels after 48 h (Fig. 5a and Supplementary Fig. 5e). Furthermore, sensitivity against AraC could be restored in a long-term low dose AraC treatment experiment (Fig. 5b,c). Additionally, we introduced DNA coding for the relapse-associated mutation EZH2/Y733LfsX6 (AA1-737, Supplementary Fig. 5c,d). Re-expression of this mutant after doxycycline induction did not result in restoration of H3K27me3 levels (Fig. 5d and Supplementary Fig. 5f). Likewise, the mutation was not able to restore sensitivity against AraC treatment, indicating an involvement of H3K27 trimethylation in the phenotype of chemoresistance (Fig. 5e,f). Doxycycline alone had no effect on the sensitivity of either *EZH2*$^{wt}$ or *EZH2*$^{-/-}$ cells towards AraC treatment (Supplementary Fig. 5g,h).

**Upregulation of EZH2 target genes desensitises cells to AraC treatment.** RNA sequencing and Proteome analysis was performed to uncover the molecular mechanism involved in EZH2-mediated chemoresistance. *EZH2* knockout in K562 cells resulted in aberrant gene and protein expression (Supplementary Fig. 7), visible in transcriptional upregulation of 216 genes and downregulation of 42 genes as well as translational upregulation of 375 genes and downregulation of 205 genes (Supplementary Table 3 and Supplementary Table 4). The change in protein and RNA expression was found to be correlated (R = 0.5, $p$ = 2.2e-16, Pearson's correla-

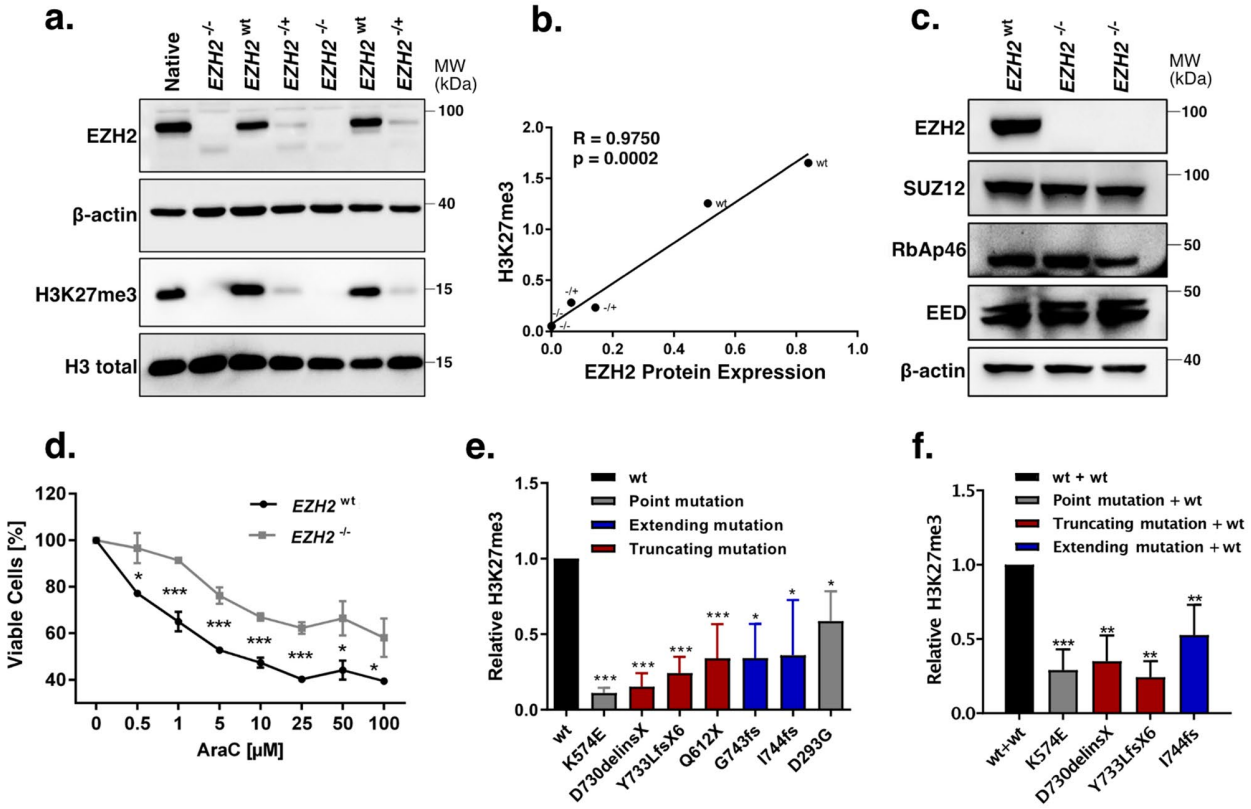



**Figure 3.** Evaluation of *EZH2* mutations found in AML patients at diagnosis. **(a)** Comparison of EZH2 expression and global H3K27me3 between $EZH2^{-/-}$, $EZH2^{+/-}$ and $EZH2^{wt}$ sc clones in 293T cells. MW, molecular weight. β-actin and H3 total, loading controls. **(b)** Correlation between EZH2 protein expression and global H3K27me3 in 293T sc clones. Pearson's correlation. **(c)** Immunoblot for EZH2, SUZ12, RbAP46 and EED expression in 293T/$EZH2^{-/-}$ sc clones. MW, molecular weight; β-actin, loading control. **(d)** AraC resistance in one 293T/$EZH2^{-/-}$ and one 293T/$EZH2^{wt}$ sc clone. Cells were treated for 72 h with different concentrations of AraC. Viable cells relative to untreated control. **(e)** H3K27me3 levels after re-expression of seven *EZH2* mutations, detected in patient diagnosis samples. Colours referring to protein structural changes caused by the mutation. 293T/$EZH2^{-/-}$ cells were transfected transiently with *EZH2* constructs 72 h before protein isolation, and global H3K27me3 was evaluated by immunoblot. Values relative to the wild type. **(f)** H3K27me3 levels after re-expression of four *EZH2* mutants in combination with wild type *EZH2*. 293T/$EZH2^{-/-}$ cells were transfected transiently with *EZH2* wildtype and *EZH2* mutant constructs 72 h before protein isolation. Values relative to the wild type. Unpaired, two-tailed Student's t-test; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. Error bars indicate mean ± s.d of at least three independent experiments.

tion, Fig. 6a) and 41 genes showed differential expression of both mRNA and protein. Most of these genes (37) were upregulated, and only two downregulated in both measures (Fig. 6b). Amongst the upregulated genes, we identified *FHL1* as well as *UBE2E1*, both involved in chemotherapy resistance and relapse in AML[43,44]. Additionally, upregulation of *CA2*, *CNN3* and *AKAP13* was found, which are suggested to be involved in chemotherapy resistance in glioblastoma, colon cancer and breast cancer, respectively. Furthermore, *PDK3*, *TPD52*, *MYO5A*, *AKT3* and *SPECC1* were upregulated, genes associated with poor prognosis in AML or involved in apoptosis. EZH2 ChIP-seq in K562 cells of a publicly available dataset (ENCSR000AQE, ENCSR000AKY) revealed peaks in the promoter region of *FHL1*, suggesting *FHL1* to be a direct target of EZH2. In the promoter region of *UBE2E1* no EZH2 peaks were found, but EZH2 binding was detected in a distal enhancer region (GeneHancer Accession: GH03J023748). Other potential direct targets of EZH2 are *CNN3*, *AKAP13*, *TPD52*, *MYO5A*, *AKT3* and *SPECC1* as EZH2 peaks were found in the respective promoter regions. Additionally, enhancers regulating *FHL1* and *TPD52* could be identified (GeneHancer Accession: GHOXJ136155 and GH08J080078). No peaks were assigned to the genes *CA2* and *PDK3*.

**Resistance of *EZH2* mutated patient-derived xenografts (PDX).** To extend our findings of *EZH2* associated AraC resistance *in vitro*, we screened relapsed AML samples for clonal outgrowth of *EZH2* mutated cells. We identified a 54-year-old patient who gained an *EZH2* mutation (p.A692G) at second relapse. A summary of the patients' course of disease including bone marrow blast counts from 9 time points is shown in Figure 7a. We established a sensitive, custom designed digital droplet PCR (ddPCR) assay to monitor the abundance of p.A692G during disease progression from first to second relapse. We detected the mutation only in the

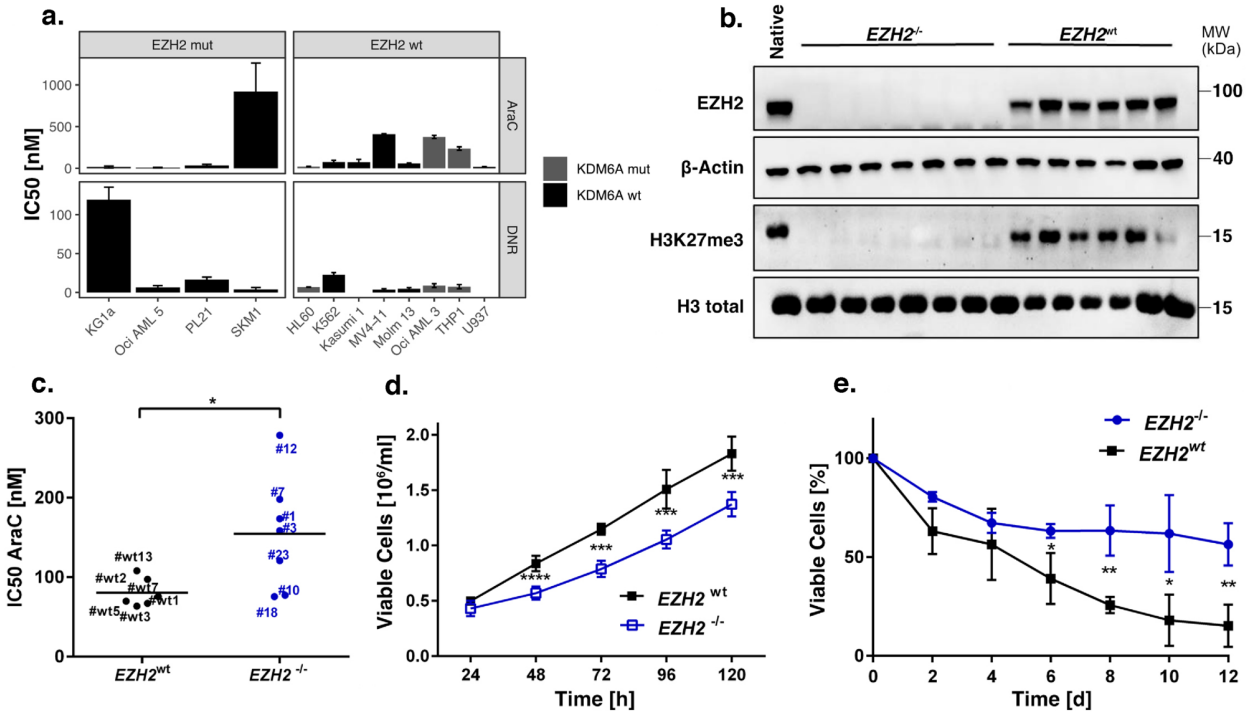



**Figure 4.** EZH2 depletion promotes resistance in the myeloid cell line K562. **(a)** Comparison of $IC_{50}$ values for DNR and AraC in twelve haematopoietic cell lines. Cells were treated with AraC/DNR for 72 h. **(b)** Immunoblot for EZH2 expression and global H3K27me3 of seven *EZH2*$^{-/-}$ and six *EZH2*$^{wt}$ sc clones in K562 cells. MW, molecular weight; β-actin and H3 total, loading controls. **(c)** Comparison of AraC $IC_{50}$ values in *EZH2*$^{wt}$ (n = 6) and *EZH2*$^{-/-}$ (n = 7) clones. Cells were treated with AraC/DMSO for 72 h. Each value represents the mean of three independent experiments. **(d)** Proliferation of *EZH2*$^{wt}$ (n = 4) and *EZH2*$^{-/-}$ (n = 7) clones for 5 d. Medium was changed every 48 h. **(e),** Long-term low dose AraC treatment in *EZH2*$^{wt}$ (n = 3) and *EZH2*$^{-/-}$ (n = 3) clones. Cells were treated with 30 nM AraC/DMSO for 12 d. Viable cells relative to untreated control. Unpaired, two-tailed Student's t-test; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. Error bars indicate mean ± s.d of three independent experiments.

second relapse with a more than 20% increase within three months (Fig. 7a). Furthermore, the patient gained a heterozygous 7q deletion at first relapse, as analysed by MLPA (Fig. 7a).

Patient cells of the first and second relapse were serially transplanted into immune-deficient mice, establishing patient-derived xenografts PDX-AML-491 and PDX-AML-661 respectively (Fig. 7a)[45]. Leukaemic cells of the initial diagnosis did not engraft in the model. PDX cells were lentivirally transduced for transgenic expression of luciferase, enabling disease monitoring *in vivo*[36]. We treated the xenograft mice with a combination of cytarabine and daunorubicin and monitored the leukaemic burden over 80 days. We observed a drastic drop in leukaemic cells in the PDX-491 mice, with a complete cure of three out of four animals (Fig. 7b). The treatment of PDX-661 mice only had minimal effect (Fig. 7c).

Targeted sequencing of a panel of 68 recurrently mutated genes[36] of patient and PDX samples revealed a strong increase of the clone harbouring the *EZH2*/p.A692G mutation in the PDX-661 samples (VAF: 98.8%) in comparison to the second relapse of the patient (VAF: 39.2%, Fig. 7d). The only other mutation illustrating an increase in variant allele frequency was a subclonal *JAK1* mutation, detectable only in the PDX-661 cells. The majority of mutations (*BCOR*, *DNMT3A*, *ETV6*, *PTPN11* and *RUNX1*) remained stable at all time points. Furthermore, two subclonal mutations in *NRAS* and *KRAS* were detected. Both were absent in the patient's first relapse. KRAS was only detectable in the PDX-491 samples, while NRAS decreased during PDX-491 passaging but was detectable again in the PDX-661 as well as in the patient's second relapse.

Dose-response analysis of PDX-491 and PDX-661 cells in vitro confirmed an increased resistance of PDX-661 towards AraC (Fig. 8a). Moreover, global H3K37me3 levels were completely depleted in PDX-661 cells, while EZH2 protein expression was stable (Fig. 8b). Transient transfection of the p.A692G mutation into our 293T/*EZH2*$^{-/-}$ model revealed decreased global H3K27me3 compared to the wild type, further confirming a LOF phenotype (Fig. 8c). To examine if the observed chemoresistance can be caused by EZH2 depletion, we established an siRNA knockdown (kd) targeting wild type *EZH2* in the PDX-491 cells. EZH2 levels could thereby be reduced by approximately 40% (Fig. 8d). Treatment of these cells for 72 h with AraC resulted in lower proliferation (Fig. 8e) and an increased resistance (Fig. 8f).

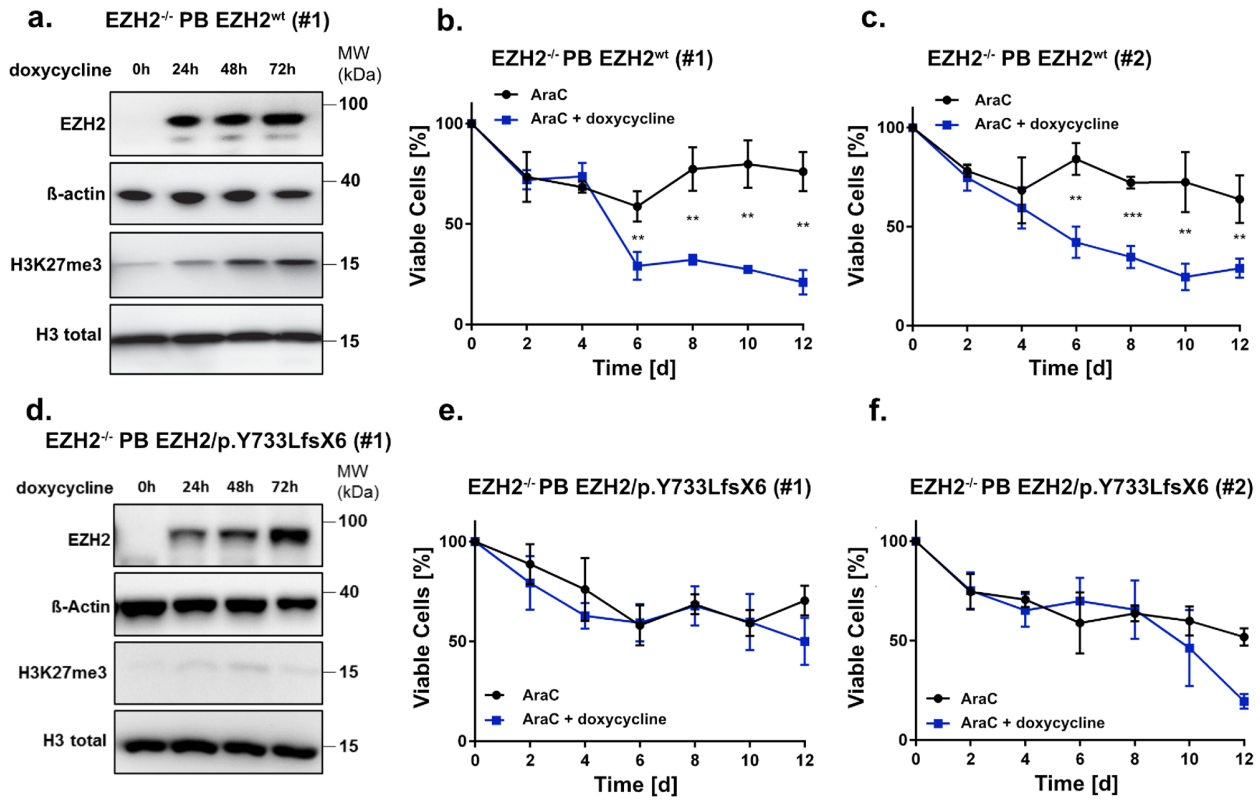



**Figure 5.** EZH2 re-expression sensitizes K562 cells to AraC treatment. **(a,d)** Immunoblot for EZH2 expression and global H3K27me3 in **(a)** *EZH2*$^{-/-}$ PB *EZH2*$^{wt}$ cells (clone #1) and **(d)** *EZH2*$^{-/-}$ PB EZH2/p.Y733LfsX6 (clone #1) after 0 h, 24 h, 48 h and 72 h of doxycycline induction. Cells were treated with 1 µg/ml doxycycline every 24 h. MW, molecular weight; β-actin and H3 total, loading controls. **(b–c, e–f),** AraC low dose long-term treatment in **(b–c)** *EZH2*$^{-/-}$ PB *EZH2*$^{wt}$ and **(e–f)** *EZH2*$^{-/-}$ PB EZH2/p.Y733LfsX6 cells. Cells were pre-treated for 3 d with doxycycline and then treated with 30 nM AraC/DMSO for 12 d. Cells were split and treated every 4 d and doxycycline was added every 48 h to ensure stable expression of EZH2. Error bars indicate mean ± s.d of three independent experiments. Unpaired, two-tailed Student's t-test; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

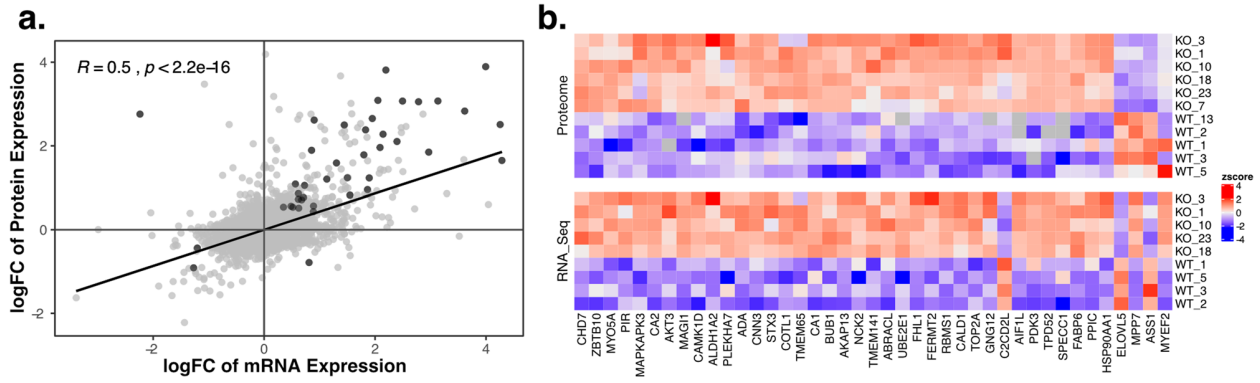



**Figure 6.** Upregulation of EZH2 target genes. **(a)** Correlation of protein and mRNA expression. Dark grey points representing genes differentially expressed in both measures (adj.P < 0.05). Pearson's correlation. **(b)** Heatmap of the 41 genes differentially expressed (adj.P < 0.05) between *EZH2*$^{-/-}$ and wild type clones, in both protein and mRNA. The colour gradient from red to blue represents high to low expression of genes. White indicates no change.

**Figure 7.** Resistance in an *EZH2* mutated PDX model. **(a)** Course of disease of an AML patient suffering from two relapses (indicated by dashed vertical lines). 7q deletion was confirmed with MLPA. Variant allele frequency (VAF) of the p.A692G mutation was monitored by digital droplet PCR (time points of samples indicated by red stars). Blast count was measured from bone marrow at the indicated time points (blue bars). Samples used for PDX engraftment are indicated with black triangles. In July 2015 two samples were taken in the same month. **(b–c)** In vivo treatment of PDX mice. NSG mice were injected with patient material of relapse 1 and 2, establishing **(b)** PDX-491 and **(c)** PDX-661. 21 d after injection, mice were treated with AraC (100 mg/kg) and DaunoXome (1 mg/kg) (treatment days indicated with red x). Leukaemic burden was monitored in vivo by bioluminescence imaging. Control mice treated only with PBS are shown in blue. **(d–e),** Variant allele frequencies in the course of first to second relapse of **(d)** EZH2/p.A692G and **(e)** other mutations identified by targeted sequencing in Patient and PDX samples. PDX samples of first engraftment (primograft) as well as first (1st re-Tx) second (2nd re-TX) and third (3rd re-TX) re-transplantation. ▶

## Discussion

Development of resistance against standard chemotherapeutics is common in AML and can be induced through various mechanisms[46]. In this study we report that loss-of-function mutations in the histone methyltransferase EZH2 are associated with increased resistance against the antimetabolite cytarabine (AraC).

*EZH2* mutations in AML are a very rare event. In fact, only 4% of our patients harboured these mutations at diagnosis, with the majority located in the catalytic SET domain, a known hotspot for *EZH2* mutations[28,47]. Seemingly, the mutations induce loss of EZH2 function, independent of the type of mutation. (Fig. 3, Supplementary Fig. 3a). Apart from the SET domain (aa 605-725), also the post-SET domain (aa 725-746), which is essential for the formation of the cofactor S-adenosyl-L-methionine (SAM) binding pocket, was found crucial in maintaining enzymatic function[48]. Two of the mutations identified in our patients, D730_delinsX and Y733LfsX6, previously described in myelodysplastic syndromes (MDS)[49], caused almost complete elimination of the post-SET domain, while two others, I744fs[50,51] and G743fs, caused frame shifts that resulted in elongated protein variants, highlighting the importance of this domain. In contrast, the missense mutation K574E, located in the CXC domain, is likely to impair the domain's binding ability to the substrate nucleosome and thereby bringing the H3 tail out of reach[50].

We found that complete loss of EZH2 promotes AraC resistance in HEK293T cells as well as the myeloid cell line K562 (Fig. 3, Fig. 4). Furthermore, increased resistance was observed in K562 cells expressing the LOF mutation Y733LfsX6 (Fig. 5d–f) and in a PDX model of a patient who gained the LOF mutation A692G at second relapse (Fig. 7). Additionally, low *EZH2* mRNA expression correlated with poor overall and relapse-free survival (Fig. 1b–c). Our findings are therefore in concordance with the study of Göllner et al.[34], who described AraC resistance in a shRNA knockdown of EZH2 in MV4-11 cells. However, elevated *EZH2* expression has also been reported in AML patients[52,53], and dual inhibition of EZH1/2 was found to eliminate quiescent leukaemic stem cells (LSCs) to prevent relapse[25]. These combined findings suggest a dual role of *EZH2* as either tumour suppressor or oncogene. In our matched diagnosis/relapse pairs, EZH2 protein and mRNA expression levels were found to be highly patient specific, and in most cases, we observed up- or downregulation in relapse (Fig. 2). EZH2 therefore appears to bear an important function in disease progression, and close monitoring of expression and mutation status seems to be crucial in choosing the best treatment approach.

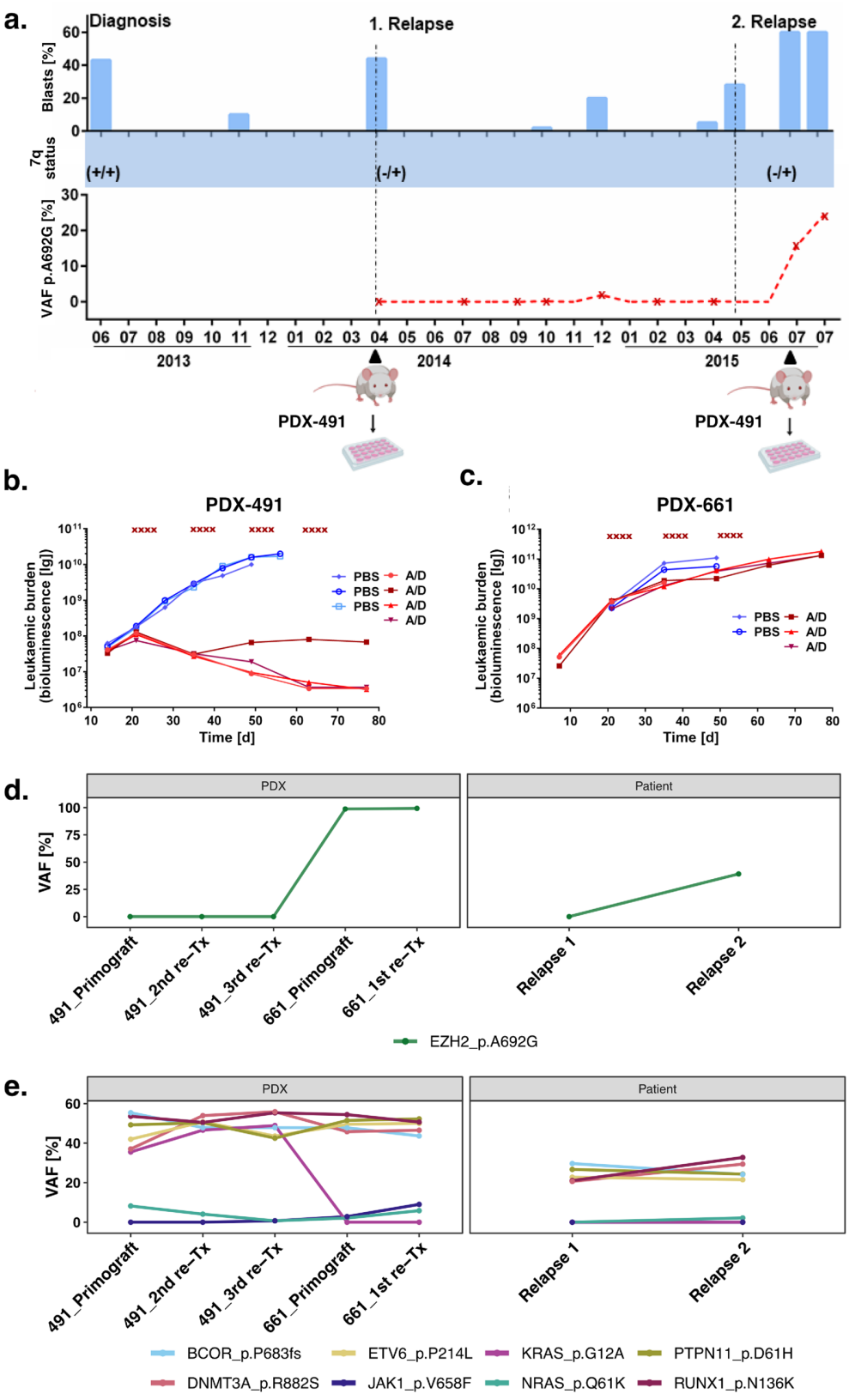
Interestingly, *RUNX1* and *ASXL1* mutations were significantly co-occurring with mutations in *EZH2*. Similar associations have been described before in myeloid malignancies including AML[28,54,55]. Therapy resistance was associated with frequent co-occurrence of *EZH2* and *RUNX1* LOF mutations[56], suggesting a cooperative role of these mutations. *ASXL1* LOF mutations on the other hand can establish an additive effect to EZH2 loss by additional reduction of H3K27 trimethylation through inhibition of PRC2 recruitment[29].

Mutations in other PRC2 subunits (EZH1, EED, SUZ12 or RbAp48) are extremely rare in AML. In the cohort of 50 AML patients of Greif et al. (2018) none could be detected, while in a study of 165 AML patients from Faber et al. (2016), only EED mutations were found with a frequency of 1.8 %. Co-occurrence of *EED* and *EZH2* mutations was found in only one of the patients. EZH2 requires direct interaction with EED to exert its enzymatic function[57]. Thus, also other mutations in the PRC2 complex like EED mutations harbour the potential to confer chemoresistance.

H3K27me3 levels can also be altered by the histone demethylase KDM6A. Loss-of-function mutations in *KDM6A* have been detected in AML and are associated with the development of chemoresistance[42]. Although we and other groups found mutations in both genes to be mutually exclusive[58], expression levels of KDM6A and EZH2 have an antagonistic effect on global H3K27 trimethylation (Supplementary Fig. 4d). Further research is needed to investigate common and specific EZH2 and KDM6A target sites.

EZH2 is responsible for the trimethylation of H3K27 and therefore inactivation of its target genes. Knockout of *EZH2* in K562 cells induced almost complete loss of H3K27me3 levels and resulted in the upregulation of 216 genes and 375 proteins (Fig. 6b). We identified *FHL1* and *UBE2E1* to be direct targets of EZH2. Overexpression of these genes has recently been described to be involved in resistance against cytarabine, and in relapse in AML patients[43,44].

