Development and application
of fluorescent reporter assays
# for the investigation of chromatin regulation

Von der Fakultät 4: Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur
Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

vorgelegt von

Sabine Larissa Pinter

Geboren am 01.08.1991 in Schwäbisch Hall

Hauptberichter: Prof. Dr. Albert Jeltsch

Mitberichter: Prof. Dr. Jörn Lausen

Prüfungsvorsitzender: Prof. Dr. Stephan Nussberger

Tag der mündlichen Prüfung: 28.10.2021
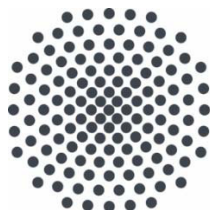
Institut für Biochemie und Technische Biochemie der Universität Stuttgart
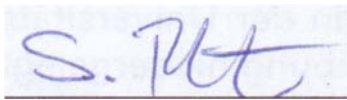
2021

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbst verfasst und dabei keine anderen als die angegeben Quellen und Hilfsmittel verwendet habe.

Declaration of Authorship

I hereby certify that this Dissertation is entirely my own work, apart from where otherwise indicated. Passages and ideas from other sources have been clearly indicated.

Sabine Pinter

10. Juni 2021, Stuttgart

Table of Contents

List of Abbreviations

5mC	5-methyl cytosine
ADD	ATRX-DNMT3-DNMT3L domain; recognizes unmodified H3K4
AMPK	adenosine monophosphate–activated protein kinase
ATP	Adenosin triphosphate
BiAD	Bimolecular Anchor Detector sensor
BiFC	Bimolecular Fluorescence Complementation
BSA	Bovine serum albumin
CD	catalytic domain
CGI	CpG island
COMPASS	Complex of Proteins Associated with Set1
CpG	Cytosine-guanine dinucleotide
CRC	Chromatin remodelling complex
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DDR	DNA damage response
DDX19A	DEAD-box RNA helicase 19A
DNMT	DNA methyltransferase
DSB	Double-strand break
DTT	Dithiothreitol
FITC	Fluorescein isothiocyanate
GQ	G-quadruplex
H2A	Histone protein 2A
HDAC	Histone deacetylase
HMT	Histone methyltransferase
HPTM	Histone post-translational modification
HRP	Horseradish peroxidase
HUSH	Human silencing hub
IDH	Isocitrate dehydrogenase
IP	immunoprecipitation
IPTG	isopropyl β-D-1-thiogalactopyranoside
KRAB	Krüppel-associated Box
lncRNA	Long noncoding RNA (>200 bp)
LSD1	Lysine-specific demethylase 1

MBD	Methyl Binding Domain
MHC	Major histocompatibility complex
MPHOSPH8	M-Phase Phosphoprotein 8
ncRNA	Noncoding RNA
NuRD	Nucleosome remodelling and deacetylase complex
ORF	open reading frame
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate-buffered saline
PEV	Position effect variegation
PFA	paraformaldehyde
PIC	Protease inhibitor cocktail
PPHLN1	Periphilin 1
PRC1/2	Polycomb repressive complex 1/2
PTM	Post-translational modification
PWWP	Pro-Trp-Trp-Pro motif containing domain; recognizes H3K36me2/3
RING	Really interesting new gene; zinc finger with ubiquitin ligase activity
RT	room temperature
RT-qPCR	Reverse-transcription coupled to qPCR
SDS	Sodium dodecyl sulfate
SET	Suppressor of variegation 3 (Su(var)), Enhancer of zeste (E(z)) and trithorax (Trx) domain; domain with lysine methyltransferase activity
SUV	Suppressor of variegation
synP	synthetic promoter sequence; constitutive EF1α promoter with 6x upstream TetR binding sites
TALE	Transcription Activator-Like Effector
tetO	Tet-operon derived Tet-repressor binding site
TF	Transcription factor
tGFP	turbo GFP
TSS	Transcription start site
UHRF1	Ubiquitin-like containing PHD and RING finger domains 1
wt	wild type
ZFP	Zinc-finger protein
ZNF	Zinc-finger
γH2AX	Serine 139 phosphorylated histone variant H2AX

Zusammenfassung

Die Transrkriptionsrate eines Gens im endogenen nukleären Kontext ist eine biologische Größe, deren Ausprägung durch eine Vielzahl von Faktoren bestimmt wird. Die Kontrolle der Gentranskriptionsrate durch eine Veränderung der umgebenden Chromatinstruktur kann durch die Anwesenheit von epigenetischen Multiproteinkomplexen, von kovalenten Chromatinmodifikationen, von RNA Strukturen und strukturbestimmenden biophysikalischen Effekten beeinflusst werden. Die Verknüpfung all dieser Faktoren über regulatorische Interaktionen bildet ein komplexes Netzwerk der epigenetischen Genregulation. Das Ziel dieser Arbeit war es, bisher unbekannte regulatorische Interaktionen von Chromatineffektorproteinen innerhalb dieses Netwerks aufzudecken und zu charakterisieren. Damit sollte das Wissen über essentielle Mechanismen der Steuerung von Genexpression erweitert werden.

Chromatineffektoren sind nukleäre Proteine, welche die Expression eines Gens durch die lokale Einführung oder Aufrechterhaltung eines entweder repressiven oder aktivierenden Chromatinzustands kontrollieren. Um funktionell verknüpfte coregulatorische Proteine umfassend identifizieren zu können, wurde ein neuartiges, dynamisch verfolgbares Reportersystem zur Messung der Aktivität epigenetischer Effektorproteine in lebenden Zellen etabliert. Durch die Kombination dieses fluoreszierenden Reportersystems mit einer leistungsstarken RNAi screening Methode wurde die DEAD-box Helikase 19A (DDX19A) als neuer Coregulator der Lysin-spezifischen Demethylase LSD1 identifiziert. LSD1 reguliert im Zusammenspiel mit vielen Interaktoren die Expression essentieller entwicklungsspezifischer Gene und spielt damit eine wichtige Rolle in der Zelldifferenzierung. Die durch LSD1 gesteuerten Prozesse sind bisher nicht vollständig verstanden und ihre Fehlsteuerung ist verbunden mit der Entwicklung verschiedener Krebsarten. Im Zuge der funktionellen Charakterisierung der Interaktion von LSD1 und DDX19A in Zellen konnte gezeigt werden, dass die Unterdrückung der Expression von *Ddx19a* zu einem Anstieg von R-loop Strukturen und der Reduzierung von LSD1 vermittelter Genrepression führt. Weiter konnte gezeigt werden, dass DDX19A an das dreifach methylierte Lysin 27 an Histon 3 (H3K27me3) bindet, und dass die Aktivität von DDX19A die Expression von Genen durch die Entfernung transkriptionsfördernder R-loops reguliert.

Zusammengenommen führten diese Beobachtungen zur Entdeckung einer bisher unbekannten regulatorischen Kaskade, die abhängig von der Demethylierung von

H3K4 durch LSD1 die Stilllegung stark aktiver Gene herbeiführt. Die H3K4-Demethylierung ermöglicht dem Polycomb Repressive Complex 2 (PRC2) aktiv zu werden und H3K27 zu methylieren, was eine Bindestelle für DDX19A erzeugt. Die Bindung von DDX19A führt zur Entfernung von co-transkriptionell auftretenden R-loops, was die Aktivität von LSD1 und PRC2 wiederum weiter steigert. Diese Beobachtungen beschreiben die Etablierung eines positiven Feedback-loops, welcher die Stilllegung eines stark transkribierten Zielgenes ermöglicht. Diese Entdeckungen liefern ein neues Beispiel für eine Kaskade aus mehreren regulatorischen Ereignissen, welche hintereinander geschaltet auftreten müssen um die epigenetischen Barrieren zu überwinden, die den Übergang eines aktiven in einen inaktiven Chromatinzustand verhindern. Die Komplexität des epigenetischen Netzwerks wird durch das Zusammenspiel solcher Erhaltungsmechanismen mit regulatorischen Feedback-Loops und der stark kontextabhängigen Aktivität von Chromatineffektorproteinen aufgebaut, und stellt nach wie vor eine große Herausforderung für die funktionelle Untersuchung epigenetischer Effektoren *in vivo* dar.

Das synthetische Reportersystem wurde im Weiteren dafür eingesetzt, das Krüppel-assoziierte Box-Protein (KRAB) zu charakterisieren. KRAB ist eine Repressordomäne, die für den Einsatz in synthetischer Biologie und in Epigenom-Editierung weit verbreitet ist. Die Interaktion von KRAB mit verschiedenen Mitgliedern des Human Silencing HUB (HUSH) Komplexes wurde in zwei verschiedenen Zelllinien untersucht. Dabei wurde eine unterschiedliche Bedeutung der verschiedenen HUSH Komponenten für die Fähigkeit von KRAB, das synthetische Reportergen in NIH/3T3 Zellen stillzulegen, festgestellt. In iMEF Zellen konnte kein Effekt der HUSH Komponenten auf die Aktivität von KRAB am Reporter festgestellt werden. Daraufhin wurde das coregulatorische Netzwerk um die KRAB Repressordomäne in iMEF Zellen mittels eines RNAi Screens umfassend analysiert. Die Ergebnisse dieser Arbeit lieferten überraschende Informationen über das beobachtbare Ausmaß an Variabilität bei der Zusammensetzung von Coregulatorkomplexen im Kontext einer einzigen Effektordomäne. Aufbauend auf den Ergebnissen aus diesem Kapitel kann eine tiefergehende Untersuchung der regulatorischen Mechanismen, die die Stilllegung von Genexpression durch KRAB vermitteln, erfolgen. Dies könnte dabei helfen, effektivere und zelltypspezifische Repressorkonstrukte für CRISPRi Ansätze zu entwickeln.

Abstract

The rate of transcription of a gene is a statistical event, that can be influenced by a multitude of nuclear parameters. A mayor factor is the nature of the local chromatin environment, which is composed of regulatory multiprotein complexes, covalent chromatin modifications, RNA structures and structure-determining biophysical effects. These factors integrate into a complex network of epigenetic gene regulation. The aim of this work was to reveal and characterize novel regulatory interactions of known chromatin effector proteins in this network, thereby enhancing the knowledge about essential mechanisms driving gene expression.

Chromatin effectors are nuclear proteins that can control gene expression by the establishment or maintenance of either a repressive or an activating chromatin state. To be able to comprehensively identify functional coregulators of these effectors, a novel tractable fluorescent reporter system to monitor the activity of epigenetic effector proteins in living cells was established. By combination of this reporter system with a state-of-the-art multiplexed RNAi screen, the DEAD-box helicase 19A (DDX19A) was identified as a novel coregulator of the lysine specific demethylase 1 (LSD1). LSD1 plays a pivotal role in cellular differentiation by regulating the expression of key developmental genes in concert with different coregulatory proteins. This process is impaired in different cancer types and incompletely understood. During the functional characterization of the interaction between LSD1 and DDX19A, it was demonstrated that suppression of *Ddx19a* results in an increase of R-loops, and a reduction of LSD1-mediated gene silencing. It further was shown that DDX19A binds to tri-methylated lysine 27 of histone 3 (H3K27me3) and that it regulates gene expression through the removal of transcription promoting R-loops. Together, these results uncovered a so far unknown transcriptional regulatory cascade, in which the downregulation of highly active genes is dependent on the LSD1 mediated demethylation of histone H3 lysine 4 (H3K4). This epigenetic trigger allows the Polycomb repressive complex 2 (PRC2) to methylate H3K27, which serves as a binding site for DDX19A. Finally, the binding of DDX19A leads to the efficient removal of R-loops at active promoters, which further de-represses LSD1 and PRC2 at these sites. The observations describe a novel mechanism that enables the robust repression of a highly expressed target gene by the establishment of a positive feedback loop. The model presents a valuable example for the cascade of multiple events that need to be activated to overcome the epigenetic

barriers preventing the transition from an active to an inactive chromatin state. The complexity of the epigenetic regulatory network that is built by the interplay of such maintenance mechanisms, of feedback-loops and the highly context-dependent activities of chromatin regulators, still poses a great challenge for the functional investigation of epigenetic effectors *in vivo*.

The fluorescent reporter system was further applied for the characterization of the Krüppel-associated box (KRAB) domain, a repressive domain that is widely used in epigenome editing approaches. The interaction of the KRAB domain with different members of the Human Silencing Hub (HUSH) complex was analysed in two different cell lines. This revealed a variability in the importance of the HUSH complex for KRAB mediated silencing of gene expression from the synthetic reporter, compared to published literature and among the two analysed cell lines. The coregulator network surrounding the KRAB repressor domain was thereupon characterized using a comprehensive RNAi screen in iMEF cells. The results from this work delivered surprising data about the observable extent of variability in the assembly of coregulator complexes around a single chromatin effector. Building up on the observations from this chapter, further analysis of the regulatory mechanisms mediating KRAB silencing will help to develop more effective, cell-type specific repressor constructs for CRISPRi approaches.

1. Introduction

1.1. The many facets of epigenetic gene regulation

1.1.1. The definition of epigenetics

The scientific field of epigenetics can look back on several decades of research, during which the definition of the term "epigenetics" has come into existence at an early stage and has been evolving gradually since. Relating to the literal meaning of the term, *epigenetic* information can be described as an additional layer of information on top of the genetic information that is stored in form of the DNA base sequence. This additional level of information plays an essential role in mediating the programmed readout of DNA. The influence becomes apparent when looking at the complexity of a mature organism. Every mammalian organism can be traced back to a single fertilized egg but is eventually composed of a variety of highly specific and phenotypically different somatic cells, each of them carrying the same genetic information. In 1942, Conrad Waddington was the first to mention the term epigenetics as "the processes involved in the mechanism by which the genes of the genotype bring about phenotypic effects" (Waddington, 1942). Back then, this definition described all events during development that translate the genetic starting material contained in a fertilized cell, to shape the final "product": a mature organism (Allis et al., 2007; Waddington, 1953). Without knowledge about the underlying molecular mechanisms, Waddington proposed the model of an epigenetic "landscape" (Waddington, 1957), which describes the developmental process from pluripotent stem cells to differentiated cell types (**Figure 1a**).

At this time, it was known that the information needed for the development of each somatic cell is stored in the nuclei of cells in the form of chromosomes. However, there was a debate on how the information was passed on after fertilization of the egg during embryonic development (Allis et al., 2007). There were two models circulating among embryologists: one acted on the assumption that each cell contains preformed elements that enlarge during development, and the other that the developmental process involves chemical reactions among soluble components that execute a complex developmental plan (Allis et al., 2007). With the description of the helical DNA structure (Watson & Crick, 1953), the nucleosome as a structural unit of chromatin (Kornberg & Thomas, 1974; Richards & Pardon, 1970) and an increasing number of

chemical modifications on chromatin (Murray, 1964; Ruiz-Carrillo et al., 1975), the second hypothesis became more conclusive.
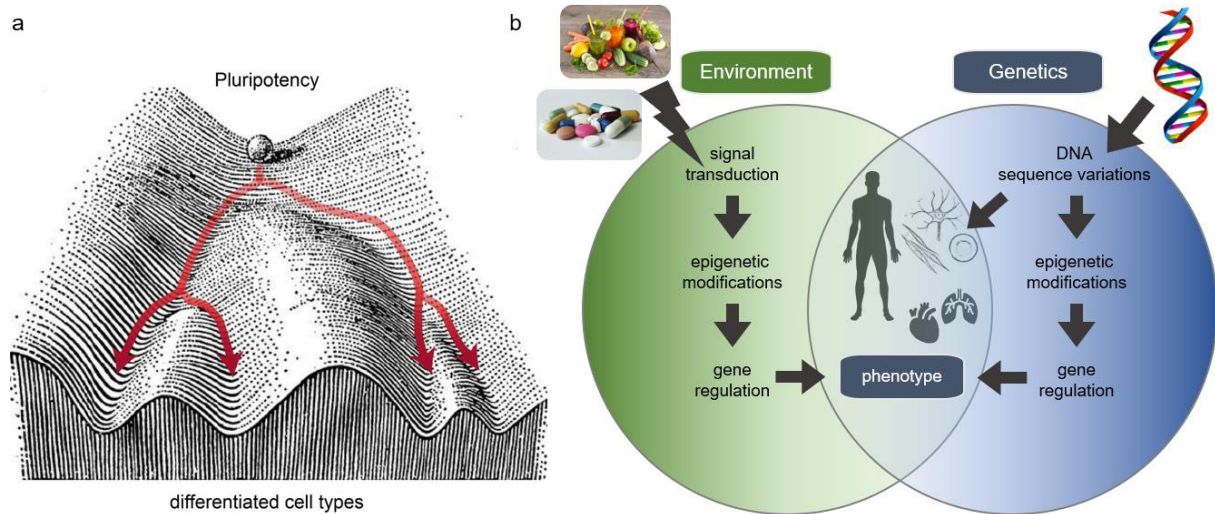
Figure 1: Epigenetic gene regulation integrates genetic information and environmental signals to produce specific phenotypes. a, The "Epigenetic landscape" as proposed by C. Waddington. Pluripotent cells are depicted as a marble at the top of a mountain ridge. The way down the slope represents cell differentiation during development. At every junction, cells further differentiate into progenitor cells of specific cell types until they reach the fully differentiated stage. The genetic material is not changed during this process. The model illustrates the unidirectionality of cell differentiation, and the epigenetic barriers preventing a transition between different cell type specific pathways (red arrows). Image adopted from Waddington, 1957. **b,** Gene regulation programs give rise to the formation of distinct phenotypes of cells, organs, and whole organisms. Variations in the genetic information can either influence the phenotypic outcome directly, or induce variations in the installation of epigenetic modification patterns that drive gene regulation. Environmental signals are integrated into gene regulation programs via signal transduction pathways during the whole life-cycle of an organism or cell.

Today, a common definition of the term epigenetics is "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (Russo et al., 1996). Importantly, DNA does not exist "naked", but is embedded into a context of structural proteins that undergo, like the DNA bases themselves, covalent and chemical modifications. The combination of DNA, RNA and structural proteins in this form is called chromatin. The modifications to chromatin can be stably inherited, thereby passing on genetic programs from one cell generation to another (Ciabrelli et al., 2017; Moussa et al., 2019; Stewart-Morgan et al., 2020; Jing Wang et al., 2009). The initial pattern of epigenetic modifications to chromatin is determined by the genetic material of a cell, which encodes for a basic configuration of chromatin regulators (**Figure 1b**). The expression and activity of these chromatin regulators, which regulate epigenetic modification patterns by a variety of mechanisms,

can undergo dynamic changes in response to environmental signals (**Figure 1b**). By linking the activity of epigenetic regulators to signal transduction pathways, cellular gene expression programs can be adapted to environmental changes (Amit et al., 2016; Bar-Sadeh et al., 2020; Cavalli & Heard, 2019; Mazzio & Soliman, 2012). To accurately control the maintenance and distribution of chromatin modifications, epigenetic regulation works in highly complex networks of chemical signals, sensory proteins and active enzymes. In the following paragraphs, the complexity of epigenetic gene regulation will be described to illustrate the importance of comprehensive studies on the underlying mechanisms.

1.1.2. Chromatin structure and the regulation of gene expression

Epigenetic regulation of gene expression is mediated by a variety of molecular and biochemical mechanisms: the post-translational modification of histone proteins, the architectural organization of DNA into 3D chromatin structures, the covalent chemical modification of DNA bases, and the presence of regulatory RNA structures. The function and significance of the different mechanisms will be explained in the following paragraphs, while taking into consideration their respective relevance for this work.

Each human cell contains about 3 billion DNA base pairs, encoding for around 20,000 protein coding genes, more than 9,000 long noncoding RNAs (lncRNAs) and almost 12,000 pseudogenes (Harrow et al., 2012). Other mammalian genomes are comparable in size (e.g. 2.7 billion base pairs in mice, according to mouse reference genome assembly GRCm39; https://www.ncbi.nlm.nih.gov/grc/mouse/data). Even eukaryotic microorganisms like the yeast *Saccharomyces cerevisiae* contain a genome of about 12 million base pairs (according to https://www.yeastgenome.org/). To store this large amount of DNA inside each nucleus and keep the information in an organized and retrievable way, the DNA strands are compacted multiple times by mechanisms that are conserved among all eukaryotes. The DNA double helix is wrapped around an octamer of structural proteins called histones (H2A, H2B, H3 and H4), forming DNA-protein complexes called nucleosomes, which are the smallest structural unit of chromatin (**Figure 2a** and **Figure 4**). Due to the phosphate backbone of DNA, the resulting chromatin is a negatively charged polymer with an innate electrostatic repulsion (Maeshima et al., 2020). This repulsion counteracting condensation of the individual nucleosomes is compensated by the positive charge of the histone proteins and the structural function of the H1 linker histones (Cutter & Hayes, 2015; Zhou et al.,

2020). The three-dimensional organization of nucleosomes along this polymer is regulated by chemical modifications of the chromatin, by local salt conditions and the presence of linker histone proteins (Maeshima et al., 2020; van Emmerik & van Ingen, 2019; Zhou et al., 2020). The open nucleosome "string" can be further condensed by stacking the nucleosomes on top of each other in a 30 nm fibre (**Figure 2a, b**) (F. Song et al., 2014; van Emmerik & van Ingen, 2019). This highly condensed ubiquitous form of chromatin is called heterochromatin, in contrast to the more accessible structure of euchromatin. Besides other intermediate forms of chromatin compaction that can form during interphase (Grigoryev, 2004), chromatin is further organized into highly condensed expanding domains during mitosis, peaking in the formation of chromosomes that are apportionable to two daughter cells (Antonin & Neumann, 2016).
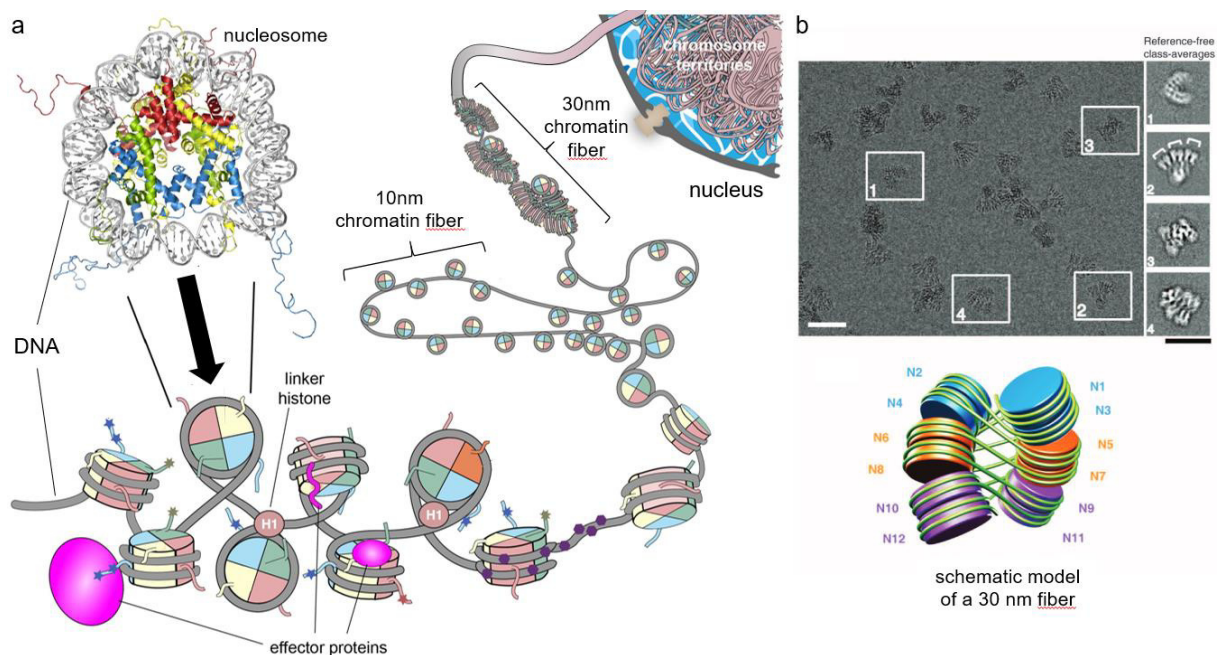
Figure 2: Structural organization of DNA in the form of chromatin. a, The organizational structure of a DNA strand wrapped around a heterocomplex of four histone proteins is called the nucleosome (top left). Nucleosomes can be organized in the nucleus in an accessible conformation that is called euchromatin, thereby forming a chromatin fiber that has a diameter of 10 nm. The chromatin strand can be further condensed into heterochromatin in the form of a highly-structured 30 nm chromatin fiber that is less accessible for histone- and DNA-interacting proteins. The degree of chromatin condensation depends on the length of the linker DNA, the presence of linker histones, of chromatin modifications (*asterisks and dots*) and chromatin effector proteins. The image was adopted from van Emmerik & van Ingen, 2019. **b,** Condensation of the beads-on-a-string 10 nm fibre into the 30 nm fibre. A Cryo-EM micrograph of 30 nm chromatin fibers reconstituted on 12 ×187 bp DNA is shown. Scale bars indicate 50 nm. *bottom:* schematic representation of the cryo-EM structure of a 30-nm chromatin fiber. The succession of nucleosomes along the DNA strand (*green*) is indicated with numbers N1-N12. The images in b were taken from F. Song et al., 2014.

The protein abundance in a cell is regulated at six major levels: at the level of transcription, of RNA processing, mRNA degradation, translation, protein modification or protein degradation. However, statistical analyses suggest that the main determinant of protein levels in the cell is the rate of transcription (J. J. Li et al., 2014; J. J. Li & Biggin, 2015). The organization of DNA sequences into hetero- and euchromatic domains is an early-acting and stable instrument to install cell-type specific protein expression profiles (Guelen et al., 2008). Focusing on the interphase manifestations of compacted chromatin, DNA can be packaged into heterochromatin to downregulate transcription either in a facultative, meaning reversible, or a constitutive and more stable way. DNA regions that are stored in the form of heterochromatin are not accessible to the transcription and replication machineries (Groth et al., 2007; B. Li et al., 2007).

To enable the expression of a specific gene, the respective genetic locus usually needs to adopt the more accessible form of chromatin called euchromatin (Bulut-Karslioglu et al., 2018). Transcriptionally active euchromatic regions require this more open conformation to enable transcription factor (TF) binding and access of RNA polymerase II. Euchromatin is characterized by variants of the histone core particles H2A (Papin et al., 2020; Raisner et al., 2005; Zilberman et al., 2008) and H3 (Ahmad & Henikoff, 2002; C. Jin et al., 2009; Loppin & Berger, 2020) and by the presence of specific histone posttranslational modifications (PTMs) that are related to actively transcribed chromatin (Ernst et al., 2011).

Facultative heterochromatin formation is a means to prevent the aberrant transcription of genes by condensation of the respective loci. Facultative heterochromatin is installed by specific chromatin remodelling complexes to fine-tune gene expression (Becker & Workman, 2013; Bornelöv et al., 2018; Narlikar et al., 2013). The installation of a facultative heterochromatin state is an epigenetic control mechanism to downregulate the expression of specific genes. It can be observed for example during dosage compensation and differentiation, and is characterized by the presence of repressive histone posttranslational modifications (HPTMs) like H3K27me3 (Bannister & Kouzarides, 2011; Gloria Mas & Di Croce, 2016; Peters et al., 2002). In contrast, the packing of non-coding and highly repetitive satellite sequences and transposons into constitutive heterochromatin is an essential mechanism to prevent illegitimate genome rearrangements and mutagenic recombination events in cells (Saksouk et al., 2015).

Both heterochromatic and euchromatic regions can form stable extensive domains in the nucleus by condensation into phase-separated droplets and by establishing functional long-range interactions (Dixon et al., 2016; Larson et al., 2017; Gloria Mas & Di Croce, 2016; Strom et al., 2017). The regulatory function of these chromatin domains will be discussed in section 1.1.3.

Transcriptional activity can further be modulated by interferences with regulatory elements that are associated with a gene. The chromatin state of a regulatory region can have effects on the transcriptional activity in *cis* i. e. on a nearby gene, or in *trans* i. e. on remotely associated genes. These regulatory *promoter* or *enhancer* regions can be several hundred up to a few thousand base pairs long, are often rich in cytosine-guanine dinucleotides (CpG) and can contain transcription factor (TF) binding sites and/or transcriptional start sites (TSSs) (Deaton & Bird, 2011; Stergachis et al., 2013). Depending on their epigenetic state, promoters and enhancers can switch between transcriptionally permissive or repressed states, thus directly regulating the expression of associated genes by a tuneable, switch-like mechanism (Deaton & Bird, 2011; H. K. Long et al., 2016). The interaction of regulatory elements with their target sequences is governed by the global three-dimensional chromatin architecture (Ibrahim & Mundlos, 2020; H. K. Long et al., 2016; Maeshima et al., 2020). Consequentially, the level of gene expression is dependent on the regulation of chromatin structure on a local as well as on a global level.

1.1.3. Domain organization and liquid-liquid phase separation (LLPS)

In addition to the regulation of gene expression by the packaging of DNA into hetero- or euchromatin, gene regulation is further influenced by a higher-order chromatin structure (**Figure 3a**) (Ibrahim & Mundlos, 2020). Hi-C chromatin interaction studies enable the comprehensive analysis of chromatin architecture on a global scale (Belton et al., 2012). These studies revealed the functional and spatial separation of chromatin in the mammalian nucleus into two different compartments called compartment A and B (Lieberman-Aiden et al., 2009; Rao et al., 2014). While genes in compartment A are rather transcriptionally active and depleted from the nuclear lamina, compartment B is characterized by facultative heterochromatin marks, late replication and association with the nuclear lamina (Rao et al., 2014). Genetic loci further associate into smaller subdomains (**Figure 3a**), which show a high frequency of intra-domain interactions, even though the genes can be sequentially distant (Ibrahim & Mundlos, 2020;

Maeshima et al., 2020). At the same time, these domains are relatively isolated from interactions with other chromatin domains. Due to their three-dimensional demarcation, these regions are called topologically associated domains (TADs). The formation of TADs in a developing organism is highly conserved, and it is suggested that they constitute an organizational unit of chromatin with functional implications for gene expression and replication (Dixon et al., 2016; Ibrahim & Mundlos, 2020). TAD formation is driven by loop extrusion of chromatin (**Figure 3a**), which is mediated by the architectural proteins cohesin and CTCF (Fudenberg et al., 2016, 2017). While it is commonly acknowledged that the dynamic three-dimensional organization of chromatin in loops, domains and compartments has gene regulatory function, the mechanisms connecting the 3D architecture with regulation of gene expression have not been conclusively described (Ibrahim & Mundlos, 2020; Maeshima et al., 2020).

In recent years, a novel mechanism for the higher order regulation of chromatin structure has come into focus. In 2017, two independent groups described the ability of the heterochromatin protein HP1α to form phase-separated liquid droplets in association with chromatin (Larson et al., 2017; Strom et al., 2017). Similar mechanisms of droplet formation have since been shown for other chromatin-associated proteins (**Figure 3b**). Liquid-liquid phase separation (LLPS) has been described as a characteristic of proteins containing intrinsically disordered regions (IDRs), the interaction of which enables the concentration of macromolecules into membrane-less substructures in the cell or nucleus (Nott et al., 2015; Owen & Shewmaker, 2019; Strom et al., 2017).

The LLPS model of heterochromatin formation is more differentiated than the previously established concept that heterochromatin compaction generally leads to a restricted access of regulatory proteins to the incorporated DNA and histones due to sterical hindrance. Phase-separation allows for the formation of plastic chromatin subdomains that can be entered by proteins interacting with the local chromatin structure or included structural proteins, while preventing the access of other proteins (B. A. Gibson et al., 2019; Sanulli et al., 2019; L. Wang et al., 2019).
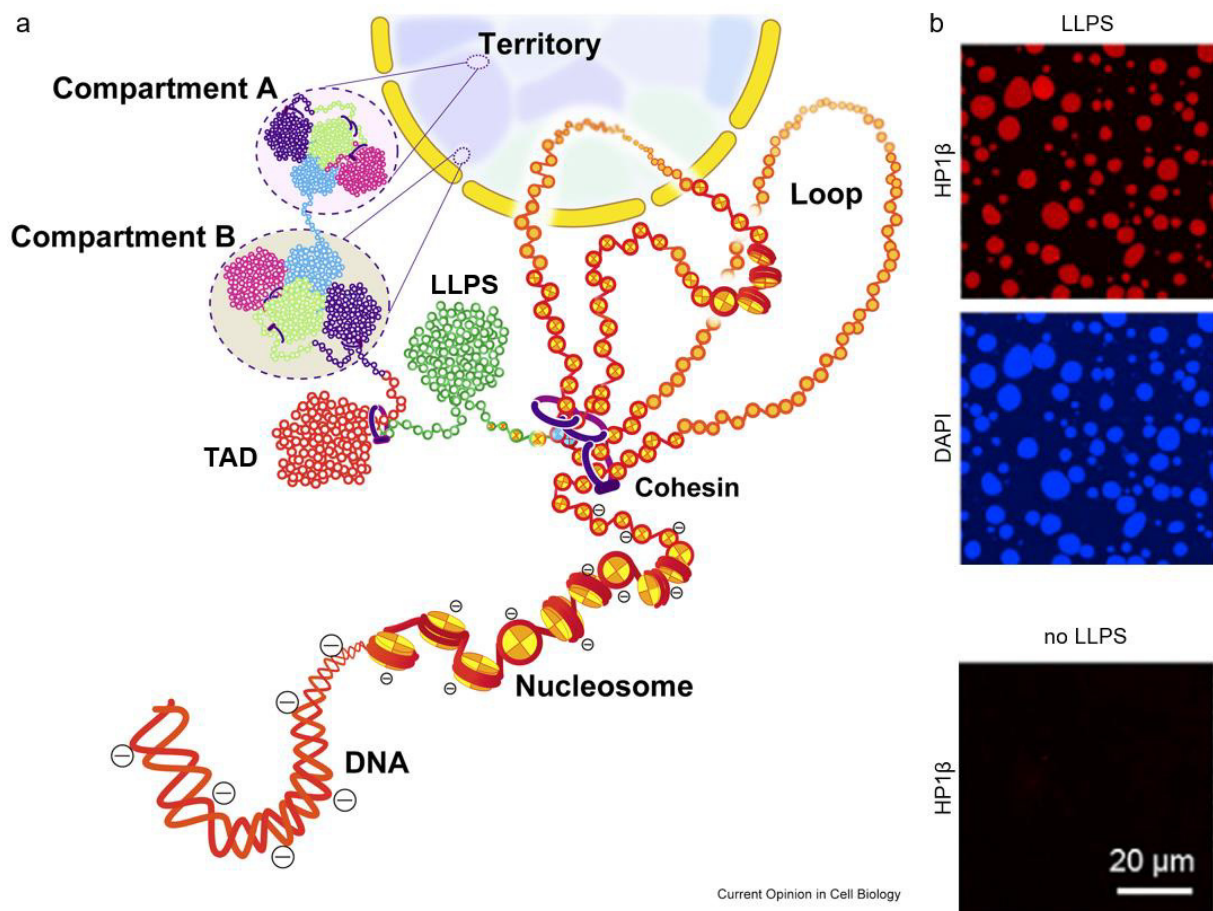
Figure 3: The levels of higher-order chromatin architecture. a, Simplified model of the hierarchical organization of chromatin in the nucleus. A chromatin string can protrude from a more condensed region by the formation of a chromatin loop. Loops are stabilized by the binding of cohesin, bringing together sequentially distant CTCF motifs at the cohesin binding site. Chromatin can further be organized into topologically associated domains (TAD) by constriction, or into liquid droplet domains by phase separation (LLPS). These domains are subunits of either chromatin compartment A, which is generally characterized by transcribed genes and active histone modifications, or of the mostly transcriptionally inactive compartment B. The differentiated compartments are spatially associated in chromosome territories that can be visualized microscopically. The image was adopted from (Maeshima et al., 2020). **b,** Visualization of droplet formation by LLPS. Synthetic nucleosomal arrays undergo LLPS depending on the presence of histone modifications and chromatin binding proteins *in vitro*. The example shows droplet formation by incubation of fluorescence-labelled HP1β with SUV39H1 and a H3K9me3 modified array (top) or an unmodified array (bottom). DAPI staining was used to visualize DNA. The image was taken from L. Wang et al., 2019.

LLPS is regulated by the presence of linker histone H1, the phosphorylation state of HP1 proteins and by the formation of multiprotein-complexes by epigenetic reader proteins, which is in turn dependent on local chromatin modifications (B. A. Gibson et al., 2019; Y. Guo et al., 2021; Larson et al., 2017; Owen & Shewmaker, 2019; L. Wang et al., 2019, 2020). Besides heterochromatin-related LLPS, the formation of euchromatic subdomains upon phase-separation of highly acetylated chromatin

domains by recruitment of bromodomain containing proteins has been described (B. A. Gibson et al., 2019). Phase-separated droplets spatially sequester and compact chromatin subdomains in a chromatin modification-dependent manner (**Figure 3b**), while allowing for further recruitment of repressive or activating chromatin factors. This droplet formation is not limited by linear proximity along the DNA strand, allowing for long-range interactions of genomic loci in LLPS compartments (Denholtz et al., 2013; B. A. Gibson et al., 2019; L. Wang et al., 2019). Thus, LLPS represents a mechanism for gene regulation that is mediated by the reorganization of chromatin architecture. Other gene regulatory mechanisms that are mediated by rearrangements of the global 3D chromatin structure have been described in the context of topologically associated domains (TADs) and the distal interactions of polycomb domains (Ibrahim & Mundlos, 2020; Gloria Mas & Di Croce, 2016). All these structural mechanisms work together to regulate gene expression, in a manner that remains to be fully described.

1.1.4. Gene regulation by histone posttranslational modifications (HPTMs)

The four histone core proteins, H2A, H2B, H3 and H4, are small basic proteins that are highly conserved in structure and function among eukaryotes. By the regulated assembly of DNA and core histones to nucleosomes (**Figure 4a**) and the further condensation of those nucleosomes to higher-order chromatin (van Emmerik & van Ingen, 2019; Zhou et al., 2020), the accessibility of a genomic DNA sequence for the transcription and replication machineries can be altered (Deaton & Bird, 2011; Saksouk et al., 2015). To enable the dynamic assembly and disassembly of nucleosome complexes, the structure has evolved to be only meta-stable (Cutter & Hayes, 2015). This means that nucleosomes can switch between a stabilized chemical state, in which the histone octamer is tightly associated with the wrapped DNA, and a modified state in which the association is more relaxed, allowing for the sliding of the histone octamer along the DNA, or even complete eviction of the nucleosome (Cosgrove et al., 2004). The stability of the nucleosome assembly can be regulated by posttranslational modification (PTM) of the histone proteins (Zentner & Henikoff, 2013). Already in 1964, Allfrey et al. showed that the acetylation of histones is an ubiquitous modification that reduces the inhibitory effect of histones on RNA synthesis (Allfrey et al., 1964). The addition of acetyl residues to positively charged lysines has a neutralizing effect on the histone, weakening the interaction of the histone with nucleosomal DNA and facilitating the access of the transcription machinery (Zentner & Henikoff, 2013).

HPTMs have regulatory effects on gene expression that exceed the mere physical regulation of the accessibility of the underlying nucleosomal DNA. In 1999, Strahl and Allis functionally correlated genomic sites in *Tetrahymena thermophila* exhibiting a specific pattern of histone acetylation in combination with methylation of lysine 4 of histone 3 (H3K4me) with sites of transcriptional activity (Strahl et al., 1999). The presence of these additional methyl-residues at the far ends of the N-terminal tail of H3 (**Figure 4a**) appeared to be essential but could not be explained with a biophysical effect on the association of the DNA-histone core complex. Their findings suggested that certain combinations of histone modifications could define distinct types of chromatin that are accessible to the transcription machinery, thus marking regions that contain active genes.

Indeed, in 2004 Schübeler et al. found a distinct binary pattern of HPTMs enriched in euchromatic regions of a *Drosophila* cell line. Transcriptionally active regions were hyperacetylated on H3 and H4 (H3ac and H4ac) and hypermethylated at lysine 4 and 79 of H3 (H3K3me2/3 and H3K79me2/3), while regions with the inverse modification pattern were transcriptionally inactive (Schübeler et al., 2004). While the acetylation of histones is a PTM invariably linked to active chromatin states, lysine methylation can either have activating or repressive effects, depending on the localization of the modified amino acid on the histone tail (**Figure 4b**). In addition to lysine residues, some arginines of the four histone tails can also be methylated, with variable effects on chromatin regulation (Fulton et al., 2018). Histone tail phosphorylation can play a role in the regulation of processes such as DNA damage repair, transcriptional regulation and chromatin compaction (Rossetto et al., 2012).

The development of massive parallel sequencing approaches has strongly advanced the progress in understanding the complexity of HPTMs, by enabling the generation of genome-wide correlation profiles of histone modifications (Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007). This approach has led to the identification of 15 different chromatin states of transcriptional regulatory elements in human cell lines (Ernst et al., 2011) (**Figure 5**). In accordance with the early findings of Strahl and Schübeler, Ernst et al. found that active promoter regions and strong enhancers were defined by the combination of H3K4 methylation and pan-H3 acetylation, and that the observation frequency of active histone marks correlated with the strength of transcriptional activation by those regulatory elements.
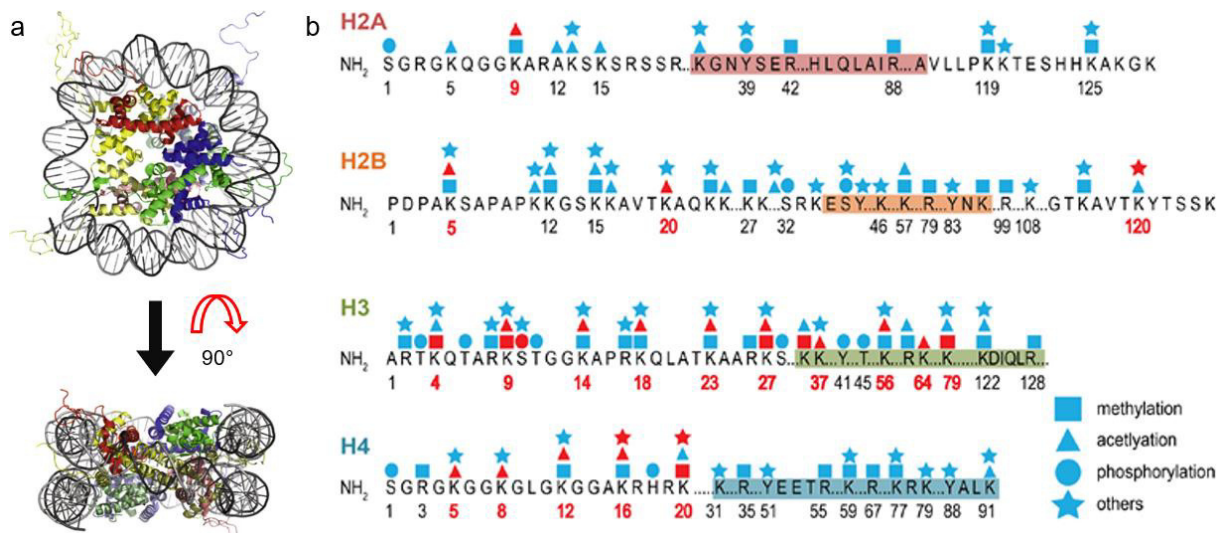
Figure 4: Nucleosome structure and histone posttranslational modifications. a, Structural model of the nucleosome core. About 1.7 superhelical turns of DNA (grey) are wrapped around the histone octamer consisting of two copies of each histone protein. N-terminal histone tails protrude from the core structure. *H2A*, green, *H2B*, blue, *H3*, yellow, *H4*, red. Image adopted from Cutter & Hayes, 2015. **b,** Selective overview of the most prevalent histone posttranslational modifications. Highlighted amino acids indicate globular histone regions, not highlighted amino acids constitute the histone tail. Red modifications have already successfully been introduced in epigenetic editing experiments by several groups. The image was taken from Tekel & Haynes, 2017.

In reference to earlier profiling studies (Guenther et al., 2007; Mikkelsen et al., 2007), they identified regions of high H3K36me3 and H4K20me1, but without methylation of H3K4 or histone acetylation, as actively transcribed genomic regions. Indeed, H3K36me3 is enriched at the 3' ends of transcribed genes (Barski et al., 2007) and is crucial for the transition from transcription initiation to transcription elongation at genes marked by active HPTMs (Guenther et al., 2007; J. Li et al., 2002; Strahl et al., 2002). During transcription, RNA Polymerase II is phosphorylated at Ser2 in its CTD, which recruits the H3K36 methyltransferase Set2 (J. H. Kim et al., 2016). Trimethylation of H3K36me3 in gene bodies has two effects preventing irregular transcription initiation in the gene body: the recruitment of histone deacetylases (HDACs) and the recruitment of DNMT3B (Neri et al., 2017). The open conformation of actively transcribed genes renders them accessible for cryptic transcriptional initiation, which is prevented by deacetylation and DNA methylation. Besides its role in protecting actively transcribed genes, H3K36me3 also regulates heterochromatin formation in pericentromeric heterochromatin (Chantalat et al., 2011; Dhayalan et al., 2010; Weinberg et al., 2019) and has several roles in DNA repair (Z. Sun et al., 2020), therefore generally regulating genomic stability.

State	CTCF	H3K27me3	H3K36me3	H4K20me1	H3K4me1	H3K4me2	H3K4me3	H3K27ac	H3K9ac	WCE	Candidate state annotation
1	16	2	2	6	17	93	99	96	98	2	Active Promoter
2	12	2	6	9	53	94	95	14	44	1	Weak Promoter
3	13	72	0	9	48	78	49	1	10	1	Inactive/poised Promoter
4	11	1	15	11	96	99	75	97	86	4	Strong enhancer
5	5	0	10	3	88	57	5	84	25	1	Strong enhancer
6	7	1	1	3	58	75	8	6	5	1	Weak/poised enhancer
7	2	1	2	1	56	3	0	6	2	1	Weak/poised enhancer
8	92	2	1	3	6	3	0	0	1	1	Insulator
9	5	0	43	43	37	11	2	9	4	1	Transcriptional transition
10	1	0	47	3	0	0	0	0	0	1	Transcriptional elongation
11	0	0	3	2	0	0	0	0	0	0	Weak transcribed
12	1	27	0	2	0	0	0	0	0	0	Polycomb-repressed
13	0	0	0	0	0	0	0	0	0	0	Heterochrom; low signal
14	22	28	19	41	6	5	26	5	13	37	Repetitive/CNV
15	85	85	91	88	76	77	91	73	85	78	Repetitive/CNV

Chromatin States — Chromatin Mark Observation Frequency (%)

Figure 5: Distinct chromatin states can be derived from epigenome correlation analysis in nine human cell types. Endogenous regulatory elements in the human genome (listed to the right) are associated with characteristic patterns of HPTMs depending on their biological function. Vice versa, the observation of a distinct combination of HPTMs allows for conclusions about the regulatory function of a genomic locus. The image was adopted from Ernst et al., 2011.

Ernst et al. found that promoter regions devoid of the earlier described active chromatin marks were statistically correlated with trimethylation of lysine 27 of H3 (H3K27me3) (**Figure 5**). H3K27me3 is a mark set by the Polycomb repressive complex 2 (PRC2) and was found enriched in two chromatin states, in repressed, so-called heterochromatic regions, and at bivalent promoters (Ernst et al., 2011). Bivalent promoters are characterized by the parallel activity of the MLL2/COMPASS complex setting H3K4me3, and the PRC2 complex setting H3K27me3 (Piunti & Shilatifard, 2016). The complexes are most likely either active at different tails of the same nucleosome or on neighbouring nucleosomes, as PRC2 is inhibited by the presence of H3K4me3 in *cis* (Schmitges et al., 2011). The co-occurrence of active and repressive chromatin marks results in an actively transcribed regulatory element that is ready to be shut down in a switch-like behaviour (Bernstein et al., 2006). Although intriguing and often-quoted, this simplified model of bivalency has also been challenged (D. Hu et al., 2013; Piunti & Shilatifard, 2016; Voigt et al., 2013). H3K27me3 is further important for the formation of facultative heterochromatin and the resulting stable silencing of genes during development (Gloria Mas & Di Croce, 2016; Wiles & Selker,

2017). The deposition of H3K27me3 is regulated by the concerted action of two polycomb complexes, PRC1 and PRC2 (Blackledge et al., 2014). The PRC complexes combine reading and writing function for H3K27me3 in a self-enhancing manner, enabling the maintenance and spreading of repressive chromatin domains (Blackledge et al., 2014; J.-R. Yu et al., 2019). This self-sustaining mechanism of PRC silencing enables inheritance of an epigenetically silenced state of genes over multiple cell generations (Ciabrelli et al., 2017).

While H3K27me3 is a histone mark correlated with the formation of facultative heterochromatin, condensation of genomic regions into stably silenced constitutive heterochromatin is dependent on deposition of H3K9me3 (Saksouk et al., 2015). Methylation of H3K9 can be performed by a variety of lysine methyltransferases in mammals (Herz et al., 2013). Consistently, H3K9 methylation is a mark interfering in multiple processes. H3K9 trimethylation by SETDB1 is able to actively silence transcription via the formation of facultative heterochromatin (Schultz et al., 2002). The formation of pericentromeric heterochromatin to secure genome integrity (Saksouk et al., 2015) is dependent on H3K9 methylation, as well as the determination of large and heritably inactive chromatin domains during lineage commitment (Guelen et al., 2008; Herz et al., 2013; Wen et al., 2009).

HPTMs are a vital for the installation and adaption of gene expression programs in cells, due to their manifold effects on nucleosome stability and turnover, their influence on local chromatin conformation changes and their ability to recruit specific epigenetic effector complexes (Zentner & Henikoff, 2013). Importantly, HPTMs also regulate cellular processes like DNA replication (Goren et al., 2008; F. Li et al., 2011; Vogelauer et al., 2002), DNA damage response (Faucher & Wellinger, 2010; Masumoto et al., 2005; van Attikum & Gasser, 2009) and mitosis (Valls et al., 2005; F. Wang & Higgins, 2013). However, as the focus of this work was the investigation of regulatory interactions inside of the epigenetic network controlling transcriptional regulation, a detailed essay of these processes would go beyond the scope of this study.

1.1.5. Gene regulation by DNA modifications

Another chemical modification of chromatin that interferes with the regulation of gene expression is the methylation of DNA bases (Jaenisch & Bird, 2003). Besides other, far less abundant DNA modifications (Deniz et al., 2019), DNA methylation usually occurs in form of the addition of a methyl group to the carbon-5 of the pyrimidine ring of a cytosine base, thereby producing 5-methyl-cytosine (5mC) (Hermann et al., 2004). In mammals, about 70% of the genome is 5mC methylated (Ehrlich et al., 1982). The remaining 30% are mostly composed of long stretches of CpG dinucleotides rich regions, so-called CpG islands (CGI), that are hypomethylated in somatic cells (Deaton & Bird, 2011). CGIs contain transcription factor binding- and transcriptional start sites and function as promoter regions. The by-default hypomethylated CGI promoters are usually associated with highly transcribed genes, and hypermethylation of high-CpG-density CGIs is correlated with stable and heritable silencing of the associated genes (Deaton & Bird, 2011; Messerschmidt et al., 2014). However, DNA methylation alone is not sufficient for silencing of gene expression, the repressive effect is permitted either by a change of transcription factor binding to the promoter (Kribelbauer et al., 2017; Yin et al., 2017) or by recruitment of repressive epigenetic effector proteins like methyl-CpG-binding domain (MBD) containing proteins (Q. Du et al., 2015).

Methylation of the cytosine-5 residue on unmethylated CpGs is introduced by the *de novo* DNA-methyltransferases DNMT3A, DNMT3B and DNMT3L, while maintenance of the DNA methylation state after replication is secured by DNMT1 (Bostick et al., 2007; Gowher et al., 2005; Hermann et al., 2004). The mechanism of function of all DNA methyltransferases is discussed in more detail in section 1.3.1.

On a global level, DNA methylation is important for the stable repression of endogenous retrotransposons, pericentromeric repeats and for X-inactivation during dosage compensation (Messerschmidt et al., 2014). Global CpG methylation is reset in a major event during embryonic development before it is reinstalled following a specific plan to determine cell-type specific expression profiles that drive differentiation (**Figure 6a**) (W. Reik et al., 2001; Xia & Xie, 2020). Only a few genomic regions called *imprinted genes* evade this global reprogramming event and retain the parental methylation status (Wolf Reik & Walter, 2001). During the somatic cell cycle, DNA methylation landscapes are rather stable. Passive dilution of DNA-methylation marks throughout somatic cell division is prevented by the maintenance DNA

methyltransferase DNMT1 (J. Song et al., 2012). For the active reprogramming of a 5mC-repressed genomic locus, the recruitment of a protein from the ten eleven translocation (TET) family is necessary (**Figure 6b**) (Kohli & Zhang, 2013). TET enzymes remove 5mC by a stepwise oxidation process that results in the removal of the oxidized base by the cellular base excision repair mechanisms (Rasmussen & Helin, 2016).
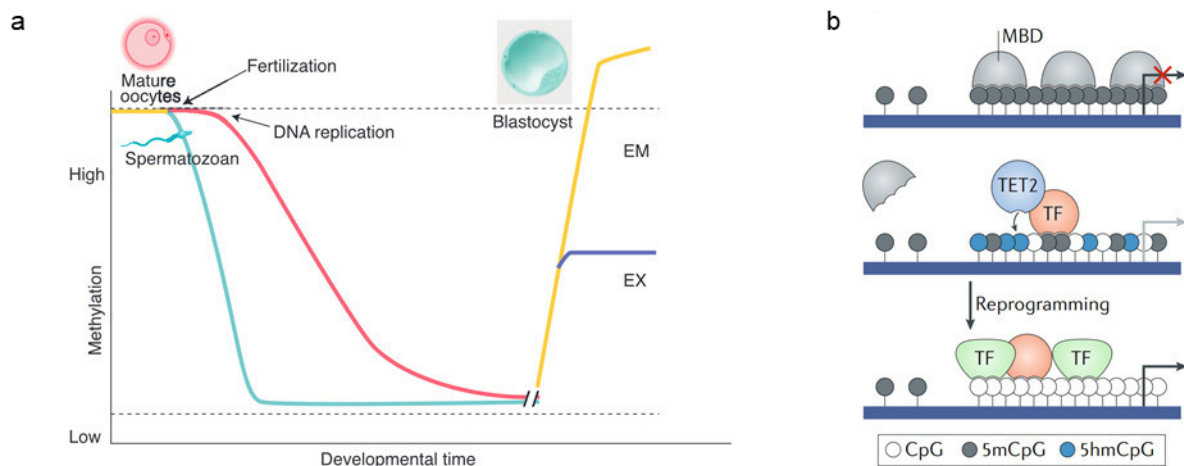
Figure 6: Epigenetic reprogramming events in mammalian cells. a, Sperm cells and mature oocytes carry distinct DNA methylation patterns inherited during gametogenesis. Immediately after fertilization, the paternal genome is actively demethylated. The maternal genome is passively demethylated during DNA replication cycles. Epigenetic reprogramming occurs around the blastocyst stage, when lineage specific remethylation of the genome is initiated. The methylation status of imprinted genes (dotted lines) is not changed during this global reprogramming event. EM = embryonic lineages; EX = extraembryonic lineages. Image taken from Reik et al., 2001. **b,** DNA methylation at CpG promoters is recognized by repressive factors like methyl-CpG binding domain containing proteins (MBD), which inhibit the transcription of genes regulated by the methylated promoter (top). During transcriptional reprogramming events, TET enzymes are recruited to the methylated promoter (centre), leading to its demethylation and enabling the binding of TFs and activating transcription (bottom). TF = transcription factor. Image taken from Greenberg & Bourc'his, 2019.

CpG methylation itself can have mutagenic capacity. The 5mC base is easily deaminated to thymine, which is not recognized as a mutation of the DNA during replication, resulting in a C to T transition mutation (Duncan & Miller, 1980). At the same time, CpG methylation is crucial for genome stability, with loss of function mutations in DNA methyltransferases resulting in aneuploidy, polyploidy, chromosomal breaks and fusions and reactivation of retroviral elements (Chédin, 2011; Dodge et al., 2005; R. S. Hansen et al., 1999; Karpf & Matsui, 2005). Consequently, DNA methylation needs to be tightly regulated to secure genome integrity and phenotypic stability.

1.1.6. Regulatory RNAs

Besides the well-described functions of RNA as a transmitter of information during transcription and translation, additional functions of RNA as a structural component of chromatin receive increasing interest in studies of epigenetic regulation. Two chromatin regulatory mechanisms that involve RNA structures and are of particular interest for the presented work will be described in the following paragraphs.

Noncoding RNAs

The common view of the mammalian genome as an assembly of ~20,000 protein coding genes embedded in a much larger environment of transcriptionally inert, often repetitive DNA that lacks any peculiar function for the survival of cells, has been revised for many years now. Diverse high-throughput approaches have revealed that messenger RNA (mRNA) transcripts make up only 3-7% of the cellular RNA mass, while 63-85% of the whole genome can be transcribed into any form of RNA (Dykes & Emanueli, 2017). Most of the RNA mass is made up from ribosomal RNA (rRNA) and transfer RNA (tRNA) that are needed for the process of translation, but there is an increasing number of highly conserved transcripts emerging, that have neither mRNA nor rRNA or tRNA function and are called "noncoding RNAs" (ncRNAs) (Dykes & Emanueli, 2017; Guttman & Rinn, 2012). The importance of long noncoding RNAs (lncRNAs) emerged in the course of genome-wide association studies (GWAS), which found that more than 90% of the SNPs that are associated with disease development lie outside of coding regions (Hindorff et al., 2009; Luo & Chen, 2020; Minotti et al., 2018). lncRNAs can regulate chromatin function and gene expression in various ways and are an integral part of the epigenetic machinery. lncRNAs can act as scaffolds for the recruitment of specific chromatin remodelling complexes to target loci (Han & Chang, 2015; Jacob et al., 2013; Tsai et al., 2010). The process of transcription of a noncoding RNA alone can reorganize chromatin architecture in *cis* or *trans* (Mishra & Kanduri, 2019; Pisignano et al., 2019). lncRNAs can regulate gene expression by directing repressive or activating epigenetic complexes to target genes (Dykes & Emanueli, 2017; O'Leary et al., 2017). The lncRNAs MALAT1 and Saf have been shown to regulate alternative splicing of pre-mRNAs of different oncogenes, thereby increasing tumorigenesis of cells (Änkö & Neugebauer, 2010; Tripathi et al., 2013; Villamizar et al., 2016). Furthermore, lncRNAs can regulate alternative splicing by targeting epigenetic complexes to the transcribed gene, thereby changing the local

chromatin environment and favouring exclusion or inclusion of the respective exons (Gonzalez et al., 2015). Finally, lncRNAs have been shown to play a role in LLPS-driven chromatin subdomain formation (Grosch et al., 2020). Consequently, the association of epigenetic complexes with three-dimensional lncRNA structures is a prevalent mechanism for the targeting and regulation of epigenetic activities at different times during transcription.

R-loops

R-loops are three-stranded nucleic acid structures consisting of a single-stranded RNA invading a DNA duplex (**Figure 7a**). R-loops are generated during transcription at characteristic sites in the mammalian genome (Sanz et al., 2016). R-loops can form at GC-rich regions where one DNA strand contains significantly more guanine (G) bases than the other, a characteristic that is called *GC-skew* (**Figure 7a, b**). If transcription runs along the cytosine (C) rich DNA template-strand, the newly formed G-rich mRNA can re-anneal with the DNA template and force the non-template strand into a single-strand conformation (Ginno et al., 2012). The DNA:RNA heterohybrid structure of the R-loop is stabilized by the establishment of a G-quadruplex (GQ) structure in the non-template DNA strand (C. Y. Lee et al., 2020).

45% of all human genes show a GC skew in their 5´ regions, between -500 bp upstream and +1500 bp downstream of the gene starting site (**Figure 7b**) (Ginno et al., 2012; Sanz et al., 2016). The GC skew is a characteristic of strong, unmethylated CpG island (CGI) promoters of human housekeeping genes (Hartono et al., 2015), and R-loops protect from *de novo* DNA methylation at these regions (**Figure 7c**), although the exact mechanism remains to be defined (Ginno et al., 2012, 2013; Sanz et al., 2016). Long after the discovery of co-transcriptional DNA:RNA hybrid formation in prokaryotes (Masukata & Tomizawa, 1990), the first regulatory mechanism including R-loop formation in mammalian cells was discovered by Daniels and Lieber, who proposed a role for DNA:RNA hybrid formation in immunoglobulin class switch recombination (Daniels & Lieber, 1995; Ribeiro de Almeida et al., 2018; K. Yu et al., 2003). Generally, R-loop formation is not only a side product of transcriptional activity (Niehrs & Luke, 2020; Sanz et al., 2016). The formation of cotranscriptional R-loops and GQs has been observed to further increase the rate of transcription (C. Y. Lee et al., 2020), a mechanism that can also explain transcriptional activation by antisense transcription (Boque-Sastre et al., 2015). Recently, further regulatory mechanisms

linking the stability of cotranscriptional R-loops with the activity of the transcription machinery have been described in *Arabidopsis thaliana* (C. Xu et al., 2021).
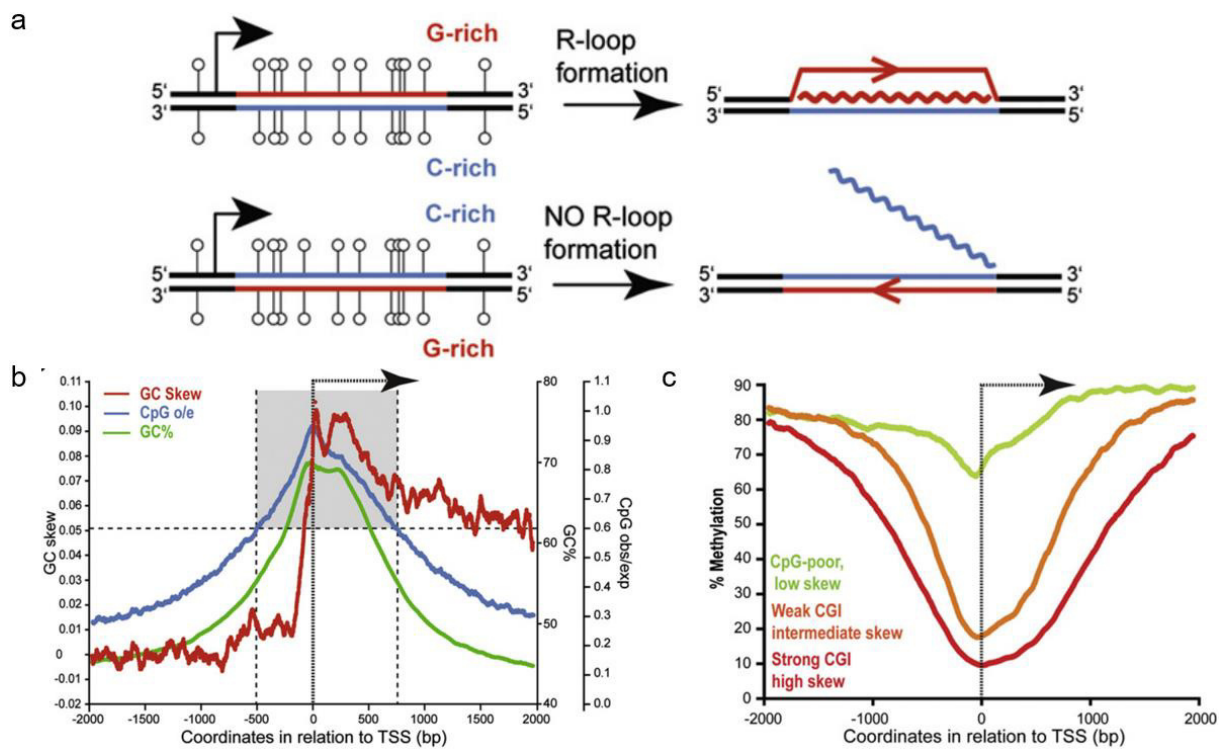
Figure 7: R-loops form over unmethylated CpG promoters. a, R-loops form at regions featuring a GC-skew during transcription of the C-rich DNA strand. The synthesized mRNA (wavy line) re-anneals with the DNA template strand to form an R-loop (top). If the G-rich DNA strand is transcribed (bottom), no R-loop can be formed. Open circles represent unmethylated CpGs. **b,** R-loop prone CpG promoters are characterized by high GC content (*GC%*), an increase in observed CpG frequency (*CpG o/e*) around ±500 bp of the TSS, and a sharp increase of *GC skew* starting at the TSS. **c,** CGI promoter strength correlates with GC skew and the amount of CpG methylation. All graphs were taken from Ginno et al., 2012.

The occurrence of R-loops at a genomic locus can be differentiated into regulatory and unscheduled R-loop formation. While the untimely formation of R-loop structures poses a threat to genome integrity (Crossley et al., 2019), their controlled formation has diverse regulatory functions on chromatin structure, transcription and the maintenance of genome integrity (Crossley et al., 2019; Niehrs & Luke, 2020). Besides the inhibition of DNA methylation, the formation of R-loops at CpG promoters can be related to several other characteristic epigenetic states. CGI promoters in different human cells that are prone to R-loop formation show an increase in H3K4 di- and trimethylation, elevated acetylation levels and low H3K27me3 levels at their TSS (Ginno et al., 2013; Sanz et al., 2016) and are linked to highly expressed genes. Consistently, there is binding of activating epigenetic complexes and depletion of repressive complexes at these sites (Sanz et al., 2016). Like the formation of R-loops, the presence of these

epigenetic modifications correlates with the extent of GC skew, R-loop potential and CpG enrichment of the CGIs (Sanz et al., 2016). This correlation depicts the tight network that links the formation of R-loops at CGI promoters with the regulation of the recruitment and activity of chromatin modifying complexes (Niehrs & Luke, 2020).

1.2. Epigenetic effector proteins

All modifications to the epigenome on histones or DNA are dynamic and can actively be changed in response to extracellular cues, as well as to signals from the local chromatin environment. This flexibility is enabled by the interaction of different chromatin effector proteins that can either set or erase those marks directly, or specifically recognize existent marks to induce the recruitment of other epigenetic modifiers. To illustrate a picture of the complex interactions controlling chromatin modification patterns genome-wide, the following paragraphs will exemplify prominent protagonists of epigenetic regulation.

1.2.1. "Writers" of chromatin modifications

Epigenetic modifications on histones are added post-translationally to their N-terminal tails. The first described "writer" of histone modifications was a histone acetyltransferase (HAT) in *Tetrahymena* (J. E. Brownell et al., 1996; James E. Brownell & Allis, 1995). The group illustrated a model in which histone acetylation, which had already been described to be associated with active genes, was actively transferred to the amino-terminal tails of histones by histone acetyltransferases in complex with the transcription machinery. Although the model was still crude at that time, it prospectively provided a mechanistic link between the active deposition of a HPTM and the process of local changes in gene expression, that was consequentially validated in later studies (Carrozza et al., 2003; Galarneau et al., 2000; Kuo et al., 1997).

The discovery of epigenetic effector proteins that write repressive chromatin modifications was more challenging. Using reverse genetics, several groups screened for genes that were essential for the suppression of position effect variegation (PEV) in *Drosophila melanogaster* and *S. Pombe* (Allshire et al., 1995; Reuter & Spierer, 1992; Wallrath, 1998), and classified the identified genes into the so-called *Su(var)*-family. However, this group of genes united only by their ability to install a repressive chromatin state was mechanistically diverse, containing histone deacetylases, protein

phosphatases and other uncharacterized heterochromatin associated genes. Rea et al. managed to isolate and characterize the human and mouse homologues of *Su(var)*3-9, the proteins SUV39H1 and Suv39h1, respectively (Rea et al., 2000) and described their SET-domain dependent methyltransferase activity on lysine 9 of histone H3. They proposed a model, in which the methylation of H3K9 induces the local formation of heterochromatin that was supported by findings of another group (Melcher et al., 2000) and describes a mechanism for the active silencing of genes following H3K9 methylation by compaction into higher order chromatin. At this point, other SET domain containing proteins like the Enhancer Of Zeste Homolog 2 (EZH2) and the Lysine Methyltransferase 2A (KMT2A, MLL1 or HRX) were known in mammals, but the group could not observe any methyltransferase activity of either protein on free histones. This observation was clarified when EZH2 (Czermin et al., 2002; Kuzmichev et al., 2002; Müller et al., 2002) and MLL1 (Dou et al., 2006) were isolated in complex with their associated proteins, upon which they showed *in vitro* catalytic activity on H3K27 and H3K4, respectively. The SET-domain containing protein family emerged as an extensive group of enzymes able to methylate histone- and non-histone proteins, with activating (e.g. MLL1) as well as repressive (e.g. EZH2) function on gene expression, depending on their specific targets (Herz et al., 2013).

Besides epigenetic writers setting histone acetylation via HAT domains or methylation marks via SET domains, there also exists a subset of mammalian protein kinases, which among other targets, also phosphorylate histone substrates. The ataxia telangiectasia mutated kinase (ATM) phosphorylates the histone variant H2A at serine 139, producing the histone variant γH2AX as an early event during DNA damage response (DDR) (Burma et al., 2001). Besides its role in energy homeostasis (Hardie, 2007), the adenosine monophosphate–activated protein kinase (AMPK) directly activates genes in response to the AMPK-signalling pathway by phosphorylating serine 36 of histone H2B (Bungard et al., 2010). There are multiple studies linking histone tail phosphorylation to histone acetylation and gene activation (Rossetto et al., 2012), but kinases have not been in the focus as epigenetic writer proteins, supposedly due to the high number of other signalling and maintenance functions regulated by kinases in the cell.

Another group of proteins adding PTMs to histones are the RING domain containing E3 ubiquitin ligases. While posttranslational poly-ubiquitination of proteins marks them

for proteasomal degradation (Glickman & Ciechanover, 2002), the addition of a single ubiquitin residue serves as a signal for protein-protein interaction. Several proteins that exhibit E3 ligase activity on histones and thus act as essential epigenetic effector proteins have been described. The Ubiquitin Like With PHD And RING Finger Domains 1 protein (UHRF1) flags hemimethylated chromatin regions for maintenance methylation by DNMT1 after replication (W. Qin et al., 2015). The RING1A/B subdomains of Polycomb repressive complex 1 (PRC1) ubiquitylate H2A to recruit PRC2, subsequently leading to the formation of a polycomb silencing complex (Blackledge et al., 2014). Finally, the E3 ubiquitin protein ligase BRE1 monoubiquitylates H2B in *Saccharomyces cerevisiae*, to recruit the Complex Of Proteins Associated With Set1 (COMPASS) and induce methylation of H3K4 (Jung Shin Lee et al., 2007). Thus, ubiquitination of histones can constitute an important signal for the installation and maintenance of chromatin states. However, ubiquitination will only play a marginal role in the experimental section of this work.

In addition to writers of HPTMs, DNA methylation writers edit the epigenome by transferring a methyl residue to the fifth carbon (C5) of cytosine bases of the genomic DNA. The DNA methyltransferases DNMT1, DNMT2, DNMT3A and DNMT3B all share a conserved catalytic domain containing six signature motifs important for their activity on C5 (Robertson, 2001; Xie et al., 1999), while the related protein DNMT3L contains only 4 of those motifs and a truncated N-terminal domain and lacks methyltransferase activity. However, this protein strongly stimulates the activity of DNMT3A and B by formation of multiprotein complexes (Aapola et al., 2000; Gowher et al., 2005; Jia et al., 2007). The different DNA methyltransferases hold distinct roles in the distribution of C5 methylation (5mC) in the genome. In all vertebrates, cytosine is mainly methylated in the context of CpG dinucleotides (Deaton & Bird, 2011), with long stretches of CpG (CGIs) mostly lacking 5mC and acting as active promoters, while "single" CpGs and CpGs in endogenous transposable elements, LINE- and LTR promoter regions are generally hypermethylated to prevent aberrant transcription (Messerschmidt et al., 2014). The cell-type specific patterns of DNA methylation are set by DNMT3A and B after a wave of global demethylation during embryonic development and are maintained by DNMT1 throughout somatic cell replication (Greenberg & Bourc'his, 2019; W. Reik et al., 2001). While the orchestrated de-novo DNA methylation by DNMT3A/B during embryogenesis and gametogenesis is crucial

for the development of a healthy organism, the DNMT3 proteins are hardly expressed in somatic tissues (Okano et al., 1998). However, mutation of the de-novo methyltransferases in somatic tissues is implicated in the development of cancer (M. S. Kim et al., 2013; Roll et al., 2008). DNA methyltransferases covalently transfer a methyl group to the C5 of a cytosine base in a three-step mechanism (S. S. Smith et al., 1992) and 5mC can actively be removed only in an even more complicated oxidation process catalysed by ten eleven translocation (TET) enzymes (Rasmussen & Helin, 2016). Generally, the writing of DNA methylation rather functions to reinforce the constant and stable silencing of repetitive sequences and genes that are dispensable for the active cellular program than for the dynamic regulation of gene expression upon external signals.

In conclusion, the writing of epigenetic modifications is conducted by a variety of effector proteins which can share common catalytic domains. Those enzymes edit their targets specifically, depending on the pre-existing chromatin context and their interaction with other proteins in multiprotein complexes. This complex formation is often guided by proteins that "read" the existing chromatin status. These chromatin readers will be explained in the following paragraph.

1.2.2. Reading chromatin modifications

The "reading" of HPTMs by protein domains is an important component of epigenetic regulation. The coupling of reading and writing of chromatin marks is often used as a mechanism for the maintenance and spreading of specific modifications. Ligand-specific reading domains for the mono-, di- and trimethylated states of lysines, for arginine methylation, lysine acetylation and serine/threonine phosphorylation on H3 have been discovered (**Figure 8a**). Reading domain and catalytic domain are often found on the same protein, with some epigenetic writers even featuring multiple reading domains (Torres & Fujimori, 2015). The catalytic activity of the protein can be influenced by the reading domain by different mechanisms (**Figure 8b**). The DNA methyltransferase DNMT3A contains an ATRX-DNMT3-DNMT3L (ADD) domain that recognizes unmodified H3K4, and a PWWP domain that reads trimethylation of H3K36 (Dhayalan et al., 2010; Yingying Zhang et al., 2010). DNMT3A is targeted to chromatin regions marked with H3K36 trimethylation (**Figure 8b**, top). In the presence of unmodified H3K4, but not in the presence of trimethylated H3K4, binding of the ADD releases DNMT3A from its autoinhibitory form (X. Guo et al., 2015) (**Figure 8b**,

bottom). The combination of both mechanisms leads to efficient DNA methylation of genomic regions carrying a combination of both marks.

The readout of HPTMs by reading domains can also confer substrate specificity to an epigenetic enzyme (Torres & Fujimori, 2015). Besides its catalytic JmJC domain, the histone demethylase KDM4A (also: JMJD2A) contains two C-terminal plant homeodomains (PHDs) and a tandem tudor domain (**Figure 8a**) (Huang et al., 2006; J. Lee et al., 2008). KDM4A catalyses the demethylation of H3K9me3 and me2. Acting as a full-length protein in *in vitro* experiments, binding of the tandem tudor domains to trimethylated H3K4 drastically increased the catalytic demethylase activity on H3K9 (Lohse et al., 2013) (**Figure 8b**, centre). This suggests that in the presence of the activating H3K4me3 mark, KDM4A efficiently stabilizes the existent active chromatin state by removing heterochromatin-associated H3K9me2/3.
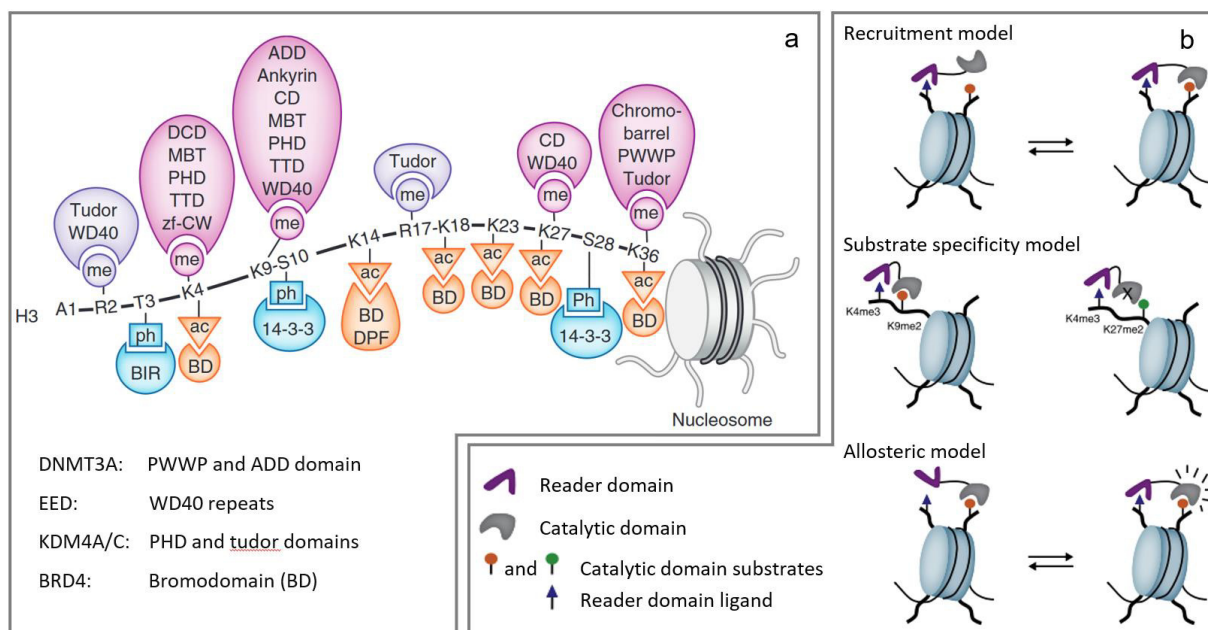
Figure 8: Diversity and function of HPTM readers. a, Representative overview of reading domains using the example of N-terminal HPTMs on H3. Readers of arginine and lysine methylation are coloured in shades of purple, acetylation readers in orange and phosphorylation readers in blue. Epigenetic effector proteins mentioned in the main text are annotated at the bottom. Image taken from Musselman et al., 2012. **b,** Different models for the coupling of reading domains to enzymatic activity. Reading domains can direct catalytic domains of the same polypeptide or complex to their site of action, modulate the substrate specificity by readout of neighbouring HPTMs or increase the catalytic activity by binding to HPTMs. The image was adopted from Torres et al., 2015.

The connection of HPTM reading and writing activity can also be mediated between different proteins in a multiprotein complex. The Polycomb repressive complex 2

(PRC2) catalyses trimethylation of H3K27 via its EZH2 subunit (Kuzmichev et al., 2002). Another subunit, the embryonic ectoderm development protein (EED), recognizes H3K27me3 via its WD40 domain and directs the PRC2 complex to regions carrying this mark, leading to the development and maintenance of extensive PRC2 repressed domains in a self-enhancing mechanism (Margueron et al., 2009). Besides directing PRC2 to modified targets, the binding of EED to H3K37me3 also allosterically activates PRC2 catalytic activity (Margueron et al., 2009) (**Figure 8b**). The essential role of HPTM reading domains in other epigenetic multiprotein complexes is described in section 1.3.

Analogous to the readers of HPTMs, the DNA methylation state of a genomic locus can be "read" by a group of proteins that contain a methyl-CpG-binding domain (MBD) (Q. Du et al., 2015). This MBD family contains four proteins that bind to methylated CpG independent of the sequence context (MeCP2, MBD1, MBD2, and MBD4), and one protein that does not bind to methylated DNA (MBD3) (Menafra & Stunnenberg, 2014; Yildirim et al., 2011). As reader proteins, MBD containing proteins exert their function over the recognition of the methylation state of a genomic locus, thereby targeting epigenetic complex proteins to methylated CpGs (Q. Du et al., 2015). MBDs play a role in processes like the regulation of histone modification patterns, chromatin remodelling, DNA repair or alternative splicing (Maunakea et al., 2013). MBD proteins establish a link between a local epigenetic signal and the directed reorganization of chromatin, either by direct interactions with chromatin or by the assembly of chromatin remodelling and deacetylation complexes (Bornelöv et al., 2018; L. Wang et al., 2020; Yildirim et al., 2011). Thus, despite the lack of an innate enzymatic function, they regulate crucial processes during development and cancer formation (Q. Du et al., 2015). Due to their unique ability to read CpG methylation specifically and in a sequence-independent manner, MBDs are used in several biotechnological applications (Jeltsch et al., 2020).

Another function of DNA methylation readers is the maintenance of DNA methylation patterns after replication. This function is conferred by the Ubiquitin Like With PHD And Ring Finger Domains (UHRF) family of proteins, which recognize hemimethylated DNA and recruit the maintenance methyltransferase DNMT1 (Arita et al., 2008; Bostick et al., 2007; Jeltsch, 2008).

Reading domains contribute to the complexity of the epigenetic gene regulation network by integrating the information that is indicated by the nature, combination, and number of chromatin modifications at a gene. Thus, even reading proteins completely lacking catalytic activity should not be underestimated in their biological function. For instance, bromodomain (BRD) containing proteins belong to an evolutionary conserved family of proteins recognizing acetylation of lysine, primarily in the context of histone lysine acetylation (Fujisawa & Filippakopoulos, 2017). The BRD containing protein BRD4 acts as a transcriptional regulator by direct recruitment of the transcriptional apparatus and reorganization of chromatin architecture (B. A. Gibson et al., 2019; Muhar et al., 2018; Rathert et al., 2015; Zuber, Shi, et al., 2011). In fact, many BRDs are potent targets for epigenetic therapies in a number of diseases (Filippakopoulos & Knapp, 2014). This emphasizes the need for a better understanding of the complex conjunctions inside the epigenetic network that sustain healthy gene expression programs in cells.

1.2.3. Erasing chromatin modifications

For all HPTMs described in the above sections, epigenetic enzymes removing those modifications have been discovered (Hsu et al., 2000; Mosammaparast & Shi, 2010; Seto & Yoshida, 2014; Ting et al., 2019). The most important epigenetic eraser enzymes which take part in the regulation of gene expression are histone deacetylases (HDACs) and lysine demethylases (KDMs).

In mammals, there are 18 different histone deacetylases divided into 4 classes that catalyse the removal of an acetyl group from the amino terminus of lysine, either in a Zn^{2+} or a NAD^+ dependent reaction (Seto & Yoshida, 2014). Of the 18 mammalian HDACs, not all are catalytically active, and only class I and class II HDACs play a major role in the regulation of gene expression (Mariadason, 2008). In general, HDACs are transcriptionally repressive epigenetic effector proteins as they reduce the accessibility of chromatin by removing histone acetylation. The family members HDAC1 and HDAC2 are responsible for about 50% of HDAC activity in the nucleus (Dovey et al., 2013; Kelly & Cowley, 2013). Their substrate specificity, their recruitment to DNA and their catalytic activity strongly depend on their association with complex partners (**Figure 12**), which will be further explained in section 1.3.1.

The lysine specific demethylase 1 (LSD1) was the first discovered enzyme shown to be able to remove the methylation of histone proteins (Shi et al., 2004). LSD1 can demethylate mono- and demethylated H3K4 and H3K9 and is part of multiple epigenetic complexes. As LSD1 was the main chromatin regulator that was investigated during this work, its function and relevance will be discussed in detail below.

Besides LSD1, histone lysine methylation can be removed by a set of proteins which contain the so-called Jumonji domain (JMJD). JMJD-containing demethylases can demethylate mono-, di- and trimethylated lysines in a Fe(II) and α-ketoglutarate dependent oxidative reaction, and specific enzymes for demethylation of each regulatory important lysine residue on H3 and H4 have been identified (Klose et al., 2006; Shin & Janknecht, 2007; Tsukada et al., 2006). More recently, JMJD containing proteins have also been associated with the demethylation of arginine residues (Meng et al., 2018).

As antagonist eraser enzymes for all HPTM writer proteins have been described, it is only consequential that there exists a group of enzymes counter-regulating DNA methyltransferase activity. However it was decades after the successful isolation of active DNA methyltransferases from mammalian cells (Roy & Weissbach, 1975; Sneider et al., 1975), when Tahiliani et al. identified three ten eleven translocase (TET) enzymes, TET1, TET2 and TET3 as proteins featuring the ability to convert 5-methyl cytosine to 5-hydroxymethyl cytosine (Tahiliani et al., 2009). TET enzymes are capable of further iterative oxidation of 5-hydroxymethyl cytosine (5hmC) to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). The oxidized products 5fC and 5caC can be removed by thymine DNA glycosylase (TDG) and the resulting abasic site is repaired by a base excision repair (BER) mechanism (Kohli & Zhang, 2013; H. Wu & Zhang, 2014), recuperating unmodified cytosine at this position. Other models for the removal of 5mC by TET propose an effect of 5hmC, 5fC and 5caC on the maintenance methylation by DNMT1 (H. Wu & Zhang, 2014). Independent of the adopted model, demethylation of 5mC is a complex mechanism involving multiple steps and protein complexes. The importance of TET enzymes becomes evident in their involvement in crucial biological processes with major rearrangements of gene expression programs. Such processes are haematopoiesis (Lio & Rao, 2019; Rasmussen & Helin, 2016),

embryonic development (Dawlaty et al., 2014) and cancer development (Rasmussen & Helin, 2016; Scourzic et al., 2015).

The lysine specific demethylase LSD1

For a long period of time during the history of histone modification studies, histone methylation was thought to be an irreversible signal once transferred to a histone protein, in contrast to histone acetylation and phosphorylation that could be enzymatically erased (Jenuwein, 2001). In 2004, Shi et al. described the lysine specific demethylase 1 (LSD1) as the first identified active histone demethylase (Shi et al., 2004). LSD1 had been detected before as part of the corepressor complexes NuRD (Tong et al., 1998), CoREST (You et al., 2001), HDAC (Hakimi et al., 2002; Humphrey et al., 2001) and CtBP (Shi et al., 2003), but so far it had been identified as the "hypothetical protein KIAA0601". Based on sequence homology analysis, this prevalent corepressor complex protein was suggested to have FAD dependent amine oxidase function. Shi then characterized LSD1 as a protein that specifically demethylated H3K4me2 peptides *in vitro*, depicting it as an evolutionary conserved transcriptional corepressor (Shi et al., 2004).

Shi et al. further investigated the catalytic properties of full-length recombinant LSD1 on methylated histone peptides and native histones isolated from HeLa. They could show that LSD1 demethylates mono- and dimethylated H3K4, but not trimethylated H3K4 or dimethylated H3K9 (Shi et al., 2004). By experiments with a C-terminal deletion mutant (amino acids 1-427), they identified the C-terminal amine oxidase (AO) domain as responsible for the catalytic activity on H3K4me.

LSD1 removes the methyl groups from H3K4 in two successive oxidative reactions, with flavin acting as a cofactor. The first step in the demethylation reaction is the FAD mediated two-electron oxidation of the methylated lysine (**Figure 9**). The resulting imine intermediate is hydrolysed in a non-enzymatic reaction to release formaldehyde, leaving a monomethylated amine at the lysine residue. The oxidation reaction is repeated to remove the remaining methyl group, producing unmodified H3K4. Reduced $FADH_2$ is re-oxidized by oxygen, producing H_2O_2 as a second side-product of the catalytic reaction (Forneris et al., 2005; Shi et al., 2004). The reaction mechanism provides one explanation for the specificity of LSD1 for mono- and di, but

not trimethylated H3K4 as a substrate, as the chemical nature of amine oxidase reactions requires a protonated nitrogen on the substrate (Shi et al., 2004).
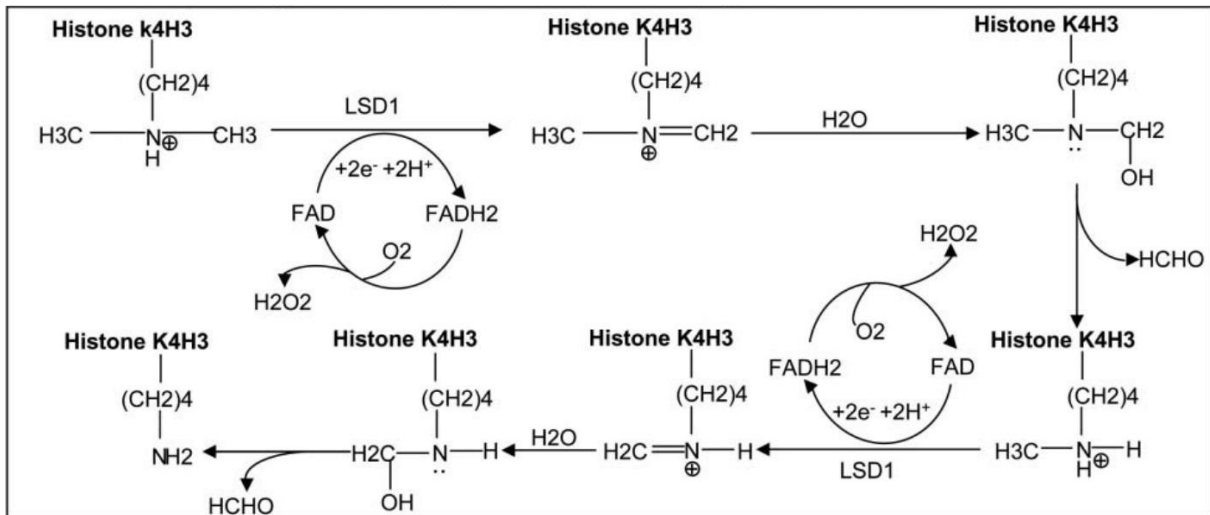
Figure 9: Reaction mechanism for the demethylation of H3K4me2 by LSD1. The methyl groups are removed successively in an oxidative reaction using FAD as a cofactor. Each removal is achieved in a two-step mechanism under the production of hydrogen peroxide and formaldehyde. The image was taken from Shi et al., 2004.

An important step towards the better understanding of the biological function of LSD1 was the crystallization of human LSD1 in complex with its cofactor FAD (Stavropoulos et al., 2006) (**Figure 10**). LSD1 consists of 3 main domains, a N-terminal *Swi3-Rsc8-Moira* (SWIRM) domain, the catalytic amine oxidase domain (AOD) and a protruding tower domain. SWIRM domains are evolutionary conserved domains found in a number of chromatin-associated proteins (Yoneyama et al., 2007). The SWIRM domain of LSD1 has been found to interact with the N-terminal tail of H3 (Tochio et al., 2006) and it is hypothesized that substrate binding by the groove between AOD and SWIRM is crucial for catalytic activity of LSD1 (Burg et al., 2016; Stavropoulos et al., 2006).
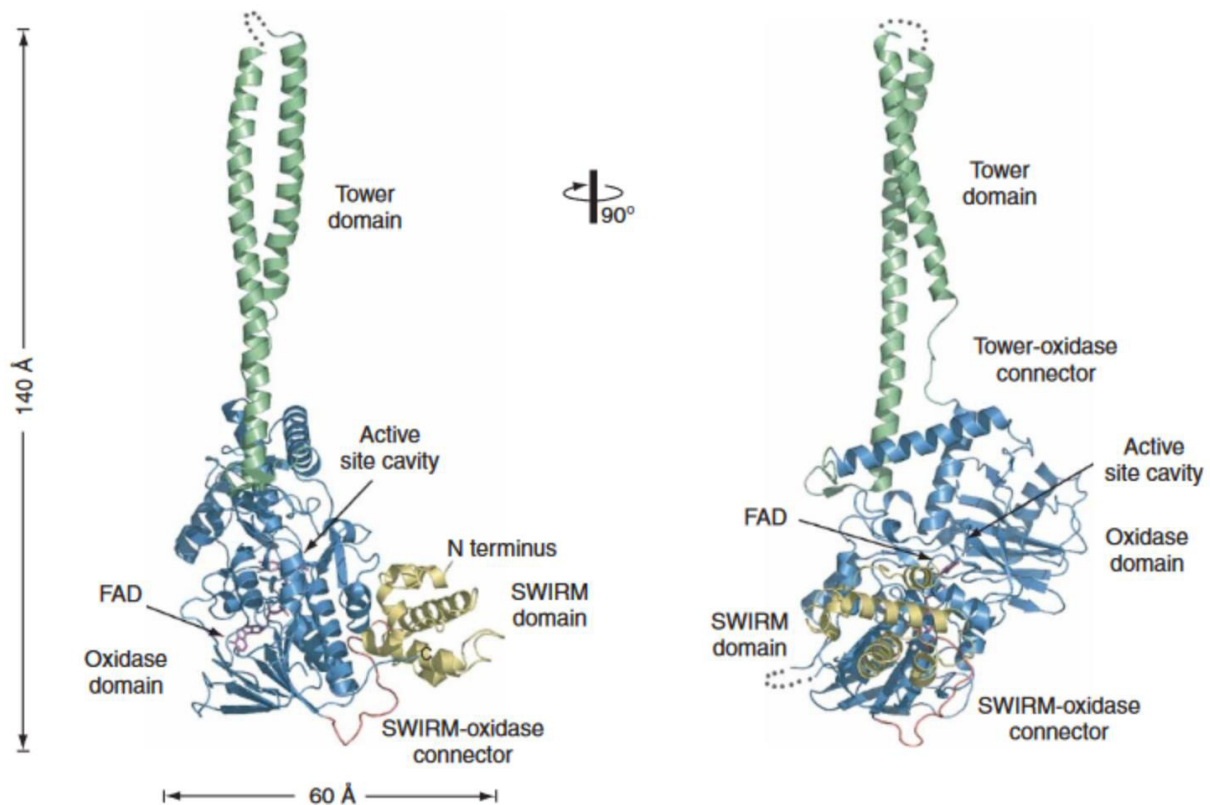
Figure 10: Crystal structure depicting the different domains of human LSD1. Two orthogonal views of the full-length protein are shown. The tower domain (green) protrudes from the amine oxidase domain (AOD, blue). The groove formed by the SWIRM (yellow) and AOD is needed for interaction of LSD1 with the H3 tail. The Flavin-adenine dinucleotide (FAD) cofactor is shown in purple. The image was taken from Stavropoulos et al., 2006.

The hydrophobic catalytic chamber of LSD1 is formed by the AOD and the tower domain and is relatively spacious, leaving room for the FAD binding and substrate binding site (Stavropoulos et al., 2006). There are three additional binding pockets close to the catalytic chamber, possibly recognizing side-chains of the H3 tail to modify LSD1 catalytic activity depending on the histone modification environment (Stavropoulos et al., 2006). Tight interactions of LSD1 and the H3 tail backbone surrounding H3K4 have been demonstrated, and are suggested to contribute to substrate specificity (M. Yang et al., 2007). The tower domain of LSD1 is formed by two antiparallel α-helices (termed TαA and TαB) that form a coiled-coil and extend approximately 100 Å from the AOD. The tower domain serves as a binding platform for regulatory proteins like CoREST (S.-A. Kim et al., 2020). Based on its role in the formation of the catalytic chamber of LSD1, it is suggested that binding of interacting proteins to the tower domain might regulate its catalytic activity (Stavropoulos et al., 2006).

Recent crystallization of LSD1 in complex with its interactor CoREST and a reconstituted nucleosome as substrate has expanded the findings made by Stavropoulos and Tochio (S.-A. Kim et al., 2020; Stavropoulos et al., 2006; Tochio et al., 2006). Kim et al. found that the interaction of LSD1 with the nucleosome in complex with CoREST is supported by the interaction of LSD1 with linker DNA, while CoREST interacts with the nucleosomal core over its SANT domain (**Figure 11**). Using biochemical assays, Kim et al. showed that the demethylase activity of LSD1 on nucleosome substrates is also dependent on its interaction with the extranucleosomal DNA and on the binding of CoREST to the nucleosome (S.-A. Kim et al., 2020). Interestingly, in their assays they also observed considerable demethylase activity of previously described catalytic mutants of LSD1 (Stavropoulos et al., 2006), suggesting that the activity of LSD1 is strongly modulated by the existent cofactor environment.
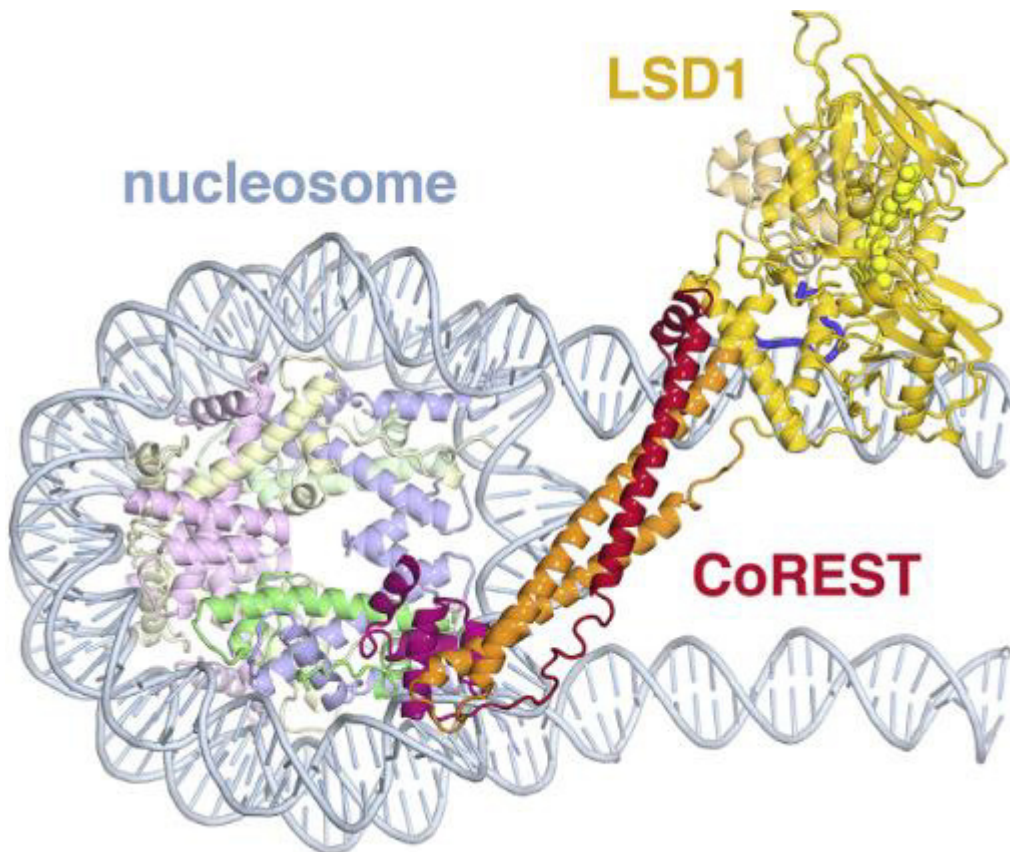
Figure 11: Crystal structure of LSD1 in complex with CoREST and the nucleosome. Truncated versions of the proteins (aa171-852 of LSD1 and aa286-440 of CoREST) and a 191 bp DNA molecule were used for crystallization. The AOD and SWIRM domain of LSD1 are shown in yellow, the tower domain is shown in orange. The CoREST SANT2 domain is pink and the linker domain red. The part of the H3 tail that inserts into the catalytic pocket of LSD1 is represented as a blue tube. The different histone proteins are shown in pastel colours. The image was taken from Kim et al., 2020.

LSD1 is mainly described as a transcriptional repressor, consistent with its catalytic activity to remove the activating H3K4me2 mark, and its association with repressive complexes (Maiques-Diaz & Somervaille, 2016; Y. Song et al., 2020). However, LSD1 has also been described to demethylate H3K9me2 and H3K9me1 in cells, thus functioning as a transcriptional activator in association with the androgen- or the estrogen-related receptor (Carnesecchi et al., 2017; Metzger et al., 2005; Wissmann et al., 2007). However, this catalytic activity could not be validated on peptide level by multiple groups and is in conflict with published crystal structures of LSD1 in association with H3 (S.-A. Kim et al., 2020; Shi et al., 2004; M. Yang et al., 2007). A splice variant of LSD1, LSD1+8a which is found in neuronal cells, has been shown to demethylate H3K9 during neuronal differentiation *in vivo* (Laurent et al., 2015). However, recombinantly expressed LSD1+8a is not able to demethylate H3K9 peptides *in vivo*, but only retains this ability when purified from neuroblastoma cells using tandem affinity purification (S.-A. Kim et al., 2020; Laurent et al., 2015). The biological meaning and functional mechanism of LSD1 activity on H3K9, as well as proposed activities on H4K20me (Jianxun Wang et al., 2015) remain diffuse, and are most probably limited to specific cell types and interactions with specific complex partners.

LSD1 is directed to its target sequences via interaction with various transcription factors and targeting of its repressive activity plays an important role during many processes that involve cellular reprogramming (Ferrari-Amorotti et al., 2013; X. Hu et al., 2009; Y. Lin et al., 2010; Maiques-Diaz & Somervaille, 2016; S.-T. Su et al., 2009; G. Sun et al., 2010). However, LSD1 also binds promiscuously at transcriptionally active sites and enhancers all over the mammalian genome (Garcia-Bassets et al., 2007; Whyte et al., 2012). It is supposed that LSD1 either resides at these loci in an inactive state, that H3K4me is protected by local factors, or that the repressive activity of LSD1 is counterbalanced by activating epigenetic complexes.

LSD1 acts as a transcriptional repressor as part of multiple epigenetic repressor complexes, like the CoREST/HDAC/LSD1 complex (M. G. Lee et al., 2005; Y. Song et al., 2020), the SIN3A/HDAC complex (Y. Yang et al., 2018) or the NuRD complex (Y. Wang et al., 2009) (see section 1.3.1). Besides its demethylase activity, LSD1 has been shown to have regulatory functions that are independent of its catalytic activity on a genome-wide scale (F. Gu et al., 2020; Maiques-Diaz et al., 2018).

LSD1 is a paradigmatic example for the networking complexity of epigenetic regulation of gene expression. Its catalytic activity is modulated by the histone code that is present on its target substrate, by its interaction with transcription factors or nuclear receptors and by accessory proteins binding to its tower domain. The recruitment of LSD1 to target loci is mediated by a multitude of transcription factors and other chromatin binding proteins, and once it is bound to a target its activity can be counterbalanced by other epigenetic effector proteins. Also, the activity of LSD1 can be regulated by several posttranslational modifications (Perillo et al., 2020). Finally, even if LSD1 is devoid of its catalytic activity, it still has regulatory function by serving as a scaffolding protein for epigenetic complexes. Consequently, although it has intensively been studied for many years now, the biological role of LSD1 has not been fully captured so far (Maiques-Diaz & Somervaille, 2016).

1.3. Formation of complex epigenetic networks

Chromatin-related processes like the regulation of transcription, DNA replication or DNA repair are regulated by the local combination of chromatin modifications that originate from a vast pool of different HPTMs and DNA modifications. The intricacy of this so-called "histone code" (Strahl & Allis, 2000) is further increased by cofactors like the nuclear localization of the chromatin locus, the number of added modifications and the combinatorial effects of different modifications, either on the same histone tail or on neighbouring histones. Thus, the connection of a specific histone modification and its regulatory outcome is not linear but has to be interpreted by the integration of all surrounding factors (Strahl & Allis, 2000). This complex task is performed by a network of epigenetic effector proteins that are connected via a multiplicity of highly specific protein-protein and protein-DNA interaction domains (Malovannaya et al., 2011; Zhao et al., 2017). Chromatin regulation can work in self-enforcing epigenetic circuits that are fuelled by the local accumulation of histone PTMs, followed by the recruitment of enzymes recognizing these HPTMs, and leading to subsequent deposition of further HPTMs by associated protein complexes (Margueron et al., 2009). These circuits enable the switching between opposing chromatin states in a way that is tightly regulated and highly dynamic at the same time. While this complexity safeguards epigenetic control mechanisms from unintended mistakes, it complicates the investigation of single units of epigenetic regulation. To illustrate the process of

chromatin regulation by chromatin effector complex formation, several complexes will be exemplified in the following paragraphs.

1.3.1. LSD1 and the different HDAC complexes

The histone deacetylases HDAC1 and HDAC2 play a major role in the regulation of chromatin structure and function (Dovey et al., 2013; Kelly & Cowley, 2013). They execute distinct roles as subunits of different multiprotein complexes (**Figure 12**). Interestingly, LSD1 can interact with all three of the major HDAC complexes (S.-A. Kim et al., 2020; Y. Wang et al., 2009; Y. Yang et al., 2018), highlighting the role of LSD1 as an important networking regulator. Although the deacetylation of histones usually is associated with gene repression, the individual outcome of HDAC binding is dependent on its association with different coregulators (Adams et al., 2018).

An unambiguously repressive HDAC complex is the HDAC/LSD1/CoREST complex (**Figure 12**) (You et al., 2001). In this complex, the physical interaction of HDAC1 and LSD1, mediated by the repressor protein CoREST, facilitates the catalytic activity of both enzymes in a co-dependent manner. This mutual activation leads to the efficient demethylation and deacetylation of target chromatin (M. G. Lee et al., 2005; Y. Song et al., 2020). LSD1 is the effector of the repressive activity of the complex, but the removal of either HDAC or CoREST severely impacts the demethylase activity of LSD1 on target genes (M. G. Lee et al., 2005; Maiques-Diaz & Somervaille, 2016). The HDAC/LSD1/CoREST complex regulates essential cell programs like hematopoietic differentiation, neuronal development and malignant proliferation (Fuentes et al., 2012; F. Gu et al., 2020; Majello et al., 2019; Saleque et al., 2007; Takagi et al., 2017).

In association with the Sin3 complex, HDAC1/2 and LSD1 regulate essential cell-type specific gene expression programs during differentiation and development (Streubel et al., 2017). Sin3A and Sin3B complexes are multifaceted corepressor complexes that can assemble around various TFs, Sin-associated proteins (SAPs) and even TETs (Adams et al., 2018; Chandru et al., 2018; Fleischer et al., 2003; Swanson et al., 2004; Yi Zhang et al., 1997). The role of the Sin3/HDAC complex in transcriptional regulation is ambivalent, depending on the context and combination of subunits, it can either have activating or repressive function on transcription (Adams et al., 2018). However, in association with LSD1 the Sin3/HDAC complex represses the expression of multiple oncogenes (Y. Yang et al., 2018).

In combination with MBD containing proteins, the chromodomain helicase CHD4 (also called Mi2-β), and different histone chaperone and zinc-finger proteins, HDAC1/2 and LSD1 can form the nucleosome remodelling and deacetylation (NuRD) complex (**Figure 12**) (Bornelöv et al., 2018). The NuRD complex has a key role in the reprogramming of ES cells during early differentiation (Reynolds et al., 2012). NuRD has an affinity for open chromatin and can be found widely distributed over the genome (Bornelöv et al., 2018; Stevens et al., 2017). The repressive activity of NuRD is executed by two catalytic mechanisms: the deacetylation of histones by HDAC1/2 and chromatin remodelling by CHD4. On a local level, NuRD induction by differentiation signals leads to the compaction of chromatin, the loss of TF binding and RNA-polymerase II accessibility and transcriptional repression (Bornelöv et al., 2018; Reynolds et al., 2012). The observation of a comprehensive reorganization of global chromatin architecture, following the clustering of NuRD-occupied domains, suggests the existence of additional and more global mechanisms of NuRD dependent reprogramming (Bornelöv et al., 2018; Stevens et al., 2017).
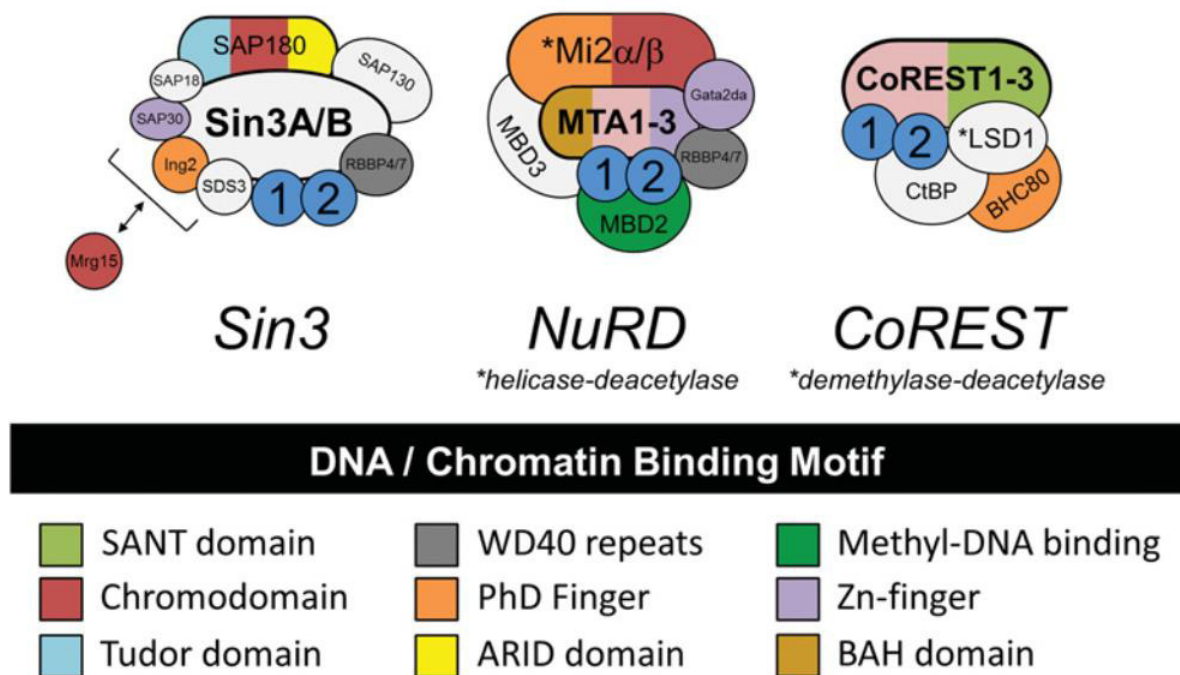
Figure 12: Composition of the different HDAC/LSD1 complexes. *Top*: HDAC1 and 2 are mostly redundant in their biological function and can associate with different epigenetic complexes. The function of HDAC1/2 in association with the Sin3 complex can either be activating or repressive, depending on the chromatin context. As subunit of the NuRD complex, HDACs contribute to repressive chromatin remodelling. In association with CoREST and LSD1, HDAC1/2 act as transcriptional repressors by promoting the demethylation of H3K4. *Bottom*: The HDAC/LSD1 complexes are targeted to DNA and chromatin by association with different reading domains. The image was taken from Kelly & Cowley, 2013.

In conclusion, the association of HDAC1/2 and LSD1 with different chromatin complexes integrates their repressive deacetylase activity into a comprehensive chromatin silencing process. This process involves the target identification by reading domains and TFs (**Figure 12**), the fine-tuning of enzymatic activity by associated proteins, the remodelling of local chromatin structure, and the formation of large repressive domains by clustering of the protein complex with chromatin (Bornelöv et al., 2018; Y. Song et al., 2020; Stevens et al., 2017). In addition to the HDAC complexes, LSD1 directly interacts with a variety of other coregulators like SNAIL/SLUG (Ferrari-Amorotti et al., 2013), the PRC2 complex (Y. Jin et al., 2017) and MLL1 (Jianxun Wang et al., 2007). This makes LSD1 a versatile chromatin regulator, which is embedded tightly into the network of epigenetic gene regulation.

1.3.2. The ZNF10-KRAB complexes

The family of Krüppel-associated box (KRAB) domain containing zinc-finger proteins (KRAB-ZFPs) is a large group of transcription factors with repressive gene regulatory function. KRAB-ZFPs have been shown to be essential for the stable silencing of transposable elements in the genome, mostly by recruitment of the Tripartite Motif Containing 28 (TRIM28 or KAP-1) protein (Jacobs et al., 2014; Wolf & Goff, 2009). Mammals contain 200-500 different KRAB-ZFPs, each targeting a different set of transposable elements (Imbeault et al., 2017). This sparked the idea that KRAB-ZFPs might have co-evolved in an arms race with new transposons invading the genome, with new KRAB-ZFPs constantly evolving by gene duplications to inhibit their transposable activity and prevent damage caused by random integration (Imbeault et al., 2017; Jacobs et al., 2014). However, this cannot be the only driver of evolution and conservation of KRAB-ZFPs, as KRAB-ZFPs still exist long after the transposons they target have lost the ability for transposition and/or transcription. KRAB-ZFPs lack an innate catalytic activity, however they represent strong regulators of transcriptional repression. The recruitment of KRAB-ZFPs to a genomic locus induces the formation of a *de novo* heterochromatic environment, leading to the silencing of nearby genes by a position-effect-variegation (PEV) like mechanism (Sripathy et al., 2006; Tchasovnikarova et al., 2015). This shows that KRAB-ZFPs contain the innate ability to repress actively transcribed genes, and this feature has extensively been used for the epigenome editing of endogenous genomic loci (Thakore et al., 2016).

49

Furthermore, KRAB-ZFPs can regulate gene expression by modulating the chromatin state of target retrotransposons in the proximity of transcription start sites (Chuong et al., 2017; Ecco et al., 2016). In doing so, KRAB-ZFPs can use transposable elements in the mammalian genome as "regulatory hubs" for the binding of different transcription factors, enabling the regulation of gene expression by global reorganization of chromatin architecture (Chuong et al., 2017; Ecco et al., 2016).

KRAB-ZFPs consist of an N-terminal ~75 amino acid KRAB domain linked to a tandem array of DNA-binding, cysteine and histidine containing C2H2-zinc-finger modules (P. Yang et al., 2017). Most KRAB-ZFPs have a KRAB domain that is composed of the two KRAB subdomains KRAB-box A and B, although it has been shown that subdomain A is sufficient for transcriptional repression, while subdomain B acts only as enhancer of silencing by KRAB-box A (Vissing et al., 1995). In this work, the focus will be on ZNF10 and its KRAB domain, as this domain is most commonly used in epigenome editing and transcriptional regulation experiments.

The repressive activity of KRAB-ZFPs is mediated by recruitment of KAP-1 (TRIM28) over the KRAB domain (Messerschmidt et al., 2012; Wolf & Goff, 2009), which subsequently acts as a scaffolding protein for different epigenetic silencing complexes (**Figure 13a**). For example, KAP-1 can recruit the Mi-2/NuRD complex (D. C. Schultz et al., 2001) and the H3K9-specific SET Domain Bifurcated 1 (SETDB1) histone lysine methyltransferase (David C Schultz et al., 2002). Consequently, sequence-specific KRAB recruitment leads to the local assembly of a large epigenetic silencing complex and the formation of a constitutive heterochromatin domain (Nishibuchi & Déjardin, 2017; P. Yang et al., 2017).

KRAB-ZFPs are targeted to their site of action by an array of zinc-finger domains, each with a recognition motif for the binding of four specific nucleotide bases in the major groove of DNA (**Figure 13b, c**) (Brayer & Segal, 2008). In humans, each KRAB protein may have 4-34 zinc-finger domains (Urrutia, 2003), theoretically enabling the specific recognition of a sequence with a length of up to 90 bp.
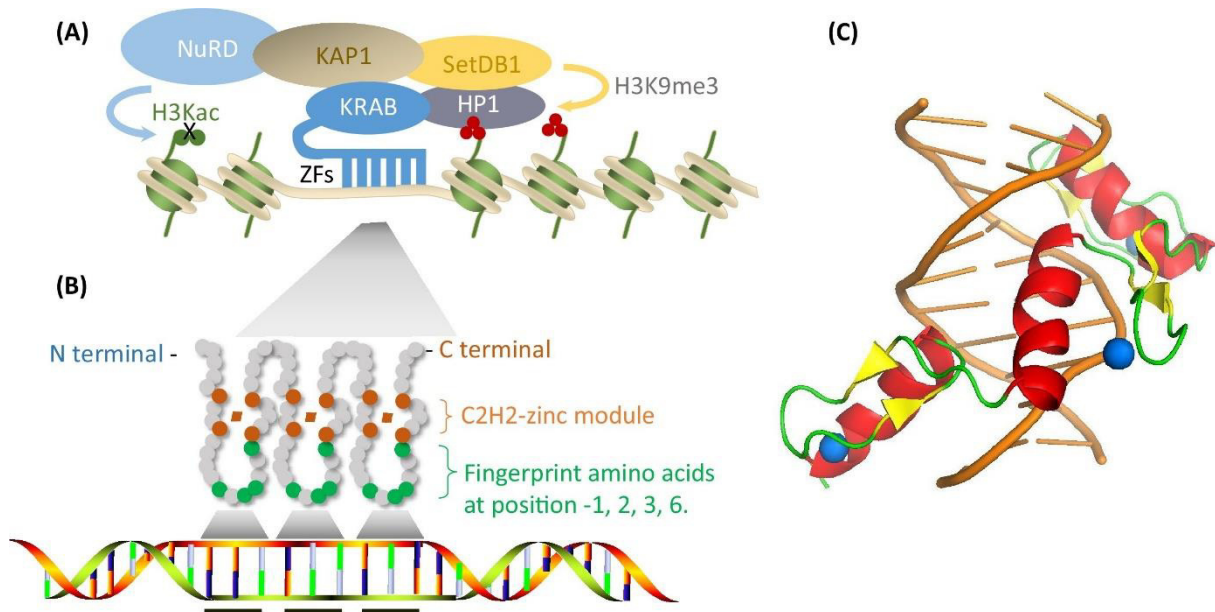
Figure 13: The mechanisms behind KRAB mediated chromatin silencing. A, Schematic depicting the silencing complex forming around KRAB-ZFPs. The KRAB domain recruits heterochromatin proteins containing a H3K9me3 reading domain (HP1) and the scaffolding protein KAP-1 (TRIM28). KAP-1 induces local histone deacetylation and H3K9 trimethylation by recruitment of NuRD and SETDB1. **B,** Schematic explaining the sequence specific DNA recognition by a three-fingered zinc-finger domain of ZFs. The tandem recognition structure is stabilized by arrays of cysteine and histidine residues (red circles), which interact with zinc ions to form the C2H2-zinc module. This structure supports the extrusion of fingerprint amino acid sequences, which specifically bind individual DNA base sequences. **C,** Structural model of the recognition of a DNA sequence by a three-fingered zinc-finger domain. Each zinc-finger consists of an α-helix (red) and two β-sheets (yellow). The C2H2-zinc module is built up by the C-terminus of the α-helices and invariant loop regions (green) wrapped around the zinc ions (blue). Sequence-specific DNA binding is performed by the α-helices in the major groove of DNA. The figure was taken from Yang et al., 2017.

Besides their ability to recruit NuRD and SETDB1 via KAP-1, KRAB-ZFPs have been shown to interact with the human silencing hub (HUSH) complex, consisting of the Transcription Activation Suppressor (TASOR or D14ABB1E), the M-Phase Phosphoprotein 8 (MPHOSPH8) and Periphilin 1 (PPHLN1). Tchasovnikarova et al. identified the HUSH components as enhancers of PEV in the haploid human leukaemia cell line KBM7 (Tchasovnikarova et al., 2015). Furthermore, they showed an interaction of the HUSH components with SETDB1, and that HUSH-mediated gene repression of transgenes integrated into sites of high H3K9me3 was targeted by MPHOSPH8. The main, evolutionary conserved mechanism for the stable silencing of endogenous retroviruses at sites of KRAB-ZFP binding is the formation of a repressive heterochromatin domain, executed by the assembly of SETBD1 and the heterochromatin protein 1 (HP1) in a complex around KAP-1 (Nishibuchi & Déjardin,

2017; David C Schultz et al., 2002). Silencing mediated by the HUSH complex is suggested to have evolved later as an additional mechanism, as only silencing of evolutionarily young LINEs is dependent on recruitment of D14ABB1E (Robbez-Masson et al., 2018). The biological relevance and detailed mechanism of this regulatory complex formation remain to be determined.

1.3.3. Chromatin remodelling complexes

An essential part of transcriptional regulation is connected to the three-dimensional remodelling of chromatin from an accessible to an inaccessible chromatin state and *vice versa* (Cairns, 2009). In interphase nuclei of eukaryotic cells, DNA is tightly packed into a 30 nm fiber consisting of stacked nucleosome core particles (Grigoryev & Woodcock, 2012). Although this form of highly organized packaging enables the orderly storage of long DNA strands using minimal space, it must be resolved during cellular processes such as transcription, replication, recombination, or DNA repair (Cairns, 2009; Venkatesh & Workman, 2015). Chromatin remodellers are organized in large multi-subunit complexes to enable target identification and activity regulation of chromatin remodelling (Längst & Manelyte, 2015). Four families of ATP-dependent chromatin remodelling complexes have been identified (Längst & Manelyte, 2015; Narlikar et al., 2013). The SWI/SNF complex in yeast was the first ATP-dependent coactivator complex described (Côté et al., 1994), and other helicases carrying a highly conserved ATPase core have since been identified in all eukaryotes (Narlikar et al., 2013). Chromatin remodellers can perform several functions to open up condensed chromatin regions, like the sliding of histone octamers along the DNA, the exchange of single histone core particles to variants or changing the nucleosomal DNA conformation (Becker & Workman, 2013). They share the ability to overcome the energetic barrier needed to break DNA-histone interactions in the nucleosome using the energy from ATP hydrolysis, thus paving the way for other proteins to access DNA (Hargreaves & Crabtree, 2011; Narlikar et al., 2013; Santoro et al., 2002). Depending on the nature of the recruited chromatin remodelling complexes (CRCs), the transcriptional output of a gene can either be active, as in the case of human SWI/SNF (Soutoglou & Talianidis, 2002), or repressive, as for the nucleolar remodelling complex NoRC (Santoro et al., 2002). Remodelling activity is strongly dependent on the local chromatin context (Längst & Manelyte, 2015), making chromatin remodelling an

important mediator between the histone code and the transcription machinery. Thus, to secure a controlled process of chromatin remodelling, proteins with catalytic remodelling activity depend on the association with HPTM readers and modifiers in multiprotein complexes, in order to find their targets and prepare the chromatin environment for the structural reorganization (Längst & Manelyte, 2015). Mutations related to subunits of chromatin remodelling complexes play a role in the development of various cancers (Hodges et al., 2016; B. Li et al., 2007), underlining the importance for the investigation of epigenetic complex formation. Chromatin remodelling is inextricably linked to the regulation of gene expression by changes in chromatin modifications. However, it will not be in the focus of this work.

1.4. Epigenetics in development and cancer

The previous paragraphs have illustrated the diversity of epigenetic mechanisms that regulate gene expression programs in cells. In the following sections, the physiological consequences of defects in the epigenetic regulatory system will be described. Following a general illustration of critical events of epigenetic reprogramming, the aftermath of epigenetic malfunction will be exemplified using the example of the lysine-specific demethylase 1 (LSD1), as this protein was in the focus of this thesis.

1.4.1. Implications of epigenetic defects during embryonic development

Somatic cells of mammalian organisms are governed by stable and heritable gene expression patterns that determine their cell-type specific morphology and functionality. These expression patterns are regulated by the targeted recruitment of transcription factors (TFs) and chromatin modifying enzymes to promoter and enhancer regions, resulting in the consequential activation or repression of associated genes (Gökbuget & Blelloch, 2019). Cells of a certain lineage commit to their specific cell-fates during embryonic development, a process that is orchestrated by epigenetic mechanisms (Mikkelsen et al., 2007; Xia & Xie, 2020). To execute this transition, extensive epigenetic reprogramming from the pluripotent chromatin state to a chromatin state driving the gene expression pattern of the respective differentiated cell type is necessary (Messerschmidt et al., 2014; Xia & Xie, 2020). Mistakes occurring during this vast reorganization event have global consequences for the developing organism, ranging from the development of chronic diseases (Jakovcevski & Akbarian,

2012; Portela & Esteller, 2010; Tough, 2016) over severe embryonic growth defects (Messerschmidt et al., 2012) to embryonic lethality (W. Lin & Dent, 2006; Z. D. Smith & Meissner, 2013).

During a healthy somatic cell life cycle, the transition from pluripotent stem cells to differentiated cell lineages is irreversible. Several important epigenetic "barriers" have been identified in adult organisms that prevent the reprogramming of somatic cells back to stem cells (Jiekai Chen et al., 2013; Soufi et al., 2012). However, it was possible to overcome these barriers to engineer induced pluripotent stem cells (iPSCs) (Papp & Plath, 2013). The deregulation of epigenetic mechanisms that trigger the reactivation of pluripotency genes or genes enabling irregular differentiation can lead to malignant transformation of cells (Flavahan et al., 2017). Mutations in a single epigenetic effector protein that may be acquired during somatic cell growth can have severe consequences for the global homeostatic chromatin state, by inducing oncogene activation, tumour suppressor silencing or adaptive cell fate transitions (Flavahan et al., 2017).

1.4.2. LSD1 and embryonic development

The lysine specific demethylase LSD1 acts as an epigenetic regulator of transcription. The complex modes of action of LSD1 are explained in detail in section 1.2.3. The expression of LSD1 during embryonic development is essential for the survival of the embryo, with homozygous genetic deletion of the LSD1 AOD in mouse preimplantation embryos being lethal at embryonic day <7.5 (Jianxun Wang et al., 2007). LSD1 orchestrates the establishment of cell-type specific developmental programs by directing co-activator (WDR5) and co-repressor (CoREST, HDAC) proteins to the promoters of developmental genes after initial organ commitment (Jianxun Wang et al., 2007). Conditional knockout of LSD1 in embryonic stem cells (ESCs) *in vitro* leads to the aberrant expression of transcription factors regulating anterior/posterior patterning and limb development (Foster et al., 2010).

Besides promoter regulation, LSD1 can associate with the nucleosome remodelling and deacetylase (NuRD) complex to decommission enhancers that drive expression of ESC genes like Oct1, Nanog and Sox2, thereby paving the way for differentiation of ESCs (Whyte et al., 2012). On a global level, LSD1 regulates the maintenance of DNA methylation during development. This can either be mediated by the removal of lysine methylation on DNMT1, or by alterations of chromatin loci that serve as a substrate for

DNMT1 (Jing Wang et al., 2009). In conclusion, the interference with the regulatory activity of LSD1 in stem- or progenitor cells leads to severe, distinct and cell-type specific defects in differentiation and development (Di Stefano et al., 2007; Fuentes et al., 2012; Vinckier et al., 2020; Jianxun Wang et al., 2007), underlining its essential role in cellular reprogramming, while the distinct mechanisms of action remain to be fully described.

1.4.3. Diseases associated with the deregulation of epigenetic factors

With the field of epigenetics growing and getting more attention over the past decades, the interest in finding epigenetic sources for diseases is increasing as well. One field of medical science trying to associate epigenetics with distinct diseases is neurobiology. Deregulation of DNA methylation and histone modification states have been found in neurodegenerative diseases like Alzheimer´s disease, Parkinson´s disease and even epilepsy (Landgrave-Gómez et al., 2015), but no mechanistic links between chromatin states and disease progression have been shown. Still, there are increasing hints that neurological disorders could benefit from epigenetic therapies (Jakovcevski & Akbarian, 2012). The Rett syndrome, a progressive neurodevelopmental disorder, is caused by mutations in the Methyl-CpG-binding protein 2 (MeCP2) (Amir et al., 1999; L. Wang et al., 2020) and the immunodeficiency, centromeric instability and facial anomalies (ICF) syndrome, a disease that also includes mental retardation of patients, is raised by mutations in the DNMT3B gene (R. S. Hansen et al., 1999). Interestingly valproic acid, a drug used to treat various types of seizure disorders and manic depression (Davis et al., 1994), acts as a histone deacetylase inhibitor and might have neuroprotective effects (Y. Su et al., 2004), although the mechanistic links still have to be discovered.

Besides the ICF syndrome, other diseases related to disfunctions of the immune system are being investigated for the influence of epigenetics on their progression (Portela & Esteller, 2010). Differential DNA methylation patterns in the MHC cluster have been shown to be a risk factor for the development of rheumatoid arthritis (Y. Liu et al., 2013) and patients with systemic lupus erythematosus (SLE) show downregulated HDAC activity in monocytes, leading to aberrant acetylation of histones (Leung et al., 2015). However, despite intensive research in this field, no epigenetic therapies for autoimmune diseases have been developed so far (Jeffries, 2018; Karagianni & Tzioufas, 2019).

Epigenetic gene regulation sustains the normal execution of gene expression programs. Consequently, defects in this regulatory system can lead to aberrant cell proliferation, cell-cycle progression, cell invasion and thus the development of cancer (P. A. Jones et al., 2016). Fortunately, significant progress in the development and establishment of epigenetic therapies for the treatment of different cancers has been made over the past decade. Drugs targeting DNMTs, HDACs, IDH1 and 2, histone methyltransferases (HMTs) and chromatin remodelling proteins have already been approved for therapeutic use by the US Food and Drug Administration (FDA), and many more epigenetic drugs currently undergo preclinical or clinical trials (Bates, 2020). DNA hypomethylating agents and histone deacetylase inhibitors are used for treatment of acute myeloid leukaemia (ALL) (Bohl et al., 2018) as well as for the combinatorial treatment of solid tumours, e.g. in breast (Connolly et al., 2017) and prostate cancer (Ferrari et al., 2019). In these hormone dependent tumours, epigenetic therapies are applied to re-sensitize resistant cancers to canonical treatments by reactivation of target gene expression. This supportive feature to overcome relapsing tumour growth is in the focus of studies searching for novel treatments for other cancers with high rates of resistance development (Brown et al., 2014; Moufarrij et al., 2019; Olino et al., 2020). The prospects for an increasing number of epigenetic therapies for cancer are promising. However, the highly plastic nature of epigenetic networks, which respond to inhibition of a single subunit with dynamic adaptions of coregulatory processes, renders epigenetic therapies susceptible to resistance development (Quagliano et al., 2020; Rathert et al., 2015). A better mechanistic understanding of the epigenetic changes leading to malignant cell growth, and of the processes that a targeted therapy could manipulate to effectively inhibit tumour growth, will be needed to boost the establishment of efficient cancer specific treatments (Quagliano et al., 2020).

1.4.4. LSD1 and cancer

LSD1 has been identified as a therapeutic target in many cancers (Verde et al., 2017). The oncogenic properties of LSD1 are mediated both by its catalytic and non-catalytic functions and can be targeted by specific antagonist molecules in cancer therapy (Fiskus et al., 2014; Vinyard et al., 2019). In acute myeloid leukemia (AML), LSD1 sustains the oncogenic potential of leukemia stem cells by inhibiting differentiation pathways (Harris et al., 2012; Schenk et al., 2012). In ER negative breast cancer, LSD1

overexpression can serve as a biomarker for an aggressive phenotype and is essential for tumour survival and growth (Lim et al., 2010; Y. Yang et al., 2018). At the same time, LSD1 can inhibit epithelial to mesenchymal transition and retain chemo-sensitivity of breast cancer cells, therefore acting as a tumour suppressor (Y. Wang et al., 2009; Y. Yang et al., 2018). Thus, the effective role of LSD1 must be determined for each clinical picture. As a coregulator of hormone signalling, LSD1 plays an important role in a variety of hormone-driven cancers like breast-, prostate- and bladder cancer (Benedetti et al., 2019; Kauffman et al., 2011; Metzger et al., 2005; Wissmann et al., 2007).

LSD1 can contribute to carcinogenesis as a transcriptional regulator by demethylating H3K4 or via its interaction with hormone receptors, but it can also change tumour pathogenicity by demethylation of non-histone targets or interaction with lncRNAs (Lian et al., 2017; Majello et al., 2019; M. Sun et al., 2016). These multiple modes of action underline the importance for LSD1 as a target for epigenetic therapies in cancer, and a number of drugs are currently undergoing clinical trials (Y. Fang et al., 2019) (**Table 1**). A promising candidate is the highly selective small-molecule inhibitor ORY-1001, which has been shown to effectively increase H3K4me2 levels at LSD1 target genes, inducing differentiation of leukemic stem cells in AML and reducing tumour growth in small cell lung cancer (SCLC) (Augert et al., 2019; Y. Fang et al., 2019; Maes et al., 2018). The inhibitor molecule GSK2879552 shows promising effects in AML, SCLC and prostate cancer (Y. Fang et al., 2019; Gao et al., 2020; Stewart & Byers, 2015), changing carcinogenic gene expression patterns by modulating hormone receptor binding and reactivating tumour suppressor signalling pathways.

Besides the development of cancer therapies based on targeting LSD1, LSD1 inhibitors are currently being tested for application in the treatment of neurodegenerative diseases like Multiple Sclerosis, Alzheimer´s disease and Huntington´s disease (Y. Fang et al., 2019) (**Table 1**). ORY-2001 has been shown to be brain-permeable, safe and effective to reduce cognitive impairment in mouse models for neurodegenerative disorders (Maes et al., 2016). Although these studies are still in a very early stage, they show great promise given the existent limitations in the treatment of neurodegenerative diseases.

Table 1: LSD1 inhibitors in clinical trials. The table represents the clinical status in October 2019. The table was taken from Fang et al., 2019. A constantly updated status can be reviewed online at www.clinicaltrials.gov.

Drugs	Phase	Trial number	Diseases
ORY-1001	Phase I/II	NA*	AML
	Phase I	NCT02913443	SCLC
	Preclinical	NA*	AML, solid tumors
TCP	Phase I	NCT02273102	AML; MDS
	Phase I/II	NCT02261779	Relapsed/refractory AML
	Phase I/II	NCT02717884	Non-M3 AML blasts
GSK2879552	Phase I	NCT02034123	Relapsed/refractory SCLC
		NCT02177812	AML
	Phase II	NCT02929498	High-risk MDS
INCB059872	Phase I/II	NCT02712905	Solid tumors and hematologic malignancy
	Phase I	NCT03514407	Relapsed Ewing sarcoma
	Phase I/II	NCT02959437	Solid tumors Advanced malignancies Metastatic cancer
	Phase I	NCT03132324	Sickle cell disease
	Phase I/II	NCT04061421	MDS/MPN
IMG-7289	Phase II	NCT03136185	Myelofibrosis
	Phase II	NCT04081220	Essential thrombocythemia
	Phase I	NCT02842827	AML and MDS
CC-90011	Phase I	NCT02875223	Relapsed/refractory solid tumors and non-Hodgkin's lymphomas
	Phase I/II	NCT03850067	SCLC
ORY-2001	Phase I	NA*	Multiple sclerosis
	Phase IIa	NCT03867253	Mild to moderate Alzheimer's disease

1.5. Methods for the investigation of chromatin effector complexes

1.5.1. Challenges

The complexity of epigenetic gene regulation poses a great challenge for the investigation of the effects of individual epigenetically active proteins on transcription. The expression of a specific gene is modulated by the recruitment of TFs and chromatin modifiers, and this recruitment is dependent on the repressive or activating state of the surrounding chromatin (Gao et al., 2020; Kribelbauer et al., 2017; Mazzio & Soliman, 2012; Tycko et al., 2017). However, not only the local chromatin state or the conformation of proximal promoter regions is involved in the regulatory process, gene expression can additionally be dependent on the epigenetic state of one or multiple distal enhancer elements, which modulate gene expression over the range of

large distances (H. K. Long et al., 2016). Thus, the translational outcome of a single gene is a stochastic event under the influence of various epigenetic and gene-regulatory processes with high spatiotemporal complexity (Nicolas et al., 2017; Tycko et al., 2017). Defining the role of a single regulatory factor inside this complex interaction network still poses a great challenge, and requires a sophisticated experimental approach.

Studies of the epigenome face the challenge of monitoring effects that can have a global impact on the chromatin environment, that are interconnected with multiple (possibly unknown) other epigenetic processes, and that are dynamically changing upon external signals or counteracting mechanisms. The easiest way to tackle these challenges is to reduce the complexity by performing *in vitro* experiments, although this will first require a good hypothesis to be able to perform a meaningful experiment with results that can be transferred to the whole cell or organism. Advances in high throughput technologies have significantly improved the investigation of global changes on protein-, RNA- or DNA-level. Still, these methods usually deliver only a snapshot of the cellular state and represent an average value of a diverse heterogenous cell population. By combination of multiple state-of-the-art methods, an approach that dynamically traces changes in epigenetic systems in a comprehensive manner can be created, aiming to define the nodes that hold together the epigenetic regulatory network. Ideally, this approach would be universally applicable to living cells of different developmental or mutagenic background, and able to detect any arising protein interactions, independent of individual predictive hypotheses.

1.5.2. *In vitro* methods

With the advances in producing modified histone peptide arrays (Rathert et al., 2008), isolating whole nucleosomes from mammalian cells (O'Neill & Turner, 2003) and even creating specifically modified recombinant designer nucleosomes (Neumann et al., 2009; Simon et al., 2007), there are many ways to obtain substrates for biochemical *in vitro* activity screening of epigenetic effector proteins. These biochemical assays can be used for the identification of epigenetic reader-, writer- or eraser-protein substrates and the characterization of mutant variants of those proteins (S.-A. Kim et al., 2020; X. Wang et al., 2017; Weirich et al., 2016). They provide a good tool for the definition of protein-substrate interactions and the impact that different combinations of histone tail modifications have on this interaction (Dou et al., 2006; S.-A. Kim et al., 2020;

Kungulovski et al., 2016; Yingying Zhang et al., 2010). However, observations from those assays sometimes fail to translate into the cellular context, and the extent to which complex interactions of proteins in this context can be investigated is very limited.

To identify interactions of epigenetic effector proteins in multiprotein complexes, coprecipitation experiments are the most common way to go (Malovannaya et al., 2011). The protein of interest can be precipitated from cellular lysate using a protein-specific antibody or a peptide tag that was biotechnologically attached to the protein before harvesting of the cells (Bantscheff et al., 2011; Elion, 2007). To identify specific proteins that form stable physical complexes with the target conditions, precipitates from the solution can be analysed using SDS-PAGE followed by detection with specific antibodies. To analyse the whole set of proteins coprecipitated with the target, the samples can be digested with specific proteases after precipitation and analysed with high-performance liquid chromatography (HPLC) and mass spectrometry (Bantscheff et al., 2011; Malovannaya et al., 2011; Weigt et al., 2016). Although these methods are powerful to identify all proteins physically interacting with a protein of interest, low abundance proteins might be missed and functional interactions without a physical component will be missed as well. Also, coprecipitation experiments require either the availability of high-quality antibodies for every target or the overexpression of a peptide-tagged protein in cells. This problem can be solved by precipitating the proteins using a covalent inhibitor molecule tagged with biotin for pulldown (Maes et al., 2018), but the availability of suitable compounds is another limiting factor.

1.5.3. *In vivo* methods

To study epigenetic effector proteins in their endogenous context of modified chromatin and interacting epigenetic complexes, methods that are working in a native cellular system or even a whole organism must be applied. Many epigenetic mechanisms are highly conserved throughout evolution and thus allow for manifold model organisms being used in epigenetic research, such as *S. cerevisiae*, *Tetrahymena*, *C. elegans* or *Drosophila melanogaster* (Allis et al., 2007). However, differences in the abundance of histone PTMs and in the usage of similar effector proteins are significant (Garcia et al., 2007; Kouzarides, 2002), leaving only the mouse model as similar enough to the human system to be used in experiments that aim to reveal therapeutically relevant data. Epigenetic processes during development and differentiation are studied in

mouse embryonic stem cells (mESCs) (Scarola et al., 2015) or induced pluripotent stem cells (iPSCs) generated from mouse or human fibroblasts (Papp & Plath, 2013). Epigenetic signal transduction, mechanisms of chromatin modification and remodelling and cancer development are studied in immortalized cell lines from various tissues and will be regarded as *in vivo* experiments in this work.

Overview of existing methods

Fluorescence microscopy-based methods to study proteins in an endogenous cell context are widely used to monitor the abundance, localization, and dynamics of molecular interaction. As an improvement of simple colocalization microscopy using fluorescent antibodies, fluorescence resonance energy transfer (FRET) microscopy (Clegg, 1995; P. G. Wu & Brand, 1994) has been developed, in which the proteins of interest are biotechnologically coupled to fluorescent proteins or dyes with concerted excitation and emission spectra. The energy emitted from one FRET partner upon excitation can excite fluorescence in the other partner only if the proteins are in proximity, allowing for the detection of colocalization by measuring the signal emitted by the second partner with a resolution down to 10-100 Å.

A major drawback of FRET microscopy is the need to genetically engineer the cells before analysis to couple fluorophores to the proteins of interest. For samples that cannot at all or hardly be altered biotechnologically, like patient samples or hard to engineer primary cells, the proximity ligation assay (PLA) can be used as an alternative to FRET (Fredriksson et al., 2002; Söderberg et al., 2008). In PLA, two proteins of a complex are targeted by specific antibodies, each coupled to a different DNA probe. The DNA probes are ligated and amplified via a fluorescence-coupled rolling-circle amplification *in situ*. While PLA is a very sensitive method to detect protein interactions directly in a cell, it is rather complex to establish this method in respect to the informational output that can be expected, and it requires the availability of suitable antibodies for each target.

An ever-growing field in epigenetic research is the biotechnological application of protein reading domains for the detection or targeted editing of epigenetic signals. To this end, either DNA sequence-specific targeting domains (zinc-fingers, TALEs or dCas9; reviewed in Rots & Jeltsch, 2018) or epigenetic modification-specific reading domains (Jeltsch et al., 2020) can be applied. Fused to fluorescent proteins, the domains can be used for visualization of epigenetic modifications via fluorescence

microscopy. A more sophisticated method for the readout of epigenetic modifications at distinct genomic loci is the combination of epigenetic reading domains and bimolecular fluorescence complementation (BiFC) microscopy (Kerppola, 2006) to engineered bimolecular anchor detector (BiAD) sensors (Lungu et al., 2017). While BiFC alone can be used to visualize interactions of proteins, the BiAD system can be applied to monitor not only the steady state, but also dynamic changes in the co-occurrence of different epigenetic signals. However, both methods require some establishment before they can be applied and have limitations as regards their sensitivity.

Besides the detection of epigenetic modifications with fluorescence microscopy, chromatin reading domains can be used to investigate the downstream effects of chromatin effector complexes. Reading domains for DNA-methylation (Serre et al., 2009; Vucic et al., 2009) or histone PTMs (Kungulovski et al., 2016) are a specific and reliable alternative to polyclonal antibodies in chromatin immunoprecipitation (ChIP) experiments followed by qPCR or Illumina sequencing (ChIP-seq), that are performed to detect global changes of epigenome modifications (Kungulovski et al., 2016).

Epigenetic reporter systems

Synthetic biology has generated the development of various reporter gene systems, which allow for the direct readout of transcriptional changes of a synthetic promoter. Synthetic reporter genes have been used to study the induced activation of transcription (Biggar & Crabtree, 2001; Fiering et al., 1990; Ho et al., 1996), as well as the induced stable repression of genes (Ayyanathan et al., 2003). Advances in lentiviral technology (Mátrai et al., 2010) have enabled the creation of stable cell lines with the reporter genes integrated into the endogenous chromatin environment of host cells or organisms. Using the dynamic readout of reporter fluorophore expression, the dynamics and stability of epigenetic silencing and heterochromatin formation have been studied in different settings and with different recruitment methods (Gilbert et al., 2013; Hathaway et al., 2012; Keung et al., 2014). The various approaches can be differentiated by the site of expression of the reporter gene in the host. Depending on their localization on artificial chromosomes (Bintu et al., 2016; Tycko et al., 2020), their targeted integration into a specific genomic locus (Braun et al., 2017; Hathaway et al., 2012; Headley et al., 2019; Keung et al., 2014) or into a random genomic site (Gilbert et al., 2013), reporter gene expression is subject to different subsets of endogenous

regulatory mechanisms. These endogenous mechanisms function as barriers hindering the intended change of gene expression (Headley et al., 2019) and alter the dynamics and stability of the introduced chromatin modulation (Bintu et al., 2016). Although reporter gene-based methods have been extensively used to study the process of the transition between active and repressed chromatin states, many questions remain unanswered. The transcriptional response of reporter gene repression occurs at different dynamics and sometimes remains incomplete (Bintu et al., 2016; Headley et al., 2019; Tycko et al., 2020; Yeo et al., 2018). How are the barriers preventing the induced silencing of transcription composed? And which endogenous coregulators need to be co-recruited to achieve efficient silencing of gene expression? These questions were challenged during the work on this thesis, by combination of the synthetic reporter gene approach with a comprehensive RNAi library targeting 1104 chromatin related genes (Rathert et al., 2015).

1.5.4. A novel approach to capture the complexity of epigenetic gene regulation

Since the proposal of the "histone code", the model that gene expression programs are encoded in the form of covalent HPTMs subsists (T. Jenuwein & Allis, 2001). However, it is now clear that the information that is packaged in the histone code is not just transmitted as a cumulative readout of single histone modifications. Instead, the biological outcome of a code can vary, depending on the individual combination with other HPTMs, with chemical modifications of DNA, the presence of regulatory RNAs and the formation of large structural domains by condensation effects. The transcriptional activity of a genomic locus is seen as a statistical event, which is controlled by this complex chromatin environment, the availability of TFs and the presence and status of *cis*-regulatory elements (Nicolas et al., 2017; Tycko et al., 2017). To complement the intricacy of the "input" information provided by the chromatin environment, epigenetic transcriptional regulators have evolved to work in complexes of several domains providing reading-, writing-, erasing- or remodelling function. Depending on the local histone code, singular complex partners can be exchanged to adapt the regulatory activity of the complex. The regulatory capacity of this vast network of epigenetic factors is enormous, and the complexity poses a great challenge for the investigation of single nodes or connections inside this network.

To set about this challenge, a comprehensive RNAi screening approach was combined with the recruitment of an epigenetic repressor domain to a synthetic reporter gene in the presented work. Several high-throughput screening approaches have been developed to capture the intricacy of biological pathways and regulatory networks. Genetic screenings can be performed using RNA interference (RNAi) with small hairpin RNAs (shRNAs), which reduce the expression of specific target genes by induction of mRNA cleavage and degradation (Martin & Caplen, 2007). To apply this method for genome-wide high-throughput screenings, the RNA scaffolds were improved and integrated into lentiviral transduction vectors, to be able to stably integrate them into the genome of mammalian cells (Brummelkamp et al., 2002; Fellmann et al., 2013; Pelossof et al., 2017; Zuber, McJunkin, et al., 2011). Combined with the lentiviral technology, extensive shRNA and CRISPR-RNA libraries are being used for the high-throughput screening of novel therapeutic targets in different cancers (S. Chen et al., 2015; Shalem et al., 2014; Tzelepis et al., 2016; Zuber, Shi, et al., 2011), of nodes in regulatory networks (Parnas et al., 2015) or of functional non-coding RNAs (S. J. Liu et al., 2017). These approaches constantly provide new insights into the tumorigenesis and resistance development of cancer cells (Rathert et al., 2015).

The measurable output of common screening methods is a change in the proliferative activity or in the grade of differentiation of the monitored cells. By introduction of a targetable fluorescent reporter gene, the key innovation of the presented work was the direct coupling of an RNAi screen to a discrete transcriptional output of a gene that is confidently regulated by the chromatin effector of interest. Transcriptional regulation is a complex and dynamic process (Tycko et al., 2017). To be able to experimentally describe isolated mechanisms of transcriptional regulation, it is essential to have robust knowledge about the surrounding influence factors. The presented approach aimed to master a balancing act between the reduction of complexity and the comprehensive analysis of regulatory interactions. On the one hand, it was aimed to reduce the variability of input factors leading to a transcriptional signal by installation of a controlled system with an easy read-out. At the same time, it was essential to sustain the integration of the reporter gene and the analysed chromatin regulator into the complex endogenous epigenetic network of eukaryotic cells. In doing so, it was anticipated to identify and characterize novel functional dependencies of chromatin regulators in their activity at a transcribed gene.

2.   Principal aims of the study

Epigenetic research faces the challenge of the high complexity and tight coregulation of effectors in epigenomic networks. Although many isolated mechanisms of epigenetic gene regulation are well-described, solid methods for the comprehensive analysis of specific processes as parts of the bigger epigenomic network are missing. The aim of this work was to expand the toolbox of existing methods by an approach that will help to capture and describe the complexity of transcriptional regulation.

2.1.   Development of a modular reporter system for the dynamic investigation of epigenetic complexes

The regulation of transcription by the induced modification of chromatin structure is a dynamic process (Amabile et al., 2016; Bintu et al., 2016; Tycko et al., 2017). To comprehensively describe this process, experimental approaches must be able to trace various types of input information, e.g. about interactions of chromatin regulators, the presence of specific epigenetic marks, transcription factor concentrations, and the location of the gene within the nucleus. However, all approaches that are available so far have strengths in the integration of certain input signals, while they fail to detect other factors completely. The extensive collection of experimental data remains an essential premise for the eventual generation of a predictive theoretical framework for eukaryotic gene regulation. To expand the available experimental toolbox, the aim of this work was to create a robust reporter system for transcriptional activity, with an easy and dynamic readout and the ability to compensate for cell-to-cell variances of transcription. The system was designed to enable the induced recruitment of a chromatin regulator to a fluorescent reporter gene, followed by the detection of transcriptional changes using flow-cytometry. It was aimed to integrate the reporter gene into an endogenous chromatin environment to allow for the detection of regulatory dependencies of the investigated chromatin regulator on endogenous cofactors. The reporter system was planned to have a modular design. Like this, the application of the method for the investigation of different chromatin regulators should be enabled. The goal was to optimize the reporter system in a way that it could be used for the characterization of catalytic mutants of epigenetic regulators. Once the sensitivity of the system to report on the effect of catalytic activities on transcriptional regulation would be validated, it was planned to analyse whether it could also be used to test small molecule inhibitors of chromatin effectors.

The most important novelty of the planned reporter system would be the ability to detect the dependency of chromatin effector proteins on the recruitment of endogenous coregulators. The capacity of the system to provide this function was planned to be validated using RNAi.

Eventually, it was aimed to obtain a robust fluorescent reporter system that could be applied for the analysis of the dynamics of chromatin conformation changes, of functional interactions of chromatin regulators inside eukaryotic regulatory networks, and for the analysis of small molecule inhibitors. If all these goals could be achieved, the reporter would present a valuable tool for the collection of experimental data that is needed to build a model for epigenetic transcriptional regulation.

2.2. Identification and characterization of novel interactors of LSD1 as a highly networking transcriptional repressor

The application of high-throughput methods represents a suitable way to meet the immense number of functional interactions that constantly take place in the epigenetic network with an equally high number of descriptive data points. The lysine specific demethylase LSD1 constituted an interesting candidate for the analysis of functional relationships in this network, as it is known to exert its repressive activity in cooperation with multiple coregulatory complexes (Maiques-Diaz & Somervaille, 2016). The aim was to identify novel coregulators of LSD1 as a transcriptional repressor by combining the reporter system with a comprehensive RNAi screen. Using this approach, it was planned to screen 1104 chromatin related genes for their connection to LSD1 induced transcriptional repression. The strongest candidate emerging as a coregulator in this screen should be validated, and the functional interaction with LSD1 should be characterized comprehensively. Using various biochemical methods, it was aimed to fully describe the succession of molecular mechanisms that need to be activated in a cell to induce the stable silencing of a reporter gene.

2.3. Analysis of the repressive complexes forming around the KRAB domain in different cell lines

The KRAB domain of ZNF10 is commonly used as a repressive domain in synthetic biology and CRISPRi approaches (Alerasool et al., 2020; Bintu et al., 2016; Tycko et al., 2020). However, the mechanisms and functional interactions mediating the installation of heterochromatin following KRAB binding to a genomic locus are ill-defined, leading to unsatisfactory or inexplicable results in epigenome editing approaches using this domain (Alerasool et al., 2020; Yeo et al., 2018; Ying et al., 2015). The aim of this project was to combine the synthetic reporter system with RNAi to characterize the mechanisms of KRAB induced transcriptional repression, and to detect variances in repressive complex formation between different cell lines. The results from this project could help to improve the efficiency of epigenome editing approaches using the KRAB domain, by revealing cell-type specific modes of KRAB induced transcriptional silencing.

3. Materials and methods

Most of the methods described in this section are also described in the methods section of **Appendix 1**. The descriptions in this thesis provide additional context and detail compared to the published methods.

3.1. Cloning

The fluorescent reporter constructs used in this work were assembled in a modular way, to build a reporter cassette consisting of an array of TetR binding sites (*tetO*), a variable promoter sequence (*synP*) and either *mCherry* or *turboGFP* as reporter fluorophore (**Figure 15** and **Appendix Figure 2**). Expression of the reporter fluorophore was coupled to an antibiotic resistance (Blasticidin or Puromycin) via a *P2A* sequence to allow for antibiotic selection of cells expressing the reporter. The reporter cassette was cloned into a pMSCV (Addgene #75085) vector backbone using Gibson Assembly® (D. G. Gibson et al., 2009) to enable retroviral integration into the target cell lines. Following optimization of the modular reporter cassette, further experiments were performed using a reporter construct with 6x *tetO* sites upstream of a constitutive *EF1A* promoter driving the expression of *mCherry*. The resulting fluorescent reporter construct is called *synP-mCherry* throughout this work and the attached publication.

The LSD1 construct was kindly provided by Tim Somervaille. All *rTetR-* and *TetR-* effector fusion constructs were cloned into a *pRRL* lentiviral vector backbone under control of a *SFFV* promoter (**Appendix Figure 1**). Expression of the fusion constructs was coupled to a Hygromycin resistance gene via a *P2A* sequence using standard cloning methods. All effector domains and full-length proteins used in this study are listed in **Table 2**.

For cloning of the plasmids used for RNAi, 97-mer oligonucleotides containing the target-specific 22 bp guide sequences were ordered as ultramers from Integrated DNA Technologies (www.idtdna.com). The oligonucleotides were amplified using Platinum Pfx DNA Polymerase (Thermo Fisher Scientific) and primers adding EcoRI and XhoI restriction sites. The PCR products were digested with EcoRI-HF and XhoI and introduced into the SGEN vector (Addgene #111171) containing a mirE shRNA cassette (Fellmann et al., 2013) via T4 DNA ligase (NEB) reaction. The SGEN vector is a construct for the lentiviral transduction of cells, on which the expression of the mirE

shRNA cassette is driven from a *SFF*V promoter and coupled to the expression of *eGFP*. The identity of all constructs was confirmed using Sanger sequencing (Microsynth SEQLAB).

Table 2: Overview of chromatin regulators used in this study. The accordance of each cloned construct with the respective sequence listed under the CCDS accession number was verified on amino acid level using Sanger sequencing.

chromatin regulator	domain	CCDS #	sequence length
LSD1	full-length	30627.1	aa 0-852
ZNF10	KRAB domain	9283.1	aa 0-98
DNMT3A	methyltransferase domain	1718.2	aa 612-912
EHMT1	preSET and SET	7050.2	aa 1002-1296
CBX3	full-length	5398.1	aa 0-182
EED	full-length	8273.1	aa 0-442

3.2. Cell culture, lentiviral transduction and flow cytometry

NIH/3T3 cells (ATCC® CRL-1658™), iMEF cells (kind gift of Prof. Thomas Jenuwein (MPI Freiburg)), and the retroviral packaging cell lines Lenti-X 293T and Platinum-E were cultivated in DMEM high glucose media (Sigma-Aldrich) supplemented with 10% FBS, 4 mM L-glutamine, 1 mM sodium pyruvate, 10 mM HEPES (pH 7.3), 50 U/ml penicillin and 50 mg/ml streptomycin in an incubator providing 37°C and 5% CO_2. For cultivation of iMEF cells, 0.1 mM β-mercaptoethanol and 1x non-essential amino acid solution (Sigma-Aldrich) were additionally added to the medium. The cell lines K562, murine MLL–AF9;NrasG12D AML cells, and MCF7 were cultivated in RPMI-1640 medium (Sigma-Aldrich) supplemented with 10% FBS, 4 mM L-glutamine, 1 mM sodium pyruvate, 10 mM HEPES (pH 7.3), 50 U/ml penicillin and 50 mg/ml streptomycin.

For retroviral packaging of pMSCV vectors, 20 µg of plasmid were precipitated for 20 min in HBS buffer (140 mM NaCl, 25 mM HEPES, 0.75 mM Na_2HPO_4, pH 7.0) together with 125 mM $CaCl_2$ and 10 µg GagPol helper plasmid. The mix was added to a 10 cm dish with Platinum-E cells growing at 75-85% confluence in supplemented DMEM. After 16 h, the media was replaced with fresh supplemented DMEM. 24 h after transfection with calcium-precipitated DNA, the medium was replaced with 5 ml of

target cell medium. Supernatant containing the virus was collected 40-50 h after transfection, filtered through a 0.45 µm filter and added to the target cells at 50-70% confluence. To increase transduction efficiency, 4 µg/ml polybrene (SIGMA) were added to the medium. To obtain single integrations of viral constructs into the target cell genome, a final multiplicity of infection (MOI) of 10% of cells was desired. The MOI was adjusted by either further dilution of the virus in target cell medium or by performing multiple rounds of infection on the same cells. MOIs were tested via flow cytometry 2 days after transduction before starting antibiotic selection. Antibiotic selection with 7.5-10 µg/ml Blasticidin or 2-3 µg/ml Puromycin (concentrations were adjusted for each cell line after selection control tests) was performed for 7 days.

For retroviral packaging of pRRL-vectors for infection of murine cell lines, plasmids were mixed with helper plasmids pCMVR8.74 (Addgene #22036) and pCAG-Eco (#35617) and 3x w/w excess of polyethyleneimine 25K (PEI) in serum free DMEM. The mix was added to Lenti-X cells residing in 2 ml supplemented DMEM at 75-90% confluence. Media exchanges and transduction of target cells were performed as described for pMSCV. Cells expressing pRRL-rTetR-effector-P2A-Hygro were selected with 500 µg/ml Hygromycin and cells expressing SGEN with 2.5 mg/ml Neomycin for 7 days. For creation of stable human cell lines, cells were first transduced with a plasmid expressing the *mouse high affinity cationic amino acid transporter 1a*, to act as a surface receptor for the ecotropic murine retroviral leukemia virus produced by the method described for murine cells. Using pCMV-VSV-G (Addgene #8454) as the envelope plasmid for transduction of human cells requires working under biosafety level S2 conditions. Virus particle production was performed as described for Lenti-X and human cells were infected 2-3 times to achieve high MOI before selection with 3 µg/ml Puromycin for 7 days, gaining cell lines that can be transduced with the murine pCAG-Eco and handled at biosafety level S1. Recruitment of the rTetR-fusion proteins was started 12 days after transduction with SGEN by treatment with 1µg/ml Doxycycline. Expression of GFP and mCherry was analysed every 1-3 days using a MACSQuant Vyb flow cytometer. Laser settings were optimized for each stable cell line and kept constant over the course of the experiment. Flow cytometry data were analysed using FlowJo[©].

Construction of stable cell lines using lentiviral transduction

Lentiviral vectors are a useful tool for the genetic manipulation of cells and have been optimized for the application in various biological settings (Mátrai et al., 2010). As opposed to γ-retroviral vectors (γ-RVs), lentiviral vectors can also be used to transduce non-dividing as well as dividing cells and have a higher packaging capacity of up to 10 kb. Commonly used lentiviral vector systems are derived from the human pathogen HIV-1, and multiple adjustments of the design have been conducted to ensure biosafety of the method (**Figure 14**). The HIV elements *Gag* and *Pol*, coding for viral matrix, capsid, and nucleocapsid components and reverse transcriptase and integrase components, respectively, are binned on an extra plasmid called packaging plasmid (**Figure 14a**) (Vigna & Naldini, 2000). In third generation constructs, the *Rev* sequence facilitating nuclear export of viral transcripts is outsourced to another plasmid (Figure 14b). The envelope plasmid expresses viral surface glycoproteins determining tropism of the final particle, e.g. the vesicular stomatitis virus G glycoprotein (VSV-G) for transduction of human cells, under a strong CMV promoter. The viral Ψ- and LTR elements needed for packaging of the sequence into a virus particle and reverse transcription and integration into the host genome, respectively, are only present on the transfer plasmid containing the gene of interest. This ensures that the viral elements coded on the packaging and envelope plasmids are not present in the viral particles used for transduction of target cell lines. Deletion of several HIV genes that are essential for replication limits the innate replicative capacity of the engineered virus. In doing so, the ability for replication of the transgene before encapsulation is restricted to the presence in an engineered cell line like HEK293T (ATCC® CRL-3216™) (Vigna & Naldini, 2000).