FHL1 might be involved in the transmembrane transport of chemotherapeutic agents. Fu et al.[43] found upregulation of *ABCC1* and *ABCC4*, encoding for the unidirectional efflux transporter proteins MRP1 and MRP4, in AML patients with high FHL1 expression. A slight upregulation of ABCC1 protein could also be detected in our data. Interestingly, Fu et al. also found expression of *FHL1* to be negatively correlated to *SLC29A1* (ENT1) expression. ENT1 is an influx transporter that mediates the uptake of chemotherapeutics and is downregulated upon loss of KDM6A[42]. Since *EZH2* and *KDM6A* mutations were found to be mutually exclusive, those findings suggest an involvement of either EZH2 or KDM6A in the regulation of transmembrane transporter proteins,


a.
Diagnosis
1. Relapse
2. Relapse
Blasts [%]
7q status
(+/+)
(-/+)
(-/+)
VAF p.A692G [%]
2013
2014
2015
PDX-491
PDX-491
b.
PDX-491
Leukaemic burden (bioluminescence [lg])
Time [d]
PBS
A/D
c.
PDX-661
d.
PDX
Patient
VAF [%]
491_Primograft
491_2nd re–Tx
491_3rd re–Tx
661_Primograft
661_1st re–Tx
Relapse 1
Relapse 2
EZH2_p.A692G
e.
BCOR_p.P683fs
ETV6_p.P214L
KRAS_p.G12A
PTPN11_p.D61H
DNMT3A_p.R882S
JAK1_p.V658F
NRAS_p.Q61K
RUNX1_p.N136K

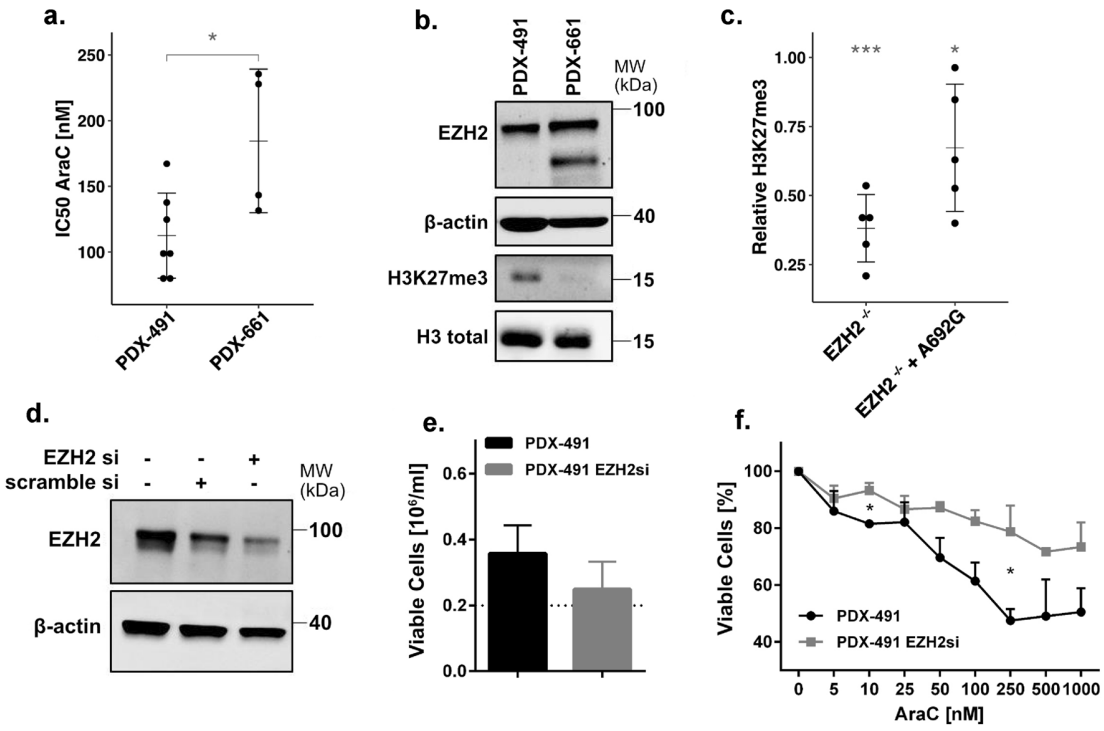



**Figure 8.** Knockdown of EZH2 in a patient-derived xenograft (PDX) model. **(a)** Comparison of $IC_{50}$ AraC values for PDX-491 and PDX-661 in vitro. **(b)** Immunoblot of EZH2 expression and global H3K27me3 in PDX-491 and PDX-661. **(c)** H3K27me3 levels of EZH2/p.A692G in 293T/*EZH2*$^{-/-}$ cells. Values normalized to H3 loading control and relative to wild type. *indicates significant difference to the wild type. **(d)** Immunoblot of EZH2 expression in PDX-491 cells treated with 10 nM siRNA. Representative blot shown for two independent experiments. **(e)** Histogram showing the proliferation of PDX-491 cells with 10 nM siRNA. Cells were pre-treated for 2 d with siRNA and then incubated for another 3 d for the proliferation assay. **(f)** AraC treatment in PDX-491 cells with 10 nM siRNA. Cells were pre-treated for 2 d with siRNA and then treated for 72 h with AraC. Unpaired, two-tailed Student's t-test; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. MW, molecular weight. β-actin and H3 total, loading controls. Error bars indicate mean ± s.d of three independent experiments.

responsible for the release or uptake of chemotherapeutic agents. The ubiquitin-conjugating enzyme UBE2E1 can regulate the expression of *HOX* genes by its ability to ubiquitinate histones[59]. Although we did not detect any aberrant expression of *HOX* genes, upregulation of HOXA9 as well as HOXB7 was reported by Göllner et al. in a resistant EZH2 negative AML cell line model. We furthermore identified upregulation of the direct EZH2 target genes *CNN3* and *AKAP13*, that are involved in chemotherapy resistance in colon cancer and breast cancer, respectively[60,61], and the genes *MYO5A*, *AKT3* and *SPECC1*, which are implicated in the evasion of apoptosis[62–64]. Additionally, upregulation of TPD52, involved in proliferation, migration, invasion and apoptosis, was found in many cancer types including AML[65].

We conclude that loss-of-function mutations in the histone methyltransferase EZH2 have the potential to confer resistance against the chemotherapeutic agent cytarabine and suggest an involvement of upregulated EZH2 target genes in apoptosis, proliferation and transmembrane transport.

## Materials and methods

**Cell culture and patient samples.** All cell lines (Supplementary Table 1) were acquired from DSMZ (Braunschweig, Germany) and cultured according to the supplier's recommendations. Patient-derived xenograft (PDX) AML samples were serially passaged in NSG mice and re-isolated for *in vitro* cultivation as previously described[23,45]. Exclusion of mycoplasma contamination was performed continuously during cell culture using the MycoAlert Mycoplasma detection kit (Lonza, Basel, Switzerland). Analysis of patient samples was based on material of AML patients from the AMLCG-99 trial (NCT00266136), AMLCG-2008 trial (NCT01382147), and the Department of Medicine III, University Hospital, LMU. Mononuclear cells were enriched from bone marrow or peripheral blood by Ficoll density gradient centrifugation. Written informed consent for scientific use of sample material was obtained from all patients. The study was performed in accordance with the ethical standards of the responsible committee on human experimentation (written approval by the Research Ethics Boards of the medical faculty of Ludwig-Maximilians-Universität, Munich, number 068-08 and 222-10) and with the Helsinki Declaration of 1975, as revised in 2000. All animal trials were performed in accordance with the current ethical standards of the official committee on animal experimentation (Regierung von Oberbayern, number 55.2-1-54-2531-95-2010 and ROB-55.2Vet-143 2532.Vet_02-16-7).

**Proliferation assay.** Suspension cells were treated with cytarabine (AraC, Selleck Chemicals, Houston, TX, USA), and daunorubicin (in-house). For short time assays, viable cells were treated once (d0) and counted after 72 h on Vi-Cell Cell Viability Analyzer (Beckman Coulter, Krefeld, Germany). For long-term proliferation assays, cells were treated three times (d0, d4, d8) and viable cells were counted every second day. Unpaired, two-tailed Student's *t*-test and calculation of $IC_{50}$ values were performed using GraphPad Prism version 6.07 (GraphPad Software, La Jolla, CA, USA). PiggyBac[23,45,66] (PB)/*EZH2* cells were pre-cultured with or without doxycycline (1µg/mL) for 72 h followed by treatment with AraC +/− doxycycline, which was added every 48 h. For knockdown experiments in PDX cells, siRNA targeting EZH2 (#s4918, Thermo Fisher Scientific, Waltham, USA) was transiently transfected (10 nM) via nucleofection (Supplementary Methods). Cells were pre-incubated for 48 h and then treated with AraC for 72 h.

**Immunoblotting.** Immunoblotting was performed as described before[3]. The following antibodies were used: anti-EZH2 (#5246, Cell Signaling Technology, Danvers, USA), anti-β-actin (A5441, Sigma Aldrich, St. Louis, USA), anti-H3 (ab1791, Abcam, Cambridge, UK), anti-H3K27me3 (#9733, Cell Signaling Technology, Danvers, USA), anti-SUZ12 (#3737, Cell Signaling Technology, Danvers, USA), anti-RbAP46 (#4522, Cell Signaling Technology, Danvers, USA), anit-EED (ab113911, Abcam, Cambridge, UK), anti-EZH1 (#42088, Cell Signaling Technology, Danvers, USA). Western blots were quantified using ImageJ version 1.50d and levels were normalized to the associated loading control (β-actin for EZH2, total H3 for H3K27me3).

***In vivo* therapy trial.** Patient-derived xenograft (PDX) cells expressing enhanced firefly luciferase and mCherry were established as described previously[45]. For *in vivo* therapy trials, $1*10^5$ PDX-AML-491 or $8*10^5$ PDX-AML-661 luciferase-positive cells were injected intravenously into 11 or 16 week old male NSG mice (NOD.Cg-*Prkdc*$^{scid}$ *Il2rg*$^{tm1Wjl}$/SzJ, The Jackson Laboratory, Bar Harbour, ME, USA), and tumour growth was regularly monitored by bioluminescence imaging (BLI) as described previously[24]. 21 days after transplantation, mice were treated with a combination of Cytarabine (AraC; 100 mg/kg, i.p., days 1-4 of therapy weeks) and liposomal daunorubicin (DaunoXome; 1mg/kg, i.v., days 1 and 4 of therapy weeks) every second week for three (AML-661, n = 3) or four (AML-491, n = 4) cycles. Tumour burden was regularly monitored by BLI and compared to untreated control mice. In total, 13 mice were included in this study; one AML-661 control mouse was sacrificed 14 days after injection due to leukaemia unrelated illness. End point of the study was end-stage leukaemia. All animal trials were performed in accordance with the current ethical standards of the official committee on animal experimentation (Regierung von Oberbayern, number 55.2-1-54-2531-95-2010 and ROB-55.2Vet-143 2532.Vet_02-16-7) and in compliance with the ARRIVE guidelines.

**Ethics approval.** We hereby confirm that all experimental protocols were approved by the Department of Medicine III, University Hospital, LMU Munich, the Department of Biology III and Center for Integrated Protein Science Munich (CIPSM); Human Biology and BioImaging, LMU Munich, Planegg Martinsried, Germany and the Helmholtz Zentrum München, Munich.

## Data availability

The RNA-seq data generated for this study is available at GEO under the accession number: GSE162623 The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[67] partner repository with the dataset identifier PXD023139.

Received: 9 October 2019; Accepted: 15 February 2021
Published online: 12 March 2021

## References

1. Bradstock, K. F. *et al.* A randomized trial of high-versus conventional-dose cytarabine in consolidation chemotherapy for adult de novo acute myeloid leukemia in first remission after induction therapy containing high-dose cytarabine. *Blood* **105**, 481–488 (2005).
2. Wang, J. *et al.* Meta-analysis of randomised clinical trials comparing idarubicin + cytarabine with daunorubicin + cytarabine as the induction chemotherapy in patients with newly diagnosed acute myeloid leukaemia. *PLoS One* **8**, e60699 (2013).
3. Hann, I. M. *et al.* Randomized Comparison of DAT Versus ADE as Induction Chemotherapy in Children and Younger Adults With Acute Myeloid Leukemia Results of the Medical Research Council's 10th AML Trial (MRC AML10). *Blood* **89**, 2311–2318 (1997).
4. Marcucci, G. *et al.* Age-related prognostic impact of different types of DNMT3A mutations in adults with primary cytogenetically normal acute myeloid leukemia. *J. Clin. Oncol.* **30**, 742–750 (2012).
5. Schenk, T. *et al.* Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia. *Nat. Med.* **18**, 605–611 (2012).
6. Kanwal, R., Gupta, K. & Gupta, S. Cancer Epigenetics: An Introduction. in *Cancer Epigenetics: Risk Assessment, Diagnosis, Treatment, and Prognosis* (ed. Verma, M.) 3–25 (Springer New York, 2015).
7. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
8. Stahl, M. *et al.* Hypomethylating agents in relapsed and refractory AML: outcomes and their predictors in a large international patient cohort. *Blood Adv* **2**, 923–932 (2018).
9. Cashen, A. F., Schiller, G. J., O'Donnell, M. R. & DiPersio, J. F. Multicenter, phase II study of decitabine for the first-line treatment of older patients with acute myeloid leukemia. *J. Clin. Oncol.* **28**, 556–561 (2010).
10. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
11. Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **20**, 1147–1155 (2013).
12. Chang, C.-J. *et al.* EZH2 promotes expansion of breast tumor initiating cells through activation of RAF1-β-catenin signaling. *Cancer Cell* **19**, 86–100 (2011).

13. Behrens, C. *et al.* EZH2 protein expression associates with the early pathogenesis, tumor progression, and prognosis of non-small cell lung carcinoma. *Clin. Cancer Res.* **19**, 6556–6565 (2013).
14. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).
15. Zingg, D. *et al.* The epigenetic modifier EZH2 controls melanoma growth and metastasis through silencing of distinct tumour suppressors. *Nat. Commun.* **6**, 6051 (2015).
16. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42**, 181–185 (2010).
17. Bödör, C. *et al.* EZH2 mutations are frequent and represent an early event in follicular lymphoma. *Blood* **122**, 3165–3168 (2013).
18. Rabello, D. *et al.* Overexpression of EZH2 associates with a poor prognosis in chronic lymphocytic leukemia. *Blood Cells Mol. Dis.* **54**, 97–102 (2015).
19. D'Angelo, V. *et al.* EZH2 is increased in paediatric T-cell acute lymphoblastic leukemia and is a suitable molecular target in combination treatment approaches. *J. Exp. Clin. Cancer Res.* **34**, 83 (2015).
20. Nishioka, C., Ikezoe, T., Yang, J. & Yokoyama, A. BCR/ABL increases EZH2 levels which regulates XIAP expression via miRNA-219 in chronic myeloid leukemia cells. *Leuk. Res.* **45**, 24–32 (2016).
21. Jankowska, A. M. *et al.* Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. *Blood* **118**, 3932–3941 (2011).
22. Simon, C. *et al.* A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. *Genes Dev.* **26**, 651–656 (2012).
23. Basheer, F. *et al.* Contrasting requirements during disease evolution identify EZH2 as a therapeutic target in AML. *J. Exp. Med.* **216**, 966–981 (2019).
24. Mochizuki-Kashio, M. *et al.* Ezh2 loss in hematopoietic stem cells predisposes mice to develop heterogeneous malignancies in an Ezh1-dependent manner. *Blood* **126**, 1172–1183 (2015).
25. Fujita, S. *et al.* Dual inhibition of EZH1/2 breaks the quiescence of leukemia stem cells in acute myeloid leukemia. *Leukemia* **32**, 855–864 (2018).
26. Neff, T. *et al.* Polycomb repressive complex 2 is required for MLL-AF9 leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5028–5033 (2012).
27. Gu, Z. *et al.* Loss of EZH2 reprograms BCAA metabolism to drive leukemic transformation. *Cancer Discov.* **9**, 1228–1247 (2019).
28. Ernst, T. *et al.* Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat. Genet.* **42**, 722–726 (2010).
29. Abdel-Wahab, O. *et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180–193 (2012).
30. Kim, E. *et al.* SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* **27**, 617–630 (2015).
31. Sashida, G. *et al.* The loss of Ezh2 drives the pathogenesis of myelofibrosis and sensitizes tumor-initiating cells to bromodomain inhibition. *J. Exp. Med.* **213**, 1459–1477 (2016).
32. Wang, J. *et al.* TET2, ASXL1 and EZH2 mutations in Chinese with myelodysplastic syndromes. *Leuk. Res.* **37**, 305–311 (2013).
33. Grossmann, V. *et al.* Molecular profiling of chronic myelomonocytic leukemia reveals diverse mutations in >80% of patients with TET2 and EZH2 being of high prognostic relevance. *Leukemia* **25**, 877–879 (2011).
34. Göllner, S. *et al.* Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. *Nat. Med.* **23**, 69–78 (2017).
35. Greif, P. A. *et al.* Evolution of cytogenetically normal acute myeloid leukemia during therapy and relapse: an exome sequencing study of 50 patients. *Clin. Cancer Res.* **24**, 1716–1726 (2018).
36. Metzeler, K. H. *et al.* Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. *Blood* **128**, 686–698 (2016).
37. Döhner, H. *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* **129**, 424–447 (2017).
38. Li, Z. *et al.* Identification of a 24-gene prognostic signature that improves the European LeukemiaNet risk classification of acute myeloid leukemia: an international collaborative study. *J. Clin. Oncol.* **31**, 1172–1181 (2013).
39. Wouters, B. J. *et al.* Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood* **113**, 3088–3091 (2009).
40. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
41. Papaemmanuil, E. *et al.* Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616–27 (2013).
42. Stief, S. M. *et al.* Loss of KDM6A confers drug resistance in acute myeloid leukemia. *Leukemia* https://doi.org/10.1038/s41375-019-0497-6 (2019).
43. Fu, Y. *et al.* Genome-wide identification of FHL1 as a powerful prognostic candidate and potential therapeutic target in acute myeloid leukaemia. *EBioMedicine* **52**, 102664 (2020).
44. Luo, H. *et al.* Microarray-based analysis and clinical validation identify ubiquitin-conjugating enzyme E2E1 (UBE2E1) as a prognostic factor in acute myeloid leukemia. *J. Hematol. Oncol.* **9**, 125 (2016).
45. Vick, B. *et al.* An advanced preclinical mouse model for acute myeloid leukemia using patients' cells of various genetic subgroups and in vivo bioluminescence imaging. *PLoS One* **10**, e0120925 (2015).
46. Zhang, J., Gu, Y. & Chen, B. Mechanisms of drug resistance in acute myeloid leukemia. *Onco. Targets. Ther.* **12**, 1937–1945 (2019).
47. Guglielmelli, P. *et al.* EZH2 mutational status predicts poor survival in myelofibrosis. *Blood* **118**, 5227–5234 (2011).
48. Wu, H. *et al.* Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. *PLoS One* **8**, e83737 (2013).
49. Nikoloski, G. *et al.* Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nat. Genet.* **42**, 665–667 (2010).
50. Poepsel, S., Kasinath, V. & Nogales, E. Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. *Nat. Struct. Mol. Biol.* **25**, 154–162 (2018).
51. Papaemmanuil, E. *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).
52. Grubach, L. *et al.* Gene expression profiling of Polycomb, Hox and Meis genes in patients with acute myeloid leukaemia. *Eur. J. Haematol.* **81**, 112–122 (2008).
53. Wen, S. *et al.* Novel combination of histone methylation modulators with therapeutic synergy against acute myeloid leukemia in vitro and in vivo. *Cancer Lett.* **413**, 35–45 (2018).
54. Gaidzik, V. I. *et al.* RUNX1 mutations in acute myeloid leukemia are associated with distinct clinico-pathologic and genetic features. *Leukemia* **30**, 2160–2168 (2016).
55. Rinke, J. *et al.* Molecular characterization of EZH2 mutant patients with myelodysplastic/myeloproliferative neoplasms. *Leukemia* **31**, 1936–1943 (2017).
56. Booth, C. A. G. *et al.* Ezh2 and Runx1 mutations collaborate to initiate lympho-myeloid leukemia in early thymic progenitors. *Cancer Cell* **33**, 274-291.e8 (2018).

57. Ueda, T. *et al.* EED mutants impair polycomb repressive complex 2 in myelodysplastic syndrome and related neoplasms. *Leukemia* **26**, 2557–2560 (2012).
58. Khan, S. N. *et al.* Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. *Leukemia* **27**, 1301–1309 (2013).
59. Zhu, B. *et al.* Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol. Cell* **20**, 601–611 (2005).
60. Nair, V. A., Al-Khayyal, N. A., Sivaperumal, S. & Abdel-Rahman, W. M. Calponin 3 promotes invasion and drug resistance of colon cancer cells. *World J. Gastrointest. Oncol.* **11**, 971–982 (2019).
61. Bentin Toaldo, C. *et al.* Protein Kinase A-induced tamoxifen resistance is mediated by anchoring protein AKAP13. *BMC Cancer* **15**, 588 (2015).
62. Alves, C. P. *et al.* Myosin-Va contributes to manifestation of malignant-related properties in melanoma cells. *J. Invest. Dermatol.* **133**, 2809–2812 (2013).
63. Hinz, N. & Jücker, M. Distinct functions of AKT isoforms in breast cancer: a comprehensive review. *Cell Commun. Signal.* **17**, 154 (2019).
64. D'Agostino, L. & Giordano, A. A novel dual signaling axis for NSP 5a3a induced apoptosis in head and neck carcinoma. *Oncotarget* **2**, 1055–1074 (2011).
65. Ha, M. *et al.* Prognostic role of TPD52 in acute myeloid leukemia: a retrospective multicohort analysis. *J. Cell. Biochem.* **120**, 3672–3678 (2019).
66. Mulholland, C. B. *et al.* Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive demethylation in mammals. *bioRxiv* 321604 (2020) https://doi.org/10.1101/321604.
67. Perez-Riverol Y, *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**(D1), D442–D450 (2019).


## Acknowledgements

The authors thank all study participants. We thank M. Fritschle for animal handling. We thank Bianka Ksienzyk for cell sorting. We thank Karl-Peter Hopfner for the nice cooperation within mutation characterisation. We thank Dr. Igor Paron and Dr. Christian Deiml from the Proteomics and Signal Transduction Group and the Clinical Proteomics Group at the Max Planck Institute of Biochemistry, Martinsried, for their technical assistance and Michael Wierer und Matthias Mann for their support. K.H.M., P.A.G., G.S., H.L., I.J., and K.S. were supported by the German Research Council (DFG) within the SFB 1243 "Cancer Evolution". J.M.K., R.M., K.V., L.E.W., M.R. and S.W. are members of the IRTG-1243 within the SFB 1243 of the DFG. TH was supported by the Wilhelm-Sander-Stiftung (Grant 2013.086.2) and the Physician Scientists Grant (G-509200-004) from the Helmholtz Zentrum München. E.U. is a fellow of the International Max Planck Research School for Molecular Life Sciences (IMPRS-LS) and is supported by the research training group 1721 (RTG 1721, DFG).

## Author contributions

K.S., J.M.K. and S.W. conceived the study. J.M.K. and S.W. wrote the manuscript and analysed the data. J.M.K. performed all experiments which are not mentioned below. M.D.B. and S.B. supported and supervised CRISPR/Cas9 mediated genome editing experiments. R.M., K.V., M.R., M.F. and H.L. supported the experiments. L.E.W. performed RNA-seq experiments. Mass spectrometry-based proteomics experiments were designed by S.W., M.R. and E.U., performed by E.U. and analysed by S.W. and E.U. B.V. supported and supervised all mice and PDX experiments. K.H.M. performed mutation analysis in AML diagnosis. T.H. performed survival analysis and provided supplementary Fig. 1d. M.S. performed and evaluated MLPA analysis. K.S., I.J., G.S., W.H. and O.W. interpreted the data and supported the project by coordinating the teams and experiments. All authors reviewed the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-84708-6.

**Correspondence** and requests for materials should be addressed to K.S.

**Reprints and permissions information** is available at www.nature.com/reprints.

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

# Supplementary Information

## LOSS-OF-FUNCTION MUTATIONS IN THE HISTONE METHYLTRANSFERASE EZH2 PROMOTE CHEMOTHERAPY RESISTANCE IN AML

Julia M. Kempf[1,†], Sabrina Weser[1,†], Michael D. Bartoschek[2], Klaus H. Metzeler[1], Binje Vick[3], Tobias Herold[1], Kerstin Völse[3], Raphael Mattes[1], Manuela Scholz[4], Lucas E. Wange[8], Moreno Festini[1], Enes Ugur[2], Maike Roas[1], Oliver Weigert[1], Sebastian Bultmann[2], Heinrich Leonhardt[2], Gunnar Schotta[5], Wolfgang Hiddemann[6,9], Irmela Jeremias[3,6,7] and Karsten Spiekermann[1,6,9,*]

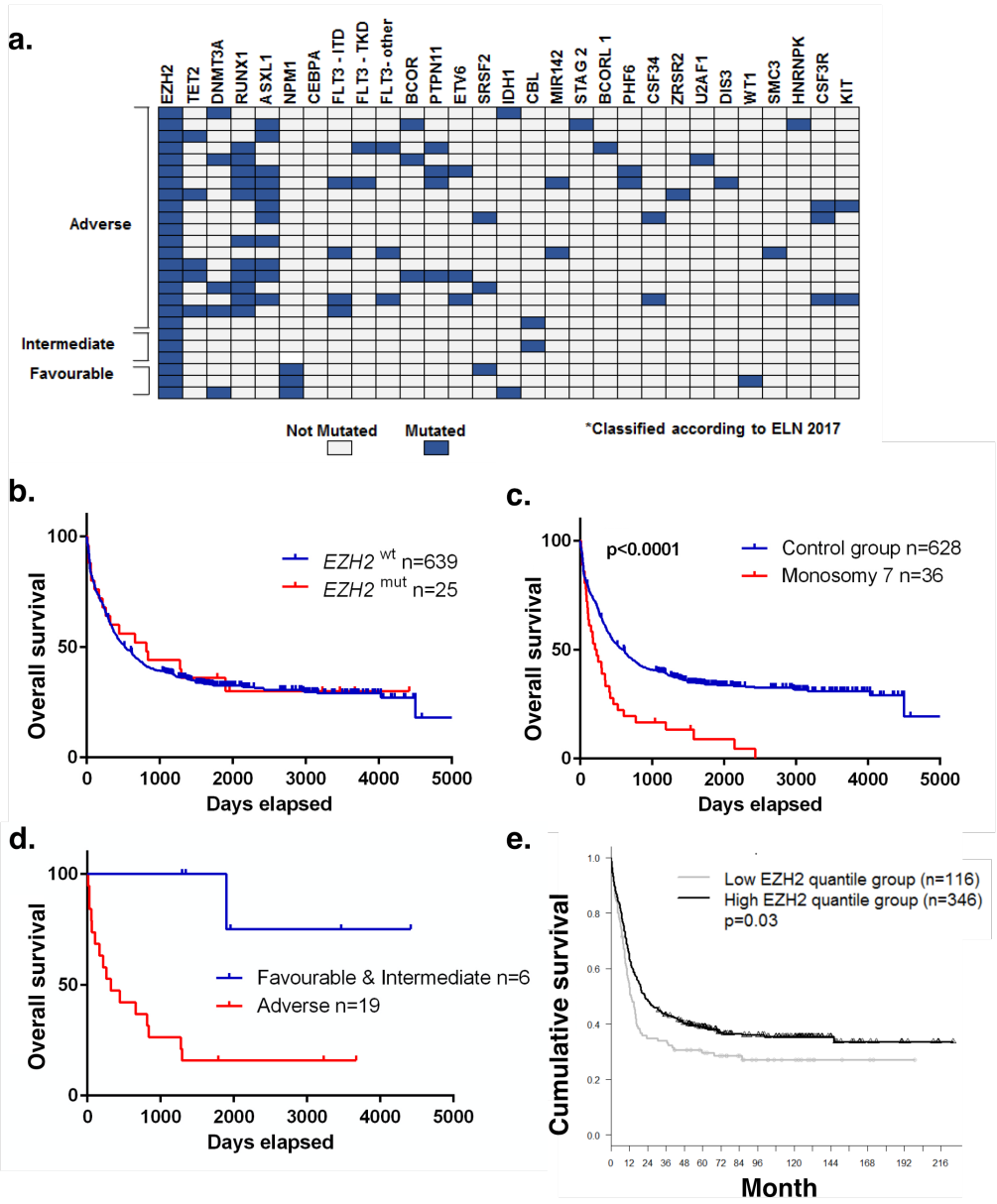
[1]Department of Medicine III, University Hospital, LMU Munich, Munich, Germany. [2]Department of Biology II and Center for Integrated Protein Science Munich (CIPSM), Human Biology and BioImaging, LMU Munich, Planegg Martinsried, Germany. [3]Research unit Apoptosis in Haematopoietic Stem Cells (AHS), Helmholtz Zentrum München, Munich, Germany. [4]Center for Human Genetics and Laboratory Diagnostic (AHC). [5]Biomedical Center and Center for Integrated Protein Science Munich, LMU Munich, Martinsried, Germany. [6]German Cancer Consortium (DKTK). [7]Department of Pediatrics, Dr. von Hauner Children's Hospital, LMU Munich, Munich, Germany. [8]Anthropology & Human Genomics, Department of Biology II, Ludwig-Maximilians-University, Martinsried, Germany. [9]German Cancer Research Center (DKFZ), Heidelberg, Germany.

† These authors contributed equally

* Corresponding author: karsten.spiekermann@med.uni-muenchen.de

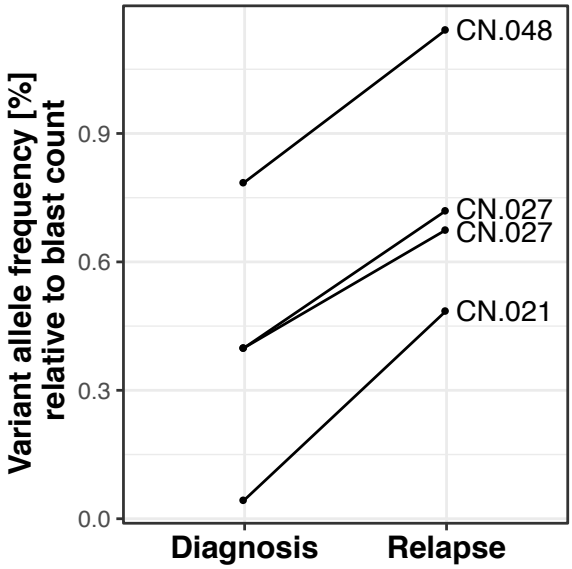
**Table of Contents**

## Supplementary Figure 1



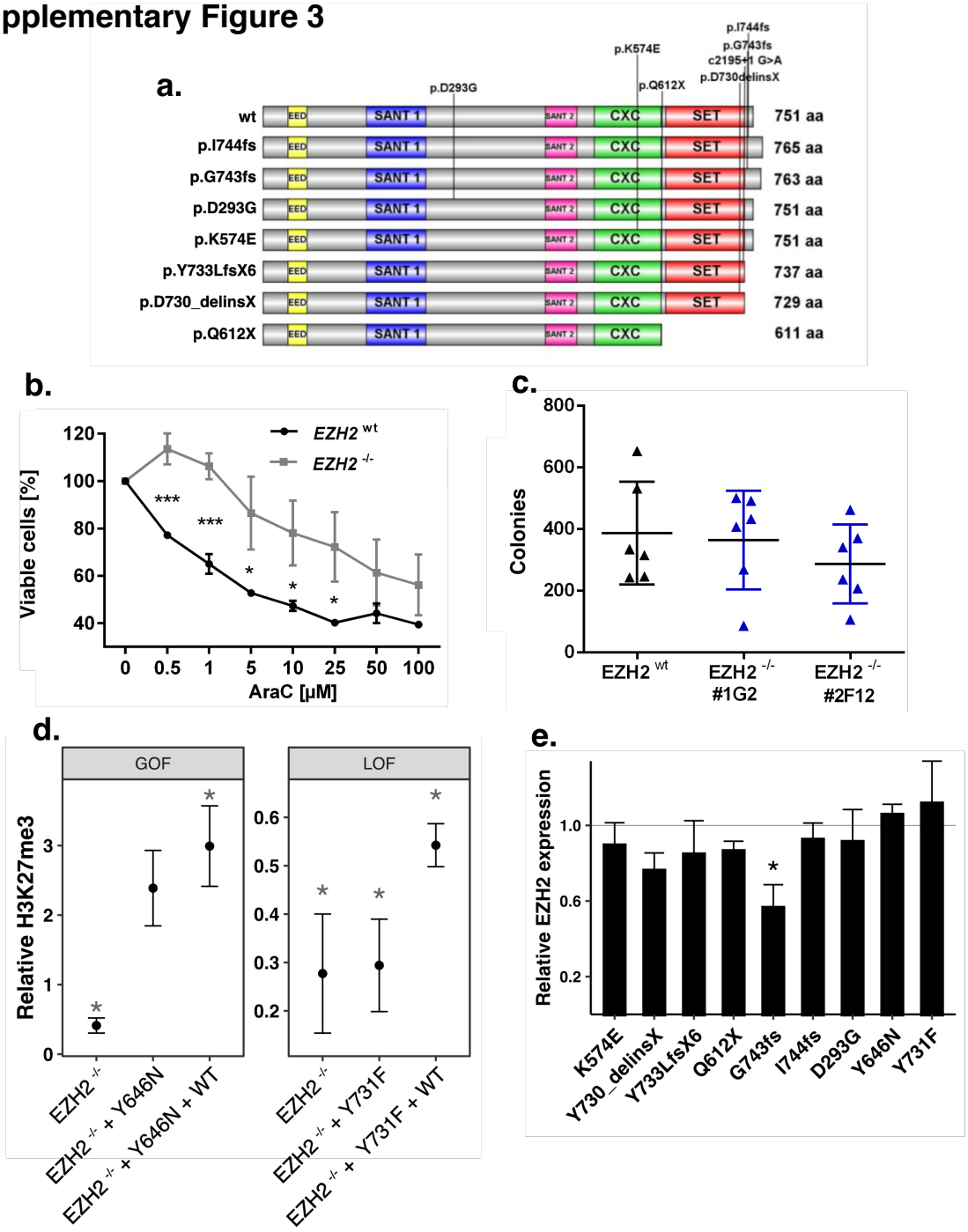

**Supplementary Figure 1: Related to Figure 1. Recurrent EZH2 mutations**
**(a),** Mutation pattern of the 25 AML patients harboring EZH2 mutations. Classification according to the ELN- 2017 recommendations. Status also includes multiple mutations in one patient. **(b-d),** Overall survival (OS) of **(b)** patients with EZH2 mutation **(c)** patients with monosomy 7 and **(d)** patients with EZH2 mutation and adverse phenotype (patients from the Metzeler et al. data set). P-value calculated with log-rank test. **(e)** Cumulative overall survival of patients with high or low EZH2 expression of the HOVON dataset (GSE14468). EZH2 high and low groups are defined by the upper and lower quantile. P-value calculated with log-rank test.