In biotechnological research, the lentiviral system is used for genetic engineering (S. Chen et al., 2015) and epigenome editing (Alerasool et al., 2020) of mammalian cells, for generation of induced pluripotent stem cells (IPSCs) (Sommer et al., 2009) and for comprehensive genetic screens (Shalem et al., 2014; Tzelepis et al., 2016; Zuber, Shi, et al., 2011). Although slightly more complex in design and handling, the creation of stable cell lines using lentiviral gene transfer is at an advantage in many ways in comparison to introduction of genes via transient transfection. Lentiviral gene transfer allows the creation of cell lines stably expressing a constant and low amount of a gene of interest, which can be used for long term experiments while minimizing toxic effects

from overexpression. mRNA copy numbers are much more comparable to endogenous expression levels than in transient transfection, with up to several thousand copies per cell (Fliedl et al., 2015). By fine-tuning transduction conditions and/or the establishment of clonal colonies by single-cell sorting, homogenous cell populations expressing an optimal amount of target protein can be established. Finally, gene expression from the viral construct can be put under the control of an inducible promoter, further reducing limitations by toxicity of the gene of interest.
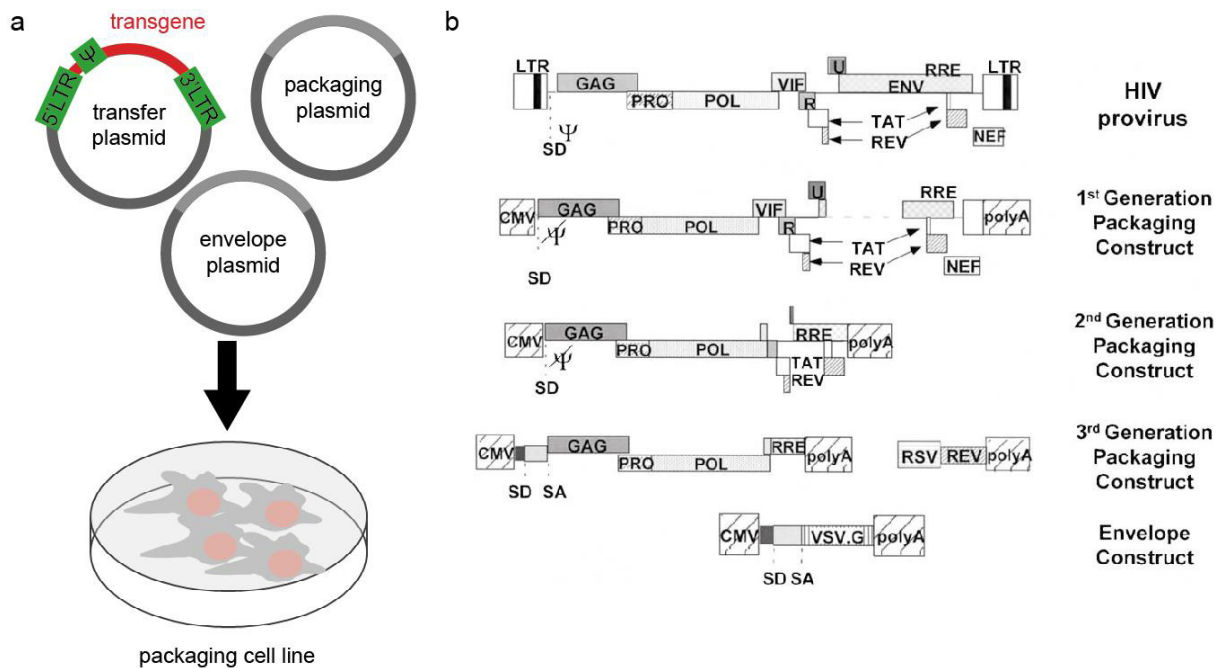
Figure 14: Overview of the lentiviral virus production system. a, Schematic view of the elements needed for the production of an infectious virus particle in a second generation lentiviral system. The transfer plasmid contains the transgene of interest and viral elements needed for expression and packaging of the sequence. Further viral precursor proteins are coded on the packaging plasmid. The envelope plasmid determines the tropism of the virus particle. Replication of the viral sequences is restricted to their expression in engineered packaging cell lines. **b,** Overview of the engineering process to increase the biosafety of viral constructs. Based on the HIV-1 provirus, several generations of packaging constructs have been developed. Packaging plasmids provide all elements needed for production of the virus particle, while LTRs and the Ψ-element are deleted. This ensures that the helper plasmid can neither be replicated, nor be packaged into the virus. The envelope plasmid provides target-cell specific surface proteins. Image taken from (Vigna & Naldini, 2000).

3.3. Pooled RNAi screening

To find novel coregulators of LSD1 during its activity as a transcriptional repressor, a comprehensive RNAi screen was performed in NIH/3T3 cells expressing the *synP-mCherry* reporter and rTetR-LSD1. The SGEN shRNA-mirE library consisting of 5459 shRNAs targeting 1010 chromatin-associated murine genes was kindly provided by Johannes Zuber. The method is also described in the methods section of **Appendix 1**. The RNAi screen for coregulators of KRAB in iMEF cells was performed accordingly, with slight adaptations that are described in the results section.

3.3.1. Cell culture and cell sorting

60 µg of the library plasmid pool was used to produce a batch of lentiviral supernatant for the transduction of NIH/3T3 *synP-mCherry* reporter cells expressing rTetR-LSD1, following the protocol described in section 3.2. To ensure a library representation of statistically >500 cells receiving each shRNA, while providing conditions that lead to the uptake of maximum one retroviral plasmid per cell, a total of 30 million cells were infected with 10% transduction efficiency. Two days after transduction, the cells were harvested, united and then split into three replicates for independent selection with 2.5 mg/ml Neomycin for 7 days. After selection, T_0 samples to test representation of shRNAs in the starting population were taken, and the treatment with 1µg/ml Doxycycline (DOX) was started. Throughout selection and DOX treatment, a minimum of 3×10^6 cells per replicate was maintained at each passage to preserve library representation. The silencing progress induced by recruitment of rTetR-LSD1 was monitored by flow cytometry analysis of the *mCherry* signals every 2-3 days.

After 14 days of DOX treatment, cells were sorted into *mCherry*-positive and *mCherry*-negative populations using a FACS Aria III. The *mCherry*- positive population was defined by gating the top 6-8% of the *mCherry* expressing population in each replicate, the *mCherry*-negative population by gating the lowest 75-80% of *mCherry* expressing cells in each replicate. To retain representation of all shRNAs enriched in the respective populations, a minimum of 5×10^5 cells was sorted for the *mCherry*-positive sample, a minimum of 6×10^6 cells for the negative sample. The 10-fold difference in the chosen cut-offs was motivated by the assumption that only a fraction of shRNAs would influence LSD1-mediated silencing and thus enrich in the *mCherry*-positive population.

A biologically and technically independent replicate was performed by repeating the whole process from virus production to cell sorting, resulting in two additional replicates that were analysed in parallel with the first three replicates in the final analysis.

3.3.2. Preparation of an Illumina sequencing library

For preparation of a DNA library compatible with standard Illumina sequencing, genomic DNA was extracted from the sorted cell populations. Cells were lysed by resuspending cell pellets in extraction buffer (10 mM Tris pH 8, 150 mM NaCl, 10 mM EDTA), adding 10% SDS and 80 µg Proteinase K (NEB) and incubating for 24 h at 55°C. DNA was isolated in three rounds of phenol extraction using Phase Lock Gel™ tubes (Quantabio) followed by precipitation of DNA overnight in 70% ethanol at -80°C. DNA was collected by centrifugation at 15,000xg for 30 min at 4°C, the pellet was washed once with 70% ethanol and air-dried at room temperature. The DNA pellet was resuspended in elution buffer (10mM Tris pH 8.5). To add random barcodes and sample barcodes to the shRNA sequences isolated with the genomic DNA of the sorted populations, primers binding in the mirE sequence were designed (MM2P51_for and MM2P71_rev, see **Table 3**).

Random barcodes were added to enable identification of DNA-sequences enriched due to amplification bias, sample barcodes to enable pooled amplification of all samples in one Illumina sequencing run. For each sample, DNA from at least 10^6 cells was used as template in multiple parallel 50µL PCR reactions, each containing 1 µg template, 1x AmpliTaq Gold buffer, 0.2 mM of each dNTP, 2 mM MgCl2, 0.3 µM of each primer and 1.25 U AmpliTaq Gold Polymerase (Life Technologies). Each sample was amplified with an individual sample barcode primer using the following cycling parameters: 95 °C for 10 min; 28 cycles of {95 °C for 30 s, 54 °C for 45 s and 72 °C for 60 s}; 72 °C for 7 min. PCR products were combined for each sample, purified from a 1% agarose gel and 20 ng per sample were transferred to a second round of PCR. In this PCR, standard Illumina P7 adaptors and the Illumina N708 index were added to all sequences using primers MM2P52_for and MM2P72_rev_N708 (total product length = 428 bp). Cycling parameters were the same as for PCR1, only using 10 ng template per reaction and 6 cycles of amplification.

The final libraries were resolved on a 1% agarose gel, cleaned up using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel), pooled in equimolar amounts and analysed on an Illumina HiSeq 3000 deep sequencer. Read length was 150 bp, covering the sample barcodes and the 22 nucleotides of the shRNA guide strand. Sequencing was performed using standard Illumina primers.

Table 3: List of primers used for the amplification of shRNA sequences for Illumina sequencing.

name	sequence	purpose
MM2P51_for	ACGACGCTCTTCCGATCT**NNNNNNYYYY**TAGT GAAGCCACAGATGTA	addition of random barcodes (**N**) and sample barcodes (**Y**) in PCR1
MM2P71_rev	CAGACGTGTGCTCTTCCGATCTTGGATGTGGA ATGTGTGCGAGG	
MM2P52_for	AATGATACGGCGACCACCGAGATCTACACTCT TT CCCTACACGACGCTCTTCCGATCT	addition of Illumina P7 adaptors and Illumina standard index N708
MM2P72_rev_ N708	CAAGCAGAAGACGGC ATACGAGATCAGAGAGGGTGACTGGAGTTCA GACGTGTGCTCTTCCGATC	

3.4. Chromatin immunoprecipitation (ChIP)

3.4.1. Preparation of *Drosophila melanogaster* spike-in chromatin

To enable sample-to-sample normalization in ChIP experiments, spike-in of chromatin from a different species before IP has been established (Bonhoure et al., 2014). Spike-in chromatin should meet the following criteria: The chromatin modification of interest has to be present as a similar epitope in both the tested and the spike-in species, the species have to be different enough to enable identification of spike-in material in the enriched sequences, and ChIPseq data for the modification of interest in the spike-in organism must be available to be able to choose control regions for the specific enrichment of sequences carrying the epitope. Chromatin isolated from the *Drosophila melanogaster* Schneider 2 cell line (S2) has been shown to serve well as a spike-in control for ChIP on mammalian cells (Egan et al., 2016). To isolate chromatin, S2 cell pellets (kindly provided by Pavel Bashtrykov) were resuspended in 250 µl TM2+ buffer (10 mM Tris pH 7.4, 2 mM MgCl$_2$, 0.5 mM PMSF, EDTA-free protease inhibitor) per 10^6 cells and incubated on ice for 1 min, before adding NP-40 to a final concentration of 0.6% and incubating on ice for another 5 min. Cells were spun down at 1000xg for 10 min at 4°C and the supernatant was discarded. The pellet was washed gently with 250 µl TM2+ and nuclei were sedimented by centrifugation at 1000xg, 10 min, 4°C.

The nuclei were resuspended in 100 µl TM2+ buffer and absorption at 260 nm was determined using NanoDrop. The suspension was pre-heated to 37°C for 2 min before adding 1 mM CaCl₂ and 400 U Micrococcal Nuclease (NEB) per 10 µg chromatin, vortexing and incubating at 37°C for 5 min. The MNase reaction was stopped by placing the samples on ice and immediately adding 2 mM EGTA, 300 mM NaCl and 0.1% TritonX. The samples were vortexed and cleared by centrifugation at 13,000xg, 10 min, 4°C. Fragmentation of chromatin was analysed by digesting a part of the sample with 40 µg Proteinase K (NEB) for 2 h at 54°C, cleaning up the DNA fragments with the Macherey Nagel Nucleospin Mini Kit and resolving on a 1.5 % agarose/TPE gel.

3.4.2. Crosslinked ChIP on mononucleosomes

For analysis of changes in histone modifications at the synthetic promoter, mononucleosomes were isolated from NIH/3T3 reporter cells after 4 and 14 days of DOX treatment. Cells were grown to 70-90% confluency in a 10 cm dish. After washing once with 10 ml PBS (Sigma-Aldrich), cells were crosslinked with 1% formaldehyde (Thermo Fisher Scientific) in PBS for 15 min at room temperature. Crosslinking was quenched with 200 mM glycine for 5 min. Cells were washed twice with 10 ml PBS and harvested with a cell scraper in 10 ml fresh PBS. Cells were pelleted by centrifugation at 700xg for 5 min and washed once more with PBS. For long-term storage, 500 nM Trichostatin A (TSA) was added to the cells before freezing at -80°C. For fractionation of chromatin into mononucleosomes, 5×10^6 cells were lysed in 125 µl lysis buffer (10 mM Tris pH 7.4, 2 mM MgCl₂, 0.6 % Igepal-Nonidet P40, 0.5 mM PMSF, 1 mM DTT, 5 mM sodium butyrate) with protease inhibitors for 15 min on ice.

Chromatin was fractionated with 300 Units Micrococcal nuclease (MNase)(NEB) for 16 min at 37°C. MNase digestion was stopped by adding 10 µM EDTA, 0.1% TritonX-100 and 0.1% sodium deoxycholate. Samples were incubated on ice for 15 min before taking samples for analysis on a LabChip® GX Touch™ Nucleic Acid Analyzer (Perkin Elmer). Samples were diluted by adding 800 µl Complete IP buffer (20 mM Tris pH 8, 2 mM EDTA, 150 mM NaCl, 0.1 % Triton X, 1 mM PMSF, 5 mM sodium butyrate) with protease inhibitors. Cellular debris was removed by spinning down the samples for 10 min, 15,000xg, 4°C. The supernatant was snap-frozen in aliquots of 70 µg chromatin and stored at -80°C.

For IP, one aliquot per reaction condition was thawed and 5% of each sample were taken aside as input. For pre-clearing of the samples, 20 µl of Dynabeads Protein G (ThermoFisher) and 5 µg rabbit or mouse IgG were added, depending on the species of the IP antibody. Samples were pre-cleared by rotation at 4°C for 2 h. The beads were removed using a magnetic rack and the supernatant was split into equal volumes for target IP and IgG control. 2.5 µg IP antibody or IgG were added to the samples and binding of the antibodies was performed overnight at 4°C. 20 µl Protein G beads per sample were pre-blocked overnight in 0.1 % BSA in Complete IP buffer. In the morning, the pre-blocked beads were collected using a magnetic rack, the blocking solution was removed, and the beads were resuspended in fresh Complete IP buffer. The beads were split equally to the IP and IgG samples and antibody-bead binding was performed for 2 h at 4°C with rotation. The target-antibody-bead complexes were collected using a magnetic rack and the unbound supernatant was discarded. The beads were washed once with IP Wash 1 buffer (20 mM Tris pH 8, 2 mM EDTA, 50 mM NaCl, 1% TritonX, 0.1% SDS), twice with High Salt buffer (20 mM Tris pH 8, 2 mM EDTA, 500 mM NaCl, 1% TritonX, 0.01% SDS), once with Lithium Chloride buffer (10 mM Tris pH 8, 1 mM EDTA, 0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate) and twice with TE buffer (10mM Tris pH 8, 1 mM EDTA). DNA fragments were eluted from the beads in two rounds of incubation with 1% SDS/100mM $NaHCO_3$. De-crosslinking and removal of RNA in the eluates and the input samples was performed at 65°C with 2 µg RNAse A for a minimum of 16 h. Proteins were removed from all samples by digestion with 60 µg Proteinase K for 2 h at 45°C. DNA fragments were purified using a Chromatin IP DNA purification Kit (Active Motif) and analysed via quantitative PCR in a CFX96 Real-time PCR detection system (Bio-Rad).

3.4.3. High sensitivity ChIP and library preparation for LSD1 ChIPseq

High sensitivity ChIP

As LSD1 binding to chromatin in the nucleus is in great parts indirect and mediated by interaction partners, a super sensitive ChIP method optimized for proteins with low binding affinities was used for LSD1 ChIPseq. To this end, the ChIP-IT High Sensitivity® Kit (Active Motif) was used following manufacturer´s protocol. In brief, 2 x 10^7 NIH/3T3 cells were fixed with 1% formaldehyde solution for 15 min. Crosslinking was stopped using the provided Stop Solution and cells were harvested using a cell scraper. Cell pellets were washed twice in ice-cold PBS and resuspended in 5 ml per

10^7 cells of Chromatin Prep Buffer supplemented with PIC and PMSF. Cell membranes were lysed in a chilled dounce homogenizer. Nuclei were separated from the lysate by centrifugation and lysed in ChIP Buffer supplemented with PIC and PMSF.

Nuclei and chromatin were fractured by sonication in a EpiShear™ probe sonicator (Active Motif) for a total of 20 min in intervals of 20 sec pulse/ 30 sec pause at 40% amplitude with incubation on ice. Fractionation of chromatin was analysed on 1.5% agarose gel and a LabChip® GX Touch™ Nucleic Acid Analyzer (Perkin Elmer) after purification of a sample fraction according to the protocol.

To maximize the amount of chromatin precipitated with the LSD1 antibody, immunoprecipitation was performed in 3 parallel reactions using 25 µg input chromatin each. The chromatin samples were incubated with 4 µg LSD1 antibody (ab17721, Abcam) each overnight at 4°C. Binding of 30 µl Protein G agarose beads per sample was performed for 3 h at 4°C.Washing and elution of the beads was performed as described in the protocol. During the column loading step, reaction 1 and 2 were combined on one column to concentrate the final DNA yield. Elution was performed twice with 25 and 30 µl of elution buffer. Eluted DNA fragments were analysed on a LabChip® GX Touch™ Nucleic Acid Analyzer.

Library preparation

For preparation of Illumina sequencing libraries, the NEBNext® Ultra™ II DNA Library Prep with Sample Purification Beads Kit (NEB) was used following the manufacturer´s protocol. In brief, 1 µg of input DNA and 50 µl of cleaned up ChIP sample were subjected to the 5´phosphorylation and dA-tailing step before ligation with the NEBNext Adaptors (use adaptors undiluted for input, diluted 1:25 for ChIP sample). No size selection was performed. For addition of Illumina indices, the input was amplified in 3 cycles of PCR with the standard Illumina primers i705 and i503 (taken from NEBNext Multiplex Oligos Kit for Illumina) and the ChIP sample was amplified in 11 cycles of PCR using i706 and i504 (taken from the NEBNext Multiplex Oligos Kit for Illumina). The PCR reaction was purified following the manufacturer´s protocol and samples were analysed on a LabChip® GX Touch™ Nucleic Acid Analyzer. The samples were analysed on an Illumina HiSeq 3000 deep sequencer using standard Illumina primers with a read length of 150 bp.

3.5.   Immunofluorescence microscopy

3.5.1. Sample preparation

For immunostaining of nuclear epitopes with specific antibodies followed by fluorescence microscopy, NIH/3T3 cells were cultivated on microscopy cover slips to a confluency of 70-90%. The cover slips were washed three times for 5 minutes with 2ml of room-temperature PBS$^{Ca2+/Mg2+}$ before incubation with 4% paraformaldehyde (PFA) in PBS for 10 min at RT to crosslink proteins and nucleic acids. Cells were washed again as described and cellular and nuclear membranes were permeabilized for the antibodies by incubating the slides in 0.5% TritonX in PBS$^{Ca2+/Mg2+}$ for 5 min at 4°C. After repeated washing, unspecific antibody binding sites were blocked by incubating the coverslips for 1h in 5% milk powder in PBS$^{Ca2+/Mg2+}$ or 5% BSA in PBS$^{Ca2+/Mg2+}$, depending on the downstream antibody. Primary antibodies were diluted in the respective blocking medium to a concentration of 0.5–3.3 µg/ml and coverslips were incubated with the antibody dilution at 4°C overnight. To remove unspecific binding of the primary antibodies, after incubation the cover slips were washed 3 times for 5 min with PBS$^{Ca2+/Mg2+}$, before incubation with the secondary Alexa-Fluor® coupled antibodies at a concentration of 1 µg/ml in the respective blocking solution for 2 h at RT. The cells were washed again as described to remove unspecific binding of the antibodies. In the second washing step, 1 µg/ml DAPI was added to the PBS$^{Ca2+/Mg2+}$ to stain heterochromatic regions as a nuclear marker. The coverslips were mounted on microscopy slides using either Mowiol® solution or Molecular Probes™ ProLong™ Gold Antifade Mountant. To obtain ideal fluorescence images, cells were imaged 18-72 h after mounting.

3.5.2. Image acquisition and analysis

For quantification of S9.6 staining, samples were analysed on a Zeiss Axio Observer.Z1 microscope equipped with a Plan-Apochromat 63x/1.40 Oil DIC M27 objective and an AxioCam MRm camera. The following excitation and emission filters were used: Blue channel: excitation filter 335-383 nm, emission filter 420-470 nm; red channel: excitation filter 538-562 nm, emission filter 570-640 nm; green channel: excitation filter 450-490 nm, emission filter 500-550 nm. Z-stacks covering the whole nucleus were acquired applying an interval of 450 nm, and images were subjected to deconvolution using a constrained iterative algorithm and the ZENblack software

(Zeiss), before generating maximum intensity projections. Quantitative image analysis was done with CellProfiler™ version 2.2. Nuclei were identified via the DAPI staining.

3.6.   DNA:RNA hybrid immunoprecipitation (DRIP)

For analysis of the presence of DNA:RNA hybrids at genomic loci, DNA:RNA hybrid immunoprecipitation (DRIP) followed by qPCR was performed. NIH/3T3 cells expressing the synthetic reporter construct were cultivated at densities ensuring constant exponential growth and harvested at a confluency of 80-100% in a 10 cm culture dish ($\sim 10 \times 10^6$ cells). Adherent cells were washed twice with 10 ml PBS to remove medium and serum, before crosslinking cellular proteins and nucleic acids by addition of 1% methanol-free formaldehyde in PBS for 10 min at RT. Crosslinking was quenched by addition of 500 mM glycine and shaking for 5 min at RT. The cells were washed again twice with 10 ml PBS to remove formaldehyde and glycine and harvested by scraping with a cell scraper in 5 ml fresh PBS. Crosslinked cells were collected by centrifugation for 8 min at 600xg and washed again carefully with 10 ml PBS to remove residual formaldehyde. Before the final collection step, cells were split into aliquots of 2-3 x 10^6 cells and spun down again at 600xg. PBS was removed and the pellets were either snap-frozen in liquid nitrogen before storing at -80°C or directly subjected to cells lysis. For extraction of nucleic acids, cells were lysed in 300 µl DRIP lysis buffer (50mM HEPES-KOH at pH 8, 140 mM NaCl, 1 mM EDTA at pH 8, 1% Triton X-100, 0.1% Na-Deoxycholate, 1% SDS) and incubated on ice for 30 min with repeated homogenization of the sample with a pipette tip to achieve complete lysis of the nuclei. Chromatin was fragmented by sonication of the samples in an Active Motif EpiShear™ probe sonicator (Active Motif) for a total of 4 min in repeated cycles of 20 sec pulse/30sec pause at 40% amplitude. For digestion of free RNA strands in the samples, chromatin was treated with 50 µg RNase A per 2 x 10^6 cells in a buffer of 300 mM NaCl, 7 mM Tris (pH 8) and 7 mM EDTA for 1.5 h at 37°C. To remove proteins and reverse the crosslinks, the samples were treated with 200 µg Proteinase K at 65°C for >16 h. Nucleic acids were extracted from the samples in two rounds of extraction with phenol using Phase Lock Gel™ tubes, followed by precipitation with 70% ethanol and 300 mM sodium acetate at -20°C overnight. Precipitated nucleic acids were washed once with Ethanol, air-dried at RT and resolved in 50 µl elution buffer (5 mM Tris pH 8.5).

Before performing immunoprecipitation, 25 µl Dynabeads Protein G (ThermoFisher) per sample were blocked in PBS with 0.5% BSA for 1 h to prevent unspecific binding of antibodies to the beads. After blocking, the beads were incubated with 2 µg S9.6 antibody per sample in 400 µl DRIP IP buffer (50 mM HEPES/KOH pH 8, 0.14 M NaCl, 5 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate) for >4 h at 4°C to form antibody-bead complexes. For DRIP, 1.5 µg nucleic acids prepared as described were diluted in 400 µl DRIP IP buffer, before adding 100 ng *Drosophila melanogaster* spike-in DNA/RNA (prepared from crosslinked *D. melanogaster* chromatin using the same protocol as for the samples), taking 10 % aside as input and adding the antibody-bead complexes. DRIP was performed by rotating the samples overnight at 4°C. To remove unspecifically bound nucleic acids, the beads were washed once with 1 ml IP Wash 1 Buffer (20 mM Tris pH 8.0, 2 mM EDTA, 50 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with 1 ml High Salt Buffer (20 mM Tris pH 8, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.01% SDS), once with 1 ml IP Wash 2 Buffer (10 mM Tris pH 8, 1 mM EDTA, 0.25 M LiCl, 1% NP-40, 1% sodium-deoxycholate) and twice in TE buffer (pH 8). Nucleic acids were eluted by heating the beads to 65°C for 15 min in 50 µl elution buffer (50 mM Tris pH 8, 10 mM EDTA, 1% SDS). The eluate was transferred to a new tube, the beads were resuspended in 50 µl fresh elution buffer and the two elution fractions were united. Nucleic acids from the IPs and the inputs were cleaned up using the Macherey-Nagel™ NucleoSpin™ Gel and PCR clean-up Kit. The amount of precipitated DNA was quantified using the ORA™SEE qPCR reagent (HighQ) and qPCR primers amplifying 120-130 bp fragments of the EF1A promoter or the annotated endogenous regions. Cq values were normalized to input and Drosophila spike-in control.

3.7. Co-precipitation of proteins with DNA:RNA hybrids

To determine association of endogenous proteins with DNA:RNA hybrid structures in cells, NIH/3T3 cells were grown to a confluency of 80-90% in a 10 cm dish and harvested with trypsin. During the whole protocol, protein low binding tubes were used. Cell pellets were washed once in PBS and lysed in 1 ml cell lysis buffer (85 mM KCl, 5 mM HEPES pH 8, 0.5 % NP-40, protease inhibitor) per 10×10^6 cells for 15 min on ice. 100 µl Dynabeads Protein G were pre-blocked in 0.5% BSA in PBS for 2 h at 4°C. After lysis of the cell membranes, nuclei were collected by spinning down the samples at 15,000x*g* for 1 min. The supernatant was removed, and nuclei were lysed in 750 µl

nuclei lysis buffer (10 mM Tris pH 7.5, 200 mM NaCl, 2.5 mM MgCl$_2$, 0.2 % sodium deoxycholate, 0.1 % SDS, 0.5 % Triton X-100) by sonication for 4 min in intervals of 20 sec pulse / 30 sec pause at 40% amplitude. 5 % v/v were taken aside as input before dilution of the sample with 2.3 ml RSB+T buffer (10 mM Tris pH 7.5, 200 mM NaCl, 2.5 mM MgCl$_2$, 0.5 % Triton X-100). Unspecific targets of mouse IgG antibodies were removed from the samples by incubating with 10 µg mouse IgG and 70 µl Dynabeads Protein G for 1 h at 4°C. The beads were locked using a magnetic rack and the pre-cleared supernatant was split into 3 Eppendorf tubes. To each tube, 32 µl of the pre-blocked beads, 0.1 ng RNase A per µg genomic DNA and either 4 µg of the S9.6 antibody or 4 µg mouse IgG were added. The third tube received no antibodies to serve as a bead-only control. IP was performed by rotation of the samples at 4°C for 2 h. To remove unspecific binding, the beads were washed 4 times with RSB+T and twice with RSB buffer without Triton X-100. Specifically bound proteins were eluted by incubating the beads in 30 µl LAP+DTT (125 mM Tris-HCl pH 6.8, 5 % SDS, 0.004 % Bromophenol Blue, 10% β-mercaptoethanol, 100 mM DTT, 20 % glycerol) at 70°C for 10 min. The beads were locked using a magnetic rack and the supernatant was transferred to a new tube. 38 µl LAP+DTT were added to the input sample and all samples were heated to 95°C for 10 min before subjecting to SDS-PAGE and Western Blot for detection of co-precipitated proteins.

3.8. SDS-PAGE and Western Blot

For the specific immunodetection of proteins, SDS-PAGE followed by Western Blot analysis was performed. To this end, cells washed with PBS to remove residual serum were lysed in cell lysis buffer (Cell Signaling Technology®) for 30 min on ice. After 10 and 20 min of incubation, lysate was sonicated with an EpiShear Probe Sonicator (Active Motif) for 2 cycles of 20 seconds to release nuclear protein. Lysate was spun down at 15,000x*g* for 10 min, the supernatant was mixed with 2x SDS sample buffer (125 mM Tris-HCl pH 6.8, 5 % SDS, 0.004 % Bromophenol Blue, 10% β-mercaptoethanol, 100 mM DTT, 20 % glycerol) and boiled at 95°C for 10 min. Proteins were resolved by SDS-PAGE on a 12% polyacrylamide gel. Proteins were transferred to an Immobilon-FL PVDF membrane at 300 mA for 90 min using a wet-tank blotting system. Proteins were detected using the mouse monoclonal rTetR antibody (#631131, TAKARA), and the rabbit polyclonal anti-DDX19A antibody (orb242165, Biorbyt) at manufacturer´s recommendations in combination with an HRP-coupled

secondary antibody. Imaging was performed on a FusionFX detection system (VILBER) and SuperSignal™ West Femto Chemiluminescence substrate (ThermoFisher).

3.9. Quantitative real-time PCR

To quantify enrichment of specific DNA sequences in IP samples, quantitative real-time PCR (qPCR) on a CFX Connect Real-Time PCR Detection System (Bio-Rad) was used. Specific primers were designed using the *NCBI Primer Blast* tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) to amplify a single target in the mouse genome with a length of 100-140 bp at an annealing temperature of 59-62°C. Samples were amplified in a total reaction volume of 15 µl using the ORA™ SEE qPCR Green ROX L (HighQu) reagent and a cycling program consisting of an initial denaturation step at 95°C for 3 min, followed by 40 cycles of denaturation at 95°C for 5 sec and 30 sec annealing/elongation at 60°C. Melt curves of the final PCR products were analysed from 65°C-95°C in steps of 1°C/10 sec and measurements showing multiple peaks in the melt curve analysis were excluded in the analysis of Cq values. For analysis of enrichment by IP, sample Cq values were normalized to input and *Drosophila* spike-in. For *Drosophila* spike-in analysis, specific regions enriched in the respective precipitated modification were used as normalization signal for each precipitated histone PTM or DNA:RNA hybrids. Measurements showing no product after 40 cycles of amplification received a Cq value of 42 in the final analysis.

For expression analysis of genes on mRNA level, joint reverse transcription (RT) of RNA and amplification of cDNA was performed using the Luna® Universal One-Step RT-qPCR Kit (NEB). Total RNA was extracted from NIH/3T3 cells expressing shRNAs for 13 days using the RNeasy Mini Kit (QIAGEN) following manufacturer´s protocol. Concentration of extracted RNA was determined by measuring absorbance at 260 nm on a NanoDrop spectrophotometer and 100 ng of RNA per 20 µl reaction volume were used for each RT-qPCR reaction performed on a CFX Connect Real-Time PCR Detection System (Bio-Rad). Samples were reverse transcribed and amplified in a single PCR program, consisting of the initial RT reaction at 55°C for 10 min, followed by 1 min denaturation at 95°C and 40 cycles of denaturation/extension at 95°C for 10 sec / 60°C for 30 sec. Fluorescence signals were read out after every elongation step, and after 40 cycles of amplification, final PCR products were analysed by recording melt curves ranging from 60-95°C to determine specificity of amplification. qPCR

primers were designed to amplify a specific product of 100-140 bp length at an annealing temperature of 60-63°C, spanning the border between two exons to ensure amplification of a correctly spliced and converted cDNA substrate. qPCR signals of the mRNAs of interest were normalized to the signal of mouse β-2 microglobulin (b2M), an ubiquitously expressed major histocompatibility complex (MHC) class I protein, for internal normalization and compared to the mRNA level in cells expressing a control shRNA targeting a luciferase sequence.

3.10. Purification of LSD1-His

For overexpression of recombinant *6xHistidin*-tagged LSD1 (LSD1-His), codon optimized *E.coli BL21(DE3)* cells were transformed with 50 ng of pET28-LSD1-His using heat-shock transformation (42°C, 45 sec). Cells were incubated on LB/2% agar, with 35 µg/ml chloramphenicol to select for expression of the optimized codons and 50 µg/ml kanamycin to select for expression of the plasmid, overnight at 37°C. 50 ml of LB medium with 50 µg/ml kanamycin were inoculated with a single colony from the overnight culture and incubated at 37°C for 6 h at 150 rpm. For the main culture, 500 ml of LB/kanamycin were inoculated with 6 ml of the pre-culture and incubated at 37°C, 150 rpm until the culture reached an OD600 of 0.7. Expression of the LSD1-His construct was induced by adding 200 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and overexpression was performed at 17°C, 150 rpm for 14 h. Cells were harvested at 5000 xg for 15 min at 4°C, the pellet was washed once in 30 ml STE buffer (100 mM NaCl, 10 mM Tris pH 8, 1mM EDTA) and frozen at -20°C until further use.

For purification of LSD1-His, pellets were resuspended in 30 ml sonication buffer (30 mM KPi pH 7.5, 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, 20 mM imidazole, 10 % glycerol) with protease inhibitor and lysed by sonication using an EpiShear™ Probe Sonicator (Active Motif). The lysate was cleared by centrifugation and filtration through a 0.45 µm CHROMAFIL GF/PET-45/25 filter (MACHEREY-Nagel). Affinity chromatography was performed using an NGC™ Chromatography system (Bio-Rad) and Ni-NTA superflow beads (Qiagen). Proteins were eluted in elution buffer (30 mM KPi pH 7.2, 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, 220 mM imidazole, 10 % glycerol) and subjected to dialysis into storage buffer (20 mM HEPES pH 7.2, 200 mM KCl, 0.2 mM DTT, 1mM EDTA, 10% glycerol). Aliquots were snap-frozen and stored at -80°C.

3.11. Peptide binding analysis

To detect binding of DDX19A to specific histone PTMs on MODified™ Histone Peptide Arrays (Active Motif) or peptide arrays synthesised in the lab containing peptides with a length of 15 amino acids were used. The arrays were synthesized by sequential spotting of amino acids on a cellulose membrane using an Autospot Peptide Array Synthesizer (Intavis AG) and the SPOT synthesis method (CITATION). MODified™ Histone Peptide Arrays (Active Motif) or synthesized peptide arrays were incubated overnight in blocking solution (5% milk powder, 1x PBS, 0.1 % Tween20) at 4°C to block unspecific binding epitopes. To remove blocking solution, arrays were washed three times for 5 min with 1x PBS/Tween20 and incubated for 10 min in interaction buffer (100 mM KCl, 20 mM HEPES pH 7.5, 1 mM EDTA pH 8, 10 % glycerol) to equilibrate the array for binding. Binding of DDX19A-GST was performed by incubation of 50 nM protein with the pre-blocked array in interaction buffer for 1 h at room temperature. The array was washed three times for 10 min in 1x PBS/Tween20 to remove unspecific binding of DDX19A to the surface. Bound proteins were detected by incubation of the array with an anti-GST antibody (GE Healthcare, #27-4577-01) in 5% milk/1x PBS/Tween20 for 1 h at room temperature. The array was washed again as described and incubated with an anti-goat horseradish peroxidase (HRP) coupled antibody in 5% milk/1x PBS/Tween20 for 1 h. The array was washed twice for 10 min in PBS/Tween20 to remove residual antibodies and once for 10 min in PBS to equilibrate for detection. Protein bound to the surface was visualized via chemiluminescence using SuperSignal™ West Femto Chemiluminescence substrate (ThermoFisher) and imaging on a FusionFX detection system (VILBER).

3.12. Co-precipitation of proteins with LSD1

NIH/3T3 were harvested by trypsinisation. The pellet was resuspended in 2x pellet volume of nuclear lysis buffer B (50 mM Tris-HCl pH 7.4, 20% glycerol, 1.5 mM MgCl2, 420 mM NaCl, 1 mM Na3VO4, 25 mM NaF) supplemented with Protease inhibitor and incubated on ice for 15 min. Lysate was homogenized with 25 strokes of a douncer (0.01-0.03 mm) and incubated with rotation at 4°C for 30 min. Lysate was cleared by spinning down at 4°C, 16 000xg, 30min and the supernatant was transferred to a new tube. 1xDP buffer (50 mM Tris-HCl pH 7.4, 5% glycerol, 1.5 mM MgCl2, 150 mM NaCl, 1 mM Na3VO4, 5 mM NaF) with 0.4% NP40 and protease inhibitors (1.8 ml buffer to 1ml lysate) was added and the sample was incubated on ice for 10 min before clearing

through ultracentrifugation (30min, 4°C / 43.000rpm / TI50.2). The supernatant was transferred to a fresh tube an 5% input were taken. The sample was incubated with 1µg anti-KDM1A antibody per 1mg protein in the lysate overnight at 4°C. 1.5-fold of loading capacity of Dynabeads protein G were added for 2h at 4°C with rotation. Beads were washed twice with DP/NP40 buffer, twice with DP buffer and twice with 150 mM NaCl. Proteins were eluted in 30 µl of 2x SDS sample buffer (125 mM Tris-HCl pH 6.8, 5 % SDS, 0.004 % Bromophenol Blue, 10% β-mercaptoethanol, 100 mM DTT, 20 % glycerol) for 10 min at 70°C and subjected to SDS-PAGE and Western Blot as described.

3.13. Analysis of the integration sites of the reporter gene into the host genome

As the retroviral transduction method described in this work does not include a targeting mechanism for the integration of the gene of interest into the host genome, integration occurs at random genomic sites. The retroactive determination of the reporter gene integration site was performed according to a published method (Ciuffi & Barr, 2011). Total genomic DNA was extracted from cells transduced with the *synP-mCherry* reporter using the DNeasy Blood & Tissue Kit (QIAGEN). 1 µg of DNA was digested with 10 U MseI for 3.5 h at 37°C. The restriction enzyme was heat-inactivated, and the fragmented DNA was ligated with phosphorylated double-stranded linker DNA (Table 4) using T4 DNA ligase (ThermoFisher Scientific). DNA fragments containing the viral 3´LTR were amplified in 3´-direction up to the next MseI site, in a first PCR reaction with primers specific for the 3´LTR and the linker sequence, using HotStarTaq DNA polymerase (QIAGEN). A second round of PCR with primers specific for the primer sequences in PCR1 (Table 4) was performed sequentially to further enrich the target fragments.

Without prior clean-up, the PCR products from PCR2 were cloned into a TOPO™ TA vector using the TOPO™ TA Cloning™ Kit for Subcloning, with One Shot™ TOP10 chemically competent E. coli cells (Invitrogen). The PCR products were sequenced using the provided sequencing primers and Sanger sequencing. Mapping of the detected sequences to the mouse mm9 genome was performed using the UCSC Mouse BLAT Search (https://genome-euro.ucsc.edu/cgi-bin/hgBlat?hgsid=252871179 _zntn6lAYEHyvUdtjxhSGmHfP7XrP&command=start).

Table 4: Oligonucleotides and primers used for the amplification of integrated viral sequences.

name	sequence	purpose
MseI_linkF	GTAATACGACTCACTATAGGGCTCCGCT	anneal and phosphorylate to form linker dsDNA
MseI_linkR	TAAGCGGAGCCCTATAGTGAGTCGTATTAC	
INT_3LTR_F	TCTTTCATTTGGGGGCTCGT	amplify sequence between linker and 3´LTR (PCR1)
INT_link_R	GTAATACGACTCACTATAGGGCTCC	
Nest_3LTR_F	GTTTCGGTGATGACGGTGAAAAC	perform nested PCR on PCR1 (PCR2)
Nest_link_R	ACTCACTATAGGGCTCCGCTT	

3.14. Electrophoretic mobility shift assay (EMSA)

To analyse binding of LSD1 to nucleic acid structures, electrophoretic mobility shift assays (EMSA) with recombinant protein and Cy5-labelled oligonucleotides were performed. To prepare nucleic acid secondary structures, unmodified DNA-, RNA- or Cy5-labelled DNA oligonucleotides were ordered from Integrated DNA Technologies (IDT)(Table 5). Oligonucleotides were diluted to concentrations of 100 µM in ddH2O and incubated in TE-50 (10 mM Tris pH 8, 1 mM EDTA, 50 mM NaCl) at equimolar amounts to form dsDNA and DNA-bubbles, or at 3x excess of RNA to form DNA:RNA hybrids or R-loops. Annealing of oligonucleotides was performed by heating to 95°C for 2 min followed by cooling down to 20°C at 5°C/min. The annealed nucleic acid structures were resolved for their molecular weight on a 6% acrylamide / 0.5x TBE gel, successfully annealed dsDNA, DNA-bubble and R-loop structures were cut out and extracted by incubation in extraction buffer (20 mM Tris pH 7.5, 150 mM NaCl, 0.1% SDS, 10 mM EDTA). Nucleic acid structures were cleaned up via phenol extraction and ethanol precipitation as described in 3.3.2. For determination of $K_D(app)$, 26 fmol of nucleic acid structures were incubated with LSD1 at concentrations of 2.6-1300 nM in EMSA binding buffer (12 mM HEPES pH 8, 0.12 mM EDTA, 120 nM KCl, 1 mM DTT, 0.01 % NP40, 12 % glycerol, 50 ng/µl BSA) for 30 min at 37°C. The samples were resolved on a 6% acrylamide/0.5x TBE gel and fluorescence of the Cy5-labelled DNA oligo was detected using a FusionFX detection system (VILBER).

Table 5: Oligonucleotides used for creation of dsDNA, DNA-bubbles and R-loop structures.

#	Name	sequence (5' to 3')	purpose
1	FP3_RNA	GAAGCUGGGACUUCCGGGAGGAGAGUG CAA	Complementary to centre of #2 (DNA:RNA hybrid, R-loop)
2	vgcoreFP3_ DNA_F	*Cy5*-CGGGTTGTCAAGAATTTTAACGGCC ATTTCTGTGTTGCACTCTCCTCCCGGAAG TCCCAGCTTCTGTGTTTGTGACAAACGCA AGCTCATGTAAGTGCTC	coupled to Cy5, centre complementary to #1 (DNA:RNA hybrid)
3	vgcoreFP3_ DNA_R	GAGCACTTACATGAGCTTGCGTTTGTCAC AAACACAGAAGCTGGGACTTCCGGGAGG AGAGTGCAACACAGAAATGGCCGTTAAAA TTCTTGACAACCCG	complementary to #2 (dsDNA)
4	vgcoreRandom_ DNA_R	GAGCACTTACATGAGCTTGCGTTTGTCAC AATTAGTAGACAGTAATTGTCGACAAATG ATGGTGAGCTATGAAATGGCCGTTAAAAT TCTTGACAACCCG	5´and 3´ends complementary to #2 (DNA bubble and R-loop)

# 4.	Results

## 4.1.	Design of the fluorescent reporter system

The aim of this work was to develop a method with the ability to meet the complexity of epigenetic gene regulatory networks. Methodical high-throughput approaches that can integrate many nodes of an interaction network in parallel have recently gained in importance. Proteomics approaches based on the isolation of multiprotein complexes from cells by using small-molecule inhibitors as probes, recombinantly tagged proteins, or target specific antibodies, followed by analysis using mass-spectrometry, constitute a powerful tool to study interactions of different chromatin regulators (Bantscheff et al., 2011; Malovannaya et al., 2011; Weigt et al., 2016). These methods are unbiased and have contributed to the discovery of novel interactions of epigenetic effector proteins in complex with coregulators, with histone PTMs or with small-molecule inhibitors (Noberini et al., 2016). However, the information gained by mass-spectrometry based proteomics approaches is conditioned by a strong physical association of the investigated proteins. Proteins with low abundance, or weak but physiologically relevant interactions are missed out in these studies. Furthermore, no conclusions about the mechanism of the interaction can be drawn.

Furthermore, the regulation of transcription is a complex event that is influenced by a multitude of factors (Nicolas et al., 2017; Tycko et al., 2017) that cannot be mirrored on the whole in an *in vitro* experiment, and is hard to track thoroughly *in vivo*. Many epigenetically active enzymes have been validated as clinical targets, and their mechanisms of action have been characterized in detail to develop specific inhibitors for the use in epigenetic therapies (Bates, 2020). However, the treatment with an epigenetic drug constitutes a major intrusion into a global, complex and close meshed network of gene regulation (X. Gu et al., 2020; Komashko & Farnham, 2010; Quagliano et al., 2020). As a response to the treatment with specific small molecule inhibitors, endogenous stabilization mechanisms lead to the rewiring of transcriptional networks, causing epigenetic cancer therapies to suffer from resistance development (B. E. Li & Ernst, 2014; Quagliano et al., 2020; Rathert et al., 2015). Thus, in addition to the detailed mechanistic biochemical characterization of individual epigenetic effector proteins, there is a persistent need for methods that can capture the function of epigenetic regulators as nodes of a complex cellular network.

Finally, the multi-subunit complexes assembled by epigenetic coregulators are variable among different cell types, and can take over various functions in the same cell depending on their local composition on chromatin (Blackledge et al., 2014; Längst & Manelyte, 2015; Streubel et al., 2017). The interference with one complex partner can induce a switch of functionality of the whole complex (Adams et al., 2018; Son et al., 2013; J.-R. Yu et al., 2019). Importantly, merely the scaffolding function of a protein as subunit of a complex can have an effect on the chromatin regulation by the complex, irrespective of the potential catalytic activity of the coregulator (Filippakopoulos & Knapp, 2014; Linhares et al., 2020; Maiques-Diaz et al., 2018; David C Schultz et al., 2002). To fully understand the regulatory function of an epigenetic effector on chromatin, a system that monitors the effects on transcription in the context of an endogenous coregulator environment is needed.

4.1.1. Basic concept of the reporter system

Targeted editing of the epigenome is an established approach in synthetic biology, which has demonstrated that the transcriptional activity of an endogenous genomic locus of interest can be modulated by the recruitment of a chromatin effector (Pickar-Oliver & Gersbach, 2019; Rots & Jeltsch, 2018; Thakore et al., 2016). The editing approach can be reversed, to use it for the functional characterization of the epigenetic effector at a chromatin locus (Bintu et al., 2016), and a similar approach was established during this work. Based on the chromatin in vivo assay (CiA) (Hathaway et al., 2012), a fluorescent reporter system was designed (**Figure 15**). The CiA assay is based on the introduction of two arrays of DNA binding sites upstream of the endogenous *Oct4* promoter in ES cells, combined with the replacement of the first exon of *Oct4* with an in-frame nuclear EGFP. This assay was used to recruit different effectors to the locus by chemically induced proximity (CIP), aiming to study the cellular dynamics of chromatin regulation. In contrast to the CiA:Oct4 allele, which is located in an endogenous repressed chromatin environment in differentiated cells (Headley et al., 2019), the reporter gene was randomly integrated into the genome of mouse fibroblast cells using lentiviral transduction. This predominantly resulted in the integration into transcriptionally active loci (Schröder et al., 2002), and thereby allowed for the characterisation of several repressive chromatin effectors using this system (**Figure 17**).

The epigenetic reporter system developed in this work was based on the genetic engineering of a cell line of interest, in a way that enabled the stable expression of two essential components: a fluorescent reporter gene cassette and a chromatin effector fused to a DNA binding domain (**Figure 15**).
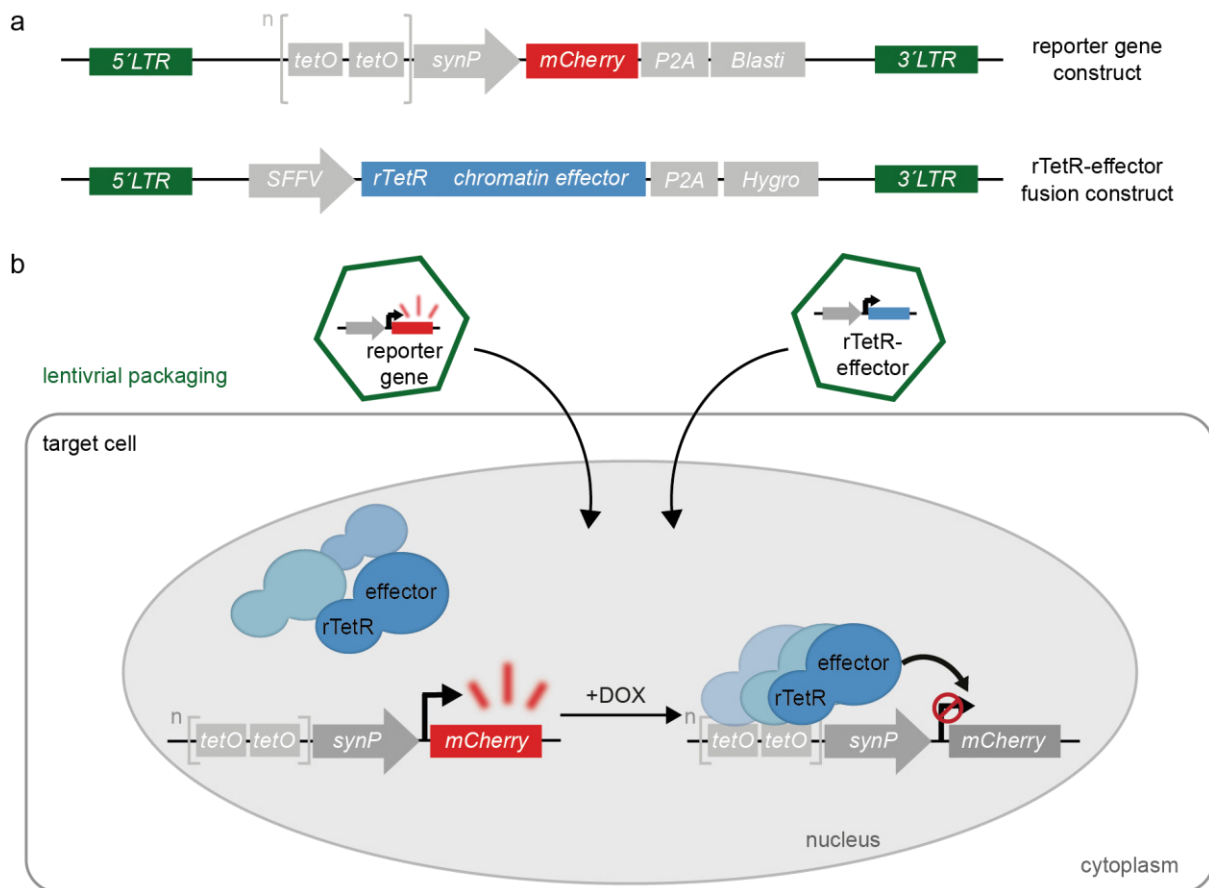
Figure 15: A modular fluorescent reporter system to dynamically track the activities of epigenetic effectors. a, Modular structure of the two components of the reporter system. Modules that were varied during this thesis are the TetR binding site array (^{n}[*tetO*]), the synthetic promoter sequence (*synP*), the reporter fluorophore (*mCherry*) and the rTetR chromatin effector fusion construct. *LTR* = viral repeat sequences, *Blasti* = Blasticidin resistance gene, *SFFV* = constitutive SFFV promoter, *Hygro* = Hygromcin resistance gene. **b,** Schematic describing the reporter system approach. The gene constructs from **a,** are transduced into target cells using lentiviral gene transfer. Both constructs are stably integrated into the host genome. Antibiotic selection results in a reporter cell population with constitutive expression of both genes. Upon DOX treatment, the rTetR-effector is recruited to the *tetO* array, and expression of the reporter fluorophore is repressed.

The sequences encoding the reporter cassette and the effector fusion protein were introduced into the cells by lentiviral transduction (Mátrai et al., 2010), resulting in random integration of the constructs into the genome. Each component of the system was coupled to the expression of an individual resistance gene, enabling antibiotic selection for cells expressing the constructs. The reporter system was designed in a modular way to enable the exchange of individual parts by standard cloning methods.

The complete fluorescent reporter gene cassette consisted of an array of two to six Tet-repressor (TetR) binding sites (*tetO*) upstream of a synthetic promoter sequence (*synP*) that drives the expression of a fluorescent protein, which could be detected using flow cytometry to dynamically monitor the activity of the reporter gene (**Figure 15a**, **b**). The second component of the reporter system was a fusion protein of the epigenetic effector protein or domain of interest fused to a TetR or reverse TetR (rTetR) DNA binding domain. This fusion enabled the dynamic recruitment of the effector protein to the *tetO*-array upstream of the *synP* element in the absence or presence of Doxycycline (DOX) in the culture medium (Gossen et al., 1995; Scholz et al., 2004). Once bound, the epigenetic effector can act on the surrounding chromatin, thereby inducing activation or repression of transcription depending on its activity. The resulting effect on transcription could be monitored and quantified by analysing the signal strength of reporter fluorophore expression using flow cytometry.

4.1.2. Optimization of the promoter and *tetO*-array modules results in a stable system for the analysis of corepressor proteins

Aiming to explore the potential and constrains of the fluorescent reporter system, preliminary experiments were conducted in mouse embryonic fibroblast cells. NIH/3T3 (ATCC® CRL-1658™) and iMEF (immortalized mouse embryonic fibroblasts, kind gift of Prof. Thomas Jenuwein, MPI Freiburg) are well characterized in literature and have low demands regarding cultivation methods.

To begin with, the choice of a suitable fluorophore to be used as the fluorescent reporter gene was essential, aiming to maximize the dynamic resolution of the system. The fluorescent protein chosen to report on promoter activity should provide stable and fast folding at 37°C, while avoiding dimerization of the fluorophores (Cranfill et al., 2016). The fluorescent protein turnover rate must be high enough to enable the tracking of changes in transcription. Finally, to report on the expression of individual components of the system in order to permit the combination of multiple fluorophores in the same cells, the fluorescence spectra of the individual proteins need to be compatible. t*urboGFP*, *mCherry* and *BFP* met the mentioned criteria and were used to monitor the individual components of the reporter in transduced cells (Evdokimov et al., 2006; Shaner et al., 2004; Zimmer, 2002). Both *turboGFP* and *mCherry* were successfully applied as reporter fluorophores of the synthetic reporter gene (**Figure 16a**).

The integration of the reporter gene via viral transduction was not targeted and occurred at random sites throughout the cell genome (Mátrai et al., 2010). During transduction, conditions were chosen that had been tested to ensure a single integration site per cell. The result was a heterogenous cell population, with the expression strength of the reporter fluorophore varying from cell to cell in the range of about 1.5 log_{10}-scales analysed by flow cytometry (**Figure 16a**).
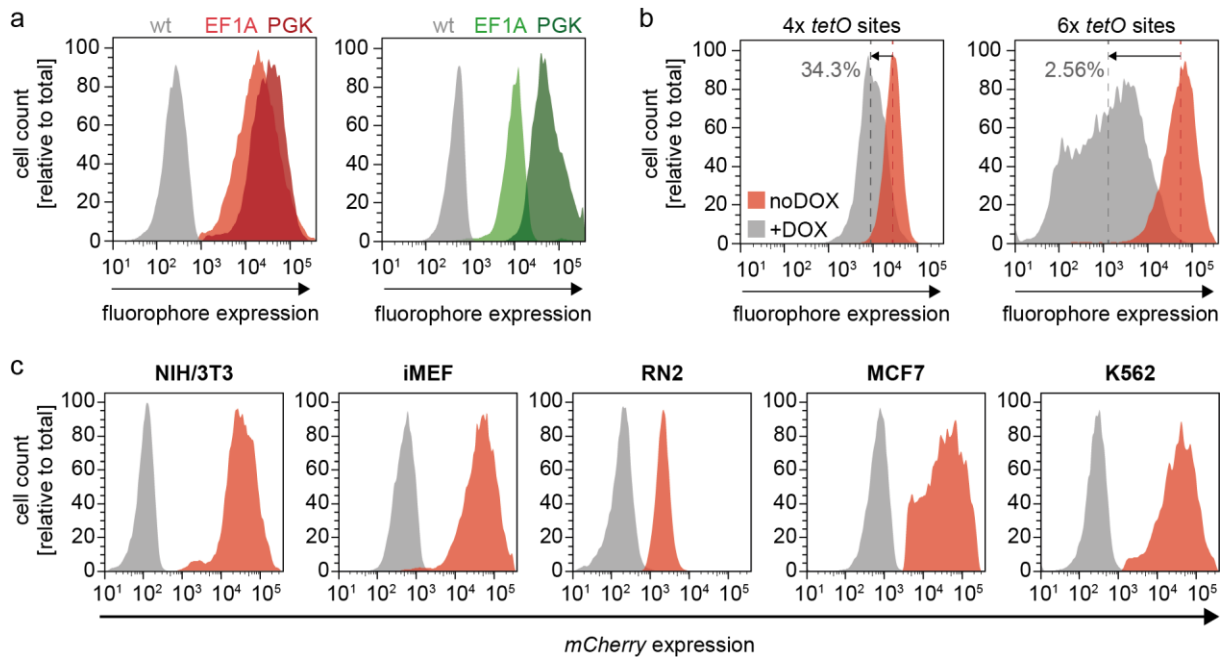
Figure 16: Establishment of a stable reporter system that can be expressed in different cell lines. a, Flow cytometry analysis of NIH/3T3 cells expressing reporter genes with different *synP* elements and reporter fluorophores. Cells expressed either *mCherry* (red) or *turboGFP* (green) under control of either an *EF1A* or a *PGK* promoter. *wt* = NIH/3T3 cells without transduction of a synthetic reporter gene. **b,** Silencing of fluorophore expression by recruitment of rTetR-KRAB to a 4x or a 6x *tetO* array. NIH/3T3 expressing an *EF1A-mCherry* reporter gene (red) were treated with DOX for 7 days (grey). Dotted lines indicate the median of the fluorescence signal. The residual median signal strength is indicated in percent (+DOX/noDOX). **c,** Flow cytometry analysis of different cell lines stably transduced with the 6x*tetO-EF1A-mCherry* reporter. Cells were analysed after antibiotic selection for 7 days. NIH/3T3 and iMEF = murine fibroblast cell lines; RN2 = murine MLL–AF9;NrasG12D AML cell line; K562 = human chronic myelogenous leukemia cell line; MCF7 = human breast adenocarcinoma cell line.

This heterogeneity was expected and desired, since the integration at different sites in the genome will result in a diverse transcriptional activity depending on the respective chromatin environment. It can be noted that even a cell population derived from a single clone would result in differences in fluorescence signal, due to the transcriptional plasticity observed in homogenous cell populations (Nicolas et al., 2017; Raj et al., 2006). The diversity in integration sites and transcriptional activities was desired in this work to exclude the identification of locus-specific or clonal effects in the final screening

experiments (Meir et al., 2020; Tycko et al., 2017). To monitor changes in transcriptional activity at the reporter in the whole cell population, the median of expression levels of at least 10,000 cells measured by flow cytometry analysis was compared (**Figure 16b**).

Due to the modular design of the reporter gene, the number of TetR binding sites (*tetO*), the synthetic promoter sequence (*synP*) and the fluorophore could be adapted to the experimental requirements in a few cloning steps. Two different constitutive promoters driving expression of the reporter fluorophore were tested. By performing individual experiments with the *EF1A* and the *PGK* promoter in NIH/3T3 cells in combination with both *turboGFP* and *mCherry* as reporter fluorophores, the observed median expression from the *PGK* promoter was consistently stronger than expression from the *EF1A* promoter (**Figure 16a**). This was contradictory to published reports of promoter strength after lentiviral transduction of mammalian cells (J. Y. Qin et al., 2010), where *PGK* was shown to result in higher transcriptional activity compared to *EF1A*. However, Qin et al. did not test NIH/3T3 cells, suggesting that NIH/3T3 show higher activity of PGK in a cell line specific manner. However, this observation was not investigated extensively. For further experiments to test the repressive activity of epigenetic effectors, the *EF1A* promoter was chosen.

In a next step, the silencing rate achieved by recruitment of an epigenetic regulator was optimized by variation of the number of *tetO*-sites upstream of the *EF1A* promoter. Recruitment of a rTetR-KRAB fusion protein to the synthetic reporter resulted in much stronger repression of fluorophore expression when the 4x *tetO* array was extended by two additional *tetO* sites (**Figure 16b**). This effect was observed in different experiments, but was not reproduced in a way that would allow for quantification of the effect. As further expansion by additional *tetO* sites increased the chance of defective recombination events during cloning and transduction of the reporter gene. Thus, the combination of six *tetO* sites in front of the *EF1A* promoter, driving the expression of *mCherry* as the reporter fluorophore was used as reporter system in further experiments, and will be abbreviated as *synP-mCherry* reporter in the remaining thesis.

By transduction of the *synP-mCherry* reporter construct, it was shown that the stable integration and expression of the reporter gene cassette could be achieved in several different cell lines (**Figure 16c**). This universal applicability is crucial, as transcriptional

landscapes and the controlling epigenetic regulatory networks show high variability between different cell types (Rahmani et al., 2019). Therefore, although general claims about the interactions of epigenetic effector proteins as transcriptional coregulators can be drawn from experiments in model cell lines, the relevance of the observed interaction must be validated for the specific cell type of interest.

### 4.1.3. Characterization of the silencing dynamics of different chromatin effectors

The cascade of changes in the chromatin environment that is required to switch a gene from an active to a repressed state and *vice versa* is complex and incompletely understood (Bintu et al., 2016; Headley et al., 2019; Tycko et al., 2017). The starting point of this work was to assess the sensitivity of the reporter gene to the recruitment of different chromatin effectors. The aim was to determine which chromatin changes were sufficient for stable gene repression, and to compare the silencing dynamics upon recruitment of functionally different chromatin regulators. Working with a reporter gene under control of a strong promoter, several chromatin regulators with the ability to assemble repressive multiprotein complexes at their target site were investigated (**Figure 17a**).

The most common domain used for transcriptional repression of target genes is the Krüppel-associated box (KRAB) domain of the zinc finger protein ZNF10 (*also*: KOX1). The KRAB domain recruits the scaffolding protein KAP-1, which in turn recruits various repressive epigenetic complexes, leading to H3K9 methylation, histone deacetylation and chromatin remodelling at the targeted locus (Alerasool et al., 2020; Beerli et al., 1998; Margolin et al., 1994). Congruously, recruitment of KRAB to the *synP-mCherry* reporter induced stable repression of the fluorescence signal (**Figure 17a**).

As a member of the heterochromatin protein 1 (HP1) family, CBX3 plays a role in transcriptional regulation by the installation and stabilization macromolecular heterochromatin complexes (Hathaway et al., 2012; Machida et al., 2018; Saksouk et al., 2015). HP1 proteins have been shown to efficiently package H3K9me3 modified nucleosomes into transcriptionally inactive heterochromatin by their scaffolding and oligomerization functions (Canzio et al., 2011, 2013; Hathaway et al., 2012; Machida et al., 2018). Thus, silencing of the *synP-mCherry* reporter by recruitment of CBX3 provided a valid control model for the maximum dynamics and extend that could be expected for induced silencing of the synthetic promoter. Indeed, complete silencing

of reporter expression could be achieved. After 7 days of CBX3 recruitment, reporter cells resembled untransduced iMEF wildtype cells in the flow cytometry analysis (**Figure 17a**). Interestingly, reporter silencing following the recruitment of the KRAB domain of ZNF10 was as efficient as the silencing by CBX3, indicating that the repressive complex assembling around KRAB was equally effective. This suggests that the ability of a recruited chromatin regulator to assemble a macromolecular heterochromatin complex at the synthetic gene constitutes a fast and efficient mechanism for the silencing of transcription.

CBX3 and KRAB both induce the maintenance and spreading of H3K9me3 modified heterochromatin domains (Saksouk et al., 2015; David C Schultz et al., 2002). The embryonic ectoderm development (EED) protein is a reader of repressive chromatin marks and acts as subunit of the repressive PRC2 complex (Margueron et al., 2009). Binding of EED to H3K27me3 allosterically activates the lysine methyltransferase EZH2 in the PRC2 complex, leading to the generation and spreading of H3K27me3 domains (Jiao & Liu, 2015; J.-R. Yu et al., 2019). Recruitment of EED to a reporter gene has been shown to silence gene expression (Bintu et al., 2016). Recruitment of rTetR-EED to the *synP-mCherry* reporter resulted in partial repression of transcription (**Figure 17a**). After 24 days of recruitment, a bimodal reporter population was observed. Overall, it can be stated that reporter silencing by assembly of H3K9me3-related heterochromatin complexes appeared to be faster and more thorough than by recruitment of EED (**Figure 17a, d**).
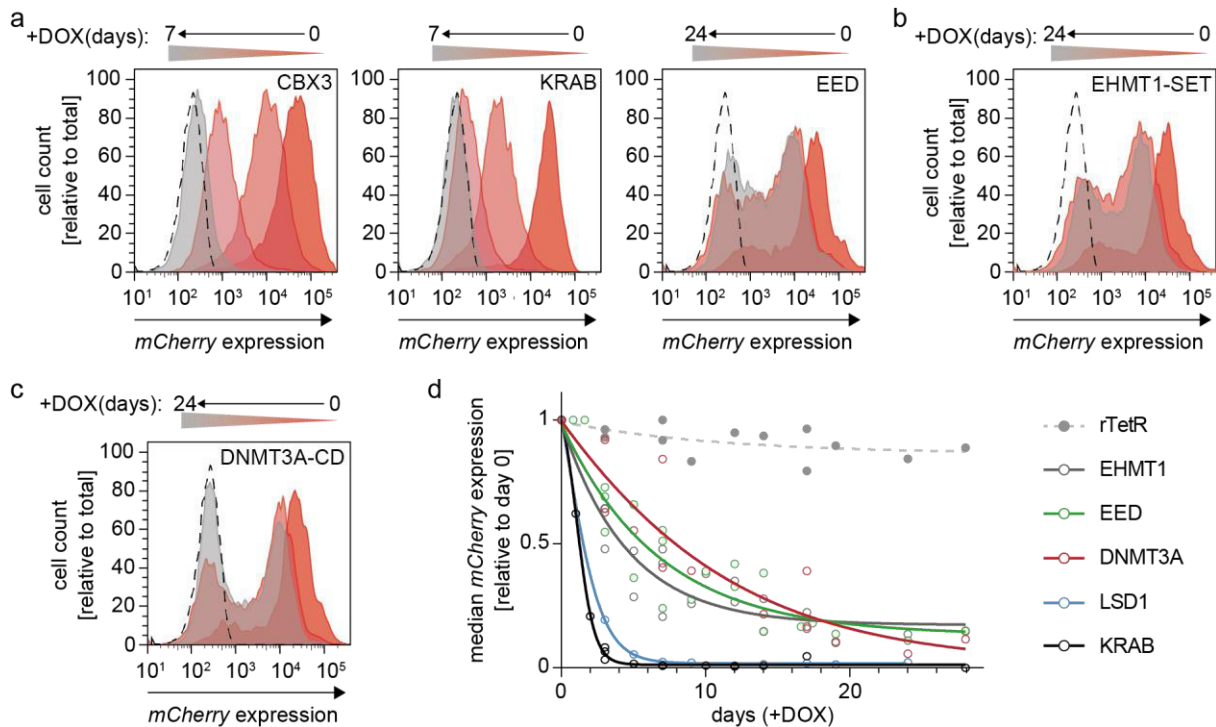
Figure 17: Different epigenetic effector proteins and protein domains show differential silencing dynamics. a-c, Representative flow cytometry analysis of iMEF *synP-mCherry* reporter expression during recruitment of rTetR-effector proteins. Scales above the diagrams indicate the duration of DOX treatment, changing their colour from day 0 (red) until the end of treatment (grey). The signal of untransduced iMEF cells is shown as dotted histogram. **a,** CBX3 and KRAB were monitored for a total of 7 days, EED for 24 days. **b,** and **c,** the SET domain of EHMT1 and the CD of DNMT3A were monitored for 24 days. **d,** Dynamics of *mCherry* expression over time upon recruitment of different repressor domains. Nonlinear regressions were interpolated from the median *mCherry* signals obtained in three independent experiments (circles) using Graphpad Prism©. Recruitment of the rTetR protein alone was performed twice and acted as a negative control. Analysis of rTetR-LSD1 recruitment was only performed once in this cell line.

To test the influence of chromatin modification writing on silencing of gene expression from the synthetic promoter, an isolated catalytic SET domain was recruited to *synP-mCherry* (**Figure 17b**). The chosen SET domain of EHMT1 (*also*: GLP) has been shown to be sufficient for the mono- and di-methylation of H3K9 *in vitro* (Tachibana et al., 2008). Recruitment of the SET domain to the *synP* element reduced *mCherry* expression to ~20% of the initial signal in the course of 14 days, but could not further silence expression upon continued recruitment (**Figure 17b, d**). The dynamics of repression were slow and incomplete compared to the effects observed with CBX3 and KRAB. This observation was especially interesting, as transcriptional repression by those regulators is mediated by trimethylation of the same histone residue (H3K9me3) that is targeted by EHMT1.

To investigate the effect of DNA methylation on gene expression from the synthetic promoter, the functional domain of DNMT3A was recruited (**Figure 17c**). The catalytic domain (CD) of DNMT3A has been used for the targeted introduction of DNA methylation *in vivo* (Vojta et al., 2016). Recruitment of the DNMT3A-CD efficiently silenced gene expression from the synthetic promoter, but with prolonged dynamics taking 14 to >24 days for the complete silencing of *mCherry* (**Figure 17c, d**). The bimodal distribution of either fully silenced or strongly expressing reporter cells suggested that transcriptional repression achieved by introduction of DNA methylation was an all-or-nothing process.

In conclusion, these data suggested that the ability of a chromatin effector to recruit endogenous complex partners had a greater impact on the efficiency of the induced transcriptional silencing than the isolated catalytic activity of an effector at the local chromatin. In line with this hypothesis, recruitment of a full-length H3K4 demethylase to the reporter was shown to efficiently silence transcription (**Figure 17d**). Fresh reporter cell lines were created each time for the independent replicates of KRAB, CBX3, EED, EHMT1 and DNMT3A repression, indicating a reproducibility of the effects. However, some data points had to be interpolated for missing values in several experiments. Although different dynamics can already be observed in the presented data, the experiments need to be reproduced extensively and with homogenous cell populations to allow for a quantification of the observed effects. Experiments with full-length EHMT1 and DNMT3A could be conducted to directly compare the augmenting effect of coregulator recruitment on gene silencing with the strength of the isolated CDs. Still, the ability of the reporter system to reflect endogenous processes of transcriptional repression underscored its applicability as a method to analyse epigenetic regulatory complexes.

4.1.4. The stability and reversibility of induced reporter silencing depend on the recruited coregulator

An interesting question that emerges during the experiments with the induced regulation of transcription by the recruitment of chromatin regulators is the stability and reversibility of the introduced epigenetic modification (Amabile et al., 2016; Bintu et al., 2016; Hathaway et al., 2012). Active and repressive chromatin states are stabilized during mitosis and meiosis by cellular maintenance machineries that work in positive feedback loops (Jaenisch & Bird, 2003). The efficient reprogramming of a chromatin locus depends on the chosen recruiting mechanism, the expression level of the epigenetic regulator, and on the ability to co-recruit endogenous chromatin modifiers to the target site (Rots & Jeltsch, 2018). Only if the spreading and maintenance of the introduced chromatin modification by endogenous coregulators is initiated by setting a strong reprogramming trigger, sustainable downregulation of transcription is achieved (Rots & Jeltsch, 2018). In reporter assays performed in CHO-K1 cells, transcriptional repression achieved either by targeted trimethylation of H3K9 or H3K27, or by deacetylation of histones, has been shown to be reversed after removal of the targeting trigger (Bintu et al., 2016). In contrast, silencing by DNA methylation, introduced by the recruitment of DNMT3B, was stably sustained after removing DNMT3B. Interestingly, transcriptional repression following the recruitment of KRAB domains has been shown to be either reversible or irreversible, depending on the endogenous epigenetic control mechanisms that were active along with the reprogramming trigger (Alerasool et al., 2020; Ying et al., 2015).

The stability of transcriptional silencing achieved by recruitment of different epigenetic effector domains to the synthetic reporter gene was tested using both the inducible Tet-ON and the reversible Tet-OFF recruitment system (**Figure 18**). Consistent with published data (Alerasool et al., 2020), DOX-induced recruitment of rTetR-KRAB (Tet-ON) induced efficient silencing of reporter expression, while fluorophore expression was reactivated upon removal of TetR-KRAB binding by the addition of DOX (Tet-OFF) (**Figure 18a**). Bintu et al. proposed a three-state-model for the silencing of transcription following repressor recruitment. The model describes the transition from an actively expressing chromatin state over a reversibly silent chromatin state to an eventually irreversibly silent state as stochastic event that is dependent on the strength and duration of the repressive trigger (Bintu et al., 2016). In contrast to their observations with this repressor, recruitment of KRAB to the synthetic reporter did not achieve the

transition into an irreversibly silent state, although the recruitment was performed for 12 days (**Figure 18a**) compared to 5 days in their system. This suggested that in the presented model system in NIH/3T3 cells, although silencing by the KRAB domain was efficiently mediated by the recruitment of endogenous coregulators, it did not succeed to induce stable silencing, perhaps because the subsequent introduction of DNA methylation was not triggered (Bintu et al., 2016; Ying et al., 2015). Analysis of DNA methylation after long term KRAB recruitment could be performed to answer this question. However, the effect might also be due to different expression levels of the effectors or different binding strengths of the DNA binder in the two systems.
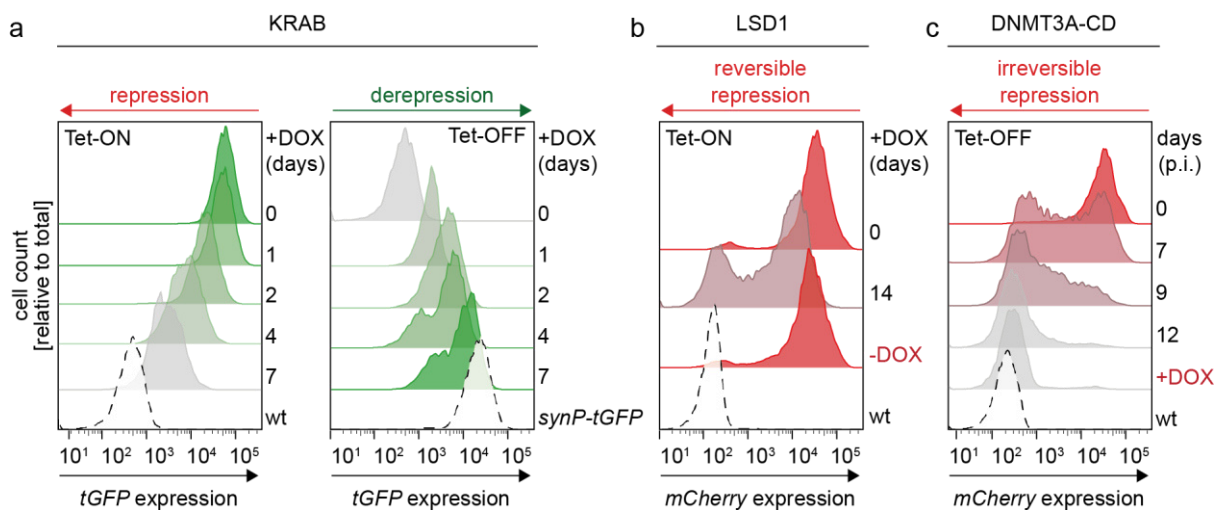
Figure 18: The stability of epigenetic silencing depends on the recruited chromatin effector. a, Representative flow cytometry analysis of NIH/3T3 *synP-tGFP* reporter cells during recruitment (left) or upon withdrawal (right) of the KRAB domain. Changes of reporter expression over time can be followed from top to bottom. Colours indicate the duration of DOX treatment. Dotted histograms indicate the maximum silenced (wt cells) or maximum active (*synP-tGFP cells without effector*) state. DOX treatment of the Tet-OFF cells was started 12 days after transduction of the TetR-KRAB construct. **b,** Transcriptional silencing of *synP-mCherry* by LSD1 is dependent on the constant recruitment with DOX. One representative experiment is shown. Numbers indicate the duration of rTetR-LSD1 recruitment. -DOX = expression profile after 14 days of LSD1 recruitment followed by 2 days without DOX. **c,** Flow cytometry analysis of iMEF *synP-mCherry* reporter cells expressing a TetR-DNMT3A-CD fusion protein (Tet-OFF system). Numbers indicate the days after infection (p.i.) with TetR-DNMT3A-CD. +DOX = Expression profile of cells after silencing for 12 days, followed by 10 days of DOX treatment. wt = untransduced iMEF wild type cells.

To investigate the stability of reporter silencing with another chromatin regulator, rTetR-LSD1 was recruited to *synP-mCherry* for 14 days (**Figure 18b**). After removal of DOX for 2 days, *mCherry* expression was fully restored to the original level. The reporter expression showed a bimodal distribution at day 14 of DOX treatment and although one fraction of the population was completely silenced to the level of untransduced NIH/3T3 cells, both populations returned to the initial actively transcribed

state in 2 days. This suggested that no sustainable chromatin remodelling towards an irreversibly silent state was established during 14 days of LSD1 recruitment.

The introduction of DNA methylation at a reporter gene has been shown to irreversibly silence gene expression (Bintu et al., 2016). To improve the slow silencing dynamics observed by recruitment of a rTetR-DNMT3A-CD fusion protein to *synP-mCherry* (**Figure 17c, d**), DNMT3A-CD was expressed as fusion to TetR. TetR has a considerably higher affinity towards *tetO* sequences in the absence of DOX than reverse TetR in the presence of DOX (Kamionka et al., 2004). Consequentially, *synP-mCherry* expression was fully silenced 12 days after transduction of the reporter cells with TetR-DNMT3A-CD (**Figure 18c**). In contrast to the Tet-OFF experiments with TetR-KRAB (**Figure 18a**), the addition of DOX after 12 days did not result in a reactivation of gene expression, even after 10 days of DOX treatment. This suggested that once DNA methylation was introduced, endogenous maintenance mechanisms stabilized the repressed state of the reporter genes. In conclusion, the results indicated that while the transient repression of transcription from the synthetic promoter could be efficiently achieved by multiple chromatin effectors, only the recruitment of the DNMT3A-CD managed to establish a heritable repressed chromatin state using this setup. The reversibility of reporter gene silencing induced by the recruitment of KRAB and LSD1 was observed consistently and repeatedly in multiple experiments. The irreversibility of reporter silencing following DNMT3A recruitment was confirmed in the lab by Katrin Collmar.