## Supplementary Figure 2



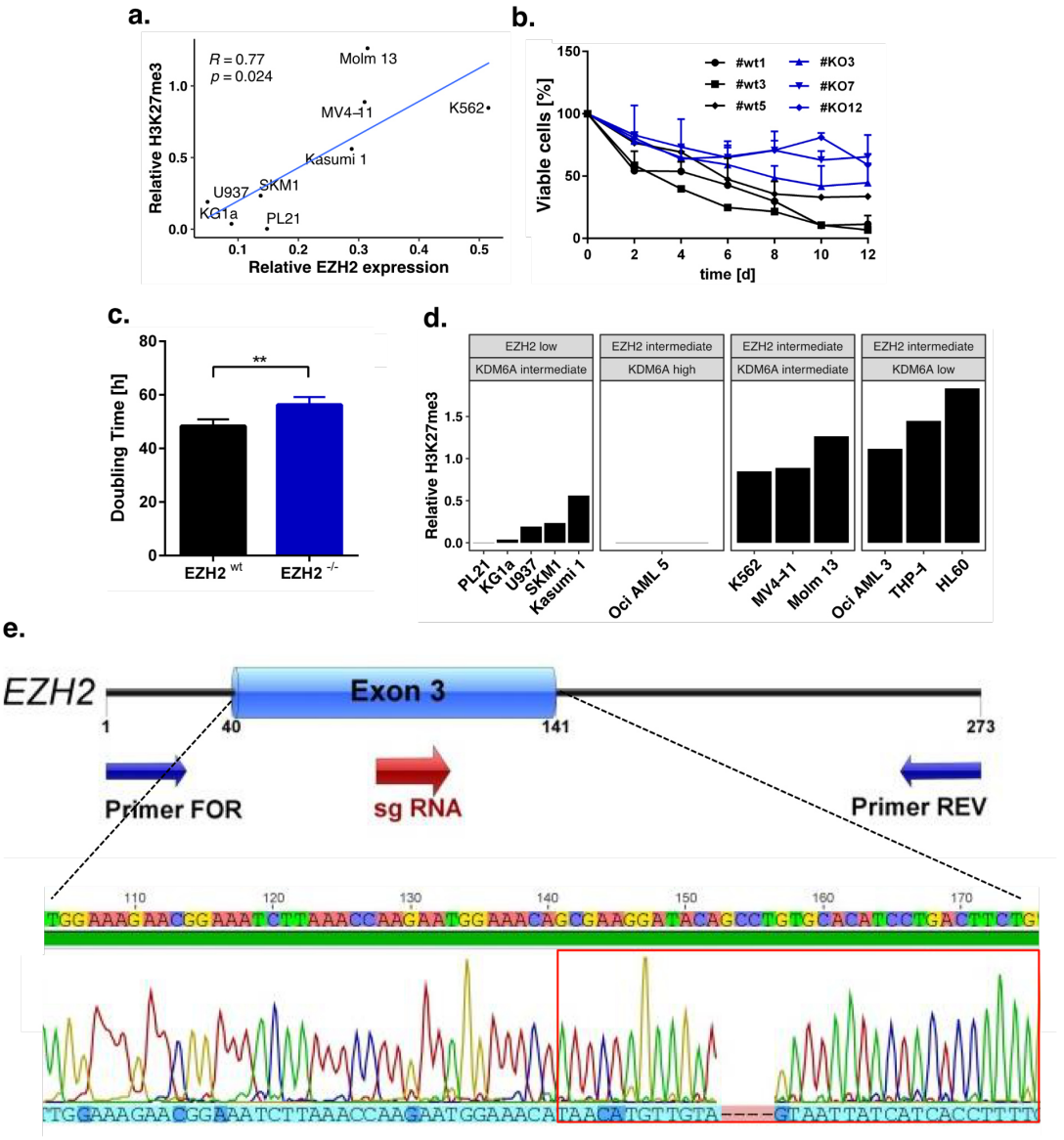

**Supplementary Figure 2: Related to Figure 2. Relevance of EZH2 mRNA Expression in AML relapse.** Variant allele frequencies (VAF) of EZH2 mutations found in three patients of the Greif et al. cohort. Patient CN.027 harbored two EZH2 mutations. Variant allele frequency was corrected by the respective blast count.

## Supplementary Figure 3



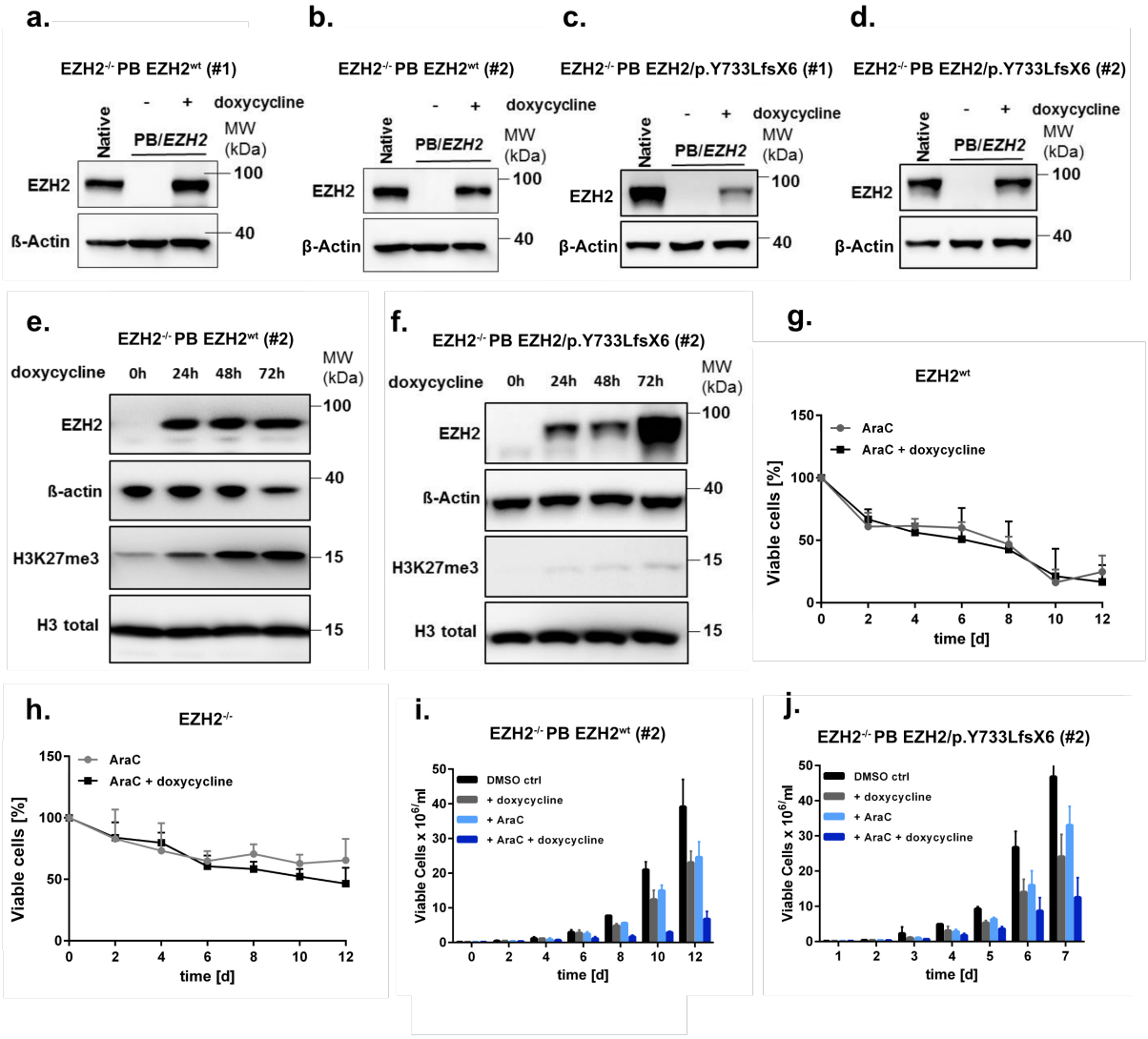

**Supplementary Figure. 3: Related to Figure. 3. Evaluation of EZH2 mutations**. **(a)** Structural aberrations of investigated EZH2 mutations. **(b)** AraC resistance in one 293T/EZH2$^{-/-}$ and one 293T/EZH2$^{wt}$ sc clone (#1G2, #1D12). Cells were treated for 72h with AraC of different concentrations. Mean for three independent experiments. Viable cells relative to untreated control. **(c)** Colony formation of three 293T sc clones. Mean for 6 independent experiments. **(d)** H3K27me3 levels of the previously described EZH2/p.Y646N and Y731F variants leading to GOF and LOF, respectively. Values relative to WT. Mean for four/three independent experiments. **(e)** EZH2 protein expression of the EZH2 variants. Mean for four independent experiments. Values relative to WT. Unpaired, two-tailed Student's t-test; *p<0.05; **p<0.01; ***p<0.001. Error bars represent standard deviation (s.d).

## Supplementary Figure 4



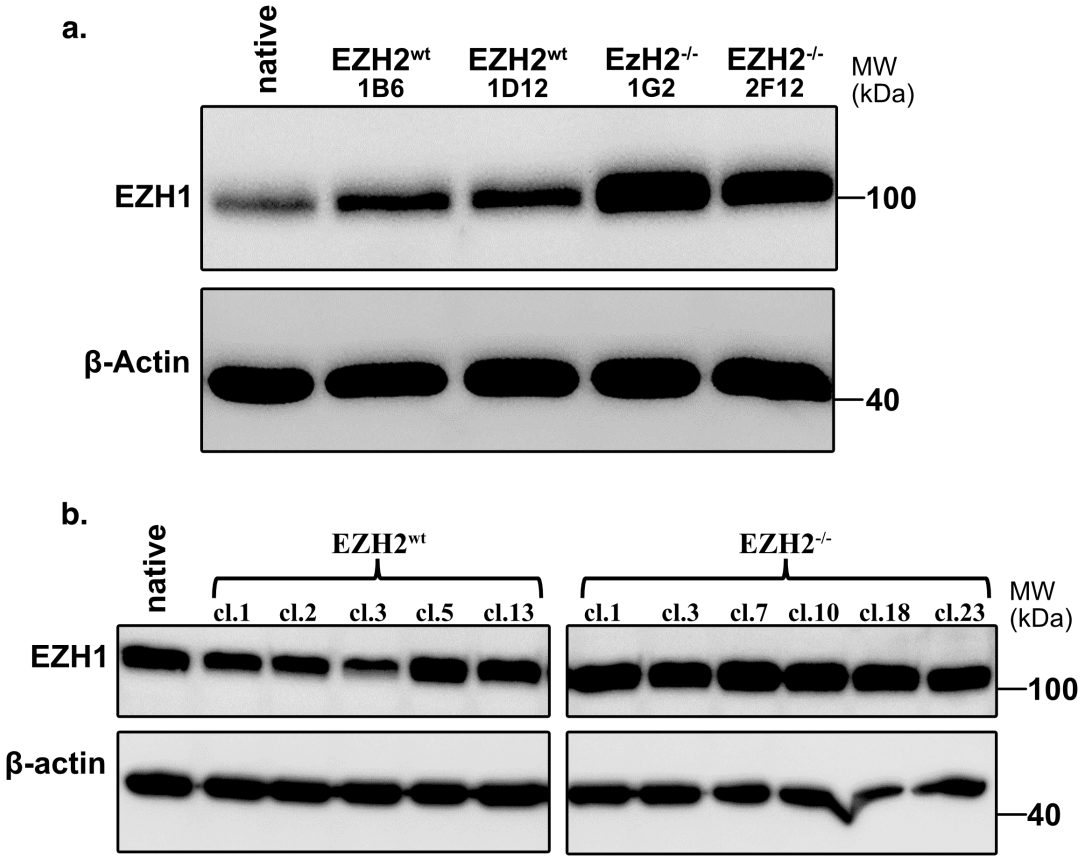

**Supplementary Figure 4: Related to Figure 4. EZH2 depletion promotes resistance in the myeloid cell line K562**. **(a)** Correlation between EZH2 protein expression and global H3K27me3 in eight hematopoietic cell lines (Pearson's correlation). **(b)** Long-term low dose AraC treatment in EZH2$^{wt}$ (n=3) and EZH2$^{-/-}$ (n=3) clones. Cells were treated with 30 nM AraC/DMSO for 12 days. Cells were split and treated every four days if necessary. Mean ± s.d are given for three independent experiments. **(c)** Doubling time of EZH2$^{wt}$ (n=6) and EZH2$^{-/-}$ (n=7) sc clones. Each clone represents the mean of three independent experiments. **(d)** Combined influence of EZH2 and KDM6A expression on the H3K27me3 levels in the 12 hematopoietic cell lines. Low and High group defined as < 0.3 and > 0.7 relative protein expression. **(e)** Schematic representation of exon 3 Sanger sequencing for CRISPR/Cas9 screening. Mutation is idicated by red box.

## Supplementary Figure 5



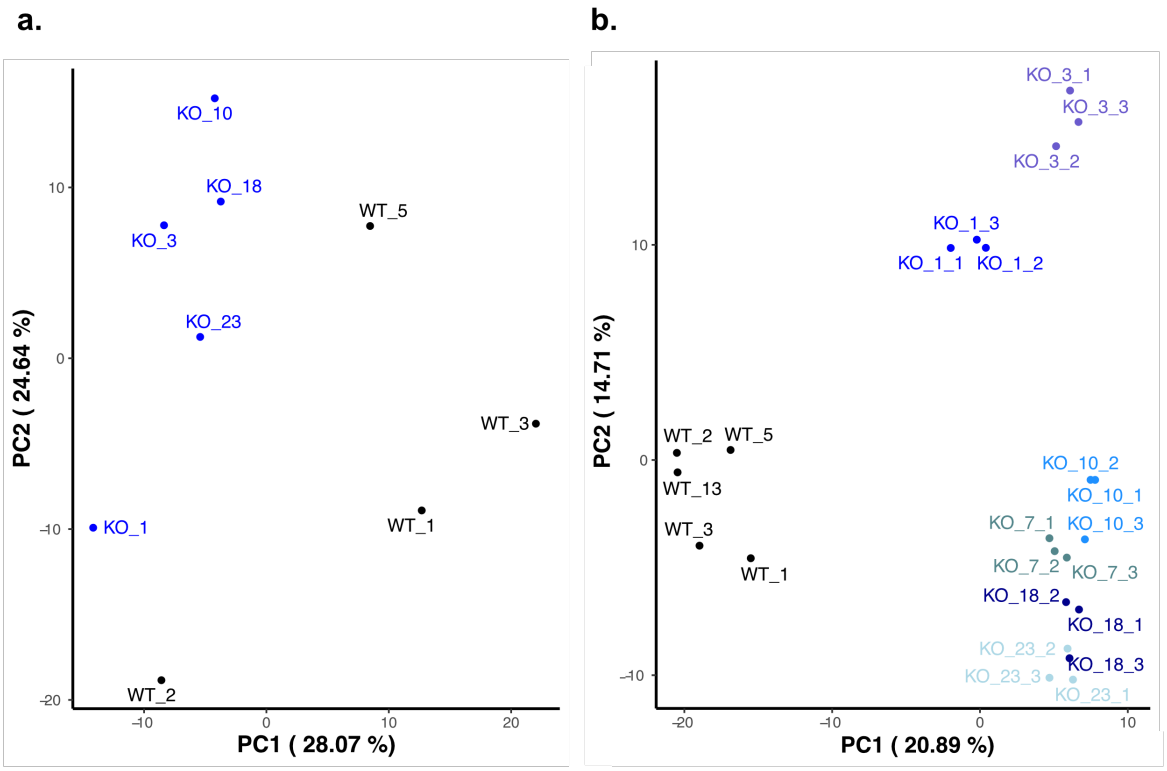

**Supplementary Figure 5: Related to Figure 5. Conditional EZH2 re-expression in the myeloid cell line K562 via PiggyBac**. **(a-d)** Immunoblot for EZH2 expression with or without doxycycline induction in **(a-b)** EZH2$^{-/-}$ PB EZH2$^{wt}$ and **(c-d)** EZH2$^{-/-}$ PB EZH2/p.Y733LfsX6 cells. Respective clone given in the plot title. **(e-f)** Immunoblot for EZH2 expression and global H3K27me3 in **(e)** EZH2$^{-/-}$ PB EZH2$^{wt}$ cells (clone #2) and **(f)** EZH2$^{-/-}$ PB EZH2/p.Y733LfsX6 (clone #2) after 0 h, 24 h, 48 h and 72 h of doxycycline induction. Cells were treated with 1 $\mu$g/ml doxycycline every 24 h. **(g-h)** Long-term proliferation under AraC treatment (30 nM) with or without doxycycline in **(g)** EZH2$^{wt}$ and **(h)** EZH2$^{-/-}$ cells. Cells were split every 96 h if necessary. Doxycycline was added every 48 h. **(i-j)** Comparison of proliferation with or without indicated treatment and DMSO control in **(i)** EZH2$^{-/-}$ PB EZH2$^{wt}$ and **(j)** EZH2$^{-/-}$ PB EZH2/p.Y733LfsX6 cells. MW, molecular weight; β-actin, loading control. Error bars indicate mean ± s.d. of three independent experiments.

## Supplementary Figure 6




**Supplementary Figure 6: Expression of EZH1. (a-b)** Immunoblot for EZH1 expression of EZH2$^{wt}$ and EZH2$^{-/-}$ clones in **(a)** HEK293T and **(b)** K562 cells. MW, molecular weight. β-actin, loading control.

## Supplementary Figure 7




**Supplementary Figure 7: Principal component analysis (PCA)** of **(a)** RNA-seq and **(b)** full proteome measurements. Altogether five wt K562 clones and six EZH2$^{-/-}$ clones were measured. Additionally, measurement of the proteome was performed in technical triplicates. Technical triplicates cluster closely together, indicating high reproducibility of the measurements. Loss of EZH2 causes aberrant transcript and protein expression, visible in principle component 1 (PC1), the component harbouring the highest variance. PCA on the basis of the 500 most variable genes or proteins. Raw gene count data was normalised with the RLE method in edgeR and scaled to upm (UMIs per million).

## Supplementary Methods

### Evaluation of global H3K27me3 levels in HEK293T

*EZH2* mutant or wildtype constructs were transfected transiently with lipofectamine 3000 (Invitrogen) in 293T/*EZH2*$^{-/-}$ cells. After 72 h, nuclear or whole cell lysates were generated and global H3K27me3 levels were evaluated by immunoblotting.

### Transient transfection with siRNA in PDX cells

PDX cells were transfected by electroporation with siRNA using the Neon Transfection System (Thermo Fisher Scientific) with the following parameters: 2200V, 20 ms, 1 pulse. The following siRNA was used: #s4918 EZH2 silencer select (pre-designed by Thermo Fisher Scientific, Waltham, US).

### CRISPR/Cas9 mediated genome editing

*EZH2* specific sgRNA was cloned into pSpCas9(BB)-2A-GFP (PX458, a gift from Feng Zhang, Addgene #48138). The sgRNA was designed using Benchling (Biology Software, 2018). After 48 h, GFP positive cells were enriched and single-cell sorted into 96-well V-bottom plates (K562) and in 30 cm dishes (HEK293T) with the FACSVantage SE. For HEK293T cells the colonies were separated manually using a 20 µL sterile pipette tip and transferring each cell clone into single wells of a 96-well plate. Cells were cultured at 37°C with 5 % $CO_2$ for several weeks, until colony formation was observed and sc clones were expanded. To screen for *EZH2* loss, cells were lysed and amplified exon 3 PCR products[1] were sequenced by Sanger sequencing. In short, gDNA was isolated with the QIAamp DNA Blood Mini Kit *(Qiagen, Hilden, Germany)*, in a 96-well plate, re-suspended in 50µL (each well) lysis buffer SC, frozen at -80°C for 30 min, incubated at 56°C for 3h and finally Proteinase K heat inactivated at 85°C for 30 min. 2.5 µL/well of this lysate were directly subjected to PCR mixture (25 µL/rxn., 0.1 µL MyTaq DNA Polymerase) and PCR (FOR: ACAATTTCTCCTTTCCTCTCCTTCA, REV: TGGACACCCTGAGGTCAATGAT) was performed under following conditions: (95°C/5 min- [95°C/30 s - 61°C/30 s - 72°C/30 s] x 45 - 72°C/40 s - 4°C/∞). *EZH2* knockout clones were identified by restriction-fragment length polymorphism (RFLP) analysis of PCR products using HpyAVI. Enzyme recognition site is lost after successful CRISPR/Cas9 targeting. Expression loss was confirmed by immunoblotting for each clone.

**PiggyBac/Transposase genome editing**

The human *EZH2* fragment was cloned into the BmtI and NotI linearized pPBtet-3xFLAG-IRES-DsRed-Express-PuroR vector[2]. (3xFLAG was removed by linearization) using the In-Fusion HD Cloning Plus Kit (Takara Bio, Saint-Germain-en-Laye, France) generating pPBtet-EZH2-IRES-DsRed-Express-PuroR. For inducible re-expression of EZH2 in *EZH2*$^{-/-}$ K562 cell lines, cells were first nucleofected with equimolar amounts of pPBtet-EZH2-IRES-DsRed-Express-PuroR and PiggyBac transposase-expression vector (Biocat, PB200PA-1). Two days after transfection, cells were subjected to puromycin selection (2 µg/mL) for 3 days. Viable cells were enriched by single-cell sorting into 96-well V-bottom plates with FACSVantage SE (BD Bioscience). Cells were cultured with puromycin (2 µg/mL) until colonies were readily visible. To confirm inducible EZH2 re-expression, clones were treated with doxycycline (1 µg/mL) for a minimum of 24 h and EZH2 expression was analyzed by Western blot.

**Generation of inducible PiggyBac *EZH2* cell lines**

To generate stable cell lines carrying doxycycline-inducible *EZH2*, cells were first nucleofected with equimolar amounts of pPBtet-EZH2-IRES-DsRed-Express-PuroR and PiggyBac transposase (Biocat, PB200PA-1). Two days after transfection, cells were subjected to puromycin selection (2 µg/mL) for 3 d. Viable cells were enriched by single-cell sorting into 96-well V-bottom plates with FACSVantage SE (BD Bioscience). Cells were cultured with puromycin (2 µg/mL) until colonies were readily visible. To screen for successful EZH2 re-expression, clones were treated with doxycycline (0.5 µg/mL) for 48 h and inducible EZH2 expression was analyzed by Western Blot.

**Library preparation and sequencing of K562 clones**

RNA sequencing was performed using prime-seq, a bulk version of the single cell RNA-seq method mcSCRB-seq[3]. Briefly, 10,000 cells per sample were lysed in 100 µL RLT Plus Buffer (Qiagen) supplemented with 1 % beta mercapto-ethanol. Following a Proteinase K (Ambion) digest, nucleic acids were isolated using SPRI Beads and a DNAseI (Thermo Scientific) digest was performed on beads. Reverse transcription of the isolated RNA was performed using barcoded oligo-dT primers and template switching oligo. Enabled by the barcoding in the RT, samples were pooled and second strand synthesis as well as amplification of cDNA was performed using Kapa HiFi HotStart polymerase (Roche, Basel,

Switzerland). After quality control of the cDNA using capillary gel electrophoresis (Bioanalyzer, Agilent, Santa Clara, USA), sequencing library preparation was performed using the NEBNext Kit (New England Biolabs, Ipswich, USA). A quarter of the cDNA was used for the fragmentation reaction and a custom adapter was used for ligation. Next, double size selection using SPRISelect beads (Beckmann Coulter, Brea, USA) was performed to only retain fragments between 300 bp and 900 bp. The full step-by-step protocol including primer sequences is accessible at protocols.io (https://www.protocols.io/view/prime-seq-s9veh66). The final library was paired-end sequenced on one HiOut lane of an Illumina HiSeq1500 instrument, with 28 bp in the first read covering the barcode and UMI and 50 bp in the second read covering the cDNA fragment. Raw fastq files were processed using the zUMIs pipeline[4].

**Sample preparation for full proteome measurements**

Full proteome measurements were performed in technical triplicates for each of the six individual *EZH2*-KOs. In addition, 5 WT K562 cells were measured as a comparison. For each technical replicate 5 x $10^6$ cells were lysed in Guanidinium Chloride-based lysis buffer (6 M Guanidinium Chloride, 100 mM Tris-HCl pH 8.5 and freshly supplemented 2 mM DTT). Cell pellets were homogenized by pipetting and boiled for 10 min at 99°C under constant shaking at 1,700 rpm. Quickly after boiling samples were addressed to sonication for 15 min (30 s on/off interval, Bioruptor Plus by Diagenode). To digest the same amounts of protein lysates, the protein concentrations were measured by a BCA assay. Meanwhile chloroacetamide was added to the samples to a final concentration of 40 mM. After at least 20 min of incubation at room temperature 30 µg of each lysate was diluted 1:10 in the digestion buffer (25 mM Tris–HCl pH 8.5 and 10% acetonitrile). Trypsin and LysC were added in a 1:100 protease to protein ratio. Samples were incubated overnight at 37°C and kept under constant shaking at 1000 rpm in a thermal shaker. The next day samples were acidified to stop the digestion with 1% trifluoroacetic acid (TFA), peptide mixtures were subsequently cleaned up on three layers of SDB-RPS[5]. After elution of peptides, samples were dried in a speedvac and resuspended in 20 µL of A* buffer (0.1% TFA and 2% acetonitrile). Prior to LC-MS/MS analysis peptide concentrations were estimated by nanodrop at 280 nm.

**Full proteome measurements based on data-independent acquisition method**

500 ng peptides of each replicate were separated by nanoflow high-pressure liquid chromatography on an Easy-nLC 1200 (Thermo Fisher Scientific) using in-house packed 50 cm C18 columns (ReproSil-Pur C18-AQ 1.9 µm resin, Dr. Maisch GmbH) and subsequently injected via a nano-electrospray ion source. The peptides were eluted from the column in an acetonitrile gradient for 120 min while the flow rate was fixed to approximately 300 nL/min and the column oven temperature to 60°C. Mass spectrometric analysis of peptides was performed on an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific, Bremen, Germany).

Data acquisition was performed in data-independent mode (DIA). For this the MS1 resolution was set to 120,000 (mass range of 300 to 1,650 m/z, maximum injection time 60 ms and AGC target 3E6) whereas the MS2 resolution was set to 30k (mass range of 361 to 1,033 m/z, 1 Da window overlap, 30 DIA windows, 1,309 ms cycle time, 22 ms MS2 injection time).

Raw files were processed by the Spectronaut Pulsar X software package (Biognosys, version 14) applying the default Biognosys factory settings for directDIA analysis. As a search basis a reference human proteome (2019) provided by Uniprot was used.

**Statistical analysis of Proteome and Transcriptome data**

Differential expression analysis was performed in R[6] using limma[7]. For RNA-seq data this was combined with voom[8]. RNA-seq count data was scaled using edgeR[9] and genes with a read count below ten in all samples were excluded. Due to insufficient library size, the samples WT_13 and KO_7 were excluded from the RNA-seq analysis.

**MLPA**

To identify microdeletions and numerical aberrations in patient samples and PDX cells, the SALSA MLPA MDS Kit from MRC Holland (Amsterdam, Netherlands) was applied. All steps were carried out as recommended by the supplier. 50 ng DNA was used as input and gDNA was generated using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). Data were analysed using the ABI 3730 XL sequencer and Genemarker V2.6.0. software from Softgenetics LLC. (State College, US).

**Statistical evaluation of Patient Survival**

The association of *EZH2* mRNA expression with clinical, genetic and outcome variables was analyzed in publicly available data sets. Patients of the AMLCG1999 trial were used as discovery cohort (GSE37642)[10,11]and results were validated in patients intensively treated in trials of the HOVON (Haemato Oncology Foundation for Adults in the Netherlands) group (GSE14468)[12,13]. Maximally selected rank statistics[14] was used to dichotomize *EZH2* expression in the discovery cohort. The identified cut point was independently validated in the HOVON data set. Kaplan-Meier estimates for overall survival (OS) were calculated using the R survival package with standard parameters. p-values were calculated using the log-rank test. Statistical analysis was performed using the R-3.4.1 software package[6]

## Supplementary References

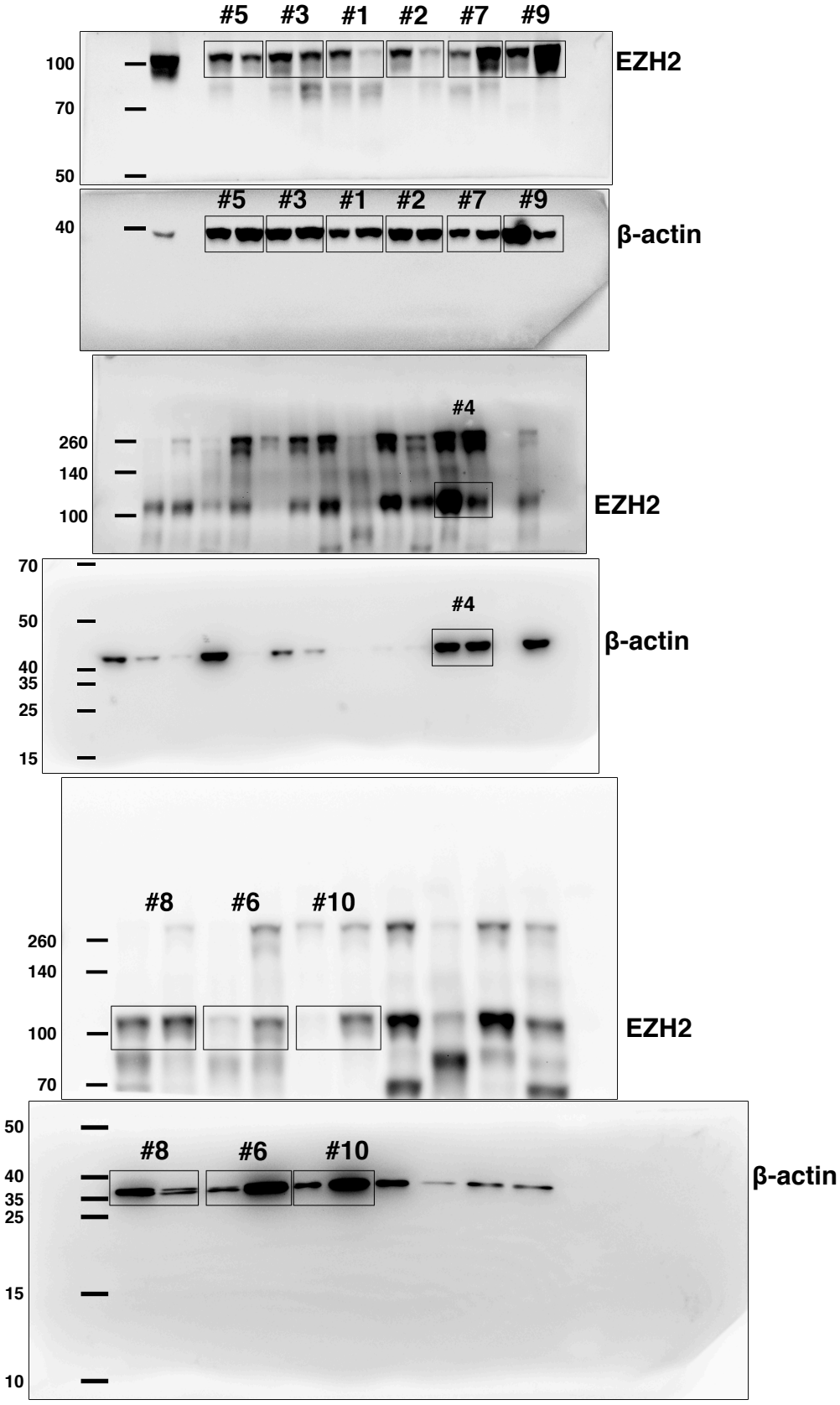

1. Mulholland, C. B. *et al.* A modular open platform for systematic functional studies under physiological conditions. *Nucleic Acids Res.* **43**, e112 (2015).
2. Stief, S. M. *et al.* Loss of KDM6A confers drug resistance in acute myeloid leukemia. *Leukemia* (2019) doi:10.1038/s41375-019-0497-6.
3. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun.* **9**, 2937 (2018).
4. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience* **7**, (2018).
5. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
6. Team, R. C. & Others. R: A language and environment for statistical computing. (2013).
7. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
8. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
9. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
10. Büchner, T. *et al.* Double induction containing either two courses or one course of high-dose cytarabine plus mitoxantrone and postremission therapy by either autologous stem-cell transplantation or by prolonged maintenance for acute myeloid leukemia. *J. Clin. Oncol.* **24**, 2480–2489 (2006).
11. Herold, T. *et al.* Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. *Blood* **124**, 1304–1311 (2014).
12. Wouters, B. J. *et al.* Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood* **113**, 3088–3091 (2009).
13. Taskesen, E. *et al.* Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients:

further evidence for CEBPA double mutant AML as a distinctive disease entity. *Blood* **117**, 2469–2475 (2011).