4.1.5. A reporter system for the analysis of coactivator proteins

While the creation of a constitutive reporter gene was relatively straightforward, establishment of an inducible system needed to report on the activating activities of epigenetic effector proteins constituted a challenge. Two different promoters reported to have minimal background expression were tested, the TRE3G promoter (Loew et al., 2010) and the minPromoter derived from pGL4 (Shimizu et al., 2012; Yoo et al., 2005). By recruitment of the strong transcriptional coactivator VP64 (Beerli et al., 1998; Hirai et al., 2011), transcription from both promoters could be upregulated in different cell lines (**Figure 19a**). However, transcriptional activation by VP64 is mediated by the direct recruitment of the transcription machinery (Hirai et al., 2011). To test whether expression of the reporter gene could be enhanced by chromatin regulation, the bromodomain containing protein 4 (BRD4) was recruited to the TRE3G promoter in

NIH/3T3 cells (**Figure 19a**). BRD4 is a histone acetylation reader that activates gene expression by recruitment of other epigenetic coregulators, but does not execute catalytic activity on chromatin (Rahman et al., 2011). Targeting of the *TRE3G-mCherry* reporter with rTetR-BRD4 failed to activate transcription of the fluorophore, while the same reporter cells could be activated by rTetR-VP64 (**Figure 19a**). By stable recruitment of TetR-BRD4 to the constitutive EF1A promoter, expression of the synthetic reporter gene could be increased by a factor of 2.8 (**Figure 19b**). However, recruitment of the TetR without a fused effector domain to the same promoter already influenced fluorophore expression.
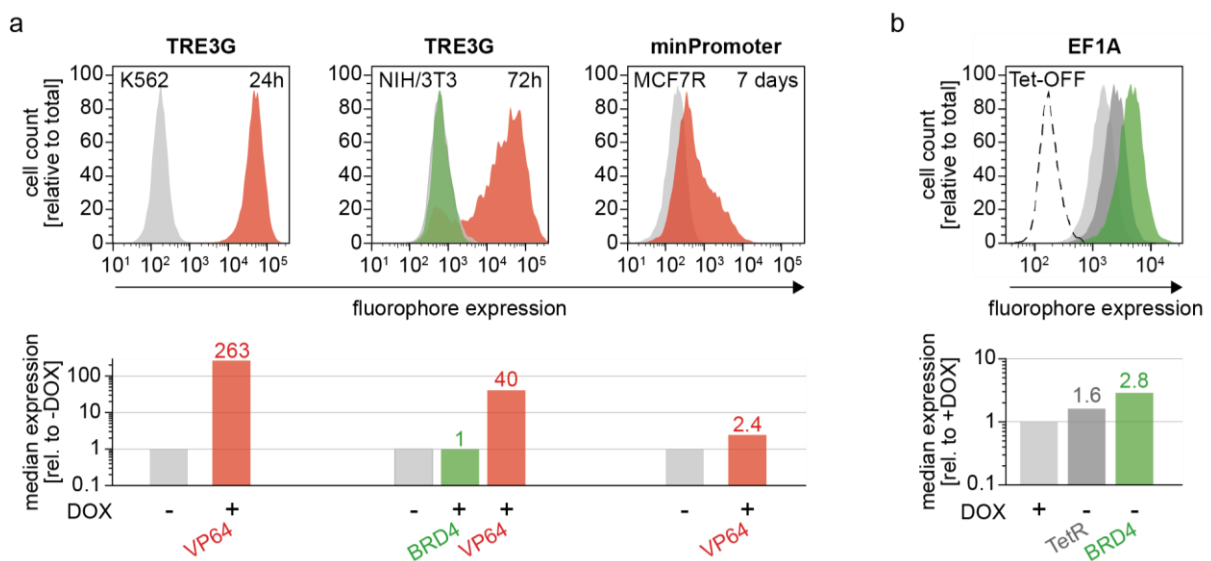
Figure 19: Different approaches for the establishment of a reporter system monitoring the activation of transcription. a, Representative flow cytometry analysis of different reporter cell lines with either a *TRE3G* or a *minimal promoter* controlling expression of the fluorophore. Top: histograms showing the induction of reporter fluorophore expression upon DOX-recruitment of rTetR-VP64 for the indicated time periods. Bottom: Bar graphs correspond to the histograms shown in the top row. Baseline reporter expression is shown in grey (-DOX). Numbers on top of the bar graphs indicate the fold-increase of expression upon recruitment of the rTetR-effector. **b,** Representative flow cytometry analysis of NIH/3T3 cells expressing *tGFP* under the constitutive *EF1A* promoter. Using the Tet-OFF system, either TetR alone (dark grey histogram) or a TetR-BRD4 fusion construct (green histogram) were stably recruited to the promoter (-DOX). Addition of DOX inhibited recruitment of either construct to *EF1A* (light grey histogram). The dotted histogram represents NIH/3T3 wt cells. Bottom: Bar graphs corresponding to the histograms. Numbers on top of the bar graphs indicate the fold increase of expression relative to +DOX.

In conclusion, a significant activation of transcription from minimal promoters could only be achieved by recruitment of VP64 to different promoter sequences (**Figure 19**). VP64 has extensively been used as transcriptional transactivator for a long time (Gossen et al., 1995; Hirai et al., 2011), but its activation mechanism does not involve the assembly of chromatin regulatory complexes. The effects of the recruitment of

BRD4 to a transcriptionally silent promoter that were observed so far were weak and hard to reproduce. To test the sensitivity of the reporter system to report on the activity of epigenetic effector proteins enhancing transcription, recruitment of catalytically active epigenetic regulators like histone acetyltransferases (HATs) might be necessary (Hilton et al., 2015).

4.2.   Proof of principle

The repressive activities of the lysine-specific demethylase LSD1 at the fluorescent reporter system were in the focus of this thesis, as this protein is a good example for a multi-functional and highly networking chromatin effector (Maiques-Diaz & Somervaille, 2016; Perillo et al., 2020). LSD1 is known to fulfil a variety of tasks in transcriptional repression, depending on its association with different multiprotein complexes (Maiques-Diaz & Somervaille, 2016). Having established a reporter system that could be silenced by the recruitment of different chromatin effectors including LSD1, the next aim was to analyse how this transcriptional repression was mechanistically mediated in the cells. In the course of this study, the relevance of the catalytic activity of LSD1 and the activity of known coregulators for the induced repression of the reporter gene by LSD1 were determined. A particular focus was set on the sensitivity of the fluorescent reporter to detect manipulations of the endogenous coregulator environment, to validate the applicability of the system to screen for functional dependencies of chromatin effectors.

4.2.1. Investigation of the effect of the catalytic activity of LSD1 on gene silencing

LSD1 acts as an epigenetic repressor of transcription via its catalytic activity by removing H3K4 mono- and dimethylation, which are a hallmark of active chromatin (Maiques-Diaz & Somervaille, 2016; Shi et al., 2004). The activity of LSD1 is further linked to its association with multiple epigenetic regulatory complexes, like the NuRD, SIN3A and CoREST complex (Bornelöv et al., 2018; S.-A. Kim et al., 2020; Y. Wang et al., 2009; Y. Yang et al., 2018). Through interaction with SNAG domain containing transcriptional repressors, LSD1 can repress gene expression independently of its own catalytic activity (Y. Lin et al., 2010; Maiques-Diaz et al., 2018).

To test the influence of the catalytic activity of LSD1 at the synthetic promoter, the dynamics of repression following the recruitment of either LSD1 wt or the catalytic mutant LSD1 K661A were compared (**Figure 20a**). The recruitment of an rTetR-LSD1

fusion protein to the *synP-mCherry* reporter resulted in the reduction of reporter fluorophore expression to ~10% of the initial signal over the course of 12 days (**Figure 20a**). The repression observed upon recruitment of a catalytically inactive mutant (K661A) (Stavropoulos et al., 2006) was much less pronounced, although both constructs were expressed at similar levels (**Figure 20b**).

To analyse the effects of rTetR-LSD1 wt or K661A recruitment on the local chromatin environment of the *synP* element, changes in histone tail modifications were analysed in chromatin immunoprecipitation (ChIP) experiments followed by qPCR for the promoter sequence (**Figure 20c**). As expected, the reduction of H3K4 dimethylation upon recruitment of LSD1 was abolished by the mutation of the catalytic domain of the demethylase.
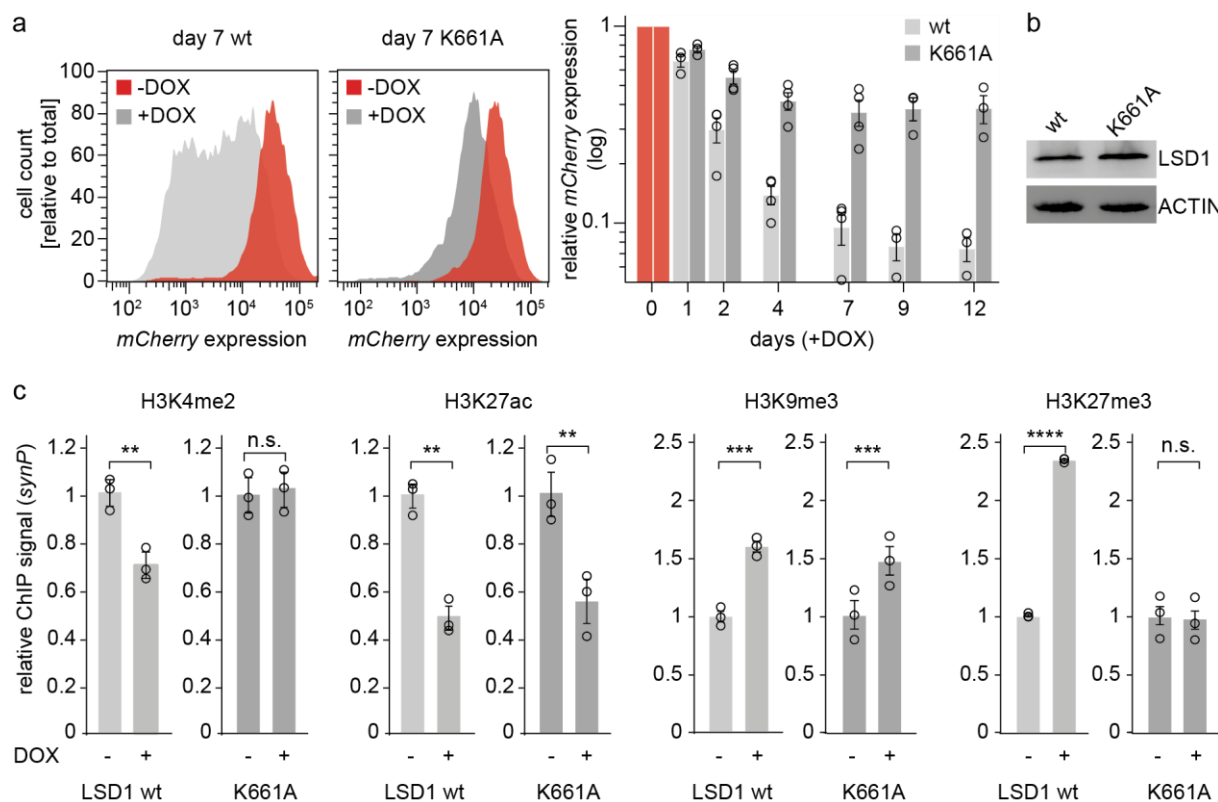
Figure 20: Characterization of LSD1 activity at the synthetic reporter. a, Flow cytometry analysis of NIH/3T3 *synP-mCherry* reporter cells during recruitment of rTetR-LSD1 wt or the inactive rTetR-LSD1 mutant (K661A). Histograms show representative reporter expression profiles at day 7 with/without recruitment of rTetR-LSD1 (±DOX). Bars represent the median *mCherry* expression relative to day 0 (n = 4, mean ± S.E.M.). **b,** Immunodetection of the expression levels of the rTetR-LSD1 wt or K661A. Whole-cell lysate of NIH/3T3 cells expressing the *synP-mCherry* reporter and the rTetR-LSD1 fusion protein was loaded. ACTIN was used as a loading control. **c,** Analysis of histone modification changes at the *synP* element upon recruitment of rTetR-LSD1 wt or K661A using ChIP-qPCR. Inputs for ChIP were isolated from NIH/3T3 reporter cells at day 4 (H3K4me2, H3K27ac, H3K27me3) or day 14 (H3K9me3) of DOX treatment. (n = 3, mean ± S.E.M., p*≤0.05, p**≤0.01, p***≤0.001, Student´s *t*-test. n.s. = *no significance*). Panel (**b**) and (**c**) were taken from Pinter et al., 2021.

Interestingly, even the LSD1 wt only reduced the level of H3K4me2 to ~70% of the initial signal. This might be due to the fact that H3K4me2 is an intermediate modification state of H3K4. Active transcription is coupled to the recruitment of H3K4 histone lysine methyltransferases (HMTs) (Krogan et al., 2003; B. Li et al., 2007; Morillon et al., 2005) which possibly counterbalance the activity of LSD1 on H3K4me2 by re-introducing methylation marks as long as the reporter gene is transcribed. The catalytic activity of LSD1 is heavily counter-regulated by endogenous factors and thus is supposedly not the only factor mediating transcriptional repression by LSD1 recruitment, which is in line with published data (Maiques-Diaz et al., 2018).

To determine which other epigenetic effectors could contribute to the observed repression of the reporter gene following LSD1 recruitment, changes in other histone tail modifications were analysed. ChIP qPCR for H3K9me3, H3K27me3 and H3K27ac revealed effects caused by co-recruitment of various other coregulatory proteins (**Figure 20c**). Recruitment of both the wild type and mutant LSD1 led to a decrease in H3K27 acetylation and a mild increase in H3K9 trimethylation, suggesting co-recruitment of HDAC- and H3K9 HMT-complexes (Maiques-Diaz & Somervaille, 2016; Y. Yang et al., 2018) to the promoter. This recruitment was independent of the catalytic activity of LSD1. A striking difference between the wt and the mutant LSD1 was found for H3K27 methylation. Recruitment of LSD1 wt led to a strong increase in H3K27 trimethylation, while recruitment of the K661A mutant failed to increase H3K27me3. This suggested that efficient silencing of expression from the synthetic promoter was dependent on the co-recruitment or co-activation of the PRC2 complex after binding of LSD1 at this locus. Further experiments supporting this hypothesis are described in the sections 4.2.2 and 4.3.3 and in the published manuscript in **Appendix 1**.

4.2.2. Application of the fluorescent reporter system for epigenetic drug screening

Many epigenetic regulators have been identified as promising targets for novel therapies in cancer and other diseases (Bates, 2020; Linhares et al., 2020; Olino et al., 2020). However, the search for efficient drugs targeting epigenetic effector proteins is still ongoing. High-throughput methods enabling the parallel screening of hundreds or thousands of chemical compounds have been used for decades, but the identification of highly specific and efficient molecules is still a challenge. Using

phenotypic screening methods in cellular systems or organisms, potent molecules preventing a specific malignant phenotype can be identified (Facciotto et al., 2019; Grüner et al., 2016; L. H. Jones & Bunnage, 2017; B. Wu et al., 2019). However, these methods do not necessarily provide information about the biological mechanisms that led to the observed phenotype, complicating the characterization of the drug's mode of action, which is a condition for its approval, rendering these methods less attractive for pharmaceutical companies (L. H. Jones & Bunnage, 2017). Pharma companies have been focussing on target-based approaches, where an identified candidate is easier to characterize. In those approaches, a known target important for the development of a specific disease is screened against a library of compounds (Drewes & Knapp, 2018; Markossian et al., 2018). Top hits are then further characterized for their pharmacological applicability as a drug.

All drug screening approaches require the combination with other methods to be able to characterize the effect of a compound molecule on the interaction profile, enzymatic activity, stability, localization, and transcription regulatory activity of a target protein. A compound screen for small-molecule inhibitors of epigenetic target proteins has been developed by Headley et al. (Headley et al., 2019), but this method represents rather a phenotypic than a target-based approach. LSD1 is an interesting drug target for many diseases (see section 1.4). To test whether the synthetic reporter system could be used for the identification of chemical compounds that interfere with LSD1's repressive activity, reporter cells were treated with several established chemical compounds targeting epigenetic effector proteins during recruitment of LSD1 (**Figure 21a-e**), and their effect on LSD1 silencing was monitored.
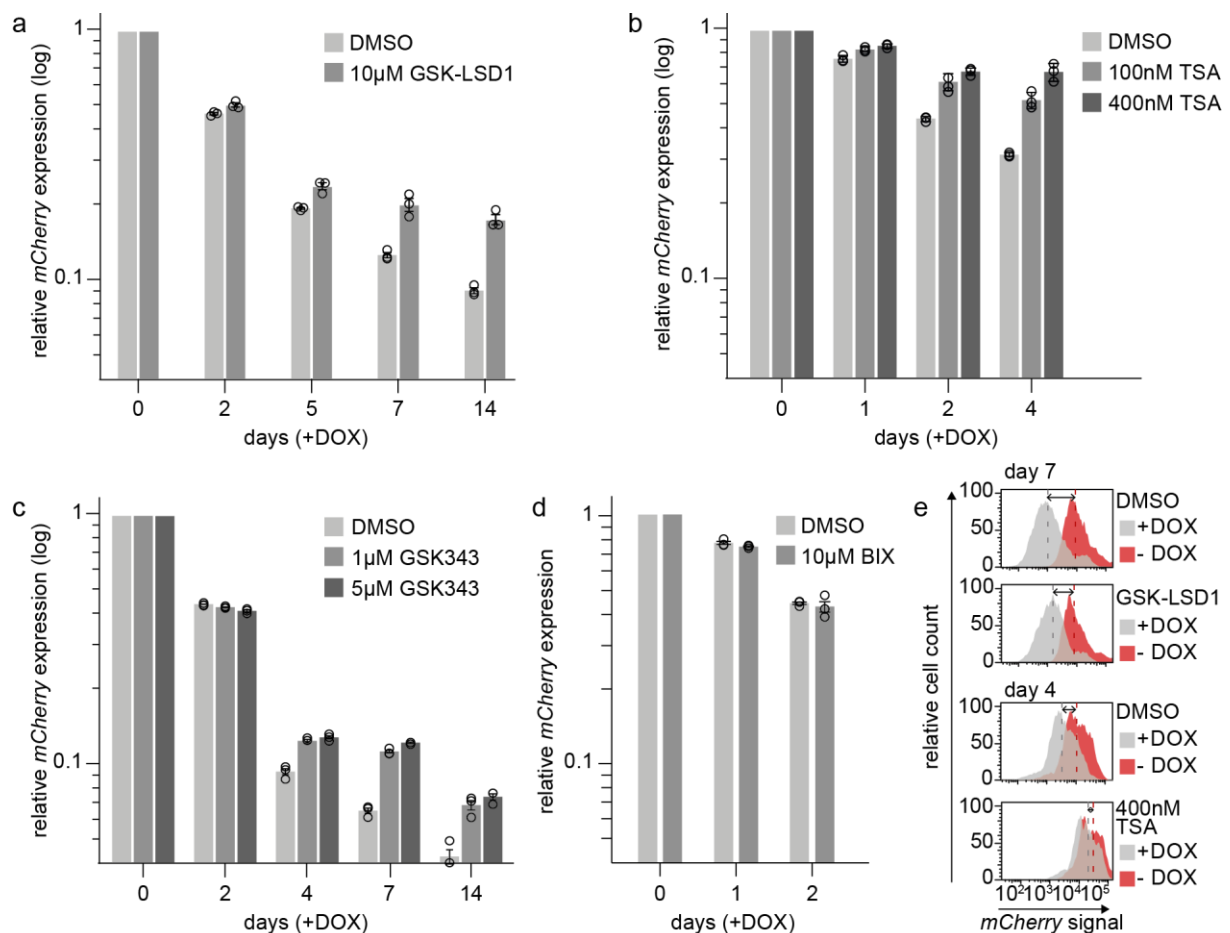
Figure 21: Chemical inhibition of endogenous coregulators impairs the silencing activity of LSD1 at *synP-mCherry*. a-d, NIH/3T3 *synP-mCherry* cells expressing an rTetR-LSD1 fusion protein were treated with small molecule inhibitors in parallel with recruitment of rTetR-LSD1. Treatment with an LSD1 inhibitor (**a**), an HDAC inhibitor (**b**), an EZH2 inhibitor (**c**) or a G9A/GLP inhibitor (**d**) was started with the addition of DOX. DMSO was used as control. Bars represent the median *mCherry* signal measured with flow cytometry, relative to cells with inhibitor treatment, but without DOX (n=3; mean±S.E.M.). Circles indicate independent replicates. Sorted cell populations were used for the experiment shown in **c. e,** Representative flow cytometry analysis of NIH/3T3 cells corresponding to the data shown in **a** and **b**. The dotted lines indicate the median *mCherry* signal.

The selective small molecule inhibitor of LSD1, GSK-LSD1 (Alsaqer et al., 2017; Takagi et al., 2017), reduced the repressive effect of LSD1-recruitment to the synthetic promoter, but did not fully prevent the silencing activity (**Figure 21a, e**). This was in line with published data, stating that the repressive activity of LSD1 is not only conferred by its catalytic activity on H3K4me1/2 (Maiques-Diaz et al., 2018) and confirmed the observations from the presented ChIP experiments (**Figure 20**). Interestingly, inhibition of HDAC activity with Trichostatin A (TSA), a pan-HDAC inhibitor (Yoshida et al., 1990), strongly interfered with LSD1-mediated silencing in a concentration-dependent manner, starting at day 2 of DOX induced recruitment (**Figure 21b, e**). This suggested that local histone deacetylase activity was an early

and important factor for the repressive function of LSD1. Histone deacetylation is a general hallmark of chromatin remodelling and heterochromatin formation (Adams et al., 2018; Bornelöv et al., 2018; Saksouk et al., 2015). Furthermore, the enzymatic activities of LSD1 and HDAC have been reported to be coupled directly by physical association of the proteins in the CoREST complex, and HDACs are found as subunits of multiple other repressive complexes containing LSD1 (Y. Song et al., 2020; Y. Wang et al., 2009; Y. Yang et al., 2018).

Having demonstrated that the reporter system was sensitive to monitor the inhibition of either LSD1 or a tightly associated complex partner, silencing in the presence of further small molecule inhibitors targeting different epigenetic complexes was tested. No influence of the H3K9-specific methyltransferase activity of G9A/GLP was observed during inhibition with the small-molecule inhibitor BIX01294 (**Figure 21d**) (Soumyanarayanan & Dymock, 2016). However, the molecule showed toxic effects after 4 days of treatment with this concentration. Titration of the compound should be considered to evaluate the effect on LSD1 reporter silencing.

Interestingly, inhibition of the PRC2 complex with GSK-343, a competitive inhibitor of EZH2 (Soumyanarayanan & Dymock, 2016), reduced silencing of reporter expression by LSD1 (**Figure 21c**). This observation further supported the hypothesis of a co-activation of PRC2 at the synthetic reporter upon recruitment of LSD1. Furthermore, the findings validated the reporter system as a tool for high-throughput screenings of unknown chemical inhibitors for their activity on LSD1 mediated silencing. Due to the modular structure of the reporter system, it could easily be adapted for the screening of small molecule inhibitors of other chromatin effectors.

4.2.3. Investigation of the roles of essential LSD1 coregulators using RNAi

Following the observation that other epigenetic coregulators were involved in the repression of the synthetic promoter by recruitment of LSD1, the effect of shRNA interference with specific known cofactors of LSD1 on reporter silencing was monitored. In theory, the suppression of an essential LSD1 interaction partner using RNAi should lead to a reduced activity of LSD1 and in result to a stabilization of *mCherry* expression in spite of the DOX-induced recruitment of LSD1 (**Figure 22a**). Different control shRNAs targeting either expression of the rTetR-LSD1 construct itself

(sh*LSD1* and sh*rTetR*) or the known interactors of LSD1, *Chd5* and *Sap130* (Kolla et al., 2015; Y. Yang et al., 2018), were stably expressed in NIH/3T3 *synP-mCherry* reporter cells. *Chd5* is a homologue of the NuRD subunits *Chd3* and *Chd4*, and has chromatin remodelling, helicase and DNA binding function (Kolla et al., 2015). *Sap130* is a poorly characterized corepressor protein that is tightly associated with the SIN3A/HDAC complex (Fleischer et al., 2003; Y. Yang et al., 2018). The dynamics of reporter silencing upon recruitment of rTetR-LSD1 under suppression of the respective coregulators was followed using flow cytometry. Compared to the silencing progress in cells expressing a neutral control shRNA (sh*Control*), silencing under suppression of the complex partners *Chd5* and *Sap130* was significantly impaired (**Figure 22b, c**). The effect of suppression of these complex partners on the strength of repression was comparable to the effect caused by a direct targeting of the rTetR-LSD1 construct by sh*LSD1* or sh*rTetR* (**Figure 22c**). This illustrated the potential of the system to screen for and characterize mechanistic connections that contribute to the concerted epigenetic repression of gene expression.
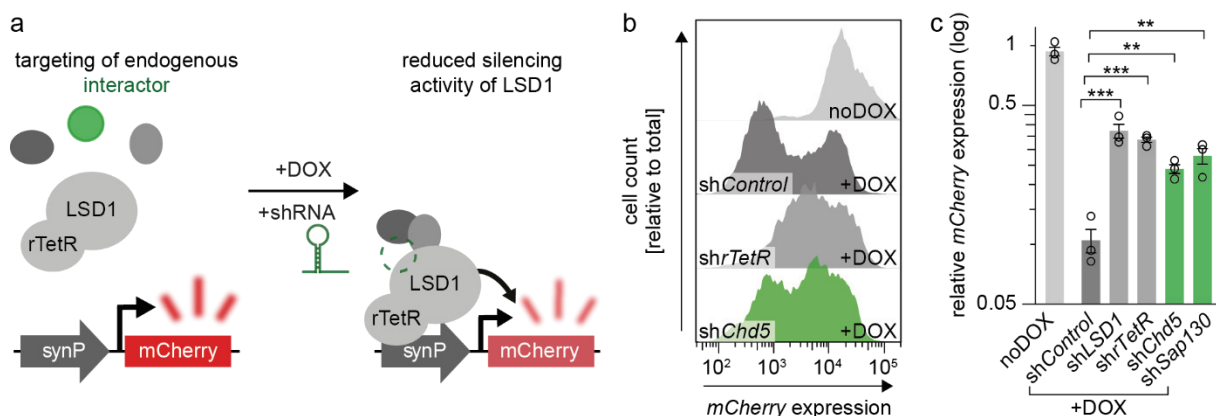
Figure 22: Investigation of coregulators of LSD1 using RNAi. a, Schematic representation of the targeted suppression of LSD1-interacting proteins using RNAi. The association of rTetR-LSD1 with endogenous complex partners (grey and green ovals) is essential for efficient reporter silencing. The presence of a known interaction partner is selectively reduced via expression of specific shRNAs (green). Endogenous proteins essential for the silencing of *mCherry* can be identified by a reduced silencing activity of LSD1 following DOX recruitment, that is observed in flow-cytometry analysis. **b,** Representative flow cytometry analysis of reporter fluorophore expression after 7 days of rTetR-LSD1 recruitment to *synP-mCherry*. NIH/3T3 cells stably expressed the reporter system and the indicated shRNAs for 10 days before addition of DOX. Untreated *synP-mCherry* cells are shown as a control (noDOX). **c,** Bar graph depicting the residual *mCherry* signal upon recruitment of rTetR-LSD1 for 7 days in cells expressing the indicated shRNAs. Bars show the median reporter expression relative to day 0, corresponding to the examples shown in (**b**). *shControl*: non-targeting shRNA; *shLSD1* and *shrTetR*: shRNAs suppressing the rTetR-LSD1 construct; *shChd5* and *shSap130*: shRNAs targeting known interactors of LSD1. The variance of expression in cells without DOX is shown representatively for 3 replicates of cells expressing a control shRNA (*noDOX*). (n=3, mean±S.E.M., Student´s *t*-test, p**≤0.01, p≤0.001).

In conclusion, the reporter system could successfully be applied to detect functional dependencies of the silencing activity of LDS1 on associated complex partners, both by chemical inhibition and by suppression with RNAi. This sensitivity validated the system for the application in high-throughput RNAi screens for the identification of coregulators of LSD1 as a transcriptional repressor.

4.3. A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade

Having established that the system can be used to identify functional dependencies of chromatin regulation, the next step was to expand this function towards a high-throughput screening method. To this end, the fluorescent reporter cell lines were combined with the expression of a chromatin-focused shRNA library, to screen for functional interaction partners of LSD1 that contribute to its repressive activity. Most of the results presented in this chapter, together with the establishment of the method, were described in a paper that was published in Nucleic Acids Research (Pinter et al., 2021) and is attached in **Appendix 1**.

As described before, LSD1 was chosen as primary candidate in the screening for functional interaction partners for multiple reasons. Although the reaction mechanism and substrate specificity of LSD1 have been studied extensively since its discovery as a histone demethylase (Culhane & Cole, 2007; Forneris et al., 2007; F. Gu et al., 2020; Maiques-Diaz & Somervaille, 2016; Metzger et al., 2005), the diversity of gene regulatory functions engaged by LSD1 remains to be fully elucidated. Its dual function as a demethylase of both repressive and active histone modifications, its integration in a multiplicity of epigenetic complexes and its regulatory function as subunit of multiprotein complexes (Forneris et al., 2007; Laurent et al., 2015; Maiques-Diaz et al., 2018; Maiques-Diaz & Somervaille, 2016; Metzger et al., 2005; Zibetti et al., 2010) indicate that LSD1 is tightly embedded in a network of epigenetic regulation mechanisms. This renders the demethylase an interesting target for epigenetic network analysis. The persistent gaps in the knowledge about the activities of LSD1 as a transcriptional regulator, despite almost two decades of intensive studies by many groups, suggest that common experimental approaches do not suffice to fully capture the different modes of action of LSD1, and possibly other chromatin effectors.

4.3.1. A chromatin focused RNAi screen detects novel and known coregulators of LSD1 activity

Aiming to gain novel insights into the epigenetic network surrounding LSD1 as a transcriptional repressor, an RNAi screen was performed. The RNAi screen was designed to analyse a comprehensive set of possible coregulators of a given chromatin effector of interest (LSD1 in this case) in a highly parallel way. To this end, the fluorescent reporter system was combined with the expression of a shRNA library (Fellmann et al., 2013) that comprised shRNAs targeting 1104 chromatin-related genes with 5-6 different hairpins per gene (Rathert et al., 2015). The screening procedure is described in section 3.3 and **Appendix 1**. In brief, NIH/3T3 cells stably expressing *synP-mCherry* and rTetR-LSD1 were virally transduced with the shRNA library in a way that led to the integration of a single shRNA construct into each cell (**Figure 23**). Cells expressing an shRNA construct were enriched using antibiotic selection. Expression of the shRNA construct could be monitored during the entire experiement by measuring the expression of GFP from the same vector. DOX treatment was started 7 days after transduction of the cells with the shRNA library, and the dynamics of *mCherry* expression was followed by flow cytometry over the course of 14 days of rTetR-LSD1 recruitment. After this time, *mCherry* expression was silenced in most of the population. Cells that showed impaired LSD1 activity (*mCherry+*) were enriched using FACS sorting (**Figure 23**). The main population, which was silenced upon LSD1 recruitment (*mCherry-*), was also collected as a control. Genomic DNA was isolated from both populations, and shRNA guide sequences present in the populations were amplified and prepared for Illumina sequencing using PCR. The pooled library plasmids used for transduction of the cells were also amplified to determine the input representation of each shRNA. The amplified shRNA guides from the sorted *positive* and *negative* populations and from the input pool were sent for Illumina sequencing and analysed for representation of each shRNA. To be able to determine the statistical significance of any observed enrichment effects, five replicates were conducted independently, and the data was analysed in parallel.
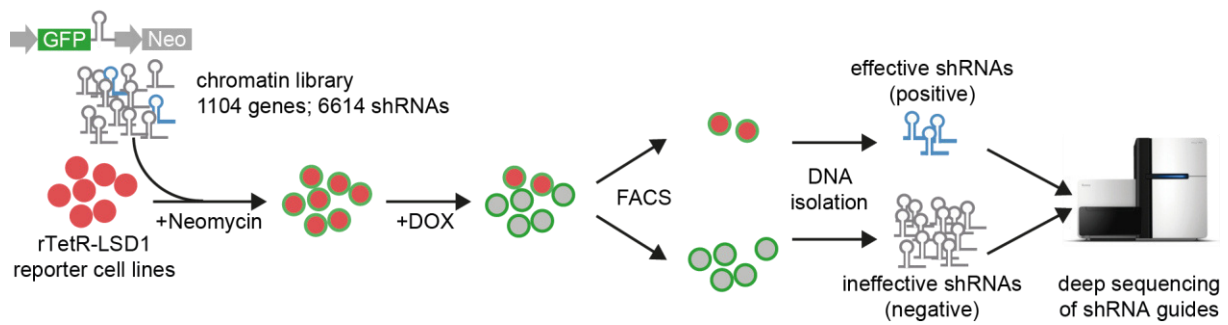
Figure 23: Workflow for the shRNA screening procedure. A chromatin-focused shRNA library is virally transduced into NIH/3T3 cells expressing the *synP-mCherry* reporter and rTetR-LSD1. After antibiotic selection for integration of the shRNA construct and DOX-induced recruitment of rTetR-LSD1 to *synP-mCherry* for 14 days, cells are FACS-sorted into silenced (grey) and active (red) populations. Genomic DNA is isolated from both populations and the enrichment of shRNA guide sequences in each population is analysed using Illumina sequencing. The figure was adopted from Pinter et al., 2021.

After filtering of shRNAs with weak representation in the input pool out of the data set, gene scores were determined for the remaining 1010 genes. To this end, the shRNA enrichment scores (read counts in the *mCherry+* over *mCherry-* population) were averaged for all shRNAs targeting the same gene in each independent replicate before calculation of the final gene score by taking the average over all replicates (**Figure 24a**). The top 100 genes exerting either a positive or a negative effect on LSD1 silencing were analysed by functional enrichment analysis using the Cytoscape plug-in ClueGO (Bindea et al., 2009), to determine the biological processes involved in the repressive activity of LSD1 at the reporter gene (**Appendix 1, Figure 2c**). In line with published data describing the biological functions of LSD1, biological pathways identified in the RNAi screen included lysine acetylation (Y. Yang et al., 2018), lysine methylation (M. G. Lee et al., 2005; Shi et al., 2004), chromatin silencing (Jing Wang et al., 2009) and chromatin remodelling (Y. Wang et al., 2009). Interestingly, "*DNA conformation change*" was also identified as a biological process influencing the silencing of the reporter gene by LSD1 recruitment The ClueGo analysis was performed by Philipp Rathert.

On the single gene level, the screen identified several known interactors of LSD1 as top scoring hits (**Figure 24b**): *Gatad2b* as part of the NuRD complex (Bornelöv et al., 2018; Y. Wang et al., 2009), *Sap25* as part of the SIN3A/HDAC complex (Shiio et al., 2006; Y. Yang et al., 2018) and *Dnmt3a* as a direct interactor (Petell et al., 2016). The candidate gene exhibiting the most significant enrichment among the top 10 hits, indicating a consistently high scoring in each of the five biological replicates, was the DEAD-box RNA helicase *Ddx19a*. The influence of *Ddx19a* suppression on the

repressive activity of LSD1 at the *synP-mCherry* reporter was validated with two independent shRNAs in single RNAi experiments (**Figure 24c, d**). However, the functional connection of this helicase with LSD1 silencing remained unclear.
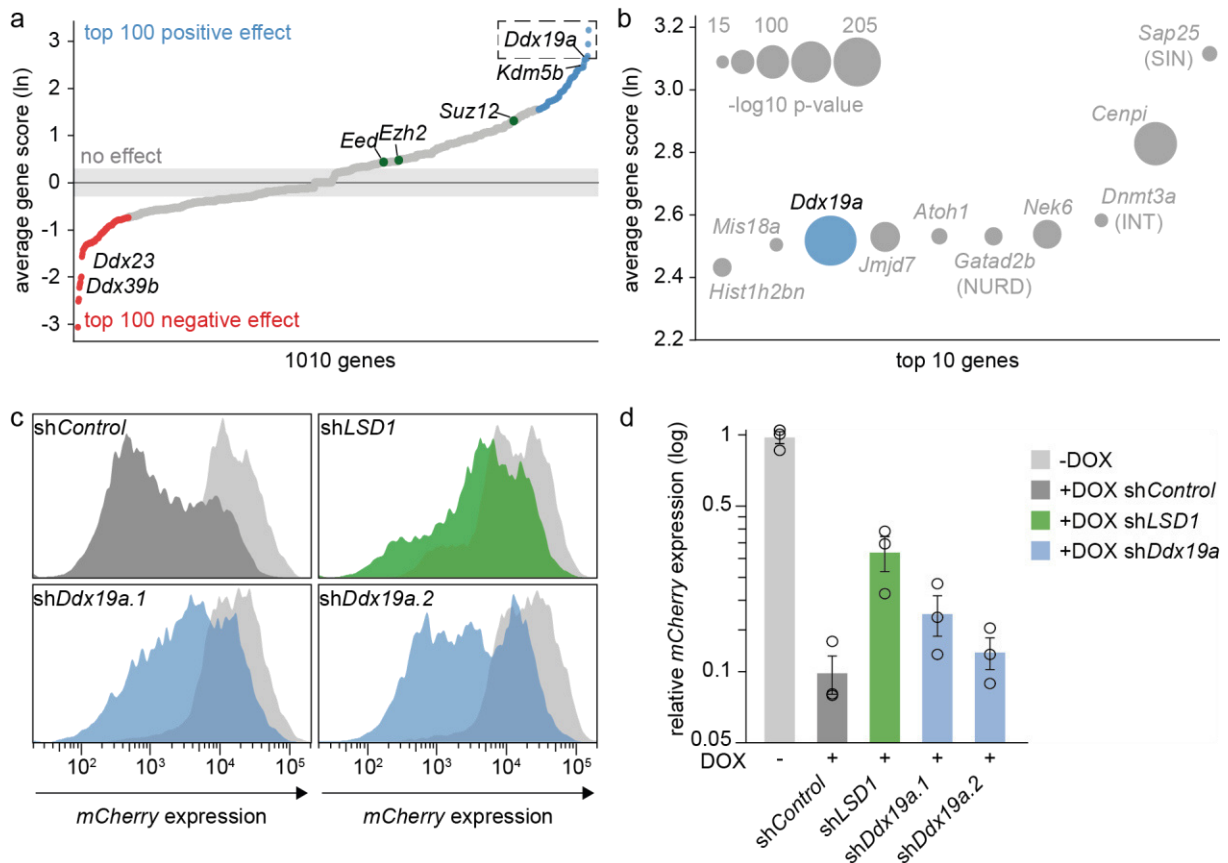

Figure 24: RNAi screening for interaction partners of LSD1 in NIH/3T3 cells. a, Scatter plot showing the effect of all genes included in the shRNA library on LSD1-induced silencing. The gene score represents the Ln of the average enrichment score (= read ratio *mCherry+*/*mCherry-* population) of all shRNAs per gene across five independent replicates. Genes imposing a positive effect on LSD1 induced silencing are coloured in blue, genes having a negative effect are highlighted in red. The position of PRC2 genes mentioned later in this thesis is highlighted. **b**, Close-up from **a,** showing the top 10 candidates identified in the screen. Spot size represents the significance of enrichment over all replicates. **c,** Validation of two shRNAs targeting *Ddx19a* (*shDdx19a.1/2*). Histograms show representative *mCherry* expression profiles in rTetR-LSD1 reporter cell populations treated with DOX for 7 days. Untreated populations are shown for comparison (light grey histograms) An shRNA directly targeting rTetR-LSD1 was used as a positive control. **d,** Bar graphs showing the median *mCherry* signals corresponding to the histograms in **c,**. Values are relative to day 0. Circles represent individual replicates (n=3; mean±s.e.m.). Figures were adopted from Pinter et al., 2021.

In general, proteins from the DEAD-box containing helicase family have RNA unwinding activity and play a role in several cellular processes involving mRNA processing and nuclear export (Alcázar-Román et al., 2006; Mikhailova et al., 2017;

Okamura et al., 2018; Zolotukhin et al., 2009). DDX19A had not yet been described as an interactor of LSD1 in interaction analyses performed by other groups (Malovannaya et al., 2011; Sehrawat et al., 2018), thus it represented an interesting candidate for further characterization. Regarding the function of DEAD-box helicases in mRNA processing and -export, the chance that the discovery of DDX19A as a top hit in the screen was a false-positive result originating from changed protein levels of parts of the reporter constructs had to be ruled out. The levels of both rTetR-LSD1 and *mCherry* in NIH/3T3 *synP-mCherry* cells stably expressing two different shRNAs targeting *Ddx19a* were compared to cells expressing a control shRNA (**Figure 25**). This revealed that neither the baseline expression of the rTetR-LSD1 effector, nor of the reporter fluorophore was changed in cells with suppression of *Ddx19a* (**Figure 25a, b**), ruling out a false-positive result on this account. The specificity of *Ddx19a* suppression for LSD1 mediated silencing was further supported by experiments with recruitment of rTetR-KRAB to the *synP* element, in which the suppression of *Ddx19a* did not influence the silencing dynamics following KRAB binding (**Figure 25c**).

As mentioned before, DDX19A had not emerged as an interactor in published coprecipitation experiments analysing interaction partners of LSD1 (Malovannaya et al., 2011; Sehrawat et al., 2018). To validate these data, coprecipitation experiments detecting DDX19A after pulldown of LSD1 from NIH/3T3 cell lysate was performed using different methods. These experiments also failed to detect a physical interaction of the two proteins (**Figure 25d**), suggesting that a direct physical interaction of LSD1 and DDX19A in NIH/3T3 cells was either too weak to detect or completely absent.
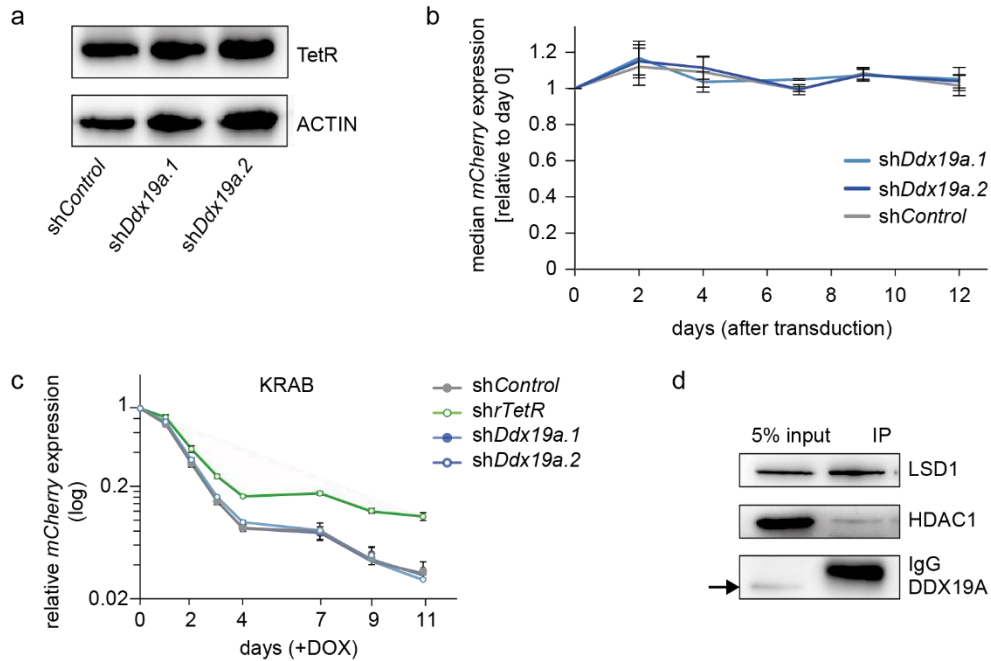
Figure 25: DDX19A does not interfere with the expression of reporter components and is not a strong physical interactor of LSD1. a, Immunodetection of the rTetR-LSD1 fusion construct in whole cell lysate of NIH/3T3 reporter cells expressing the indicated shRNAs. Cells were harvested 13 days after transduction with the shRNA construct. ACTIN was used as a loading control. **b,** Stability of reporter fluorophore expression in NIH/3T3 *synP-mCherry* cells expressing the indicated shRNAs without DOX treatment. Median *mCherry* signals are shown relative to the day of transduction with the shRNA constructs (n = 3, mean±*s.e.m.*). **c,** Dynamics of reporter silencing under recruitment of rTetR-KRAB with suppression of *Ddx19a*. NIH/3T3 *synP-mCherry* cells stably expressing the indicated shRNAs were treated with DOX and analysed using flow cytometry every 2-3 days. The median *mCherry* signal is plotted (n=3; mean±s.e.m.). **d,** Representative Western-Blot analysis of proteins coprecipitated with LSD1 from NIH/3T3 whole cell lysate. HDAC1 is shown as known interactor of LSD1. *IgG* indicates the signal originating from the antibody used for IP. The arrow indicates the signal for DDX19A. The images were taken from Pinter et al., 2021.

The homologue protein of DDX19A, DDX19B, has been shown to unwind hetero-hybrids of one DNA strand and one RNA strand *in vitro,* as well as and in the nucleus upon DNA-damage signalling (Hodroj et al., 2017). This DNA:RNA hybrid unwinding activity was confirmed for DDX19A *in vitro* in our lab by Franziska Knodel (**Appendix 1, Figure 3E**). As LSD1 has been described to bind and be regulated by structured forms of RNA (Hirschi et al., 2016; Porro et al., 2014), the possibility of a functional axis connecting LSD1 and DDX19A over DNA:RNA hybrid structures was further investigated.

4.3.2. DDX19A interferes with LSD1 activity by regulating R-loop homeostasis

R-loops are specific three-stranded DNA:RNA hybrid structures that arise during transcription, when the nascent mRNA hybridizes with the template DNA strand (Ginno et al., 2012). Various mechanisms have been proposed in which R-loops can influence transcription, *e.g.* by enhancing RNA Pol II termination, by regulation of the local epigenetic environment leading to chromatin decondensation, or by recruitment of both repressive and activating epigenetic complexes (Boque-Sastre et al., 2015; P. B. Chen et al., 2015; Ginno et al., 2013; Sanz et al., 2016; Skourti-Stathaki et al., 2019; Skourti-Stathaki & Proudfoot, 2014; C. Xu et al., 2021). Supposing that the formation of R-loops at the synthetic promoter might constitute the mechanistic link between DDX19A activity and the silencing activity of LSD1 at the reporter gene, the regulatory function of R-loops on the repressive activity of LSD1 was characterized.

To determine the physical association of LSD1 and DDX19A with R-loops, pulldown experiments were performed using the S9.6 antibody, which is a monoclonal antibody specifically raised against DNA:RNA hybrid structures (Sanz et al., 2016). Using this antibody for IP, both LSD1 and DDX19A were enriched from NIH/3T3 whole cell lysate (**Figure 26a**), suggesting a stable association of both proteins with R-loops under physiological conditions. Electrophoretic shift assays (EMSA) performed with recombinant LSD1 and different oligonucleotide structures indicated an increased affinity of LSD1 towards R-loops compared to double-stranded DNA (**Figure 26b**). To investigate the effect of *Ddx19a* gene suppression on global R-loop levels, immunofluorescence microscopy experiments were performed in NIH/3T3 cells expressing shRNAs targeting *Ddx19a* (**Figure 26c, d**). The experiment confirmed the observation that DDX19A can resolve R-loops *in vitro*, with lower levels of DDX19A in the cells leading to an increase in nuclear R-loop foci compared to the control. The global increase of R-loop formation as result of *Ddx19a* suppression was further confirmed on DNA level by DNA:RNA hybrid immunoprecipitation (DRIP) followed by qPCR for endogenous genomic loci (**Figure 26e**). Together, these results indicated a role of DDX19A in the regulation of global R-loop formation in NIH/3T3 cells.
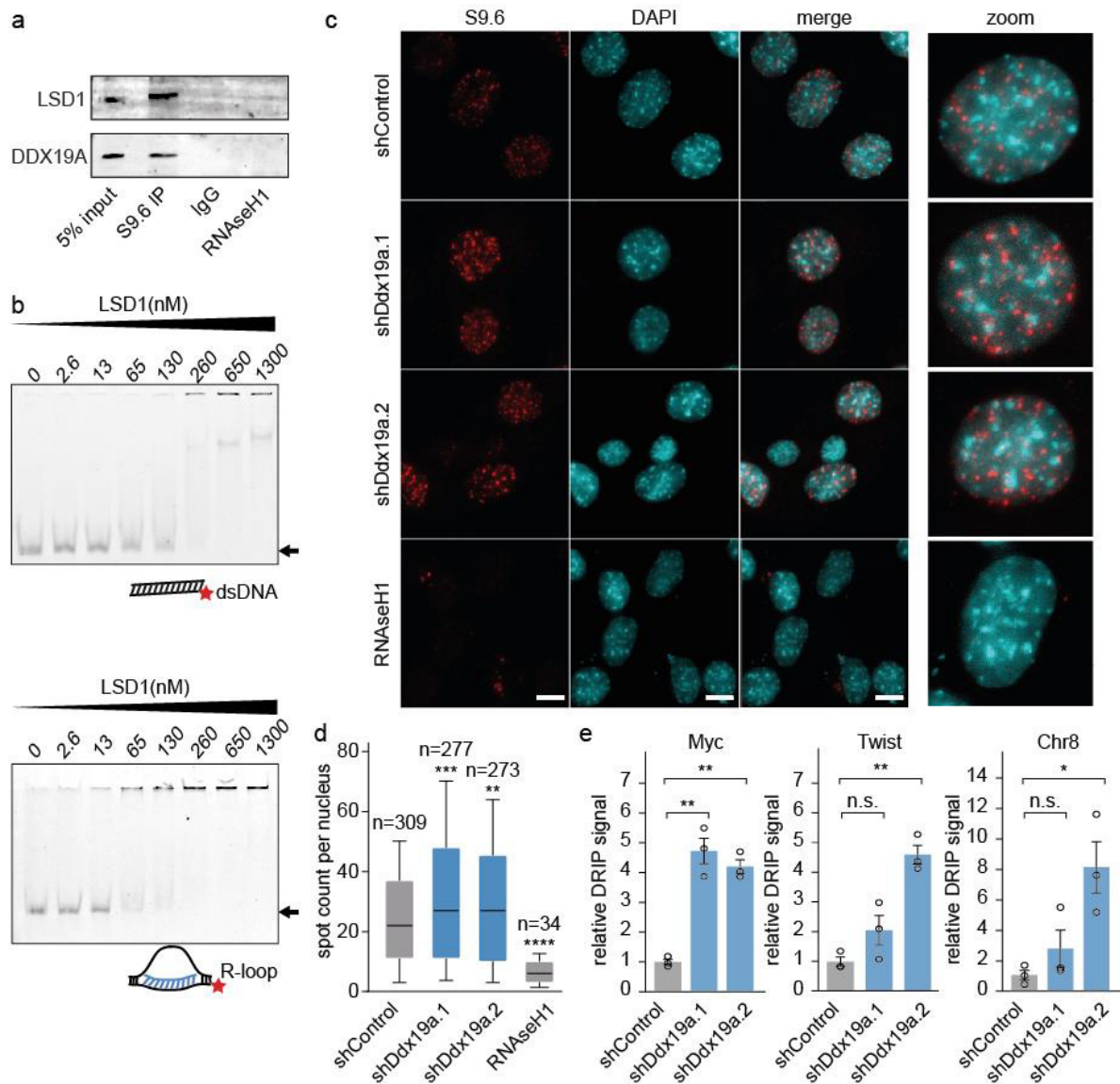
Figure 26: LSD1 and DDX19A show multiple modes of interaction with R-loops. a, Immunodetection of LSD1 and DDX19A in precipitate from NIH/3T3 whole cell lysate. The R-loop specific S9.6 antibody was used for IP. Mouse IgG was used as a negative control. Samples treated with RNAseH1 before IP with the S9.6 antibody serve as a control for antibody specificity. **b,** Representative polyacrylamide gels showing the binding of recombinant LSD1 to dsDNA or R-loops during EMSA. 26 fmol of a Cy5-labelled DNA oligo hybridized to either a complementary DNA strand (dsDNA) or a semi-complementary DNA strand and an intercalated complementary ssRNA (R-loop) were incubated with the indicated concentrations of LDS1 in a 20 µl reaction before gel electrophoresis. **c,** Representative immunofluorescence images of NIH/3T3 cells expressing the indicated shRNAs and stained with the S9.6 antibody. DAPI was used for visualization of nuclei. Transfection with RNAseH1 24 h before fixation was used as a negative control. Scale bar is 10 µm. **d,** Quantitative analysis of the images shown in **c,**. Nuclear spots were counted using the Cellprofiler software. Box-and-Whisker plots indicate the median and the 10-90 percentile from three independent experiments (p**≤0.01, p***≤0.001, p****≤0.0001, Student´s t-test). **e,** Analysis of R-loop formation at three endogenous loci upon suppression of *Ddx19a* expression using DRIP. Total nucleic acids were extracted from NIH/3T3 cells expressing the indicated shRNAs for 10 days and used as input for the IP with the S9.6 antibody. qPCR signals are shown relative to *shControl*. Circles represent independent replicates (n=3, mean±s.e.m.; p*≤0.05, p**≤0.01, n.s.=non-significant, Student´s *t*-test). The figure was adopted from Pinter et al., 2021, with exception of panel (**b**).

119

Having observed the physical association of LSD1 with R-loops *in vitro,* the next step was to characterize the nature of endogenous genomic loci that featured both LSD1 binding and R-loop structures in NIH/3T3. R-loops can form over actively transcribed, hypomethylated CpG-promoters that are associated with active chromatin-marks and -complexes (Ginno et al., 2012; Sanz et al., 2016). This seemed contradictory to the presence of LSD1 as an epigenetic repressor protein. To dissect this phenomenon, genomic loci that were both bound by LSD1 and were sites of R-loop formation were identified in NIH/3T3. To this end, LSD1 ChIP-seq was performed and the results were compared with publicly available data for genome-wide R-loop distribution (DRIP-seq, DRIPc-seq and RNAse-seq data; see **Appendix 1, Figure 4** and **Figure 27**). The final data analysis was performed by Philipp Rathert. A subset of LDS1 peaks correlated with regions of high DRIP signal (**Figure 27a, b**). Interestingly, these regions were associated with actively transcribed genes despite the repressive nature of LSD1 (**Figure 27b**). This phenomenon has already been discussed in literature (Garcia-Bassets et al., 2007; Maiques-Diaz et al., 2018; Whyte et al., 2012), suggesting a somehow "poised" or inactive state of LSD1 at these loci. Correlation of publicly available DRIPc-seq and RNA-seq data with the LSD1 ChIP-seq data further revealed, that R-loop formation and LSD1 binding co-occurred genome-wide at a subset of particularly highly transcribed genes (**Appendix 1, Figure 4D** and **Figure 27c**). All these observations were confirmed in the leukemia cell line K562, by analysis of publicly available data performed by Philipp Rathert (**Appendix 1, Figure 4**).

Together, these data suggested that the formation of R-loop structures at a subset of highly transcribed genes might serve as a regulatory OFF-switch for the activity of LSD1, poising it for fast and efficient repression of the respective gene upon R-loop removal following a trigger signal. This hypothesis was further supported by a Gene Ontology Analysis performed by Philipp Rathert. Correlation of the genes that were both identified in the LSD1 ChIP-seq and in the DRIP-seq data in K562 revealed, that these genes were associated with fundamental developmental signalling pathways (**Appendix 2, Figure S4**).
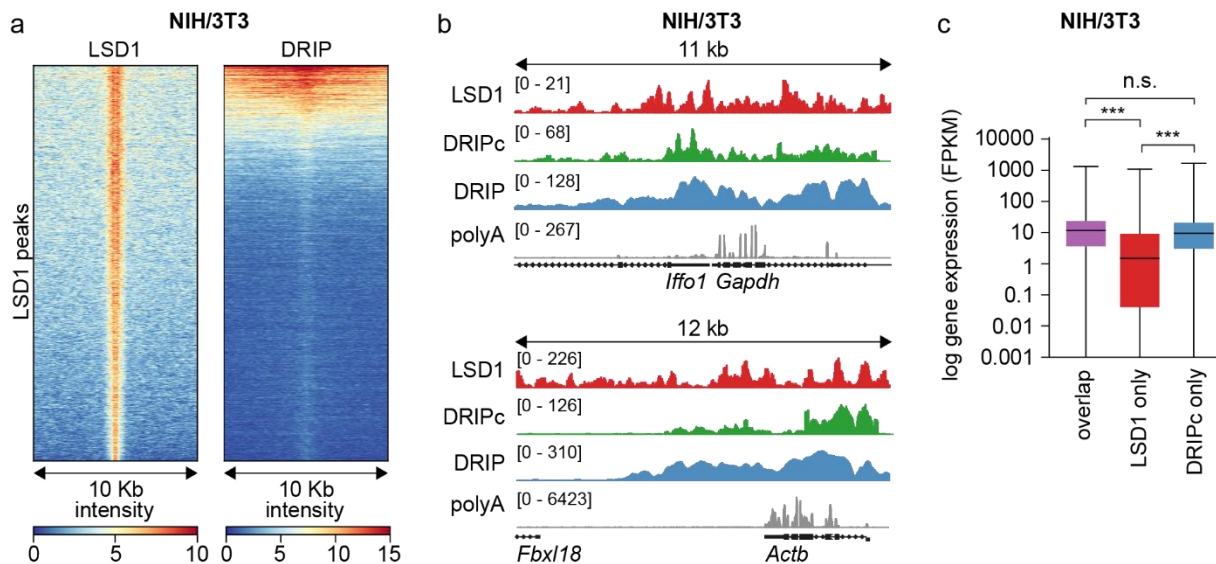
Figure 27: Regions of LSD1 occupancy and R-loop related features correlate genome wide in NIH/3T3 and K562 cells. a, Heatmaps of LSD1-ChIP and DRIP signals in NIH/3T3 cells. Signals are plotted on the heatmap within a 10 kb window around the peak centre. **b,** Representative genomic regions showing the occupancy of R-loop related features and LSD1 in NIH/3T3 cells. Signal tracks were obtained from published LSD1 ChIP-seq, DRIPc-seq, DRIP-seq and R-ChIP-seq data and compared to the LSD1 ChIP-seq performed in NIH/3T3. **c,** Expression levels (mRNA, FPKM) of genes associated with LSD1 and DRIP/DRIPc signals in NIH/3T3 cells (*P ≤ 0.05, **P ≤ 0.001, n.s. = non-significant, ordinary one-way ANOVA with multiple comparisons test). The figure was adopted from Pinter et al., 2021.

To gain further insights into the mechanisms connecting LSD1, DDX19A and R-loops with the regulation of gene expression, the effects of LSD1 and DDX19A on R-loop formation were investigated at the synthetic reporter gene. DRIP-qPCR experiments confirmed that the inhibition of LSD1 silencing upon *Ddx19a* suppression was accompanied by a stabilization of R-loops at the synthetic promoter (**Figure 28a**). This effect was dependent on the catalytic activity of LSD1, as R-loops were removed less efficiently from the promoter by recruitment of the mutant LSD1 compared to the wild type (**Figure 28a**). Consistently, no effect of *Ddx19a* suppression on reporter silencing was observed when recruiting the catalytically inactive K661A mutant of LSD1 to the reporter gene (**Figure 28b**). The importance of H3K4 demethylation by LSD1 for the genome-wide regulation of R-loop formation was confirmed by chemical inhibition of the catalytic activity of LSD1 in NIH/3T3 cells, which led to a global increase in R-loops (**Appendix 1, Figure 5D**).

Early experiments with the synthetic reporter had shown that the repression of reporter expression following the recruitment of rTetR-LSD1 was accompanied by a local decrease in H3K4me2 and H3K27ac, while H3K27me3 and H3K9me3 were increased (**Figure 20**). To determine the effect of the stabilization of R-loops at the *synP* element

on changes in the local chromatin environment, similar ChIP-qPCR experiments were performed in NIH/3T3 rTetR-LSD1 reporter cell lines with suppression of *Ddx19a* expression (**Figure 28c-f**). Demethylation of H3K4 and trimethylation of H3K27 following LSD1 recruitment were completely abolished in the cells upon suppression of *Ddx19a* (**Figure 28 c, d**), while the increase in H3K9me3 and the deacetylation of H3K27 were unchanged compared to the control (**Figure 28 e, f**). These results further supported the existence of a reciprocal effect between the demethylase activity of LSD1 and the formation of R-loops at a transcribed locus. They further suggested a connection of R-loop stabilization with the local activity of the PRC2 complex. This was in line with the existence of regulatory effects of RNA structures on PRC2 recruitment and activity that have been described in literature. However, the molecular mechanisms behind and the regulatory outcome of these interactions are still being widely discussed (Alecki et al., 2020; Almeida et al., 2020; Laugesen et al., 2019; Mocavini & Di Croce, 2020).
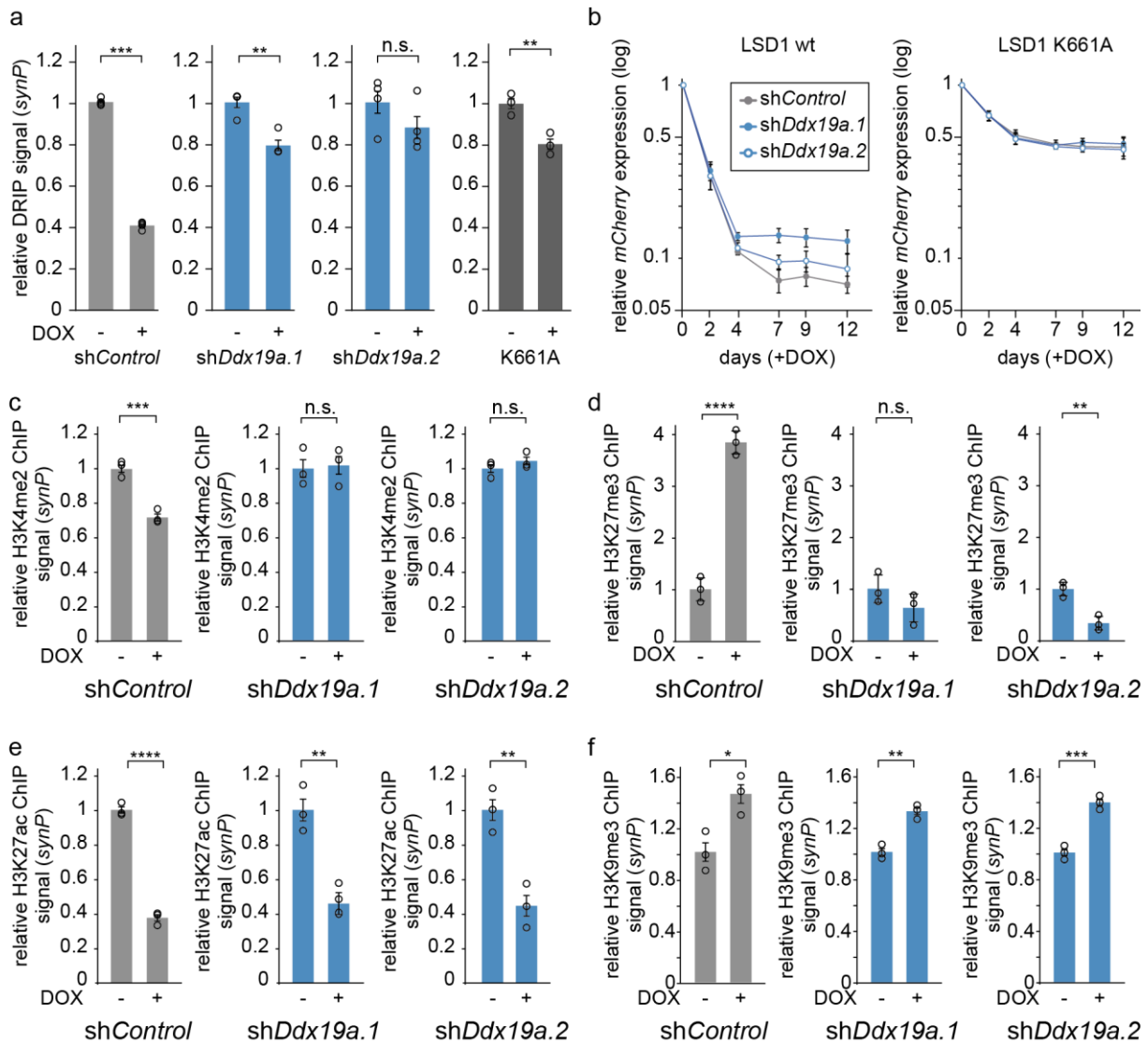
Figure 28: Effects of *Ddx19a* suppression on local chromatin regulation. a, DRIP-qPCR analysis of DNA:RNA hybrid structures at the *synP* element. DRIP was performed after recruitment of rTetR-LSD1 wt under suppression of *Ddx19a,* or after recruitment of the LSD1 K661A mutant. Total nucleic acids were extracted from NIH/3T3 cells expressing the indicated shRNAs at day 14 with and without DOX treatment, and were used as input for IP with the S9.6 antibody. qPCR signals are shown relative to -DOX. Circles represent independent replicates (n=4, mean±s.e.m.; p**≤0.01, p***≤0.001, n.s.=non-significant, Student´s t-test). **b,** Dynamics of fluorescent reporter silencing upon recruitment of LSD1 wt or K661A. NIH/3T3 reporter cells expressing the indicated shRNAs were treated with DOX and analysed every 2-3 days using flow cytometry. Circles indicate the median *mCherry* signal relative to day 0 (n=3, mean±s.e.m.). **c,-f,** ChIP-qPCR analysis of the indicated histone modifications at the *synP* element in cells expressing the LSD1 reporter system and the indicated shRNAs. IPs were performed with mononucleosomes isolated from NIH/3T3 reporter cell lines after 4 days (H3K4me2, H3K27ac) or 14 days (H3K9me3) of LSD1 recruitment. Bars are relative to -DOX (n=3; mean±s.e.m.; p*≤0.05, p**≤0.01, p***≤0.001, p****≤0.0001, n.s.=non-significant; Student´s t-test). The figure was adopted from Pinter et al., 2021.

4.3.3. The role of DDX19A in LSD1 mediated silencing of actively transcribed genes

Aiming to further characterize the role of DDX19A in the regulation of LSD1-bound promoter regions on a biochemical level, recombinantly expressed DDX19A was analyzed for its affinity towards different chromatin modifications. Incubation of MODified™ Histone Peptide Arrays (Active Motif) with DDX19A showed that the DEAD-box RNA helicase specifically recognized two repressive chromatin marks, H3K27me3 and H4K20me3 (**Figure 29a, b**). The methylation specific binding could further be confirmed on peptide SPOT arrays (**Figure 29c**). This was an interesting observation, as trimethylation of H3K27 had also been detected as a histone modification influencing LSD1-mediated silencing of the reporter system (**Figure 20b**) and PRC2 components scored as positive regulators of LSD1 in the screen (**Figure 24a**). The recognition of histone modifications by DEAD-box helicases had not been described so far. The binding affinity of DDX19A towards K27-methylated and -unmodified H3 peptides was confirmed in equilibrium peptide binding assays in the lab by M. Choudalakis, revealing a strikingly low K_D for DDX19A and H3K27me3 (**Appendix 1, Figure 6C**).

In mammalian cells, trimethylation of H3K27 is introduced by the Enhancer of Zeste (EZH2) protein as a member of the PRC2 complex (K. H. Hansen et al., 2008; Qian & Zhou, 2006) and an effect of the PRC2 complex on LSD1 mediated repression was observed earlier in this work (**Figure 21** and **Figure 20**). To determine the influence of EZH2/PRC2 activity on LSD1 activity at the reporter gene, LSD1 silencing of the *synP-mCherry* reporter was performed in the presence of the specific EZH2 inhibitor GSK343 (Soumyanarayanan & Dymock, 2016). As expected, the repression of the reporter gene by LSD1 was reduced when PRC2 was not functional (**Figure 29d**). Furthermore, the resolution of R-loops at the *synP* element and at endogenous genomic loci was shown to be dependent on the catalytic activity of EZH2 (**Figure 29e**). The PRC2 complex has been shown to interact with R-loops *in vivo* and *in vitro* (Alecki et al., 2020; Skourti-Stathaki et al., 2019), but how their presence influences PRC2 activity remains controversial.
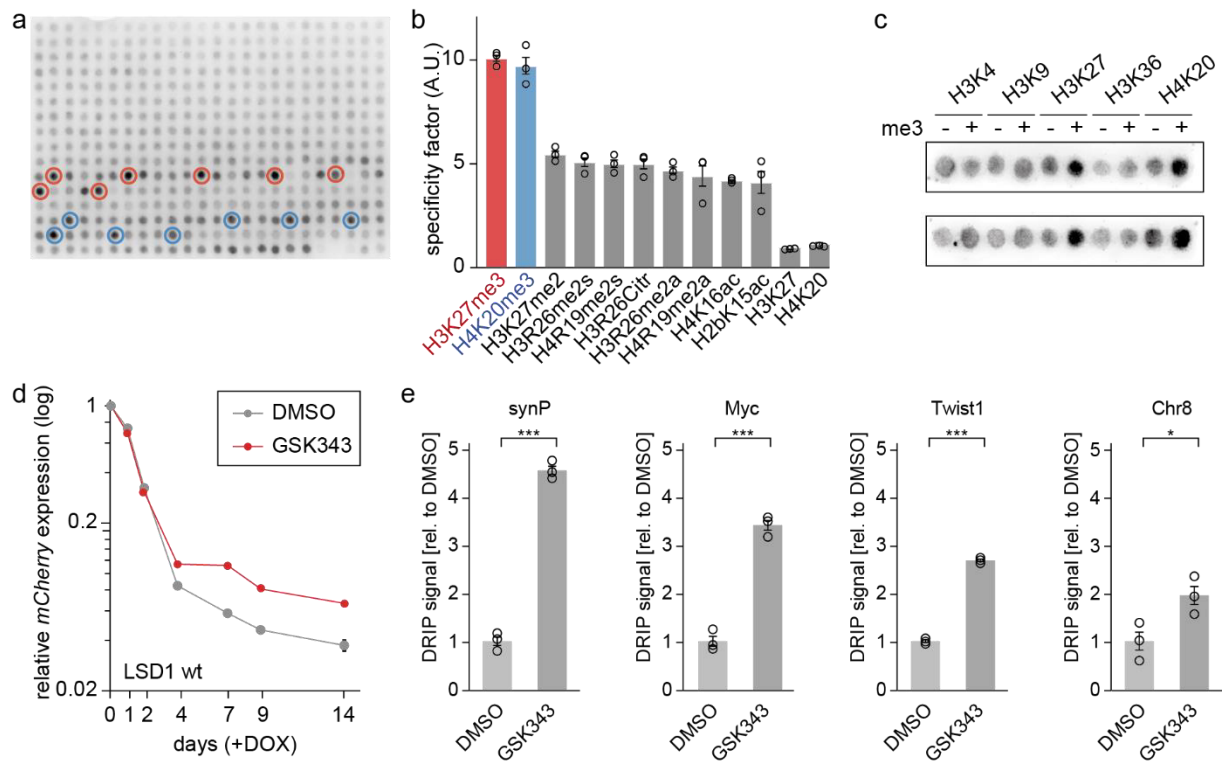
Figure 29: DDX19A is a reader of H3K27 trimethylation. a, Representative MODified™ histone peptide array incubated with recombinant GST-DDX19A. Binding of DDX19A was detected using an HRP-coupled secondary antibody for GST. Modified peptides containing H3K27me3 are highlighted in red, peptides containing H4K20me3 are highlighted in blue. **b,** Quantification of DDX19A binding to modified histone peptides. Signals from three independent replicates of the experiment shown in (**a**) were quantified using the Array Analyze Software (Active Motif). The specificity factor describes the average signal of all peptides carrying the indicated mark compared to the background signal. Circles indicate independent replicates (n=3, mean±s.e.m.). **c,** Two representative peptide SPOT arrays featuring the indicated unmodified and trimethylated histone peptides. Arrays were incubated with 50 nM DDX19A-GST. **d,** Dynamics of fluorescent reporter silencing upon recruitment of rTetR-LSD1 wt. NIH/3T3 reporter cells were treated with DOX and either DMSO or the EZH2 inhibitor GSK343 and analysed every 2-3 days using flow cytometry. Circles indicate the median *mCherry* signal relative to day 0 (n=3, mean±s.e.m.). **e,** DRIP-qPCR analysis of R-loop formation at the *synP* and three endogenous loci. Total nucleic acids for IP were isolated from NIH/3T3 cells treated with DMSO or GSK343 for 3 days. Signals are normalized to DMSO. (n=3; mean±s.e.m.; p*≤0.05, p***≤0.001, Student´s t-test). The figure was adopted from Pinter et al., 2021.

In conclusion, the fluorescent reporter system developed in this study was successfully combined with a comprehensive chromatin-focused RNAi screen, leading to the expansion of the chromatin regulatory network surrounding LSD1 (**Figure 30**). The fluorescent reporter was further used for the mechanistic characterization of the regulatory links connecting the activities of LSD1, PRC2 and DDX19A. The functional interaction of changes in chromatin that were known to mediate gene repression downstream of LSD1 recruitment, namely H3K4 demethylation and H3K27 methylation, could be validated using this system (**Figure 30**). Most importantly, the discovery of DDX19A as a coregulator of LSD1 led to the integration of R-loop formation as a novel regulatory factor in the network surrounding LSD1. Regarding the growing interest in the determination of the effect of R-loops on transcriptional regulation and chromatin, this observation holds promise for further interesting discoveries (Niehrs & Luke, 2020; C. Xu et al., 2021).
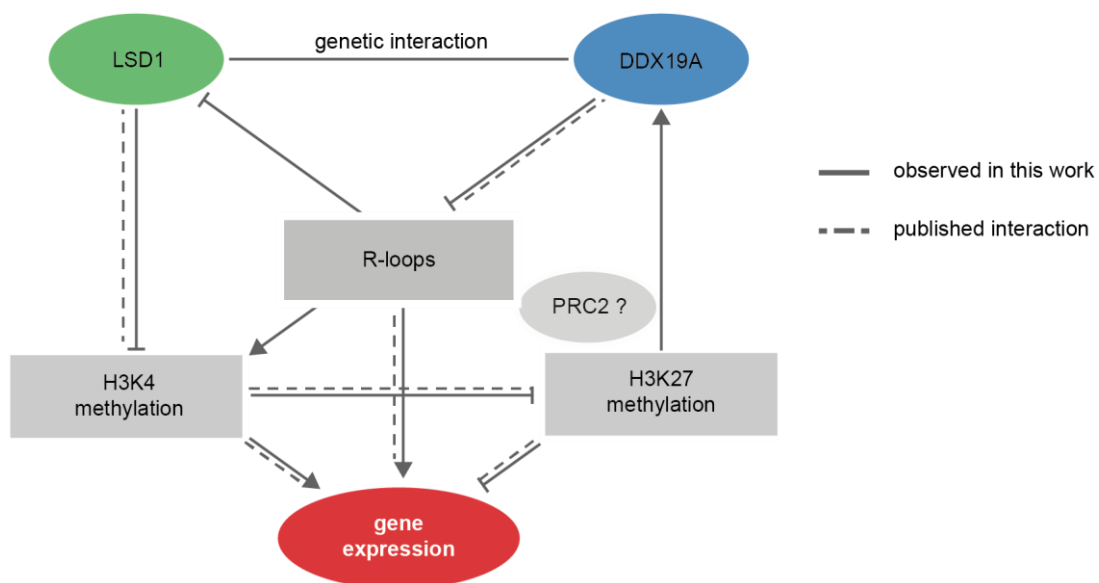
Figure 30: Schematic overview of the functional interactions connecting LSD1 and DDX19A with transcriptional regulation. Continuous lines represent connections that were observed in this work, dotted lines represent interactions that are known from literature. T-shaped lines indicate an inhibitory effect. The functional integration of the PRC2 complex into this network remains to be characterized. The figure was adopted from Pinter et al., 2021.

4.3.4. Analysis of the impact of the genomic integration site on the repressive activity of epigenetic effector proteins

In early experiments with different epigenetic repressor proteins, a pronounced variance in the responsiveness of the reporter populations towards the targeted silencing was observed (section 4.1.3). The experiments were conducted using heterogenous cell populations, with the synthetic reporter gene randomly integrated into the genome. Thus, the variable response was attributed to the impact of local chromatin factors influencing the efficiency of the epigenetic editor. The introduction of a transgene into the genome of an eukaryotic organism or cell line implies that the expression of the transgene is put under the regulatory control of the local chromatin context (Clark et al., 1996; Dobie et al., 1996; Ramírez et al., 2001; Tchasovnikarova et al., 2015). To segregate the expression of a reporter gene from the influence of *cis* or *trans* regulatory sequences like promoters or enhancers inherent to the host organism, reporter genes can be separated by insulators or expressed from plasmids (K. H. Hansen et al., 2008; N. C. Lee et al., 2013; Markstein et al., 2008). However, the integration of the reporter gene into an endogenous chromatin environment was a key premise to enable the investigation of coregulation in the epigenetic network in this thesis.

Aiming to determine the extent to which the genomic integration site affected the silencing of the reporter gene by recruitment of rTetR-LSD1, monoclonal reporter populations were created using fluorescence activated cell sorting (FACS). After transduction and selection of the *synP-mCherry* reporter and the rTetR-LSD1 fusion construct into NIH/3T3 cells, single-cell sorting was performed, and the resulting monoclonal populations were analysed using flow cytometry (**Figure 31a**). The initial medians of the *mCherry* signals varied in a range of ~0.5 log_{10}-scales between the individual clones, meaning that the strongest baseline expression was about five times higher than the weakest. The *mCherry* signals stretched over the range of almost one log_{10}-scale within each monoclonal population. This residual scatter is expected due to variations in gene expression during the cell cycle and to stochastic transcriptional bursting (Nicolas et al., 2017), while variations in between the population medians can be attributed to effects of the reporter integration site.

127

Upon recruitment of rTetR-LSD1 to *synP-mCherry*, different silencing dynamics were observed for each clone (**Figure 31b**). Interestingly, clones with higher initial expression levels appeared to be relatively resistant to LSD1-induced silencing, while clones with lower baseline expression could be strongly repressed (**Figure 31a, b**). The only exception to this trend was clone #3, which showed much stronger silencing dynamics that clone #24, despite comparable expression levels.
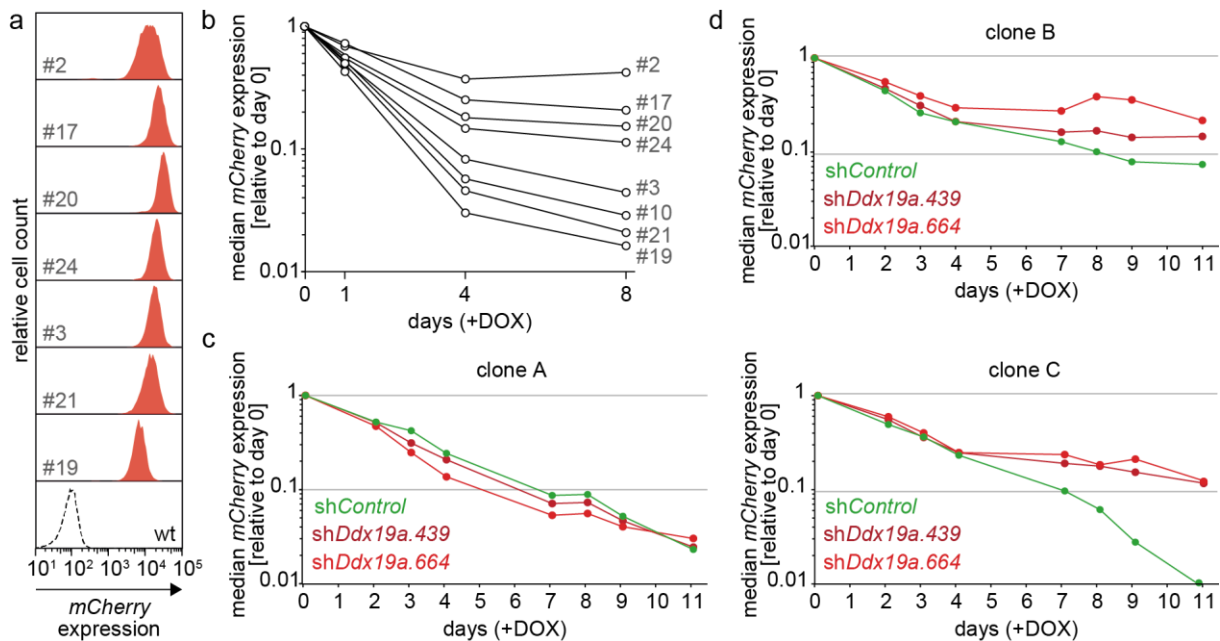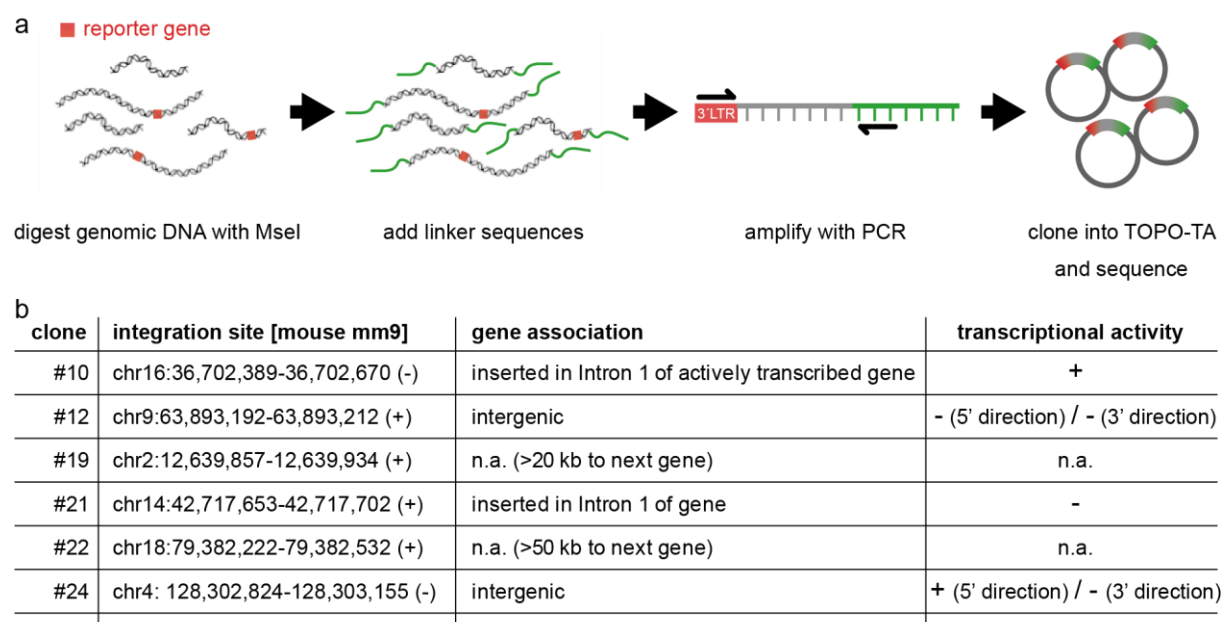

Figure 31: Influence of the genomic integration site of the reporter gene on transcriptional silencing by rTetR-LSD1. a, Expression profiles of representative monoclonal NIH/3T3 *synP-mCherry* reporter cell populations. Numbers indicate individual clones. wt = wild type NIH/3T3 cells without reporter. **b,** Silencing dynamics of the individual clonal populations shown in (**a**) upon recruitment of rTetR-LSD1. Reporter cells were treated with DOX and analysed using flow cytometry at the indicated timepoints. **c,d,** Silencing dynamics of representative clonal reporter populations expressing rTetR-LSD1 and shRNAs suppressing *Ddx19a*. The median *mCherry* expression upon recruitment of rTetR-LSD1 in reporter cells expressing the indicated shRNAs is shown.

Having observed that the extent of the repressive activity of LSD1 was dependent on the genomic locus, the dependence of LSD1-mediated reporter silencing on the DEAD-box RNA helicase DDX19A was analysed in different clonal populations. Fresh monoclonal populations expressing *synP-mCherry*, rTetR-LSD1 and either a control shRNA or one of two shRNAs targeting *Ddx19a* were created. Interestingly, some of those clones did not show an effect of DDX19A suppression (**Figure 31c**), while suppression of DDX19A strongly interfered with LSD1 silencing in other clones (**Figure**

31d). This supported the hypothesis that the mechanism of LSD1 mediated transcriptional silencing can vary depending on the existing chromatin environment.

To be able to characterize the genomic environment into which the reporter gene integrated in each clone, a published method for integration site mapping (Ciuffi & Barr, 2011) was adopted (**Figure 32a**). To this end, genomic DNA was extracted from each clone, fragmented using the restriction enzyme MseI and ligated a linker sequence to the newly formed ends. The ligated fragments were amplified in two rounds of PCR with primers binding in the 3'LTR of the viral construct and the linker sequence. Using this method, the reporter integration site of six clones from the experiment shown in **Figure 31a** could be determined (**Figure 32b**).

clone	integration site [mouse mm9]	gene association	transcriptional activity
#10	chr16:36,702,389-36,702,670 (-)	inserted in Intron 1 of actively transcribed gene	+
#12	chr9:63,893,192-63,893,212 (+)	intergenic	- (5' direction) / - (3' direction)
#19	chr2:12,639,857-12,639,934 (+)	n.a. (>20 kb to next gene)	n.a.
#21	chr14:42,717,653-42,717,702 (+)	inserted in Intron 1 of gene	-
#22	chr18:79,382,222-79,382,532 (+)	n.a. (>50 kb to next gene)	n.a.
#24	chr4: 128,302,824-128,303,155 (-)	intergenic	+ (5' direction) / - (3' direction)

Figure 32: Mapping of the genomic integration sites of six reporter cell clones reveals integration-site sensitive reaction to repression by LSD1. a, Schematic describing the workflow for integration site mapping. Genomic DNA is isolated from NIH/3T3 reporter cells and fragmented with MseI. Linker oligonucleotides are ligated to the MseI-ends and the fragments are amplified using primers binding in the reporter 3'LTR and the linker sequence. Amplified fragments of genomic DNA are cloned into TOPO®-TA vectors, sequenced and mapped to the mouse genome. **b,** Annotation of the genomic integration sites according to the mouse mm9 genome annotation. Association of the integration site with a gene was defined for either direct integration into a gene sequence, or integration less than 10 kb away from the next gene. **+** = associated gene is transcribed in NIH/3T3 cells, **-** = gene is not transcribed in NIH/3T3 cells, n.a. = not applicable (integration site >10 kb away from next gene).

In the six clones that could be analysed, the reporter gene had integrated into a different genomic site for each clone. The lentiviral construct had integrated into intragenic as well as intergenic regions (**Figure 32b**). The expression levels of endogenous genes in less than 10 kb distance from the identified integration sites were

analysed using a publicly available RNAseq dataset for NIH/3T3 (GEO accession #GSM970853). Only in one clone (#10), the reporter sequence was integrated into an intron of a gene that is stably expressed in NIH/3T3 (**Figure 32b**). Interestingly, in this clone the reporter gene was highly responsive to silencing with LSD1 (**Figure 31b**). The silencing dynamics were even more pronounced in clones with intergenic integration (#19) or integration into the intron of a transcriptionally silent gene (#21). In contrast, a clone in which the reporter gene was integrated downstream of an actively transcribed gene (#24) appeared to be relatively resistant to complete silencing by rTetR-LSD1 (**Figure 31b**). Unfortunately, the clone that was strongly resistant to LSD1 silencing could not be mapped using the described approach.

In conclusion, the experiments with monoclonal reporter cell populations confirmed that gene expression from the synthetic reporter was dependent on the endogenous chromatin environment surrounding its integration site. More importantly, the endogenous mechanisms preventing LSD1-induced silencing appeared to be different depending on the genomic locus.

4.4.   Analysis of the interaction between HUSH complex and KRAB/KAP-1

Besides the published project that dealt with the identification of coregulators of LSD1, the RNAi reporter system was used to investigate the coregulator network surrounding the transcriptional repressor ZNF10. KRAB-ZFPs play an essential role in controlling transposable elements, in the formation of extensive heterochromatic regions and in the protection of genome integrity (Ecco et al., 2017; Imbeault et al., 2017). The KRAB domain of ZNF10 is further commonly used as a repressor domain in synthetic biology and epigenome editing approaches (Alerasool et al., 2020; Tycko et al., 2020; Yeo et al., 2018). Recruitment of this domain to chromatin induces the formation of a heterochromatin-inducing complex comprising SETDB1, HP1 and NuRD around the corepressor protein KAP-1 (David C Schultz et al., 2002; P. Yang et al., 2017; Yeo et al., 2018). More recently, KAP-1 has been described to interact with the human silencing hub (HUSH) complex, an epigenetic repressor complex identified as mediator of heterochromatin spreading (Tchasovnikarova et al., 2015). The HUSH complex is composed of the H3K9me3 binding protein MPHOSPH8 and the less characterized proteins PPHLN1 and D14ABB1E (Douse et al., 2020), and interacts with a smaller complex of the H3K9-specific methyltransferase SETDB1 and ATF7IP to trimethylate

lysine 9 of histone 3 (Robbez-Masson et al., 2018; Tchasovnikarova et al., 2015; Timms, Tchasovnikarova, Antrobus, et al., 2016). Like the KRAB domain of ZNF10, which is frequently used for targeted repression of transcription (Alerasool et al., 2020), the HUSH complex has been shown to cooperate with the scaffolding protein KAP-1 (*also*: TRIM28) in the silencing of retrotransposons and evolutionary young genes (Robbez-Masson et al., 2018). Using the fluorescent reporter system in combination with shRNA mediated suppression of genes, the network connecting KRAB-ZFPs, SETDB1 and the HUSH among different cell lines was functionally characterized.

### 4.4.1. Characterization of the HUSH complex in NIH/3T3 cells

In a first approach, knockdown-validated shRNAs targeting the HUSH components *Mphosph8, D14abb1e, Atf7ip* and *Setdb1* (unpublished data from the Zuber lab, IMP, Vienna) were stably expressed in NIH/3T3 cells (**Figure 33**). To determine the relevance of the HUSH complex in NIH/3T3 cells, immunostaining for global H3K9 trimethylation (H3K9me3) followed by fluorescence microscopy was performed. This showed, that the dense heterochromatic foci typical for H3K9me3 localization in NIH/3T3 cells (Lungu et al., 2017) were strongly reduced upon suppression of any HUSH complex member (**Figure 33a,b**). The reduction of global levels of H3K9me3 under expression of shRNAs targeting the HUSH complex was further confirmed by Western Blot of NIH/3T3 whole cell lysate (**Figure 33c**). In accordance with published data for human cells (Tchasovnikarova et al., 2015; Timms et al., 2016), this suggested that the HUSH complex was essential for the establishment and/or maintenance of H3K9me3 levels in heterochromatic regions of NIH/3T3 cells, and that all tested HUSH members were needed for the complex to be operational. D14ABB1E has been described as the core subunit of HUSH that directs HUSH assembly at repetitive elements to ensure stable repression (Douse et al., 2020). The strong effects of *D14Abb1e* suppression on global H3K9me3 levels observed in immunofluorescence and western blot analysis confirmed this central role in NIH/3T3 (**Figure 33**).
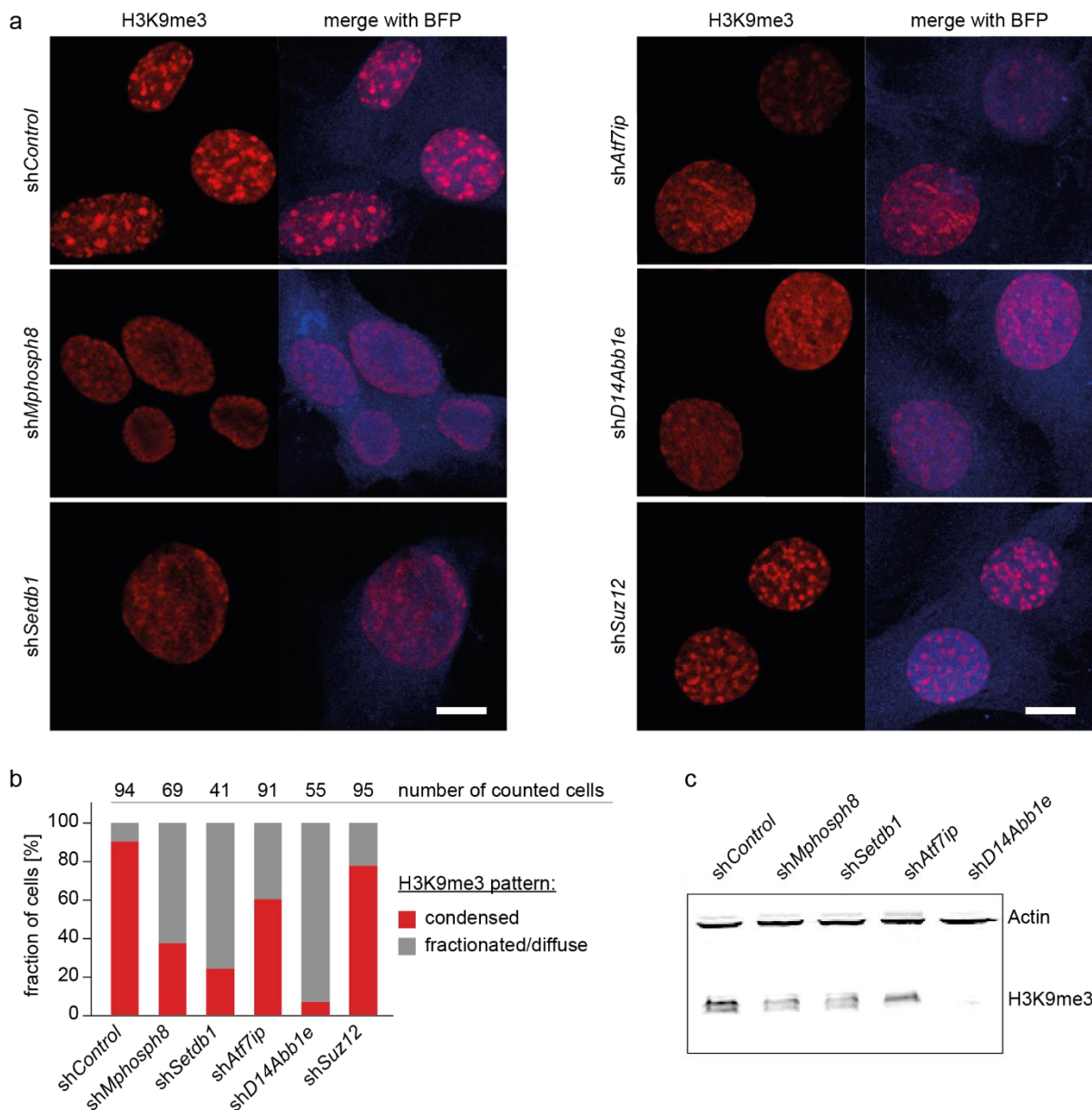
Figure 33: HUSH components are important for the regulation of H3K9me3 levels in heterochromatic foci in NIH/3T3. a, Representative immunofluorescence microscopy images of NIH/3T3 cells expressing the respective shRNAs for 14 days. sh*Suz12* was used as an additional control. Images are single cuts through the nuclei of fixed cells. red: antibody staining for H3K9me3, blue: BFP expression from the shRNA vector. Scale bar is 10 µM. **b,** Quantification of cells showing a phenotype with condensed H3K9me3 foci in the staining shown in (**a**). Numbers above the bars indicate the total number of cells counted for each shRNA. **c,** Immunodetection blot for H3K9me3 in whole-cell lysate from NIH/3T3 cells expressing the indicated shRNAs. Actin was used as a loading control.

SETDB1 and D14ABB1E have been shown to be recruited to euchromatic regions by KAP-1, inducing the local formation of heterochromatin leading to stable transcriptional repression (Robbez-Masson et al., 2018; David C Schultz et al., 2002). KAP-1 is described as the main scaffolding protein assembling a repressive complex at sites of ZNF10 binding (Alerasool et al., 2020; Ecco et al., 2017), thus representing an

essential mediator of KRAB-ZFP induced epigenetic silencing. To test whether KRAB-induced silencing of the synthetic reporter system in NIH/3T3 was dependent on recruitment of the HUSH complex by KAP-1, and to determine which component of the HUSH complex was most important in this context, NIH/3T3 reporter cell lines were transduced with shRNAs targeting the expression of *Atf7ip, Setdb1* or *Mphosph8* before DOX-induced recruitment of the KRAB domain to the *synP* element (**Figure 34**). The repressive activity of the KRAB domain on reporter gene expression was impaired by suppression of each tested HUSH component (**Figure 34a,b**). This effect was cumulative when *Mphosph8* and *Setdb1* were targeted in parallel with RNAi (**Figure 35a, b**). These observations are in line with published data about the importance of individual HUSH components for KAP-1 related gene repression (Robbez-Masson et al., 2018), where the suppression of *Mphosph8* and *Setdb1* gene expression induced the strongest derepression of KRAB target genes (**Figure 34b** and **Figure 35b**).
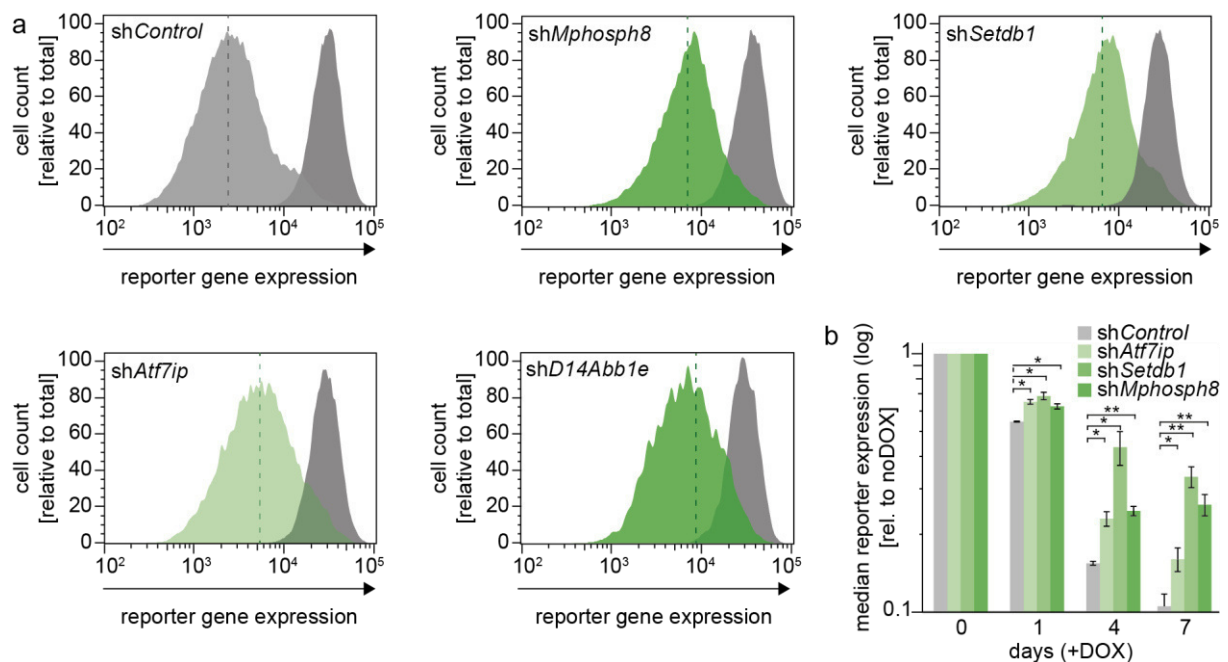
Figure 34: Expression of HUSH components is necessary for reporter silencing by KRAB in NIH/3T3. a, Representative flow cytometry analysis of reporter gene expression upon rTetR-KRAB recruitment. NIH/3T3 reporter cells stably expressing the indicated shRNAs were analysed after 7 days of DOX treatment (light grey and green) and without DOX (dark grey). Dotted lines indicate the median of the +DOX populations. **b,** Comparison of reporter silencing dynamics upon rTetR-KRAB recruitment, with and without suppression of HUSH complex partners. Bar graphs indicate the median reporter expression relative to noDOX, from independent experiments as shown representatively in (**a**) (n=3, mean±SEM, p*≤0.05, p**≤0.01, Student´s *t*-test.). Data for *D14Abb1e* are not shown in this panel, as this experiment was only performed once.

On a global level, *Mphosph8* and *Setdb1* also showed strong effects in maintaining H3K9me3 (**Figure 33**), while *D14Abb1e* suppression had an even greater impact on this mark. Preliminary experiments at the synthetic promoter supported a role of *D14Abb1e* in KRAB-induced transcriptional silencing (**Figure 34**), but combinatorial experiments with other HUSH components remain to be performed. However, *D14Abb1e* has also been observed to be a minor transcriptional coregulator of KRAB/KAP-1 at reporter genes, compared to the influence of *Mphosph8* and *Setdb1* on gene repression, by other groups (Robbez-Masson et al., 2018).
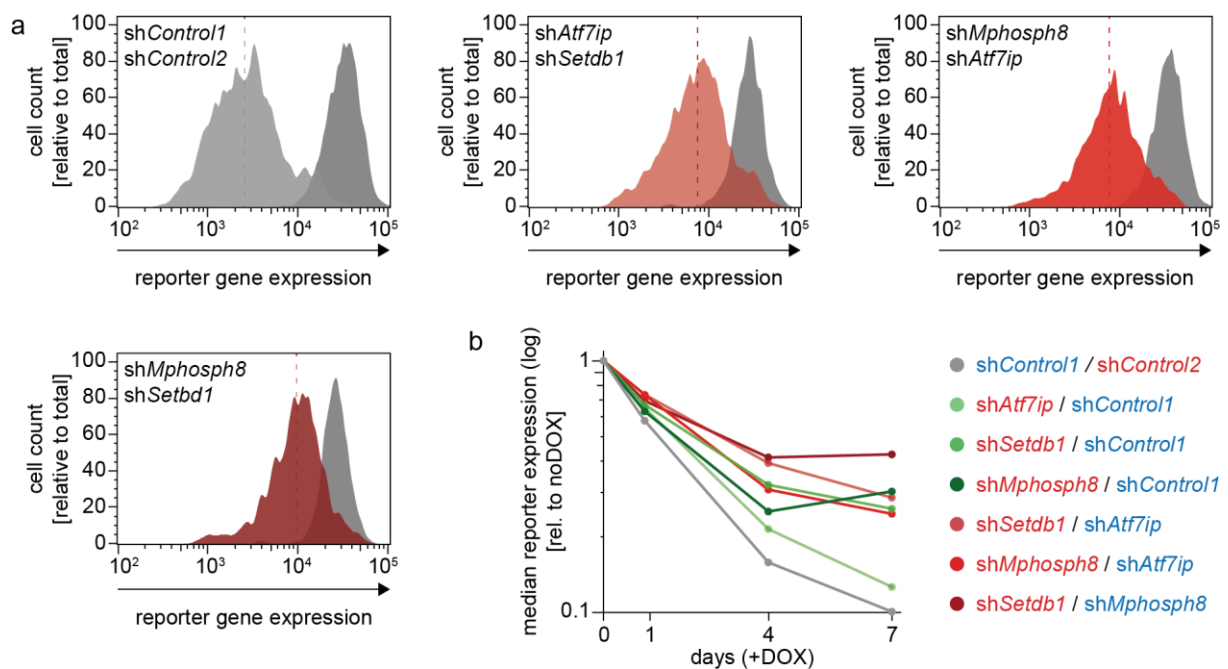
Figure 35: Suppression of multiple HUSH subunits in parallel increases the effect on KRAB induced silencing. Representative flow cytometry profiles of NIH/3T3 reporter cell lines co-expressing two HUSH shRNAs. Cells were analysed after 7 days of DOX-induced rTetR-KRAB recruitment (light grey and red) and without DOX (dark grey). Dotted lines indicate the median of the +DOX population. **b,** Comparison of reporter silencing dynamics upon rTetR-KRAB recruitment with combined suppression of multiple HUSH genes. The medians at day 7 (+DOX) were derived from the expression profiles shown in (**a**). Blue and red labelling indicates the type of fluorophore coupled to expression of the respective shRNA. Double-knockdown experiments were only performed once.

Suppression of *Atf7ip* showed the weakest effects on both the maintenance of global H3K9me3 levels (**Figure 33**) and the repression of reporter gene expression by KRAB (**Figure 34**). The functions of ATF7IP in the HUSH complex are to protect SETDB1 from proteasomal degradation (Timms, Tchasovnikarova, Antrobus, et al., 2016) and

possibly to enhance the interaction between MPHOSPH8 and SETDB1 (Tsusaka et al., 2018). While the loss of ATF7IP has been shown to phenocopy the loss of SETDB1 itself (Timms, Tchasovnikarova, Antrobus, et al., 2016), the observations presented in **Figure 33** and **Figure 34** appeared less distinct.

4.4.2. Analysis of the repressive complex around ZNF10 KRAB in iMEF cells

Previous publications have shown that the strength of the repressive activity of KRAB domains from various zinc finger proteins can vary independent of their ability to recruit KAP-1 (Murphy et al., 2016). It was hypothesized that this is dependent on the association with other corepressor proteins. Having established that the function of the transcriptional repressor domain KRAB is dependent on members of the HUSH complex in NIH/3T3 cells, this hypothesis was further tested in a different cell line. To this end, iMEF cells were transduced with the *synP-mCherry* reporter, the rTetR-KRAB fusion protein and the same shRNAs that were tested in NIH/3T3. The KRAB domain was recruited to the reporter by addition of DOX and fluorophore expression was monitored using flow cytometry (**Figure 36 a-f**). Interestingly, in iMEF no effect of the suppression of any of the HUSH complex partners tested in NIH/3T3 was observed (**Figure 36 a-f**). This suggested that in this cell line, the repressive protein complex forming around the KRAB domain might be different from the complex that is published in literature and was observed in the experiments with NIH/3T3.

To confirm this hypothesis, an RNAi-screen targeting 1104 genes with 5-6 shRNAs per gene was performed in iMEF rTetR-KRAB *synP-mCherry* reporter cells. The coregulator screen was performed as described for rTetR-LSD1 in NIH/3T3 (section 4.3.1 and **Appendix 1**), with a few exceptions. First, to account for the fast-silencing dynamics of the KRAB domain, the *mCherry+* and *mCherry-* populations were sorted after 7 days of DOX treatment. Second, only two independent replicates of the screen have been performed so far. Further replicates will be needed to reinforce the data with statistical power, however, the preliminary observations show an interesting trend (**Figure 37** and **Figure 39**).
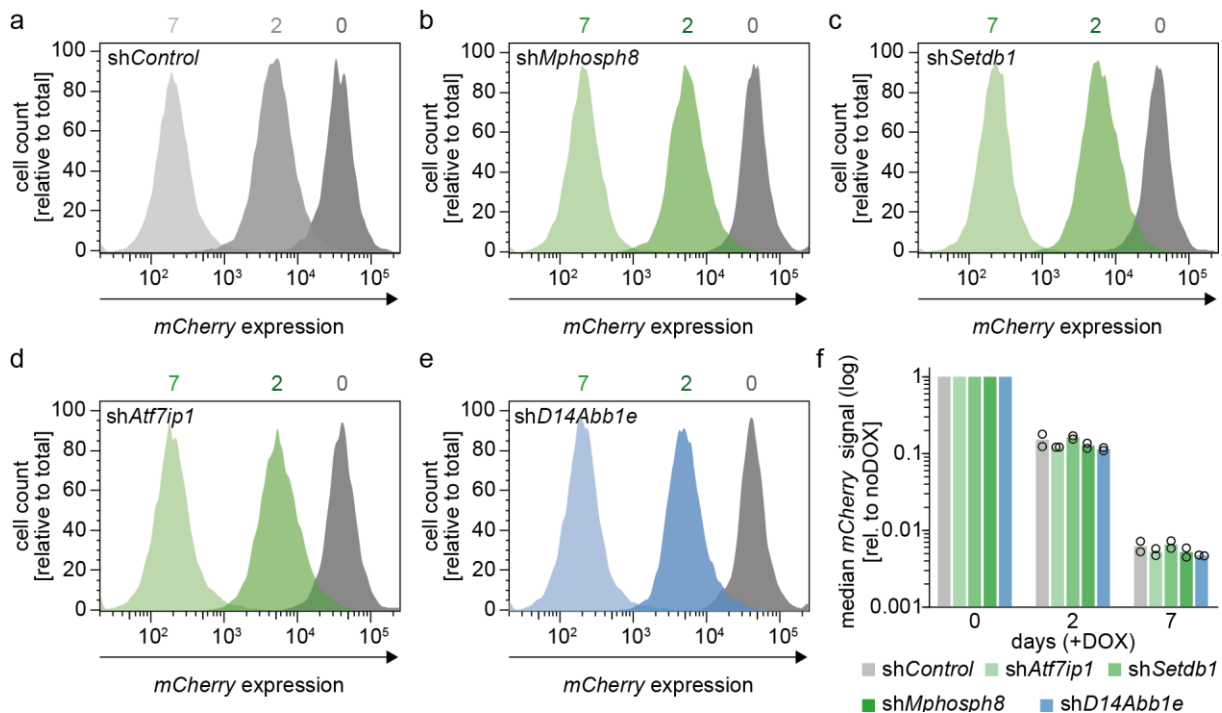
Figure 36: The HUSH complex is not essential for the induced silencing of the reporter expression by KRAB in iMEF cells. a-e, Representative flow cytometry analysis of iMEF reporter cell lines stably expressing the indicated shRNAs. Cells were monitored over 7 days of DOX-induced recruitment of rTetR-KRAB to the *synP* element. Numbers above the histograms indicate the duration of DOX treatment (days). **f,** Bar graphs illustrating the reporter silencing progress following rTetR-KRAB recruitment. Data correspond to the histograms shown in (**a**)-(**e**). The medians of the *mCherry* expression of the respective populations are shown relative to noDOX. Bar graphs represent the average of two biologically independent replicates. Circles indicate the values of the replicates, which were created in independent rounds of viral transduction of the shRNA vectors.

By comparing the obtained shRNA read counts in the sorted *mCherry+* and *mCherry-* populations among the two replicates, the reproducibility of the screen was analysed (**Figure 37a, b**). After Illumina sequencing, the raw read counts of each shRNA guide sequence were normalized to the respective total sample barcode read count, to allow for comparison of the representation between samples. The overall performance of the screen was validated by the analysis of control shRNAs that were spiked into the library pool DNA in equimolar plasmid amounts before virus production for the screen. Spike-in control shRNAs were represented reproducibly in the replicates of the sorted positive and negative population (**Figure 37a, b**). A spike-in shRNA targeting the production of the essential replication protein RPA3 (Fellmann et al., 2013) was completely depleted from all sorted cell populations 14 days after transduction (**Figure 37c, d**).
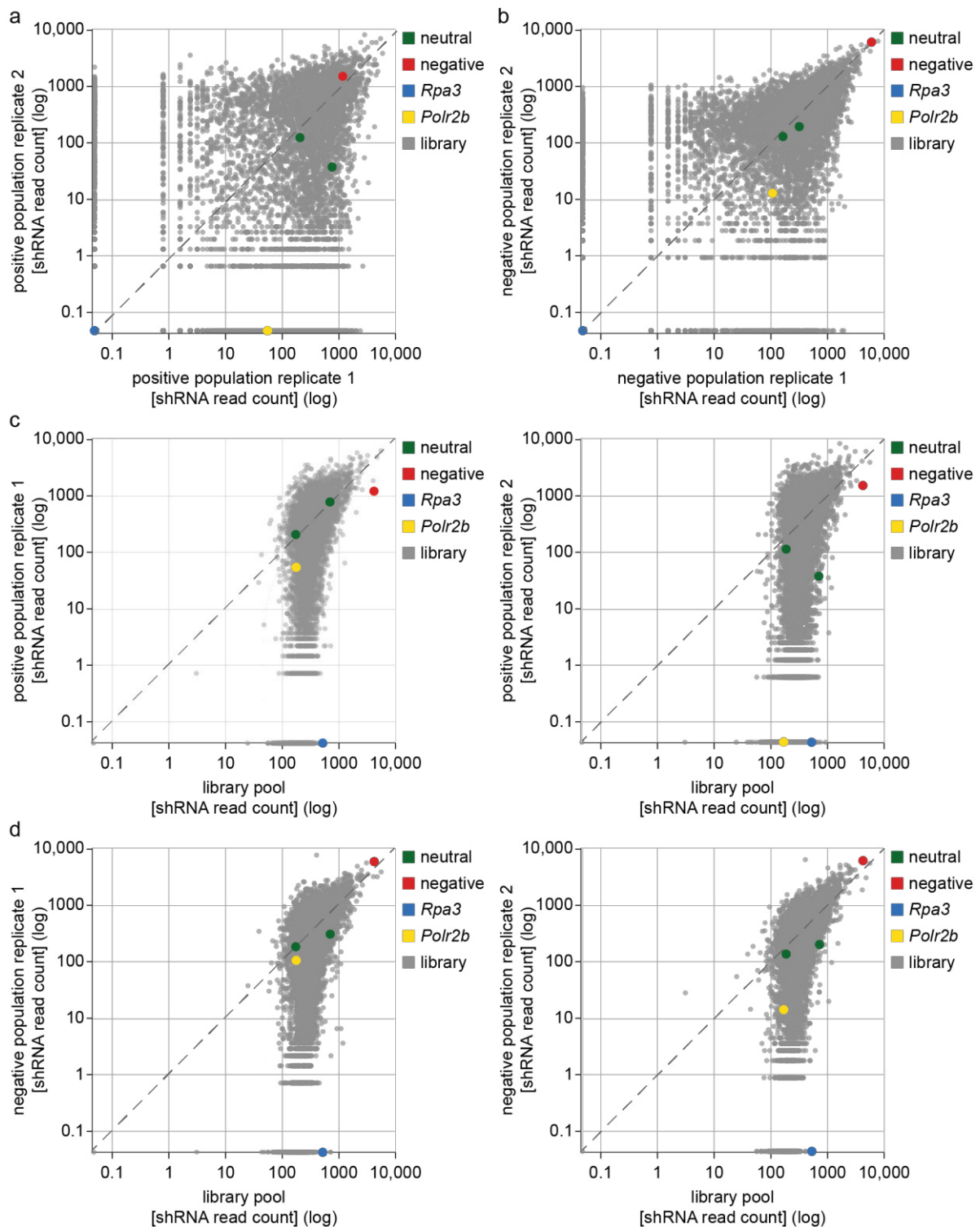
Figure 37: Comparison of shRNA representation in the sorted cell populations and the input DNA library. a,-d, Scatter plots showing the shRNA guide read counts obtained by Illumina sequencing. All 6614 shRNAs originally present in the library were analysed. Spike-in control shRNAs are highlighted in the indicated colours. The read counts were normalized to the total read count of the respective samples and are shown on a logarithmic axis for better visualization. **a,** shRNA read counts in the *mCherry+* sorted populations of both screen replicates. **b,** shRNA read counts in the *mCherry-* sorted populations of both screen replicates. **c,d,** Correlation of the read counts in the (**c**) *mCherry+* and (**d**) *mCherry-* populations with the representation in the input library. One outlier data point is not shown in (**a**)-(**d**) to allow for better scaling of the graphs.

In comparison, an internal control shRNA targeting the essential RNA polymerase II subunit B (sh*Polr2b*), which was already present in the initial library pool, consistently depleted from all populations, although the effect was less pronounced than for the spike-in shRNA targeting RPA3 (**Figure 37c, d**). Neutral control spike-in shRNAs that were validated as non-interfering with the reporter system in the lab did not show a trend towards enrichment in the *mCherry+* or *mCherry-* population (**Figure 37c, d**). A negative control shRNA that directly targets the reporter gene mRNA over a recognition sequence (Rathert et al., 2015) depleted from both positive populations (**Figure 37c**) and slightly enriched in the negative populations (**Figure 37d**).

As the RNAi screen was performed to investigate the influence of the HUSH complex members on KRAB-induced transcriptional silencing, the representation of all HUSH shRNAs in the library pool and in the sorted cell populations was analysed. shRNAs with an inhibitory effect on cell proliferation deplete over the time of cultivation and DOX treatment, resulting in a reduced read count in all sorted populations. To exclude an integration of these secondary effect in the analysis of the HUSH representation in the sorted populations, depletion scores were calculated for all HUSH-targeting shRNAs in each of the four *mCherry+*/*mCherry-* populations (**Figure 38**). The average of those depletion scores was taken to visualize a general anti-proliferative effect of the shRNA. Only the shRNAs targeting *Mphosph8* and *Pphln1* appeared to have an anti-proliferative effect. However, this effect was not very pronounced compared to the observed depletion in the neutral control shRNA and in the anti-proliferative shRNA *Rpa3.455* (**Figure 38**). For most shRNAs, the depletion could be attributed to the strong depletion of this hairpin in a single sorted population.
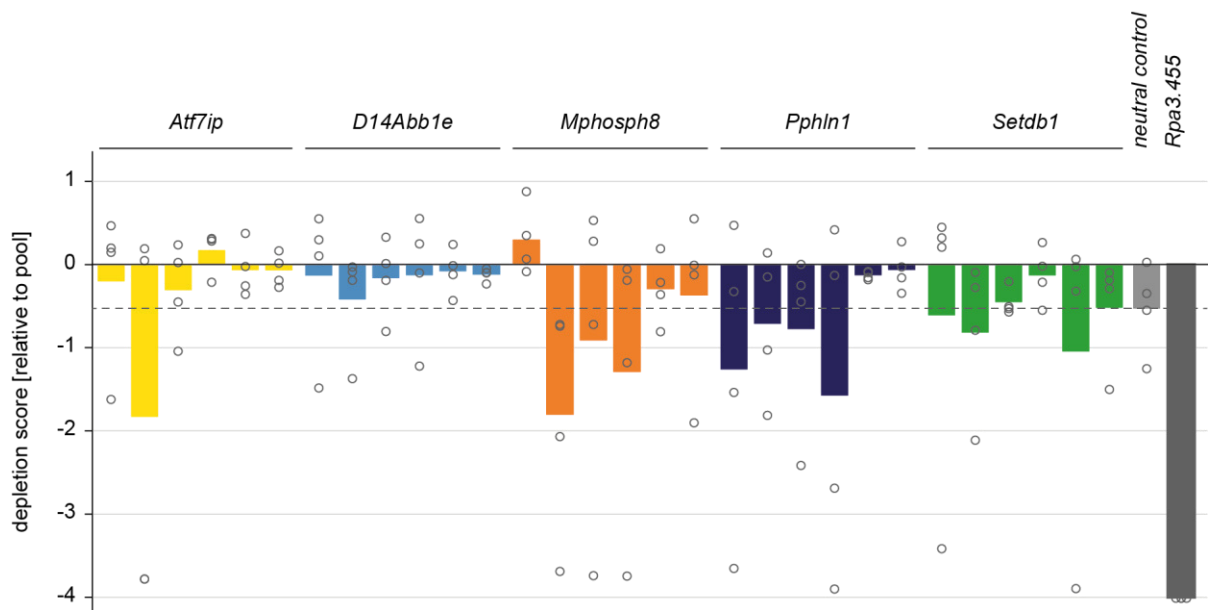
Figure 38: Analysis of the depletion of HUSH shRNAs during the cultivation of cells before sorting. Each bar represents one shRNA present in the screen. Individual depletion scores were calculated as the logarithmic ratio of the normalized read count in each sorted population over the read count in the library pool (empty circles). Bars are the average of the individual depletion scores calculated for each sorted population (two replicates of *mCherry+* and *mCherry-* each). Colours indicate the six individual shRNAs targeting the same gene. Depletion scores of the neutral control shRNA *Ren.660* and *Rpa3.455* are shown for comparison.

After analysis of the raw data, the gene scores, which describe the effect of enrichment of all shRNAs targeting the same gene, were calculated as described for the LSD1 screen in section 4.3.1 (**Figure 39a**). This consolidation was performed to summarize the individual effect of the six shRNAs targeting the same gene (**Figure 39b**). Like this, the influence of the expression of each gene in the library on KRAB-induced silencing could be evaluated. Genes sustaining a positive gene score (in total: 146 genes) were identified as positive coregulators of KRAB mediated silencing (**Figure 39a**). The results from the RNAi screen confirm the observations from the single-shRNA experiments (**Figure 36**) and support the hypothesis that the repressive function of KRAB is independent of HUSH activity in iMEF cells. Of all members of the HUSH complex that were included in the shRNA library, only the transcription activation suppressor *D14abb1e* appeared as a positive regulator of KRAB, with 3 out of 6 shRNAs enriching in the positive population (**Figure 39a, b**). *D14Abb1e* plays a central role as a subunit of the HUSH complex (Douse et al., 2020; Gresakova et al., 2019). The high scoring of *D14Abb1e* in combination with the low scores of the other HUSH members suggests that this gene might have additional regulatory functions or the

ability to assemble another silencing complex in the context of transcriptional repression by KRAB. Analysis of published RNAseq data from NIH/3T3 (GEO accession #GSM970853) and iMEF (GEO accession #GSE90108) cells supported the hypothesis that the HUSH complex plays a less significant role in iMEF cells (**Figure 39d**) than it was observed for NIH/3T3 (**Figure 33**), because it was found to be expressed only weakly. The role of SETDB1 in the methylation of H3K9 downstream of KRAB recruitment could be adopted by other H3K9 specific methyltransferases like the high-scoring *Suv39h1* (García-Cao et al., 2004), which is expressed at comparable levels in both cell lines (**Figure 39c, d**). However, the reporter gene could also have been integrated into genomic loci that were under control of *Suv39h1* as a H3K9 methyltransferase.

Following the assumption that the essential function of the HUSH complex to stably repress repetitive regions in association with KRAB might be adopted by another regulatory complex in MEF cells, the top scoring hits in the KRAB screen were analysed (**Figure 39c**). It must be mentioned that the transcriptional regulator *Myc* was removed from the top-scoring gene view (**Figure 39d**). *Myc* is an essential gene that is embedded in a complex network of chromatin regulation (Seruggia et al., 2019). Regarding the anti-proliferative effects of *Myc* suppression with RNAi (Fellmann et al., 2013) and its role as a master regulator of gene expression (Poole & van Riggelen, 2017; Takahashi et al., 2007), it was suggested that the high scoring might be due to unknown secondary effects. However, it cannot be excluded that it also interferes with KRAB induced transcriptional silencing.

With *Mecp2* and *Sap30*, two members of the SIN3A/HDAC complex scored among the top 10 hits, and another two subunits of this complex (*Sap130* and *Sap25*) scored as positive regulators of KRAB (**Figure 39a, c**). In contrast to this, none of the NuRD/HDAC complex members were identified as positive regulators in this screen. Furthermore, the zinc-finger protein ZMYM4, which has been described to build a corepressor complex with LSD1 and RCOR1 (Cibis et al., 2020) scored at position five, and *Rcor1* was also identified as a positive co-regulator of KRAB (**Figure 39a**).

It remains to be mentioned that all hairpins targeting the scaffolding protein KAP-1, which has been described to recruit the HUSH complex to KRAB binding sites, were depleted in the cell populations that were analysed after 7 days of KRAB recruitment (**Figure 39e**). This indicated that the expression of KAP-1 was essential in iMEF cells,

and the effect precluded a qualitative analysis of the importance of KAP-1 for KRAB induced silencing in this system. However, as no toxicity was observed for the suppression of the HUSH complex members (**Figure 38**), it was assumed that an impairment of the functionality of the HUSH complex was not the source of the antiproliferative effect of KAP-1 depletion.
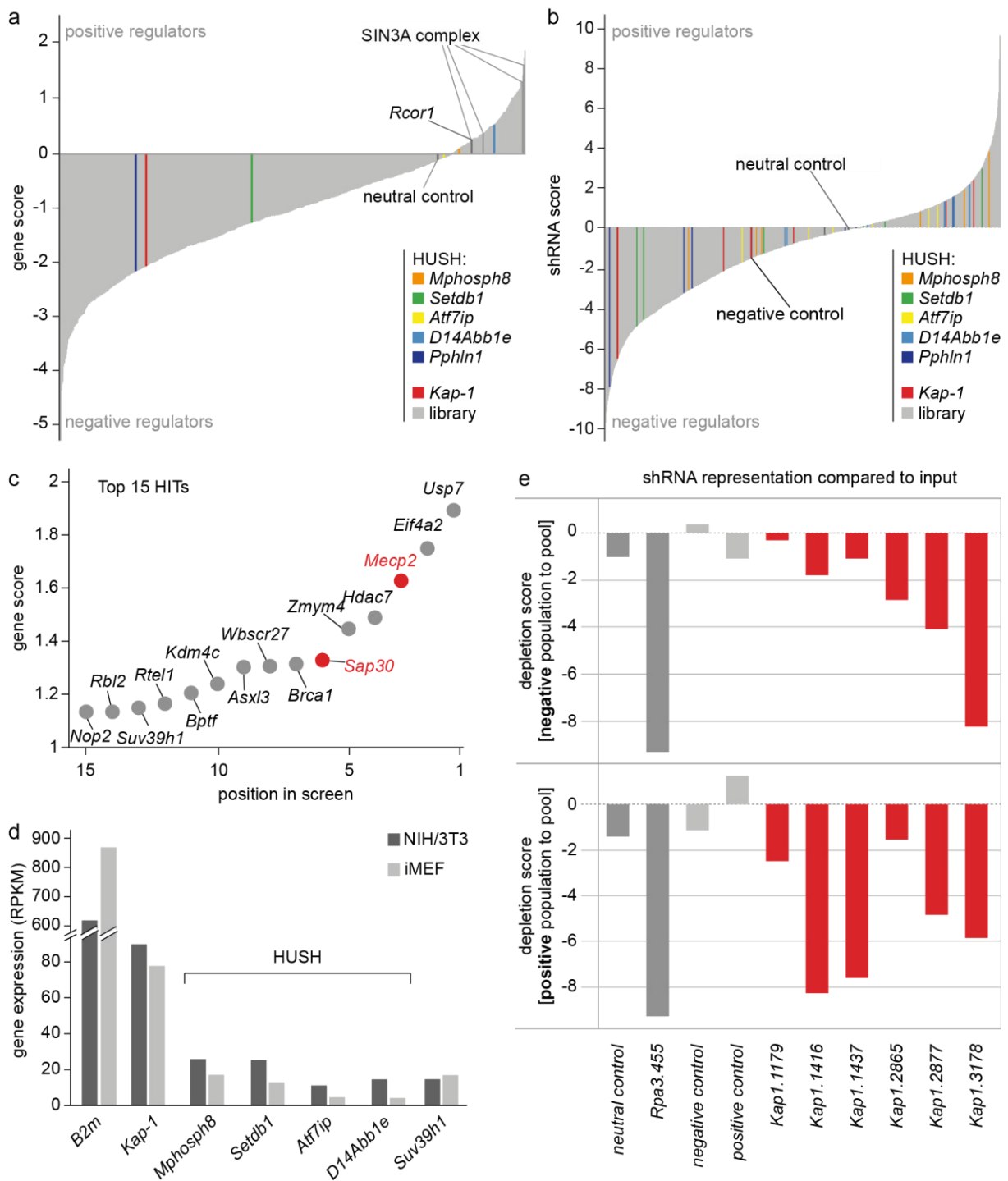
Figure 39: The HUSH complex is dispensable for KRAB mediated silencing in iMEF. a, Distribution of the gene scores for all genes included in the library. The gene score represents the average shRNA score of all shRNAs targeting the same gene. Gene scores >0 indicate positive regulators of KRAB, gene scores <0 indicate negative regulators. HUSH and *Kap-1* shRNAs are highlighted in the indicated colours. The position of the SIN3A complex members that are mentioned in the main text is indicated. **b,** Distribution of shRNA scores for all shRNAs included in the library. The shRNA score represents the logarithmic ratio of the read counts in the positive over the negative population, averaged over both replicates. **c,** Scatter plot showing the top 15 HITs of the RNAi screen. Gene scores are the same as in (**a**). Genes of the SIN3A complex are highlighted in red. **d,** Gene expression levels of the HUSH components in NIH/3T3 and iMEF cell lines. RPKM counts were taken from publicly available RNA-seq data. *B2m, Kap-1* and *Suv39h1* are shown for comparison. **e,** Depletion of *Kap-1* shRNAs in

the reporter cells during the screening process. Depletion scores represent the logarithm of the ratio of the normalized shRNA read counts in the sorted cell populations to the read count in the library pool. In contrast to Figure 37, the depletion is shown separately for the *mCherry+* and *mCherry-* populations. Read counts were averaged over the two replicates before calculating the ratio.

In conclusion, the fluorescent reporter system was used to validate the critical interaction of the KRAB domain with HUSH components for the induction of a repressive chromatin state in NIH/3T3 cells. Suppression of HUSH complex partners led to a global decrease in H3K9me3 levels in these cells (**Figure 33**). Furthermore, expression of the HUSH subunits *Atf7ip*, *Setdb1*, *Mphosph8* and *D14Abb1e* was shown to be essential for the efficient silencing of a reporter gene by the recruitment of KRAB (**Figure 34** and **Figure 35**). As this effect could not be observed in iMEF cells (**Figure 36**), a comprehensive RNAi screen was performed to probe the regulatory network surrounding KRAB as a transcriptional repressor. In combination with the single-RNAi data, the screen delivered preliminary, yet interesting hints on the variability of epigenetic complex formation around the same repressor protein (**Figure 39**). The data suggest that the composition of regulatory complexes can vary significantly between different cell lines, even among two cell lines of comparable origin and phenotype.

5. Discussion

Trying to seize the universal importance of epigenetic mechanisms for the development of multicellular organisms, it is helpful to visualize that every cell in an adult organism contains the complete DNA-sequence information that would be needed to form any other cell type of this organism. In vertebrates, the development of phenotypically and functionally specialized cell types from the fertilized egg cell is directed by the concerted action of a plurality of epigenetic effector proteins. The epigenetic installation of cell-type specific gene expression programs during embryonic development is tightly regulated in a stepwise and highly time-resolved manner (Soshnikova & Duboule, 2009). In differentiated cells, epigenetic control of gene expression enables the integration of external signals to adapt the phenotype of cells or whole organisms to environmental cues (Bar-Sadeh et al., 2020; Mazzio & Soliman, 2012). Finally, epigenetic mechanisms ensure the accurate inheritance of epigenetic patterns from the mother to the daughter cells during mitosis (Stewart-Morgan et al., 2020). Consequently, epigenetic regulation of gene expression is a process that must be maintained throughout the whole life cycle of multicellular organisms. The integration of epigenetic modifications and epigenetic effector proteins into a dense network of feedback-loops and verification mechanisms has proven beneficial during evolution, preventing the development of diseases that arise from misdirected epigenetic control (Bates, 2020; Greenberg & Bourc'his, 2019; Landgrave-Gómez et al., 2015; Shilatifard, 2012). The aim of this work was to improve the understanding of epigenetic mechanisms as parts of a complex network, and to provide a new tool for the investigation of singular epigenetic effector proteins as integral components of this network.

5.1. Development of a fluorescent reporter system for the dynamic analysis of epigenetic effectors in live cells

5.1.1. The fluorescent reporter system as a valuable expansion to established methods

Targetable reporter genes have been used as a biotechnological tool for the analysis of gene expression levels for almost 30 years now (Gossen et al., 1995; Ho et al., 1996; Vissing et al., 1995). Early in the history of transgene experiments, scientists observed that if the gene of interest was stably integrated into the host genome, the genomic site of integration had an influence on the expression levels of the transgene (Dobie et al., 1996). This effect had first been described in experiments with *Drosophila melanogaster* fruit flies, where the term *position effect variegation* (PEV) was introduced to describe the phenomenon that an active gene could become inactivated by the spreading of adjacent heterochromatic regions (Reuter & Spierer, 1992). Today, it is understood that transgenes cannot only be repressed by the H3K9me3-dependent spreading of heterochromatin, but are susceptible to a variety of epigenetic regulatory mechanisms that apply to the locus of integration. This embedment of reporter transgenes into the host epigenetic network can be used to study epigenetic mechanisms regulating chromatin modifications and transcription (Bintu et al., 2016; Hathaway et al., 2012; Headley et al., 2019; Keung et al., 2014; Tycko et al., 2020).

The idea behind this work was to create a synthetic reporter gene that would deliver constant and easily accessible readout of transcription from a targetable promoter. Furthermore, this system should neither be limited to a single integration site or cell line, nor should it be outsourced to an artificial chromosome, increasing the chance to comprehensively monitor all endogenous mechanisms of epigenetic regulation. It was shown that the system responded dynamically to induced recruitment of various epigenetic repressor proteins fused to a rTetR protein. Silencing of the reporter gene expression was not only dependent on the recruitment of an epigenetic effector protein (**Figure 17**), but also on the co-recruitment of endogenous corepressor proteins, as was shown by drug treatments and RNAi experiments (**Figure 21** and **Figure 22**). This was an essential observation for the progress of this project, as similar reporter gene approaches have already been used to screen for epigenetic barriers preventing transcriptional activation (Headley et al., 2019) or for novel regulatory protein domains (Tycko et al., 2020). To expand to those approaches, the goal of this work was to

screen for novel interactors of already known epigenetic effectors using a comprehensive RNAi library, aiming to expand the knowledge about regulatory multiprotein complexes.

Outline of interesting observations during the establishment of the system

An early aim during the establishment of the reporter gene approach was to evaluate the sensitivity of the reporter system to the regulation by different chromatin effectors, and to compare the significances of different chromatin modifications for the formation of a repressed chromatin state. Transcription from the strong EF1A promoter was rapidly and efficiently silenced by the recruitment of a KRAB domain and CBX3 (**Figure 17**). In contrast to the recruitment of those effectors, the recruitment of the SET domain of EHMT1 did not achieve efficient silencing of reporter expression (**Figure 17b**). This might suggest that the conversion from di- to trimethylation at H3K9 was inhibited by endogenous factors stabilizing the existing chromatin state at the reporter locus (Headley et al., 2019). Another explanation would be that the mere introduction of H3K9me3, although this is also part of the repressive cascade downstream of KRAB and CBX3, is not sufficient for the downregulation of gene expression. EHMT1 acts in a heterodimeric complex (Tachibana et al., 2005) and mediates the recruitment of additional corepressors *in vivo* (Chang et al., 2011; Fritsch et al., 2010; Shinkai & Tachibana, 2011). Even though the isolated SET domain is active on histone substrates in vitro (Tachibana et al., 2008), presumably many of these interactions were lost when the protein was reduced to its catalytic domain, leading to a decrease in its activity on chromatin.

The repressive activity of the PRC2 subunit EED at the reporter gene was also limited (**Figure 17b**). However, some cells in the heterogeneous reporter population were fully silenced by recruitment of EED. This suggests that in some cells, the reporter gene was integrated into PRC2-controlled loci which can efficiently be silenced by the formation of a *de novo* repressive polycomb domain (J.-R. Yu et al., 2019), while at other regions gene expression was only slightly reduced by the presence of the complex. At these regions, the existing active chromatin environment might inhibit the activity of PRC2 (Laugesen et al., 2019; Mocavini & Di Croce, 2020).

An interesting observation on the chromatin modification level was the limited ability of LDS1 to reduce H3K4me2 at the promoter element (**Figure 20**). This might be due to the fact that H3K4me2 is an intermediate modification between the fully active

trimethylated, and the repressed unmethylated state of H3K4. LSD1 itself is unable to demethylate trimethylated H3K4, but the presence of HDACs co-recruited by LSD1 recruits the Jumonji demethylase JARID1B (Barrett et al., 2007), which can demethylate H3K4me3, leading to a constant increase of H3K4me2. However, to exclude the effects caused by the heterogeneity of the reporter population and allow for better quantitative analysis of the ChIP signals, experiments with a monoclonal reporter population should be considered.

Another open question that could be addressed using the synthetic reporter system would be the resolution of the temporal dynamics connecting chromatin modulation and transcriptional repression. In the presented work, changes in chromatin modifications were analysed long after an initial repression of transcription was observed. Studies investigating the chromatin remodelling complex NuRD have shown that the decrease in H3K27ac, induced by co-recruitment of HDACs, follows long after the initial decline in transcriptional activity (Bornelöv et al., 2018), and thus rather functions to stabilize the repressed state than to induce it. Other groups targeted bivalent genes in mESCs with the BAF remodelling complex and observed, that the loss of H3K27me3 and the gain in H3K4me3 following complex recruitment precedes transcriptional activation (Braun et al., 2017). These observations support a model stating that certain histone modifications present a barrier that holds up the present chromatin state, which needs to be overcome to remodel the chromatin structure and allow for changes in transcriptional activity (Hathaway et al., 2012; Headley et al., 2019). However, this model cannot explain the transcriptional activity that is observed in heterochromatic regions in multiple organisms (Saksouk et al., 2015), and omits the effects of chromatin modifications on TF binding, transcriptional initiation, elongation and termination (Jonkers et al., 2014; B. Li et al., 2007; Stasevich et al., 2014; Tycko et al., 2017). Time-resolved analysis of transcriptional changes and the onset of chromatin marks using the presented reporter system could be used for determining the chronological order of events that control epigenetic repression.

The three-state model for transcriptional repression
Studies with similar reporter gene approaches also observed that the dynamics of the onset and stability of chromatin modification changes vary depending on the recruited chromatin modifier (Bintu et al., 2016). By investigating the temporal nature of different chromatin modifications, the reporter system could be used to study the mechanisms

of epigenetic spreading and memory. With their respective approaches, Bintu and Hathaway showed that artificially introduced chromatin modifications are lost following modification-specific kinetics once the initial trigger is removed (Bintu et al., 2016; Hathaway et al., 2012) This observation was confirmed for multiple epigenetic domains in this work (**Figure 18**). Persistent silencing of gene expression was only achieved by recruitment of a DNA methyltransferase domain, but repression by this domain could not be achieved in all cells (**Figure 17**). Bintu et al. propose a three-state model for the transcriptional repression following chromatin regulator recruitment. According to this model, the recruitment of a repressive regulator to a transcribed locus causes cells to stochastically advance from an actively expressing state to a reversibly silent state and then to an irreversibly silent state (Bintu et al., 2016). The stochastic transition to the respective states is dependent on the duration and strength of the recruitment. Furthermore, transcriptional repression is described as an all-or-nothing process. Silencing of the reporter gene can either be achieved completely or not at all. This model explains the formation of a bimodal population that was observed by recruitment of EHMT1, EED and DNMT3A (**Figure 17**). The bimodal form indicates that depending on the locus of reporter integration, the cells either respond to the chromatin effector recruitment or not. Observed intermediate *mCherry* signals can be explained by transcriptional bursting, cell cycle dependent changes of expression, and the locus-dependent baseline expression levels in the heterogenous reporter cell populations that were used in this experiment (Nicolas et al., 2017; Tycko et al., 2017). Interestingly, recruitment of the KRAB domain to the synthetic reporter did not achieve a transition of the cells into an irreversibly silent state, although the time of recruitment was longer than described in Bintu et al. (**Figure 18**). The repressed chromatin state established by KRAB recruitment was independent of DNA methylation in the CHO system published by Bintu et al. (Bintu et al., 2016). However, the stable repression of a gene after recruitment of KRAB has been shown to require the deposition of DNA methylation in other cells (Ying et al., 2015). For both assumptions, it can be suggested that the recruitment of KRAB to the synthetic reporter failed to efficiently activate the endogenous repressive network of the NIH/3T3 cells at this locus. To install a stable and heritable heterochromatic domain, the activity of H3K9 methyltransferases, histone deacetylases and heterochromatin binding proteins is needed (Canzio et al., 2013; Hathaway et al., 2012). ChIP-qPCR and bisulfite sequencing (Darst et al., 2010) experiments at the synthetic promoter could reveal at which stage the establishment

of a heritable heterochromatin domain upon KRAB recruitment failed in this system, and whether DNA methylation was involved in this process.

The catalytic domains of DNMT3A and B have widely been used for the targeted introduction of DNA methylation at reporter genes or genomic loci of interest (Amabile et al., 2016; Bintu et al., 2016; Broche et al., 2021; Rots & Jeltsch, 2018). The introduction of DNA methylation is always correlated with a reduction in gene expression, but the stability of the signal is dependent on the nature of the targeted locus (Broche et al., 2021). The bimodal response of the reporter cells to the recruitment of DNMT3A-CD observed in this work (**Figure 17**), together with the long-term stability of the repression after removal of DNMT3A (**Figure 18**), suggest that the reporter gene is not effectively methylated in the irresponsive cells. The isolated DNMT3A-CD has been shown to be very efficient in depositing DNA methylation (Broche et al., 2021), is independent of accessory proteins and is not inhibited by active chromatin marks (Yingying Zhang et al., 2010). Recruitment of the DNMT3A-CD by TetR strongly improved the efficiency of gene repression, suggesting that the binding strength of the targeting module was the limiting factor for efficient DNA methylation (Kamionka et al., 2004). Furthermore, it is possible that TET enzymes are actively recruited to the strong CpG promoter of the reporter gene (Rasmussen & Helin, 2016).

The stability of an induced change in chromatin conformation strongly depends on the primary state of the genetic locus, with the system tending to maintain or return to its original chromatin state (Hathaway et al., 2012). With these observations, two pivotal questions remain to be answered: What are the barriers presented by the epigenetic regulatory network that need to be overcome to achieve stable silencing of a primarily active locus? And which mechanisms in this network need to be triggered to shape epigenetic memory? (Bintu et al., 2016; Broche et al., 2021; Headley et al., 2019).

One way to address these questions could be to repurpose the screening approach presented in this work to a negative-coregulator RNAi screen. This could either be done by screening for early-silencing reporter cells, or for cells that remain silenced after withdrawal of the effector domain from the synthetic reporter gene.

5.1.2. Possible optimizations of the fluorescent reporter system

In recent years, the application of CRISPR-Cas, CRISPRi and CRISPRa technologies have replaced RNAi approaches as the primary tool for high-throughput screening methods in cells (Pickar-Oliver & Gersbach, 2019). However, CRISPR-Cas knockout methods suffer from an increase in cell-death rates caused by the massive introduction of DNA double-strand breaks (Pickar-Oliver & Gersbach, 2019), and the complete knockout of target genes increases the risk of false-negative results in interaction screens due to anti-proliferative effects. Both risks could be reduced by the application of CRISPRi instead of active CRISPR-Cas9, but these methods were not well-established at the time of the beginning of this thesis. Considering these drawbacks, CRISPR knockout screens qualify for the use in drop-out screens investigating cell proliferation, differentiation, or drug-sensitivity, but do not provide a big advantage over RNAi screens in a reporter-based gene regulatory screen. The drawbacks of CRISPR screening approaches have been addressed by the optimization of gRNA libraries and Cas9-variants (Ong et al., 2017; Pickar-Oliver & Gersbach, 2019; Sanson et al., 2018) in the past years. Consequently, the transition from combining the synthetic reporter with an RNAi library screen to combination with a CRISPR-based screen could be considered for future experiments.

Regarding the synthetic reporter cassette itself, the system could be optimized by exchanging the TetR-derived targeting module fused to the chromatin regulator. Using this DOX dependent recruitment system, the options for combination of the reporter with a conditional system driving the expression of the RNAi library are limited. However, the possibility to precisely time the beginning of shRNA expression would reduce the number of shRNAs dropping out and improve the reproducibility and comparability among all shRNAs in the library. Synthetic biology has brought forth a multitude of methods to achieve the conditional regulation of gene expression (Kirkpatrick et al., 2020; Lienert et al., 2014; Smanski et al., 2016). By fusing the chromatin modifier and the DNA-binding anchor domain to different parts of either a light-sensitive or a rapamycin-sensitive system, recruitment to the synthetic reporter gene can be achieved by induced dimerization of the chimeric proteins (Kallunki et al., 2019; Lienert et al., 2014). However, these methods are critical in their establishment due to their increased complexity, high costs for rapamycin in large culture volumes and the risk of background expression of light-inducible systems. Alternatively, the

TetR-dependent expression system could be combined with a cumate-controlled operator system (Kallunki et al., 2019), thereby uncoupling the triggers controlling expression of the chromatin regulator and the shRNA library.

Another challenge that could not be completed during this thesis was the establishment of a reporter system for the investigation of transcriptionally activating epigenetic effectors. Only a slight increase in transcriptional activity could be observed in experiments with recruitment of BRD4 to different promoter sequences (**Figure 19**). To boost the measurable effect of transcriptional activation, a two part reporter system with a TEV protease under the control of the targetable minimal promoter, and a TEV-inducible flipGFP protein under control of a constitutive promoter could be used (Qiang Zhang et al., 2019). During the establishment of a functional reporter, transcriptional activation could be achieved by recruitment of a histone acetyltransferase, as acetyltransferases are more direct activators of transcription than BRD4 and have successfully been used in CRISPRa experiments (Hilton et al., 2015)

Finally, during the characterization of different epigenetic regulators, variances in the response to recruitment of the same regulatory domain inside the heterogeneous reporter cell populations were observed. While reporter expression could be fully silenced by recruitment of EED or the DNMT3A-CD in some cells, DOX treatment had little to no effect on *mCherry* expression in other cells. This observation was attributed to locus specific epigenetic maintenance effects preventing induced silencing in the non-responsive cells. In CHO cells, it has been described that the main variance factor for the transcriptional regulation of a transgene is locus-dependent and caused by the random integration into the host (Grav et al., 2018; Jae Seong Lee et al., 2018). By working with monoclonal reporter cell populations, locus-specific mechanisms of chromatin maintenance and regulation upon an epigenetic trigger could be analysed, greatly contributing to the understanding of transcriptional control. Primary experiments in this thesis have shown that the creation of monoclonal reporter populations is possible, and that those reporter populations react in a locus-specific manner to LSD1 induced silencing (**Figure 31**). Furthermore, the dependency of transcriptional repression following LSD1 recruitment on the activity of the helicase DDX19A was not observed at all genetic integration sites (**Figure 31**). This suggests that the creation of monoclonal reporter populations represents a valuable approach for the investigation of epigenetic barriers preventing the targeted transcriptional

repression of specific genomic loci. However, the creation of a monoclonal cell population is very time-consuming and cannot be achieved successfully with all types of cells. The arguments that still speak in favour of working with monoclonal populations will be discussed in the following section.

5.1.3. Challenges on the way towards the computational modelling of gene regulation

All methodical approaches that try to describe the regulation of eukaryotic gene expression face the problem of the high complexity of this process. Transcription from a specific DNA locus is controlled by multiple factors, including the affinity of the RNA polymerase and TFs to the promoter and enhancer sequences, local nucleosome occupancy, chromatin modifications and chromatin structure, and the three-dimensional nuclear localization of the sequence (Nicolas et al., 2017). This is further complicated by the fact that said impact factors also regulate each other, e.g. local levels of chromatin modifications influence the local chromatin structure and the affinity of RNA polymerase and TFs (Tycko et al., 2017). Furthermore, gene expression occurs in waves of transcriptional bursting, alternating between "on" and "off" states of transcription (Sanchez & Golding, 2013), which leads to high cell-to-cell heterogeneity on a temporal scale. This implies that the transcriptional output measured from a reporter gene must be considered as a statistical event that is controlled by a variety of factors. Aiming to extract the impact of an individual regulatory factor (e.g. the deposition of a specific chromatin modification) on transcriptional activity from the analysis of the transcriptional output of a reporter gene, a statistical noise created by all those other factors must always be anticipated in the experimental design and considered in the final evaluation of the results.

One approach to reduce the complexity of systems for the analysis of transcriptional regulation is to conduct the measurements on a single-cell level. Expression levels can be visualized on the single-cell level using microscopy (Bintu et al., 2016; Paige et al., 2011; Raj et al., 2006), and several sequencing approaches to analyse chromatin-structure and modifications in single-cells have been developed (Buenrostro et al., 2015; Cusanovich et al., 2015; Rotem et al., 2015). However, all these methods suffer from high costs, the need for elaborate equipment and limitations concerning their dynamic sensitivity and the achievable sequence coverage.

In this work, by creating monoclonal fluorescent reporter cell populations, the locus-dependent variances in the gene regulatory environment (Grav et al., 2018; Jae Seong Lee et al., 2018) were minimized under retention of the temporal dynamics of the system and of a high number of available data points. The dynamics of reporter fluorophore expression upon recruitment of chromatin regulator domains have been studied on the single-cell level using time-lapse microscopy (Bintu et al., 2016). However, the method presented by Bintu et *al.* requires a sophisticated quantification method to account for the distribution of the fluorescent protein on daughter cells during cell divisions. Furthermore, the number of acquired data points ranges in a magnitude of 10^2 cells per experiment, opposed to 10^4-10^5 cells that can easily be analysed in one run using flow cytometry. This high number of data points allows for the quantification of fluorescence signals by analysis of population medians, thereby compensating for temporal variances of the fluorescence signal that arise from transcriptional bursting or during cell-cycle.

Comprehensive computational models for the regulation of gene expression upon chromatin factor binding must be able to integrate the effects of transcriptional bursting, spreading of chromatin states and of three-dimensional chromatin looping (**Figure 40**). Synthetic biology approaches deliver the experimental data that are essential to set up computational models (Tycko et al., 2017). Using the monoclonal fluorescent reporter system presented in this thesis, quantifiable changes in transcriptional output of a reporter gene can be reassigned to the recruitment of a specific effector domain. In a monoclonal population, many gene regulatory factors are similar among all cells and can be regarded as constants in simplified models describing the events at the synthetic gene. Residual variances in the expression levels of endogenous TFs, the association of the reporter gene with endogenous promoters and enhancers (**Figure 40a**), endogenous chromatin maintenance mechanisms (**Figure 40b**) and effects of the nuclear localization of the reporter gene (**Figure 40c**) are compensated by the high number of data points acquired in flow cytometry. This leaves the recruitment and the activity of a chromatin regulator at the synthetic gene as the major factor influencing gene expression.
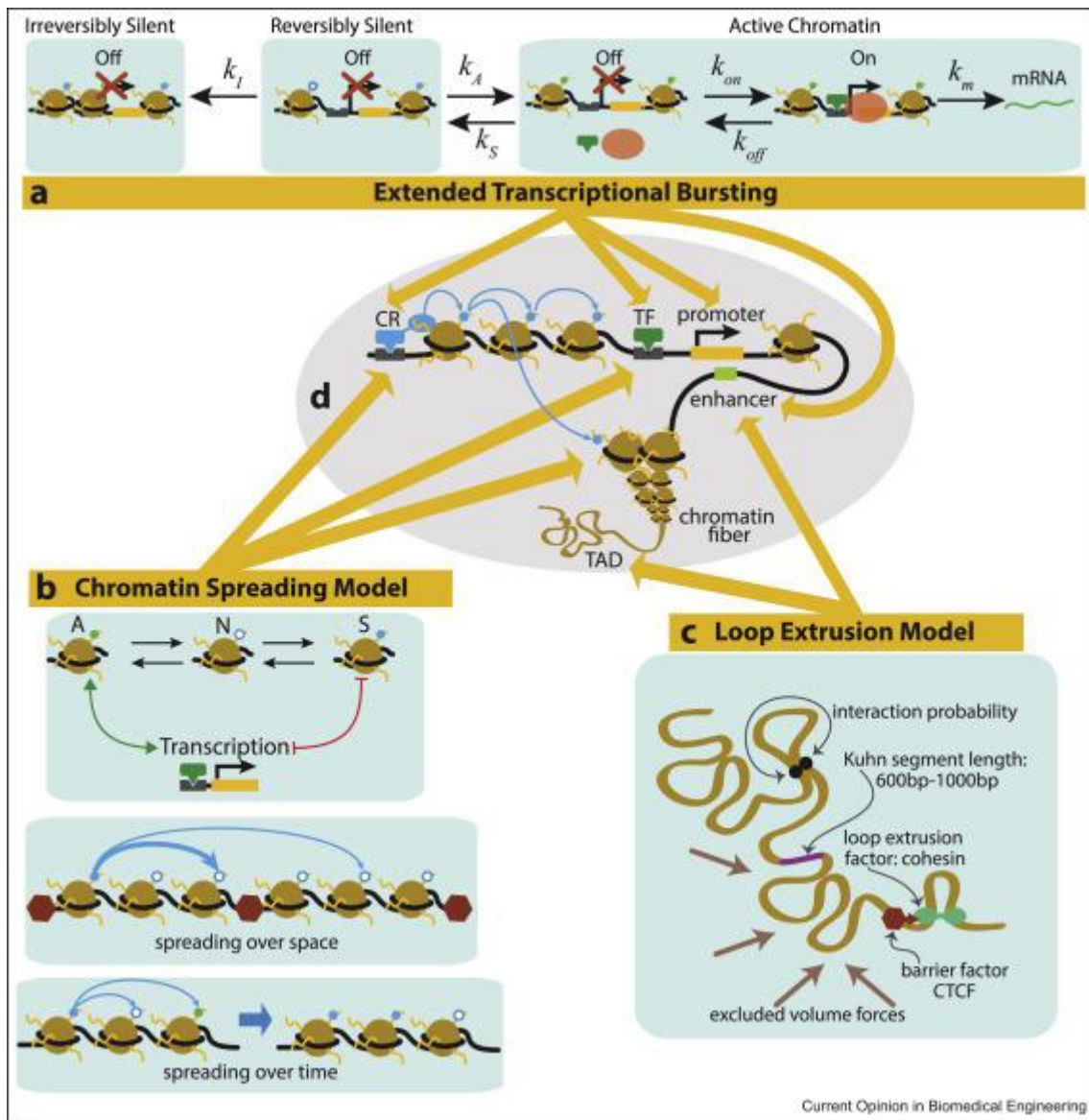
Figure 40: Different models for the regulation of gene expression by chromatin factors. Transcription from an endogenous promoter is dependent on the nature of the local chromatin, which can switch between three states: *active*, *reversibly silent* and *irreversibly silent*. Transition between these states occurs at statistical rates *k*, that can be simulated based on different models. **a,** The Extended Transcriptional Bursting model proposes that gene expression is a statistical event regulated by the temporal association of a genomic locus with chromatin regulators (CR), transcription factors (TF), and promoter- and enhancer-sequences. **b,** The Chromatin Spreading Model predicts the dynamics of spreading and maintenance of a chromatin state over space and time, leading to inhibition or activation of transcription. **c,** The Loop Extrusion Model simulates the probability of chromatin contacts in the three-dimensional space of chromatin organization, helping to predict the association of regulatory sequences with transcriptional start sites over long linear distances. **d,** The models (**a-c**) can be integrated to simulate the regulatory environment of a genomic locus. Image was taken from Tycko et al. (2017).

The production of monoclonal reporter cell populations is beneficial for the intended use of the reporter system to quantitatively analyse the dynamics of transcriptional repression induced by different epigenetic effector domains. Furthermore, the resulting homogenization of the fluorescent signal enhances observed effects of perturbations on transcriptional repression, thus improving the identification and validation of chromatin regulator dependencies. On the other hand, RNAi experiments with the DEAD-box helicase DDX19A and LSD1 as a repressor domain showed that the dependency of LSD1 on DDX19A activity was locus-dependent. This suggests that for loss-of-function screening approaches (RNAi/CRISPR) with the aim to identify all possible functional genetic interactions of an epigenetic regulator, the application of heterogenous reporter cell populations should be considered to avoid false-negative results. Furthermore, mammalian cells hold the potential to form a clonally stable transcriptional phenotype that is driven by epigenetic memory mechanisms (Meir et al., 2020; Shaffer et al., 2020). The performance of a comprehensive RNAi screen in a monoclonal population might detect regulatory effects that are specific to the clonal transcriptional environment in this cell population, but do not constitute a generally and globally active regulatory mechanism.

5.2. Identification and characterization of novel interactors of the lysine-specific demethylase 1 (LSD1)

The concept that transcriptional regulation is driven by the formation of a permissive or repressive chromatin state, which is the result of the concerted action of epigenetic effector proteins, has been established many years ago (Deaton & Bird, 2011; Grewal & Moazed, 2003). Epigenetic effector proteins assemble in multiprotein complexes in a combinatorial way to install cell-type specific gene expression programs (Lange et al., 2011), and mistakes in the configuration of these regulatory complexes can lead to the development of diseases (Garay et al., 2016; Hodges et al., 2016; Jiang, 2020). Statistical approaches to model the regulatory mechanisms controlling gene expression are receiving growing attention, aiming to eventually predict transcriptional outcomes on a global level, *e.g.* in response to an epigenetic drug treatment (Angelin-Bonnet et al., 2019; Tycko et al., 2017). These mathematical approaches rely on the existence of reliable experimental data from high-throughput approaches that comprehensively describe interactions in a gene regulatory network. Powerful

proteomics approaches studying the composition of gene regulatory complexes in association with modified chromatin have been developed (Wierer & Mann, 2016). However, these data need to be supported by functional information about the dynamic processes that shape gene regulatory networks in a living biological system. Using the reporter system presented in this work, the transcriptional response to exogenous triggers of the epigenetic network controlling expression of the fluorescent gene was analysed. In doing so, novel functional interactions inside the regulatory network surrounding LSD1 were identified, characterized, and annotated to a so far unknown signalling cascade. The application of this method to investigate further regulatory proteins holds the potential to vastly expand the knowledge about different epigenetic regulatory cascades.

5.2.1. A novel regulatory cascade downstream of LSD1 silencing

Since its discovery as the first histone demethylase (Shi et al., 2004) and its successive rise as an interesting target for epigenetic cancer therapies (Augert et al., 2019; Y. Fang et al., 2019; Fiskus et al., 2014), the diverse roles of LSD1 as an epigenetic regulator have intensively been studied. LSD1 acts as a transcriptional repressor in many repressive multiprotein complexes (Bornelöv et al., 2018; S.-A. Kim et al., 2020; Y. Yang et al., 2018). LSD1 has also been described to act as transcriptional activator by removing repressive H3K9me2 marks in neuronal cells (Laurent et al., 2015) and in association with steroid hormone receptors (Bennesch et al., 2016; Carnesecchi et al., 2017; Metzger et al., 2005). However, this function appears to be rather cell-type specific, with the splice variant of LSD1 acting in neuronal differentiation being restricted to expression in neuronal cells (Laurent et al., 2015; Zibetti et al., 2010), and the associated androgen and estrogen receptors being mainly expressed in reproductive tissues (Fagerberg et al., 2014; Hua et al., 2018). It has been shown that the activity of LSD1 is strongly dependent on the binding of cofactors to its TOWER domain, suggesting that structural changes to the catalytic cavity induced by substrate- and cofactor-binding determine the specificity of its catalytic activity (S.-A. Kim et al., 2020; Stavropoulos et al., 2006). This might explain why the *in vitro* activity of LSD1 on H3K9me2 can only be observed if LSD1 is either copurified with its associated factors, or if those factors are additionally added to the reaction (Carnesecchi et al., 2017; Laurent et al., 2015). More generally, it reveals the importance of the interaction

of LSD1 with different coregulators in determining its regulatory activity at a genetic locus.

Despite its mostly repressive role, LSD1 is associated with core promoters and enhancers of actively transcribed genes in mESCs, together with ESC-specific TFs like *Oct4* and *Nanog* and the transcription machinery (Petell et al., 2016; Whyte et al., 2012). Similarly, the PRC2 complex can bind to transcriptionally active promoters in mESCs without being catalytically active at these loci (Kaneko et al., 2013). How LSD1 and PRC2 are recruited to and kept inactive at these loci has remained unclear.

In this work, PRC2 activity was discovered to be involved in the silencing of a strong synthetic promoter by LSD1 recruitment (**Figure 20**, **Figure 21** and **Figure 28**). LSD1 has been coprecipitated with PRC2 components (Y. Jin et al., 2017), and LSD1 and the PRC2 complex subunit EZH2 have been shown to be corecruited by lncRNAs, suggesting an involvement of RNA at loci coregulated by LSD1 and PRC2 (Lian et al., 2017; M. Sun et al., 2016; Tsai et al., 2010; T. peng Xu et al., 2018). This is in line with the observation that the formation of R-loop structures at the synthetic reporter gene had a regulatory effect on both LSD1 and PRC2 activity. In literature, the presence of RNA structures at a genomic locus has been described to have either an inhibitory, recruiting or activating effect on PRC2 (Almeida et al., 2020; Beltran et al., 2016; Y. Long et al., 2020; Mocavini & Di Croce, 2020; X. Wang et al., 2017; Qi Zhang et al., 2019). G4-RNA structures have been shown to act as non-competitive inhibitors of LSD1 demethylase activity *in vitro*, and a regulatory role of RNA structures for LSD1 activity at genomic sites has been proposed (Hirschi et al., 2016). Despite all these observations, the regulatory effects of RNA structures on the activities of LSD1 and PRC2 on chromatin remain ill-defined.

A controversially discussed biological example for the coregulation of gene expression by LSD1 and PRC2 is their co-occupancy of so-called bivalent regions, which are genetic elements that control the expression of developmental genes in human ESCs (Adamo et al., 2011). These bivalent promoters have been proposed to serve as "poised" gene-expression switches, with the ability to rapidly initiate novel transcriptional programs upon pro-differentiation stimuli in mESCs, but the exact nature of this switch remains controversial (D. Hu et al., 2013; Piunti & Shilatifard, 2016; Voigt et al., 2013). Several publications propose that the balance of H3K4 trimethylation and H3K27 trimethylation at bivalent regions regulates the establishment

of a Polycomb-domain contact network inside the euchromatic chromatin compartment (Denholtz et al., 2013; Glòria Mas et al., 2018; Gloria Mas & Di Croce, 2016; Vieux-Rochas et al., 2015). Disturbance of this balance in ESCs, e.g. by repression of the H3K4 methyltransferase MLL2, leads to the global re-organization of chromatin architecture and consequently in the severe impairment of the onset of developmental programs after induction of differentiation (Glòria Mas et al., 2018). The formation of RNA-structures might also play a role in the regulation of these bivalent loci. Both sense- and antisense transcription can be stabilized by the presence of H3K4me3, which might lead to an increase in R-loop formation (Boque-Sastre et al., 2015). The model of bivalent regions acting as a developmental switch is intriguing, and it is possible that toggling of this switch might be regulated by the presence of RNA structures. Several non-coding RNAs are under the control of pluripotency-related TFs, and are differentially expressed during the stages of differentiation, indicating an essential role of lncRNAs in pluripotency maintenance and lineage-specific differentiation (Jingcheng Chen et al., 2020; Z. Du et al., 2018; Fico et al., 2019).