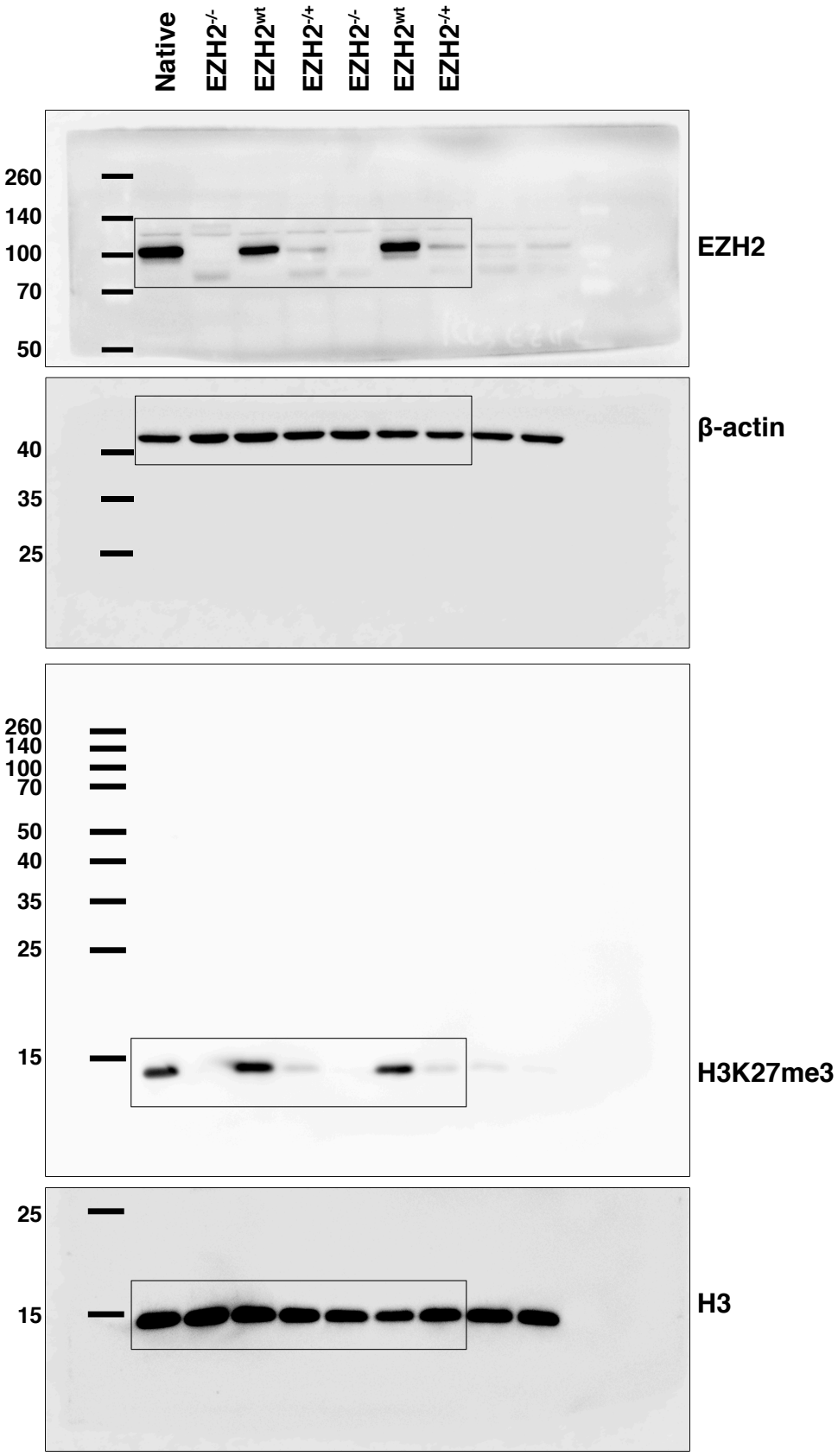
14. Hothorn, T. & Lausen, B. On the exact distribution of maximally selected rank statistics. *Comput. Stat. Data Anal.* **43**, 121–137 (2003).

## Uncropped images of western blots

## Figure 2a



## Figure 3a


Native
EZH2-/-
EZH2wt
EZH2-/+
EZH2-/-
EZH2wt
EZH2-/+
260
140
100
70
50
EZH2
40
35
25
β-actin
260
140
100
70
50
40
35
25
15
H3K27me3
25
15
H3

## Figure 3c

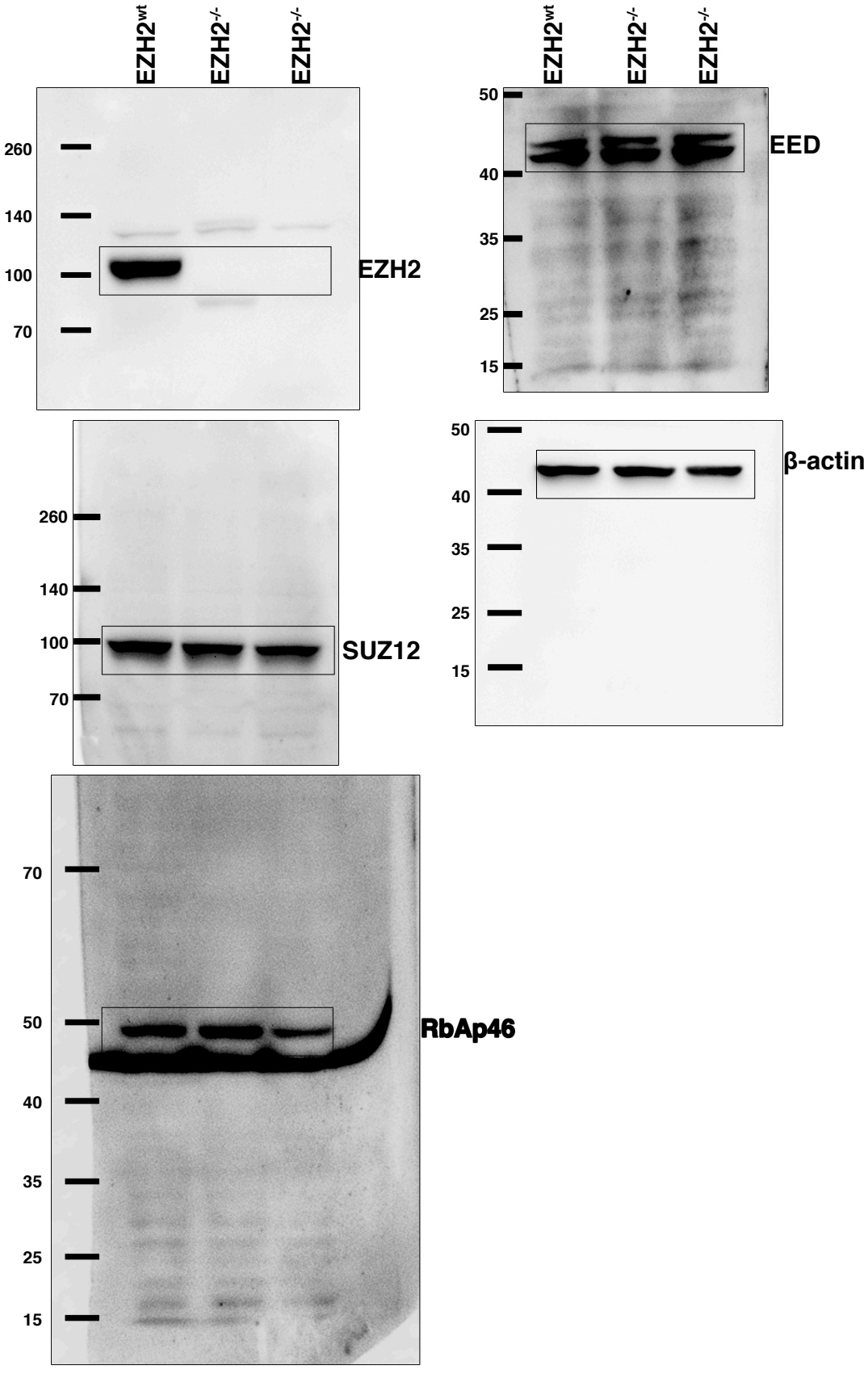
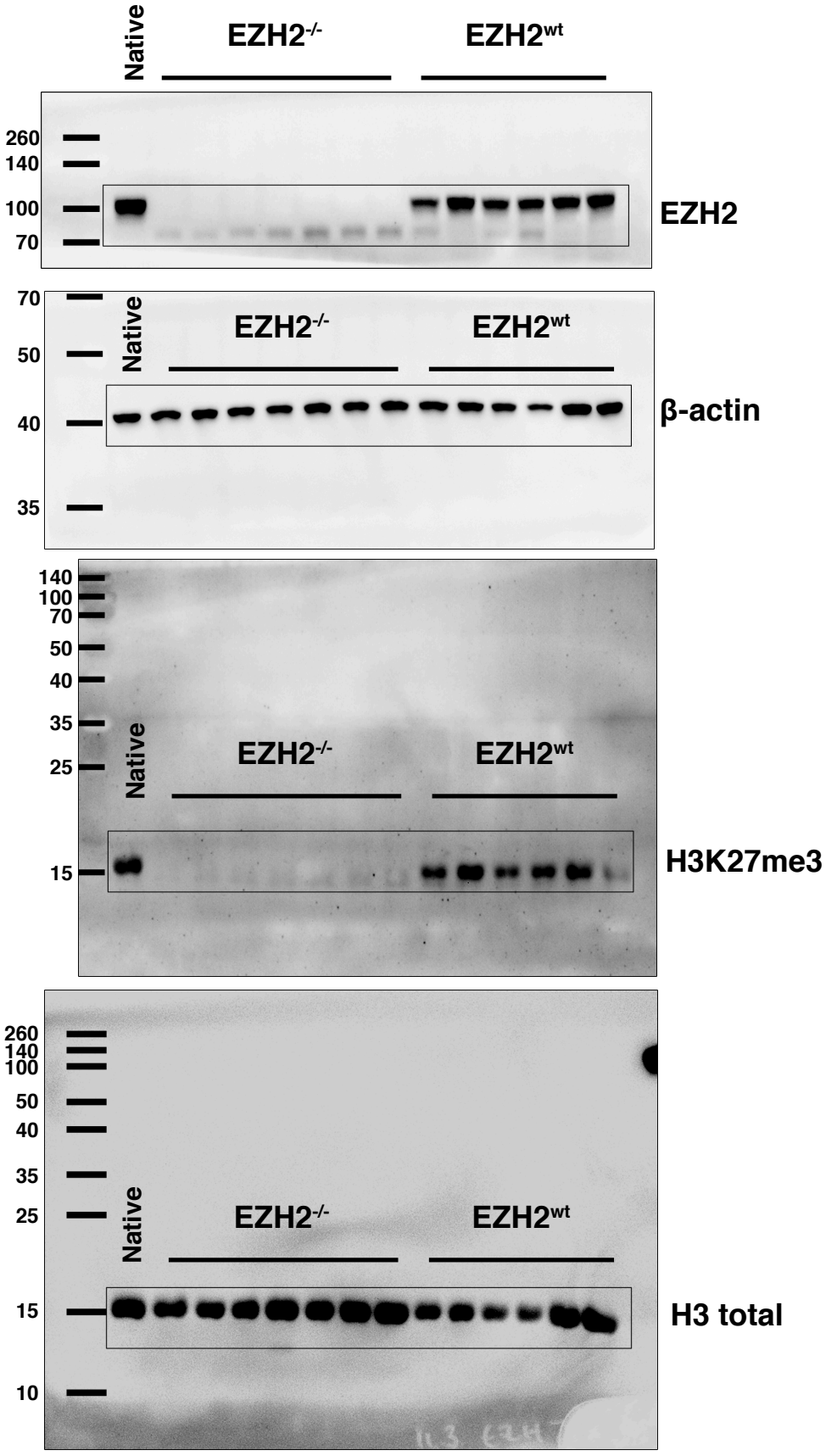
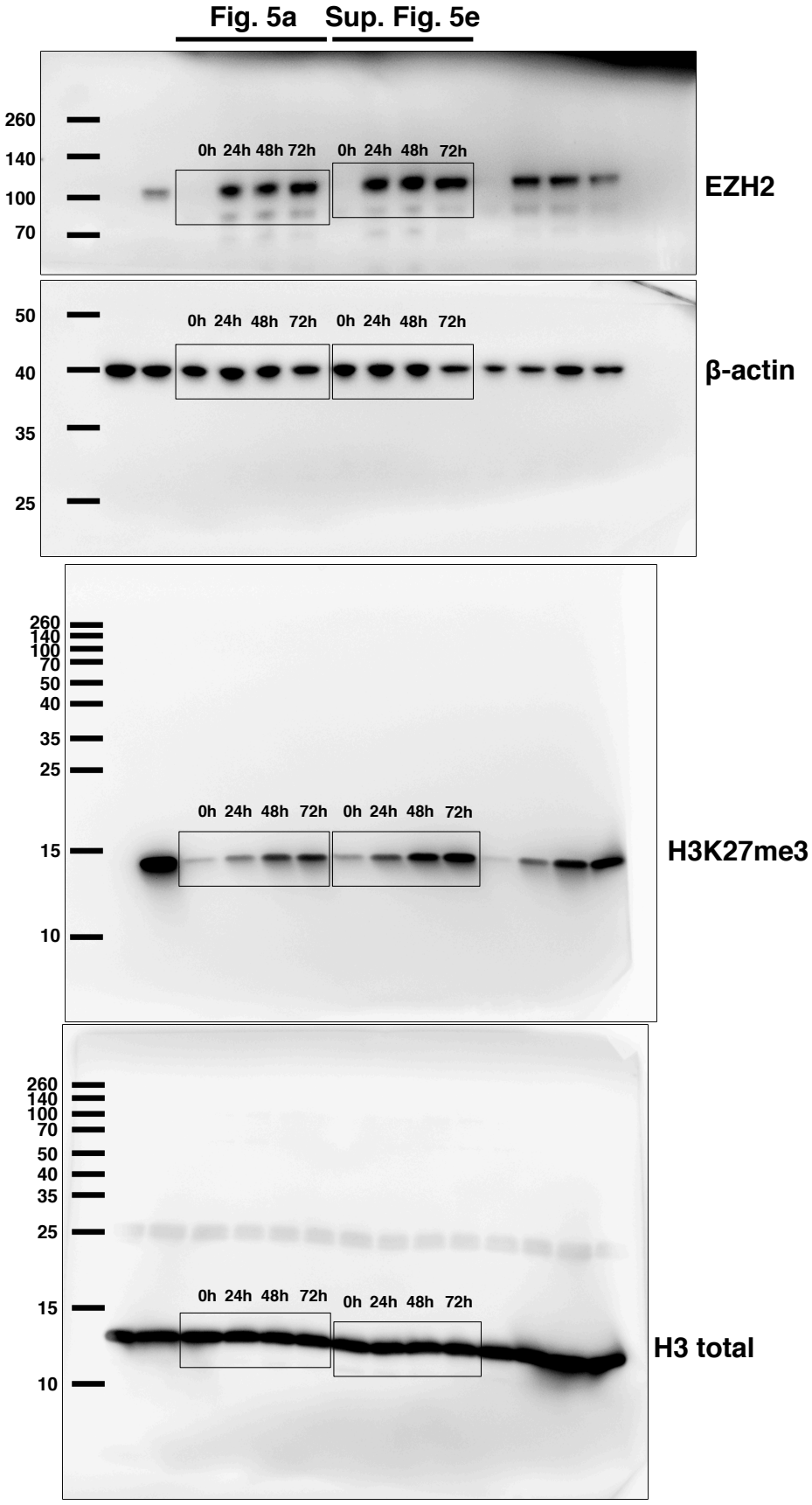

EZH2wt
EZH2-/-
EZH2-/-
260
140
100
70
EZH2
EZH2wt
EZH2-/-
EZH2-/-
50
40
35
25
15
EED
260
140
100
70
SUZ12
50
40
35
25
15
β-actin
70
50
40
35
25
15
RbAp46

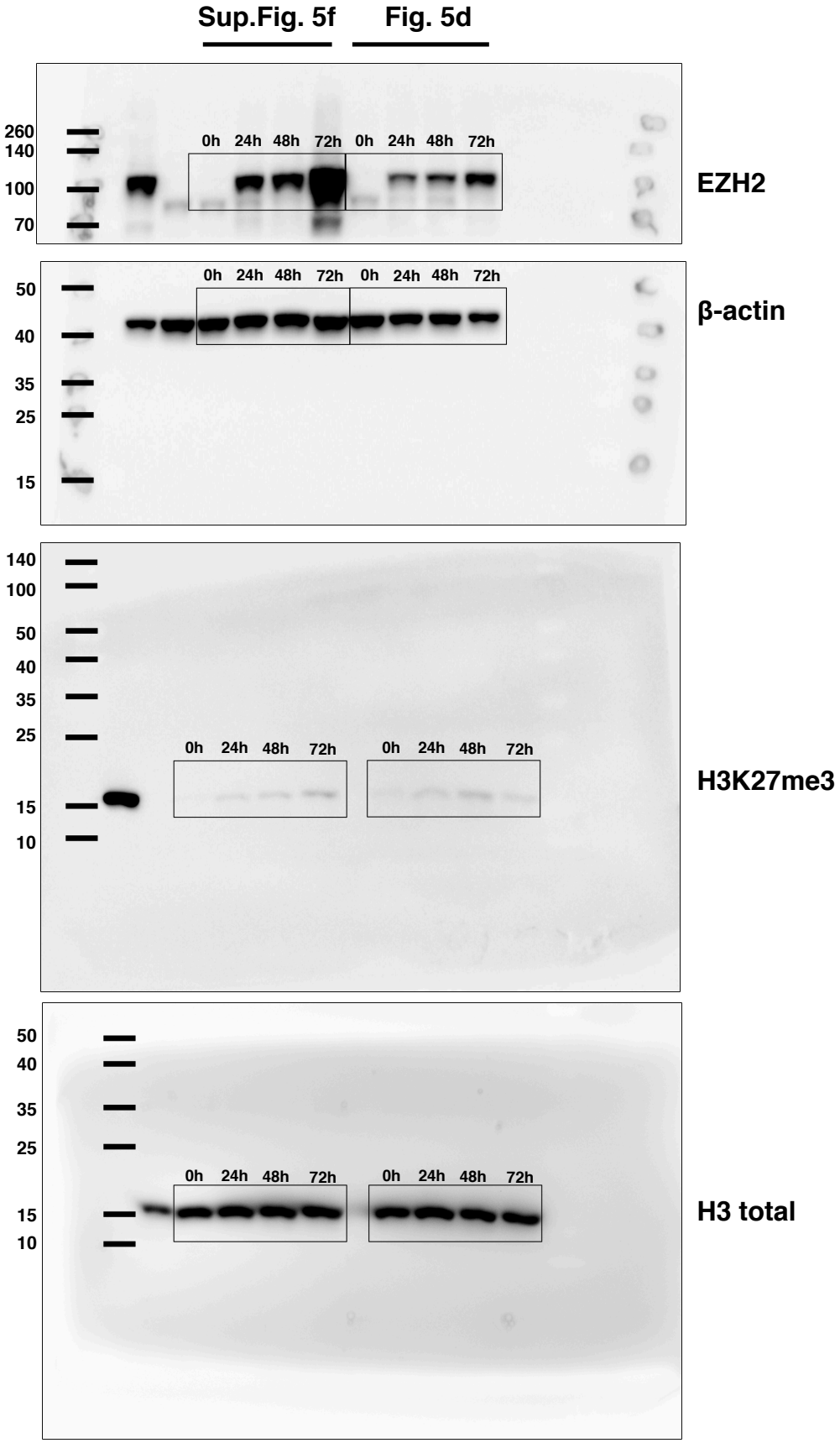
**Figure 4b**


Native
EZH2-/-
EZH2wt
260
140
100
70
EZH2
70
50
40
35
β-actin
140
100
70
50
40
35
25
15
H3K27me3
260
140
100
50
40
35
25
15
10
H3 total

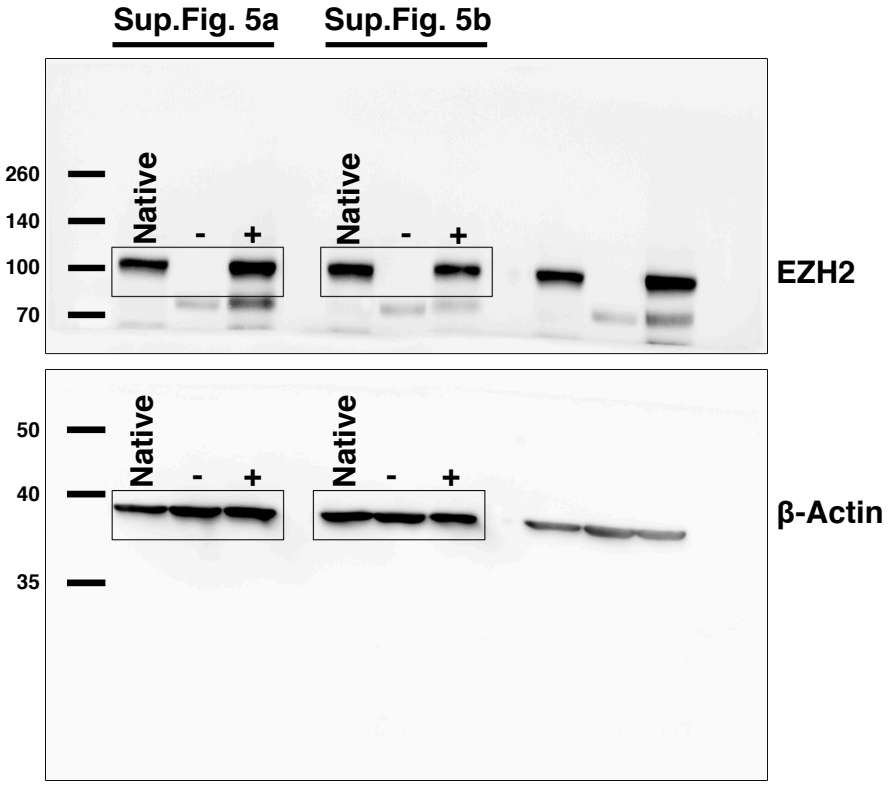
## Figure 5a & Supplementary Figure 5e



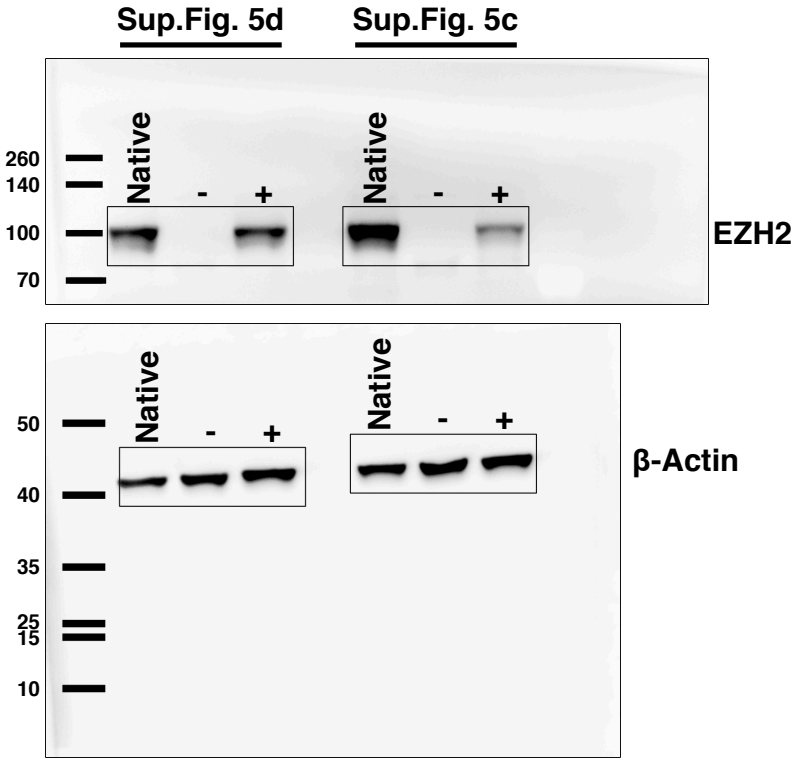
## Figure 5d & Supplementary Figure 5f



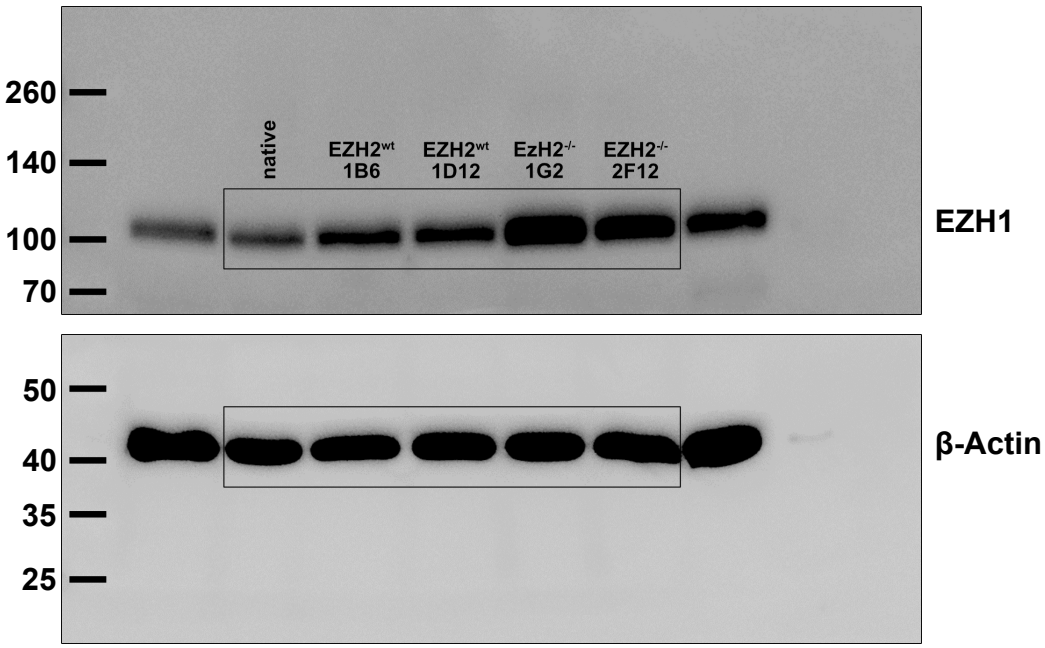
## Supplementary Figure 5a & Supplementary Figure 5b