Despite the huge number of articles published on LSD1 and the PRC2 complex, the idea of their regulatory interaction remains diffuse. It is assured that there is a so-called "poised" state of EZH2- and LSD1-binding at highly expressed genes, and that RNA has a regulatory role on the activity of both enzymes. The bivalent model of poised genetic regions assumes that both H3K4me3 and H3K27me3 are present at these loci, indicating an inactive state of LSD1 and an active state of PRC2. Mathematical models integrating experimental data of poised loci delivered an alternative explanation, suggesting that these regions constantly switch between a fully active and a repressed chromatin state (Sneppen & Ringrose, 2019). Although these models are highly simplified, they are in line with the observations of transcriptional regulation made in this work. By integration of the findings that followed the discovery of DDX19A as a positive regulator of LSD1 activity (**Figure 30**), a model for the regulatory activity of LSD1 at highly transcribed genomic loci was proposed (**Figure 41**). LSD1 and PRC2 can co-occupy the *cis*-regulatory regions of actively transcribed genes. The persistent transcriptional activity leads to the local formation of R-loops (Boque-Sastre et al., 2015; Ginno et al., 2012) and stabilizes active histone marks by the recruitment of HMTs and HATs (Herz et al., 2013; Shilatifard, 2012; Z. Wang et al., 2009) that counterbalance the activity of LSD1. PRC2 activity is inhibited by both the presence of

H3K4 trimethylation and the formation of cotranscriptional R-loops (Laugesen et al., 2019; Mocavini & Di Croce, 2020). The resolution of R-loops is regulated by the presence of topoisomerases and helicases like the DNA:RNA helicase DDX19A to prevent stalling of the transcription machinery (Skourti-Stathaki & Proudfoot, 2014). Upon an external trigger, LSD1 activity is enhanced, overcoming the methyltransferase activity of HMTs and leading to the local decrease in H3K4 methylation. After the removal of H3K4 methylation, PRC2 can methylate H3K27 to form H3K27me3, which is a binding signal for DDX19A. Either by allosteric activation or enhanced recruitment, DDX19A then efficiently resolves R-loops at the promoter region. Thus, the repressive activity of the DNA:RNA hybrids on LSD1 and PRC2 is removed and their catalytic activities are increased, forming a positive feedback loop for stable epigenetic repression of the genetic locus.
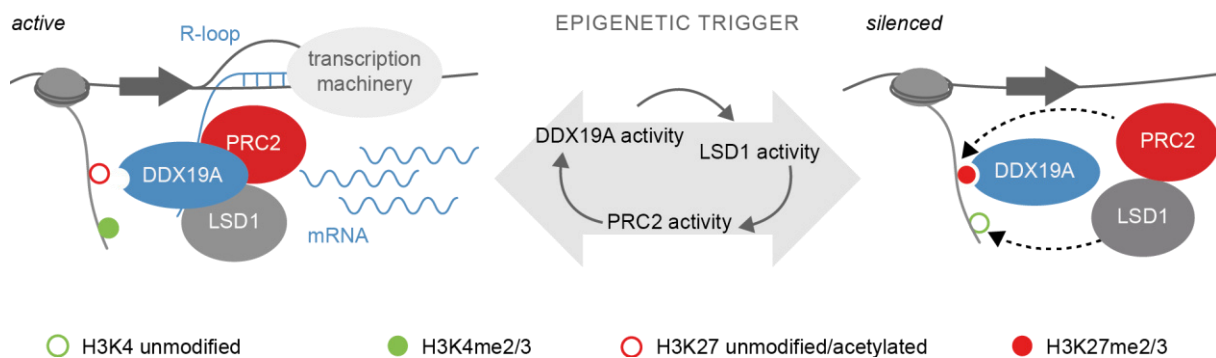
Figure 41: LSD1, DDX19A and PRC2 cooperate in the regulation of transcription at highly expressed genes. Schematic depicting the regulatory cascade leading to stable silencing of active genes downstream of LSD1 activity. LSD1 and PRC2 are localized at highly transcribed genes. Active transcription promotes the local presence of H3K4me2/3 and the formation of R-loops. These R-loops are balanced by specific helicases (e.g. DDX19A). Upon an external repressive stimulus LSD1 activity is increased leading to a reduction in H3K4 methylation, which enables the PRC2 complex to methylate H3K27. H3K27me3 serves as a binding motif for DDX19A, which then efficiently removes local R-loops, enabling the further increase of LSD1 activity and thus closing a repressive feedback loop that leads to robust silencing of transcription.

By inhibition or suppression of a single component of this regulatory cascade, silencing of reporter expression remained incomplete. Furthermore, the initial expression level of the reporter fluorophore was reinstalled quickly after removal of the repressive trigger. Together, these observations indicate that the model of bistable chromatin states probably is better suited to describe the transcriptional regulation of poised loci by LSD1 and PRC2 than the model of bivalency (Sneppen & Ringrose, 2019).

All experiments with suppression of DDX19A or inhibitor treatment for LSD1/PRC2 were performed in heterogeneous cell populations. This resulted in distinct, but not in all-or-none effects of the disruption of the cascade on the efficiency of reporter silencing. By utilizing the ability of the synthetic reporter system to resolve locus-specific effects when analysing multiple monoclonal reporter populations, it would be possible to determine the nature of the chromatin environment that is responsive to changes in the LSD1-PRC2-R-loop cascade. ChIPseq experiments for PRC2 and LSD1 binding and the presence of H3K4me2 and H3K27me3 in DDX19A-supressed or RNAseH1 treated cells could help to define the effect of R-loop formation on LSD1/PRC2 recruitment and activity on a global scale. ATACseq and Hi-C-seq after stabilization or destabilization of R-loop formation could reveal a possible role of RNA structures in chromatin organization. In combination with further RNAi or inhibitor treatment approaches, the role of other epigenetic modulators like DNMT3A/B for the final silencing of expression at poised active regions could be comprehensively defined. In conclusion, application of the reporter system for the further functional analysis of the newly discovered cascade holds promise for the detailed characterization of the processes mediating stable silencing of chromatin.

5.2.2. The DEAD-box helicase DDX19A as a regulator of transcription

The comprehensive RNAi screen for functional interactors of LSD1 identified the DEAD-box containing helicase DDX19A as an essential coregulatory protein. DDX19A is an ATP-dependent RNA helicase, which was mainly regarded as a nuclear-pore associated mRNA export factor (Tran et al., 2007), before it was found to relocalize to the nucleoplasm to resolve DNA:RNA hybrid structures upon an ATR signalling trigger (Hodroj et al., 2017). Regarding these earlier annotated functions, the question of how DDX19A is connected to LSD1 and to its repressive function on transcriptional regulation was intriguing.

The human family of DEAD-box helicases is highly conserved, but the individual molecular functions of each family member are poorly characterized and can only be consolidated under the broad function of *mRNA processing* (Awasthi, Verma, et al., 2018; Linder & Jankowsky, 2011; Ma et al., 2016; Tran et al., 2007). DEAD-box helicases can remodel nuclear complex structures of RNA and proteins. By either inhibiting or enhancing the binding of proteins to nascent RNA, several DEAD-box helicases regulate transcription (Linder & Jankowsky, 2011; Ma et al., 2016; H. Zhang

et al., 2020). The yeast homologue of DDX19A (DBP5) has also been shown to directly interact with TFs, and has consequently been proposed to interfere with transcription at the level of transcriptional initiation (Estruch & Cole, 2003). Such a general model for transcriptional regulation by DDX19A is unlikely regarding the presented results, as it was shown that the knockdown of DDX19A did not interfere with KRAB-induced silencing of transcription or with the basal expression levels of synthetic reporter constructs. It is further unlikely that LSD1 or its associated cofactors are sterically repelled from the synthetic promoter due to an enrichment of R-loops in the absence of DDX19A activity. LSD1 is strongly recruited by its fusion to the rTetR, and ChIP-qPCR experiments showed that the subsequent recruitment of HDACs and H3K9 methyltransferases was not affected by suppression of DDX19A.

It has been proposed here that the function of DDX19A at the synthetic promoter was the regulation of transcription by controlling the local R-loop formation. R-loops have emerged as variable modulators of local transcriptional activity and chromatin modification patterns (Ginno et al., 2012, 2013; Sanz et al., 2016). One example for a similar control mechanism is the yeast DEAD-box helicase DBP2, which directly controls the transcriptional induction of the GAL gene cluster by regulating the formation of lncRNA:DNA hybrid structures. These hybrids repel transcriptional repressor proteins and allow for further chromatin looping that further activates transcription (Cloutier et al., 2016). There are both publications linking R-loop formation to transcriptional activation (Boque-Sastre et al., 2015; Sanz et al., 2016) and to transcriptional repression (Skourti-Stathaki et al., 2019), but many of these observations are purely correlative, and the molecular mechanisms that discriminate between transcriptional activation or repression by the presence of R-loops are unclear. Recently, a mechanism for the regulatory function of cotranscriptional R-loops in chromatin silencing has been proposed in *Arabidopsis thaliana* (C. Xu et al., 2021). This mechanism requires the concerted action of a H3K4 demethylase, the PRC2 complex, and proteins that mediate R-loop resolution to install a repressed chromatin state at an actively transcribed locus (X. Fang et al., 2020; C. Xu et al., 2021). Interestingly, a novel H3K27me3/H4K20me3 reader function could be attributed to the R-loop resolving DEAD-box helicase DDX19A. This discovery might represent an important puzzle piece on the way to understand transcriptional regulation by R-loops,

functionally linking the chromatin modification environment to the local stability of R-loop formation in mammals.

To further determine the function of DDX19A in transcriptional regulation, the biochemical characterization of DDX19A as an epigenetic effector protein will be of interest. The ATPase activity of DDX19A has been shown to be regulated by the presence of ssRNA in an autoinhibitory way (Collins et al., 2009). How would the activity be altered by the binding of modified histones? Where is the binding domain of DDX19A? How is the activity of DDX19A regulated by the published association with TFs? The combination of ChIP-seq and DRIP-seq experiments could reveal if the activity of DDX19A is restricted to PRC2-repressed loci, or a general feature of R-loop prone regions.

Two other DEAD-box helicases, DDX39B and DDX23, promoted the silencing of the *synP* element by LSD1 (**Figure 24**). However, this observation might be ascribed to LSD1 effects unrelated to LSD1 activity. Both DDX39B and DDX23 have essential functions in mRNA transport and splicing, and DDX39B can promote translation by regulation of pre-ribosomal RNA levels (Awasthi, Chakrapani, et al., 2018; Mathew et al., 2008; Nakata et al., 2017). Thus, the low scoring of those DEAD-box helicases is likely due to severe deregulation of the expression levels of reporter system components.

In conclusion, the discovery of DDX19A as a reader of histone *modifications* opens a whole new field of regulatory mechanisms, directly linking the presence of a specific histone code to the stability of local RNA structures.

5.3. Variability in the multi-protein complex assembling around the KRAB repressor domain of ZNF10

Besides the LSD1 project, the synthetic reporter system was applied for the analysis of cell line specific variances in the repressive complex surrounding the KRAB domain of ZNF10. KRAB-ZFPs play an essential role in the regulation of retroviral elements and other repetitive regions in mammalian genomes (Ecco et al., 2017; Imbeault et al., 2017). Due to its repressive function, the ZNF10 KRAB domain is very commonly used for the directed silencing of gene expression in synthetic biology and CRISPRi approaches (Alerasool et al., 2020; Bintu et al., 2016; Pickar-Oliver & Gersbach, 2019;

Tycko et al., 2020). However, artificially induced silencing by the KRAB domain can vary between reversible and irreversible repression, and suffers from being incomplete in some cases (Alerasool et al., 2020; Yeo et al., 2018; Ying et al., 2015). A better understanding of the epigenetic mechanisms mediating transcriptional repression downstream of KRAB-binding is essential for the improvement of the predictability and efficiency of CRISPRi and other KRAB-related approaches.

In this work, the KRAB domain was initially chosen to validate the sensitivity of the reporter system towards RNAi. Silencing by this domain was published to be dependent on its association with the HUSH complex, and efficient HUSH-targeting shRNAs were available in the laboratory (Robbez-Masson et al., 2018; Tchasovnikarova et al., 2015).

First, the interaction of KRAB and the HUSH complex was validated in NIH/3T3 cells. KRAB recruits the scaffolding protein KAP-1, which in turn recruits the members of the HUSH complex. The HUSH components ATF7IP, D14ABB1E, MPHOSPH8 and SETDB1 were all validated to be important for KRAB-induced silencing of the fluorescent reporter in NIH/3T3 (**Figure 34** and **Figure 35**). However, the impact of the individual HUSH components on reporter silencing and the global maintenance of H3K9me3 in this cell line varied in its extent. The H3K9me3 reader protein MPHOSPH8 and the H3K9 methyltransferase SETDB1 were identified as strongest coregulators of KRAB-induced silencing (**Figure 34**), and suppression of both proteins had a strong impact on the level of global H3K9me3 (**Figure 33**). This was in line with published data, which propose that the H3K9me3 reader and writer proteins cooperate in the maintenance and spreading of heterochromatic regions in vertebrates, in a mechanism comparable to PEV in *Drosophila* (Cruz-Tapias et al., 2019; Timms, Tchasovnikarova, & Lehner, 2016).

Suppression of the central HUSH subunit D14ABB1E led to a severe decrease in global H3K9me3 levels (**Figure 33**). However, upon recruitment of KRAB to the synthetic reporter it did not appear to be a stronger transcriptional coregulator than the other HUSH components (**Figure 34**), an observation that was in line with similar published experiments using reporter genes (Robbez-Masson et al., 2018). D14ABB1E has been described to be essential for HUSH activity at transcribed repetitive regions, with a mechanism that is dependent on RNA binding by an inactive PARP domain and its interaction with MPHOSPH8 (Douse et al., 2020). It is assumed

that the D14ABB1E dependent assembly of HUSH can function to control the rate of untimely transcription of repetitive elements, thus preventing genomic damage (Douse et al., 2020). The essential function of *D14Abb1e* as a member of the HUSH complex might be specific for the targeting and maintenance of heterochromatin at repetitive genomic regions, as suppression of *D14Abb1e* did not manage to reactivate ZNF91 repressed reporter genes, in contrast to suppression of other HUSH components (Gresakova et al., 2019; Robbez-Masson et al., 2018). The HUSH components SETDB1 and ATF7IP have been shown to be no physical interactors of D14ABB1E (Douse et al., 2020). Thus it can be suggested that the KRAB mediated reporter silencing observed in NIH/3T3 cells was initiated by the direct recruitment of SETDB1 by KAP-1 (Sripathy et al., 2006), while D14ABB1E acted as an enhancer of H3K9me3 spreading by recruiting MPHOSPH8 and PPHLN1 (Douse et al., 2020).

Suppression of *Atf7ip* had only small effects on the silencing of the fluorescent reporter, as well as on the maintenance of global H3K9me3 levels in NIH/3T3 compared to the other tested HUSH subunits (**Figure 33**). ATF7IP is a chaperone of SETDB1 and has been proposed to enhance the recruitment of MPOSPH8 (Timms et al., 2016; Tsusaka et al., 2018). ATF7IP is essential for the protection of SETDB1 from proteasomal degradation, and knockout of ATF7IP has been shown to phenocopy the loss of SETDB1 (Timms et al., 2016). Regarding the crucial role of SETDB1 for HUSH activity in NIH/3T3, it was suggested that the reduced levels of ATF7IP following RNAi were still sufficient to secure a functional HUSH complex in this cell line. Thus, the effect of ATF7IP reduction appeared less pronounced than the effect by direct suppression of SETDB1.

5.3.1. A comprehensive RNAi screen probing the regulatory network surrounding KRAB

In contrast to the findings in NIH/3T3, the interaction of KRAB and the HUSH complex could not be observed in iMEF (**Figure 36**). An RNAi screen in iMEF cells targeting 1104 chromatin-associated genes, including the HUSH complex members with 6 different shRNAs per gene, was performed to confirm the results from the single-shRNA experiments (**Figure 37**). Although the data need to be substantiated with additional experiments, several interesting findings are beginning to become apparent already.

The scaffolding protein KAP-1 that mediates the interaction of KRAB and the HUSH complex (Robbez-Masson et al., 2018) could not be analysed in the screen, as suppression of this gene was toxic in the cells (**Figure 39**). However, the downstream regulators of KRAB silencing could further be investigated. The only HUSH subunit scoring as a strong positive regulator of KRAB was *D14Abb1e*, supporting its essential role in chromatin repression downstream of KRAB recruitment, in contrast to the observations made in NIH/3T3. Gresakova et al. showed that D14ABB1E changes its protein interaction profile during development and differentiation (Gresakova et al., 2019). D14ABB1E has been shown to have common but non-redundant functions with KAP-1 (Robbez-Masson et al., 2018), and KRAB-ZFPs can assemble repressive complexes independently of their association with KAP-1 (Murphy et al., 2016). This suggests that in iMEF cells, D14ABB1E might be a mediator of KRAB induced silencing by assembling an alternative repressive complex to the KAP-1/HUSH complex. This hypothesis is supported by the fact that while *Setdb1* was a clear negative regulator of KRAB repression in iMEF, *Mphosph8* scored weakly as a positive regulator (**Figure 39**). The indistinct effect of *Mphosph8* might be due to the general depletion of some of the *Mphosph8* shRNAs in all sorted populations (**Figure 38**). D14ABB1E exerts its repressive function over the direct interaction with MPHOSPH8 and PPHLN1 (Douse et al., 2020), and the hitherto existing data of the screen suggest that the H3K9 methyltransferase activity of SETDB1 in complex with KRAB and D14ABB1E can be replaced by another methyltransferase (Fritsch et al., 2010; Herz et al., 2013).

The lower expression of all HUSH complex members in iMEF cells compared to NIH/3T3 further supported the minor role of this regulatory complex in this cell line (**Figure 39**). Interestingly, the H3K9 methyltransferase SUV39H1 was expressed at comparable levels in both cell lines and scored under the top 15 positive regulators of KRAB activity in iMEF, highlighting this methyltransferase as a possible replacement of SETDB1. The individual functions of the H3K9 methyltransferases SUV39H1, SUV39H2, G9A, GLP, SETDB1, and SETDB2 have been found hard to determine, and it is assumed that they act redundantly depending on the local regulatory context, recruiting mechanism and expression pattern (Fritsch et al., 2010; Herz et al., 2013).

Many epigenetic multiprotein complexes consist of an initial core complex that can associate with a choice of protein subunits, depending on the current developmental stage or differentiated cell type (Adams et al., 2018; Hodges et al., 2016; Jiang, 2020;

Malovannaya et al., 2011). Analysis of the top scoring genes in the screen (**Figure 39d**) provided an impulse for alternative silencing mechanisms active downstream of rTetR-KRAB recruitment in this system. KRAB/KAP-1 has been described to induce the formation of a heterochromatic multiprotein complex consisting of HP1, SETDB1 and NuRD/CoREST (P. Yang et al., 2017). However, like SETDB1 the NuRD complex associated proteins CHD3 and CHD4, GATAD2B, MTA1, MTA2 and MTA3 and MBD2 also were not identified as positive coregulators of KRAB in iMEF cells. The CoREST subunit LSD1 (Kim et al., 2020) also scored as negative regulator, suggesting that the recruitment of NuRD/CoREST was not the main repressive mechanism responsible for KRAB reporter silencing. However, *Rcor1* was identified as a positive coregulator of KRAB (**Figure 39a**). Furthermore, the zinc finger protein ZMYM4, which has been shown to recruit RCOR1 and LSD1 (Cibis et al., 2020), also scored as a top coregulator (**Figure 39d**). The demethylation of H3K4 is an essential step in the formation of a transcriptionally repressed chromatin state (Broche et al., 2021; Hathaway et al., 2012; Saksouk et al., 2015). Further experiments will have to determine whether LSD2, an orthologue of LSD1 that scored high at position 27 in the evaluation of the current KRAB screening data, could replace LSD1 as a H3K4 demethylase in this context, although it has not been shown to interact with RCOR1 (R. Fang et al., 2013). Finally, four members of the SIN3A complex (*Sap130, Sap30, Mecp2* and *Sap25*) were enriched in the positive population (**Figure 39a**). Especially the high scoring of *Mecp2* and *Sap30* under the top 15 coregulators of KRAB (**Figure 39d**) indicated that the SIN3A complex might replace NuRD as main corepressor complex of KRAB in iMEF cells.

5.3.2. Future experiments to substantiate the observations from the iMEF KRAB screen

The presented data from the shRNA-mirE screen in iMEF show some interesting trends and offer an impulse for several downstream experiments. However, several experiments remain to be done to validate the observations presented so far. Although all shRNAs that were used were validated for their knockdown efficiency, their specific efficiency in iMEF cells should be determined. The RNAi experiments should further be reproduced with additional shRNAs targeting the same genes. Biological replication of the screening procedure could help to substantiate the impact of individual potential interactors of KRAB emerging in the existing data by providing statistical significance.

Multiple neutral control shRNAs should be implemented in further replicates of the screen to account for exceptional secondary effects and gain a better picture of the reproducibility of the screening procedure. Also, positive control shRNAs targeting the KRAB domain or the rTetR-KRAB fusion constructs should be established and included in future screens. Statistically significant coregulators could then be characterized in cellular and biochemical experiments. By performing the RNAi screen in NIH/3T3 cells, the results from the single-shRNA experiments could be validated and the different modes of action of KRAB silencing in both cell lines could be compared in side-by-side experiments. Although the KRAB domain of ZNF10 has been used as a repressive tool in synthetic biology for decades, unanticipated inefficiencies in the repression of target gene expression have been observed and condoned so far. A better understanding of variant KRAB functions, besides its well-described recruitment of cofactors via KAP-1, might help in the design of more effective KRAB repressor domains for the application in CRISPRi (Alerasool et al., 2020; Yeo et al., 2018). Furthermore, the development of specific tandem fusions of dCas-KRAB and additional transcriptional regulators for the application in different cell lines should be considered as already done by (Alerasool et al., 2020; Yeo et al., 2018).

5.4. Conclusions and outlook

In conclusion, a novel reporter system that expands the tool kit for the analysis of epigenetic mechanisms by a modular and dynamic high-throughput screening platform was successfully established during this doctoral thesis. It was applied to characterize the epigenetic network surrounding the relevant oncogene LSD1, culminating in the discovery of an unknown epigenetic regulatory cascade. Multiple novel coregulators of LSD1 activity were discovered. Further biochemical characterization of those proteins might help to understand unanticipated responses to epigenetic treatments that are already in clinical trials, and could allow for combinatorial treatments that prevent resistance development. From a fundamental research perspective, this work led to the expansion of the regulatory network surrounding an important epigenetic effector protein by two novel, so-far unrelated functional mechanisms. Future experiments characterizing the functions of LSD1 will define the mechanistic effects of local R-loop formation on LSD1 repression.

The presented data lead the way for additional experiments aiming to answer several interesting research questions. First, the regulatory function of RNA on epigenetic effector proteins is still poorly defined, although it draws increasing attention as a regulatory factor, especially at genetic loci in an equilibrium state between transcriptional activity and silence. During this work, the role of R-loops as important regulators of LSD1 and the PRC2 complex was described. It will be interesting to further characterize the influence of DNA- and RNA-secondary structures on the catalytic activity and repressive potential of different epigenetic effector proteins using the original synthetic promoter and the designed variants.

By application of the synthetic reporter and the RNAi system in different cell lines, a variability in the repressive complex forming around the KRAB domain of ZNF10 was discovered. Multiple groups are trying to improve KRAB domains used for epigenome editing with CRISPRi (Alerasool et al., 2020; O'Geen et al., 2017; Yeo et al., 2018). A more detailed analysis of the coregulator network of KRAB using the reporter system might help to understand the inconsistent response of different genomic targets or cell lines to KRAB recruitment. This knowledge will enable the engineering of more effective KRAB domains for epigenome editing in different cellular backgrounds.

As a long-term goal, the screening of various epigenetic repressor domains will provide information about a multitude of coregulator interactions in transcriptional networks. Integration of these data will deliver insights into the nodes and intersections of different repressive epigenetic pathways. Using computational methods, a comprehensive network of epigenetic regulation could be simulated based on these data. Making use of the dynamic properties of the system, the chronological order of events leading from open to closed chromatin and vice versa can be analysed, providing another layer of information for the network simulation. Bioinformatic approaches hold promise to eventually reveal the complex mechanisms behind epigenetic regulation, and solid and comprehensive experimental data are the foundation on which the computational analysis will be built.

6.   References

Aapola, U., Shibuya, K., Scott, H. S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Kawasaki, K., Minoshima, S., Krohn, K., Antonarakis, S. E., Shimizu, N., Kudoh, J., & Peterson, P. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics*, *65*(3), 293–298. https://doi.org/10.1006/geno.2000.6168

Adamo, A., Sesé, B., Boue, S., Castaño, J., Paramonov, I., Barrero, M. J., & Belmonte, J. C. I. (2011). LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nature Cell Biology*, *13*(6), 652–659. https://doi.org/10.1038/ncb2246

Adams, G. E., Chandru, A., & Cowley, S. M. (2018). Co-repressor, co-activator and general transcription factor: The many faces of the Sin3 histone deacetylase (HDAC) complex. *Biochemical Journal*, *475*(24), 3921–3932. https://doi.org/10.1042/BCJ20170314

Ahmad, K., & Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Molecular Cell*, *9*(6), 1191–1200. https://doi.org/10.1016/S1097-2765(02)00542-7

Alcázar-Román, A. R., Tran, E. J., Guo, S., & Wente, S. R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nature Cell Biology*. https://doi.org/10.1038/ncb1427

Alecki, C., Chiwara, V., Sanz, L. A., Grau, D., Arias Pérez, O., Boulier, E. L., Armache, K. J., Chédin, F., & Francis, N. J. (2020). RNA-DNA strand exchange by the Drosophila Polycomb complex PRC2. *Nature Communications*, *11*(1), 1–14. https://doi.org/10.1038/s41467-020-15609-x

Alerasool, N., Segal, D., Lee, H., & Taipale, M. (2020). An efficient KRAB domain for CRISPRi applications in human cells. *Nature Methods*, *17*(11), 1093–1096. https://doi.org/10.1038/s41592-020-0966-x

Allfrey, V. G., Faulkner, R., & Mirsky, A. E. (1964). Acetylation and Methylation of Histones and their possible Role in the Regulation of RNA Synthesis. *Proceedings of the National Academy of Sciences of the U.S.A.*, *51*(5), 786–794. https://doi.org/10.1073/pnas.51.5.786

Allis, C. D., Jenuwein, T., & Reinberg, D. (2007). *Epigenetics* (First Edit). Cold Spring Harbor Laboratory Press.

Allshire, R. C., Nimmo, E. R., Ekwall, K., Javerzat, J.-P., & Cranston, G. (1995). *Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation*.

Almeida, M., Bowness, J. S., & Brockdorff, N. (2020). The many faces of Polycomb regulation by RNA. In *Current Opinion in Genetics and Development* (Vol. 61, pp. 53–61). Elsevier Ltd. https://doi.org/10.1016/j.gde.2020.02.023

Alsaqer, S. F., Tashkandi, M. M., Kartha, V. K., Yang, Y. T., Alkheriji, Y., Salama, A., Varelas, X., Kukuruzinska, M., Monti, S., & Bais, M. V. (2017). Inhibition of LSD1 epigenetically attenuates oral cancer growth and metastasis. *Oncotarget*, *8*(43), 73372–73386. https://doi.org/10.18632/oncotarget.19637

Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L., & Lombardo, A. (2016). Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell*, *167*(1), 219-232.e14. https://doi.org/10.1016/j.cell.2016.09.006

Amir, R. E., Van Den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. *Nature Genetics*, *23*(2), 185–188. https://doi.org/10.1038/13810

Amit, I., Winter, D. R., & Jung, S. (2016). The role of the local environment and epigenetics in shaping macrophage identity and their effect on tissue homeostasis. In *Nature Immunology* (Vol. 17, Issue 1, pp. 18–25). Nature Publishing Group. https://doi.org/10.1038/ni.3325

Angelin-Bonnet, O., Biggs, P. J., & Vignes, M. (2019). Gene Regulatory Networks: A Primer in Biological Processes and Statistical Modelling. In *Methods in Molecular Biology* (Vol. 1883, pp. 347–383). Humana Press Inc. https://doi.org/10.1007/978-1-4939-8882-2_15

Änkö, M. L., & Neugebauer, K. M. (2010). Long noncoding RNAs add another layer to pre-mRNA splicing regulation. In *Molecular Cell* (Vol. 39, Issue 6, pp. 833–834). Cell Press. https://doi.org/10.1016/j.molcel.2010.09.003

Antonin, W., & Neumann, H. (2016). Chromosome condensation and decondensation during mitosis. In *Current Opinion in Cell Biology* (Vol. 40, pp. 15–22). Elsevier Ltd. https://doi.org/10.1016/j.ceb.2016.01.013

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., & Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature*, *455*(7214), 818–821. https://doi.org/10.1038/nature07249

Augert, A., Eastwood, E., Ibrahim, A. H., Wu, N., Grunblatt, E., Basom, R., Liggitt, D., Eaton, K. D., Martins, R., Poirier, J. T., Rudin, C. M., Milletti, F., Cheng, W. Y., Mack, F., & MacPherson, D. (2019). Targeting NOTCH activation in small cell lung cancer through LSD1 inhibition. *Science Signaling*, *12*(567), 2922. https://doi.org/10.1126/scisignal.aau2922

Awasthi, S., Chakrapani, B., Mahesh, A., Chavali, P. L., Chavali, S., & Dhayalan, A. (2018). DDX39B promotes translation through regulation of pre-ribosomal RNA levels. *RNA Biology*. https://doi.org/10.1080/15476286.2018.1517011

Awasthi, S., Verma, M., Mahesh, A., Khan, M. I. K., Govindaraju, G., Rajavelu, A., Chavali, P. L., Chavali, S., & Dhayalan, A. (2018). DDX49 is an RNA helicase that affects translation by regulating mRNA export and the levels of pre-ribosomal RNA. *Nucleic Acids Research*, *46*(12), 6304–6317. https://doi.org/10.1093/nar/gky231

Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K., & Rauscher, F. J. (2003). Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: A mammalian cell culture model of gene variegation. *Genes and Development*, *17*(15), 1855–1869. https://doi.org/10.1101/gad.1102803

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. In *Cell Research* (Vol. 21, Issue 3, pp. 381–395). Nature Publishing Group. https://doi.org/10.1038/cr.2011.22

Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A. M., Schlegl, J., Abraham, Y., Becher, I., Bergamini, G., Boesche, M., Delling, M., D??mpelfeld, B., Eberhard, D., Huthmacher, C., Mathieson, T., Poeckel, D., Reader, V., Strunk, K., … Drewes, G. (2011). Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nature Biotechnology*, *29*(3), 255–268. https://doi.org/10.1038/nbt.1759

Bar-Sadeh, B., Rudnizky, S., Pnueli, L., Bentley, G. R., Stöger, R., Kaplan, A., & Melamed, P. (2020). Unravelling the role of epigenetics in reproductive adaptations to early-life environment. In *Nature Reviews Endocrinology* (Vol. 16, Issue 9, pp. 519–533). Nature Research. https://doi.org/10.1038/s41574-020-0370-8

Barrett, A., Santangelo, S., Tan, K., Catchpole, S., Roberts, K., Spencer-Dene, B., Hall, D., Scibetta, A., Burchell, J., Verdin, E., Freemont, P., & Taylor-Papadimitriou, J. (2007). Breast cancer associated transcriptional repressor PLU-1/JARID1B interacts directly with histone deacetylases. *International Journal of Cancer*, *121*(2), 265–275. https://doi.org/10.1002/ijc.22673

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., & Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, *129*(4), 823–837. https://doi.org/10.1016/j.cell.2007.05.009

Bates, S. E. (2020). Epigenetic Therapies for Cancer. *New England Journal of Medicine*, *383*(7), 650–663. https://doi.org/10.1056/nejmra1805035

Becker, P. B., & Workman, J. L. (2013). Nucleosome remodeling and epigenetics. *Cold Spring Harbor Perspectives in Biology*, *5*(9). https://doi.org/10.1101/cshperspect.a017905

Beerli, R. R., Segal, D. J., Dreier, B., & Barbas, C. F. (1998). Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences*, *95*(25).

Belton, J. M., McCord, R. P., Gibcus, J. H., Naumova, N., Zhan, Y., & Dekker, J. (2012). Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods*, *58*(3), 268–276. https://doi.org/10.1016/j.ymeth.2012.05.001

Beltran, M., Yates, C. M., Skalska, L., Dawson, M., Reis, F. P., Viiri, K., Fisher, C. L., Sibley, C. R., Foster, B. M., Bartke, T., Ule, J., & Jenner, R. G. (2016). The interaction of PRC2 with RNA or chromatin s mutually antagonistic. *Genome*

Research, *26*(7), 896–907. https://doi.org/10.1101/gr.197632.115

Benedetti, R., Dell'Aversana, C., De Marchi, T., Rotili, D., Liu, N. Q., Novakovic, B., Boccella, S., Di Maro, S., Cosconati, S., Baldi, A., Niméus, E., Schultz, J., Höglund, U., Maione, S., Papulino, C., Chianese, U., Iovino, F., Federico, A., Mai, A., … Altucci, L. (2019). Inhibition of Histone Demethylases LSD1 and UTX Regulates ERα Signaling in Breast Cancer. *Cancers*, *11*(12), 2027. https://doi.org/10.3390/cancers11122027

Bennesch, M. A., Segala, G., Wider, D., & Picard, D. (2016). LSD1 engages a corepressor complex for the activation of the estrogen receptor α by estrogen and cAMP. *Nucleic Acids Research*, *44*(18), 8655–8670. https://doi.org/10.1093/nar/gkw522

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., & Lander, E. S. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*, *125*(2), 315–326. https://doi.org/10.1016/j.cell.2006.02.041

Biggar, S. R., & Crabtree, G. R. (2001). Cell signaling can direct either binary or graded transcriptional responses. *EMBO Journal*, *20*(12), 3167–3176. https://doi.org/10.1093/emboj/20.12.3167

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W. H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btp101

Bintu, L., Yong, J., Antebi, Y. E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., & Elowitz, M. B. (2016). Dynamics of epigenetic regulation at the single-cell level. In *Science* (Vol. 351, Issue 6274, pp. 720–724). American Association for the Advancement of Science. https://doi.org/10.1126/science.aab2956

Blackledge, N. P., Farcas, A. M., Kondo, T., King, H. W., McGouran, J. F., Hanssen, L. L. P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., Ishikura, T., Long, H. K., Sheahan, T. W., Brockdorff, N., Kessler, B. M., Koseki, H., & Klose, R. J. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell*, *157*(6), 1445–1459. https://doi.org/10.1016/j.cell.2014.05.004

Bohl, S. R., Bullinger, L., & Rücker, F. G. (2018). Epigenetic therapy: azacytidine and decitabine in acute myeloid leukemia. In *Expert Review of Hematology* (Vol. 11, Issue 5, pp. 361–371). Taylor and Francis Ltd. https://doi.org/10.1080/17474086.2018.1453802

Bonhoure, N., Bounova, G., Bernasconi, D., Praz, V., Lammers, F., Canella, D., Willis, I. M., Herr, W., Hernandez, N., Delorenzi, M., Deplancke, B., Desvergne, B., Guex, N., Naef, F., Rougemont, J., Schibler, U., Andersin, T., Cousin, P., Gilardi, F., … Symul, L. (2014). Quantifying ChIP-seq data: A spiking method providing an internal reference for sample-to-sample normalization. *Genome Research*, *24*(7), 1157–1168. https://doi.org/10.1101/gr.168260.113

Boque-Sastre, R., Soler, M., Oliveira-Mateos, C., Portela, A., Moutinho, C., Sayols, S., Villanueva, A., Esteller, M., & Guil, S. (2015). Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1421197112

Bornelöv, S., Reynolds, N., Xenophontos, M., Gharbi, S., Johnstone, E., Floyd, R., Ralser, M., Signolet, J., Loos, R., Dietmann, S., Bertone, P., & Hendrich, B. (2018). The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Molecular Cell*, *71*(1), 56-72.e4. https://doi.org/10.1016/j.molcel.2018.06.003

Bostick, M., Jong, K. K., Estève, P. O., Clark, A., Pradhan, S., & Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science*, *317*(5845), 1760–1764. https://doi.org/10.1126/science.1147939

Braun, S. M. G., Kirkland, J. G., Chory, E. J., Husmann, D., Calarco, J. P., & Crabtree, G. R. (2017). Rapid and reversible epigenome editing by endogenous chromatin regulators. *Nature Communications*, *8*(1), 1–8. https://doi.org/10.1038/s41467-017-00644-y

Brayer, K. J., & Segal, D. J. (2008). Keep your fingers off my DNA: Protein-protein interactions mediated by C2H2 zinc finger domains. In *Cell Biochemistry and Biophysics* (Vol. 50, Issue 3, pp. 111–131). Cell Biochem Biophys. https://doi.org/10.1007/s12013-008-9008-5

Broche, J., Kungulovski, G., Bashtrykov, P., Rathert, P., & Jeltsch, A. (2021). Genome-wide investigation of the dynamic changes of epigenome modifications after global DNA methylation editing. *Nucleic Acids Research*, *49*(1), 158–176. https://doi.org/10.1093/nar/gkaa1169

Brown, R., Curry, E., Magnani, L., Wilhelm-Benartzi, C. S., & Borley, J. (2014). Poised epigenetic states and acquired drug resistance in cancer. *Nature Reviews Cancer*, *14*(11), 747–753. https://doi.org/10.1038/nrc3819

Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., & Allis, C. D. (1996). Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, *84*(6), 843–851. https://doi.org/10.1016/S0092-8674(00)81063-6

Brownell, James E., & Allis, C. D. (1995). An activity gel assay detects a single, catalytically active his tone acetyltransferase subunit in Tetrahymena macronuclei. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(14), 6364–6368. https://doi.org/10.1073/pnas.92.14.6364

Brummelkamp, T. R., Bernards, R., & Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science*, *296*(5567), 550–553. https://doi.org/10.1126/science.1068999

Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., Chang, H. Y., & Greenleaf, W. J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, *523*(7561), 486–490. https://doi.org/10.1038/nature14590

Bulut-Karslioglu, A., Macrae, T. A., Oses-Prieto, J. A., Covarrubias, S., Percharde, M., Ku, G., Diaz, A., McManus, M. T., Burlingame, A. L., & Ramalho-Santos, M. (2018). The Transcriptionally Permissive Chromatin State of Embryonic Stem Cells Is Acutely Tuned to Translational Output. *Cell Stem Cell*, *22*(3), 369-383.e8. https://doi.org/10.1016/j.stem.2018.02.004

Bungard, D., Fuerth, B. J., Zeng, P.-Y., Faubert, B., Maas, N. L., Viollet, B., Carling, D., Thompson, C. B., Jones, R. G., & Berger, S. L. (2010). Signaling Kinase AMPK Activates Stress-Promoted Transcription via Histone H2B Phosphorylation. *Science*, *329*(5996), 1201 LP – 1205. https://doi.org/10.1126/science.1191241

Burg, J. M., Gonzalez, J. J., Maksimchuk, K. R., & McCafferty, D. G. (2016). Lysine-Specific Demethylase 1A (KDM1A/LSD1): Product Recognition and Kinetic Analysis of Full-Length Histones. *Biochemistry*, *55*(11), 1652–1662. https://doi.org/10.1021/acs.biochem.5b01135

Burma, S., Chen, B. P., Murphy, M., Kurimasa, A., & Chen, D. J. (2001). ATM Phosphorylates Histone H2AX in Response to DNA Double-strand Breaks. *Journal of Biological Chemistry*, *276*(45), 42462–42467. https://doi.org/10.1074/jbc.C100466200

Cairns, B. R. (2009). The logic of chromatin architecture and remodelling at promoters. In *Nature* (Vol. 461, Issue 7261, pp. 193–198). Nature Publishing Group. https://doi.org/10.1038/nature08450

Canzio, D., Chang, E. Y., Shankar, S., Kuchenbecker, K. M., Simon, M. D., Madhani, H. D., Narlikar, G. J., & Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Molecular Cell*, *41*(1), 67–81. https://doi.org/10.1016/j.molcel.2010.12.016

Canzio, D., Liao, M., Naber, N., Pate, E., Larson, A., Wu, S., Marina, D. B., Garcia, J. F., Madhani, H. D., Cooke, R., Schuck, P., Cheng, Y., & Narlikar, G. J. (2013). A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature*, *496*(7445), 377–381. https://doi.org/10.1038/nature12032

Carnesecchi, J., Forcet, C., Zhang, L., Tribollet, V., Barenton, B., Boudra, R., Cerutti, C., Billas, I. M. L., Sérandour, A. A., Carroll, J. S., Beaudoin, C., & Vanacker, J. M. (2017). ERRα induces H3K9 demethylation by LSD1 to promote cell invasion. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1614664114

Carrozza, M. J., Utley, R. T., Workman, J. L., & Côté, J. (2003). The diverse functions of histone acetyltransferase complexes. In *Trends in Genetics* (Vol. 19, Issue 6, pp. 321–329). Elsevier Ltd. https://doi.org/10.1016/S0168-9525(03)00115-X

Cavalli, G., & Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. In *Nature* (Vol. 571, Issue 7766, pp. 489–499). Nature Publishing Group. https://doi.org/10.1038/s41586-019-1411-0

Chandru, A., Bate, N., Vuister, G. W., & Cowley, S. M. (2018). Sin3A recruits Tet1 to the PAH1 domain via a highly conserved Sin3-Interaction Domain. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32942-w

Chang, Y., Sun, L., Kokura, K., Horton, J. R., Fukuda, M., Espejo, A., Izumi, V., Koomen, J. M., Bedford, M. T., Zhang, X., Shinkai, Y., Fang, J., & Cheng, X. (2011). MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. *Nature Communications*, *2*(1), 1–10. https://doi.org/10.1038/ncomms1549

Chantalat, S., Depaux, A., Héry, P., Barral, S., Thuret, J. Y., Dimitrov, S., & Gérard, M. (2011). Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. *Genome Research*, *21*(9), 1426–1437. https://doi.org/10.1101/gr.118091.110

Chédin, F. (2011). Chapter 7 - The DNMT3 Family of Mammalian De Novo DNA Methyltransferases. In X. Cheng & R. M. B. T.-P. in M. B. and T. S. Blumenthal (Eds.), *Modifications of Nuclear DNA and its Regulatory Proteins* (Vol. 101, pp. 255–285). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-387685-0.00007-X

Chen, Jiekai, Liu, H., Liu, J., Qi, J., Wei, B., Yang, J., Liang, H., Chen, Y., Chen, J., Wu, Y., Guo, L., Zhu, J., Zhao, X., Peng, T., Zhang, Y., Chen, S., Li, X., Li, D., Wang, T., & Pei, D. (2013). H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature Genetics*, *45*(1), 34–42. https://doi.org/10.1038/ng.2491

Chen, Jingcheng, Wang, Y., Wang, C., Hu, J. F., & Li, W. (2020). LncRNA Functions as a New Emerging Epigenetic Factor in Determining the Fate of Stem Cells. In *Frontiers in Genetics* (Vol. 11, p. 277). Frontiers Media S.A. https://doi.org/10.3389/fgene.2020.00277

Chen, P. B., Chen, H. V., Acharya, D., Rando, O. J., & Fazzio, T. G. (2015). R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.3122

Chen, S., Sanjana, N. E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D. A., Song, J., Pan, J. Q., Weissleder, R., Lee, H., Zhang, F., & Sharp, P. A. (2015). Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell*, *160*(6), 1246–1260. https://doi.org/10.1016/j.cell.2015.02.038

Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics*, *18*(2), 71–86. https://doi.org/10.1038/nrg.2016.139

Ciabrelli, F., Comoglio, F., Fellous, S., Bonev, B., Ninova, M., Szabo, Q., Xuéreb, A., Klopp, C., Aravin, A., Paro, R., Bantignies, F., & Cavalli, G. (2017). Stable Polycomb-dependent transgenerational inheritance of chromatin states in Drosophila. *Nature Genetics*, *49*(6), 876–886. https://doi.org/10.1038/ng.3848

Cibis, H., Biyanee, A., Dörner, W., Mootz, H. D., & Klempnauer, K. H. (2020). Characterization of the zinc finger proteins ZMYM2 and ZMYM4 as novel B-MYB binding proteins. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-65443-w

Ciuffi, A., & Barr, S. D. (2011). Identification of HIV integration sites in infected host genomic DNA. *Methods*, *53*(1), 39–46.

https://doi.org/10.1016/j.ymeth.2010.04.004

Clark, A. J., Harold, G., & Yull, F. E. (1996). Mammalian cDNA and Prokaryotic Reporter Sequences Silence Adjacent Transgenes in Transgenic Mice. *Nucleic Acids Research*, *25*(5), 1009–1014. https://doi.org/10.1093/nar/25.5.1009

Clegg, R. M. (1995). Fluorescence resonance energy transfer. *Current Opinion in Biotechnology*, *6*(1), 103–110. https://doi.org/10.1016/0958-1669(95)80016-6

Cloutier, S. C., Wang, S., Ma, W. K., Al Husini, N., Dhoondia, Z., Ansari, A., Pascuzzi, P. E., & Tran, E. J. (2016). Regulated Formation of lncRNA-DNA Hybrids Enables Faster Transcriptional Induction and Environmental Adaptation. *Molecular Cell*, *61*(3), 393–404. https://doi.org/10.1016/j.molcel.2015.12.024

Collins, R., Karlberg, T., Lehtiö, L., Schütz, P., van den Berg, S., Dahlgren, L. G., Hammarström, M., Weigelt, J., & Schüler, H. (2009). The DEXD/H-box RNA helicase DDX19 is regulated by an α-helical switch. *Journal of Biological Chemistry*, *284*(16), 10296–10300. https://doi.org/10.1074/jbc.C900018200

Connolly, R. M., Li, H., Jankowitz, R. C., Zhang, Z., Rudek, M. A., Jeter, S. C., Slater, S. A., Powers, P., Wolff, A. C., Fetting, J. H., Brufsky, A., Piekarz, R., Ahuja, N., Laird, P. W., Shen, H., Weisenberger, D. J., Cope, L., Herman, J. G., Somlo, G., … Stearns, V. (2017). Combination epigenetic therapy in advanced breast cancer with 5-azacitidine and entinostat: A phase II national cancer institute/stand up to cancer study. *Clinical Cancer Research*, *23*(11), 2691–2701. https://doi.org/10.1158/1078-0432.CCR-16-1729

Cosgrove, M. S., Boeke, J. D., & Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. In *Nature Structural and Molecular Biology* (Vol. 11, Issue 11, pp. 1037–1043). Nature Publishing Group. https://doi.org/10.1038/nsmb851

Côté, J., Quinn, J., Workman, J. L., & Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science*, *265*(5168), 53–60. https://doi.org/10.1126/science.8016655

Cranfill, P. J., Sell, B. R., Baird, M. A., Allen, J. R., Lavagnino, Z., De Gruiter, H. M., Kremers, G. J., Davidson, M. W., Ustione, A., & Piston, D. W. (2016). Quantitative assessment of fluorescent proteins. *Nature Methods*, *13*(7), 557–562. https://doi.org/10.1038/nmeth.3891

Crossley, M. P., Bocek, M., & Cimprich, K. A. (2019). R-Loops as Cellular Regulators and Genomic Threats. In *Molecular Cell*. https://doi.org/10.1016/j.molcel.2019.01.024

Cruz-Tapias, P., Robin, P., Pontis, J., Del Maestro, L., & Ait-Si-Ali, S. (2019). The H3K9 methylation writer SETDB1 and its reader MPP8 cooperate to silence satellite DNA repeats in mouse embryonic stem cells. *Genes*, *10*(10). https://doi.org/10.3390/genes10100750

Culhane, J. C., & Cole, P. A. (2007). LSD1 and the chemistry of histone demethylation. In *Current Opinion in Chemical Biology* (Vol. 11, Issue 5, pp. 561–568). Elsevier Current Trends. https://doi.org/10.1016/j.cbpa.2007.07.014

Cusanovich, D. A., Daza, R., Adey, A., Pliner, H. A., Christiansen, L., Gunderson, K. L., Steemers, F. J., Trapnell, C., & Shendure, J. (2015). Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*, *348*(6237), 910–914. https://doi.org/10.1126/science.aab1601

Cutter, A. R., & Hayes, J. J. (2015). A brief review of nucleosome structure. In *FEBS Letters* (Vol. 589, Issue 20, pp. 2914–2922). Elsevier. https://doi.org/10.1016/j.febslet.2015.05.016

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., & Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, *111*(2), 185–196. https://doi.org/10.1016/S0092-8674(02)00975-3

Daniels, G. A., & Lieber, M. R. (1995). RNA: DNA complex formation upon transcription of immunoglobulin switch regions: Implications for the mechanism and regualtion of class switch recombination. *Nucleic Acids Research*, *23*(24), 5006–5011. https://doi.org/10.1093/nar/23.24.5006

Darst, R. P., Pardo, C. E., Ai, L., Brown, K. D., & Kladde, M. P. (2010). Bisulfite Sequencing of DNA. *Current Protocols in Molecular Biology*, *91*(1), 7.9.1-7.9.17. https://doi.org/10.1002/0471142727.mb0709s91

Davis, R., Peters, D. H., & McTavish, D. (1994). Valproic acid: A reappraisal of its pharmacological properties and clinical efficacy in epilepsy. In *Drugs* (Vol. 47, Issue 2, pp. 332–372). Springer International Publishing. https://doi.org/10.2165/00003495-199447020-00008

Dawlaty, M. M., Breiling, A., Le, T., Barrasa, M. I., Raddatz, G., Gao, Q., Powell, B. E., Cheng, A. W., Faull, K. F., Lyko, F., & Jaenisch, R. (2014). Loss of tet enzymes compromises proper differentiation of embryonic stem cells. *Developmental Cell*, *29*(1), 102–111. https://doi.org/10.1016/j.devcel.2014.03.003

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes and Development*, *25*(10), 1010–1022. https://doi.org/10.1101/gad.2037511

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., & Plath, K. (2013). Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell*, *13*(5), 602–616. https://doi.org/10.1016/j.stem.2013.08.013

Deniz, Ö., Frost, J. M., & Branco, M. R. (2019). Regulation of transposable elements by DNA modifications. In *Nature Reviews Genetics* (Vol. 20, Issue 7, pp. 417–431). Nature Publishing Group. https://doi.org/10.1038/s41576-019-0106-6

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R. Z., Ragozin, S., & Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *Journal of Biological Chemistry*, *285*(34), 26114–26120. https://doi.org/10.1074/jbc.M109.089433

Di Stefano, L., Ji, J. Y., Moon, N. S., Herr, A., & Dyson, N. (2007). Mutation of Drosophila Lsd1 Disrupts H3-K4 Methylation, Resulting in Tissue-Specific Defects during Development. *Current Biology*, *17*(9), 808–812. https://doi.org/10.1016/j.cub.2007.03.068

Dixon, J. R., Gorkin, D. U., & Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. In *Molecular Cell* (Vol. 62, Issue 5, pp. 668–680). Cell Press. https://doi.org/10.1016/j.molcel.2016.05.018

Dobie, K. W., Lee, M., Fantes, J. A., Graham, E., Clark, A. J., Springbett, A., Lathe, R., & McClenaghan, M. (1996). Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(13), 6659–6664. https://doi.org/10.1073/pnas.93.13.6659

Dodge, J. E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., & Li, E. (2005). Inactivation of Dnmt3b in Mouse Embryonic Fibroblasts Results in DNA Hypomethylation, Chromosomal Instability, and Spontaneous Immortalization*. *Journal of Biological Chemistry*, *280*(18), 17986–17991. https://doi.org/https://doi.org/10.1074/jbc.M413246200

Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., & Roeder, R. G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nature Structural & Molecular Biology*, *13*(8), 713–719. https://doi.org/10.1038/nsmb1128

Douse, C. H., Tchasovnikarova, I. A., Timms, R. T., Protasio, A. V., Seczynska, M., Prigozhin, D. M., Albecka, A., Wagstaff, J., Williamson, J. C., Freund, S. M. V., Lehner, P. J., & Modis, Y. (2020). TASOR is a pseudo-PARP that directs HUSH complex assembly and epigenetic transposon control. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-18761-6

Dovey, O. M., Foster, C. T., Conte, N., Edwards, S. A., Edwards, J. M., Singh, R., Vassiliou, G., Bradley, A., & Cowley, S. M. (2013). Histone deacetylase 1 and 2 are essential for normal T-cell development and genomic stability in mice. *Blood*, *121*(8), 1335–1344. https://doi.org/10.1182/blood-2012-07-441949

Drewes, G., & Knapp, S. (2018). Chemoproteomics and Chemical Probes for Target Discovery. In *Trends in Biotechnology* (Vol. 36, Issue 12, pp. 1275–1286). Elsevier Ltd. https://doi.org/10.1016/j.tibtech.2018.06.008

Du, Q., Luu, P. L., Stirzaker, C., & Clark, S. J. (2015). Methyl-CpG-binding domain proteins: Readers of the epigenome. In *Epigenomics* (Vol. 7, Issue 6, pp. 1051–1073). Future Medicine Ltd. https://doi.org/10.2217/epi.15.39

Du, Z., Jia, L., Wang, Y., Wang, C., Wen, X., Chen, J., Zhu, Y., Yu, D., Zhou, L., Chen, N., Zhang, S., Celik, I., Ay, F., Gao, S., Zhang, S., Li, W., Hoffman, A. R., Cui, J., & Hu, J. F. (2018). Combined RNA-seq and RAT-seq mapping of long noncoding RNAs in pluripotent reprogramming. *Scientific Data*, *5*. https://doi.org/10.1038/sdata.2018.255

Duncan, B. K., & Miller, J. H. (1980). Mutagenic deamination of cytosine residues in DNA. *Nature*, *287*(5782), 560–561. https://doi.org/10.1038/287560a0

Dykes, I. M., & Emanueli, C. (2017). Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. *Genomics, Proteomics & Bioinformatics*, *15*(3), 177–186. https://doi.org/10.1016/j.gpb.2016.12.005

Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., Imbeault, M., Rowe, H. M., Turelli, P., & Trono, D. (2016). Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Developmental Cell*, *36*(6), 611–623. https://doi.org/10.1016/j.devcel.2016.02.024

Ecco, G., Imbeault, M., & Trono, D. (2017). KRAB zinc finger proteins. *Development*, *144*(15), 2719–2729. https://doi.org/10.1242/dev.132605

Egan, B., Yuan, C.-C., Craske, M. L., Labhart, P., Guler, G. D., Arnott, D., Maile, T. M., Busby, J., Henry, C., Kelly, T. K., Tindell, C. A., Jhunjhunwala, S., Zhao, F., Hatton, C., Bryant, B. M., Classon, M., & Trojer, P. (2016). An Alternative Approach to ChIP-Seq Normalization Enables Detection of Genome-Wide Changes in Histone H3 Lysine 27 Trimethylation upon EZH2 Inhibition. *PLOS ONE*, *11*(11), e0166438. https://doi.org/10.1371/journal.pone.0166438

Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., Mccune, R. A., & Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Research*, *10*(8), 2709–2721. https://doi.org/10.1093/nar/10.8.2709

Elion, E. A. (2007). Detection of Protein-Protein Interactions by Coprecipitation. *Current Protocols in Protein Science*, *49*(1), 19.4.1-19.4.10. https://doi.org/10.1002/0471140864.ps1904s49

Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M., & Bernstein, B. E. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, *473*(7345), 43–49. https://doi.org/10.1038/nature09906

Estruch, F., & Cole, C. N. (2003). An Early Function during Transcription for the Yeast mRNA Export Factor Dbp5p/Rat8p Suggested by Its Genetic and Physical Interactions with Transcription Factor IIH Components. *Molecular Biology of the Cell*, *14*(4), 1664–1676. https://doi.org/10.1091/mbc.e02-09-0602

Evdokimov, A. G., Pokross, M. E., Egorov, N. S., Zaraisky, A. G., Yampolsky, I. V., Merzlyak, E. M., Shkoporov, A. N., Sander, I., Lukyanov, K. A., & Chudakov, D. M. (2006). Structural basis for the fast maturation of Arthropoda green fluorescent protein. *EMBO Reports*, *7*(10), 1006–1012. https://doi.org/10.1038/sj.embor.7400787

Facciotto, C., Casado, J., Turunen, L., Leivonen, S. K., Tumiati, M., Rantanen, V., Kauppi, L., Lehtonen, R., Leppä, S., Wennerberg, K., & Hautaniemi, S. (2019). Drug screening approach combines epigenetic sensitization with immunochemotherapy in cancer. *Clinical Epigenetics*, *11*(1), 192. https://doi.org/10.1186/s13148-019-0781-3

Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., Asplund, A., Sjöstedt, E., Lundberg, E., Szigyarto, C. A.-K., Skogs, M., Takanen, J. O., Berling, H., Tegel, H., Mulder, J., … Uhlén, M. (2014). Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Molecular & Cellular Proteomics : MCP*, *13*(2), 397–406. https://doi.org/10.1074/mcp.M113.035600

Fang, R., Chen, F., Dong, Z., Hu, D., Barbera, A. J., Clark, E. A., Fang, J., Yang, Y., Mei, P., Rutenberg, M., Li, Z., Zhang, Y., Xu, Y., Yang, H., Wang, P., Simon, M. D., Zhou, Q., Li, J., Marynick, M. P., … Shi, Y. G. (2013). LSD2/KDM1B and Its Cofactor NPAC/GLYR1 Endow a Structural and Molecular Model for Regulation of H3K4 Demethylation. *Molecular Cell*, *49*(3), 558–570. https://doi.org/10.1016/j.molcel.2012.11.019

Fang, X., Wu, Z., Raitskin, O., Webb, K., Voigt, P., Lu, T., Howard, M., & Dean, C. (2020). The 3 processing of antisense RNAs physically links to chromatin-based transcriptional control. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(26), 15316–15321. https://doi.org/10.1073/pnas.2007268117

Fang, Y., Liao, G., & Yu, B. (2019). LSD1/KDM1A inhibitors in clinical trials: Advances and prospects. In *Journal of Hematology and Oncology* (Vol. 12, Issue 1, pp. 1–14). BioMed Central Ltd. https://doi.org/10.1186/s13045-019-0811-9

Faucher, D., & Wellinger, R. J. (2010). Methylated H3K4, a Transcription-Associated Histone Modification, Is Involved in the DNA Damage Response Pathway. *PLOS Genetics*, *6*(8), e1001082. https://doi.org/10.1371/journal.pgen.1001082

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D. Y., Barbosa, I. A. M., Kwon, J. S., Guan, Y., Sinha, N., & Zuber, J. (2013). An optimized microRNA backbone for effective single-copy RNAi. *Cell Reports*, *5*(6), 1704–1713. https://doi.org/10.1016/j.celrep.2013.11.020

Ferrari-Amorotti, G., Fragliasso, V., Esteki, R., Prudente, Z., Soliera, A. R., Cattelani, S., Manzotti, G., Grisendi, G., Dominici, M., Pieraccioli, M., Raschellà, G., Chiodoni, C., Colombo, M. P., & Calabretta, B. (2013). Inhibiting interactions of lysine demethylase LSD1 with snail/slug blocks cancer cell invasion. *Cancer Research*, *73*(1), 235–245. https://doi.org/10.1158/0008-5472.CAN-12-1739

Ferrari, A. C., Alumkal, J. J., Stein, M. N., Taplin, M. E., Babb, J., Barnett, E. S., Gomez-Pinillos, A., Liu, X., Moore, D., DiPaola, R., & Beer, T. M. (2019). Epigenetic therapy with panobinostat combined with bicalutamide rechallenge in castration-resistant prostate cancer. *Clinical Cancer Research*, *25*(1), 52–63. https://doi.org/10.1158/1078-0432.CCR-18-1589

Fico, A., Fiorenzano, A., Pascale, E., Patriarca, E. J., & Minchiotti, G. (2019). Long non-coding RNA in stem cell pluripotency and lineage commitment: functions and evolutionary conservation. In *Cellular and Molecular Life Sciences* (Vol. 76, Issue 8, pp. 1459–1471). Birkhauser Verlag AG. https://doi.org/10.1007/s00018-018-3000-z

Fiering, S., Northrop, J. P., Nolan, G. P., Mattila, P. S., Crabtree, G. R., & Herzenberg, L. A. (1990). Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the T-cell antigen receptor. *Genes and Development*, *4*(10), 1823–1834. https://doi.org/10.1101/gad.4.10.1823

Filippakopoulos, P., & Knapp, S. (2014). Targeting bromodomains: Epigenetic readers of lysine acetylation. In *Nature Reviews Drug Discovery* (Vol. 13, Issue 5, pp. 337–356). Nature Publishing Group. https://doi.org/10.1038/nrd4286

Fiskus, W., Sharma, S., Shah, B., Portier, B. P., Devaraj, S. G. T., Liu, K., Iyer, S. P., Bearss, D., & Bhalla, K. N. (2014). Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. *Leukemia*, *28*(11), 2155–2164. https://doi.org/10.1038/leu.2014.119

Flavahan, W. A., Gaskell, E., & Bernstein, B. E. (2017). Epigenetic plasticity and the hallmarks of cancer. In *Science*. https://doi.org/10.1126/science.aal2380

Fleischer, T. C., Yun, U. J., & Ayer, D. E. (2003). Identification and Characterization of Three New Components of the mSin3A Corepressor Complex. *Molecular and Cellular Biology*, *23*(10), 3456–3467. https://doi.org/10.1128/mcb.23.10.3456-3467.2003

Fliedl, L., Kast, F., Grillari, J., Wieser, M., & Grillari-Voglauer, R. (2015). Optimization of a quantitative PCR based method for plasmid copy number determination in human cell lines. *New Biotechnology*, *32*(6), 716–719. https://doi.org/10.1016/j.nbt.2015.03.004

Forneris, F., Binda, C., Adamo, A., Battaglioli, E., & Mattevi, A. (2007). Structural basis of LSD1-CoREST selectivity in histone H3 recognition. *Journal of Biological Chemistry*, *282*(28), 20070–20074. https://doi.org/10.1074/jbc.C700100200

Forneris, F., Binda, C., Vanoni, M. A., Mattevi, A., & Battaglioli, E. (2005). Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Letters*, *579*(10), 2203–2207. https://doi.org/10.1016/j.febslet.2005.03.015

Foster, C. T., Dovey, O. M., Lezina, L., Luo, J. L., Gant, T. W., Barlev, N., Bradley, A., & Cowley, S. M. (2010). Lysine-Specific Demethylase 1 Regulates the Embryonic Transcriptome and CoREST Stability. *Molecular and Cellular Biology*, *30*(20), 4851–4863. https://doi.org/10.1128/mcb.00521-10

Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gústafsdóttir, S. M., Östman, A., & Landegren, U. (2002). Protein detection using proximity-dependent DNA ligation assays. *Nature Biotechnology*, *20*(5), 473–477. https://doi.org/10.1038/nbt0502-473

Fritsch, L., Robin, P., Mathieu, J. R. R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., & Ait-Si-Ali, S. (2010). A Subset of the Histone H3 Lysine 9 Methyltransferases Suv39h1, G9a, GLP, and SETDB1 Participate in a Multimeric Complex. *Molecular Cell*, *37*(1), 46–56. https://doi.org/10.1016/j.molcel.2009.12.017

Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A., & Mirny, L. A. (2017). Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harbor Symposia on Quantitative Biology*, *82*, 45–55. https://doi.org/10.1101/sqb.2017.82.034710

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports*, *15*(9), 2038–2049. https://doi.org/10.1016/j.celrep.2016.04.085

Fuentes, P., Cánovas, J., Berndt, F. A., Noctor, S. C., & Kukuljan, M. (2012). CoREST/LSD1 Control the Development of Pyramidal Cortical Neurons. *Cerebral Cortex*, *22*(6), 1431–1441. https://doi.org/10.1093/cercor/bhr218

Fujisawa, T., & Filippakopoulos, P. (2017). Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. In *Nature Reviews Molecular Cell Biology* (Vol. 18, Issue 4, pp. 246–262). Nature Publishing Group. https://doi.org/10.1038/nrm.2016.143

Fulton, M. D., Brown, T., & Zheng, Y. G. (2018). Mechanisms and Inhibitors of Histone Arginine Methylation. *The Chemical Record*, *18*(12), 1792–1807. https://doi.org/10.1002/tcr.201800082

Galarneau, L., Nourani, A., Boudreault, A. A., Zhang, Y., Héliot, L., Allard, S., Savard, J., Lane, W. S., Stillman, D. J., & Côté, J. (2000). Multiple links between the NuA4 histone acetyltransferase complex and epigenetic control of transcription. *Molecular Cell*, *5*(6), 927–937. https://doi.org/10.1016/S1097-2765(00)80258-0

Gao, S., Chen, S., Han, D., Wang, Z., Li, M., Han, W., Besschetnova, A., Liu, M., Zhou, F., Barrett, D., Luong, M. P., Owiredu, J., Liang, Y., Ahmed, M., Petricca, J., Patalano, S., Macoska, J. A., Corey, E., Chen, S., … Cai, C. (2020). Chromatin binding of FOXA1 is promoted by LSD1-mediated demethylation in prostate cancer. *Nature Genetics*, *52*(10), 1011–1017. https://doi.org/10.1038/s41588-020-0681-7

Garay, P. M., Wallner, M. A., & Iwase, S. (2016). Yin-yang actions of histone methylation regulatory complexes in the brain. In *Epigenomics* (Vol. 8, Issue 12, pp. 1689–1708). Future Medicine Ltd. https://doi.org/10.2217/epi-2016-0090

Garcia-Bassets, I., Kwon, Y. S., Telese, F., Prefontaine, G. G., Hutt, K. R., Cheng, C. S., Ju, B. G., Ohgi, K. A., Wang, J., Escoubet-Lozach, L., Rose, D. W. W., Glass, C. K., Fu, X. D., & Rosenfeld, M. G. (2007). Histone Methylation-Dependent Mechanisms Impose Ligand Dependency for Gene Activation by Nuclear Receptors. *Cell*. https://doi.org/10.1016/j.cell.2006.12.038

García-Cao, M., O'Sullivan, R., Peters, A. H. F. M., Jenuwein, T., & Blasco, M. A. (2004). Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nature Genetics*, *36*(1), 94–99. https://doi.org/10.1038/ng1278

Garcia, B. A., Hake, S. B., Diaz, R. L., Kauer, M., Morris, S. A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B. D., Allis, C. D., & Hunt, D. F. (2007). Organismal differences in post-translational modifications in histones H3 and H4. *Journal of Biological Chemistry*, *282*(10), 7641–7655.

https://doi.org/10.1074/jbc.M607900200

Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S., & Rosen, M. K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell*, *179*(2), 470-484.e21. https://doi.org/10.1016/j.cell.2019.08.037

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*(2), 442. https://doi.org/10.1016/j.cell.2013.06.044

Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I., & Chédin, F. (2013). GC skew at the 59 and 39 ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Research*, *23*(10), 1590–1600. https://doi.org/10.1101/gr.158436.113

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell*, *45*(6), 814–825. https://doi.org/10.1016/J.MOLCEL.2012.01.017

Glickman, M. H., & Ciechanover, A. (2002). The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiological Reviews*, *82*(2), 373–428. https://doi.org/10.1152/physrev.00027.2001

Gökbuget, D., & Blelloch, R. (2019). Epigenetic control of transcriptional regulation in pluripotency and early differentiation. In *Development (Cambridge)* (Vol. 146, Issue 19). Company of Biologists Ltd. https://doi.org/10.1242/dev.164772

Gonzalez, I., Munita, R., Agirre, E., Dittmer, T. A., Gysling, K., Misteli, T., & Luco, R. F. (2015). A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature. *Nature Structural and Molecular Biology*, *22*(5), 370–376. https://doi.org/10.1038/nsmb.3005

Goren, A., Tabib, A., Hecht, M., & Cedar, H. (2008). DNA replication timing of the human β-globin domain is controlled by histone modification at the origin. *Genes & Development*, *22*(10), 1319–1324. https://doi.org/10.1101/gad.468308

Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science*, *268*(5218), 1766–1769. https://doi.org/10.1126/science.7792603

Gowher, H., Liebert, K., Hermann, A., Xu, G., & Jeltsch, A. (2005). Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *Journal of Biological Chemistry*, *280*(14), 13341–13348. https://doi.org/10.1074/jbc.M413412200

Grav, L. M., Sergeeva, D., Lee, J. S., Marin De Mas, I., Lewis, N. E., Andersen, M. R., Nielsen, L. K., Lee, G. M., & Kildegaard, H. F. (2018). Minimizing Clonal Variation during Mammalian Cell Line Engineering for Improved Systems Biology Data Generation. *ACS Synthetic Biology*, *7*(9), 2148–2159. https://doi.org/10.1021/acssynbio.8b00140

Greenberg, M. V. C., & Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology*, *20*(10), 590–607. https://doi.org/10.1038/s41580-019-0159-6

Gresakova, V., Novosadova, V., Prochazkova, M., Bhargava, S., Jenickova, I., Prochazka, J., & Sedlacek, R. (2019). Fam208a orchestrates interaction protein network essential for early embryonic development and cell division. *Experimental Cell Research*, *382*(1), 111437. https://doi.org/10.1016/j.yexcr.2019.05.018

Grewal, S. I. S., & Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. In *Science* (Vol. 301, Issue 5634, pp. 798–802). American Association for the Advancement of Science. https://doi.org/10.1126/science.1086887

Grigoryev, S. A. (2004). Keeping fingers crossed: heterochromatin spreading through interdigitation of nucleosome arrays. *FEBS Letters*, *564*(1–2), 4–8. https://doi.org/10.1016/S0014-5793(04)00258-3

Grigoryev, S. A., & Woodcock, C. L. (2012). Chromatin organization - The 30nm fiber. In *Experimental Cell Research* (Vol. 318, Issue 12, pp. 1448–1455). Academic Press Inc. https://doi.org/10.1016/j.yexcr.2012.02.014

Grosch, M., Ittermann, S., Shaposhnikov, D., & Drukker, M. (2020). Chromatin-Associated Membraneless Organelles in Regulation of Cellular Differentiation. In *Stem Cell Reports* (Vol. 15, Issue 6, pp. 1220–1232). Cell Press. https://doi.org/10.1016/j.stemcr.2020.10.011

Groth, A., Rocha, W., Verreault, A., & Almouzni, G. (2007). Chromatin Challenges during DNA Replication and Repair. In *Cell* (Vol. 128, Issue 4, pp. 721–733). Cell Press. https://doi.org/10.1016/j.cell.2007.01.030

Grüner, B. M., Schulze, C. J., Yang, D., Ogasawara, D., Dix, M. M., Rogers, Z. N., Chuang, C. H., McFarland, C. D., Chiou, S. H., Brown, J. M., Cravatt, B. F., Bogyo, M., & Winslow, M. M. (2016). An in vivo multiplexed small-molecule screening platform. *Nature Methods*, *13*(10), 883–889. https://doi.org/10.1038/nmeth.3992

Gu, F., Lin, Y., Wang, Z., Wu, X., Ye, Z., Wang, Y., & Lan, H. (2020). Biological roles of LSD1 beyond its demethylase activity. In *Cellular and Molecular Life Sciences* (Vol. 77, Issue 17, pp. 3341–3350). Springer. https://doi.org/10.1007/s00018-020-03489-9

Gu, X., Tohme, R., Tomlinson, B., Sakre, N., Hasipek, M., Durkin, L., Schuerger, C., Grabowski, D., Zidan, A. M., Radivoyevitch, T., Hong, C., Carraway, H., Hamilton, B., Sobecks, R., Patel, B., Jha, B. K., Hsi, E. D., Maciejewski, J., & Saunthararajah, Y. (2020). Decitabine- and 5-azacytidine resistance emerges from adaptive responses of the pyrimidine metabolism network. *Leukemia*, 1–14. https://doi.org/10.1038/s41375-020-1003-x

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W., & van Steensel, B. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, *453*(7197), 948–951. https://doi.org/10.1038/nature06947

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., & Young, R. A. (2007). A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell*, *130*(1), 77–88. https://doi.org/10.1016/j.cell.2007.05.042

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., Tian, C., Wang, J., Cong, Y., & Xu, Y. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature*, *517*(7536), 640–644. https://doi.org/10.1038/nature13899

Guo, Y., Zhao, S., & Wang, G. G. (2021). Polycomb Gene Silencing Mechanisms: PRC2 Chromatin Targeting, H3K27me3 "Readout", and Phase Separation-Based Compaction. *Trends in Genetics*. https://doi.org/10.1016/j.tig.2020.12.006

Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. In *Nature* (Vol. 482, Issue 7385, pp. 339–346). Nature. https://doi.org/10.1038/nature10887

Hakimi, M. A., Bochar, D. A., Chenoweth, J., Lane, W. S., Mandel, G., & Shiekhattar, R. (2002). A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(11), 7420–7425. https://doi.org/10.1073/pnas.112008599

Han, P., & Chang, C.-P. (2015). Long non-coding RNA and chromatin remodeling. *RNA Biology*, *12*(10), 1094–1098. https://doi.org/10.1080/15476286.2015.1063770

Hansen, K. H., Bracken, A. P., Pasini, D., Dietrich, N., Gehani, S. S., Monrad, A., Rappsilber, J., Lerdrup, M., & Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. *Nature Cell Biology*, *10*(11), 1291–1300. https://doi.org/10.1038/ncb1787

Hansen, R. S., Wijmenga, C., Luo, P., Stanek, A. M., Canfield, T. K., Weemaes, C. M. R., & Gartler, S. M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proceedings of the National Academy of Sciences*, *96*(25), 14412 LP – 14417. https://doi.org/10.1073/pnas.96.25.14412

Hardie, D. G. (2007). AMP-activated/SNF1 protein kinases: Conserved guardians of cellular energy. In *Nature Reviews Molecular Cell Biology* (Vol. 8, Issue 10, pp. 774–785). Nature Publishing Group. https://doi.org/10.1038/nrm2249

Hargreaves, D. C., & Crabtree, G. R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Research*, *21*(3), 396–420. https://doi.org/10.1038/cr.2011.32

Harris, W. J., Huang, X., Lynch, J. T., Spencer, G. J., Hitchin, J. R., Li, Y., Ciceri, F., Blaser, J. G., Greystoke, B. F., Jordan, A. M., Miller, C. J., Ogilvie, D. J., & Somervaille, T. C. P. (2012). The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell*, *21*(4), 473–487. https://doi.org/10.1016/j.ccr.2012.03.014

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B. L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., … Hubbard, T. J. (2012). GENCODE: The reference human genome annotation for the ENCODE project. *Genome Research*, *22*(9), 1760–1774. https://doi.org/10.1101/gr.135350.111

Hartono, S. R., Korf, I. F., & Chédin, F. (2015). GC skew is a conserved property of unmethylated CpG island promoters across vertebrates. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkv811

Hathaway, N. A. A., Bell, O., Hodges, C., Miller, E. L. L., Neel, D. S. S., & Crabtree, G. R. R. (2012). Dynamics and memory of heterochromatin in living cells. *Cell*, *149*(7), 1447–1460. https://doi.org/10.1016/j.cell.2012.03.052

Headley, K. M., Kedziora, K. M., Alejo, A., Lai, E. Z.-X., Purvis, J. E., & Hathaway, N. A. (2019). Chemical screen for epigenetic barriers to single allele activation of Oct4. *Stem Cell Research*, *38*, 101470. https://doi.org/10.1016/j.scr.2019.101470

Hermann, A., Gowher, H., & Jeltsch, A. (2004). Biochemistry and biology of mammalian DNA methyltransferases. In *Cellular and Molecular Life Sciences* (Vol. 61, Issues 19–20, pp. 2571–2587). Springer. https://doi.org/10.1007/s00018-004-4201-1

Herz, H. M., Garruss, A., & Shilatifard, A. (2013). SET for life: Biochemical activities and biological functions of SET domain-containing proteins. In *Trends in Biochemical Sciences* (Vol. 38, Issue 12, pp. 621–639). Elsevier. https://doi.org/10.1016/j.tibs.2013.09.004

Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*, *33*(5), 510–517. https://doi.org/10.1038/nbt.3199

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., & Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9362–9367. https://doi.org/10.1073/pnas.0903103106

Hirai, H., Tani, T., & Kikyo, N. (2011). Structure and functions of powerful transactivators: VP16, MyoD and FoxA. In *International Journal of Developmental Biology* (Vol. 54, Issues 11–12, pp. 1589–1596). NIH Public Access. https://doi.org/10.1387/ijdb.103194hh

Hirschi, A., Martin, W. J., Luka, Z., Loukachevitch, L. V., & Reiter, N. J. (2016). G-quadruplex RNA binding and recognition by the lysine-specific histone demethylase-1 enzyme. *RNA, 22*(8), 1250–1260. https://doi.org/10.1261/rna.057265.116

Ho, S. N., Biggar, S. R., Spencer, D. M., Schreiber, S. L., & Crabtree, G. R. (1996). Dimeric ligands define a role for transcriptional activation domains in reinitiation. *Nature, 382*(6594), 822–826. https://doi.org/10.1038/382822a0

Hodges, C., Kirkland, J. G., & Crabtree, G. R. (2016). The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harbor Perspectives in Medicine, 6*(8), a026930. https://doi.org/10.1101/cshperspect.a026930

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Abou Merhi, R., & Maiorano, D. (2017). An <scp>ATR</scp> -dependent function for the Ddx19 <scp>RNA</scp> helicase in nuclear R-loop metabolism. *The EMBO Journal, 36*(9), 1182–1198. https://doi.org/10.15252/embj.201695131

Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F., Lin, R., Smith, M. M., & Allis, C. D. (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell, 102*(3), 279–291. https://doi.org/10.1016/S0092-8674(00)00034-9

Hu, D., Garruss, A. S., Gao, X., Morgan, M. A., Cook, M., Smith, E. R., & Shilatifard, A. (2013). The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. *Nature Structural and Molecular Biology, 20*(9), 1093–1097. https://doi.org/10.1038/nsmb.2653

Hu, X., Li, X., Valverde, K., Fu, X., Noguchi, C., Qiu, Y., & Huang, S. (2009). LSD1-mediated epigenetic modification is required for TAL1 function and hematopoiesis. *Proceedings of the National Academy of Sciences of the United States of America, 106*(25), 10141–10146. https://doi.org/10.1073/pnas.0900437106

Hua, H., Zhang, H., Kong, Q., & Jiang, Y. (2018). Mechanisms for estrogen receptor expression in human cancer. In *Experimental Hematology and Oncology* (Vol. 7, Issue 1, p. 24). BioMed Central Ltd. https://doi.org/10.1186/s40164-018-0116-7

Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., & Xu, R. M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science, 312*(5774), 748–751. https://doi.org/10.1126/science.1125162

Humphrey, G. W., Wang, Y., Russanova, V. R., Hirai, T., Qin, J., Nakatani, Y., & Howard, B. H. (2001). Stable Histone Deacetylase Complexes Distinguished by the Presence of SANT Domain Proteins CoREST/kiaa0071 and Mta-L1. *Journal of Biological Chemistry, 276*(9), 6817–6824. https://doi.org/10.1074/jbc.M007372200

Ibrahim, D. M., & Mundlos, S. (2020). The role of 3D chromatin domains in gene regulation: a multi-facetted view on genome organization. *Current Opinion in Genetics & Development, 61*, 1–8. https://doi.org/10.1016/j.gde.2020.02.015

Imbeault, M., Helleboid, P. Y., & Trono, D. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature*, *543*(7646), 550–554. https://doi.org/10.1038/nature21683

Jacob, M. D., Audas, T. E., Uniacke, J., Trinkle-Mulcahy, L., & Lee, S. (2013). Environmental cues induce a long noncoding RNA-dependent remodeling of the nucleolus. *Molecular Biology of the Cell*, *24*(18), 2943–2953. https://doi.org/10.1091/mbc.E13-04-0223

Jacobs, F. M. J., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A. D., Katzman, S., Paten, B., Salama, S. R., & Haussler, D. (2014). An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, *516*(7530), 242–245. https://doi.org/10.1038/nature13760

Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. In *Nature Genetics* (Vol. 33, Issue 3S, pp. 245–254). Nature Publishing Group. https://doi.org/10.1038/ng1089

Jakovcevski, M., & Akbarian, S. (2012). Epigenetic mechanisms in neurological disease. *Nature Medicine*, *18*(8), 1194–1204. https://doi.org/10.1038/nm.2828

Jeffries, M. A. (2018). Epigenetic editing: How cutting-edge targeted epigenetic modification might provide novel avenues for autoimmune disease therapy. *Clinical Immunology*, *196*, 49–58. https://doi.org/10.1016/j.clim.2018.02.001

Jeltsch, A. (2008). Reading and writing DNA methylation. In *Nature Structural and Molecular Biology* (Vol. 15, Issue 10, pp. 1003–1004). Nature Publishing Group. https://doi.org/10.1038/nsmb1008-1003

Jeltsch, A., Broche, J., Lungu, C., & Bashtrykov, P. (2020). Biotechnological Applications of MBD Domain Proteins for DNA Methylation Analysis. In *Journal of Molecular Biology* (Vol. 432, Issue 6, pp. 1816–1823). Academic Press. https://doi.org/10.1016/j.jmb.2019.08.020

Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. In *Science* (Vol. 293, Issue 5532, pp. 1074–1080). American Association for the Advancement of Science. https://doi.org/10.1126/science.1063127

Jenuwein, Thomas. (2001). Re-SET-ting heterochromatin by histone methyltransferases. In *Trends in Cell Biology* (Vol. 11, Issue 6, pp. 266–273). Elsevier Ltd. https://doi.org/10.1016/S0962-8924(01)02001-3

Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., & Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, *449*(7159), 248–251. https://doi.org/10.1038/nature06146

Jiang, H. (2020). The complex activities of the SET1/MLL complex core subunits in development and disease. In *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* (Vol. 1863, Issue 7, p. 194560). Elsevier B.V. https://doi.org/10.1016/j.bbagrm.2020.194560

Jiao, L., & Liu, X. (2015). Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science*, *350*(6258). https://doi.org/10.1126/science.aac4383

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., & Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark "nucleosome-free regions" of active promoters and other regulatory regions. *Nature Genetics*, *41*(8), 941–945. https://doi.org/10.1038/ng.409

Jin, Y., Huo, B., Fu, X., Hao, T., Zhang, Y., Guo, Y., & Hu, X. (2017). LSD1 collaborates with EZH2 to regulate expression of interferon-stimulated genes. *Biomedicine and Pharmacotherapy*, *88*, 728–737. https://doi.org/10.1016/j.biopha.2017.01.055

Jones, L. H., & Bunnage, M. E. (2017). Applications of chemogenomic library screening in drug discovery. In *Nature Reviews Drug Discovery* (Vol. 16, Issue 4, pp. 285–296). Nature Publishing Group. https://doi.org/10.1038/nrd.2016.244

Jones, P. A., Issa, J.-. P. J., & Baylin, S. (2016). Targeting the cancer epigenome for therapy. *Nat Rev Genet*, *15*.

Jonkers, I., Kwak, H., & Lis, J. T. (2014). Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *ELife*, *2014*(3). https://doi.org/10.7554/eLife.02407

Kallunki, Barisic, Jäättelä, & Liu. (2019). How to Choose the Right Inducible Gene Expression System for Mammalian Studies? *Cells*, *8*(8), 796. https://doi.org/10.3390/cells8080796

Kamionka, A., Bogdanska-Urbaniak, J., Scholz, O., & Hillen, W. (2004). Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor. *Nucleic Acids Research*, *32*(2), 842–847. https://doi.org/10.1093/nar/gkh200

Kaneko, S., Son, J., Shen, S. S., Reinberg, D., & Bonasio, R. (2013). PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nature Structural and Molecular Biology*, *20*(11), 1258–1264. https://doi.org/10.1038/nsmb.2700

Karagianni, P., & Tzioufas, A. G. (2019). Epigenetic perspectives on systemic autoimmune disease. In *Journal of Autoimmunity* (Vol. 104, p. 102315). Academic Press. https://doi.org/10.1016/j.jaut.2019.102315

Karpf, A. R., & Matsui, S. (2005). Genetic Disruption of Cytosine DNA Methyltransferase Enzymes Induces Chromosomal Instability in Human Cancer Cells. *Cancer Research*, *65*(19), 8635–8639. https://doi.org/10.1158/0008-5472.CAN-05-1961

Kauffman, E. C., Robinson, B. D., Downes, M. J., Powell, L. G., Lee, M. M., Scherr, D. S., Gudas, L. J., & Mongan, N. P. (2011). Role of androgen receptor and associated lysine-demethylase coregulators, LSD1 and JMJD2A, in localized and advanced human bladder cancer. *Molecular Carcinogenesis*, *50*(12), 931–944. https://doi.org/10.1002/mc.20758

Kelly, R. D. W., & Cowley, S. M. (2013). The physiological roles of histone deacetylase (HDAC) 1 and 2: Complex co-stars with multiple leading parts. *Biochemical Society Transactions*, *41*(3), 741–749. https://doi.org/10.1042/BST20130010

Kerppola, T. K. (2006). Visualization of molecular interactions by fluorescence complementation. In *Nature Reviews Molecular Cell Biology* (Vol. 7, Issue 6, pp. 449–456). Nature Publishing Group. https://doi.org/10.1038/nrm1929

Keung, A. J., Bashor, C. J., Kiriakov, S., Collins, J. J., & Khalil, A. S. (2014). Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell*, *158*(1), 110–120. https://doi.org/10.1016/j.cell.2014.04.047

Kim, J. H., Lee, B. B., Oh, Y. M., Zhu, C., Steinmetz, L. M., Lee, Y., Kim, W. K., Lee, S. B., Buratowski, S., & Kim, T. S. (2016). Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. *Nature Communications*, *7*(1), 1–11. https://doi.org/10.1038/ncomms13534

Kim, M. S., Kim, Y. R., Yoo, N. J., & Lee, S. H. (2013). Mutational analysis of DNMT3A gene in acute leukemias and common solid cancers. *APMIS*, *121*(2), 85–94. https://doi.org/10.1111/j.1600-0463.2012.02940.x

Kim, S.-A., Zhu, J., Yennawar, N., Eek, P., & Tan, S. (2020). Crystal Structure of the LSD1/CoREST Histone Demethylase Bound to Its Nucleosome Substrate. *Molecular Cell*, *78*(5), 903-914.e4. https://doi.org/10.1016/j.molcel.2020.04.019

Kirkpatrick, R. L., Lewis, K., Langan, R. A., Lajoie, M. J., Boyken, S. E., Eakman, M., Baker, D., & Zalatan, J. G. (2020). Conditional Recruitment to a DNA-Bound CRISPR-Cas Complex Using a Colocalization-Dependent Protein Switch. *ACS Synthetic Biology*, *9*(9), 2316–2323. https://doi.org/10.1021/acssynbio.0c00012

Klose, R. J., Kallin, E. M., & Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics*, *7*(9), 715–727. https://doi.org/10.1038/nrg1945

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, *502*(7472), 472–479. https://doi.org/10.1038/nature12750

Kolla, V., Naraparaju, K., Zhuang, T., Higashi, M., Kolla, S., Blobel, G. A., & Brodeur, G. M. (2015). The tumour suppressor CHD5 forms a NuRD-type chromatin remodelling complex. *Biochemical Journal*, *468*(2), 345–352. https://doi.org/10.1042/BJ20150030

Komashko, V. M., & Farnham, P. J. (2010). 5-azacytidine treatment reorganizes genomic histone modification patterns. *Epigenetics*, *5*(3), 229–240. https://doi.org/10.4161/epi.5.3.11409

Kornberg, R. D., & Thomas, J. O. (1974). Chromatin Structure: Oligomers of the Histones. *Science*, *184*(4139), 865–868. http://www.jstor.org/stable/1738169

Kouzarides, T. (2002). Histone methylation in transcriptional control. In *Current Opinion in Genetics and Development* (Vol. 12, Issue 2, pp. 198–209). Elsevier Ltd. https://doi.org/10.1016/S0959-437X(02)00287-3

Kribelbauer, J. F., Laptenko, O., Chen, S., Martini, G. D., Freed-Pastor, W. A., Prives, C., Mann, R. S., & Bussemaker, H. J. (2017). Quantitative Analysis of the DNA Methylation Sensitivity of Transcription Factor Complexes. *Cell Reports*, *19*(11), 2383–2395. https://doi.org/10.1016/j.celrep.2017.05.069

Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., Greenblatt, J. F., & Shilatifard, A. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: Linking transcriptional elongation to histone methylation. *Molecular Cell*, *11*(3), 721–729. https://doi.org/10.1016/S1097-2765(03)00091-1

Kungulovski, G., Mauser, R., Reinhardt, R., & Jeltsch, A. (2016). Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody. *Epigenetics and Chromatin*, *9*(1), 11. https://doi.org/10.1186/s13072-016-0061-9

Kuo, Min-Hao; Zhou, Jianxin; Jambeck, Per; Churchill, Mair E.A.; Allis, C. D. (1997). *Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo*. Genes and Development. http://genesdev.cshlp.org/content/12/5/627.short

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., & Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes and Development*, *16*(22), 2893–2905. https://doi.org/10.1101/gad.1035902

Landgrave-Gómez, J., Mercado-Gómez, O., & Guevara-Guzmán, R. (2015). Epigenetic mechanisms in neurological and neurodegenerative diseases. *Frontiers in Cellular Neuroscience*, *9*(FEB), 58. https://doi.org/10.3389/fncel.2015.00058

Lange, M., Demajo, S., Jain, P., & Di Croce, L. (2011). Combinatorial assembly and function of chromatin regulatory complexes. In *Epigenomics* (Vol. 3, Issue 5, pp. 567–580). https://doi.org/10.2217/epi.11.83

Längst, G., & Manelyte, L. (2015). Chromatin Remodelers: From Function to Dysfunction. *Genes*, *6*(2), 299–324. https://doi.org/10.3390/genes6020299

Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., & Narlikar, G. J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*, *547*(7662), 236–240. https://doi.org/10.1038/nature22822

Laugesen, A., Højfeldt, J. W., & Helin, K. (2019). Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. In *Molecular Cell* (Vol. 74, Issue 1, pp. 8–18). Cell Press. https://doi.org/10.1016/j.molcel.2019.03.011

Laurent, B., Ruitu, L., Murn, J., Hempel, K., Ferrao, R., Xiang, Y., Liu, S., Garcia, B. A., Wu, H., Wu, F., Steen, H., & Shi, Y. (2015). A Specific LSD1/KDM1A Isoform Regulates Neuronal Differentiation through H3K9 Demethylation. *Molecular Cell*, *57*(6), 957–970. https://doi.org/10.1016/j.molcel.2015.01.010

Lee, C. Y., McNerney, C., Ma, K., Zhao, W., Wang, A., & Myong, S. (2020). R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17176-7

Lee, J., Thompson, J. R., Botuyan, M. V., & Mer, G. (2008). Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. *Nature Structural & Molecular Biology*, *15*(1), 109–111. https://doi.org/10.1038/nsmb1326

Lee, Jae Seong, Park, J. H., Ha, T. K., Samoudi, M., Lewis, N. E., Palsson, B. O., Kildegaard, H. F., & Lee, G. M. (2018). Revealing Key Determinants of Clonal Variation in Transgene Expression in Recombinant CHO Cells Using Targeted Genome Editing. *ACS Synthetic Biology*, *7*(12), 2867–2878. https://doi.org/10.1021/acssynbio.8b00290

Lee, Jung Shin, Shukla, A., Schneider, J., Swanson, S. K., Washburn, M. P., Florens, L., Bhaumik, S. R., & Shilatifard, A. (2007). Histone Crosstalk between H2B Monoubiquitination and H3 Methylation Mediated by COMPASS. *Cell*, *131*(6), 1084–1096. https://doi.org/10.1016/j.cell.2007.09.046

Lee, M. G., Wynder, C., Cooch, N., & Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature*, *437*(7057), 432–435. https://doi.org/10.1038/nature04021

Lee, N. C., Kononenko, A. V., Lee, H. S., Tolkunova, E. N., Liskovykh, M. A., Masumoto, H., Earnshaw, W. C., Tomilin, A. N., Larionov, V., & Kouprina, N. (2013). Protecting a transgene expression from the HAC-based vector by different chromatin insulators. *Cellular and Molecular Life Sciences*, *70*(19), 3723–3737. https://doi.org/10.1007/s00018-013-1362-9

Leung, Y. T., Shi, L., Maurer, K., Song, L., Zhang, Z., Petri, M., & Sullivan, K. E. (2015). Interferon regulatory factor 1 and histone H4 acetylation in systemic lupus erythematosus. *Epigenetics*, *10*(3), 191–199. https://doi.org/10.1080/15592294.2015.1009764

Li, B., Carey, M., & Workman, J. L. (2007). The Role of Chromatin during Transcription. In *Cell* (Vol. 128, Issue 4, pp. 707–719). Cell Press. https://doi.org/10.1016/j.cell.2007.01.015

Li, B. E., & Ernst, P. (2014). Two decades of leukemia oncoprotein epistasis: The MLL1 paradigm for epigenetic deregulation in leukemia. In *Experimental Hematology* (Vol. 42, Issue 12, pp. 995–1012). Elsevier Inc. https://doi.org/10.1016/j.exphem.2014.09.006

Li, F., Martienssen, R., & Cande, W. Z. (2011). Coordination of DNA replication and histone modification by the Rik1–Dos2 complex. *Nature*, *475*(7355), 244–248. https://doi.org/10.1038/nature10161

Li, J. J., Bickel, P. J., & Biggin, M. D. (2014). System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ*, *2014*(1), 1–26. https://doi.org/10.7717/peerj.270

Li, J. J., & Biggin, M. D. (2015). Statistics requantitates the central dogma. *Science*, *347*(6226), 1066 LP – 1067. https://doi.org/10.1126/science.aaa8332

Li, J., Moazed, D., & Gygi, S. P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *Journal of Biological Chemistry*, *277*(51), 49383–49388. https://doi.org/10.1074/jbc.M209294200

Lian, Y., Li, Z., Fan, Y., Huang, Q., Chen, J., Liu, W., Xiao, C. X., & Xu, H. Z. (2017). The lncRNA-HOXA-AS2/EZH2/LSD1 oncogene complex promotes cell proliferation in pancreatic cancer. *American Journal of Translational Research*, *9*(12), 5496–5506. www.ajtr.org

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293. https://doi.org/10.1126/science.1181369

Lienert, F., Lohmueller, J. J., Garg, A., & Silver, P. A. (2014). Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nature Reviews Molecular Cell Biology*, *15*(2), 95–107. https://doi.org/10.1038/nrm3738

Lim, S., Janzer, A., Becker, A., Zimmer, A., Schüle, R., Buettner, R., & Kirfel, J. (2010). Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. *Carcinogenesis*, *31*(3), 512–520. https://doi.org/10.1093/carcin/bgp324

Lin, W., & Dent, S. Y. R. (2006). Functions of histone-modifying enzymes in development. In *Current Opinion in Genetics and Development* (Vol. 16, Issue 2, pp. 137–142). Elsevier Current Trends. https://doi.org/10.1016/j.gde.2006.02.002

Lin, Y., Wu, Y., Li, J., Dong, C., Ye, X., Chi, Y. I., Evers, B. M., & Zhou, B. P. (2010). The SNAG domain of snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. *EMBO Journal*, *29*(11), 1803–1816. https://doi.org/10.1038/emboj.2010.63

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping: the DEAD box RNA helicase family. In *Nature Reviews Molecular Cell Biology* (Vol. 12, Issue 8, pp. 505–516). https://doi.org/10.1038/nrm3154

Linhares, B. M., Grembecka, J., & Cierpicki, T. (2020). Targeting epigenetic protein-protein interactions with small-molecule inhibitors. In *Future Medicinal Chemistry* (Vol. 12, Issue 14, pp. 1305–1326). Future Medicine Ltd. https://doi.org/10.4155/fmc-2020-0082

Lio, C. W. J., & Rao, A. (2019). TET enzymes and 5hMC in adaptive and innate immune systems. In *Frontiers in Immunology* (Vol. 10, Issue FEB, p. 210). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2019.00210

Liu, S. J., Horlbeck, M. A., Cho, S. W., Birk, H. S., Malatesta, M., He, D., Attenello, F. J., Villalta, J. E., Cho, M. Y., Chen, Y., Mandegar, M. A., Olvera, M. P., Gilbert, L.

A., Conklin, B. R., Chang, H. Y., Weissman, J. S., & Lim, D. A. (2017). CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, *355*(6320), eaah7111. https://doi.org/10.1126/science.aah7111

Liu, Y., Aryee, M. J., Padyukov, L., Daniele Fallin, M., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., Shchetynsky, K., Scheynius, A., Kere, J., Alfredsson, L., Klareskog, L., Ekstr, T. J., & Feinberg, A. P. (2013). Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nature Biotechnology*. https://doi.org/10.1038/nbt.2487

Loew, R., Heinz, N., Hampf, M., Bujard, H., & Gossen, M. (2010). Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnology*, *10*(1), 81. https://doi.org/10.1186/1472-6750-10-81

Lohse, B., Helgstrand, C., Kristensen, J. B. L., Leurs, U., Cloos, P. A. C., Kristensen, J. L., & Clausen, R. P. (2013). Posttranslational Modifications of the Histone 3 Tail and Their Impact on the Activity of Histone Lysine Demethylases In Vitro. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0067653

Long, H. K., Prescott, S. L., & Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. In *Cell* (Vol. 167, Issue 5, pp. 1170–1187). Cell Press. https://doi.org/10.1016/j.cell.2016.09.018

Long, Y., Hwang, T., Gooding, A. R., Goodrich, K. J., Rinn, J. L., & Cech, T. R. (2020). RNA is essential for PRC2 chromatin occupancy and function in human pluripotent stem cells. *Nature Genetics*, *52*(9), 931–938. https://doi.org/10.1038/s41588-020-0662-x

Loppin, B., & Berger, F. (2020). Histone Variants: The Nexus of Developmental Decisions and Epigenetic Memory. *Annual Review of Genetics*, *54*(1). https://doi.org/10.1146/annurev-genet-022620-100039

Lungu, C., Pinter, S., Broche, J., Rathert, P., & Jeltsch, A. (2017). Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-00457-z

Luo, J., & Chen, R. (2020). Genetic variations associated with long noncoding RNAs. *Essays in Biochemistry*, *64*(6), 867–873. https://doi.org/10.1042/EBC20200033

Ma, W. K., Paudel, B. P., Xing, Z., Sabath, I. G., Rueda, D., & Tran, E. J. (2016). Recruitment, Duplex Unwinding and Protein-Mediated Inhibition of the Dead-Box RNA Helicase Dbp2 at Actively Transcribed Chromatin. *Journal of Molecular Biology*, *428*(6), 1091–1106. https://doi.org/10.1016/j.jmb.2016.02.005

Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J. ichi, Wolf, M., & Kurumizaka, H. (2018). Structural Basis of Heterochromatin Formation by Human HP1. *Molecular Cell*, *69*(3), 385-397.e8. https://doi.org/10.1016/j.molcel.2017.12.011

Maes, T., Mascaró, C., Rotllant, D., Cavalcanti, F., Carceller, E., Ortega, A., Molinero, C., & Buesa, C. (2016). P4-404: Ory-2001: An Epigenetic Drug for the Treatment of Cognition Defects in Alzheimer'S Disease and Other Neurodegenerative Disorders. *Alzheimer's & Dementia*, *12*, P1192–P1192. https://doi.org/10.1016/j.jalz.2016.07.149

Maes, T., Mascaró, C., Tirapu, I., Estiarte, A., Ciceri, F., Lunardi, S., Guibourt, N., Perdones, A., Lufino, M. M. P., Somervaille, T. C. P., Wiseman, D. H., Duy, C., Melnick, A., Willekens, C., Ortega, A., Martinell, M., Valls, N., Kurz, G., Fyfe, M., … Buesa, C. (2018). ORY-1001, a Potent and Selective Covalent KDM1A Inhibitor, for the Treatment of Acute Leukemia. *Cancer Cell*, *33*(3), 495-511.e12. https://doi.org/10.1016/J.CCELL.2018.02.002

Maeshima, K., Tamura, S., Hansen, J. C., & Itoh, Y. (2020). Fluid-like chromatin: Toward understanding the real chromatin organization present in the cell. *Current Opinion in Cell Biology*, *64*, 77–89. https://doi.org/10.1016/j.ceb.2020.02.016

Maiques-Diaz, A., & Somervaille, T. C. (2016). LSD1: biologic roles and therapeutic targeting. *Epigenomics*, *8*(8), 1103–1116. https://doi.org/10.2217/epi-2016-0009

Maiques-Diaz, A., Spencer, G. J., Lynch, J. T., Ciceri, F., Williams, E. L., Amaral, F. M. R., Wiseman, D. H., Harris, W. J., Li, Y., Sahoo, S., Hitchin, J. R., Mould, D. P., Fairweather, E. E., Waszkowycz, B., Jordan, A. M., Smith, D. L., & Somervaille, T. C. P. (2018). Enhancer Activation by Pharmacologic Displacement of LSD1 from GFI1 Induces Differentiation in Acute Myeloid Leukemia. *Cell Reports*, *22*(13), 3641–3659. https://doi.org/10.1016/j.celrep.2018.03.012

Majello, B., Gorini, F., Saccà, C., & Amente, S. (2019). Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer. *Cancers*, *11*(3), 324. https://doi.org/10.3390/cancers11030324

Malovannaya, A., Lanz, R. B., Jung, S. Y., Bulynko, Y., Le, N. T., Chan, D. W., Ding, C., Shi, Y., Yucer, N., Krenciute, G., Kim, B.-J., Li, C., Chen, R., Li, W., Wang, Y., O'Malley, B. W., & Qin, J. (2011). Analysis of the Human Endogenous Coregulator Complexome. *Cell*, *145*(5), 787–799. https://doi.org/10.1016/j.cell.2011.05.006

Margolin, J. F., Friedman, J. R., Meyer, W. K. H., Vissing, H., Thiesen -, H. J., & Rauscher, F. J. (1994). Kruppel-associated boxes are potent transcriptional repression domains. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(10), 4509–4513. https://doi.org/10.1073/pnas.91.10.4509

Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., Pirrotta, V., Reinberg, D., & Gamblin, S. J. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*, *461*(7265), 762–767. https://doi.org/10.1038/nature08398

Mariadason, J. M. (2008). HDACs and HDAC inhibitors in colon cancer. *Epigenetics*, *3*(1), 28–37. https://doi.org/10.4161/epi.3.1.5736

Markossian, S., Ang, K. K., Wilson, C. G., & Arkin, M. R. (2018). Small-molecule screening for genetic diseases. In *Annual Review of Genomics and Human Genetics* (Vol. 19, pp. 263–288). Annual Reviews Inc.

https://doi.org/10.1146/annurev-genom-083117-021452

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E., & Perrimon, N. (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genetics*, *40*(4), 476–483. https://doi.org/10.1038/ng.101

Martin, S. E., & Caplen, N. J. (2007). Applications of RNA Interference in Mammalian Systems. *Annual Review of Genomics and Human Genetics*, *8*(1), 81–108. https://doi.org/10.1146/annurev.genom.8.080706.092424

Mas, Glòria, Blanco, E., Ballaré, C., Sansó, M., Spill, Y. G., Hu, D., Aoi, Y., Le Dily, F., Shilatifard, A., Marti-Renom, M. A., & Di Croce, L. (2018). Promoter bivalency favors an open chromatin architecture in embryonic stem cells. *Nature Genetics*, *50*(10), 1452–1462. https://doi.org/10.1038/s41588-018-0218-5

Mas, Gloria, & Di Croce, L. (2016). The role of Polycomb in stem cell genome architecture. In *Current Opinion in Cell Biology* (Vol. 43, pp. 87–95). Elsevier Ltd. https://doi.org/10.1016/j.ceb.2016.09.006

Masukata, H., & Tomizawa, J. ichi. (1990). A mechanism of formation of a persistent hybrid between elongating RNA and template DNA. *Cell*, *62*(2), 331–338. https://doi.org/10.1016/0092-8674(90)90370-T

Masumoto, H., Hawke, D., Kobayashi, R., & Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature*, *436*(7048), 294–298. https://doi.org/10.1038/nature03714

Mathew, R., Hartmuth, K., Möhlmann, S., Urlaub, H., Ficner, R., & Lührmann, R. (2008). Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.1415

Mátrai, J., Chuah, M. K., & Vandendriessche, T. (2010). Recent advances in lentiviral vector development and applications. In *Molecular Therapy* (Vol. 18, Issue 3, pp. 477–490). Nature Publishing Group. https://doi.org/10.1038/mt.2009.319

Maunakea, A. K., Chepelev, I., Cui, K., & Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Research*, *23*(11), 1256–1269. https://doi.org/10.1038/cr.2013.110

Mazzio, E. A., & Soliman, K. F. A. (2012). Basic concepts of epigenetics impact of environmental signals on gene expression. In *Epigenetics* (Vol. 7, Issue 2, pp. 119–130). Taylor and Francis Inc. https://doi.org/10.4161/epi.7.2.18764

Meir, Z., Mukamel, Z., Chomsky, E., Lifshitz, A., & Tanay, A. (2020). Single-cell analysis of clonal maintenance of transcriptional and epigenetic states in cancer cells. *Nature Genetics*, *52*(7), 709–718. https://doi.org/10.1038/s41588-020-0645-y

Melcher, M., Schmid, M., Aagaard, L., Selenko, P., Laible, G., & Jenuwein, T. (2000). Structure-Function Analysis of SUV39H1 Reveals a Dominant Role in Heterochromatin Organization, Chromosome Segregation, and Mitotic Progression. *Molecular and Cellular Biology*, *20*(10), 3728–3741. https://doi.org/10.1128/mcb.20.10.3728-3741.2000

Menafra, R., & Stunnenberg, H. G. (2014). MBD2 and MBD3: elusive functions and mechanisms. *Frontiers in Genetics*, *5*(DEC), 428. https://doi.org/10.3389/fgene.2014.00428

Meng, Y., Li, H., Liu, C., Zheng, L., & Shen, B. (2018). Jumonji domain-containing protein family: The functions beyond lysine demethylation. In *Journal of Molecular Cell Biology* (Vol. 10, Issue 4, pp. 371–373). Oxford University Press. https://doi.org/10.1093/jmcb/mjy010

Messerschmidt, D. M., De Vries, W., Ito, M., Solter, D., Ferguson-Smith, A., & Knowles, B. B. (2012). Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science*, *335*(6075), 1499–1502. https://doi.org/10.1126/science.1216154

Messerschmidt, D. M., Knowles, B. B., & Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. In *Genes and Development* (Vol. 28, Issue 8, pp. 812–828). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/gad.234294.113

Metzger, E., Wissmann, M., Yin, N., Müller, J. M., Schneider, R., Peters, A. H. F. M., Günther, T., Buettner, R., & Schüle, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor- dependent transcription. *Nature*, *437*(7057), 436–439. https://doi.org/10.1038/nature04020

Mikhailova, T., Shuvalova, E., Ivanov, A., Susorov, D., Shuvalov, A., Kolosov, P. M., & Alkalaeva, E. (2017). RNA helicase DDX19 stabilizes ribosomal elongation and termination complexes. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw1239

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., … Bernstein, B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, *448*(7153), 553–560. https://doi.org/10.1038/nature06008

Minotti, L., Agnoletto, C., Baldassari, F., Corrà, F., & Volinia, S. (2018). SNPs and Somatic Mutation on Long Non-Coding RNA: New Frontier in the Cancer Studies? *High-Throughput*, *7*(4), 34. https://doi.org/10.3390/ht7040034

Mishra, K., & Kanduri, C. (2019). Understanding long noncoding RNA and chromatin interactions: What we know so far. In *Non-coding RNA* (Vol. 5, Issue 4, p. 54). MDPI AG. https://doi.org/10.3390/ncrna5040054

Mocavini, I., & Di Croce, L. (2020). RNA closing the Polycomb circle. In *Nature Genetics* (Vol. 52, Issue 9, pp. 866–867). Nature Research. https://doi.org/10.1038/s41588-020-0683-5

Morillon, A., Karabetsou, N., Nair, A., & Mellor, J. (2005). Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. *Molecular Cell*, *18*(6), 723–734. https://doi.org/10.1016/j.molcel.2005.05.009

Mosammaparast, N., & Shi, Y. (2010). Reversal of histone methylation: Biochemical and molecular mechanisms of histone demethylases. In *Annual Review of Biochemistry* (Vol. 79, pp. 155–179). Annual Reviews . https://doi.org/10.1146/annurev.biochem.78.070907.103946

Moufarrij, S., Dandapani, M., Arthofer, E., Gomez, S., Srivastava, A., Lopez-Acevedo, M., Villagra, A., & Chiappinelli, K. B. (2019). Epigenetic therapy for ovarian cancer: Promise and progress. *Clinical Epigenetics*, *11*(1), 1–11. https://doi.org/10.1186/s13148-018-0602-0

Moussa, H. F., Bsteh, D., Yelagandula, R., Pribitzer, C., Stecher, K., Bartalska, K., Michetti, L., Wang, J., Zepeda-Martinez, J. A., Elling, U., Stuckey, J. I., James, L. I., Frye, S. V., & Bell, O. (2019). Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing. *Nature Communications*. https://doi.org/10.1038/s41467-019-09628-6

Muhar, M., Ebert, A., Neumann, T., Umkehrer, C., Jude, J., Wieshofer, C., Rescheneder, P., Lipp, J. J., Herzog, V. A., Reichholf, B., Cisneros, D. A., Hoffmann, T., Schlapansky, M. F., Bhat, P., von Haeseler, A., Köcher, T., Obenauf, A. C., Popow, J., Ameres, S. L., & Zuber, J. (2018). SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science (New York, N.Y.)*, eaao2793. https://doi.org/10.1126/science.aao2793

Müller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., & Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*, *111*(2), 197–208. https://doi.org/10.1016/S0092-8674(02)00976-5

Murphy, K. E., Shylo, N. A., Alexander, K. A., Churchill, A. J., Copperman, C., & García-García, M. J. (2016). The Transcriptional Repressive Activity of KRAB Zinc Finger Proteins Does Not Correlate with Their Ability to Recruit TRIM28. *PLOS ONE*, *11*(9), e0163555. https://doi.org/10.1371/journal.pone.0163555

Murray, K. (1964). The Occurrence of N-Methyl Lysine in Histones. *Biochemistry*, *3*(1). https://pubs.acs.org/sharingguidelines

Nakata, D., Nakao, S., Nakayama, K., Araki, S., Nakayama, Y., Aparicio, S., Hara, T., & Nakanishi, A. (2017). The RNA helicase DDX39B and its paralog DDX39A regulate androgen receptor splice variant AR-V7 generation. *Biochemical and Biophysical Research Communications*. https://doi.org/10.1016/j.bbrc.2016.12.153

Narlikar, G. J., Sundaramoorthy, R., & Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. In *Cell* (Vol. 154, Issue 3, pp. 490–503). Cell Press. https://doi.org/10.1016/j.cell.2013.07.011

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., & Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. *Nature*, *543*(7643), 72–77. https://doi.org/10.1038/nature21373

Neumann, H., Hancock, S. M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., & Chin, J. W. (2009). A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. *Molecular Cell*, *36*(1), 153–163. https://doi.org/10.1016/j.molcel.2009.07.027

Nicolas, D., Phillips, N. E., & Naef, F. (2017). What shapes eukaryotic transcriptional bursting? In *Molecular BioSystems* (Vol. 13, Issue 7, pp. 1280–1290). Royal Society of Chemistry. https://doi.org/10.1039/c7mb00154a

Niehrs, C., & Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. *Nature Reviews Molecular Cell Biology*, *21*(3), 167–178. https://doi.org/10.1038/s41580-019-0206-3

Nishibuchi, G., & Déjardin, J. (2017). The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Research*, *25*(1), 77–87. https://doi.org/10.1007/s10577-016-9547-3

Noberini, R., Sigismondo, G., & Bonaldi, T. (2016). The contribution of mass spectrometry-based proteomics to understanding epigenetics. In *Epigenomics* (Vol. 8, Issue 3, pp. 429–445). Future Medicine Ltd. https://doi.org/10.2217/epi.15.108

Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D., & Baldwin, A. J. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell*, *57*(5), 936–947. https://doi.org/10.1016/j.molcel.2015.01.013

O'Geen, H., Ren, C., Nicolet, C. M., Perez, A. A., Halmai, J., Le, V. M., MacKay, J. P., Farnham, P. J., & Segal, D. J. (2017). DCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Research*, *45*(17), 9901–9916. https://doi.org/10.1093/nar/gkx578

O'Leary, V. B., Hain, S., Maugg, D., Smida, J., Azimzadeh, O., Tapio, S., Ovsepian, S. V., & Atkinson, M. J. (2017). Long non-coding RNA PARTICLE bridges histone and DNA methylation. *Scientific Reports*, *7*(1), 1790. https://doi.org/10.1038/s41598-017-01875-1

O'Neill, L. P., & Turner, B. M. (2003). Immunoprecipitation of native chromatin: NChIP. *Methods*, *31*(1), 76–82. https://doi.org/10.1016/S1046-2023(03)00090-2

Okamura, M., Yamanaka, Y., Shigemoto, M., Kitadani, Y., Kobayashi, Y., Kambe, T., Nagao, M., Kobayashi, I., Okumura, K., & Masuda, S. (2018). Depletion of mRNA export regulator DBP5/DDX19, GLE1 or IPPK that is a key enzyme for the production of IP6, resulting in differentially altered cytoplasmic mRNA expression and specific cell defect. *PLOS ONE*, *13*(5), e0197165. https://doi.org/10.1371/journal.pone.0197165

Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, *19*(3), 219–220. https://doi.org/10.1038/890

Olino, K., Park, T., & Ahuja, N. (2020). Exposing Hidden Targets: Combining epigenetic and immunotherapy to overcome cancer resistance. In *Seminars in Cancer Biology* (Vol. 65, pp. 114–122). Academic Press. https://doi.org/10.1016/j.semcancer.2020.01.001

Ong, S. H., Li, Y., Koike-Yusa, H., & Yusa, K. (2017). Optimised metrics for CRISPR-KO screens with second-generation gRNA libraries. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-07827-z

Owen, I., & Shewmaker, F. (2019). The role of post-translational modifications in the phase transitions of intrinsically disordered proteins. In *International Journal of Molecular Sciences* (Vol. 20, Issue 21, p. 5501). MDPI AG. https://doi.org/10.3390/ijms20215501

Paige, J. S., Wu, K. Y., & Jaffrey, S. R. (2011). RNA mimics of green fluorescent protein. *Science, 333*(6042), 642–646. https://doi.org/10.1126/science.1207339

Papin, C., Le Gras, S., Ibrahim, A., Salem, H., Karimi, M. M., Stoll, I., Ugrinova, I., Schröder, M., Fontaine-Pelletier, E., Omran, Z., Bronner, C., Dimitrov, S., & Hamiche, A. (2020). CpG Islands Shape the Epigenome Landscape. *Journal of Molecular Biology*. https://doi.org/10.1016/j.jmb.2020.09.018

Papp, B., & Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. In *Cell* (Vol. 152, Issue 6, pp. 1324–1343). Cell Press. https://doi.org/10.1016/j.cell.2013.02.043

Parnas, O., Jovanovic, M., Eisenhaure, T. M., Herbst, R. H., Dixit, A., Ye, C. J., Przybylski, D., Platt, R. J., Tirosh, I., Sanjana, N. E., Shalem, O., Satija, R., Raychowdhury, R., Mertins, P., Carr, S. A., Zhang, F., Hacohen, N., & Regev, A. (2015). A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell*, *162*(3), 675–686. https://doi.org/10.1016/j.cell.2015.06.059

Pelossof, R., Fairchild, L., Huang, C. H., Widmer, C., Sreedharan, V. T., Sinha, N., Lai, D. Y., Guan, Y., Premsrirut, P. K., Tschaharganeh, D. F., Hoffmann, T., Thapar, V., Xiang, Q., Garippa, R. J., Rätsch, G., Zuber, J., Lowe, S. W., Leslie, C. S., & Fellmann, C. (2017). Prediction of potent shRNAs with a sequential classification algorithm. *Nature Biotechnology*, *35*(4), 350–353. https://doi.org/10.1038/nbt.3807

Perillo, B., Tramontano, A., Pezone, A., & Migliaccio, A. (2020). LSD1: more than demethylation of histone lysine residues. In *Experimental and Molecular Medicine* (Vol. 52, Issue 12, pp. 1936–1947). Springer Nature. https://doi.org/10.1038/s12276-020-00542-2

Petell, C. J., Alabdi, L., He, M., San Miguel, P., Rose, R., & Gowher, H. (2016). An epigenetic switch regulates de novo DNA methylation at a subset of pluripotency gene enhancers during embryonic stem cell differentiation. *Nucleic Acids Research*, *44*(16), 7605–7617. https://doi.org/10.1093/nar/gkw426

Peters, A. H. F. M., Mermoud, J. E., O'carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., & Jenuwein, T. (2002). Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nature Genetics*, *30*(1), 77–80. https://doi.org/10.1038/ng789

Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology*, *20*(8), 490–507. https://doi.org/10.1038/s41580-019-0131-5

Pinter, S., Knodel, F., Choudalakis, M., Schnee, P., Kroll, C., Fuchs, M., Broehm, A., Weirich, S., Roth, M., Eisler, S. A., Zuber, J., Jeltsch, A., & Rathert, P. (2021). A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkab180

Pisignano, G., Pavlaki, I., & Murrell, A. (2019). Being in a loop: How long non-coding RNAs organise genome architecture. In *Essays in Biochemistry* (Vol. 63, Issue 1, pp. 177–186). Portland Press Ltd. https://doi.org/10.1042/EBC20180057

Piunti, A., & Shilatifard, A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science*, *352*(6290), aad9780. https://doi.org/10.1126/science.aad9780

Poole, C. J., & van Riggelen, J. (2017). MYC—master regulator of the cancer epigenome and transcriptome. In *Genes* (Vol. 8, Issue 5). MDPI AG. https://doi.org/10.3390/genes8050142

Porro, A., Feuerhahn, S., & Lingner, J. (2014). TERRA-Reinforced Association of LSD1 with MRE11 Promotes Processing of Uncapped Telomeres. *Cell Reports*, *6*(4), 765–776. https://doi.org/10.1016/J.CELREP.2014.01.022

Portela, A., & Esteller, M. (2010). *Epigenetic modifications and human disease*. https://doi.org/10.1038/nbt.1685

Qian, C., & Zhou, M. M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. In *Cellular and Molecular Life Sciences* (Vol. 63, Issue 23, pp. 2755–2763). Springer. https://doi.org/10.1007/s00018-006-6274-5

Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., Xiang, A. P., Ren, B.-Z., & Lahn, B. T. (2010). Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLoS ONE*, *5*(5), e10611. https://doi.org/10.1371/journal.pone.0010611

Qin, W., Wolf, P., Liu, N., Link, S., Smets, M., Mastra, F. La, Forné, I., Pichler, G., Hörl, D., Fellinger, K., Spada, F., Bonapace, I. M., Imhof, A., Harz, H., & Leonhardt, H. (2015). DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Research*, *25*(8), 911–929. https://doi.org/10.1038/cr.2015.72

Quagliano, A., Gopalakrishnapillai, A., & Barwe, S. P. (2020). Understanding the Mechanisms by Which Epigenetic Modifiers Avert Therapy Resistance in Cancer. In *Frontiers in Oncology* (Vol. 10, p. 992). Frontiers Media S.A.

https://doi.org/10.3389/fonc.2020.00992

Rahman, S., Sowa, M. E., Ottinger, M., Smith, J. A., Shi, Y., Harper, J. W., & Howley, P. M. (2011). The Brd4 Extraterminal Domain Confers Transcription Activation Independent of pTEFb by Recruiting Multiple Proteins, Including NSD3. *Molecular and Cellular Biology*, *31*(13), 2641–2652. https://doi.org/10.1128/mcb.01341-10

Rahmani, E., Schweiger, R., Rhead, B., Criswell, L. A., Barcellos, L. F., Eskin, E., Rosset, S., Sankararaman, S., & Halperin, E. (2019). Cell-type-specific resolution epigenetics without the need for cell sorting or single-cell biology. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-11052-9

Raisner, R. M., Hartley, P. D., Meneghini, M. D., Bao, M. Z., Liu, C. L., Schreiber, S. L., Rando, O. J., & Madhani, H. D. (2005). Histone variant H2A.Z Marks the 5′ ends of both active and inactive genes in euchromatin. *Cell*, *123*(2), 233–248. https://doi.org/10.1016/j.cell.2005.10.002

Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., & Tyagi, S. (2006). Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biology*, *4*(10), e309. https://doi.org/10.1371/journal.pbio.0040309

Ramírez, A., Milot, E., Ponsa, I., Marcos-Gutiérrez, C., Page, A., Santos, M., Jorcano, J., & Vidal, M. (2001). Sequence and chromosomal context effects on variegated expression of keratin 5/lacZ constructs in stratified epithelia of transgenic mice. *Genetics*, *158*(1), 341–350. https://doi.org/10.1093/genetics/158.1.341

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., & Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, *159*(7), 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

Rasmussen, K. D., & Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. In *Genes and Development* (Vol. 30, Issue 7, pp. 733–750). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/gad.276568.115

Rathert, P., Roth, M., Neumann, T., Muerdter, F., Roe, J. S., Muhar, M., Deswal, S., Cerny-Reiterer, S., Peter, B., Jude, J., Hoffmann, T., Boryn, L. M., Axelsson, E., Schweifer, N., Tontsch-Grunt, U., Dow, L. E., Gianni, D., Pearson, M., Valent, P., … Zuber, J. (2015). Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature*, *525*(7570), 543–547. https://doi.org/10.1038/nature14898

Rathert, P., Zhang, X., Freund, C., Cheng, X., & Jeltsch, A. (2008). Analysis of the Substrate Specificity of the Dim-5 Histone Lysine Methyltransferase Using Peptide Arrays. *Chemistry and Biology*, *15*(1), 5–11. https://doi.org/10.1016/j.chembiol.2007.11.013

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtier, K., Ponting, C. P., Allis, C. D., & Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, *406*(6796), 593–599. https://doi.org/10.1038/35020506

Reik, W., Dean, W., & Walter, J. (2001). Epigenetic reprogramming in mammalian development. In *Science* (Vol. 293, Issue 5532, pp. 1089–1093). https://doi.org/10.1126/science.1063443

Reik, Wolf, & Walter, J. (2001). Evolution of imprinting mechanisms: The battle of the sexes begins in the zygote. In *Nature Genetics* (Vol. 27, Issue 3, pp. 255–256). Nature Publishing Group. https://doi.org/10.1038/85804

Reuter, G., & Spierer, P. (1992). Position effect variegation and chromatin proteins. In *BioEssays* (Vol. 14, Issue 9, pp. 605–612). John Wiley & Sons, Ltd. https://doi.org/10.1002/bies.950140907

Reynolds, N., Latos, P., Hynes-Allen, A., Loos, R., Leaford, D., O'Shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P., Kalkan, T., Costello, I., Humphreys, P., Mansfield, W., Nakagawa, K., Strouboulis, J., Behrens, A., Bertone, P., & Hendrich, B. (2012). NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. *Cell Stem Cell*, *10*(5), 583–594. https://doi.org/10.1016/j.stem.2012.02.020

Ribeiro de Almeida, C., Dhir, S., Dhir, A., Moghaddam, A. E., Sattentau, Q., Meinhart, A., & Proudfoot, N. J. (2018). RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Molecular Cell*, *70*(4), 650-662.e8. https://doi.org/10.1016/j.molcel.2018.04.001

Richards, B. M., & Pardon, J. F. (1970). The molecular structure of nucleohistone (DNH). *Experimental Cell Research*, *62*(1), 184–196. https://doi.org/10.1016/0014-4827(79)90519-6

Robbez-Masson, L., Tie, C. H. C., Conde, L., Tunbak, H., Husovsky, C., Tchasovnikarova, I. A., Timms, R. T., Herrero, J., Lehner, P. J., & Rowe, H. M. (2018). The hush complex cooperates with trim28 to repress young retrotransposons and new genes. *Genome Research*, *28*(6), 836–845. https://doi.org/10.1101/gr.228171.117

Robertson, K. D. (2001). DNA methylation, methyltransferases, and cancer. *Oncogene*, *20*(24), 3139–3155. https://doi.org/10.1038/sj.onc.1204341

Roll, J. D., Rivenbark, A. G., Jones, W. D., & Coleman, W. B. (2008). DNMT3b overexpression contributes to a hypermethylator phenotype in human breast cancer cell lines. *Molecular Cancer*, *7*. https://doi.org/10.1186/1476-4598-7-15

Rossetto, D., Avvakumov, N., & Côté, J. (2012). Histone phosphorylation: A chromatin modification involved in diverse nuclear events. In *Epigenetics* (Vol. 7, Issue 10, pp. 1098–1108). Taylor and Francis Inc. https://doi.org/10.4161/epi.21975

Rotem, A., Ram, O., Shoresh, N., Sperling, R. A., Goren, A., Weitz, D. A., & Bernstein, B. E. (2015). Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nature Biotechnology*, *33*(11), 1165–1172. https://doi.org/10.1038/nbt.3383

Rots, M. G., & Jeltsch, A. (2018). Editing the Epigenome: Overview, Open Questions, and Directions of Future Development. In A. Jeltsch & M. G. Rots (Eds.), *Epigenome Editing: Methods and Protocols* (pp. 3–18). Springer New York.

https://doi.org/10.1007/978-1-4939-7774-1_1

Roy, P. H., & Weissbach, A. (1975). DNA methylase from hela cell nuclei. *Nucleic Acids Research*, *2*(10), 1669–1684. https://doi.org/10.1093/nar/2.10.1669

Ruiz-Carrillo, A., Wangh, L. J., & Allfrey, V. G. (1975). Processing of Newly Synthesized Histone Molecules. *Science*, *190*(4210), 117–128. http://www.jstor.org/stable/1740932

Russo, V. E. A. (Vincenzo E. A. ., Martienssen, R. A., & Riggs, A. D. (1996). *Epigenetic mechanisms of gene regulation*. Cold Spring Harbor Laboratory Press. https://agris.fao.org/agris-search/search.do?recordID=US201300304567

Saksouk, N., Simboeck, E., & Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. In *Epigenetics and Chromatin* (Vol. 8, Issue 1, pp. 1–17). BioMed Central Ltd. https://doi.org/10.1186/1756-8935-8-3

Saleque, S., Kim, J., Rooke, H. M., & Orkin, S. H. (2007). Epigenetic Regulation of Hematopoietic Differentiation by Gfi-1 and Gfi-1b Is Mediated by the Cofactors CoREST and LSD1. *Molecular Cell*, *27*(4), 562–572. https://doi.org/10.1016/j.molcel.2007.06.039

Sanchez, A., & Golding, I. (2013). Genetic determinants and cellular constraints in noisy gene expression. In *Science* (Vol. 342, Issue 6163, pp. 1188–1193). American Association for the Advancement of Science. https://doi.org/10.1126/science.1242975

Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., Vaimberg, E. W., Goodale, A., Root, D. E., Piccioni, F., & Doench, J. G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications*, *9*(1), 1–15. https://doi.org/10.1038/s41467-018-07901-8

Santoro, R., Li, J., & Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nature Genetics*, *32*(3), 393–396. https://doi.org/10.1038/ng1010

Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., Griffin, P. R., Gross, J. D., & Narlikar, G. J. (2019). HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature*, *575*(7782), 390–394. https://doi.org/10.1038/s41586-019-1669-2

Sanz, L. A., Hartono, S. R., Lim, Y. W., Steyaert, S., Rajpurkar, A., Ginno, P. A., Xu, X., & Chédin, F. (2016). Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2016.05.032

Scarola, M., Comisso, E., Pascolo, R., Chiaradia, R., Maria Marion, R., Schneider, C., Blasco, M. A., Schoeftner, S., & Benetti, R. (2015). Epigenetic silencing of Oct4 by a complex containing SUV39H1 and Oct4 pseudogene lncRNA. *Nature Communications*, *6*(1), 1–13. https://doi.org/10.1038/ncomms8631

Schenk, T., Chen, W. C., Göllner, S., Howell, L., Jin, L., Hebestreit, K., Klein, H. U., Popescu, A. C., Burnett, A., Mills, K., Casero, R. A., Marton, L., Woster, P., Minden, M. D., Dugas, M., Wang, J. C. Y., Dick, J. E., Müller-Tidow, C., Petrie, K.,

& Zelent, A. (2012). Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia. *Nature Medicine*, *18*(4), 605–611. https://doi.org/10.1038/nm.2661

Schmitges, F. W., Prusty, A. B., Faty, M., Stützer, A., Lingaraju, G. M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., Bunker, R. D., Wirth, U., Bouwmeester, T., Bauer, A., Ly-Hartig, N., Zhao, K., Chan, H., Gu, J., Gut, H., … Thomä, N. H. (2011). Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2011.03.025

Scholz, O., Henßler, E.-M., Bail, J., Schubert, P., Bogdanska-Urbaniak, J., Sopp, S., Reich, M., Wisshak, S., Köstner, M., Bertram, R., & Hillen, W. (2004). Activity reversal of Tet repressor caused by single amino acid exchanges. *Molecular Microbiology*, *53*(3), 777–789. https://doi.org/10.1111/j.1365-2958.2004.04159.x

Schröder, A. R. W., Shinn, P., Chen, H., Berry, C., Ecker, J. R., & Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, *110*(4), 521–529. https://doi.org/10.1016/S0092-8674(02)00864-4

Schübeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., Van Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J., Bell, S. P., & Groudine, M. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes and Development*, *18*(11), 1263–1271. https://doi.org/10.1101/gad.1198204

Schultz, D. C., Friedman, J. R., & Rauscher, F. J. (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: The PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2α subunit of NuRD. *Genes and Development*, *15*(4), 428–443. https://doi.org/10.1101/gad.869501

Schultz, David C, Ayyanathan, K., Negorev, D., Maul, G. G., & Rauscher Iii, F. J. (2002). SETDB1 : a novel KAP - 1 - associated histone H3 , lysine 9 - specific methyltransferase that contributes to HP1 - mediated silencing of euchromatic genes by KRAB zinc - finger proteins. *Genes and Development*, *16*, 919–932. https://doi.org/10.1101/gad.973302

Scourzic, L., Mouly, E., & Bernard, O. A. (2015). TET proteins and the control of cytosine demethylation in cancer. *Genome Medicine*, *7*(1), 9. https://doi.org/10.1186/s13073-015-0134-6

Sehrawat, A., Gao, L., Wang, Y., Bankhead, A., McWeeney, S. K., King, C. J., Schwartzman, J., Urrutia, J., Bisson, W. H., Coleman, D. J., Joshi, S. K., Kim, D. H., Sampson, D. A., Weinmann, S., Kallakury, B. V. S., Berry, D. L., Haque, R., Van Den Eeden, S. K., Sharma, S., … Alumkal, J. J. (2018). LSD1 activates a lethal prostate cancer gene network independently of its demethylase function. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1719168115

Serre, D., Lee, B. H., & Ting, A. H. (2009). MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Research*, *38*(2), 391–399. https://doi.org/10.1093/nar/gkp992

Seruggia, D., Oti, M., Tripathi, P., Canver, M. C., LeBlanc, L., Di Giammartino, D. C., Bullen, M. J., Nefzger, C. M., Sun, Y. B. Y., Farouni, R., Polo, J. M., Pinello, L., Apostolou, E., Kim, J., Orkin, S. H., & Das, P. P. (2019). TAF5L and TAF6L Maintain Self-Renewal of Embryonic Stem Cells via the MYC Regulatory Network. *Molecular Cell*, *74*(6), 1148-1163.e7. https://doi.org/10.1016/j.molcel.2019.03.025

Seto, E., & Yoshida, M. (2014). Erasers of histone acetylation: The histone deacetylase enzymes. *Cold Spring Harbor Perspectives in Biology*, *6*(4), a018713. https://doi.org/10.1101/cshperspect.a018713

Shaffer, S. M., Emert, B. L., Reyes Hueros, R. A., Cote, C., Harmange, G., Schaff, D. L., Sizemore, A. E., Gupte, R., Torre, E., Singh, A., Bassett, D. S., & Raj, A. (2020). Memory Sequencing Reveals Heritable Single-Cell Gene Expression Programs Associated with Distinct Cellular Behaviors. *Cell*, *182*(4), 947-959.e17. https://doi.org/10.1016/j.cell.2020.07.003

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., & Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, *343*(6166), 84–87. https://doi.org/10.1126/science.1247005

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology*, *22*(12), 1567–1572. https://doi.org/10.1038/nbt1037

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., & Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, *119*(7), 941–953. https://doi.org/10.1016/j.cell.2004.12.012

Shi, Y., Sawada, J. I., Sui, G., Affar, E. B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P. S., Nakatani, Y., & Shi, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature*, *422*(6933), 735–738. https://doi.org/10.1038/nature01550

Shiio, Y., Rose, D. W., Aur, R., Donohoe, S., Aebersold, R., & Eisenman, R. N. (2006). Identification and Characterization of SAP25, a Novel Component of the mSin3 Corepressor Complex. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.26.4.1386-1397.2006

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: Mechanisms of regulation in development and disease pathogenesis. *Annual Review of Biochemistry*, *81*, 65–95. https://doi.org/10.1146/annurev-biochem-051710-134100

Shimizu, N., Kawakami, K., & Ishitani, T. (2012). Visualization and exploration of Tcf/Lef function using a highly responsive Wnt/B-catenin signaling-reporter transgenic zebrafish. *Developmental Biology*, *370*(1), 71–85. https://doi.org/10.1016/j.ydbio.2012.07.016

Shin, S., & Janknecht, R. (2007). Diversity within the JMJD2 histone demethylase family. *Biochemical and Biophysical Research Communications*, *353*(4), 973–977. https://doi.org/10.1016/j.bbrc.2006.12.147

Shinkai, Y., & Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. In *Genes and Development* (Vol. 25, Issue 8, pp. 781–788). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/gad.2027411

Simon, M. D., Chu, F., Racki, L. R., de la Cruz, C. C., Burlingame, A. L., Panning, B., Narlikar, G. J., & Shokat, K. M. (2007). The Site-Specific Installation of Methyl-Lysine Analogs into Recombinant Histones. *Cell*, *128*(5), 1003–1012. https://doi.org/10.1016/j.cell.2006.12.041

Skourti-Stathaki, K., & Proudfoot, N. J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. In *Genes and Development*. https://doi.org/10.1101/gad.242990.114

Skourti-Stathaki, K., Torlai Triglia, E., Warburton, M., Voigt, P., Bird, A., & Pombo, A. (2019). R-Loops Enhance Polycomb Repression at a Subset of Developmental Regulator Genes. *Molecular Cell*, *73*(5), 930-945.e4. https://doi.org/10.1016/j.molcel.2018.12.016

Smanski, M. J., Zhou, H., Claesen, J., Shen, B., Fischbach, M. A., & Voigt, C. A. (2016). Synthetic biology to access and expand nature's chemical diversity. *Nature Reviews Microbiology*, *14*(3), 135–149. https://doi.org/10.1038/nrmicro.2015.24

Smith, S. S., Kaplan, B. E., Sowers, L. C., & Newman, E. M. (1992). Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(10), 4744–4748. https://doi.org/10.1073/pnas.89.10.4744

Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature Reviews Genetics*, *14*(3), 204–220. https://doi.org/10.1038/nrg3354

Sneider, T. W., Teague, W. M., & Rogachevsky, L. M. (1975). S-adenosylmethionine: Dna-cytosine 5-methyltransferase from a novikoff rat hepatoma cell line. *Nucleic Acids Research*, *2*(10), 1685–1700. https://doi.org/10.1093/nar/2.10.1685

Sneppen, K., & Ringrose, L. (2019). Theoretical analysis of Polycomb-Trithorax systems predicts that poised chromatin is bistable and not bivalent. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10130-2

Söderberg, O., Leuchowius, K. J., Gullberg, M., Jarvius, M., Weibrecht, I., Larsson, L. G., & Landegren, U. (2008). Characterizing proteins and their interactions in cells and tissues using the in situ proximity ligation assay. *Methods*, *45*(3), 227–232. https://doi.org/10.1016/j.ymeth.2008.06.014

Sommer, C. A., Stadtfeld, M., Murphy, G. J., Hochedlinger, K., Kotton, D. N., & Mostoslavsky, G. (2009). Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette. *Stem Cells*, *27*(3), 543–549. https://doi.org/10.1634/stemcells.2008-1075

Son, J., Shen, S. S., Margueron, R., & Reinberg, D. (2013). Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. *Genes and Development*, *27*(24), 2663–2677.

https://doi.org/10.1101/gad.225888.113

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., & Li, G. (2014). Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. *Science*, *344*(6182), 376–380. https://doi.org/10.1126/science.1251413

Song, J., Teplova, M., Ishibe-Murakami, S., & Patel, D. J. (2012). Structure-Based Mechanistic Insights into DNMT1-Mediated Maintenance DNA Methylation. *Science*, *335*(6069), 709 LP – 712. https://doi.org/10.1126/science.1214453

Song, Y., Dagil, L., Fairall, L., Robertson, N., Wu, M., Ragan, T. J., Savva, C. G., Saleh, A., Morone, N., Kunze, M. B. A., Jamieson, A. G., Cole, P. A., Hansen, D. F., & Schwabe, J. W. R. (2020). Mechanism of Crosstalk between the LSD1 Demethylase and HDAC1 Deacetylase in the CoREST Complex. *Cell Reports*, *30*(8), 2699-2711.e8. https://doi.org/10.1016/j.celrep.2020.01.091

Soshnikova, N., & Duboule, D. (2009). Epigenetic temporal control of mouse hox genes in vivo. *Science*, *324*(5932), 1321–1323. https://doi.org/10.1126/science.1171468

Soufi, A., Donahue, G., & Zaret, K. S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*, *151*(5), 994–1004. https://doi.org/10.1016/j.cell.2012.09.045

Soumyanarayanan, U., & Dymock, B. W. (2016). Recently discovered EZH2 and EHMT2 (G9a) inhibitors. In *Future Medicinal Chemistry* (Vol. 8, Issue 13, pp. 1635–1654). Future Science. https://doi.org/10.4155/fmc-2016-0096

Soutoglou, E., & Talianidis, I. (2002). Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science*, *295*(5561), 1901–1904. https://doi.org/10.1126/science.1068356

Sripathy, S. P., Stevens, J., & Schultz, D. C. (2006). The KAP1 Corepressor Functions To Coordinate the Assembly of De Novo HP1-Demarcated Microenvironments of Heterochromatin Required for KRAB Zinc Finger Protein-Mediated Transcriptional Repression. *Molecular and Cellular Biology*, *26*(22), 8623–8638. https://doi.org/10.1128/mcb.00487-06

Stasevich, T. J., Hayashi-Takanaka, Y., Sato, Y., Maehara, K., Ohkawa, Y., Sakata-Sogawa, K., Tokunaga, M., Nagase, T., Nozaki, N., McNally, J. G., & Kimura, H. (2014). Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature*, *516*(7530), 272–275. https://doi.org/10.1038/nature13714

Stavropoulos, P., Blobel, G., & Hoelz, A. (2006). *Crystal structure and mechanism of human lysine-specific demethylase-1*. https://doi.org/10.1038/nsmb1113

Stergachis, A. B., Haugen, E., Shafer, A., Fu, W., Vernot, B., Reynolds, A., Raubitschek, A., Ziegler, S., LeProust, E. M., Akey, J. M., & Stamatoyannopoulos, J. A. (2013). Exonic transcription factor binding directs codon choice and affects protein evolution. *Science*, *342*(6164), 1367–1372. https://doi.org/10.1126/science.1243490

Stevens, T. J., Lando, D., Basu, S., Atkinson, L. P., Cao, Y., Lee, S. F., Leeb, M., Wohlfahrt, K. J., Boucher, W., O'Shaughnessy-Kirwan, A., Cramard, J., Faure, A. J., Ralser, M., Blanco, E., Morey, L., Sansó, M., Palayret, M. G. S., Lehner, B., Di Croce, L., … Laue, E. D. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*, *544*(7648), 59–64. https://doi.org/10.1038/nature21429

Stewart-Morgan, K. R., Petryk, N., & Groth, A. (2020). Chromatin replication and epigenetic cell memory. In *Nature Cell Biology* (Vol. 22, Issue 4, pp. 361–371). Nature Research. https://doi.org/10.1038/s41556-020-0487-y

Stewart, C. A., & Byers, L. A. (2015). Altering the Course of Small Cell Lung Cancer: Targeting Cancer Stem Cells via LSD1 Inhibition. In *Cancer Cell* (Vol. 28, Issue 1, pp. 4–6). Cell Press. https://doi.org/10.1016/j.ccell.2015.06.011

Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, *403*(6765), 41–45. https://doi.org/10.1038/47412

Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z.-W., Bone, J. R., Caldwell, J. A., Mollah, S., Cook, R. G., Shabanowitz, J., Hunt, D. F., & Allis, C. D. (2002). Set2 Is a Nucleosomal Histone H3-Selective Methyltransferase That Mediates Transcriptional Repression. *Molecular and Cellular Biology*, *22*(5), 1298–1306. https://doi.org/10.1128/mcb.22.5.1298-1306.2002

Strahl, B. D., Ohba, R., Cook, R. G., & Allis, C. D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(26), 14967–14972. https://doi.org/10.1073/pnas.96.26.14967

Streubel, G., Fitzpatrick, D. J., Oliviero, G., Scelfo, A., Moran, B., Das, S., Munawar, N., Watson, A., Wynne, K., Negri, G. L., Dillon, E. T., Jammula, S., Hokamp, K., O'Connor, D. P., Pasini, D., Cagney, G., & Bracken, A. P. (2017). Fam60a defines a variant Sin3a-Hdac complex in embryonic stem cells required for self-renewal. *The EMBO Journal*, *36*(15), 2216–2232. https://doi.org/10.15252/embj.201696307

Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X., & Karpen, G. H. (2017). Phase separation drives heterochromatin domain formation. *Nature*, *547*(7662), 241–245. https://doi.org/10.1038/nature22989

Su, S.-T., Ying, H.-Y., Chiu, Y.-K., Lin, F.-R., Chen, M.-Y., & Lin, K.-I. (2009). Involvement of Histone Demethylase LSD1 in Blimp-1-Mediated Gene Repression during Plasma Cell Differentiation. *Molecular and Cellular Biology*, *29*(6), 1421–1431. https://doi.org/10.1128/mcb.01158-08

Su, Y., Ryder, J., Li, B., Wu, X., Fox, N., Solenberg, P., Brune, K., Paul, S., Zhou, Y., Liu, F., & Ni, B. (2004). Lithium, a common drug for bipolar disorder treatment, regulates amyloid-β precursor protein processing. *Biochemistry*, *43*(22), 6899–6908. https://doi.org/10.1021/bi035627j

Sun, G., Alzayady, K., Stewart, R., Ye, P., Yang, S., Li, W., & Shi, Y. (2010). Histone Demethylase LSD1 Regulates Neural Stem Cell Proliferation. *Molecular and Cellular Biology*, *30*(8), 1997–2005. https://doi.org/10.1128/mcb.01116-09

Sun, M., Nie, F., Wang, Y., Zhang, Z., Hou, J., He, D., Xie, M., Xu, L., De, W., Wang, Z., & Wang, J. (2016). LncRNA HOXA11-AS promotes proliferation and invasion of gastric cancer by scaffolding the chromatin modification factors PRC2, LSD1, and DNMT1. *Cancer Research*, *76*(21), 6299–6310. https://doi.org/10.1158/0008-5472.CAN-16-0356

Sun, Z., Zhang, Y., Jia, J., Fang, Y., Tang, Y., Wu, H., & Fang, D. (2020). H3K36me3, message from chromatin to DNA damage repair. In *Cell and Bioscience* (Vol. 10, Issue 1, pp. 1–9). BioMed Central Ltd. https://doi.org/10.1186/s13578-020-0374-z

Swanson, K. A., Knoepfler, P. S., Huang, K., Kang, R. S., Cowley, S. M., Laherty, C. D., Eisenman, R. N., & Radhakrishnan, I. (2004). HBP1 and Mad1 repressors bind the Sin3 corepressor PAH2 domain with opposite helical orientations. *Nature Structural and Molecular Biology*, *11*(8), 738–746. https://doi.org/10.1038/nsmb798

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., & Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *The EMBO Journal*, *27*(20), 2681–2690. https://doi.org/10.1038/emboj.2008.192

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., & Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes and Development*, *19*(7), 815–826. https://doi.org/10.1101/gad.1284005

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., & Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, *324*(5929), 930–935. https://doi.org/10.1126/science.1170116

Takagi, S., Ishikawa, Y., Mizutani, A., Iwasaki, S., Matsumoto, S., Kamada, Y., Nomura, T., & Nakamura, K. (2017). LSD1 inhibitor T-3775440 inhibits SCLC cell proliferation by disrupting LSD1 interactions with SNAG domain proteins INSM1 and GFI1B. *Cancer Research*, *77*(17), 4652–4662. https://doi.org/10.1158/0008-5472.CAN-16-3502

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Tchasovnikarova, I. A., Timms, R. T., Matheson, N. J., Wals, K., Antrobus, R., Göttgens, B., Dougan, G., Dawson, M. A., & Lehner, P. J. (2015). Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science*, *348*(6242), 1481–1485. https://doi.org/10.1126/science.aaa7227

Tekel, S. J., & Haynes, K. A. (2017). Molecular structures guide the engineering of chromatin. In *Nucleic Acids Research* (Vol. 45, Issue 13, pp. 7555–7570). Oxford University Press. https://doi.org/10.1093/nar/gkx531

Thakore, P. I., Black, J. B., Hilton, I. B., & Gersbach, C. A. (2016). Editing the epigenome: technologies for programmable transcription and epigenetic modulation. In *Nature Methods* (Vol. 13, Issue 2, pp. 127–137). Nature Publishing Group. https://doi.org/10.1038/nmeth.3733

Timms, R. T., Tchasovnikarova, I. A., Antrobus, R., Dougan, G., & Lehner, P. J. (2016). ATF7IP-Mediated Stabilization of the Histone Methyltransferase SETDB1 Is Essential for Heterochromatin Formation by the HUSH Complex. *Cell Reports*, *17*(3), 653–659. https://doi.org/10.1016/j.celrep.2016.09.050

Timms, R. T., Tchasovnikarova, I. A., & Lehner, P. J. (2016). Position-effect variegation revisited: HUSHing up heterochromatin in human cells. *BioEssays*, *38*(4), 333–343. https://doi.org/10.1002/bies.201500184

Ting, X., Xia, L., Yang, J., He, L., Si, W., Shang, Y., & Sun, L. (2019). USP11 acts as a histone deubiquitinase functioning in chromatin reorganization during DNA repair. *Nucleic Acids Research*, *47*(18), 9721–9740. https://doi.org/10.1093/nar/gkz726

Tochio, N., Umehara, T., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Watanabe, S., Tomo, Y., Hanada, M., Ikari, M., Sato, M., Terada, T., Nagase, T., Ohara, O., Shirouzu, M., Tanaka, A., Kigawa, T., & Yokoyama, S. (2006). Solution structure of the SWIRM domain of human histone demethylase LSD1. *Structure*, *14*(3), 457–468. https://doi.org/10.1016/j.str.2005.12.004

Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E., & Schreiber, S. L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature*, *395*(6705), 917–921. https://doi.org/10.1038/27699

Torres, I. O., & Fujimori, D. G. (2015). Functional coupling between writers, erasers and readers of histone and DNA methylation. In *Current Opinion in Structural Biology* (Vol. 35, pp. 68–75). Elsevier Ltd. https://doi.org/10.1016/j.sbi.2015.09.007

Tough, D. F. (2016). Immune disease-associated variants in gene enhancers point to BET epigenetic mechanisms for therapeutic intervention. *Epigenomics*. https://doi.org/10.2217/epi-2016-0144

Tran, E. J., Zhou, Y., Corbett, A. H., & Wente, S. R. (2007). The DEAD-Box Protein Dbp5 Controls mRNA Export by Triggering Specific RNA:Protein Remodeling Events. *Molecular Cell*, *28*(5), 850–859. https://doi.org/10.1016/j.molcel.2007.09.019

Tripathi, V., Shen, Z., Chakraborty, A., Giri, S., Freier, S. M., Wu, X., Zhang, Y., Gorospe, M., Prasanth, S. G., Lal, A., & Prasanth, K. V. (2013). Long Noncoding RNA MALAT1 Controls Cell Cycle Progression by Regulating the Expression of Oncogenic Transcription Factor B-MYB. *PLoS Genetics*, *9*(3), e1003368. https://doi.org/10.1371/journal.pgen.1003368

Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E., & Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, *329*(5992), 689–693. https://doi.org/10.1126/science.1192002

Tsukada, Y. I., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., & Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, *439*(7078), 811–816. https://doi.org/10.1038/nature04433

Tsusaka, T., Kikuchi, M., Shimazu, T., Suzuki, T., Sohtome, Y., Akakabe, M., Sodeoka, M., Dohmae, N., Umehara, T., & Shinkai, Y. (2018). Tri-methylation of ATF7IP by G9a/GLP recruits the chromodomain protein. *Epigenetics and Chromatin*, *11*(1). https://doi.org/10.1186/s13072-018-0231-z

Tycko, J., DelRosso, N., Hess, G. T., Aradhana, Banerjee, A., Mukund, A., Van, M. V., Ego, B. K., Yao, D., Spees, K., Suzuki, P., Marinov, G. K., Kundaje, A., Bassik, M. C., & Bintu, L. (2020). High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell*, *183*(7), 2020-2035.e16. https://doi.org/10.1016/j.cell.2020.11.024

Tycko, J., Van, M. V., Elowitz, M. B., & Bintu, L. (2017). Advancing towards a global mammalian gene regulation model through single-cell analysis and synthetic biology. *Current Opinion in Biomedical Engineering*, *4*, 174–193. https://doi.org/10.1016/j.cobme.2017.10.011

Tzelepis, K., Koike-Yusa, H., De Braekeleer, E., Li, Y., Metzakopian, E., Dovey, O. M., Mupo, A., Grinkevich, V., Li, M., Mazan, M., Gozdecka, M., Ohnishi, S., Cooper, J., Patel, M., McKerrell, T., Chen, B., Domingues, A. F., Gallipoli, P., Teichmann, S., … Yusa, K. (2016). A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Reports*, *17*(4), 1193–1205. https://doi.org/10.1016/j.celrep.2016.09.079

Urrutia, R. (2003). KRAB-containing zinc-finger repressor proteins. In *Genome Biology* (Vol. 4, Issue 10). Genome Biol. https://doi.org/10.1186/gb-2003-4-10-231

Valls, E., Sánchez-Molina, S., & Martínez-Balbás, M. A. (2005). Role of Histone Modifications in Marking and Activating Genes through Mitosis*. *Journal of Biological Chemistry*, *280*(52), 42592–42600. https://doi.org/https://doi.org/10.1074/jbc.M507407200

van Attikum, H., & Gasser, S. M. (2009). Crosstalk between histone modifications during the DNA damage response. *Trends in Cell Biology*, *19*(5), 207–217. https://doi.org/https://doi.org/10.1016/j.tcb.2009.03.001

van Emmerik, C. L., & van Ingen, H. (2019). Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR. In *Progress in Nuclear Magnetic Resonance Spectroscopy* (Vol. 110, pp. 1–19). Elsevier B.V. https://doi.org/10.1016/j.pnmrs.2019.01.002

Venkatesh, S., & Workman, J. L. (2015). Histone exchange, chromatin structure and the regulation of transcription. In *Nature Reviews Molecular Cell Biology* (Vol. 16, Issue 3, pp. 178–189). Nature Publishing Group. https://doi.org/10.1038/nrm3941

Verde, G., Querol-Paños, J., Cebrià-Costa, J., Pascual-Reguant, L., Serra-Bardenys, G., Iturbide, A., & Peiró, S. (2017). Lysine-Specific Histone Demethylases Contribute to Cellular Differentiation and Carcinogenesis. *Epigenomes*, *1*(1), 4. https://doi.org/10.3390/epigenomes1010004

Vieux-Rochas, M., Fabre, P. J., Leleu, M., Duboule, D., & Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(15), 4672–4677. https://doi.org/10.1073/pnas.1504783112

Vigna, E., & Naldini, L. (2000). Lentiviral vectors: Excellent tools for experimental gene transfer and promising candidates for gene therapy. In *Journal of Gene Medicine* (Vol. 2, Issue 5, pp. 308–316). John Wiley and Sons Ltd. https://doi.org/10.1002/1521-2254(200009/10)2:5<308::aid-jgm131>3.0.co;2-3

Villamizar, O., Chambers, C. B., Riberdy, J. M., Persons, D. A., & Wilber, A. (2016). Long noncoding RNA Saf and splicing factor 45 increase soluble Fas and resistance to apoptosis. *Oncotarget*, *7*(12), 13810–13826. https://doi.org/10.18632/oncotarget.7329

Vinckier, N. K., Patel, N. A., Geusz, R. J., Wang, A., Wang, J., Matta, I., Harrington, A. R., Wortham, M., Wetton, N., Wang, J., Jhala, U. S., Rosenfeld, M. G., Benner, C. W., Shih, H. P., & Sander, M. (2020). LSD1-mediated enhancer silencing attenuates retinoic acid signalling during pancreatic endocrine cell development. *Nature Communications*, *11*(1), 1–15. https://doi.org/10.1038/s41467-020-16017-x

Vinyard, M. E., Su, C., Siegenfeld, A. P., Waterbury, A. L., Freedy, A. M., Gosavi, P. M., Park, Y., Kwan, E. E., Senzer, B. D., Doench, J. G., Bauer, D. E., Pinello, L., & Liau, B. B. (2019). CRISPR-suppressor scanning reveals a nonenzymatic role of LSD1 in AML. *Nature Chemical Biology*, *15*(5), 529–539. https://doi.org/10.1038/s41589-019-0263-0

Vissing, H., Meyer, W. K. H., Aagaard, L., Tommerup, N., & Thiesen, H. J. (1995). Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. *FEBS Letters*, *369*(2–3), 153–157. https://doi.org/10.1016/0014-5793(95)00728-R

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J., & Grunstein, M. (2002). Histone Acetylation Regulates the Time of Replication Origin Firing. *Molecular Cell*, *10*(5), 1223–1233. https://doi.org/https://doi.org/10.1016/S1097-2765(02)00702-5

Voigt, P., Tee, W. W., & Reinberg, D. (2013). A double take on bivalent promoters. In *Genes and Development* (Vol. 27, Issue 12, pp. 1318–1338). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/gad.219626.113

Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., Klasic, M., & Zoldos, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research*, *44*(12), 5615–5628. https://doi.org/10.1093/nar/gkw159

Vucic, E. A., Wilson, I. M., Campbell, J. M., & Lam, W. L. (2009). Methylation analysis by DNA immunoprecipitation (MeDIP). *Methods in Molecular Biology*, *556*, 141–

153. https://doi.org/10.1007/978-1-60327-192-9_10

Waddington, C. H. (1942). The Epigenotype. *Endeavour*, 18–20.

Waddington, C. H. (1953). Epigenetics and evolution. *Symp. Soc. Exp. Biol.*, *7*, 186–199.

Waddington, C. H. (1957). *The strategy of the genes*. MacMillan.

Wallrath, L. L. (1998). Unfolding the mysteries of heterochromatin. *Current Opinion in Genetics and Development*, *8*(2), 147–153. https://doi.org/10.1016/S0959-437X(98)80135-4

Wang, F., & Higgins, J. M. G. (2013). Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends in Cell Biology*, *23*(4), 175–184. https://doi.org/https://doi.org/10.1016/j.tcb.2012.11.005

Wang, Jianxun, Scully, K., Zhu, X., Cai, L., Zhang, J., Prefontaine, G. G., Krones, A., Ohgi, K. A., Zhu, P., Garcia-Bassets, I., Liu, F., Taylor, H., Lozach, J., Jayes, F. L., Korach, K. S., Glass, C. K., Fu, X. D., & Rosenfeld, M. G. (2007). Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature*, *446*(7138), 882–887. https://doi.org/10.1038/nature05671

Wang, Jianxun, Telese, F., Tan, Y., Li, W., Jin, C., He, X., Basnet, H., Ma, Q., Merkurjev, D., Zhu, X., Liu, Z., Zhang, J., Ohgi, K., Taylor, H., White, R. R., Tazearslan, C., Suh, Y., Macfarlan, T. S., Pfaff, S. L., & Rosenfeld, M. G. (2015). LSD1n is an H4K20 demethylase regulating memory formation via transcriptional elongation control. *Nature Neuroscience*, *18*(9), 1256–1264. https://doi.org/10.1038/nn.4069

Wang, Jing, Hevi, S., Kurash, J. K., Lei, H., Gay, F., Bajko, J., Su, H., Sun, W., Chang, H., Xu, G., Gaudet, F., Li, E., & Chen, T. (2009). The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature Genetics*, *41*(1), 125–129. https://doi.org/10.1038/ng.268

Wang, L., Gao, Y., Zheng, X., Liu, C., Dong, S., Li, R., Zhang, G., Wei, Y., Qu, H., Li, Y., Allis, C. D., Li, G., Li, H., & Li, P. (2019). Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. *Molecular Cell*, *76*(4), 646-659.e6. https://doi.org/10.1016/j.molcel.2019.08.019

Wang, L., Hu, M., Zuo, M. Q., Zhao, J., Wu, D., Huang, L., Wen, Y., Li, Y., Chen, P., Bao, X., Dong, M. Q., Li, G., & Li, P. (2020). Rett syndrome-causing mutations compromise MeCP2-mediated liquid–liquid phase separation of chromatin. *Cell Research*, *30*(5), 393–407. https://doi.org/10.1038/s41422-020-0288-7

Wang, X., Paucek, R. D., Gooding, A. R., Brown, Z. Z., Ge, E. J., Muir, T. W., & Cech, T. R. (2017). Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nature Structural and Molecular Biology*, *24*(12), 1028–1038. https://doi.org/10.1038/nsmb.3487

Wang, Y., Zhang, H., Chen, Y., Sun, Y., Yang, F., Yu, W., Liang, J., Sun, L., Yang, X., Shi, L., Li, R., Li, Y., Zhang, Y., Li, Q., Yi, X., & Shang, Y. (2009). LSD1 Is a Subunit of the NuRD Complex and Targets the Metastasis Programs in Breast Cancer.