## Supplementary Figure 5c & Supplementary Figure 5d



## Supplementary Figure 6a



## Supplementary Figure 6b

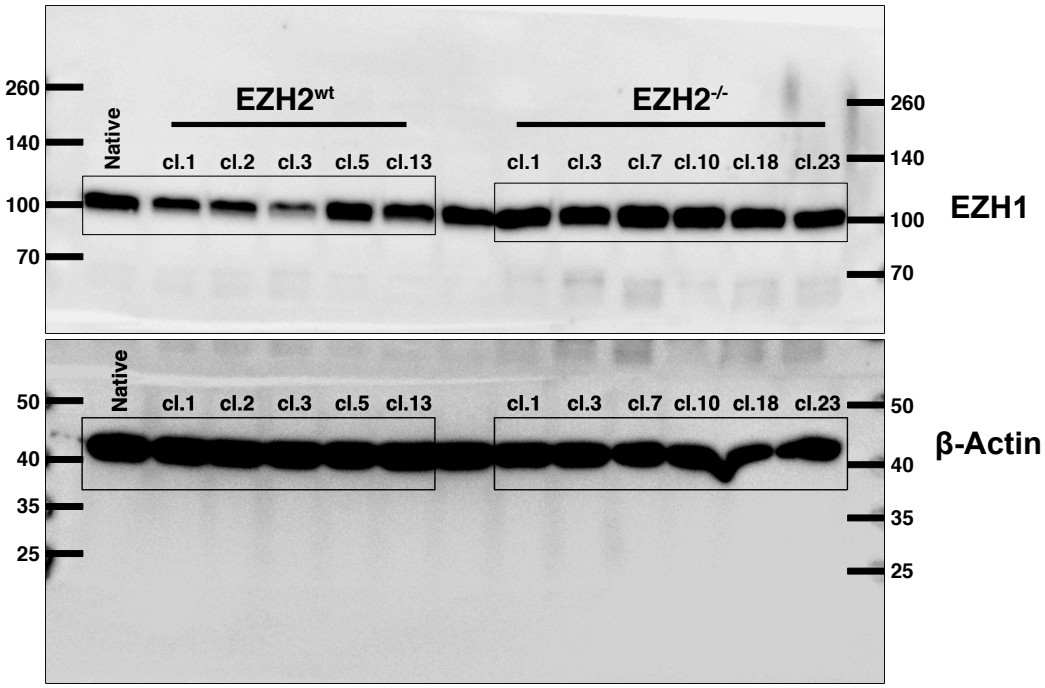
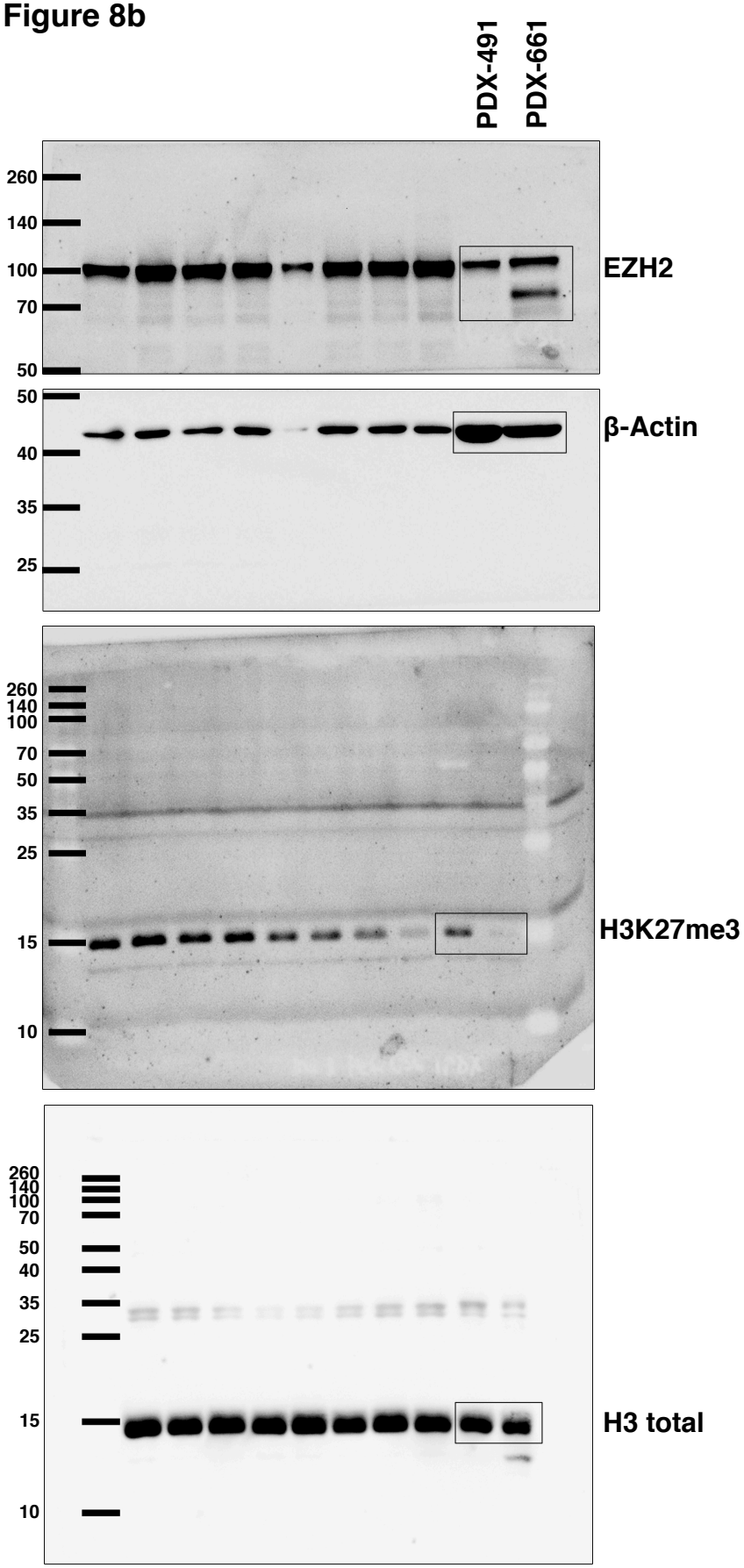
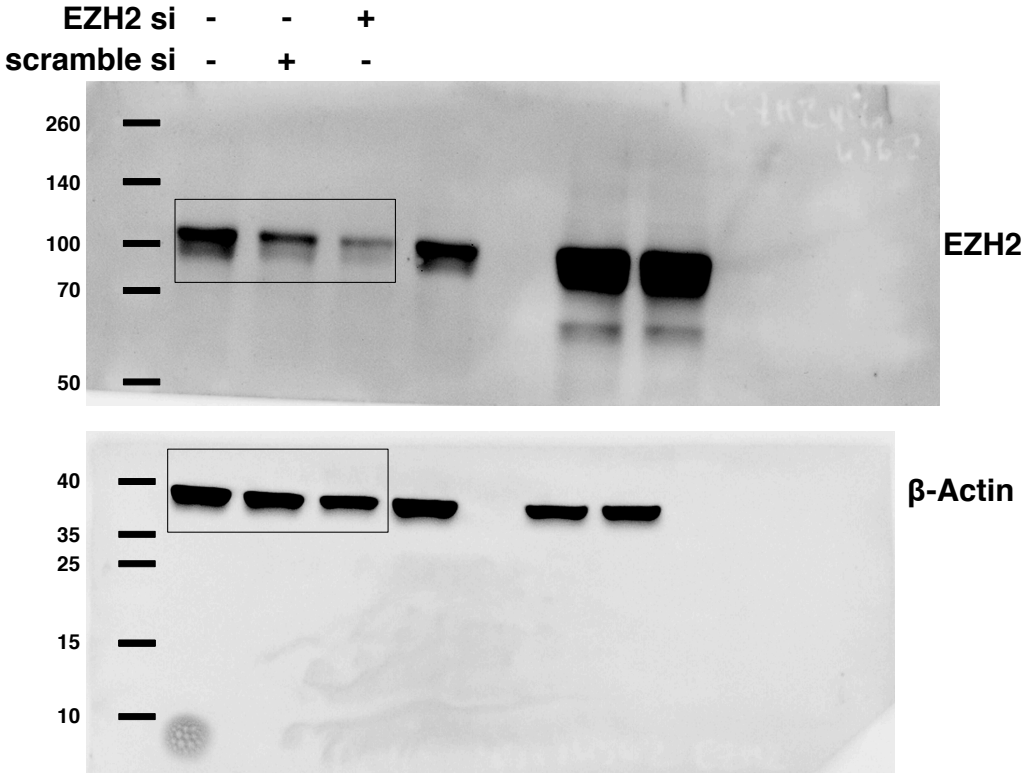


**Figure 8b**


PDX-491
PDX-661
260
140
100
70
50
EZH2
50
40
35
25
β-Actin
260
140
100
70
50
35
25
15
10
H3K27me3
260
140
100
70
50
40
35
25
15
10
H3 total

## Figure 8d



# Disease progression in cytogenetically normal AML is driven by loss of function mutations in DNMT3A

Manuscript in preparation

# Disease progression in cytogenetically normal AML is driven by loss of function mutations in DNMT3A

Sabrina Weser[1], Swati Parekh[4], Luise Hartmann[1,2,3,] Karsten Spiekermann[1,2,3*]

**1 Department of Medicine III, University Hospital, LMU Munich, Munich, Germany**

**2 German Cancer Consortium (DKTK), Heidelberg, Germany**

**3 German Cancer Research Center (DKFZ), Heidelberg, Germany**

**4 Max Planck Research Group 'Chromatin and Aging', Max Planck Institute for Biology of Aging, Cologne, Germany**

* Karsten.Spiekermann@med.uni-muenchen.de

## Abstract

*DNMT3A* mutations are highly prevalent in AML and have been associated with poor prognosis. Despite recent treatment advances, AML relapse rate remains high and is the major reason for treatment failure. Here, we analyzed RNA-seq and DNA methylation data of matched diagnosis and relapse samples from patients harboring DNMT3A mutations. Co-expression analysis revealed up-regulation of immune system and metabolic processes in relapse, as well as regulation of pathways inducing increased DNA replication, DNA repair and autophagy inhibition. Our differential methylation analysis revealed elevated levels of hypermethylated CpGs upon relapse. Such hypermethylated CpGs were found enriched in regions related to myeloid differentiation. Moreover, leukemia stem cell (LSC) signatures were found to be enriched in differentially expressed genes between diagnosis and relapse. Finally, we observed that *DNMT3B* whose expression is associated with hypermethylation is also overexpressed upon relapse. Collectively, our results indicate aberrant DNA methylation to be involved in regulation of leukemia cell differentiation, resulting in increased stemness at relapse and identified a putative role of DNMT3B in this process.

## Introduction

Cytogenetically normal acute myeloid leukemia (CN-AML) accounts for approximately 50% of all AML cases and is characterized by a heterogeneous mutation spectrum (Deneberg et al. 2011). Many patients achieve complete remission after initial chemotherapy, however most of them will eventually suffer from relapse, constituting a major obstacle for treatment. Persisting leukemia stem cells (LSCs) with repopulating potential are thought to be the major reason for treatment failure (Ishikawa et al. 2007; Shlush et al. 2017; Fong et al. 2015). In AML without genetic aberrations, epigenetic changes play a major role in relapse progression and chemotherapy resistance and mutations in epigenetic modifiers like *TET2*, *EZH2*, *KDM6A*, *IDH1/2* and *DNMT3A* have been reported repeatedly (Stief et al. 2019; Kempf et al. 2021; Figueroa et al. 2010; Roboz et al. 2020; Tulstrup et al. 2021).

Mutations in the DNA methyltransferase 3 Alpha DNMT3A are highly prevalent in AML, approximately 37% of patients with a cytogenetically normal karyotype (CN) are affected (Ley et al. 2010). *DNMT3A* mutations have been classified as an unfavorable prognostic marker, associated with reduced time to relapse and shorter overall survival (Ribeiro et al. 2012). Occurring early in pre-leukemic cells (Shlush et al. 2014), *DNMT3A* mutations are extensively stable during disease progression and have been identified as persistent during chemotherapy (Corces-Zimmerman et al. 2014). About 60% of *DNMT3A* mutations are found at the residue R882, located in the catalytic region (Ley et al. 2010). Leading to loss of DNMT3A function, R882- and several other variants have been linked to a global reduction of DNA methylation (Russler-Germain et al. 2014). Besides DNMT3A, the DNMT family comprises three more members, the enzymatically active DNMT1 and DNMT3B and the regulatory factor DNMT3L (Challen et al. 2014). While DNMT1 acts almost exclusively as a maintenance DNA methyltransferase, methylating hemimethylated CpGs (Choi et al. 2011), DNMT3B is a de novo methyltransferase and has been identified as an important

regulator of AML progression (Wong et al. 2019). In contrast to DNMT3A, mutations in DNMT3B are rare events in AML, however overexpression of DNMT3B has been associated to inferior disease-free and overall survival as well as higher relapse rates (Hayette et al. 2012; Niederwieser et al. 2015; Lamba et al. 2018).

In our previous study we analyzed 50 CN-AML patients and found frequent occurrences of genetic alterations in epigenetic modifiers (Greif et al. 2018). Several patients even gained mutations in epigenetic modifiers at relapse, suggesting an involvement of epigenetic regulation in disease progression. The present work aims to characterize the influence of DNMT3A mutations on disease progression from the initial diagnosis to relapse after chemotherapy. We analyzed gene expression as well as DNA methylation in paired diagnosis and relapse samples of CN-AML patients, characterizing changes between relapse and diagnosis as well as differences between DNMT3A mutated and WT patients. We found DNA hypermethylation at relapse to be involved in increased stemness and identified a possible role of DNMT3B in disease progression.

## Results

### DNMT3A mutations persist at relapse

In our cohort of 50 CN-AML patients (Greif et al. 2018) *DNMT3A* mutations were detected in 27 patients, with one patient harboring two *DNMT3A* mutations. The most frequently detected *DNMT3A* mutation was an arginine to histidine substitution located at residue R882 (12 patients), in the catalytic region at the C-terminus. R882C, and G543C were found in three and two patients respectively, all other *DNMT3A* mutations occurred in a single patient only. Most mutations (10) were located in the methyltransferase (MTase) region, three were located in the AAD regulatory domain and one right outside the PWWP regulatory domain (Figure 1). All detected *DNMT3A* mutations were classified as stable, meaning they were present in both diagnosis and relapse samples and therefore neither

gained nor lost at relapse. Additionally, pre-leukemic *DNMT3A* mutations could be identified in 19 patients, by their presence in the remission samples (Greif et al. 2018). *DNMT3A* has frequently been found co-mutated with *NPM1*, *FLT3* or *IDH1* (Renneville et al. 2012; Roller et al. 2013; Cancer Genome Atlas Research Network ...), while in this cohort co-mutation with *DNMT3A* could not be detected (Fisher's Exact Test, $p < 0.05$). However, *ASXL1*, *RUNX1* and *SRF2* were found mutated more often in the group of patients without *DNMT3A* mutation (Fisher's Exact Test, $p < 0.05$) (Figure S1).


**Figure1: Location of identified DNMT3A mutations.** Schematic overview of DNMT3A protein structure and identified mutations in 50 CN-AML patients. Canonical isoform (Q9Y6K1-1) shown. **PWWP**: Pro-Trp-Trp-Pro motif domain, **AAD**: ATRX, DNMT3, DNMT3L zinc finger domain, **MTase**: Methyltransferase, **green**: regulatory domain, **blue**: catalytic domain

## DNMT3A mutations cause overexpression of poor prognosis marker genes

*DNMT3A* mutations like R882H, R882C, R882P, R729W, and S714C have been described as causing loss of methyltransferase activity (Holz-Schietinger et al. 2012; Bera et al. 2020; Emperle et al. 2019; Sandoval et al. 2019). In accordance with those findings, we observed lower global DNA methylation level in patients with *DNMT3A* mutation (DNMT3A-Mut) than in patients without *DNMT3A* mutation (DNMT3A-WT), indicating a reduction of DNMT3A function (Figure 2a, Figure S2). To investigate the transcriptional changes induced by *DNMT3A* mutations, differential expression as well as methylation analysis was performed between DNMT3A-Mut and DNMT3A-WT patients at diagnosis and relapse.

Differential expression analysis resulted in 694 up-regulated and 707 down-regulated genes in the diagnosis samples ($p < 0.05$), while no change could be detected for the relapse samples ($p < 0.1$) (Figure S2b, c, Table S1, Table S2). Notably, *DNMT3A* gene expression was not affected by *DNMT3A* mutations. The group of up-regulated genes was found enriched for genes previously identified as up-regulated in transplanted *Dnmt3a*-null versus control hematopoietic stem cells (HSCs) (Challen et al. 2011), confirming loss of DNMT3A function to be the reason for transcriptional up-regulation (hypergeometric test, $p = 0.032$). Overrepresentation analysis of the up-regulated genes identified TNF superfamily cytokine production as a positively regulated process in DNMT3A-Mut patients (Figure 2b). Up-regulated genes in this pathway included the toll-like receptors 2 and 4 (TLR2 and TLR4), previously associated with poor overall survival and chemoresistance (Rybka et al. 2015). Additionally, up-regulation of the antiapoptotic genes *BCL6* and *BCL2A1* as well as the membrane receptor *CLEC7A* was observed, all associated to chemotherapy resistance in AML (Kawabata et al. 2021; Zhang et al. 2020; Cucchi et al. 2020). Since *DNMT3A* mutations have been found to ablate hematopoietic stem cell (HSC) and leukemia stem cell (LSC) differentiation (Koya et al. 2016), the differentially expressed genes were tested for enrichment of HSC and LSC expression patterns (Eppert et al. 2011; Gentles et al. 2010; Chambers et al. 2007). No enrichment could be detected for the mentioned data sets, indicating a similar differentiation state in the groups of DNMT3A-Mut and DNMT3A-WT patients. Additionally, differential methylation analysis between the DNMT3A-Mut and DNMT3A-WT patients was performed for both diagnosis and relapse samples. In the diagnosis differential methylation was detected in 4,753 genes with 11,060 differentially methylated CpGs, most of them hypomethylated (10,141 hypo, 919 hyper, $adj.p < 0.05$, $|db| > 0.1$) (Table S3). Genes of the relapse samples were mainly hypomethylated as well but only comprised 313 genes, including 564 differentially methylated CpGs (536 hypo, 28 hyper, $adj.p < 0.05$, $|db| > 0.1$) (Table S4).

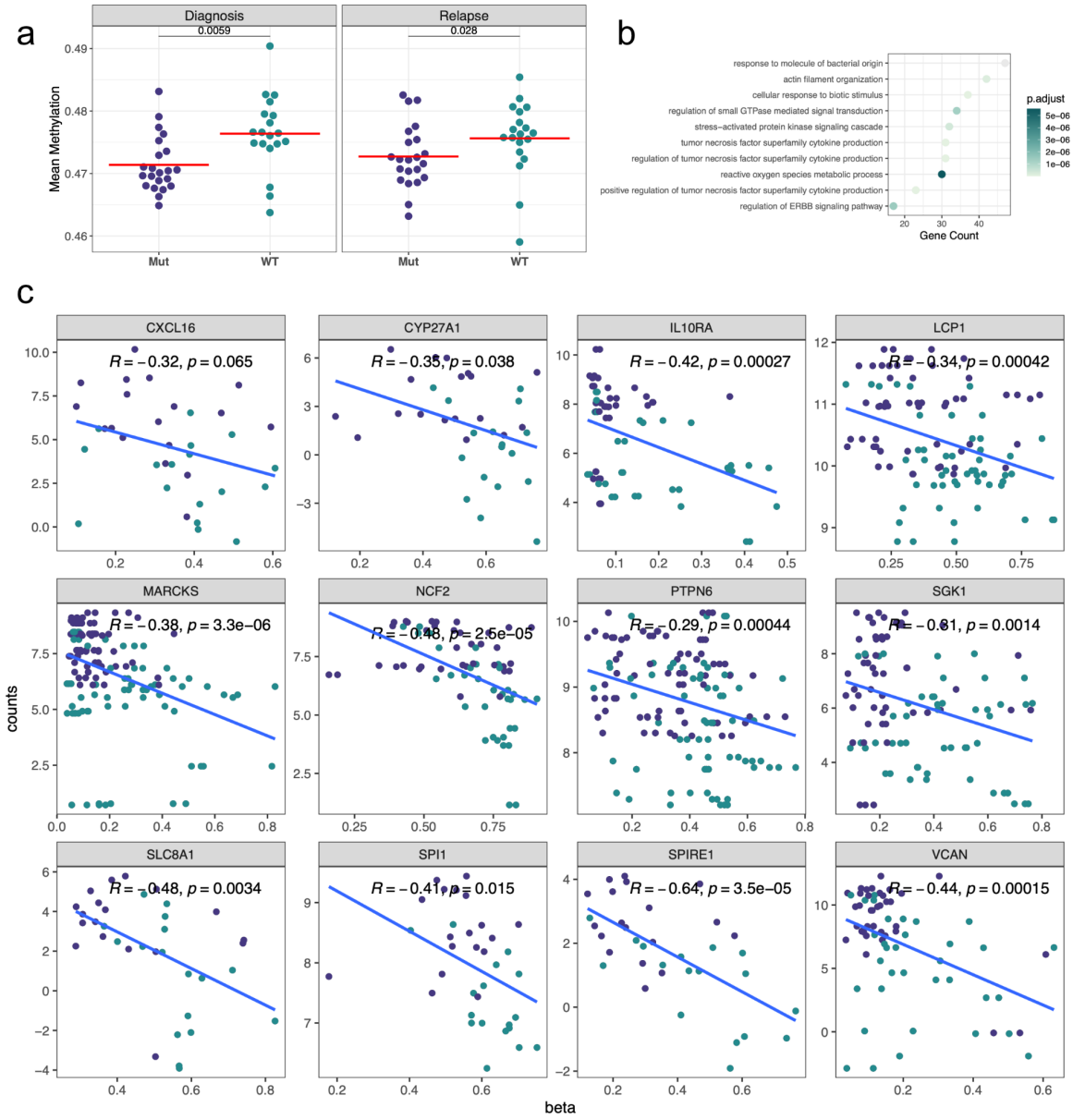



**Figure2: Hypomethylation in DNMT3A-Mut patients. a,** Mean methylation of all CpGs analyzed in DNMT3A-Mut and DNMT3A-WT patients for relapse and diagnosis samples, wilcoxon test. **b,** Overrepresentation analysis of down-regulated genes in DNMT3A-Mut patients, top 10 categories shown, $p < 0.05$. **c,** Genes correlated between gene expression and Methylation in diagnosis samples of DNMT3A-Mut versus DNMT3A-WT patients, Spearman correlation.

To determine the transcriptional changes induced directly by aberrant DNA methylation, 369 genes with both differential expression and methylation in diagnosis samples were identified. As expected, most of those genes (93%) were found to be hypomethylated. Hypomethylated CpGs were mainly found in the gene body and TSS regions outside of CpG-Islands. Those regions were also found to be enriched over the background of all CpGs in the analysis. Correlation coefficients were calculated for the 369 differentially methylated and expressed genes and revealed methylation level of 134 genes correlated to gene expression (spearman $p < 0.05$) (Figure 2c, Table S5). Those included genes previously described as poor prognostic markers in AML like *LCP1*, *PTPN6/SHP1*, *MARCKS*, *CXCL16*, *VCAN*, *IL10RA*, *SLC8A1* and *CYP27A1* *(Shin et al. 2020; Reich et al. 2020; Lebedev et al. 2019; Yan et al. 2019; Xu et al. 2021; Huang et al. 2019)*. Furthermore, up-regulated genes included the autophagy inhibitor *SGK1*, *SPIRE1* involved in DNA repair, the hematopoietic master regulator *SPI1/PU.1* and *NCF2*, which is linked to chemoresistance (Plessner and Grosse 2019; Liu et al. 2017; Pham et al. 2013; Rosa et al. 2021).

### Gene co-expression networks depict relapse progression

To obtain more insight on the influence of DNMT3A on disease progression, matched diagnosis and relapse samples of *DNMT3A* mutated patients were analyzed. Differential expression analysis resulted in 548 up-regulated and 816 down-regulated genes, with a tendency to higher fold changes in the down-regulated ones (Table S6). Enriched biological processes were detected by overrepresentation analysis and included GOs involved in cancer cell differentiation such as "mononuclear cell differentiation", "cilium assembly" and "toll-like receptor signaling pathway" in the top 10 categories (Figure S3a).

To deepen this insight in the molecular mechanism of disease progression, weighted gene co-expression network analysis (WGCNA) was performed to identify clusters of highly

correlated genes. Seven co-expression modules were identified that could be associated with relapse samples. The modules skyblue, midnightblue, lightgreen and lightcyan passed the correlation p-value cutoff of 0.05 (Figure S3c). Gene Ontology Analysis was performed for those modules to reveal the biological associations of the co-expressed genes. Genes in the skyblue module were found to be over-represented in GOs involved in DNA replication and repair. Those genes were positively correlated to relapse and included several differentially up-regulated genes (Figure 3a, Figure S3e) such as the oncogenes *BRCA1*, *TRIP13* and *DSCC1* as well as *POLA1*, which encodes for the catalytic subunit of DNA polymerase alpha. The lightgreen module was found to be positively correlated to relapse as well and comprised genes commonly involved in the up-regulation of metabolic processes (Figure 3b, Figure S3f). Differentially up-regulated genes in this module included *LARS1*, recently described as a key player in senescence escape as well as the two other Aminoacyl-tRNA synthetases *FARS2* and *EARS2* (Guillon et al. 2021). Furthermore, we observed up-regulation of genes involved in amino acid and nucleotide synthesis, such as *PYCR3* and *HIBCH* and *GART* and *CAD*, which have been previously linked to chemotherapy resistance (Bogner et al. 2021; Shan et al. 2019; Li et al. 2015; Givechian et al. 2018). Immune system activation was observed in the midnightblue module. Genes in this module were negatively correlated to relapse and uniformly associated to activation of the immune system (Figure 3c, Figure S3g). Examples for differentially expressed genes in this module included the leukocyte immunoglobulin-like receptors *LILRB4* and *LILRB1*. Those inhibitory receptors are expressed on several immune cell types, including macrophages and monocytes, playing a key regulator role in immune response (Liu et al. 2020). The lightcyan module included the GOs vesicle and endosome organization (Figure 3d, Figure S3h). Genes in this module were uniformly down-regulated and included the differentially expressed genes *PI4K2A*, *VPS18* and *RAB1A*, involved in autophagosome formation as well as autophagosome-lysosome fusion (Albanesi et al. 2015; Peng et al.

2012; Webster et al. 2016). Down-regulation of those genes facilitates cancer proliferation through autophagy inhibition (Watson et al. 2015).

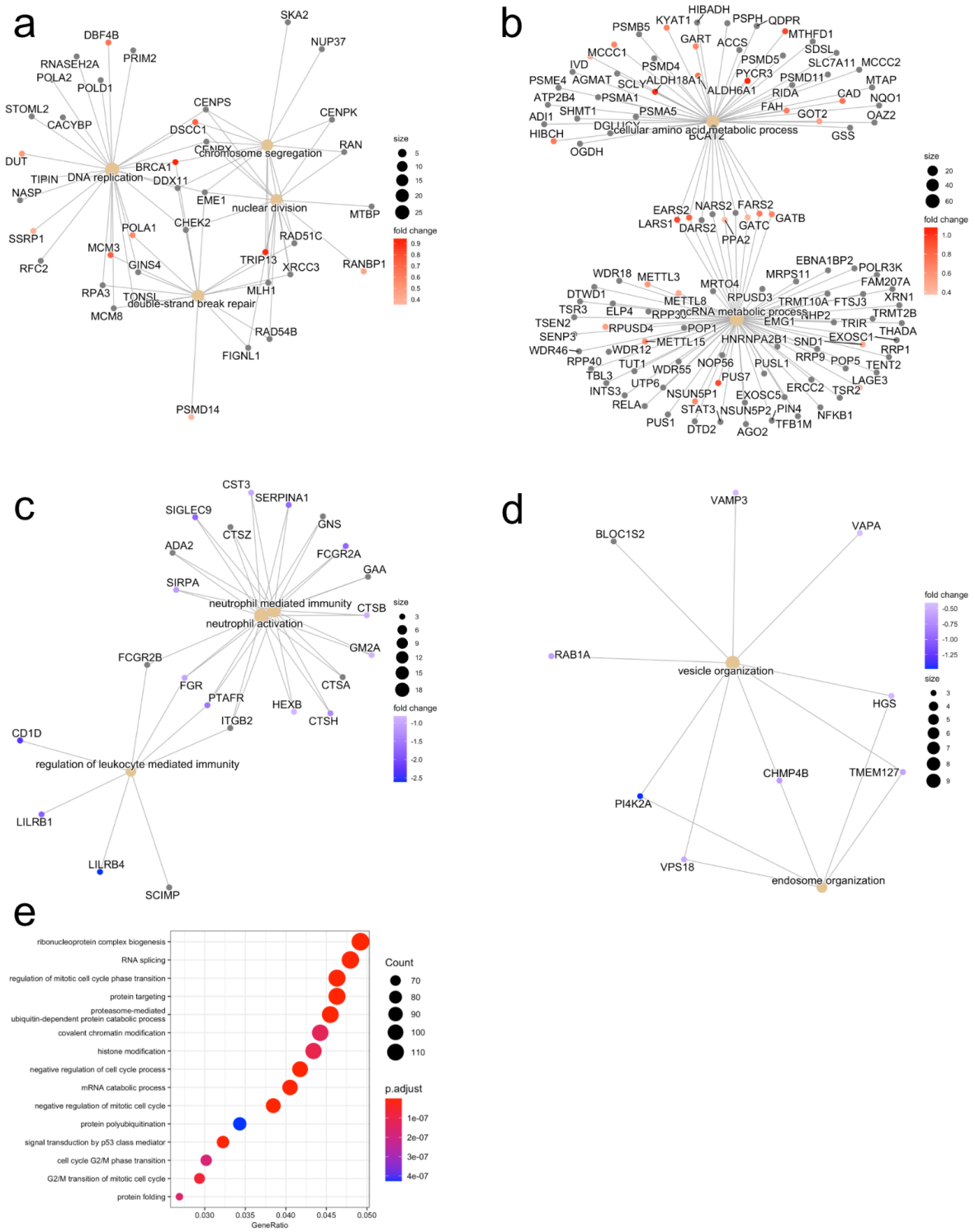


**Figure 3: Regulation of biological processes in relapse. a,** Cnet of selected GOs depicting DNA replication and repair in the skyblue module, differentially expressed genes are colored by their fold change level. **b,** Cnet of selected GOs depicting metabolic processes in the lightgreen module, differentially expressed genes colored. **c,** Cnet of selected GOs depicting Immune System up-regulation in the midnightblue module, differentially expressed genes colored. **d,** Cnet of selected GOs depicting vesicle and endosome organization in the lightcyan module, differentially expressed genes colored. **e,** GO analysis of co-methylation module pink.

Additionally, the WGCNA method was employed to identify relapse-related co-methylation modules of individual CpGs. With a correlation p-value cutoff of 0.05 the two modules tan and pink were found to be negatively correlated to relapse (Figure S3d). Gene ontology analysis of the pink module revealed over-representation of co-methylation associated genes involved in cell cycle regulation as well as epigenetic and post-translational modifications (Figure 3e). However, none of those CpGs were found significant in the differential methylation analysis. Methylation of this module was restricted to CpG islands and shores (55% and 45% respectively), suggesting a role in transcriptional regulation of associated target genes. In the tan module GO analysis did not indicate any over-representation.

## Hypermethylation is involved in increased stemness of relapse

We found global methylation to be elevated at relapse in comparison to diagnosis of DNMT3A-Mut patients (Figure 4a). Mean methylation level of CpG islands, shores and shelves depicted a significant increase in relapse samples in comparison to diagnosis (Figure S4). As expected, differential methylation analysis in those patients revealed predominantly hypermethylated CpGs (1,158 hyper, 91 hypo) (Table S6). Overrepresentation analysis was performed to assess biological processes affected by the hypermethylated CpGs. In accordance with the expression data, we observed the terms associated with myeloid differentiation in the top 10 categories (Figure 4c). We therefore asked if relapsed cells are in general in a less differentiated state and indeed found

significant enrichment for gene expression signatures of leukemia stem cells (LSCs) (Eppert et al. 2011). Of the 42 down-regulated genes (FDR < 0.1) in LSC-enriched versus LSC-depleted AML cell populations comprised in this LSC signature (Eppert et al. 2011), 19 were also down-regulated at relapse (p = 5.4e-14). In contrast, only 7 of the genes up-regulated in relapse were found in the signature of 163 genes up-regulated in LSCs and no enrichment could be detected. To determine the influence of the observed methylation changes on transcriptional activity, correlation between DNA methylation and gene expression was examined. In total, 79 differentially expressed and methylated genes were identified, of which 31 were significantly correlated (spearman, adj.P < 0.05, (Table S10). Most genes harbored DMCs located in either Promoter (5 genes) or Enhancer (19 genes) regions. Transcriptional down-regulation was associated with hypermethylation only (21 genes), while up-regulation was correlated to either hypo or hypermethylation (5 hypo, 5 hyper). Assuming an inhibitory effect of hypermethylated CpGs in TSS or enhancer regions, genes down-regulated in response to methylation were identified. Those included regulators of myeloid cell differentiation *CCR1*, *LILRB4* and *PLEKHO2* (Figure 4b).

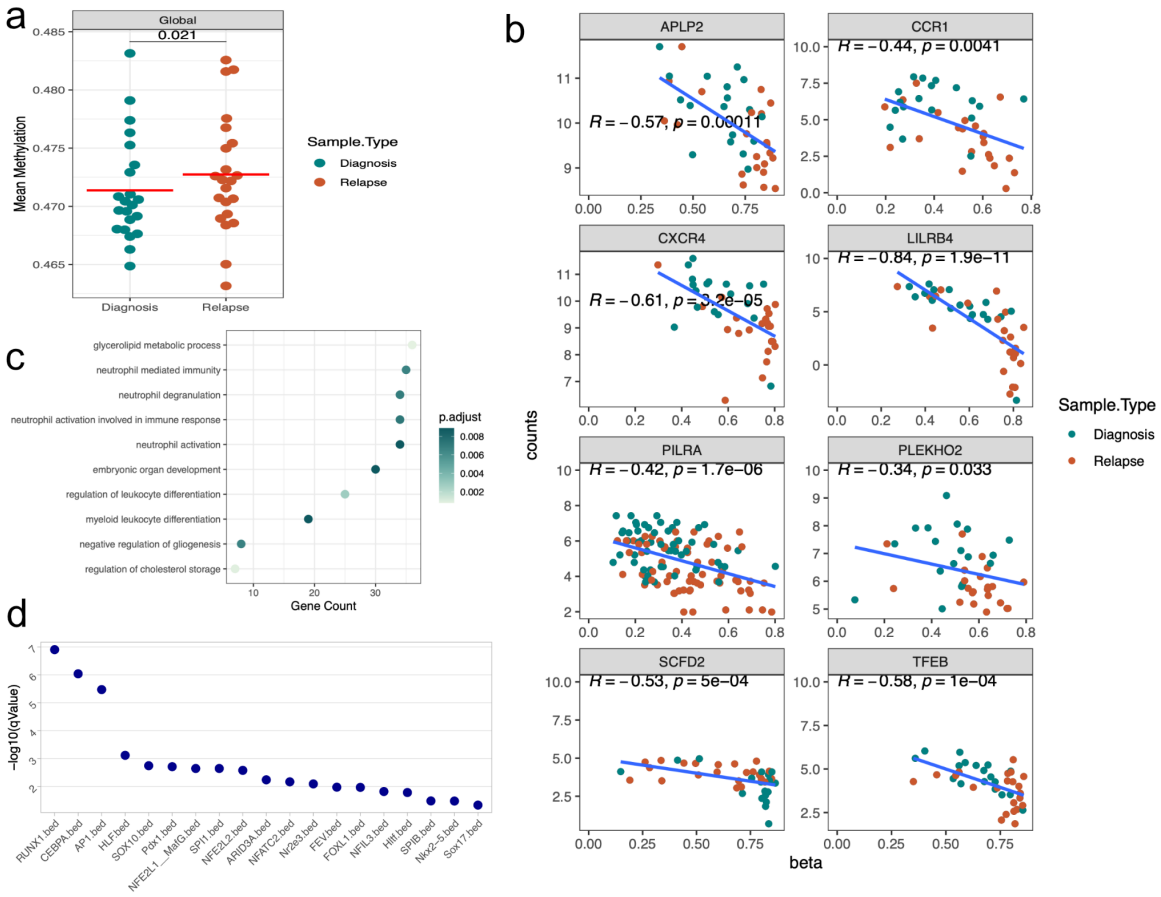



**Figure 4: Hypermethylation in relapse. a,** Mean methylation of all CpGs analyzed in relapse and diagnosis samples of DNMT3A-Mut patients. wilcox test. **b,** Genes correlated between gene expression and Methylation in DNT3A-Mut patients, spearman correlation. **c,** Overrepresentation analysis of genes harboring hypermethylated CpGs, top 10 categories. **d,** Locus overlap analysis (LOLA), Transcription factor motif sites from Jaspar database for hypermethylated sites, qValue < 0.05.

## DNMT3B overexpression contributes to relapse progression

*DNMT3B* overexpression has been reported repeatedly as a marker of poor disease progression, higher relapse rate and chemoresistance in AML (Niederwieser et al. 2015; Monteferrario et al. 2014; Lamba et al. 2018; Hayette et al. 2012). In accordance with those findings we detected *DNMT3B*, but none of the other *DNMT*s, to be up-regulated in the relapse of DNMT3A-Mut patients (Table S6). None of the CpGs located in the *DNMT3B* gene or in any known *DNMT3B* enhancer regions (GeneHancer Version 4-4) was

differentially methylated, indicating a methylation-independent gene expression change. To investigate the involvement of DNMT3B on the observed increase of DNA methylation in relapsed patients, correlation of *DNMT3B* as well as *DNMT3A* gene expression with global methylation levels was analyzed. In the context of normal DNMT3A function (DNMT3A-WT patients) a strong positive correlation with global methylation was observed for both *DNMT3B* as well as *DNMT3A* (Figure 5a). Aberrant DNMT3A function (DNMT3A-Mut patients) however was not able to maintain this relation. Neither *DNMT3A* nor *DNMT3B* was found correlated to global methylation. Since DNMT3B is known to possess lower methylation activity than DNMT3A (Takeshima et al. 2006), its influence on global methylation could be clouded by aberrant DNMT3A function. Therefore, the analysis was repeated for mean methylation levels of all hypermethylated CpGs. Here, a significant positive correlation with *DNMT3B* expression, but not *DNMT3A* was detected, suggesting an involvement of DNMT3B on the observed hypermethylation phenotype (Figure 5b). Overexpression of *DNMT3B* has been frequently observed in AML patients and an associated gene expression profile has been published (Niederwieser et al. 2015). Indeed, we found an enrichment for this gene set among the differentially expressed genes between relapse and diagnosis of DNMT3A-Mut patients. Of the 168 down-regulated genes in *DNMT3B* overexpressors (Niederwieser et al. 2015), 125 were also down-regulated at relapse. Correspondingly, 60 of the 195 up-regulated genes in DNMT3B overexpressors were found up-regulated in relapse. Notably, none of the genes were regulated in the opposite direction.

Intriguingly, eight of the down-regulated genes correlated to hypermethylation *CCR1*, *CRISPLD2*, *FAM49A*, *LILRB4*, *TFEB*, *TMEM104*, *SLC11A1* and *PLEKHO2*) were moreover linked to DNMT3B overexpression. Furthermore, overlapping genes included *CD14*, *TLR4*, *CEBPB* and *TLR8*, involved in macrophage differentiation. This suggests an involvement of DNMT3B in cell differentiation.

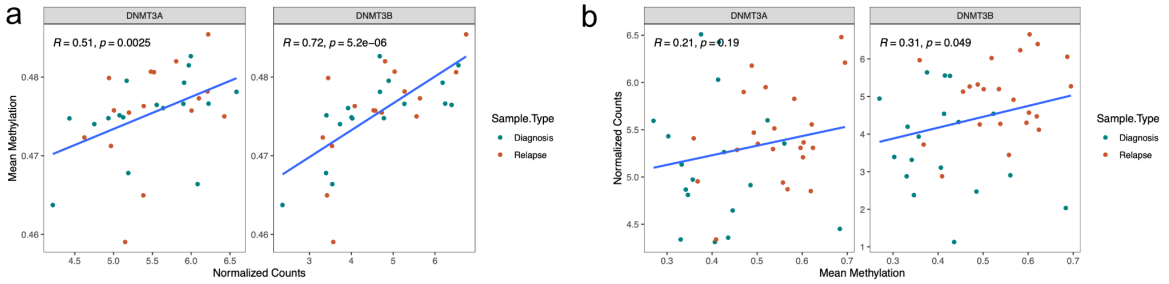



**Figure 5: Influence of DNMT3B on methylation levels. a,b,** Correlation of DNMT3B and DNMT3A expression with mean methylation. **a,** Mean methylation of all CpGs analyzed. **b,** Mean methylation of hypermethylated CPGs between relapse and diagnosis in DNMT3A-Mut patients. **R**: spearman's rank correlation coefficient.

## Discussion

In AML, *DNMT3A* ranks amongst the most frequently mutated genes and is associated with poor prognosis. Although *DNMT3A* mutations do not seem to impair complete remission after initial chemotherapy, relapse rate is very high. Since we found *DNMT3A* mutations to be stable during disease progression (Greif et al. 2018), the high relapse rate can likely be explained by *DNMT3A*-mutated LSCs, holding the potential to survive chemotherapy and repopulate.

In accordance with this, we found overexpression of multiple poor prognosis marker genes in DNMT3A-Mut patients. Negative correlation of methylation and expression of those genes suggested a regulatory role of methylation with transcriptional activation to be a likely result of reduced methyltransferase activity. Those genes included *LCP1*, *PTPN6/SHP1*, *MARCKS*, *CXCL16*, *VCAN*, *IL10RA*, SLC8A1 and CYP27A1, all associated with inferior overall survival in AML (Figure 2c) (Shin et al. 2020; Reich et al. 2020; Lebedev et al. 2019; Yan et al. 2019; Xu et al. 2021; Huang et al. 2019). Furthermore, we observed a significant enrichment for LSC expression signatures (Eppert et al. 2011) in the relapse of DNMT3A-Mut patients. This increased "stemness" was further confirmed by the downregulation of CCR1, *LILRB4* and *PLEKHO2*, involved in myeloid cell differentiation.

Additionally, hypermethylation of promoter or enhancer regions suggested involvement of methylation in silencing of those genes (Figure 4b).

DNA hypermethylation at CpG islands of tumor suppressor genes and its involvement in AML pathogenesis and aggressiveness has been described thoroughly and has led to the approval of several DNA hypomethylating agents (Wouters and Delwel 2016). In addition, indicating epigenetic involvement in AML disease progression, hypermethylation has been detected in relapsed AML patients (Kroeger et al. 2008). Here, we report a global increase of methylation in the relapse of DNMT3A-Mut patients (Figure 4a). As a consequence of *DNMT3B* upregulation, hypermethylation could possibly result from de novo methylation independent of DNMT3A function and assist the downregulation of genes involved in cell differentiation (Figure 4b). Indeed, Dnmt3b was found to generate hypermethylation in Dnmt3a-null HSCs (Challen et al. 2014). Furthermore, genes involved in differentiation were found to be downregulated in a group of AML patients expressing high DNMT3B levels, although no correlation to methylation could be detected (Niederwieser et al. 2015).

*SPI1*, encoding the hematopoietic master transcription factor PU.1 is of major importance in HSC self-renewal and differentiation (Pham et al. 2013). We found upregulation of *SPI1*, likely due to hypomethylation of its promoter, in DNMT3A-Mut patients (Figure 2c). Additionally, upregulation of *BCL2A1*, a known SPI1 target (Jenal et al. 2010) has recently been described to confer resistance against the BCL2 inhibitor venetoclax in AML (Zhang et al. 2020). DNMT3A loss of function mutations might thus be involved in the emergence of chemotherapy resistance. Controversially, *SPI1* was found downregulated in the relapse of DNMT3A-Mut patients. Additional hypermethylation at SPI1 binding sites in enhancer regions of downregulated genes involved in differentiation, suggests an involvement of SPI1 in the stemness at relapse (Figure 4b, Table S10). Taken together, these data reveal an involvement of DNA methylation in the increase of stemness in AML relapse and suggest a possible role of DNMT3B in disease progression.

## Methods

### Patients

Paired diagnosis and relapse samples of patients with CN-AML were collected at the Department of Medicine III, University Hospital, LMU in collaboration with the German Cancer Consortium (DKTK). Informed consent for scientific use of sample material was received from all study participants in accordance with the Declaration of Helsinki. Patients were treated intensively in agreement with the Acute Myeloid Leukemia Cooperative Group (AMLCG) and the Study Alliance Leukemia (SAL). The median age of patients at diagnosis was 66 years (range: 21-89 years). Patient characteristics are provided in our previous (Greif et al. 2018).

### RNA Sequencing

Libraries were prepared according to the TruSeq RNA Sample Preparation protocol and sequenced with 101 bp read length as paired-end layout on an Illumina HiSeq 2000 Instrument. On average ~30 million reads were obtained per sample. The reads were mapped to human reference genome hg19 by allowing only one alignment per read as 2-pass mapping using the RNA-seq aligner STAR (version 2.4.0.1) (Dobin et al. 2013). Genes were annotated with the Ensembl annotation version GRCh37.75 in R using Rsubread (Liao et al. 2019).

### DNA Methylation

Methylation analysis was performed by Infinium® HumanMethylation450 BeadChip Arrays. Raw idat files were processed using the package minfi (version 1.38.0). Poor performing probes (detP value < 0.01) were excluded prior to quantile normalization.

Probes on sex chromosomes and cross reactive probes were excluded. The remaining probes were used for further analysis. Beta and M-values were calculated.

### Differential Expression and Methylation Analysis

Differential Expression and Methylation analysis was performed using the R package limma. For RNA-seq data this was combined with voom. Tumor burden was included as a covariate in the model matrix and duplicate correlation was performed to assess the inter patient variability.

### Region Set Enrichment Analysis

Locus overlap analysis (LOLA) (Sheffield and Bock 2016) was used to identify enrichment of transcription factor binding sites from the Jaspar database amongst the sets of hyper and hypo methylated CpGs. The full set of CpGs considered in the differential methylation analysis was utilized as background (universe) and all enrichments with a qvalue < 0.05 were considered significant.

### Enrichment Analysis

GO enrichment analysis was performed using the function enrichGO and enrichment analysis for LSC, HSC, Dnmt3a-null and DNMT3B gene signatures was performed using the function enricher from clusterProfiler (version 4.0.5).

## Data Availability

The RNA-Seq and 450k array data generated for this study is available at GEO under the accession number.

## Author Contributions

K.S. and S.W. conceived the study. S.W. wrote the manuscript and analyzed the data. S.P. was involved in preprocessing of RNA-seq data and supported the analysis. L.H. designed the methylation experiment. K.S. interpreted the data and supervised the project. All authors reviewed the manuscript.

## Conflict of Interest

The authors do not declare any conflict of interest

## Acknowledgments

The authors thank all study participants. We thank I.Valtierra and J.Hilgareda for critical reading of the manuscript. We thank I.Hellmann, and B.Vieth for advice in computational analysis and statistics. K.S. was supported by the German Research Council (DFG) within the SFB 1243 “Cancer Evolution''. S.W. is a member of the IRTG-1243 within the SFB 1243 of the DFG.

# Supplementary Information

## Supplementary Figures

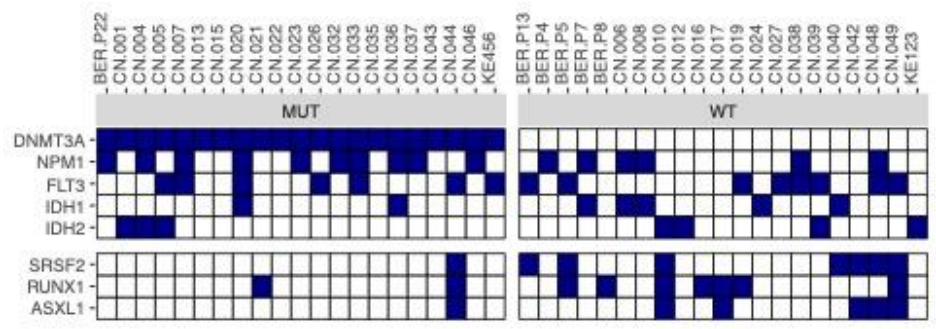



**Figure S1: Associated with Figure 1.** Schematic overview of selected mutations in DNMT3A-Mut or DNMT3A-WT patients. Upper panel: Common mutations co-muted with DNMT3A. Lower panel: Mutations occurring more often in DNMT3A-WT patients. Cohort from Greif et al. [(Greif et al. 2018)]

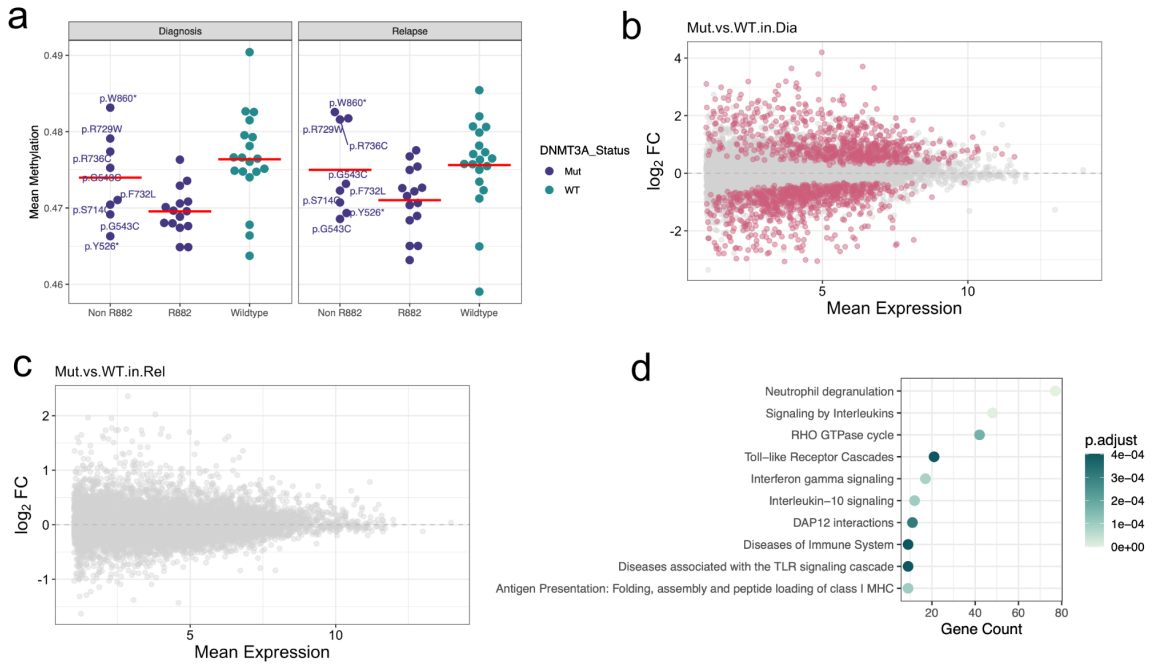



**Figure S2: Associated with Figure 2. a,** Mean methylation of all CpGs analyzed in DNMT3A-Mut and WT patients for relapse and diagnosis samples, wilcox test. **b,** MA-plot of differential expression analysis between DNMT3A-Mut and DNMT3A-WT diagnosis samples, red: significant genes with adj.P < 0.05, gray: not significant genes. **c,** MA-plot of differential expression analysis between DNMT3A-Mut and DNMT3A-WT relapse samples, red: significant genes with adj.P < 0.05, gray: not significant genes. **d,** Overrepresentation analysis, reactome, top10 categories shown.

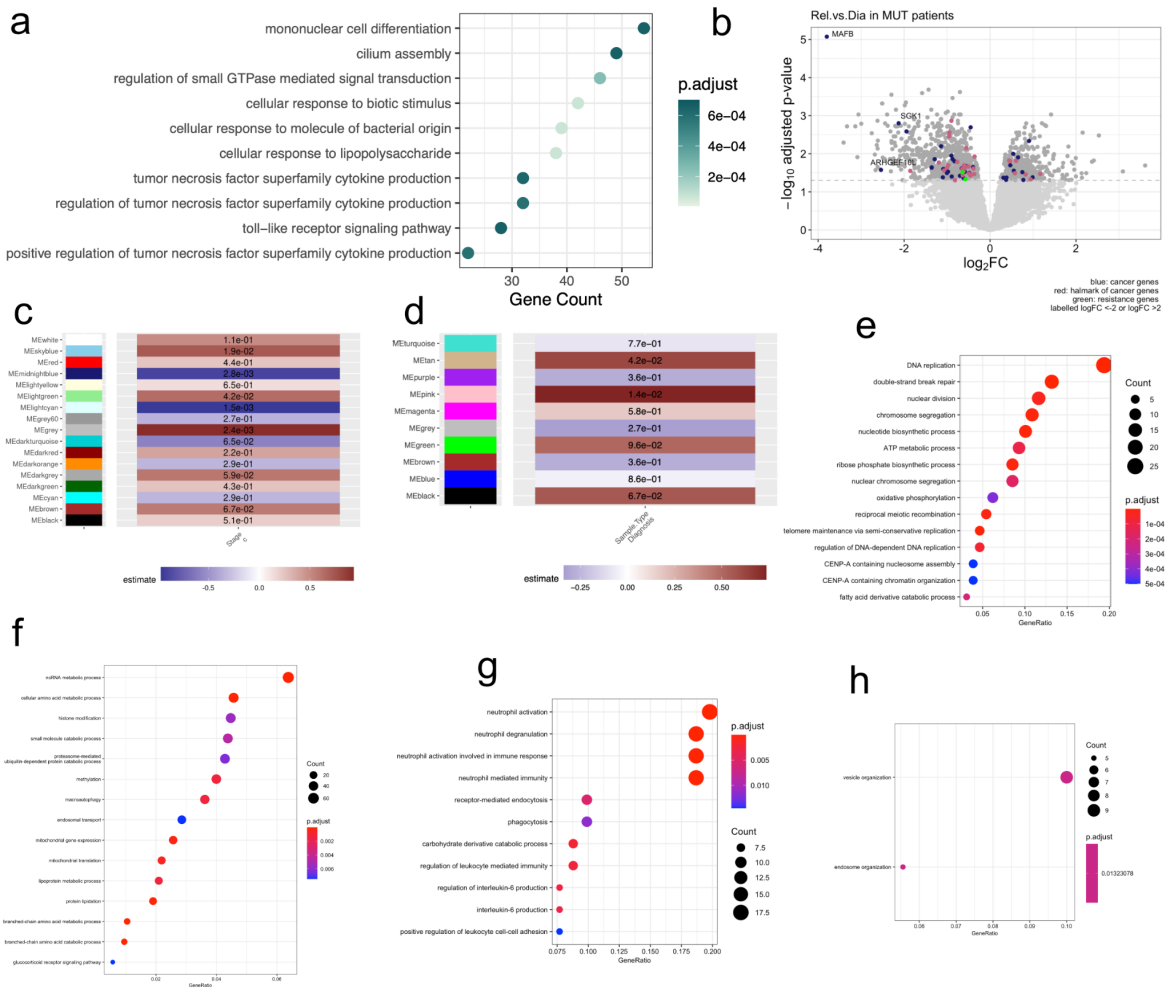



**Figure S3: Associated with Figure 3. a,** Overrepresentation analysis of differentially upregulated genes between relapse and diagnosis in DNMT3A-Mut patients, top10 categories shown. **b,** volcano plot of differential expression analysis between relapse and diagnosis in DNMT3A-Mut patients, red: significant genes with adj.P < 0.05, gray: no significance. **c,d,** Module Sample Associations **c,** Association of Relapse with co-expression modules. **d,** Association of relapse with co-methylation modules. **e-h,** GO analysis of co-expression modules. **e,** skyblye module. **f,** lightgreen module. **g,** midnightblue module. **h,** lightcyan module.

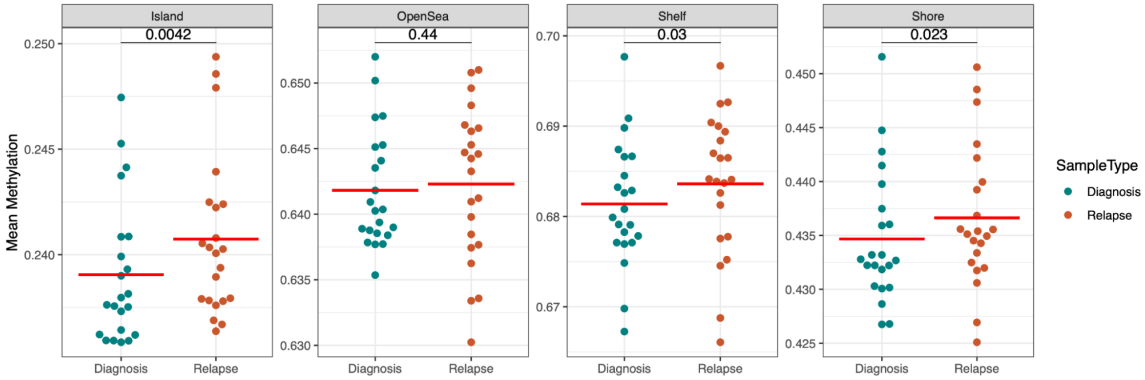



**Figure S4: Associated with Figure 4.** Mean methylation of all CpGs analyzed in relapse and diagnosis samples of DNMT3A-Mut patients, per genomic location relative to CpG Islands

# Discussion

Acute myeloid leukemia (AML) is a highly heterogeneous hematologic malignancy, characterized by the occurrence of varying genetic drivers and clonal evolution (Arnone et al. 2020). Despite this genetic complexity, mutation frequencies are low in comparison to other cancers (Kandoth et al. 2013) and the distribution of mutations before and after relapse remains mainly constant (Li et al. 2016). Relapse-specific mutations, gained or lost exclusively at relapse, have been reported in individual studies but could not be verified across multiple patient cohorts. However, mutations recurrently gained or lost at relapse were frequently detected in genes associated with epigenetic regulation, including *KDM6A*, *DNMT3A*, *IDH1/2* or *TET1/2* (Vosberg and Greif 2019; Greif et al. 2018). While genetic aberrations are essential for AML initiation and development, those results suggest disease progression to be mainly independent of genetic factors and highlight the importance of epigenetic regulation in the evolution of relapse (Hassan et al. 2017). Consonantly, Li et al. found that epigenetic allele (epiallele) burden, defined as the shift of DNA methylation pattern at specific loci against a normal bone marrow control, is of higher prognostic relevance for relapse prediction than the mutation burden. In this cohort, time to relapse was notably shorter in patients with high epiallele burden as compared to those with low epiallele burden (Li et al. 2016). Furthermore, distinctive DNA methylation and histone acetylation signatures have been identified that can distinguish promyelocytic leukemia (APL) patients with high risk of relapse (routinely defined by WBC and platelet count) from a low/intermediate risk group. Notably, differences in the mutation spectrum between high and low/intermediate risk could not be found in this cohort (Singh et al. 2018). Chemotherapy resistance is prevalent in relapsed AML and represents the major form of treatment failure (Döhner et al. 2015). Epigenomic alterations therefore pose promising

new targets for the development of anti-cancer drugs. In fact, several new drugs targeting DNA hypermethylation have been discovered in recent years, including the IDH1 and IDH2 inhibitors Ivosidenib and Enasidenib, already approved for the treatment of relapsed patients (Tang et al. 2021). Despite recent advances in the treatment of AML, relapse rate remains very high and only few options for the treatment of relapsed AML are available. A better understanding of the mechanisms involved in relapse progression is therefore crucial and the development of new targeted therapies is desperately needed.

## Loss of the histone modifiers KDM6A and EZH2

Here, we investigate the impact of the histone demethylase KDM6A and the histone methyltransferase EZH2 on disease progression and chemoresistance. We found that the loss of each of those epigenetic regulators confers drug-resistance in AML. Monitoring resistance to chemotherapeutic agents with the use of proliferation assays in several leukemic cell lines, we observed KDM6A$^{-/-}$ cells to be more resistant to cytarabine and daunorubicin, while EZH2$^{-/-}$ cells displayed increased resistance only against cytarabine. This relation could be confirmed in vivo by the establishment of patient derived xenograft (PDX) models of AML patients carrying mutations in either *EZH2* or *KDM6A*. The tumor suppressive role of KDM6A during leukemia development seems to be conferred by different mechanisms dependent on the type of leukemia. While leukemogenesis in T-ALL relies on the demethylase function of KDM6A, it was found to be independent of this process in AML (Van der Meulen et al. 2015; Gozdecka et al. 2018).

In conformity with those findings, we discovered KDM6A-mediated chemoresistance to be conducted by H3K27 acetylation. Analyzing RNA-Seq and ChIP-Seq data of KDM6A$^{-/-}$ or mutant cell lines, we found loss of KDM6A to be responsible for the methylation-independent acetylation of the membrane transporter gene *ENT1*, resulting in its downregulation and subsequent reduction of the cells' drug intake. Zhang et al. added to

those discoveries by establishing an oncogenic connection of KDM6A expression and imatinib resistance in CML, likewise conferred by demethylase-independent functions (Zhang et al. 2021). Equivalently, EZH2 has been described to possess a dual role as either tumor suppressor or oncogene, and both phenotypes have been associated with chemoresistance in AML. While our study and Göllner et al. found AraC resistance to be the result of EZH2 loss (Göllner et al. 2017), Jiang et al. in contrast, described reduced EZH2 activity to enhance the effect of adriamycin (Jiang et al. 2021). Aberrant expression of EZH2 as well as KDM6A therefore appear to have a crucial influence on treatment success and sensitivity to various chemotherapeutics. In our cohort of matched diagnosis and relapse patient samples, we observed KDM6A as well as EZH2 gene and protein expression to be heterogeneously regulated. In most patients, mRNA and protein levels of both genes were either comparatively higher or lower expressed in relapse than in diagnosis, with similar distributions. Screening for expression of those genes in addition to the mutation status might therefore be beneficial for new treatment approaches. Mutations in *KDM6A* or *EZH2* are relatively rare. In a study of 664 AML patients only 2% carried *KDM6A* and only 4% carried *EZH2* mutations (Metzeler et al. 2016). In our matched diagnosis and relapse patient samples, mutations in *EZH2* were gained at relapse in 1 patient and in *KDM6A* in 2 patients (Greif et al. 2018). It therefore stands to reason that the mutation status alone cannot explain the observed change of EZH2 or KDM6A expression in most patients. Consistently, Göllner et al. were able to explain the reduction of EZH2 levels in some patients by proteasomal degradation, provoked by EZH2 phosphorylation at threonine 487 (Göllner et al. 2017). In those patients, EZH2 levels were restored by treatment with the proteasomal inhibitor bortezomib, previously proposed for the treatment of AML in combination with cytarabine and daunorubicin (Attar et al. 2013). In case of KDM6A, bortezomib in vivo treatment of a PDX model bearing reduced KDM6A expression did not result in an elevation of KDM6A levels, suggesting proteasome-independent regulation in

this model (Stief 2018). A likely mechanism for *KDM6A* regulation could be DNA methylation. In a publicly available dataset of 344 AML patients (Figueroa et al. 2010) we found high *KDM6A* methylation level (upper quartile) to be correlated to shorter overall survival (OS). This is in accordance with the data of Greif et al. that furthermore proposed a correlation of low *KDM6A* gene expression and poor overall survival (Greif et al. 2018). Further studies will be needed to determine the exact effect of *KDM6A* methylation on gene regulation. Next to the effect of KDM6A loss on overall survival, we similarly observed poor overall and relapse-free survival in patients with low EZH2 gene expression. This relation seems surprising, given the antagonistic role of both genes on H3K27 trimethylation. However, a dual role, possessing both tumor suppressor as well as oncogenic functions, has been described in both genes (Basheer et al. 2019; Hua et al. 2021). We furthermore observed H3K27 trimethylation to be regulated by a combined effect of KDM6A and EZH2 in several cell lines. This connection was exploited to propose a new treatment approach for multiple myeloma (MM) patients with a loss of KDM6A. Treating *KDM6A*$^{-/-}$ cell lines with the EZH2 inhibitor GSK343, Ezponda et al. demonstrated that a reduction of EZH2 was able to rebalance global H3K27me3 level and led to a decrease in viability of those cells (Ezponda et al. 2017). These results furthermore highlight the close interaction of both genes and the importance of H3K27me3 level regulation in disease progression and therapy. Decreased H3K27me3 levels established by EZH2 knockout were also observed in our K562 cell line model. Using RNA-Seq and proteomics as well as a publicly available EZH2 ChIP-Seq data set we identified *FHL1* and *UBE2E1* to be direct targets of EZH2 and observed their upregulation upon EZH2 loss. Either of those genes has been associated with resistance against cytarabine in AML. Fu et al. found FHL1 expression to be elevated in leukemia stem cells (LSCs) and observed overexpression of FHL1 to confer cytarabine resistance in an AML cell line model. The group furthermore established a connection between high *FHL1* expression and

upregulation of the efflux transporter genes *ABCC1* and *ABCC4* (encoding MRP1 and MRP2) as well as downregulation of the influx transporter gene *SLC29A1* (ENT1), all responsible for the transmembrane transport of cytarabine (Fu et al. 2020). Downregulation of ENT1 was likewise observed in our KDM6A KO/KD cell line models. Since KDM6A and EZH2 have antagonistic roles on the trimethylation of H3K27, those results suggest different mechanisms leading to ENT1 downregulation by either EZH2 or KDM6A loss. Indeed, while we assume reduced H3K27me3 due to EZH2 loss to be the reason for FHL1 upregulation, we found evidence of a H3K27me3-independent mechanism for ENT1 downregulation due to KDM6A loss. In a H3K27me3 and H3K27ac ChIP-Seq experiment of the KDM6A WT MM-1 and KDM6A mutant MM-6 sister cell lines, we observed differential H3K27ac in the ENT1 promoter as well as an enhancer region but could not detect any changes in the H3K27me3 levels. Similarly, we observed lower levels of H3K27ac in ENT1 for the knockdown of *KDM6A* in the cell lines K562 and THP-1. Those results are in concordance with several studies highlighting a methylation-independent function of KDM6A in normal embryonic stem cell (ESC) differentiation as well as in KDM6A's tumor suppressor function in pancreatic cancer and in AML (Wang et al. 2012; Andricovich et al. 2018; Gozdecka et al. 2018). However, we also obtained opposing results. In our K562 KDM6A KO cells we observed WT KDM6A to re-sensitize cells to AraC, while a catalytically dead KDM6A variant did not induce any sensitivity changes, suggesting drug response to be dependent on the histone H3K27 demethylase function. As mentioned before, EZH2 inhibitors have shown promising results in the treatment of several KDM6A-depleted cancer types by restoring normal H3K27me3 levels (Van der Meulen et al. 2015; Ezponda et al. 2017; Wang et al. 2018). It therefore stands to reason that the mechanism conferring chemoresistance in KDM6A-depleted cells is not completely independent of H3K27me3, however its influence might vary depending on the cancer type, cell line or even on specific genomic locations. In contrast to AraC resistance, we observed

resistance against daunorubicin (DNR) in KDM6A-depleted but not in EZH2-depleted cells. ENT1 is a nucleoside transporter and the main facilitator of import for the pyrimidine analog cytarabine. However, it is incapable of transporting the anthracycline daunorubicin (Zhang et al. 2007). For this reason, daunorubicin resistance in KDM6A-depleted cells needs to be conferred by other mechanisms. Indeed, treatment with the ENT1 inhibitor NBMPR did not result in resistance against daunorubicin. Using the same data, upregulation of the aldo-keto reductases AKR1C1 and AKR1C2 was found in KDM6A-depleted cells (Stief 2018). Involved in the metabolism of chemotherapeutic drugs, members of the hydroxysteroid dehydrogenase family (AKR1Cs) have been linked to chemotherapy resistance including resistance against daunorubicin in several cancer types (Bortolozzi et al. 2018; Matsunaga et al. 2014; Verma et al. 2019). Although the efflux transporter MRP1 (Multidrug Resistance Protein 1) has been described to transport daunorubicin and overexpression of MRP1 has been associated with chemotherapy resistance in AML, we did not observe daunorubicin resistance in our EZH2$^{-/-}$ K562 cell line model (Hooijberg et al. 2000; Fu et al. 2020). Göllner et al. however described daunorubicin resistance in EZH2-depleted MV4-11 cells (Göllner et al. 2017). Daunorubicin resistance due to a loss of EZH2 therefore seems to be dependent on the cell line and could even be patient-specific. Further investigation is needed to determine the exact mechanism of EZH2-dependent daunorubicin resistance as well as the specific involvement of FHL1 and its role in ABCC1 upregulation. Similar to many cancer types, a gender bias for the benefit of females can be observed in AML. Males are diagnosed more often with AML than females and have in general shorter overall survival (Acharya et al. 2018; Hossain and Xie 2015; Cartwright et al. 2002). In a study including 21 tumor types, X-chromosomal tumor suppressor genes (TSGs) including KDM6A, escaping X-inactivation, have been proposed as a possible reason for this discrepancy (Dunford et al. 2017). Possessing two functional versions of these TSGs, females are to some extent protected from gene loss. In accordance, mutations in

*KDM6A* have been observed more frequently in males than in females (Van der Meulen et al. 2015; van Haaften et al. 2009) and low *KDM6A* expression correlating with poor prognosis was observed in male AML patients (Greif et al. 2018). Besides *KDM6A*, several other genes have been identified to harbor mutations more often in either male or female patients, despite not being located on the X-chromosome. Those included mutations in *RUNX1*, *ASXL1* and *EZH2*, more prevalent in male patients, and *NPM1* and *DNMT3A* mutations, more prevalent in female patients (Metzeler et al. 2016). Since many gender-specific mutations have been identified to be preleukemic or associated to preleukemic mutations, gender differences could be applicable in clinical risk stratification or even disease preventive measures (De-Morgan et al. 2021).

## Loss of the DNA methyltransferase DNMT3A

In a study of 188 women, healthy at study baseline but developed AML with a median of 9.6 years later, DNMT3A mutations were identified as a predictor of increased risk and accelerated time of AML onset (Desai et al. 2018). Insufficient to cause AML on their own, DNMT3A mutations have been observed to promote HSC self-renewal and cause reduced differentiation (Challen et al. 2011). This suggests a potential role of DNMT3A mutations as drivers of clonal hematopoiesis, facilitating the acquisition of further genetic lesions like FLT3-ITD that in turn cause the transformation to AML (Yang et al. 2016; Buscarlet et al. 2017). Those studies clearly indicate the importance of DNMT3A mutations for disease onset. Disease maintenance however, has been found to function independently of DNMT3A expression (Yang et al. 2016). HSCs harboring preleukemic mutations, including DNMT3A have been observed to survive induction chemotherapy and persist in remission, contributing to healthy hematopoiesis (Corces-Zimmerman et al. 2014; Greif et al. 2018). Persisting DNMT3A mutations in remission have furthermore been observed to correlate with shorter time to relapse and DNMT3A mutations in this cohort were never lost at

relapse (Greif et al. 2018). This suggests an involvement of DNMT3A in relapse progression, possibly facilitating the growth of leukemia cells ultimately leading to the reintroduction of AML.

Here we investigate changes in gene expression as well as DNA methylation in the relapse of AML patients harboring DNMT3A mutations. In comparison to diagnosis samples of the same patient, we found DNA methylation to be involved in the downregulation of genes controlling myeloid differentiation and observed an enrichment of leukemia stem cell (LSC) signatures. Our results are therefore in accordance with Hackl et al. who similarly found increased "stemness" of relapse by describing an LSC signature enrichment in the relapse of 11 CN-AML patients. (Hackl et al. 2015). Those findings support the concept of a leukemia stem cell model, in which AML relapse is induced by a subpopulation of chemoresistant cells with self-renewal capacity and the ability to reinitiate leukemic growth (Batlle and Clevers 2017; Senft and Jeremias 2019). In a recent study of Duy et al. the authors were able to extend this concept by analyzing AML cells surviving chemotherapy directly after treatment. They found AML cells to enter a senescence-like phenotype to survive genotoxic stress induced by chemotherapy. Initially depleted of LSC populations, those senescence-like cells were able to repopulate leukemia giving rise to relapse populations with an enrichment of LSC signatures (Duy et al. 2021). To reinstate tumor growth, senescent cells would need to somehow overcome their dormant condition and re-enter the cell cycle. Indeed, we observed upregulation of the aminoacyl-tRNA synthetase (ARS) LARS1, recently described as a key player in senescence escape. Responsible for the ligation of the amino acids leucine and tyrosine to their corresponding tRNA, LARS1 and YARS1 have been found to promote senescence escape by the reactivation of tRNA synthesis (Guillon et al. 2021). Besides LARS1, we furthermore observed upregulation of two other aminoacyl-tRNA synthetases, FARS2 and EARS2, encoded by the nuclear genome but transported to the mitochondrion, which were not

covered in the study of Guillon et al. who focused on the genomic ARSs only. However, they could exert similar functions by assisting the upregulation of metabolic processes required for tumor repopulation. Although we observed global hypermethylation in relapse, no significant change in DNMT3A expression could be found, suggesting an involvement of other DNA methylation modulators. The family of DNMTs comprises 4 more members (DNMT1, DNMT2, DNMT3B and DNMT3L), out of which only DNMT1 and DNMT3B are catalytically active (Lyko 2018). Overexpression of either gene has been associated with DNA hypermethylation (Kobayashi et al. 2011; Biniszkiewicz et al. 2002) as well as AML pathogenesis (Monteferrario et al. 2014; Lamba et al. 2018; Solly et al. 2017; Mizuno et al. 2001). Our data revealed upregulation of DNMT3B in relapse and is therefore in accordance with the study of Challen et al, who found Dnmt3b to generate aberrant CpG-Island (CGI) hypermethylation in HSCs with a Dnmt3a knock-out (Challen et al. 2014). Both DNMT3A and DNMT3B are involved in normal hematopoietic stem cell differentiation and loss of function mutations in either gene have been associated with a differentiation block in AML (Koya et al. 2016; Challen et al. 2011; Challen et al. 2014. However, we found upregulation of DNMT3B in relapse samples, depicting increased “stemness”. This result is in accordance with the study of Niederwieser et al. describing association of high DNMT3B expression with a gene expression profile comprising genes in differentiation, proliferation and survival pathways in primary CN-AML patients (Niederwieser et al. 2015). This profile included the downregulated genes CD14, TLR4, CEBPB and TLR8, involved in macrophage differentiation, that were likewise found downregulated in our relapse samples, suggesting DNMT3B to be involved in the impairment of differentiation. Furthermore, we found an enrichment of this gene set amongst genes differentially expressed between relapse and diagnosis in our data, confirming the relation of high DNMT3B expression and regulation of gene expression. Niederwieser et al. furthermore proposed a methylation-independent role of DNMT3B, as the groups of high and low DNMT3B-expressing patients in their data did

not depict any differential methylation. In contrast, we found upregulation of DNMT3B and hypermethylation in relapse and confirmed a correlation of methylation and gene expression in 31 genes. Eight of which (CCR1, CRISPLD2 FAM49A, LILRB4, TFEB, TMEM104, SLC11A1 and PLEKHO2) have been furthermore found associated with DNMT3B expression by Niederwieser et al. This suggests DNMT3B to regulate gene expression by its methyltransferase activity in our data. Similar results were obtained by a study of Poole et al. describing aberrant DNA methylation upon DNMT3B knock-down in a T-ALL cell line model (Poole et al. 2017). The combined results of the mentioned studies argue for an involvement of DNMT3B on the establishment of DNA methylation and the resulting regulation of gene expression. However, this relation might be patient- or cell type-specific and is likely dependent on the expression of other DNA methylation modulators. Further investigations are therefore necessary to unravel the effect of DNMT3B on AML progression as well as its interplay with DNMT3A. Besides upregulation of *DNMT3A*, we additionally detected downregulation of *APOBEC3A* and observed a negative correlation of APOBEC3A and DNA methylation levels. Belonging to a family of cytosine deaminating enzymes, APOBEC3A has been found involved in the demethylation of 5-methylcytosine (5mC) (Schutsky et al. 2017). Demethylation by those enzymes is provoked by the deamination of modified cytosines, resulting in a genomic mismatch that subsequently is repaired by base excision repair (BER) creating unmodified cytosines (Bhutani et al. 2011). Cytosine deaminases have been found to play a role in active demethylation in several cell types (Dominguez et al. 2015; Milagre et al. 2017) and their deficiency has furthermore been linked to increased methylation levels (Popp et al. 2010). In CLL they have furthermore been linked to poor prognosis (Schubert et al. 2018). However, the function of APOBEC3A in hematopoietic and AML cells as well as AML disease progression still needs to be investigated.

# Conclusions and Outlook

Despite recent treatment advances, relapse rate in AML remains very high and is the main reason for treatment failure. Epigenetic regulation has recently been identified as a key player in relapse progression and chemoresistance and therefore poses a promising target for drug discovery. New insights on the influence of epigenetic regulation on AML disease progression and chemoresistance are therefore desperately needed. I contributed to this effort by investigating the impact of loss-of-function mutations in the epigenetic regulators KDM6A, EZH2 and DNMT3A. By analyzing RNA-Seq, ChIP-Seq and proteomics data I contributed to the identification of two mechanisms of drug resistance conferred by the loss of either KDM6A or EZH2. I furthermore studied disease progression in a cohort of DNMT3A mutated patients and described an involvement of hypermethylation on "stemness" reduction by analyzing RNA-Seq as well as methylation BeadChip data.

At present, approved epigenetic therapies comprise IDH1- and IDH2 inhibitors as well as hypomethylating agents (HMA), however, many more are currently undergoing clinical trials. The involvement of EZH2 in numerous biological processes and its dual tumor suppressive and oncogenic role, have made it a promising target in AML therapy. Several EZH2 inhibitors have already shown substantial antileukemic activity and the proteasome inhibitor bortezomib has been found to stabilize EZH2 levels in patients with reduced EZH2 expression. EZH2 inhibitors have furthermore achieved good results in the treatment of KDM6A-depleted cells, restoring normal H3K27me3 levels and the equivalent application of KDM6A inhibitors on EZH2-depleted cells might be possible (Rinke et al. 2020). Modulating epigenetic processes in AML therapy therefore is a complex endeavor that needs to be handled cautiously. Ongoing clinical trials will soon shed light on the potential and scope of application of many more epigenetic therapies.

# References


Abelson, Sagi, Grace Collord, Stanley W. K. Ng, Omer Weissbrod, Netta Mendelson Cohen, Elisabeth Niemeyer, Noam Barda, et al. 2018. "Prediction of Acute Myeloid Leukaemia Risk in Healthy Individuals." *Nature* 559 (7714): 400–404.

Acharya, Utkarsh H., Anna B. Halpern, Qian Vicky Wu, Jenna M. Voutsinas, Roland B. Walter, Seongseok Yun, Mohammed Kanaan, and Elihu H. Estey. 2018. "Impact of Region of Diagnosis, Ethnicity, Age, and Gender on Survival in Acute Myeloid Leukemia (AML)." *Journal of Drug Assessment* 7 (1): 51–53.

Agger, Karl, Paul A. C. Cloos, Jesper Christensen, Diego Pasini, Simon Rose, Juri Rappsilber, Irina Issaeva, Eli Canaani, Anna Elisabetta Salcini, and Kristian Helin. 2007. "UTX and JMJD3 Are Histone H3K27 Demethylases Involved in HOX Gene Regulation and Development." *Nature* 449 (7163): 731–34.

Albrecht, Tara A. 2014. "Physiologic and Psychological Symptoms Experienced by Adults with Acute Leukemia: An Integrative Literature Review." *Oncology Nursing Forum* 41 (3): 286–95.

Andricovich, Jaclyn, Stephanie Perkail, Yan Kai, Nicole Casasanta, Weiqun Peng, and Alexandros Tzatsos. 2018. "Loss of KDM6A Activates Super-Enhancers to Induce Gender-Specific Squamous-like Pancreatic Cancer and Confers Sensitivity to BET Inhibitors." *Cancer Cell* 33 (3): 512–26.e8.

Arber, Daniel A., Attilio Orazi, Robert Hasserjian, Jürgen Thiele, Michael J. Borowitz, Michelle M. Le Beau, Clara D. Bloomfield, Mario Cazzola, and James W. Vardiman. 2016. "The 2016 Revision to the World Health Organization Classification of Myeloid Neoplasms and Acute Leukemia." *Blood* 127 (20): 2391–2405.

Arnone, Marlon, Martina Konantz, Pauline Hanns, Anna M. Paczulla Stanger, Sarah Bertels, Parimala Sonika Godavarthy, Maximilian Christopeit, and Claudia Lengerke. 2020. "Acute Myeloid Leukemia Stem Cells: The Challenges of Phenotypic Heterogeneity." *Cancers* 12 (12). https://doi.org/10.3390/cancers12123742.

Aryee, Martin J., Andrew E. Jaffe, Hector Corrada-Bravo, Christine Ladd-Acosta, Andrew P. Feinberg, Kasper D. Hansen, and Rafael A. Irizarry. 2014. "Minfi: A Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium DNA Methylation Microarrays." *Bioinformatics* 30 (10): 1363–69.

Attar, Eyal C., Jeffrey L. Johnson, Philip C. Amrein, Gerard Lozanski, Martha Wadleigh, Daniel J. DeAngelo, Jonathan E. Kolitz, et al. 2013. "Bortezomib Added to Daunorubicin and Cytarabine during Induction Therapy and to Intermediate-Dose Cytarabine for Consolidation in Patients with Previously Untreated Acute Myeloid Leukemia Age 60 to 75 Years: CALGB (Alliance) Study 10502." *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 31 (7): 923–29.

Baccarella, Alyssa, Claire R. Williams, Jay Z. Parrish, and Charles C. Kim. 2018. "Empirical Assessment of the Impact of Sample Number and Read Depth on RNA-Seq Analysis Workflow Performance." *BMC Bioinformatics* 19 (1): 1–12.

Bannister, Andrew J., and Tony Kouzarides. 2011. “Regulation of Chromatin by Histone Modifications.” *Cell Research* 21 (3): 381–95.

Barski, Artem, Suresh Cuddapah, Kairong Cui, Tae-Young Roh, Dustin E. Schones, Zhibin Wang, Gang Wei, Iouri Chepelev, and Keji Zhao. 2007. “High-Resolution Profiling of Histone Methylations in the Human Genome.” *Cell* 129 (4): 823–37.

Basheer, Faisal, George Giotopoulos, Eshwar Meduri, Haiyang Yun, Milena Mazan, Daniel Sasca, Paolo Gallipoli, et al. 2019. “Contrasting Requirements during Disease Evolution Identify EZH2 as a Therapeutic Target in AML.” *The Journal of Experimental Medicine* 216 (4): 966–81.

Batlle, Eduard, and Hans Clevers. 2017. “Cancer Stem Cells Revisited.” *Nature Medicine* 23 (10): 1124–34.

Bennett, J. M., D. Catovsky, M. T. Daniel, G. Flandrin, D. A. Galton, H. R. Gralnick, and C. Sultan. 1985. “Proposed Revised Criteria for the Classification of Acute Myeloid Leukemia. A Report of the French-American-British Cooperative Group.” *Annals of Internal Medicine* 103 (4): 620–25.

Bentley, David R., Shankar Balasubramanian, Harold P. Swerdlow, Geoffrey P. Smith, John Milton, Clive G. Brown, Kevin P. Hall, et al. 2008. “Accurate Whole Human Genome Sequencing Using Reversible Terminator Chemistry.” *Nature* 456 (7218): 53–59.

Bhutani, Nidhi, David M. Burns, and Helen M. Blau. 2011. “DNA Demethylation Dynamics.” *Cell* 146 (6): 866–72.

Bibikova, Marina, Bret Barnes, Chan Tsan, Vincent Ho, Brandy Klotzle, Jennie M. Le, David Delano, et al. 2011. “High Density DNA Methylation Array with Single CpG Site Resolution.” *Genomics* 98 (4): 288–95.

Bibikova, Marina, Jennie Le, Bret Barnes, Shadi Saedinia-Melnyk, Lixin Zhou, Richard Shen, and Kevin L. Gunderson. 2009. “Genome-Wide DNA Methylation Profiling Using Infinium® Assay.” *Epigenomics* 1 (1): 177–200.

Biniszkiewicz, Detlev, Joost Gribnau, Bernard Ramsahoye, François Gaudet, Kevin Eggan, David Humpherys, Mary-Ann Mastrangelo, Zhan Jun, Jörn Walter, and Rudolf Jaenisch. 2002. “Dnmt1 Overexpression Causes Genomic Hypermethylation, Loss of Imprinting, and Embryonic Lethality.” *Molecular and Cellular Biology* 22 (7): 2124–35.

Booth, Michael J., Tobias W. B. Ost, Dario Beraldi, Neil M. Bell, Miguel R. Branco, Wolf Reik, and Shankar Balasubramanian. 2013. “Oxidative Bisulfite Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine.” *Nature Protocols* 8 (10): 1841–51.

Bortolozzi, Roberta, Silvia Bresolin, Elena Rampazzo, Maddalena Paganin, Francesca Maule, Elena Mariotto, Daniele Boso, et al. 2018. “AKR1C Enzymes Sustain Therapy Resistance in Paediatric T-ALL.” *British Journal of Cancer* 118 (7): 985–94.

Buenrostro, Jason D., Paul G. Giresi, Lisa C. Zaba, Howard Y. Chang, and William J. Greenleaf. 2013. “Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position.” *Nature Methods* 10 (12): 1213–18.

Buscarlet, Manuel, Sylvie Provost, Yassamin Feroz Zada, Amina Barhdadi, Vincent Bourgoin, Guylaine Lépine, Luigina Mollica, Natasha Szuber, Marie-Pierre Dubé, and Lambert Busque. 2017. “DNMT3A and TET2 Dominate Clonal Hematopoiesis and Demonstrate

Benign Phenotypes and Different Genetic Predispositions." *Blood* 130 (6): 753–62.

Cancer Genome Atlas Research Network, Timothy J. Ley, Christopher Miller, Li Ding, Benjamin J. Raphael, Andrew J. Mungall, A. Gordon Robertson, et al. 2013. "Genomic and Epigenomic Landscapes of Adult de Novo Acute Myeloid Leukemia." *The New England Journal of Medicine* 368 (22): 2059–74.

Cartwright, Ray A., Karen A. Gurney, and Anthony V. Moorman. 2002. "Sex Ratios and the Risks of Haematological Malignancies." *British Journal of Haematology* 118 (4): 1071–77.

Chakraborty, Sajib, Md Ismail Hosen, Musaddeque Ahmed, and Hossain Uddin Shekhar. 2018. "Onco-Multi-OMICS Approach: A New Frontier in Cancer Research." *BioMed Research International* 2018 (October): 9836256.

Challen, Grant A., Deqiang Sun, Mira Jeong, Min Luo, Jaroslav Jelinek, Jonathan S. Berg, Christoph Bock, et al. 2011. "Dnmt3a Is Essential for Hematopoietic Stem Cell Differentiation." *Nature Genetics* 44 (1): 23–31.

Challen, Grant A., Deqiang Sun, Allison Mayle, Mira Jeong, Min Luo, Benjamin Rodriguez, Cates Mallaney, et al. 2014. "Dnmt3a and Dnmt3b Have Overlapping and Distinct Functions in Hematopoietic Stem Cells." *Cell Stem Cell* 15 (3): 350–64.

Chen, Kent T. J., Roger Gilabert-Oriol, Marcel B. Bally, and Ada W. Y. Leung. 2019. "Recent Treatment Advances and the Role of Nanotechnology, Combination Products, and Immunotherapy in Changing the Therapeutic Landscape of Acute Myeloid Leukemia." *Pharmaceutical Research* 36 (9): 125.

Chhangawala, Sagar, Gabe Rudy, Christopher E. Mason, and Jeffrey A. Rosenfeld. 2015. "The Impact of Read Length on Quantification of Differentially Expressed Genes and Splice Junction Detection." *Genome Biology* 16 (June): 131.

Chodavarapu, Ramakrishna K., Suhua Feng, Yana V. Bernatavichute, Pao-Yang Chen, Hume Stroud, Yanchun Yu, Jonathan A. Hetzel, et al. 2010. "Relationship between Nucleosome Positioning and DNA Methylation." *Nature* 466 (7304): 388–92.

Clapier, Cedric R., Janet Iwasa, Bradley R. Cairns, and Craig L. Peterson. 2017. "Mechanisms of Action and Regulation of ATP-Dependent Chromatin-Remodelling Complexes." *Nature Reviews. Molecular Cell Biology* 18 (7): 407–22.

Conerly, Melissa L., Sheila S. Teves, Daniel Diolaiti, Michelle Ulrich, Robert N. Eisenman, and Steven Henikoff. 2010. "Changes in H2A.Z Occupancy and DNA Methylation during B-Cell Lymphomagenesis." *Genome Research* 20 (10): 1383–90.

Conesa, Ana, Pedro Madrigal, Sonia Tarazona, David Gomez-Cabrero, Alejandra Cervera, Andrew McPherson, Michał Wojciech Szcześniak, et al. 2016. "A Survey of Best Practices for RNA-Seq Data Analysis." *Genome Biology* 17 (January): 13.

Corces-Zimmerman, M. Ryan, Wan-Jen Hong, Irving L. Weissman, Bruno C. Medeiros, and Ravindra Majeti. 2014. "Preleukemic Mutations in Human Acute Myeloid Leukemia Affect Epigenetic Regulators and Persist in Remission." *Proceedings of the National Academy of Sciences of the United States of America* 111 (7): 2548–53.

Daver, Naval, Andrew H. Wei, Daniel A. Pollyea, Amir T. Fathi, Paresh Vyas, and Courtney D. DiNardo. 2020. "New Directions for Emerging Therapies in Acute Myeloid Leukemia:

The next Chapter." *Blood Cancer Journal* 10 (10): 107.

Dedeurwaerder, S., M. Defrance, M. Bizet, E. Calonne, G. Bontempi, and F. Fuks. 2014. "A Comprehensive Overview of Infinium HumanMethylation450 Data Processing." *Briefings in Bioinformatics* 15 (6): 929–41.

De Kouchkovsky, I., and M. Abdul-Hay. 2016. "Acute Myeloid Leukemia: A Comprehensive Review and 2016 Update." *Blood Cancer Journal* 6 (7): e441.

De-Morgan, Aviv, Manja Meggendorfer, Claudia Haferlach, and Liran Shlush. 2021. "Male Predominance in AML Is Associated with Specific Preleukemic Mutations." *Leukemia* 35 (3): 867–70.

Deneberg, Stefan, Philippe Guardiola, Andreas Lennartsson, Ying Qu, Verena Gaidzik, Odile Blanchet, Mohsen Karimi, et al. 2011. "Prognostic DNA Methylation Patterns in Cytogenetically Normal Acute Myeloid Leukemia Are Predefined by Stem Cell Chromatin Marks." *Blood* 118 (20): 5573–82.

Desai, Pinkal, Nuria Mencia-Trinchant, Oleksandr Savenkov, Michael S. Simon, Gloria Cheang, Sangmin Lee, Michael Samuel, et al. 2018. "Somatic Mutations Precede Acute Myeloid Leukemia Years before Diagnosis." *Nature Medicine* 24 (7): 1015–23.

Dijk, Erwin L. van, Hélène Auger, Yan Jaszczyszyn, and Claude Thermes. 2014. "Ten Years of next-Generation Sequencing Technology." *Trends in Genetics: TIG* 30 (9): 418–26.

Dijk, Erwin L. van, Yan Jaszczyszyn, and Claude Thermes. 2014. "Library Preparation Methods for next-Generation Sequencing: Tone down the Bias." *Experimental Cell Research* 322 (1): 12–20.

Ding, Li, Timothy J. Ley, David E. Larson, Christopher A. Miller, Daniel C. Koboldt, John S. Welch, Julie K. Ritchey, et al. 2012. "Clonal Evolution in Relapsed Acute Myeloid Leukaemia Revealed by Whole-Genome Sequencing." *Nature* 481 (7382): 506–10.

Dobin, Alexander, and Thomas R. Gingeras. 2015. "Mapping RNA-Seq Reads with STAR." *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]* 51 (September): 11.14.1–19.

Döhner, Hartmut, Elihu Estey, David Grimwade, Sergio Amadori, Frederick R. Appelbaum, Thomas Büchner, Hervé Dombret, et al. 2017. "Diagnosis and Management of AML in Adults: 2017 ELN Recommendations from an International Expert Panel." *Blood* 129 (4): 424–47.

Döhner, Hartmut, Daniel J. Weisdorf, and Clara D. Bloomfield. 2015. "Acute Myeloid Leukemia." *The New England Journal of Medicine* 373 (12): 1136–52.

Domcke, Silvia, Anaïs Flore Bardet, Paul Adrian Ginno, Dominik Hartl, Lukas Burger, and Dirk Schübeler. 2015. "Competition between DNA Methylation and Transcription Factors Determines Binding of NRF1." *Nature* 528 (7583): 575–79.

Dominguez, Pilar M., Matt Teater, Nyasha Chambwe, Matthias Kormaksson, David Redmond, Jennifer Ishii, Bao Vuong, et al. 2015. "DNA Methylation Dynamics of Germinal Center B Cells Are Mediated by AID." *Cell Reports* 12 (12): 2086–98.

Dunford, Andrew, David M. Weinstock, Virginia Savova, Steven E. Schumacher, John P. Cleary, Akinori Yoda, Timothy J. Sullivan, et al. 2017. "Tumor-Suppressor Genes That

Escape from X-Inactivation Contribute to Cancer Sex Bias." *Nature Genetics* 49 (1): 10–16.

Dupree, Emmalyn J., Madhuri Jayathirtha, Hannah Yorkey, Marius Mihasan, Brindusa Alina Petre, and Costel C. Darie. 2020. "A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of This Field." *Proteomes* 8 (3). https://doi.org/10.3390/proteomes8030014.

Duy, Cihangir, Meng Li, Matt Teater, Cem Meydan, Francine E. Garrett-Bakelman, Tak C. Lee, Christopher R. Chin, et al. 2021. "Chemotherapy Induces Senescence-Like Resilient Cells Capable of Initiating AML Recurrence." *Cancer Discovery* 11 (6): 1542–61.

Esteller, Manel. 2007. "Epigenetic Gene Silencing in Cancer: The DNA Hypermethylome." *Human Molecular Genetics* 16 Spec No 1 (April): R50–59.

Ezponda, Teresa, Daphné Dupéré-Richer, Christine M. Will, Eliza C. Small, Nobish Varghese, Tej Patel, Behnam Nabet, et al. 2017. "UTX/KDM6A Loss Enhances the Malignant Phenotype of Multiple Myeloma and Sensitizes Cells to EZH2 Inhibition." *Cell Reports* 21 (3): 628–40.

Faulk, Christopher. 2019. "Chapter 2-2 - Implications of DNA Methylation in Toxicology." In *Toxicoepigenetics*, edited by Shaun D. McCullough and Dana C. Dolinoy, 153–71. Academic Press.

Fedurco, Milan, Anthony Romieu, Scott Williams, Isabelle Lawrence, and Gerardo Turcatti. 2006. "BTA, a Novel Reagent for DNA Attachment on Glass and Efficient Generation of Solid-Phase Amplified DNA Colonies." *Nucleic Acids Research* 34 (3): e22.

Feng, Qin, Hengbin Wang, Huck Hui Ng, Hediye Erdjument-Bromage, Paul Tempst, Kevin Struhl, and Yi Zhang. 2002. "Methylation of H3-Lysine 79 Is Mediated by a New Family of HMTases without a SET Domain." *Current Biology: CB* 12 (12): 1052–58.

Figueroa, Maria E., Sanne Lugthart, Yushan Li, Claudia Erpelinck-Verschueren, Xutao Deng, Paul J. Christos, Elizabeth Schifano, et al. 2010. "DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia." *Cancer Cell* 17 (1): 13–27.

Fu, Yue, Man Xu, Zelong Cui, Zongcheng Yang, Zhiyong Zhang, Xiaolin Yin, Xiangnan Huang, Minran Zhou, Xiaoming Wang, and Chunyan Chen. 2020. "Genome-Wide Identification of FHL1 as a Powerful Prognostic Candidate and Potential Therapeutic Target in Acute Myeloid Leukaemia." *EBioMedicine* 52 (February): 102664.

Giani, Alice Maria, Guido Roberto Gallo, Luca Gianfranceschi, and Giulio Formenti. 2020. "Long Walk to Genomics: History and Current Approaches to Genome Sequencing and Assembly." *Computational and Structural Biotechnology Journal* 18: 9–19.

Göllner, Stefanie, Thomas Oellerich, Shuchi Agrawal-Singh, Tino Schenk, Hans-Ulrich Klein, Christian Rohde, Caroline Pabst, et al. 2017. "Loss of the Histone Methyltransferase EZH2 Induces Resistance to Multiple Drugs in Acute Myeloid Leukemia." *Nature Medicine* 23 (1): 69–78.

Gozdecka, Malgorzata, Eshwar Meduri, Milena Mazan, Konstantinos Tzelepis, Monika Dudek, Andrew J. Knights, Mercedes Pardo, et al. 2018. "UTX-Mediated Enhancer and Chromatin Remodeling Suppresses Myeloid Leukemogenesis through Noncatalytic Inverse Regulation of ETS and GATA Programs." *Nature Genetics* 50 (6): 883–94.

Greer, Eric L., and Yang Shi. 2012. “Histone Methylation: A Dynamic Mark in Health, Disease and Inheritance.” *Nature Reviews. Genetics* 13 (5): 343–57.

Greif, Philipp A., Luise Hartmann, Sebastian Vosberg, Sophie M. Stief, Raphael Mattes, Ines Hellmann, Klaus H. Metzeler, et al. 2018. “Evolution of Cytogenetically Normal Acute Myeloid Leukemia During Therapy and Relapse: An Exome Sequencing Study of 50 Patients.” *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 24 (7): 1716–26.

Grimwade, David, Adam Ivey, and Brian J. P. Huntly. 2016. “Molecular Landscape of Acute Myeloid Leukemia in Younger Adults and Its Clinical Relevance.” *Blood* 127 (1): 29–41.

Guillon, Jordan, Bertrand Toutain, Coralie Petit, Hugo Coquelet, Cécile Henry, Alice Boissard, Catherine Guette, and Olivier Coqueret. 2021. “tRNA Biogenesis and Specific Aminoacyl-tRNA Synthetases Regulate Senescence Stability Under the Control of mTOR.” *bioRxiv*. https://doi.org/10.1101/2020.04.30.068114.

Gujral, Palak, Vishakha Mahajan, Abbey C. Lissaman, and Anna P. Ponnampalam. 2020. “Histone Acetylation and the Role of Histone Deacetylases in Normal Cyclic Endometrium.” *Reproductive Biology and Endocrinology: RB&E* 18 (1): 84.

Haaften, Gijs van, Gillian L. Dalgliesh, Helen Davies, Lina Chen, Graham Bignell, Chris Greenman, Sarah Edkins, et al. 2009. “Somatic Mutations of the Histone H3K27 Demethylase Gene UTX in Human Cancer.” *Nature Genetics* 41 (5): 521–23.

Hackl, Hubert, Katarina Steinleitner, Karin Lind, Sybille Hofer, Natasa Tosic, Sonja Pavlovic, Nada Suvajdzic, Heinz Sill, and Rotraud Wieser. 2015. “A Gene Expression Profile Associated with Relapse of Cytogenetically Normal Acute Myeloid Leukemia Is Enriched for Leukemia Stem Cell Genes.” *Leukemia & Lymphoma* 56 (4): 1126–28.

Hasin, Yehudit, Marcus Seldin, and Aldons Lusis. 2017. “Multi-Omics Approaches to Disease.” *Genome Biology* 18 (1): 83.

Hassan, Ciaran, Ebrahim Afshinnekoo, Sheng Li, Shixiu Wu, and Christopher E. Mason. 2017. “Genetic and Epigenetic Heterogeneity and the Impact on Cancer Relapse.” *Experimental Hematology* 54 (October): 26–30.

Héberlé, Éléa, and Anaïs Flore Bardet. 2019. “Sensitivity of Transcription Factors to DNA Methylation.” *Essays in Biochemistry* 63 (6): 727–41.

Hong, Mingye, Shuang Tao, Ling Zhang, Li-Ting Diao, Xuanmei Huang, Shaohui Huang, Shu-Juan Xie, Zhen-Dong Xiao, and Hua Zhang. 2020. “RNA Sequencing: New Technologies and Applications in Cancer Research.” *Journal of Hematology & Oncology* 13 (1): 166.

Hooijberg, J. H., H. M. Pinedo, C. Vrasdonk, W. Priebe, J. Lankelma, and H. J. Broxterman. 2000. “The Effect of Glutathione on the ATPase Activity of MRP1 in Its Natural Membranes.” *FEBS Letters* 469 (1): 47–51.

Hossain, Md Jobayer, and Li Xie. 2015. “Sex Disparity in Childhood and Young Adult Acute Myeloid Leukemia (AML) Survival: Evidence from US Population Data.” *Cancer Epidemiology* 39 (6): 892–900.

Hrdlickova, Radmila, Masoud Toloue, and Bin Tian. 2017. “RNA-Seq Methods for Transcriptome Analysis.” *Wiley Interdisciplinary Reviews. RNA* 8 (1).

https://doi.org/10.1002/wrna.1364.

Hua, Chunyan, Jiaqing Chen, Shuting Li, Jianan Zhou, Jiahong Fu, Weijian Sun, and Wenqian Wang. 2021. “KDM6 Demethylases and Their Roles in Human Cancers.” *Frontiers in Oncology* 11 (December): 779918.

Jaiswal, Siddhartha, Pierre Fontanillas, Jason Flannick, Alisa Manning, Peter V. Grauman, Brenton G. Mar, R. Coleman Lindsley, et al. 2014. “Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes.” *The New England Journal of Medicine* 371 (26): 2488–98.

Jang, Hyun Sik, Woo Jung Shin, Jeong Eon Lee, and Jeong Tae Do. 2017. “CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function.” *Genes* 8 (6). https://doi.org/10.3390/genes8060148.

Jan, Max, Thomas M. Snyder, M. Ryan Corces-Zimmerman, Paresh Vyas, Irving L. Weissman, Stephen R. Quake, and Ravindra Majeti. 2012. “Clonal Evolution of Preleukemic Hematopoietic Stem Cells Precedes Human Acute Myeloid Leukemia.” *Science Translational Medicine* 4 (149): 149ra118.

Jiang, Xuejie, Ling Jiang, Jiaying Cheng, Fang Chen, Jinle Ni, Changxin Yin, Qiang Wang, et al. 2021. “Inhibition of EZH2 by Chidamide Exerts Antileukemia Activity and Increases Chemosensitivity through Smo/Gli-1 Pathway in Acute Myeloid Leukemia.” *Journal of Translational Medicine* 19 (1): 117.

Jjingo, Daudi, Andrew B. Conley, Soojin V. Yi, Victoria V. Lunyak, and I. King Jordan. 2012. “On the Presence and Role of Human Gene-Body DNA Methylation.” *Oncotarget* 3 (4): 462–74.

Kandoth, Cyriac, Michael D. McLellan, Fabio Vandin, Kai Ye, Beifang Niu, Charles Lu, Mingchao Xie, et al. 2013. “Mutational Landscape and Significance across 12 Major Cancer Types.” *Nature* 502 (7471): 333–39.

Kantarjian, Hagop, Tapan Kadia, Courtney DiNardo, Naval Daver, Gautam Borthakur, Elias Jabbour, Guillermo Garcia-Manero, Marina Konopleva, and Farhad Ravandi. 2021. “Acute Myeloid Leukemia: Current Progress and Future Directions.” *Blood Cancer Journal* 11 (2): 41.

Karpievitch, Yuliya V., Ashoka D. Polpitiya, Gordon A. Anderson, Richard D. Smith, and Alan R. Dabney. 2010. “Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects.” *The Annals of Applied Statistics* 4 (4): 1797–1823.

Kempf, Julia M., Sabrina Weser, Michael D. Bartoschek, Klaus H. Metzeler, Binje Vick, Tobias Herold, Kerstin Völse, et al. 2021. “Loss-of-Function Mutations in the Histone Methyltransferase EZH2 Promote Chemotherapy Resistance in AML.” *Scientific Reports* 11 (1): 5838.

Kircher, Martin, and Janet Kelso. 2010. “High-Throughput DNA Sequencing--Concepts and Limitations.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 32 (6): 524–36.

Kivioja, Teemu, Anna Vähärautio, Kasper Karlsson, Martin Bonke, Martin Enge, Sten Linnarsson, and Jussi Taipale. 2011. “Counting Absolute Numbers of Molecules Using

Unique Molecular Identifiers." *Nature Methods* 9 (1): 72–74.

Kobayashi, Yuya, Devin M. Absher, Zulfiqar G. Gulzar, Sarah R. Young, Jesse K. McKenney, Donna M. Peehl, James D. Brooks, Richard M. Myers, and Gavin Sherlock. 2011. "DNA Methylation Profiling Reveals Novel Biomarkers and Important Roles for DNA Methyltransferases in Prostate Cancer." *Genome Research* 21 (7): 1017–27.

Kouzarides, Tony. 2007. "Chromatin Modifications and Their Function." *Cell* 128 (4): 693–705.

Koya, Junji, Keisuke Kataoka, Tomohiko Sato, Masashige Bando, Yuki Kato, Takako Tsuruta-Kishino, Hiroshi Kobayashi, et al. 2016. "DNMT3A R882 Mutants Interact with Polycomb Proteins to Block Haematopoietic Stem and Leukaemic Cell Differentiation." *Nature Communications* 7 (March): 10924.

Lamba, Jatinder K., Xueyuan Cao, Susana C. Raimondi, Roya Rafiee, James R. Downing, Shi Lei, Tanja Gruber, Raul C. Ribeiro, Jeffrey E. Rubnitz, and Stanley B. Pounds. 2018. "Integrated Epigenetic and Genetic Analysis Identifies Markers of Prognostic Significance in Pediatric Acute Myeloid Leukemia." *Oncotarget* 9 (42): 26711–23.

Lauber, Chris, Nádia Correia, Andreas Trumpp, Michael A. Rieger, Anna Dolnik, Lars Bullinger, Ingo Roeder, and Michael Seifert. 2020. "Survival Differences and Associated Molecular Signatures of DNMT3A-Mutant Acute Myeloid Leukemia Patients." *Scientific Reports* 10 (1): 12761.

Lavarone, Elisa, Caterina M. Barbieri, and Diego Pasini. 2019. "Dissecting the Role of H3K27 Acetylation and Methylation in PRC2 Mediated Control of Cellular Identity." *Nature Communications* 10 (1): 1679.

Ledergerber, Christian, and Christophe Dessimoz. 2011. "Base-Calling for next-Generation Sequencing Platforms." *Briefings in Bioinformatics* 12 (5): 489–97.

Leisch, Michael, Bettina Jansko, Nadja Zaborsky, Richard Greil, and Lisa Pleyer. 2019. "Next Generation Sequencing in AML-On the Way to Becoming a New Standard for Treatment Initiation And/or Modulation?" *Cancers* 11 (2). https://doi.org/10.3390/cancers11020252.

Levin, Joshua Z., Moran Yassour, Xian Adiconis, Chad Nusbaum, Dawn Anne Thompson, Nir Friedman, Andreas Gnirke, and Aviv Regev. 2010. "Comprehensive Comparative Analysis of Strand-Specific RNA Sequencing Methods." *Nature Methods* 7 (9): 709–15.

Liao, Yang, and Wei Shi. 2020. "Read Trimming Is Not Required for Mapping and Quantification of RNA-Seq Reads at the Gene Level." *NAR Genomics and Bioinformatics* 2 (3): lqaa068.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. "featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features." *Bioinformatics* 30 (7): 923–30.

Li, Sheng, Francine E. Garrett-Bakelman, Stephen S. Chung, Mathijs A. Sanders, Todd Hricik, Franck Rapaport, Jay Patel, et al. 2016. "Distinct Evolution and Dynamics of Epigenetic and Genetic Heterogeneity in Acute Myeloid Leukemia." *Nature Medicine* 22 (7): 792–99.

Lister, Ryan, Mattia Pelizzola, Robert H. Dowen, R. David Hawkins, Gary Hon, Julian

Tonti-Filippini, Joseph R. Nery, et al. 2009. “Human DNA Methylomes at Base Resolution Show Widespread Epigenomic Differences.” *Nature* 462 (7271): 315–22.

Lyko, Frank. 2018. “The DNA Methyltransferase Family: A Versatile Toolkit for Epigenetic Regulation.” *Nature Reviews. Genetics* 19 (2): 81–92.

Marchesi, Francesco, Ombretta Annibali, Elisabetta Cerchiara, Maria Cristina Tirindelli, and Giuseppe Avvisati. 2011. “Cytogenetic Abnormalities in Adult Non-Promyelocytic Acute Myeloid Leukemia: A Concise Review.” *Critical Reviews in Oncology/hematology* 80 (3): 331–46.

Margueron, Raphaël, and Danny Reinberg. 2011. “The Polycomb Complex PRC2 and Its Mark in Life.” *Nature* 469 (7330): 343–49.

Marioni, John C., Christopher E. Mason, Shrikant M. Mane, Matthew Stephens, and Yoav Gilad. 2008. “RNA-Seq: An Assessment of Technical Reproducibility and Comparison with Gene Expression Arrays.” *Genome Research* 18 (9): 1509–17.

Martignoles, Jean-Alain, François Delhommeau, and Pierre Hirsch. 2018. “Genetic Hierarchy of Acute Myeloid Leukemia: From Clonal Hematopoiesis to Molecular Residual Disease.” *International Journal of Molecular Sciences* 19 (12). https://doi.org/10.3390/ijms19123850.

Matsunaga, Toshiyuki, Ayano Yamaguchi, Yoshifumi Morikawa, Chihiro Kezuka, Hiroaki Takazawa, Satoshi Endo, Ossama El-Kabbani, Kazuo Tajima, Akira Ikari, and Akira Hara. 2014. “Induction of Aldo-Keto Reductases (AKR1C1 and AKR1C3) Abolishes the Efficacy of Daunorubicin Chemotherapy for Leukemic U937 Cells.” *Anti-Cancer Drugs* 25 (8): 868–77.

McEwen, Lisa M., Meaghan J. Jones, David Tse Shen Lin, Rachel D. Edgar, Lucas T. Husquin, Julia L. MacIsaac, Katia E. Ramadori, et al. 2018. “Systematic Evaluation of DNA Methylation Age Estimation with Common Preprocessing Methods and the Infinium MethylationEPIC BeadChip Array.” *Clinical Epigenetics* 10 (1): 123.

Meissner, Alexander, Tarjei S. Mikkelsen, Hongcang Gu, Marius Wernig, Jacob Hanna, Andrey Sivachenko, Xiaolan Zhang, et al. 2008. “Genome-Scale DNA Methylation Maps of Pluripotent and Differentiated Cells.” *Nature* 454 (7205): 766–70.

Metzeler, Klaus H., Tobias Herold, Maja Rothenberg-Thurley, Susanne Amler, Maria C. Sauerland, Dennis Görlich, Stephanie Schneider, et al. 2016. “Spectrum and Prognostic Relevance of Driver Gene Mutations in Acute Myeloid Leukemia.” *Blood* 128 (5): 686–98.

Milagre, Inês, Thomas M. Stubbs, Michelle R. King, Julia Spindel, Fátima Santos, Felix Krueger, Martin Bachman, et al. 2017. “Gender Differences in Global but Not Targeted Demethylation in iPSC Reprogramming.” *Cell Reports* 18 (5): 1079–89.

Mizuno, S., T. Chijiwa, T. Okamura, K. Akashi, Y. Fukumaki, Y. Niho, and H. Sasaki. 2001. “Expression of DNA Methyltransferases DNMT1, 3A, and 3B in Normal Hematopoiesis and in Acute and Chronic Myelogenous Leukemia.” *Blood* 97 (5): 1172–79.

Monteferrario, D., S. M. Noordermeer, S. Bergevoet, G. Huls, J. H. Jansen, and B. A. van der Reijden. 2014. “High DNA-Methyltransferase 3B Expression Predicts Poor Outcome in Acute Myeloid Leukemia, Especially among Patients with Co-Occurring NPM1 and

FLT3 Mutations." *Blood Cancer Journal* 4 (August): e233.

Morganti, Stefania, Paolo Tarantino, Emanuela Ferraro, Paolo D'Amico, Giulia Viale, Dario Trapani, Bruno Achutti Duso, and Giuseppe Curigliano. 2020. "Role of Next-Generation Sequencing Technologies in Personalized Medicine." In *P5 eHealth: An Agenda for the Health Technologies of the Future*, edited by Gabriella Pravettoni and Stefano Triberti, 125–54. Cham: Springer International Publishing.

Morris, Tiffany J., and Stephan Beck. 2015. "Analysis Pipelines and Packages for Infinium HumanMethylation450 BeadChip (450k) Data." *Methods* 72 (January): 3–8.

Mortazavi, Ali, Brian A. Williams, Kenneth McCue, Lorian Schaeffer, and Barbara Wold. 2008. "Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq." *Nature Methods* 5 (7): 621–28.

Nakabayashi, Kazuhiko. 2017. "Illumina HumanMethylation BeadChip for Genome-Wide DNA Methylation Profiling: Advantages and Limitations." In *Handbook of Nutrition, Diet, and Epigenetics*, edited by Vinood Patel and Victor Preedy, 1–15. Cham: Springer International Publishing.

Niederwieser, C., J. Kohlschmidt, S. Volinia, S. P. Whitman, K. H. Metzeler, A-K Eisfeld, K. Maharry, et al. 2015. "Prognostic and Biologic Significance of DNMT3B Expression in Older Patients with Cytogenetically Normal Primary Acute Myeloid Leukemia." *Leukemia* 29 (3): 567–75.

Ooi, Steen K. T., Chen Qiu, Emily Bernstein, Keqin Li, Da Jia, Zhe Yang, Hediye Erdjument-Bromage, et al. 2007. "DNMT3L Connects Unmethylated Lysine 4 of Histone H3 to de Novo Methylation of DNA." *Nature* 448 (7154): 714–17.

Ostrander, Elizabeth L., Ashley C. Kramer, Cates Mallaney, Hamza Celik, Won Kyun Koh, Jake Fairchild, Emily Haussler, Christine R. C. Zhang, and Grant A. Challen. 2020. "Divergent Effects of Dnmt3a and Tet2 Mutations on Hematopoietic Progenitor Cell Fitness." *Stem Cell Reports* 14 (4): 551–60.

Ozsolak, Fatih, and Patrice M. Milos. 2011. "RNA Sequencing: Advances, Challenges and Opportunities." *Nature Reviews. Genetics* 12 (2): 87–98.

Parekh, Swati, Christoph Ziegenhain, Beate Vieth, Wolfgang Enard, and Ines Hellmann. 2016. "The Impact of Amplification on Differential Expression Analyses by RNA-Seq." *Scientific Reports* 6 (May): 25533.

Parker, Carol E., Maria R. Warren, and Viorel Mocanu. 2011. "Mass Spectrometry for Proteomics." In *Neuroproteomics*, edited by Oscar Alzate. Boca Raton (FL): CRC Press/Taylor & Francis.

Picelli, Simone, Asa K. Björklund, Björn Reinius, Sven Sagasser, Gösta Winberg, and Rickard Sandberg. 2014. "Tn5 Transposase and Tagmentation Procedures for Massively Scaled Sequencing Projects." *Genome Research* 24 (12): 2033–40.

Pidsley, Ruth, Elena Zotenko, Timothy J. Peters, Mitchell G. Lawrence, Gail P. Risbridger, Peter Molloy, Susan Van Djik, Beverly Muhlhausler, Clare Stirzaker, and Susan J. Clark. 2016. "Critical Evaluation of the Illumina MethylationEPIC BeadChip Microarray for Whole-Genome DNA Methylation Profiling." *Genome Biology* 17 (1): 208.

Pinney, Sara E. 2014. "Mammalian Non-CpG Methylation: Stem Cells and Beyond." *Biology*

3 (4): 739–51.

Poole, Candace J., Wenli Zheng, Atul Lodh, Aleksey Yevtodiyenko, Daniel Liefwalker, Honglin Li, Dean W. Felsher, and Jan van Riggelen. 2017. “DNMT3B Overexpression Contributes to Aberrant DNA Methylation and MYC-Driven Tumor Maintenance in T-ALL and Burkitt’s Lymphoma.” *Oncotarget* 8 (44): 76898–920.

Popp, Christian, Wendy Dean, Suhua Feng, Shawn J. Cokus, Simon Andrews, Matteo Pellegrini, Steven E. Jacobsen, and Wolf Reik. 2010. “Genome-Wide Erasure of DNA Methylation in Mouse Primordial Germ Cells Is Affected by AID Deficiency.” *Nature* 463 (7284): 1101–5.

Portela, Anna, and Manel Esteller. 2010. “Epigenetic Modifications and Human Disease.” *Nature Biotechnology* 28 (10): 1057–68.

Rea, S., F. Eisenhaber, D. O’Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, et al. 2000. “Regulation of Chromatin Structure by Site-Specific Histone H3 Methyltransferases.” *Nature* 406 (6796): 593–99.

Renaud, Gabriel, Udo Stenzel, Tomislav Maricic, Victor Wiebe, and Janet Kelso. 2015. “deML: Robust Demultiplexing of Illumina Sequences Using a Likelihood-Based Approach.” *Bioinformatics* 31 (5): 770–72.

Rhee, I., K. W. Jair, R. W. Yen, C. Lengauer, J. G. Herman, K. W. Kinzler, B. Vogelstein, S. B. Baylin, and K. E. Schuebel. 2000. “CpG Methylation Is Maintained in Human Cancer Cells Lacking DNMT1.” *Nature* 404 (6781): 1003–7.

Rinke, Jenny, Andrew Chase, Nicholas C. P. Cross, Andreas Hochhaus, and Thomas Ernst. 2020. “EZH2 in Myeloid Malignancies.” *Cells* 9 (7). https://doi.org/10.3390/cells9071639.

Robert, Marie-France, Steves Morin, Normand Beaulieu, France Gauthier, Ian C. Chute, Annie Barsalou, and A. Robert MacLeod. 2003. “DNMT1 Is Required to Maintain CpG Methylation and Aberrant Gene Silencing in Human Cancer Cells.” *Nature Genetics* 33 (1): 61–65.

Romer-Seibert, Jennifer S., and Sara E. Meyer. 2021. “Genetic Heterogeneity and Clonal Evolution in Acute Myeloid Leukemia.” *Current Opinion in Hematology* 28 (1): 64–70.

Russler-Germain, David A., David H. Spencer, Margaret A. Young, Tamara L. Lamprecht, Christopher A. Miller, Robert Fulton, Matthew R. Meyer, et al. 2014. “The R882H DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers.” *Cancer Cell* 25 (4): 442–54.

Santos, F. P. S., S. Faderl, G. Garcia-Manero, C. Koller, M. Beran, S. O’Brien, S. Pierce, et al. 2009. “Adult Acute Erythroleukemia: An Analysis of 91 Patients Treated at a Single Institution.” *Leukemia* 23 (12): 2275–80.

Schmidt, Andreas, Ignasi Forne, and Axel Imhof. 2014. “Bioinformatic Analysis of Proteomics Data.” *BMC Systems Biology* 8 Suppl 2 (March): S3.

Schones, Dustin E., Kairong Cui, Suresh Cuddapah, Tae-Young Roh, Artem Barski, Zhibin Wang, Gang Wei, and Keji Zhao. 2008. “Dynamic Regulation of Nucleosome Positioning in the Human Genome.” *Cell* 132 (5): 887–98.

Schubert, Maria, Hubert Hackl, Franz Josef Gassner, Richard Greil, and Roland Geisberger. 2018. “Investigating Epigenetic Effects of Activation-Induced Deaminase in Chronic Lymphocytic Leukemia.” *PloS One* 13 (12): e0208753.

Schutsky, Emily K., Christopher S. Nabel, Amy K. F. Davis, Jamie E. DeNizio, and Rahul M. Kohli. 2017. “APOBEC3A Efficiently Deaminates Methylated, but Not TET-Oxidized, Cytosine Bases in DNA.” *Nucleic Acids Research* 45 (13): 7655–65.

Senft, Daniela, and Irmela Jeremias. 2019. “A Rare Subgroup of Leukemia Stem Cells Harbors Relapse-Inducing Potential in Acute Lymphoblastic Leukemia.” *Experimental Hematology* 69 (January): 1–10.

Sheng, Quanhu, Kasey Vickers, Shilin Zhao, Jing Wang, David C. Samuels, Olivia Koues, Yu Shyr, and Yan Guo. 2017. “Multi-Perspective Quality Control of Illumina RNA Sequencing Data Analysis.” *Briefings in Functional Genomics* 16 (4): 194–204.

Shlush, Liran I., Amanda Mitchell, Lawrence Heisler, Sagi Abelson, Stanley W. K. Ng, Aaron Trotman-Grant, Jessie J. F. Medeiros, et al. 2017. “Tracing the Origins of Relapse in Acute Myeloid Leukaemia to Stem Cells.” *Nature* 547 (7661): 104–8.

Short, Nicholas J., Shouhao Zhou, Chenqi Fu, Donald A. Berry, Roland B. Walter, Sylvie D. Freeman, Christopher S. Hourigan, et al. 2020. “Association of Measurable Residual Disease With Survival Outcomes in Patients With Acute Myeloid Leukemia: A Systematic Review and Meta-Analysis.” *JAMA Oncology* 6 (12): 1890–99.

Singh, Abhishek A., Francesca Petraglia, Angela Nebbioso, Guoqiang Yi, Mariarosaria Conte, Sergio Valente, Amit Mandoli, et al. 2018. “Multi-Omics Profiling Reveals a Distinctive Epigenome Signature for High-Risk Acute Promyelocytic Leukemia.” *Oncotarget* 9 (39): 25647–60.

Sinha, Ankit, and Matthias Mann. 2020. “A Beginner’s Guide to Mass Spectrometry–based Proteomics.” *The Biochemist* 42 (5): 64–69.

Slatko, Barton E., Andrew F. Gardner, and Frederick M. Ausubel. 2018. “Overview of Next-Generation Sequencing Technologies.” *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]* 122 (1): e59.

Solly, Françoise, Catherine Koering, Aminetou Mint Mohamed, Delphine Maucort-Boulch, Guillaume Robert, Patrick Auberger, Pascale Flandrin-Gresta, et al. 2017. “An miRNA-DNMT1 Axis Is Involved in Azacitidine Resistance and Predicts Survival in Higher-Risk Myelodysplastic Syndrome and Low Blast Count Acute Myeloid Leukemia.” *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 23 (12): 3025–34.

Spencer, David H., Bilal Al-Khalil, David Russler-Germain, Tamara Lamprecht, Nicole Havey, Robert S. Fulton, Michelle O’Laughlin, Catrina Fronick, Richard K. Wilson, and Timothy J. Ley. 2014. “Whole-Genome Bisulfite Sequencing of Primary AML Cells with the DNMT3A R882H Mutation Identifies Regions of Focal Hypomethylation That Are Associated with Open Chromatin.” *Blood* 124 (21): 608–608.

Stief, Sophie M. 2018. “The Role of KDM6A in the Clonal Evolution of Acute Myeloid Leukemia.” Ludwig-Maximilians-Universität München. https://edoc.ub.uni-muenchen.de/24668/1/Stief_Sophie_Maria.pdf.

Stief, Sophie M., Anna-Li Hanneforth, Sabrina Weser, Raphael Mattes, Michela Carlet,

Wen-Hsin Liu, Michael D. Bartoschek, et al. 2019. “Loss of KDM6A Confers Drug Resistance in Acute Myeloid Leukemia.” *Leukemia*, June. https://doi.org/10.1038/s41375-019-0497-6.

Swerdlow, Steven H., Elias Campo, N. Lee Harris, Elaine S. Jaffe, Stefano A. Pileri, Harald Stein, Jurgen Thiele, James W. Vardiman, and Others. 2017. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Vol. 2. WORLD HEALTH ORGN.

Talbert, Paul B., and Steven Henikoff. 2021. “Histone Variants at a Glance.” *Journal of Cell Science* 134 (6). https://doi.org/10.1242/jcs.244749.

Tang, Kenny, Andre C. Schuh, and Karen Wl Yee. 2021. “3+7 Combined Chemotherapy for Acute Myeloid Leukemia: Is It Time to Say Goodbye?” *Current Oncology Reports* 23 (10): 120.

Timp, Winston, and Gregory Timp. 2020. “Beyond Mass Spectrometry, the next Step in Proteomics.” *Science Advances* 6 (2): eaax8978.

Tulstrup, Morten, Mette Soerensen, Jakob Werner Hansen, Linn Gillberg, Maria Needhamsen, Katja Kaastrup, Kristian Helin, Kaare Christensen, Joachim Weischenfeldt, and Kirsten Grønbæk. 2021. “TET2 Mutations Are Associated with Hypermethylation at Key Regulatory Enhancers in Normal and Malignant Hematopoiesis.” *Nature Communications* 12 (1): 6061.

Van der Meulen, Joni, Viraj Sanghvi, Konstantinos Mavrakis, Kaat Durinck, Fang Fang, Filip Matthijssens, Pieter Rondou, et al. 2015. “The H3K27me3 Demethylase UTX Is a Gender-Specific Tumor Suppressor in T-Cell Acute Lymphoblastic Leukemia.” *Blood* 125 (1): 13–21.

Verma, Kshitij, Tianzhu Zang, Trevor M. Penning, and Paul C. Trippier. 2019. “Potent and Highly Selective Aldo-Keto Reductase 1C3 (AKR1C3) Inhibitors Act as Chemotherapeutic Potentiators in Acute Myeloid Leukemia and T-Cell Acute Lymphoblastic Leukemia.” *Journal of Medicinal Chemistry* 62 (7): 3590–3616.

Vosberg, Sebastian, and Philipp A. Greif. 2019. “Clonal Evolution of Acute Myeloid Leukemia from Diagnosis to Relapse.” *Genes, Chromosomes & Cancer* 58 (12): 839–49.

Wang, Chaochen, Ji-Eun Lee, Young-Wook Cho, Ying Xiao, Qihuang Jin, Chengyu Liu, and Kai Ge. 2012. “UTX Regulates Mesoderm Differentiation of Embryonic Stem Cells Independent of H3K27 Demethylase Activity.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (38): 15324–29.

Wang, Lu, Zibo Zhao, Patrick A. Ozark, Damiano Fantini, Stacy A. Marshall, Emily J. Rendleman, Kira A. Cozzolino, et al. 2018. “Resetting the Epigenetic Balance of Polycomb and COMPASS Function at Enhancers for Cancer Therapy.” *Nature Medicine* 24 (6): 758–69.

Wang, Zhenxing, Xiaoliang Wu, and Yadong Wang. 2018. “A Framework for Analyzing DNA Methylation Data from Illumina Infinium HumanMethylation450 BeadChip.” *BMC Bioinformatics* 19 (Suppl 5): 115.

Watanabe, Daisuke, Isao Suetake, Takashi Tada, and Shoji Tajima. 2002. “Stage- and Cell-Specific Expression of Dnmt3a and Dnmt3b during Embryogenesis.” *Mechanisms*

*of Development* 118 (1-2): 187–90.

Weber, Michael, Ines Hellmann, Michael B. Stadler, Liliana Ramos, Svante Pääbo, Michael Rebhan, and Dirk Schübeler. 2007. “Distribution, Silencing Potential and Evolutionary Impact of Promoter DNA Methylation in the Human Genome.” *Nature Genetics* 39 (4): 457–66.

“White Blood Cell Count.” 2021. A.D.A.M. Medical Encyclopedia. January 19, 2021. https://medlineplus.gov/ency/presentations/100151_3.htm.

Wilkinson, Alex W., Jonathan Diep, Shaobo Dai, Shuo Liu, Yaw Shin Ooi, Dan Song, Tie-Mei Li, et al. 2019. “SETD3 Is an Actin Histidine Methyltransferase That Prevents Primary Dystocia.” *Nature* 565 (7739): 372–76.

Williams, Claire R., Alyssa Baccarella, Jay Z. Parrish, and Charles C. Kim. 2016. “Trimming of Sequence Reads Alters RNA-Seq Gene Expression Estimates.” *BMC Bioinformatics* 17 (February): 103.

Wu, Michael C., Bonnie R. Joubert, Pei-Fen Kuan, Siri E. Håberg, Wenche Nystad, Shyamal D. Peddada, and Stephanie J. London. 2014. “A Systematic Assessment of Normalization Approaches for the Infinium 450K Methylation Platform.” *Epigenetics: Official Journal of the DNA Methylation Society* 9 (2): 318–29.

Xu, Zongli, Liang Niu, and Jack A. Taylor. 2021. “The ENmix DNA Methylation Analysis Pipeline for Illumina BeadChip and Comparisons with Seven Other Preprocessing Pipelines.” *Clinical Epigenetics* 13 (1): 216.

Yang, Liubin, Benjamin Rodriguez, Allison Mayle, Hyun Jung Park, Xueqiu Lin, Min Luo, Mira Jeong, et al. 2016. “DNMT3A Loss Drives Enhancer Hypomethylation in FLT3-ITD-Associated Leukemias.” *Cancer Cell* 29 (6): 922–34.

Yates, J. R., 3rd. 2000. “Mass Spectrometry. From Genomics to Proteomics.” *Trends in Genetics: TIG* 16 (1): 5–8.

Zhang, Chengwan, Li Shen, Yifu Zhu, Ran Xu, Zhikui Deng, Xiaoning Liu, Yihan Ding, et al. 2021. “KDM6A Promotes Imatinib Resistance through YY1-Mediated Transcriptional Upregulation of TRKA Independently of Its Demethylase Activity in Chronic Myelogenous Leukemia.” *Theranostics* 11 (6): 2691–2705.

Zhang, Jing, Li Jing, Menghan Li, Lingfeng He, and Zhigang Guo. 2019. “Regulation of Histone Arginine Methylation/demethylation by Methylase and Demethylase (Review).” *Molecular Medicine Reports* 19 (5): 3963–71.

Zhang, Jing, Frank Visser, Karen M. King, Stephen A. Baldwin, James D. Young, and Carol E. Cass. 2007. “The Role of Nucleoside Transporters in Cancer Chemotherapy with Nucleoside Drugs.” *Cancer Metastasis Reviews* 26 (1): 85–110.

Zhang, Yaoyang, Bryan R. Fonslow, Bing Shan, Moon-Chang Baek, and John R. Yates 3rd. 2013. “Protein Analysis by Shotgun/bottom-up Proteomics.” *Chemical Reviews* 113 (4): 2343–94.

# List of Figures



# List of Tables

# Acknowledgments

First and foremost, I would like to express my special appreciation and thanks to my mentor **Prof. Karsten Spiekermann** who provided me the opportunity to work in his lab. Thank you, Karsten, for encouraging my research and for your continuous support.

I would like to extend my deepest gratitude to my supervisor **Prof. Wolfgang Enard** who supported my project with insightful input. Thank you, Wolfgang, for your invaluable advice and useful critiques.

I am also grateful to **Ines Hellmann** for her unconditional guidance in bioinformatic analysis and statistics. Thank you, Ines, for sharing your expertise and insights.

Many thanks to my Enard Lab office buddies **Beate Vieth** and **Ilse Valtierra** whose enthusiastic and humorous way has always been contagious.

I would also like to thank my colleagues from the Spiekermann Lab: **Julia Kempf**, **Sophie Stief**, **Anna Vetter**, **Raphael Mattes**, **Harald Polzer** and **Belay Tizazu** for a wonderful work atmosphere, brilliant discussions and many great conversations.

Special thanks to **Swati Parekh** whose passionate and patient teaching was the reason I ended up in bioinformatics and who has supported me in every matter ever since. Swati, I couldn't have done it without you.

Many thanks to **Ines Bliesener** whose expertise in the wet lab is beyond compare. Thank you, Ines, for teaching me all your tips and tricks for succeeding in the lab.

My sincere thanks to the rest of the **AG Enard** as well as all the members of the **SFB 1243** who have been great colleagues and excellent scientists.

I very much appreciate the hard work of the Enard Lab K9 Squad, tirelessly committed to providing emotional support. Thank you, **Elly**, **Daisy,** and **Leyla**.

Last but not least I want to thank my family **Britta Weser**, **Frank Weser** and **Volker Kittke** for always believing in me especially when I myself could not. Thank you for always being there for me.