Cell, *138*(4), 660–672. https://doi.org/10.1016/j.cell.2009.05.050

Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W., & Zhao, K. (2009). Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. *Cell.* https://doi.org/10.1016/j.cell.2009.06.049

Watson, J. D., & Crick, F. H. (1953). The structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology*, *18*, 123–131. https://doi.org/10.1101/SQB.1953.018.01.020

Weigt, D., Hopf, C., & Médard, G. (2016). Studying epigenetic complexes and their inhibitors with the proteomics toolbox. In *Clinical Epigenetics* (Vol. 8, Issue 1, pp. 1–16). Springer Verlag. https://doi.org/10.1186/s13148-016-0244-z

Weinberg, D. N., Papillon-Cavanagh, S., Chen, H., Yue, Y., Chen, X., Rajagopalan, K. N., Horth, C., McGuire, J. T., Xu, X., Nikbakht, H., Lemiesz, A. E., Marchione, D. M., Marunde, M. R., Meiners, M. J., Cheek, M. A., Keogh, M. C., Bareke, E., Djedid, A., Harutyunyan, A. S., … Lu, C. (2019). The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature*, *573*(7773), 281–286. https://doi.org/10.1038/s41586-019-1534-3

Weirich, S., Kudithipudi, S., & Jeltsch, A. (2016). Specificity of the SUV4-20H1 and SUV4-20H2 protein lysine methyltransferases and methylation of novel substrates. *Journal of Molecular Biology*, *428*(11), 2344–2358. https://doi.org/10.1016/j.jmb.2016.04.015

Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A., & Feinberg, A. P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature Genetics*, *41*(2), 246–250. https://doi.org/10.1038/ng.297

Whyte, W. A., Bilodeau, S., Orlando, D. A., Hoke, H. A., Frampton, G. M., Foster, C. T., Cowley, S. M., & Young, R. A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature*, *482*(7384), 221–225. https://doi.org/10.1038/nature10805

Wierer, M., & Mann, M. (2016). Proteomics to study DNA-bound and chromatin-associated gene regulatory complexes. In *Human Molecular Genetics* (Vol. 25, Issue R2, pp. R106–R114). Oxford University Press. https://doi.org/10.1093/hmg/ddw208

Wiles, E. T., & Selker, E. U. (2017). H3K27 methylation: a promiscuous repressive chromatin mark. *Current Opinion in Genetics and Development*, *43*, 31–37. https://doi.org/10.1016/j.gde.2016.11.001

Wissmann, M., Yin, N., Müller, J. M., Greschik, H., Günther, T., Metzger, E., Schüle, R., Fodor, B. D., Jenuwein, T., Vogler, C., Schneider, R., & Buettner, R. (2007). Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nature Cell Biology*, *9*(3), 347–353. https://doi.org/10.1038/ncb1546

Wolf, D., & Goff, S. P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature*, *458*(7242), 1201–1204. https://doi.org/10.1038/nature07844

Wu, B., Pan, X., Chen, X., Chen, M., Shi, K., Xu, J., Zheng, J., Niu, T., Chen, C., Shuai, X., & Liu, Y. (2019). Epigenetic drug library screening identified an LSD1 inhibitor to target UTX-deficient cells for differentiation therapy. *Signal Transduction and Targeted Therapy*, *4*(1), 1–10. https://doi.org/10.1038/s41392-019-0040-2

Wu, H., & Zhang, Y. (2014). Reversing DNA methylation: Mechanisms, genomics, and biological functions. In *Cell* (Vol. 156, Issues 1–2, pp. 45–68). Cell. https://doi.org/10.1016/j.cell.2013.12.019

Wu, P. G., & Brand, L. (1994). Resonance energy transfer: Methods and applications. In *Analytical Biochemistry* (Vol. 218, Issue 1, pp. 1–13). Academic Press. https://doi.org/10.1006/abio.1994.1134

Xia, W., & Xie, W. (2020). Rebooting the Epigenomes during Mammalian Early Embryogenesis. *Stem Cell Reports*. https://doi.org/10.1016/j.stemcr.2020.09.005

Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., Okumura, K., & Li, E. (1999). Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene*, *236*(1), 87–95. https://doi.org/10.1016/S0378-1119(99)00252-8

Xu, C., Wu, Z., Duan, H. C., Fang, X., Jia, G., & Dean, C. (2021). R-loop resolution promotes co-transcriptional chromatin silencing. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22083-6

Xu, T. peng, Wang, W. yu, Ma, P., Shuai, Y., Zhao, K., Wang, Y. fen, Li, W., Xia, R., Chen, W. ming, Zhang, E. bao, & Shu, Y. qian. (2018). Upregulation of the long noncoding RNA FOXD2-AS1 promotes carcinogenesis by epigenetically silencing EphB3 through EZH2 and LSD1, and predicts poor prognosis in gastric cancer. *Oncogene*, *37*(36), 5020–5036. https://doi.org/10.1038/s41388-018-0308-y

Yang, M., Culhane, J. C., Szewczuk, L. M., Gocke, C. B., Brautigam, C. A., Tomchick, D. R., MacHius, M., Cole, P. A., & Yu, H. (2007). Structural basis of histone demethylation by LSD1 revealed by suicide inactivation. *Nature Structural and Molecular Biology*, *14*(6), 535–539. https://doi.org/10.1038/nsmb1255

Yang, P., Wang, Y., & Macfarlan, T. S. (2017). The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. In *Trends in Genetics* (Vol. 33, Issue 11, pp. 871–881). Elsevier Ltd. https://doi.org/10.1016/j.tig.2017.08.006

Yang, Y., Huang, W., Qiu, R., Liu, R., Zeng, Y., Gao, J., Zheng, Y., Hou, Y., Wang, S., Yu, W., Leng, S., Feng, D., Wang, Y., & Shen, Z. (2018). LSD1 coordinates with the SIN3A/HDAC complex and maintains sensitivity to chemotherapy in breast cancer. *Journal of Molecular Cell Biology*. https://doi.org/10.1093/jmcb/mjy021

Yeo, N. C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C. C., Guo, X., Sharma, S., Tung, A., Cecchi, R. J., Tuttle, M., Pradhan, S., Lim, E. T., Davidsohn, N., Ebrahimkhani, M. R., Collins, J. J., Lewis, N. E., Kiani, S., & Church, G. M. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nature Methods*, *15*(8), 611–616. https://doi.org/10.1038/s41592-018-0048-5

Yildirim, O., Li, R., Hung, J. H., Chen, P. B., Dong, X., Ee, L. S., Weng, Z., Rando, O. J., & Fazzio, T. G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell*, *147*(7), 1498–1510. https://doi.org/10.1016/j.cell.2011.11.054

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P. K., Kivioja, T., Dave, K., Zhong, F., Nitta, K. R., Taipale, M., Popov, A., Ginno, P. A., Domcke, S., Yan, J., Schübeler, D., Vinson, C., & Taipale, J. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science*, *356*(6337). https://doi.org/10.1126/science.aaj2239

Ying, Y., Yang, X., Zhao, K., Mao, J., Kuang, Y., Wang, Z., Sun, R., & Fei, J. (2015). The Krüppel-associated box repressor domain induces reversible and irreversible regulation of endogenous mouse genes by mediating different chromatin states. *Nucleic Acids Research*, *43*(3), 1549–1561. https://doi.org/10.1093/nar/gkv016

Yoneyama, M., Tochio, N., Umehara, T., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Watanabe, S., Tomo, Y., Nishimura, Y., Harada, T., Terada, T., Shirouzu, M., Hayashizaki, Y., Ohara, O., Tanaka, A., Kigawa, T., & Yokoyama, S. (2007). Structural and Functional Differences of SWIRM Domain Subtypes. *Journal of Molecular Biology*, *369*(1), 222–238. https://doi.org/10.1016/j.jmb.2007.03.027

Yoo, S. H., Ko, C. H., Lowrey, P. L., Buhr, E. D., Song, E. J., Chang, S., Yoo, O. J., Yamazaki, S., Lee, C., & Takahashi, J. S. (2005). A noncanonical E-box enhancer drives mouse Period2 circadian oscillations in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2608–2613. https://doi.org/10.1073/pnas.0409763102

Yoshida, M., Kijima, M., Akita, M., & Beppu, T. (1990). Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *Journal of Biological Chemistry*, *265*(28), 17174–17179. https://doi.org/10.1016/s0021-9258(17)44885-x

You, A., Tong, J. K., Grozinger, C. M., & Schreiber, S. L. (2001). CoREST is an integral component of the CoREST-human histone deacetylase complex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(4), 1454–1458. https://doi.org/10.1073/pnas.98.4.1454

Yu, J.-R., Lee, C.-H., Oksuz, O., Stafford, J. M., & Reinberg, D. (2019). PRC2 is high maintenance. *Genes & Development*, *33*(15–16), 903–935. https://doi.org/10.1101/gad.325050.119

Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E., & Lieber, M. R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunology*, *4*(5), 442–451. https://doi.org/10.1038/ni919

Zentner, G. E., & Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. In *Nature Structural and Molecular Biology* (Vol. 20, Issue 3, pp. 259–266). https://doi.org/10.1038/nsmb.2470

Zhang, H., Wu, Z., Lu, J. Y., Huang, B., Zhou, H., Xie, W., Wang, J., & Shen, X. (2020). DEAD-Box Helicase 18 Counteracts PRC2 to Safeguard Ribosomal DNA in Pluripotency Regulation. *Cell Reports*, *30*(1), 81-97.e7. https://doi.org/10.1016/j.celrep.2019.12.021

Zhang, Qi, McKenzie, N. J., Warneford-Thomson, R., Gail, E. H., Flanigan, S. F., Owen, B. M., Lauman, R., Levina, V., Garcia, B. A., Schittenhelm, R. B., Bonasio, R., & Davidovich, C. (2019). RNA exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2. *Nature Structural and Molecular Biology*, *26*(3), 237–247. https://doi.org/10.1038/s41594-019-0197-y

Zhang, Qiang, Schepis, A., Huang, H., Yang, J., Ma, W., Torra, J., Zhang, S. Q., Yang, L., Wu, H., Nonell, S., Dong, Z., Kornberg, T. B., Coughlin, S. R., & Shu, X. (2019). Designing a Green Fluorogenic Protease Reporter by Flipping a Beta Strand of GFP for Imaging Apoptosis in Animals. *Journal of the American Chemical Society*, *141*(11), 4526–4530. https://doi.org/10.1021/jacs.8b13042

Zhang, Yi, Iratni, R., Erdjument-Bromage, H., Tempst, P., & Reinberg, D. (1997). Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell*, *89*(3), 357–364. https://doi.org/10.1016/S0092-8674(00)80216-0

Zhang, Yingying, Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., & Jeltsch, A. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Research*, *38*(13), 4246–4253. https://doi.org/10.1093/nar/gkq147

Zhao, S., Yang, M., Zhou, W., Zhang, B., Cheng, Z., Huang, J., Zhang, M., Wang, Z., Wang, R., Chen, Z., Zhu, J., & Li, H. (2017). Kinetic and high-throughput profiling of epigenetic interactions by 3D-carbene chip-based surface plasmon resonance imaging technology. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(35), E7245–E7254. https://doi.org/10.1073/pnas.1704155114

Zhou, B.-R., Feng, H., Kale, S., Fox, T., Khant, H., de Val, N., Ghirlando, R., Panchenko, A. R., & Bai, Y. (2020). Distinct Structures and Dynamics of Chromatosomes with Different Human Linker Histone Isoforms. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2020.10.038

Zibetti, C., Adamo, A., Binda, C., Forneris, F., Toffolo, E., Verpelli, C., Ginelli, E., Mattevi, A., Sala, C., & Battaglioli, E. (2010). Alternative splicing of the histone demethylase LSD1/KDM1 contributes to the modulation of neurite morphogenesis in the mammalian nervous system. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5500-09.2010

Zilberman, D., Coleman-Derr, D., Ballinger, T., & Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature*, *456*(7218), 125–129. https://doi.org/10.1038/nature07324

Zimmer, M. (2002). Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. *Chemical Reviews*, *102*(3), 759–781. https://doi.org/10.1021/cr010142r

Zolotukhin, A. S., Uranishi, H., Lindtner, S., Bear, J., Pavlakis, G. N., & Felber, B. K. (2009). Nuclear export factor RBM15 facilitates the access of DBP5 to mRNA. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkp782

Zuber, J., McJunkin, K., Fellmann, C., Dow, L. E., Taylor, M. J., Hannon, G. J., & Lowe, S. W. (2011). Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nature Biotechnology*, *29*(1), 79–83. https://doi.org/10.1038/nbt.1720

Zuber, J., Shi, J., Wang, E., Rappaport, A. R., Herrmann, H., Sison, E. A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., Taylor, M. J., Johns, C., Chicas, A., Mulloy, J. C., Kogan, S. C., Brown, P., Valent, P., Bradner, J. E., Lowe, S. W., & Vakoc, C. R. (2011). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature*, *478*(7370), 524–528. https://doi.org/10.1038/nature10334

7. Author´s contribution

I want to thank all participants for their contribution to the published manuscript

The following contributions were made:

8. Acknowledgements

9. Appendix

9.1. Appendix 1

Pinter, S., Knodel, F., Choudalakis, M., Schnee, P., Kroll, C., Fuchs, M., Broehm, A., Weirich, S., Roth, M., Eisler, S. A., Zuber, J., Jeltsch, A., & Rathert, P. (2021). A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkab180

Published online 6 April 2021

A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation

Sabine Pinter[1], Franziska Knodel[1], Michel Choudalakis[1], Philipp Schnee[1], Carolin Kroll[1], Marina Fuchs[1], Alexander Broehm[1], Sara Weirich[1], Mareike Roth[2], Stephan A. Eisler[3], Johannes Zuber[2,4], Albert Jeltsch[1] and Philipp Rathert ®[1,*]

[1]Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, 70569 Stuttgart, Germany, [2]Research Institute of Molecular Pathology, Vienna BioCenter, Vienna, Austria, [3]Stuttgart Research Center Systems Biology (SRCSB), University of Stuttgart, 70569 Stuttgart, Germany and [4]Medical University of Vienna, Vienna BioCenter (VBC), Vienna, Austria

ABSTRACT

The lysine specific demethylase 1 (LSD1) plays a pivotal role in cellular differentiation by regulating the expression of key developmental genes in concert with different coregulatory proteins. This process is impaired in different cancer types and incompletely understood. To comprehensively identify functional coregulators of LSD1, we established a novel tractable fluorescent reporter system to monitor LSD1 activity in living cells. Combining this reporter system with a state-of-the-art multiplexed RNAi screen, we identify the DEAD-box helicase 19A (DDX19A) as a novel coregulator and demonstrate that suppression of *Ddx19a* results in an increase of R-loops and reduced LSD1-mediated gene silencing. We further show that DDX19A binds to tri-methylated lysine 27 of histone 3 (H3K27me3) and it regulates gene expression through the removal of transcription promoting R-loops. Our results uncover a novel transcriptional regulatory cascade where the downregulation of genes is dependent on the LSD1 mediated demethylation of histone H3 lysine 4 (H3K4). This allows the polycomb repressive complex 2 (PRC2) to methylate H3K27, which serves as a binding site for DDX19A. Finally, the binding of DDX19A leads to the efficient removal of R-loops at active promoters, which further de-represses LSD1 and PRC2, establishing a positive feedback loop leading to a robust repression of the target gene.

GRAPHICAL ABSTRACT
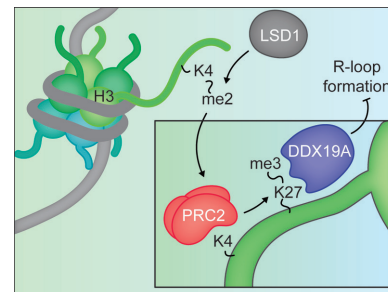
INTRODUCTION

The lysine specific demethylase 1 (LSD1, also known as KDM1A) has emerged as a critical regulator of essential physiological processes including the regulation of hormone receptor–mediated transcription (1), pluripotency and stem cell differentiation (2–5), cell cycle control (6) and DNA damage response (7). In agreement with the central role of LSD1 in such essential regulatory programs, LSD1 has been implicated in malignant transformation and maintenance of tumour pathogenesis in various ways. Overexpression of LSD1 has been observed in various tumour types (8–14) and imbalanced histone modifications, due to elevated LSD1 expression, are significantly associated with increased cellular growth and suppression of cell cycle regulatory proteins in a broad array of tissues. High levels of LSD1 have been shown to promote epithelial-to-mesenchymal transition (EMT) in breast cancer (BC) (15–17) and neuroblastoma (18), thereby contributing to cancer progression. Knockdown (KD) or inhibition of LSD1 reduces both the invasiveness and proliferative capacity of BC cells *in vitro* (19,20) and small molecules targeting LSD1 induce terminal differentiation of leukaemia cells (21,22).

Thus, LSD1 represents a critical oncogene and potential therapeutic target in different cancer subtypes.

Most biological functions of LSD1 are associated with its activity to regulate the lysine methylation state of histones and non-histone proteins. LSD1 has been highlighted for its dual ability to stimulate or suppress gene expression (23–25) and was reported to demethylate lysine residues on histones as well as non-histone substrates such as p53 and DNMT1 (26,27). LSD1 mediates the demethylation of histone H3K4me1 and H3K4me2, thereby conducting a transcriptional repression (28–30), in part through downregulation of enhancer function (22). Contradictory to its corepressor function, LSD1 can directly activate the expression of target genes through demethylation of histone H3K9me2 (29–32). The exact molecular mechanism of its dual substrate specificity remains unclear, but recent publications support the hypothesis that a newly discovered alternative LSD1 splice variant (LSD1+8a) restricted to neuronal tissues is responsible for demethylation of H3K9 (33–35). LSD1 has been shown to be associated with actively transcribed genes in many cell types (22,23,28), which suggests that its H3K4 demethylation activity is blocked at these loci. In fact, the activity of LSD1 is tightly controlled and counterbalanced by associated coregulators and the interaction of LSD1 with coregulatory complexes, e.g. CoREST or the NuRD histone deacetylase (HDAC) transcription corepressor complexes, represents an important regulatory feature (1,32,36,37). Additionally, LSD1 activity was shown to be negatively regulated by the interaction with specific RNA structures (38), a feature also shown for other coregulator complexes, e.g. PRC2 (39,40). Finally, LSD1 can be subject to post-translational modifications (PTMs) which regulate its transcriptional activity (41).

This highlights the immense complexity of LSD1 regulation on different levels, which creates highly specific and tightly controlled LSD1 transcriptional outputs regulated by coordinated fine-tuning of the binding affinity of LSD1 to target loci and complex partners. Understanding the dependence of LSD1 function on accessory proteins will shed light on several signaling pathways and provide new therapeutic avenues by targeting factors that modulate LSD1 activity instead of or additionally to targeting LSD1 itself (42). Understanding how LSD1 evokes specific transcriptional profiles depending on its association with defined coregulators in distinct cellular contexts will be critical for the development of novel and more efficient LSD1-focused therapies. To date no comprehensive strategy to identify LSD1 coregulators and unravel their molecular function has been devised.

Recent methodologic advances introduced the chromatin *in vivo* assay (CiA) system, a variation of chemical induced proximity (CIP), as a novel method to investigate the consequences of locally induced alterations of the chromatin landscape after controlled recruitment of an epigenetic effector (43). CiA has successfully been applied to study the dynamics of heterochromatin formation at the Oct4 locus in mouse embryonic stem cells (mESCs) after the recruitment of HP1α (44), components of the PRC2 complex (45) and, combined with a high throughput small molecule library screen, to identify compounds inducing the formation of euchromatin (46). Additionally, CiA has been used to in-

vestigate the opposing effect of the BAF complex on PRC-induced heterochromatin formation, leading to the formation of accessible chromatin (47,48).

We aimed to identify and characterize functional coregulators that are required for LSD1 activity and adopted the CiA concept to generate a time-resolved fluorescent reporter system to monitor the activity of LSD1 in cells. To identify essential and novel coregulators of LSD1, we combined our fluorescent reporter system with a microRNA-embedded short hairpin RNA (shRNAmir) library focused on epigenetic effectors to perform a chromatin effector coregulator screen (ChECS). Our results provide a detailed functional view on the coregulator network of LSD1 in a multiplexed manner. Deeper characterization of one of the top hits from the screen, the DEAD-box helicase 19A (DDX19A) showed that RNA:DNA hybrid structures (also called R-loops) strongly interfered with the activity of LSD1. Our data reveal a novel regulatory cascade, which enables LSD1 induced transcriptional repression via a three-step mechanism. The decrease of H3K4 methylation at a particular genomic region induced by the activity of LSD1 leads to the recruitment of PRC2 to introduce H3K27 methylation. This modification serves as a signal for DDX19A, which binds to H3K27me3 via a yet unknown motif and removes R-loops. This de-represses LSD1 and PRC2 establishing a positive feedback loop leading to a strong repression of transcription at the targeted region.

MATERIALS AND METHODS

Plasmids

The fluorescent reporter expressing *mCherry* from a synthetic promoter (*synP*), consisting of six *tetO* binding sites upstream of an *EF1a* promoter, was cloned into pMSCV vector, on which expression of *mCherry* was driven by the *synP* promoter element and coupled to *Blasticidin* resistance *via* a *P2A* (pMSCV-tetO-EF1a-mCherry-2A-Blasti). The *rTetR-LSD1* fusion construct was cloned into a *pRRL* backbone by standard cloning methods. The LSD1 construct was kindly provided by Tim Somervaille. Expression was driven from an *SFFV* promoter and coupled to *Hygromycin* resistance via a *P2A* sequence (pRRL-rTetR-LSD1-P2A-Hygro). shRNA guides were cloned into the SGEN vector (49).

Antibodies

Antibodies used for ChIP were H3K9me3 (ab8898, Abcam), H3K4me2 (ab7766, Abcam and #39141, Active Motif), H3K27me3 (#39155, Active Motif), H3K27Ac (ab4729, Abcam) and KDM1/LSD1 (ab17721, Abcam). Primary antibodies used for immunodetection after Western Blot were TetR monoclonal antibody 9G9 (#631131, TAKARA), DDX19A (orb242165, Biorbyt or ab108462, Abcam), RNA:DNA Hybrid Antibody, clone S9.6 (MABE1095, MERCK/Sigma-Aldrich) and KDM1A (#61607 and #39186, Active Motif). Secondary antibody used for immunofluorescence was the Goat Anti-Mouse IgG H&L (Alexa Fluor® 594, ab150116, Abcam). Antibody used for detection of DDX19A-GST was the goat anti-GST antibody (GE Healthcare, #27–4577-01).

Secondary antibodies for analysis of Western Blots were either coupled to horseradish peroxidase (GE Healthcare) or to IRDye® 800CW (ab216773, Abcam).

Pooled RNAi screening

After spiking in control shRNAs at equimolar amounts, the shRNA-mirE library (5451 shRNAs and 8 control shRNAs) targeting 1010 chromatin-associated murine genes was transduced into NIH/3T3 cells expressing the *synP-mCherry* reporter and rTetR-LSD1. To ensure library representation, a total of 30 million cells were infected with 10% transduction efficiency using conditions that predominantly lead to a single retroviral integration and represent each shRNA in a calculated number of >500 cells. Cells were split into replicates and selected with 2.5 mg/ml Neomycin for 7 days before starting treatment with 1 µg/ml doxycycline. Throughout selection >3 × 10^6 cells per replicate were maintained at each passage to preserve library representation. After 14 days of DOX treatment, cells were sorted into *mCherry*-positive (top 6–8%, minimum of 5 × 10^5 cells) and *mCherry*-negative (lowest 75–80%, minimum of 6 × 10^6 cells, see Supplementary Figure S2) populations using a FACS Aria III. Genomic DNA for both populations and 5 replicates was isolated with phenol-extraction using PhaseLock tubes, followed by ethanol precipitation. For each sample, DNA from at least 10^6 cells was used as template in multiple parallel 50-µl PCR reactions, each containing 1 µg template, 1× AmpliTaq Gold buffer, 0.2 mM of each dNTP, 2 mM MgCl$_2$, 0.3 µM of each primer and 1.25 U AmpliTaq Gold Polymerase (Life Technologies). In a first round of PCRs, random barcodes and sample barcodes were added to the shRNA sequences using the following cycling parameters: 95°C for 10 min; 28 cycles of (95°C for 30 s, 54°C for 45 s and 72°C for 60 s); 72°C for 7 min and primers *MM2P51_for* and *MM2P71_rev* (Supplementary Table S3). PCR products were combined for each sample, purified from a 1% agarose gel and 20 ng per sample were transferred to a second round of PCR, using similar cycling parameters as for PCR1, but with only 10 ng template per reaction, 6 cycles of amplification and primers *MM2P52_for* and *MM2P72_rev_N708* (Supplementary Table S3). In the second PCR, standard Illumina P7 adaptors and the Illumina N708 index were added to the sequences (total product length = 428 bp). All primers used for the library preparation are listed in Supplementary Table S2. The final libraries were cleaned up from a 1% agarose gel, pooled and analysed on an Illumina HiSeq 3000 deep sequencer (150 bp read length including the 22 nucleotides of the guide strand), using standard Illumina primers. Sequence processing was performed using a public Galaxy server (www.usegalaxy.eu). All primary screen data are provided in Supplementary Data Table S1. For each shRNA, the number of matching reads was normalized to the total number of library-specific reads per lane and imported into Microsoft Excel for further analysis. Completely depleted shRNAs (0 reads at T0) obtained a fold depletion value of 1 × 10^{-3}. The average enrichment score for each individual shRNA was calculated by dividing the geometric mean of the normalized reads of the mCherry+ population by the respective normalized reads mCherry− pop-

ulations (mCherry+/mCherry−) across five replicates. The gene score was derived by summarizing the average enrichment score of all shRNAs per gene. *P*-values are based on a Poisson distribution of each shRNA in each individual replicate followed by the combination of all *P*-values across all replicates using Fisher´s method (cumulative χ^2). χ^2 is a chi-squared distribution with 2k degrees of freedom, where k is the number of tests being combined. This fact was used to determine the *P*-value for χ^2 followed by a Bonferroni correction for multiple comparison to obtain a *P*-value for each investigated gene in the library.

Cell culture, retroviral transduction and flow cytometry

NIH/3T3, Lenti-X 293T and Platinum-E retroviral packaging cell lines were cultivated in DMEM high glucose media (Sigma-Aldrich) supplemented with 10% FBS, 20 mM glutamate, 10 mM sodium pyruvate, 10 mM HEPES (pH 7.3), 100 U/ml penicillin and 100 mg/ml streptomycin in an incubator providing 37°C and 5% CO$_2$. For retroviral packaging of pMSCV vectors, 20 µg of plasmid were precipitated for 20 min in HBS buffer (140 mM NaCl, 25 mM HEPES, 0.75 mM Na$_2$HPO$_4$, pH 7.0) together with 125 mM CaCl$_2$ and 10 µg GagPol helper plasmid. The mix was added to a 10 cm dish with Platinum-E cells growing at 75–85% confluence in supplemented DMEM. After 16 and 24 h, the media was replaced with fresh DMEM. Supernatant containing the virus was gathered 40–50 h after transfection, filtered through a 0.45 µm filter and added to the target cells at 50–70% confluence. Antibiotic selection for *pMSCV-tetO-EF1a-mCherry-2A-Blasti* with 10 µg/ml Blasticidin was started 2 days after transduction and kept up for 7 days. For retroviral packaging of pRRL-vectors, plasmids were mixed with helper plasmids *pCMVR8.74* (pCMVR8.74 was a gift from Didier Trono, Addgene #22036) and *pCAG-Eco* (pCAG-Eco was a gift from Arthur Nienhuis & Patrick Salmon, Addgene #35617) and 3× (w/w) excess of polyethyleneimine 25K in serum free DMEM. The mix was added to Lenti-X cells residing in supplemented DMEM at 75–90% confluence. Media exchanges and transduction of target cells was performed as described for pMSCV. Cells expressing *pRRL-rTetR-LSD1-P2A-Hygro* were selected with 500 µg/ml Hygromycin and cells expressing SGEN with 2.5 mg/ml Neomycin for 7 days. Recruitment of rTetR-LSD1 was started 12 days after transduction with SGEN by treatment with 1µg/ml Doxycycline. Inhibitor treatment with 10 µM GSK-LSD1, 5 µM GSK343 (Sigma-Aldrich) or 400 nM TSA (Sigma-Aldrich) was started in parallel with the first DOX induction and maintained for the indicated number of days. Expression of *GFP* and *mCherry* was analysed every 1–3 days using a MACSQuant Vyb flow cytometer.

ChIP-qPCR

For H3K4me2-, H3K27me3- and H3K27ac-ChIP experiments, NIH/3T3 stably expressing the *synP-mCherry* reporter, rTetR-LSD1 wt or K661A and if indicated, the respective shRNA, were treated with 1 µg/ml doxycycline for 4 days. Cells were washed once with 1× PBS, before incubation with 1% formaldehyde in 1× PBS for 15 min

at room temperature. Crosslinking was quenched with 225 mM glycine for 5 min. Cells were washed twice with 1× PBS and harvested with a Corning® cell scraper in 10 ml 1× PBS per 10 × 10⁶ cells. Cells were centrifuged for 8 min at 600 × *g* and the pellet was washed again with 10 ml 1× PBS, 500 nM TSA per 10 × 10⁶ cells. Pellets were split into aliquots of 5 × 10⁶ cells, snap frozen and stored at –80°C until use. For preparation of mononucleosomes, each pellet was lysed in 125 μl lysis buffer (10 mM Tris–HCl pH 7.4, 2 mM MgCl₂, 0.6% Igepal-Nonidet P40, 0.5 mM PMSF, 1 mM DTT, cOmplete™ EDTA-free PIC, 5 mM sodium-butyrate) for 15 min on ice. Samples were digested with 300 U micrococcal nuclease for 16 min at 37°C. The reaction was put on ice and stopped by addition of 8 μM EDTA, 0.1% Triton X-100 and 0.1% sodium deoxycholate. Samples were diluted by addition of 800 μl Complete IP buffer (20 mM Tris–HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.1% Triton X-100, 1 mM PMSF, cOmplete™ EDTA-free PIC, 5 mM sodium-butyrate) and clarified by centrifugation at 15 000 × *g* for 10 min at 4°C. The supernatant was split into aliquots of 40–70 μg chromatin and snap frozen. Before IP, *Drosophila melanogaster* mononucleosomes were added to the NIH/3T3 chromatin samples as spike-in control (2–3.5 μg = 5% of total chromatin). Ten percent of the sample was taken as input. For pre-clearing, 2.5 μg of rabbit/mouse IgG (depending on the species of antibody used for IP) and 10 μl of Dynabeads® Protein G were incubated with the sample for 2 h at 4°C with constant rotation. The beads were removed using a magnetic rack and the sample was split into halves for IP/IgG control. 2.5 μg of ChIP antibody or IgG were added to the samples and incubated over night at 4°C with constant rotation. 20 μl Dynabeads® Protein G per sample were blocked overnight in Complete IP buffer with 0.1 mg/ml BSA. Pre-blocked beads were incubated with the samples for 2 h at 4°C with rotation to bind antibodies. Beads were washed twice with low salt buffer (20 mM Tris–HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with high salt buffer (20 mM Tris–HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS) and once with TE buffer (10 mM Tris–HCl pH 8.0, 1 mM EDTA) to remove unspecific binding. Bound chromatin was eluted from the beads by resuspending in 100 μl SDS elution buffer (1% SDS, 100 mM NaHCO₃) and rotating for 30 min at room temperature. Elution was performed twice and the eluates were combined. Samples were de-crosslinked for 16 h at 65°C with 2 μg RNase A and 270 mM NaCl. Proteinase digest was performed for 2 h at 45°C with 60 μg Proteinase K. DNA fragments were extracted using the ChIP DNA Purification Kit (Active Motif), and amplified using ORA™SEE qPCR reagent (HighQ) and qPCR primers amplifying a 120 bp fragment of the *EF1a* promoter. Cq values were normalized to input and *Drosophila* spike-in control. ChIP for H3K9me3 was performed the same way, except that the samples were harvested after 14 days of DOX treatment and not crosslinked before fragmentation and IP.

ChIP-seq

LSD1 ChIP-seq was performed using the ChIP-IT High Sensitivity® Kit (Active Motif) following manufacturer's

instructions. In brief, 2 × 10⁷ NIH/3T3 cells were fixed and harvested as described in the protocol. Fragmentation was performed in aliquots of 5 × 10⁶ cells using an EpiShear Probe Sonicator (Active Motif) for 39 cycles (20 s pulse, 30 s pause, 40% amplitude), aliquots were united again and 5% were taken for input. To increase amount of precipitated chromatin, 3 × 25 μg chromatin were used for three independent IPs with 4 μg LSD1 antibody (ab17721, Abcam) each, following manufacturer's protocol. During DNA clean-up, the three samples were loaded onto two columns and the final eluates were united. Library preparation was performed using the NEBNext® Ultra™ II DNA Library Prep Kit following manufacturer´s protocol. 1 μg of input DNA and 50% of precipitated ChIP DNA were used. After end repair and adapter ligation, the input was amplified using standard Illumina primers i705 and i503 for 3 cycles, the ChIP sample was amplified using i706 + i504 for 11 cycles. Libraries were analysed on a LabChip® GX Touch™ Nucleic Acid Analyzer. Fragments with a size of 250–700 bp were sequenced on an Illumina HiSeq3000 using standard Illumina protocols.

ChIP-Seq data analysis

Data analysis was performed on a public Galaxy server (www.usegalaxy.eu). After quality control, the remaining reads were aligned to the respective genome (mouse: mm9, human: hg19) using Bowtie2. Reads with same start and end position on the same strand were removed from the alignment. To identify ChIP-seq peaks, we used the MACS2 peak finding algorithm (50). A threefold enrichment relative to input control samples was used for peak calling as well as the option to call broad peaks. Building a shifting model was disabled and the small nearby and large nearby region parameters were set to 5000 and 20 000, respectively. The extension size was set to the respective median insert size of the ChIP-seq treatment sample for paired-end data and the estimated fragment size for single-end data. Downstream analysis was performed using the deepTools2 (51) suite using the multiBigwigSummary function to compute the average scores for each of the bigWig files in every genomic region. This analysis was performed for the entire genome by running the program in bins mode. Subsequently the result was plotted using the plotPCA and plotCorrelation functions. Peaks were assigned to the respective genes using ChIP-enrich (52) by assigning peaks to the closest upstream/downstream TSS.

The NIH/3T3 LSD1 ChIP-seq data of our study is available at Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) entry GSE158441.

The following ChIP-seq tracks were obtained from published data sets in K562 cells and mapped to hg19: H3K27ac (Encode sample ENCFF384ZZM), H3K27me3 (Encode sample ENCFF936BVT), H3K9me3 (Encode sample ENCFF700FQH), H3K36me3 (Encode sample ENCFF223BKS), H3K4me1 (Encode sample ENCFF463AQS), H3K4me2 (Encode sample ENCFF778DNU), LSD1 (GEO sample GSM831002), R-ChIP (GEO sample GSM2551007/8), DRIP-seq (GEO sample GSM1720619), GQ-seq (GEO sample GSM2876090/1). K562 RNA-seq (GEO sample

GSM1557077). The following ChIP-seq tracks were obtained from published data sets in NIH/3T3 cells and mapped to mm9: H3K9ac (GEO sample GSM1246687), H3K27me3 (GEO sample GSM1246690), H3K36me3 (GEO sample GSM1246692), H3K9me2 (GEO sample GSM1246688), H4ac (GEO sample GSM1418787), H3K4me3 (GEO sample GSM879920), DRIPc-seq (GEO sample GSM2104456), DRIP-seq (GEO sample GSM1720621). NIH/3T3 RNA-seq (Encode sample ENCFF001QSC).

RNA:DNA hybrid IP (DRIP)

Based on a detailed assessment of various DRIP protocols (53), DRIP was performed as described, with slight adaptations to workflow #19 (53). In brief, NIH/3T3 cells were harvested, washed once with 1X PBS and cross-linked in 1% formaldehyde/PBS for 10 min at room temperature. Crosslinking was quenched with 500 mM glycine for 5 min at room temperature. Cells were lysed in 300 μl of ChIP lysis buffer (50mM HEPES–KOH at pH 7.5, 140 mM NaCl, 1 mM EDTA at pH 8, 1% Triton X-100, 0.1% sodium-deoxycholate, 1% SDS) per 2 million cells for 30 min on ice and homogenized with a syringe every 10 min. Chromatin was fragmented by sonication using an EpiShear Probe Sonicator (Active Motif) for 2×12 cycles (20 s pulse, 30 s pause, 40% amplitude). The fragmented chromatin was supplemented with 300 mM NaCl and 50 μg of RNase A in 450 μl TE buffer (10 mM Tris–HCl pH 8, 10 mM EDTA pH 8) and incubated at 37°C for 1 h. The cross-linked *D. melanogaster* mononucleosomes were treated in parallel to obtain DNA for spike-in controls. To remove proteins and reverse the cross-links, the samples were treated with 15 μl of Proteinase K (20 mg/ml; Thermo Fisher Scientific) at 65°C for 16 h. Nucleic acids were extracted in two rounds of phenol extraction using PhaseLock tubes (Eppendorf), followed by ethanol precipitation at –20°C overnight. 15 μl Dynabeads® Protein G per sample were blocked overnight in 1% BSA in 1× PBS. Nucleic acid precipitate was collected by centrifugation at 15 000 × g for 30 min, 4°C, the pellet was washed once with 70% ethanol and air-dried at 25–30°C. The pellet was resuspended in 5 mM Tris–HCl, pH 8.5 and concentration was determined by NanoDrop. To immobilize the S9.6 antibody, pre-blocked Dynabeads® Protein G were resuspended in 1 ml IP buffer (50 mM HEPES/KOH at pH 7.5; 0.14 M NaCl; 5 mM EDTA; 1% Triton X-100; 0.1% sodium-deoxycholate) and incubated with 2 μg of S9.6 antibody per sample for 4 h, 4°C. of Fragmented nucleic acids (1.5 μg) and 50 ng of *Drosophila* spike-in were added to the antibody/bead complexes and IP was performed overnight at 4°C. For the RNAseH1 controls, 1.5 μg of nucleic acids were digested with 40 U RNase H (NEB) at 37°C overnight. The enzyme was inactivated by incubation for 20 min at 65°C and the sample was taken as input for the IP in parallel with the untreated samples. Beads were washed once with 1 ml IP wash 1 buffer (20 mM Tris pH 8.0, 2 mM EDTA, 50 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with 1 ml high salt buffer (20 mM Tris pH 8, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.01% SDS), once with 1 ml IP wash buffer 2 (10 mM Tris pH 8, 1 mM EDTA, 0.25 M LiCl, 1% NP-40, 1% sodium-deoxycholate)

and twice in TE buffer (pH 8). Nucleic acids were eluted in 50 μl elution buffer (50 mM Tris pH 8, 10 mM EDTA, 1% SDS) for 15 min at 65°C and further purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), nucleic acids were eluted in 50 μl of elution buffer (5 mM Tris, pH 8.5). DNA fragments were amplified using ORA™SEE qPCR reagent (HighQ) and qPCR primers amplifying a 120 bp fragment of the *EF1a* promoter. Cq values were normalized to input and Drosophila spike-in control.

S9.6 antibody IP

100 μl Dynabeads® Protein G were pre-blocked with 0.5% BSA/PBS for 2 h at 4°C. 10×10^6 non-crosslinked NIH/3T3 cells were harvested, washed once in 1× PBS and lysed in 1 ml Cell Lysis Buffer (85 mM KCl, 5 mM HEPES pH 8, 0.5% NP-40, cOmplete™ EDTA-free PIC) for 15 min on ice. Nuclei were collected by spinning 1 min at 15 000 × g, 4°C. The pellet was resuspended in 750 μl RSB buffer (10 mM Tris–HCl pH 7.5, 200 mM NaCl, 2.5 mM MgCl₂, cOmplete™ EDTA-free PIC) supplemented with 0.2% sodium-deoxycholate, 0.1% SDS, 0.05% sodium-lauroyl-sarcosinate and 0.5% Triton X-100. Cells were sonicated for 4 min in an EpiShear Probe Sonicator (20 s pulse, 30 s pause, 40%). After taking 5% as input, samples were transferred to a 15 ml falcon tube and diluted 1:4 by addition of 2.3 ml RSB with 0.5% Triton X-100 (RSB+T). The samples were subjected to pre-clearing with 5 μg mouse IgG and 35 μl Dynabeads® Protein G for 1 h at 4°C. Magnetic beads were locked and the supernatant was split into three Eppendorf tubes. For the S9.6 antibody-specificity control, 40 U of RNAse H (NEB) were added to one of the tubes and all samples were incubated for 10 min at 37°C before adding the IP antibodies. The samples were subjected to IP with 4 μg of either the S9.6 antibody or mouse IgG and 32 μl pre-blocked beads per sample. 10 ng RNase A was added to each tube before rotating at 4°C for 2.5 h. Beads were washed 4× with 500 μl RSB+T and 2× with RSB. With each buffer change, beads were transferred to fresh low-binding tubes to minimize leftover unspecific binding. Proteins were eluted in 40 μl of 2× SDS sample buffer (125 mM Tris–HCl pH 6.8, 5% SDS, 0.004% Bromophenol Blue, 10% β-mercaptoethanol, 100 mM DTT, 20% glycerol) for 10 min at 70°C. The supernatant was transferred to fresh tubes and boiled at 95°C for 10 min for denaturation, along with the input sample mixed with 2× sample buffer. SDS-PAGE and immunodetection of proteins were performed as described for immunodetection of proteins.

Immunodetection of proteins after SDS-PAGE

For the analysis of protein levels, cells were harvested 13 days after transduction and antibiotic selection. Pellets were lysed in cell lysis buffer (Cell Signaling Technology®) for 30 min on ice. After 10 and 20 min of incubation, the lysate was sonicated with an EpiShear Probe Sonicator (Active Motif) for 2 cycles of 20 s to release nuclear protein. The lysate was centrifuged at 15 000 × g for 10 min, the supernatant was mixed with 2× SDS sample buffer (125 mM Tris–HCl pH 6.8, 5% SDS, 0.004% Bromophenol Blue, 10% β-mercaptoethanol, 100 mM DTT, 20% glycerol) and boiled

at 95°C for 10 min. Proteins were resolved by SDS-PAGE on a 12% polyacrylamide gel. Proteins were transferred to an Immobilon-FL PVDF membrane at 300 mA for 90 min using a wet-tank blotting system (BioRad). Proteins were detected using a target specific primary antibody at manufacturer´s recommendations in combination with a species-specific HRP- or IRDye® 800CW-coupled secondary antibody. Imaging was performed on a FusionFX detection system (VILBER) using SuperSignal™ West Femto Chemiluminescence substrate (ThermoFisher Odyssey® CLx imaging system (LI-COR).

Gene expression analysis

For analysis of mRNA expression levels of *Ddx19a*, cells were harvested 13 days after transduction with the specific shRNAs (Supplementary Table S3) and antibiotic selection. RNA was extracted using the *RNeasy Plus mini Kit* (QIAGEN). Reverse transcription and quantitative PCR were performed in one step using the Luna® Universal One-Step RT-qPCR Kit (NEB) and a CFX Real-Time PCR detection system (Bio-Rad). Beta-2-Microglobulin was used for normalization. qRT-PCR primers are described in Supplementary Table S5.

Protein purification

For GST-tag purification of GST-DDX19A, *E.coli BL21(DE3)* cells were transformed with 50 ng of pGEX-DDX19A plasmid and plated on LB agar with 35 μg/ml Chloramphenicol and 50 μg/ml Kanamycin. Subsequently, 50 ml LB/Kanamycin were inoculated with one colony and the starter-culture was cultivated at 37°C, 150 rpm for 6 h. 500 ml LB/Kanamycin were inoculated with 6 ml of starter culture and cultivated at 37°C, 150 rpm until $OD_{600} = 0.7$. Expression of GST-DDX19A was induced by addition of 500 μM IPTG and overexpression was performed at 20°C, 150 rpm for 14 h. Cells were harvested at $5000 \times g$ for 15 min, 4°C. Pellets were washed once in 30 ml STE buffer (100 mM NaCl, 10 mM Tris–HCl pH 8, 1mM EDTA) and frozen at –20°C until use. For purification, pellets were resuspended in 30 ml sonication buffer (20 mM HEPES pH 7.5, 0.2 mM DTT, 500 mM KCl, 1 mM EDTA, 10% glycerol) with protease inhibitor and lysed by sonication using an EpiShear Probe Sonicator (Active Motif). The lysate was cleared by centrifugation and filtration through a 0.45 μm CHROMAFIL GF/PET-45/25 filter (MACHEREY-Nagel). Affinity chromatography was performed using an NGC™ Chromatography system (BIO-Rad) and Protino® Glutathione Agarose 4B beads (MACHEREY-Nagel). Proteins were eluted in elution buffer (20 mM HEPES pH 7.5, 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, 10% glycerol, 40 mM glutathione) and subjected to dialysis into storage buffer (20 mM HEPES pH 7.5, 200 mM KCl, 0.2 mM DTT, 1 mM EDTA, 10% glycerol). Aliquots were snap-frozen and stored at –80°C. For storage at –20°C, proteins were transferred to a different storage buffer (20 mM HEPES pH 7.5, 200 mM KCl, 0.2 mM DTT, 1 mM EDTA, 60% glycerol) by another round of dialysis.

RNA:DNA unwinding assay

RNA:DNA unwinding assay was performed as described (54). In brief: RNA:DNA hybrids were annealed *in vitro* in 5 mM Tris/HCl pH 8.5. The sequence of the top RNA strand was: 5′-GAAGCUGGGACUUCCGGGAGGAGAGUGCAA-3′, and the sequence of the bottom DNA strand was 5′-CGGGTTGTCAAGAATTTTAACGGCCATTTCTGTGTTGCACTCTCCTCCCGGAAGTCCCAGCTTCTGTGTTTGTGACAAACGCAAGCTCATGTAAGTGCTC-3′. The annealed RNA:DNA hybrid has a 5′ ssDNA overhang and is labeled with Cy-5. Unwinding experiments were carried out at 30°C for 60 min in 30 mM Tris–HCl (pH 7.5), 50 mM NaCl, 5 mM MgCl₂, 2 mM DTT, 0.01% NP-40, 0.1 mg/ml BSA, 4 mM ATP, 1 nM Cy-5-labeled RNA:DNA hybrid substrate, in the presence of 4.79 μM recombinant DDX19A. The reaction was stopped by the addition of SDS to a final concentration of 0.5% and 20 ng proteinase K (20 mg/ml). The reaction was afterward loaded onto a 10% non-denaturing polyacrylamide gel and analysed using a FusionFX detection system (VILBER).

Celluspot and peptide arrays

Peptide arrays containing peptides with a length of 15 amino acids were synthesized by spotting on a cellulose membrane using an Autospot peptide array synthesizer (Intavis AG) and the SPOT synthesis method (55). MODified™ Histone Peptide Arrays (Active Motif) or synthesized peptide arrays were blocked overnight in blocking solution (5% milk powder, 1× PBS, 0.1% Tween20) at 4°C. Both arrays were washed three times for 5 min with 1× PBS/Tween-20 and pre-incubated for 10 min in interaction buffer (100 mM KCl, 20 mM HEPES pH 7.5, 1 mM EDTA pH 8, 10% glycerol). Binding of DDX19A-GST was performed by incubation of 50 nM protein with the pre-blocked array in interaction buffer for 1 h at room temperature. The array was washed three times for 10 min in 1× PBS/Tween20 and incubated with an anti-GST antibody (GE Healthcare, #27-4577-01) in 5% non-fat dried milk/1x PBS/Tween-20 for 1 h at room temperature. The array was washed again as described and incubated with an anti-goat-HRP antibody in 5% milk/1× PBS/Tween-20 for 1 h. After repeated washing, twice for 10 min in PBS/Tween-20 and once for 10 min in PBS, the array was imaged using a FusionFX detection system (VILBER) and SuperSignal™ West Femto Chemiluminescence substrate (ThermoFisher). Synthesized peptide sequences were: H3K4_ARTKQTARKSTGGKA; H3K9 RTKQTARKSTGGKAP; H3K27 LATKAARKSAPATGG; H3K36 APATGGVKKPHRYRP; H4K20 GGAKRHRK VLRDNIQ

Immunofluorescence microscopy

NIH/3T3 cells were cultivated until 70–90% confluency on microscopy coverslips. Cells were washed three times for 5 min with 2 ml PBS$^{Ca2+ Mg2+}$ (Sigma-Aldrich). Cells were fixed for 10 min at room temperature in 4% paraformaldehyde. Cells were washed as described and permeabilized with 0.2% ice-cold TritonX-100 in PBS for 5 min. Cells were blocked in 2 ml 5% non-fat dried milk in PBS$^{Ca2+ Mg2+}$ for

1h at room temperature. Primary antibody binding was performed overnight at 4°C with a concentration of 4 μg/ml antibody in PBS$^{Ca2+ Mg2+}$/5% non-fat dried milk powder. Secondary antibody binding was performed at room temperature for 2 h with a concentration of 0.5 μg/ml antibody in PBS$^{Ca2+ Mg2+}$/5% milk powder. Cells were stained with 1μg/ml DAPI in PBS$^{Ca2+ Mg2+}$ for 3 min, washed again with PBS and mounted on microscopy slides using Mowiol® 4–88 (Sigma-Aldrich).

Image acquisition and analysis

For the quantification of the S9.6 staining, samples were analysed on a Zeiss Axio Observer.Z1 microscope equipped with a Plan-Apochromat 63×/1.40 Oil DIC M27 objective and an AxioCam MRm camera. The following excitation and emission filters were used: Blue channel: excitation filter 335–383 nm, emission filter 420–470 nm; red channel: excitation filter 538–562 nm, emission filter 570–640 nm; green channel: excitation filter 450–490 nm, emission filter 500–550 nm. Z-stacks covering the whole nucleus were acquired applying an interval of 450 nm, and images were subjected to deconvolution using a constrained iterative algorithm and the ZENblue version 2.3 software (Zeiss), before generating maximum intensity projections. Quantitative image analysis was done with CellProfiler™ version 2.2 (56). Nuclei were identified via the DAPI staining.

Co-immunoprecipitation of LSD1

NIH/3T3 were harvested by trypsinization. The pellet was resuspended in 2× pellet volume of nuclear lysis buffer B (50 mM Tris–HCl pH 7.4, 20% glycerol, 1.5 mM MgCl$_2$, 420 mM NaCl, 1 mM Na$_3$VO$_4$, 25 mM NaF) supplemented with Protease inhibitor and incubated on ice for 15 min. Lysate was homogenized with 25 strokes of a douncer (0.01–0.03 mm) and incubated with rotation at 4°C for 30 min. Lysate was cleared by spinning down at 4°C, 16 000 × g, 30min and the supernatant was transferred to a new tube. 1× DP buffer (50 mM Tris–HCl pH 7.4, 5% glycerol, 1.5 mM MgCl$_2$, 150 mM NaCl, 1 mM Na$_3$VO$_4$, 5 mM NaF) with 0.4% NP40 and protease inhibitors (1.8 ml buffer to 1ml lysate) was added and the sample was incubated on ice for 10 min before clearing through ultracentrifugation (30 min, 4°C/43 000 rpm/TI50.2). The supernatant was transferred to a fresh tube and 5% input were taken. The sample was incubated with 1μg anti-KDM1A antibody per 1mg protein in the lysate over night at 4°C. 1.5-fold of loading capacity of Dynabeads protein G were added for 2 h at 4°C with rotation. Beads were washed twice with DP/NP40 buffer, twice with DP buffer and twice with 150 mM NaCl. Proteins were eluted in 30 μl of 2× SDS sample buffer (125 mM Tris–HCl pH 6.8, 5% SDS, 0.004% Bromophenol Blue, 10% β-mercaptoethanol, 100 mM DTT, 20% glycerol) for 10 min at 70°C and subjected to SDS-PAGE and immunodetection as described for gene expression analysis.

Equilibrium peptide binding experiments

Determination of K_D of DDX19A and H3K27 was conducted using H3.1 peptide labeled with FITC. The H3.1 peptide comprising residues 16–34 of the H3.1 tail was unmodified or trimethylated at K27. Binding was analysed using a Jasco FP-8300 spectrofluorometer with an automatic polarizer (FDP-837). Acquisitions were performed at 23°C, with excitation at 495.0 nm and emission measured at 520 nm. Slit width was set to 5 nm. 50 nM of peptide were dissolved in 0.5 ml of anisotropy buffer (20 mM HEPES pH 7.5, 100 mM KCl, 0.1 mM DTT, 10% v/v glycerol). DDX19A diluted in dialysis buffer (20 mM HEPES pH 7.5, 200 mM KCl, 1 mM Na2-EDTA, 60% v/v glycerol, 0.2 mM DTT) was added stepwise. Titrations were conducted in at least 3 technical replicates. Control experiments were conducted with dialysis buffer without protein and the fluorescence anisotropy values were corrected accordingly. For determination of the K_D-values for H3.1, the data were fitted to a simple binding equilibrium:

$$\text{Signal} = \text{BL} + F * \frac{c_{\text{DDX19A}}}{c_{\text{DDX19A}} + K_D}$$

With K_D = equilibrium dissociation constant, F = signal factor and BL = baseline.

Statistical analyses

Results are presented as means ± standard error of the mean [s.e.m.]. Matching sets of samples (treated vs control) were normalized to the average of all samples in this replicate. Values were scaled to the average of all untreated replicates set to average = 1. If not stated otherwise, statistical significance was calculated by one-tailed unpaired t-test on two experimental conditions with p≤0.05 considered statistically significant. Statistical significance levels are denoted as follows: ****$P \leq 0.0001$; ***$P \leq 0.001$; **$P \leq 0.01$; *$P \leq 0.05$; n.s. = non-significant. No statistical methods were used to predetermine sample size.

RESULTS

Development of a novel fluorescent reporter system to investigate LSD1 and associated cofactors in living cells

The transcriptional output of LSD1 is highly dependent on its associated complex partners. In order to measure the activity of LSD1 in a cellular context and in association with its coregulators in a time-resolved manner, we established a novel reporter system, which can be transduced into cell lines of interest (Figure 1 and Supplementary Figure S1). In this system, the expression of a fluorescent reporter protein (*mCherry*) is driven by a synthetic promoter (*synP*) consisting of six tetracycline repressor (TetR) binding elements (*tetO*) introduced upstream of a strong *EF1a* promoter (Figure 1A).

Following transduction and antibiotic selection, the designed reporter construct exhibited a strong *mCherry* fluorescence signal in different cell lines (Supplementary Figure S1A). Next, we generated a fusion construct of full length human LSD1 with the reverse tetracycline repressor protein (rTetR), which was transduced into NIH/3T3 cells expressing the *synP-mCherry* reporter (Figure 1B). This allowed to induce spatial proximity of rTetR-LSD1 to the *synP* element by the addition of Doxycycline (DOX). Af-
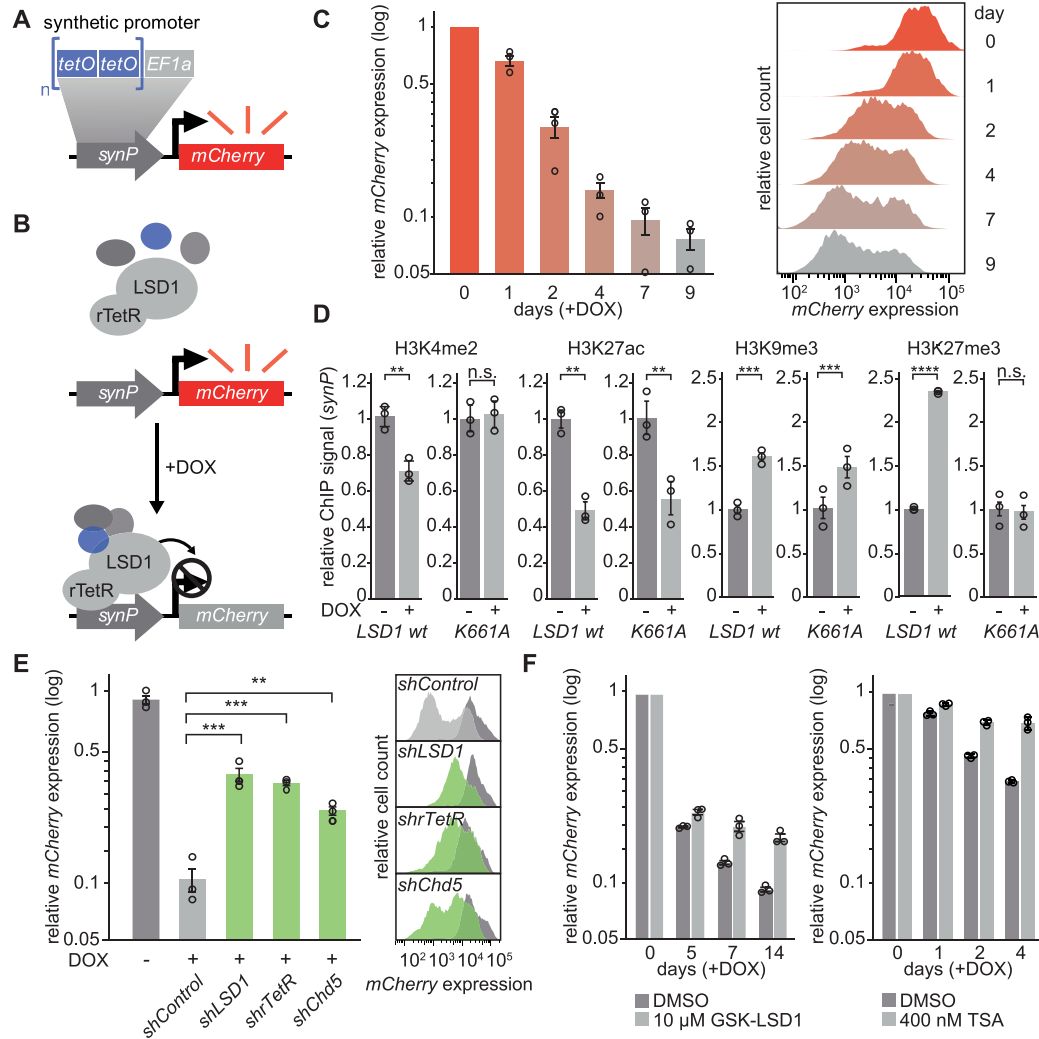
Figure 1. Generation of a fluorescent reporter system to investigate LSD1 and associated cofactors in living cells. (**A**) Illustration depicting the core components of the fluorescent reporter system, which is transduced into cell lines of interest. Stable expression of *mCherry* is driven by a synthetic promoter (*synP*), which consists of 6 Tet repressor binding sites (*tetO*) upstream of an *EF1a* promoter. (**B**) Cell lines expressing the *synP-mCherry* reporter are further transduced with a vector expressing a fusion protein of human LSD1 and the reverse tetracycline repressor protein (rTetR) under a constitutive promoter. Upon DOX treatment, the rTetR-LSD1 fusion protein is recruited to *synP* together with endogenous complex partners, leading to the suppression of *mCherry* expression. (**C**) Flow-cytometric analysis of the *mCherry* expression in NIH/3T3 cells co-expressing the *synP-mCherry* reporter and the rTetR-LSD1 fusion protein after treatment with DOX for the indicated number of days. Left: Bar graphs showing the median *mCherry* signal relative to day 0. Circles represent individual replicates (n = 3, mean±*s.e.m.*). Right: Histograms depicting the distribution of *mCherry* signals of one representative replicate over time (y-axis normalized to highest cell count). (**D**) Bar graphs depicting changes of the indicated histone marks at the *synP* element, analysed by ChIP-qPCR. IPs were performed with mononucleosomes isolated from reporter cell lines at day 4 (H3K4me2, H3K27ac, H3K27me3) or day 14 (H3K9me3) of either LSD1 wt or K661A recruitment. Bar graphs are relative to -DOX. Circles represent independent replicates (n = 3, mean ± s.e.m.; **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$, n.s. = non-significant; Student's *t*-test). (**E**) Flow-cytometric analysis of the *mCherry* signal in NIH/3T3 cells expressing the *synP-mCherry* reporter, rTetR-LSD1 and the indicated shRNAs at day 7 of DOX treatment. Dark grey: -DOX, light grey: neutral control shRNA, green: positive control shRNAs +DOX. Left: Bar graphs show the median *mCherry* signal relative to day 0. -DOX is shown for the control shRNA (*shControl*). Circles represent independent replicates (n = 3, mean±*s.e.m.*; **$P \leq 0.01$, ***$P \leq 0.001$; Student's *t*-test). Right: Histograms showing the *mCherry* expression profiles at day 7 ±DOX of one representative replicate. (**F**) Bar graphs depicting the median *mCherry* expression of +DOX NIH/3T3 reporter cells relative to -DOX in the presence of GSK-LSD1, TSA or DMSO. Treatment of cells was started in parallel with the addition of DOX and maintained for the indicated number of days. Circles represent independent replicates (n = 3, mean ± s.e.m.).

terwards, LSD1 mediated effects on the *synP-mCherry* reporter gene expression over time can be detected by flow cytometry or fluorescence microscopy (Figure 1B, C and Supplementary Figure S1C, D). ChIP analysis revealed dynamic changes of the chromatin environment at the *synP* promoter element after recruitment of the rTetR-LSD1 fusion protein. We observed a mild reduction in histone H3

lysine 4 dimethylation (H3K4me2) (Figure 1D), which is in agreement with recent literature, stating that the activity of LSD1 is highly dependent on coregulatory effector proteins (22) and probably heavily counterbalanced by H3K4 specific methyltransferases (KMTs) (24,57). Furthermore, this can be explained by the fact that H3K4me2 is an intermediate mark and we observed co-recruitment of Kdm5b (see

below), which is an H3K4me3 demethylase that continuously generates H3K4me2 at the locus. Histone H3 lysine 27 acetylation (H3K27ac), another modification characteristic for active chromatin, showed a more pronounced reduction, whereas histone H3 lysine 9 trimethylation (H3K9me3) and H3K27me3, both associated with inactive genomic regions, were increased (Figure 1D). In contrast, the recruitment of rTetR alone to the *synP-mCherry* reporter induced no change in reporter gene expression (Supplementary Figure S1C). Similarly, the recruitment of a catalytically inactive mutant of LSD1 (K661A) (58) to the *synP* promoter did not lead to a strong reduction of fluorescent reporter gene expression when compared to the LSD1 wt (Supplementary Figure S1C). We also investigated the consequences of K661A recruitment on selected histone modifications at the *synP* promoter element. As expected, recruitment of K661A did not lead to a reduction of H3K4me2, whereas changes in H3K27ac or H3K9me3 were comparable to those induced by recruitment of the LSD1 wt (Figure 1D). Interestingly, we did not observe an increase in H3K27me3 when recruiting LSD1 K661A (Figure 1D). These results demonstrated that the observed change in reporter fluorescence is a direct effect of LSD1 recruitment, actively changing the chromatin environment at the promoter and indicate that known coregulators of LSD1, like HDACs and lysine methyltransferases (KMTs) such as G9a/GLP and PRC2, are co-recruited and active at the *synP* element. Furthermore, we observed a mechanistic connection between the enzymatic activity of LSD1 and its ability to induce a strong reduction in gene expression and an increase in H3K27me3, the latter probably related to the inability of PRC2 to methylate histone H3 methylated at K4 (59,60).

To test if the reporter system was sensitive to perturbations and could thus be applied to study the influence of LSD1 associated coregulators, we suppressed the expression of the rTetR-LSD1 fusion construct using shRNAs (Supplementary Data Table S4) targeting the rTetR or LSD1 (Supplementary Figure S1B) parts of the rTetR-LSD1 fusion protein and of the known LSD1 complex partner *Chd5* (61). Silencing of any functional part or *Chd5* resulted in a substantially impaired ability of LSD1 to induce effective silencing of the *synP-mCherry* reporter (Figure 1E). Furthermore, treatment with the LSD1 inhibitor GSK-LSD1 or the pan-HDAC inhibitor Trichostatin A (TSA) demonstrated that the activity of rTetR-LSD1 at the *synP* element is dependent on the catalytic activity of LSD1 and HDACs (Figure 1F and Supplementary Figure S1E). Consequently, silencing of *mCherry* expression by recruitment of rTetR-LSD1 to *synP* is conditioned by the presence and activity of additional endogenous coregulators.

A multiplexed shRNAmir screen identifies essential and novel LSD1 coregulators

We applied our novel reporter system for LSD1 activity to systematically probe a comprehensive selection of chromatin coregulators for their requirement to enable LSD1-mediated silencing. To this end, we screened a focused shRNA library targeting 1010 chromatin-associated murine genes (4–6 shRNAs per gene) in a multiplexed format in NIH/3T3 cells expressing the *synP-mCherry* reporter and the rTetR-LSD1 fusion protein (Figure 2A and Supplementary Table S1). After an initial antibiotic selection for successful integration of the constitutively expressed shRNA constructs, we induced recruitment of the rTetR-LSD1 fusion protein to the *synP* element by the addition of DOX. During subsequent cultivation for 14 days under constant treatment with DOX, cells expressing effective shRNAs targeting regulators of LSD1, which are critically required for LSD1-mediated gene silencing accumulated in a cell population that showed persistent expression of *mCherry* (positive population). Using FACS, these cells were separated from the major population, which exhibited the usual reduction in fluorescent reporter signal (negative population) and expressed ineffective shRNAs (Figure 2A and Supplementary Figure S2A). The representation of each shRNA in the input library and in the sorted positive and negative cell populations was quantified using deep-sequencing of the shRNA guide strands amplified from genomic DNA using established protocols (Supplementary Figure S2B and Table S1) (62). To rank all genes represented in the shRNA library for their effect on LSD1 activity, gene scores reflecting the enrichment of multiple shRNAs per gene in the positive cell population compared to the negative population were calculated (Figure 2B, and Supplementary Table S2). As expected, the screen managed to identify coregulators of LSD1 that were already described as complex partners in the literature (Figure 2B) and suppression of several genes known to be associated with LSD1, e.g. *Dnmt3a* (63), *Gatad2b* (NuRD) (64) and *Sap25* (SIN) (65,66), were confirmed to be especially important for LSD1-mediated gene silencing. Interestingly, the H3K4me3 demethylase Kdm5b was ranked at position 20 in the screen, suggesting that LSD1 requires the demethylase activity of KDM5B (67) to generate K4me2, which then can be demethylated further by LSD1 leading to stable silencing of the *synP-mCherry* reporter in NIH/3T3 cells. In addition to already known factors, the screen also identified novel coregulators not associated with LSD1 biology so far (Figure 2B and Supplementary Figure S2B).

We selected the top 100 positive and negative hits identified in the screen and performed an enrichment analysis using ClueGO (68) to visualize functionally related coregulators influencing LSD1 activity as a clustered network of the associated Gene Ontology (GO) pathways (Figure 2C). These data show that LSD1 mainly cooperates with proteins linked to biological pathways associated with lysine methylation including pathways regulating H3K4, K9 or K27 methylation, which constitutes the largest cluster in the network (Figure 2C). Other clusters comprise pathways associated with chromatin silencing, lysine acetylation and chromatin remodeling. Furthermore, the network is enriched for pathways related to DNA damage response and nuclear receptor signaling, functions which were already described to be regulated by LSD1 (7,17,32).

Interestingly, one cluster in the network is connected to conformational changes of DNA (Figure 2C) and indeed the screen identified three ATP-dependent RNA helicases among the top 10 positive and negative coregulators of LSD1 (Figure 2B). DDX39B, DDX23 and DDX19A belong to the so-called DEAD-box family of RNA-dependent ATPases that have RNA unwinding activity and are in-
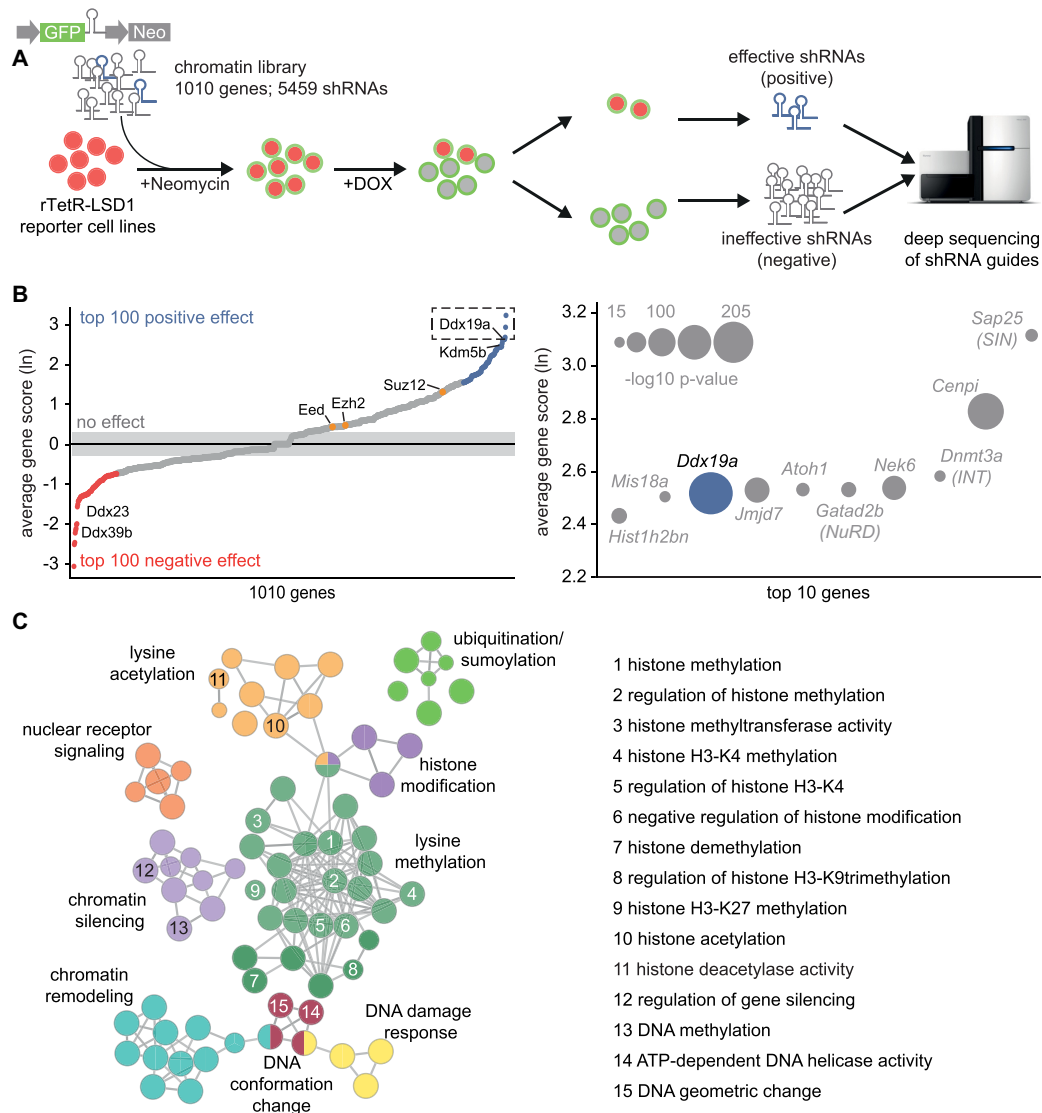
Figure 2. A chromatin-focused shRNA screen identifies novel and known coregulators of LSD1 activity. (**A**) Workflow describing the ChECS screening strategy. A library composed of 5459 shRNAs targeting 1010 chromatin-related genes (*GFP+*) was virally transduced into rTetR-LSD1 reporter cell lines (*mCherry+*). After antibiotic selection, cells were treated with DOX for 14 days and FACS sorted for high (*mCherry+*) or low *(mCherry-)* mCherry expression. Genomic DNA was isolated from both populations and the shRNA guide sequences were amplified for Illumina sequencing. (**B**) Scatter plots ranking all genes according to their effect on LSD1 activity (gene score). Left: Gene scores of all genes present in the shRNA library. The gene score represents the Ln of the average enrichment score (read ratio *mCherry+/mCherry-*) of all shRNAs per gene across five replicates. Genes imposing a positive effect on LSD1 induced silencing are coloured in blue, genes having a negative effect are highlighted in red. The position of Ddx19a, Kdm5b and the PRC2 core components Suz12, Ezh2 and Eed is indicated. Right: Top ten genes identified in the screening procedure to positively influence LSD1 activity. Significance is represented by spot size ($-\log_{10}$ *P*-value). (**C**) *ClueGo* network clustering the top 100 positive and negative regulators of LSD1 for their biological function (GO-annotated biological process). Biological processes of selected clusters are highlighted on the right. The statistical test used for the enrichment was based on a two-sided hypergeometric test with a Bonferroni correction and kappa score of 0.4. Only pathways with p≤0.01 are shown.

volved in pre-mRNA splicing, mRNA export from the nucleus to the cytoplasm or translation (69–73). Loss of DDX39B and DDX23 promoted silencing of the *synP* element, likely due to LSD1 unrelated effects caused by deregulation of the transport, splicing or translation of components of the reporter system, functions reported for both helicases previously (72,74). Among the three DEAD-box helicases identified as coregulators, DDX19A was the only helicase identified as a positive regulator of LSD1 mediated silencing and it scored with the highest significance in the screen (Figure 2B). The homolog of DDX19A, DDX19B has recently been shown to be involved in the removal of RNA:DNA hybrid structures (so-called R-loops) and the activity of DDX19B was shown to be dependent on the DNA damage response induced by the ATR-Chk1 pathway (54).

R-loops are highly dynamic structures that occur at different regions in the eukaryotic genome and exhibit critical regulatory functions during replication, transcription, and recombination (75–78). R-loops preferentially form at

GC-rich regions, where the newly synthesized G-rich RNA hybridizes to the C-rich DNA template (79). They have been described to be associated with both up- and down-regulation of transcription (80) and occur at unmethylated human CpG island promoters (81). Interestingly, R-loops have been shown to colocalize with H3K4 methylation on a genome-wide scale (82).

Aiming to characterize the role of DDX19A in LSD1 induced silencing of gene expression, we further investigated the effects of *Ddx19a* suppression on R-loop regulation and gene expression on a global and local level. Two top-scoring shRNAs from our screen (*shDdx19a*.1/2) showed only mild effects on cell viability, suppressed DDX19A expression (Supplementary Figure S2C, D) and were validated to interfere with LSD1 silencing activity at the *synP* promoter (Supplementary Figure S2E). We next sought to determine the mechanism by which DDX19A influences LSD1-induced silencing and analysed rTetR-LSD1 protein levels after suppression of *Ddx19a* to rule out that the effect of *Ddx19a* is merely a consequence of reduced expression or defects in mRNA processing of the rTetR-LSD1 fusion protein. However, we did not observe any reduction in expression of the rTetR-LSD1 fusion protein and no alteration of the expression of the *synP-mCherry* reporter without the addition of DOX, which could lead to a false-positive result and be responsible for the observed remaining fluorescence signal (Supplementary Figure S2F, G). To confirm that the negative effect of *Ddx19a* suppression of LSD1-mediated silencing is specific to LSD1 function, we investigated the effects of *Ddx19a* suppression on the repressive activity of KRAB, which we recruited to the *synP* element using a rTetR-KRAB fusion protein. Suppression of Ddx19a expression did not influence the activity of KRAB, suggesting that DDX19A is not a general requirement for gene silencing and the observed function is specific to LSD1 (Supplementary Figure S2H).

DDX19A has not been described as an interactor of LSD1 (83,84) and to test if LSD1 recruits DDX19A by a direct interaction, we performed co-immunoprecipitation (co-IP) experiments from NIH/3T3 cell lysate. Although other known complex partners of LSD1 like HDAC1 co-precipitated with LSD1, we could not detect a direct interaction of LSD1 and DDX19A (Supplementary Figure S2I), which could be due to low abundance of DDX19A or a weak interaction with LSD1. Thus, we aimed to further characterize the effects of *Ddx19a* suppression on LSD1 activity in more detail.
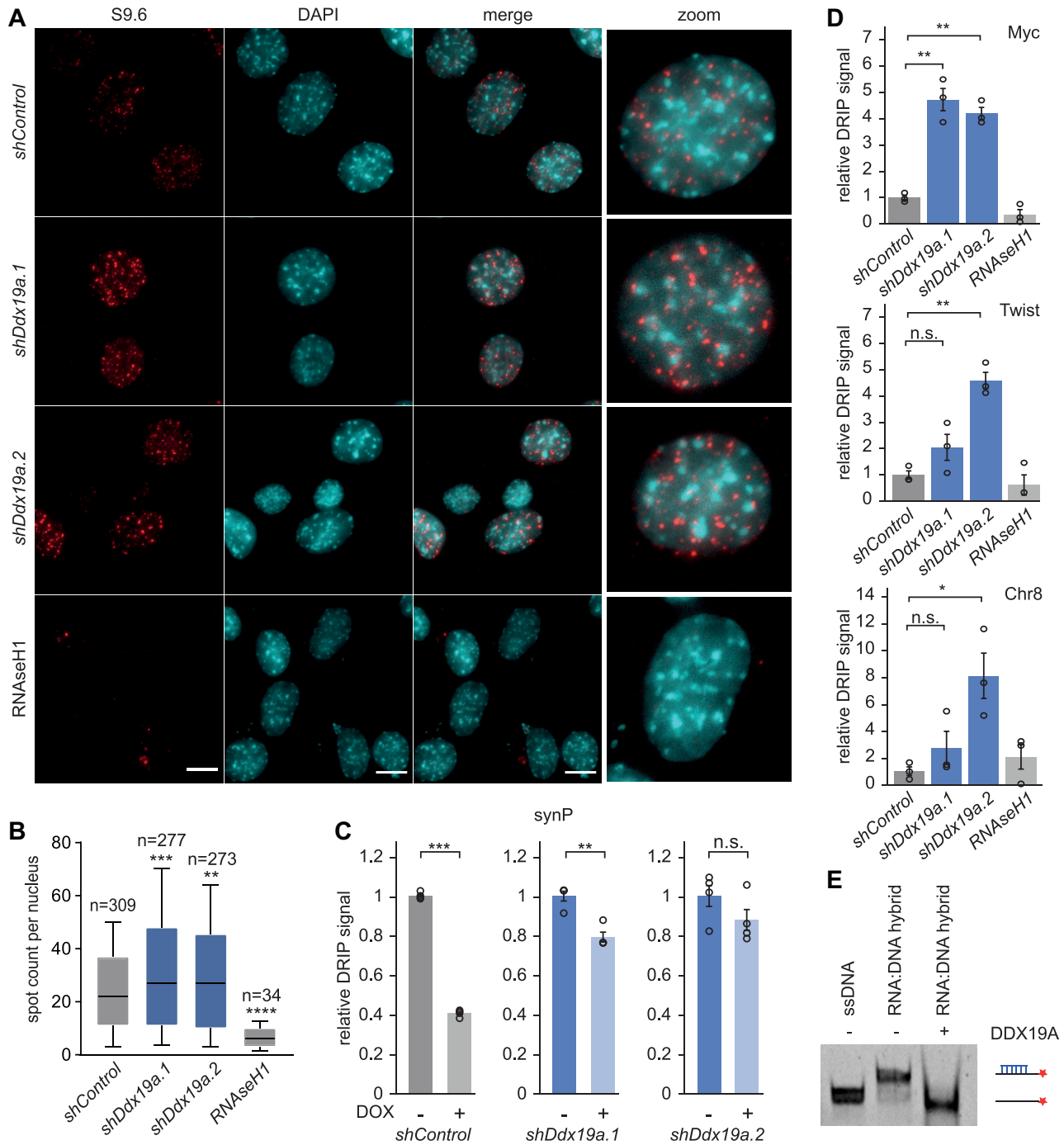
DDX19A is involved in R-loop homeostasis

Since the homologue of DDX19A, DDX19B, actively participates in the removal of R-loops (54), we investigated the dynamics of R-loop formation in NIH/3T3 cells after suppression of *Ddx19a* using immunofluorescence microscopy. We employed the S9.6 antibody, which specifically recognizes R-loops (85). To test for the specificity of the signal detected with the S9.6 antibody, we used Ribonuclease H1 (RNAseH1), known to specifically degrade the RNA of RNA:DNA hybrid structures. As shown in Figure 3A, transient expression of RNAseH1 prior to the immunostaining with the S9.6 antibody led to a strong reduc-

tion of the R-loop signal (Figure 3A). Upon suppression of *Ddx19a* we noticed a significant increase in R-loop spot counts per nucleus compared to a neutral control shRNA (Figure 3A, B and Supplementary Figure S3A) as well as an enhancement of R-loop spot intensity (Figure 3A and Supplementary Figure S3A, B). Subsequently, we studied the alterations of R-loops at the *synP* element before and after DOX induced LSD1 recruitment. As expected, R-loops at the *synP* promoter element were reduced following DOX induced recruitment of LSD1 and this effect was strongly attenuated under suppression of *Ddx19a* (Figure 3C). Furthermore, we investigated R-loop dynamics at representative endogenous genomic loci (Supplementary Figure S3E) upon suppression of *Ddx19a* expression. Two regions are associated with the developmental master regulators Myc (86,87) and Twist1 (88,89), which are highly expressed and associated with extensive H3K4 methylation and R-loop signal. In addition, we investigated R-loop dynamics at an intergenic region on chromosome 8 (Chr8) characterized by very low R-loop and high H3K27me3 signal (Supplementary Figure S3E). Using DNA-RNA immunoprecipitation (DRIP) followed by qPCR, we observed an increase in R-loops upon *Ddx19a* suppression in all cases (Figure 3D) showing that also master regulators like Myc and Twist1 as well as regions decorated with high levels of H3K27me3 respond in a similar way to suppression of *Ddx19a* expression as the *synP* element of our artificial reporter construct. As a control, we included an RNAseH1 incubation step, which reduced the DRIP signal confirming the specificity of the antibody (Figure 3D and Supplementary Figure S3C, D). This is in line with our observations in the immunofluorescence experiments and implies that the effect of suppression of *Ddx19a* expression on R-loops is not restricted to our artificial reporter construct but also occurs at endogenous regions.

In addition, we purified recombinant DDX19A and confirmed its ability to resolve RNA:DNA hybrids *in vitro* (Figure 3E and Supplementary Figure S3F). Our data indicate that suppression of Ddx19a expression leads to the accumulation of R-loops and interferes with the silencing activity of LSD1 at the *synP-mCherry* reporter. This observation is in agreement with the model that R-loops forming over the *synP* element stabilize its transcriptional activity, which is supported by recent publications showing that R-loops impose various effects on chromatin regulators to promote transcription (81,82,90–93) including the inhibition of LSD1 (38) and PRC2 (39,40).

LSD1 and R-loops colocalize and occupy regions associated with highly transcribed genes

We wanted to explore if the interaction of LSD1 and DDX19A could be mediated through R-loops. In order to investigate if LSD1 localizes to genomic regions decorated with R-loops, we analysed the genome-wide distribution of LSD1 in two different cell lines (K562 and NIH/3T3) and compared it to features associated with R-loops and histone modifications signaling either active or repressed gene expression (Figure 4 and Supplementary Figure S4). To this end, we performed an LSD1 ChIP-seq in NIH/3T3 cells, employed publicly available LSD1 ChIP-seq data from

Figure 3. DDX19A is involved in R-loop homeostasis. (**A**) Representative immunofluorescence microscopy images of NIH/3T3 cells expressing the indicated shRNAs for 10 days and stained for R-loops with the S9.6 antibody. Transfection with RNAseH1 24 h before fixation serves as a control for the S9.6 antibody specificity. Images are maximum intensity projections of a Z-stack covering the whole nucleus. Scale bars are 10 μm. (**B**) Box and whisker plot showing the quantification of the R-loop spot count per nucleus in images from (**A**). Images were analysed using CellProfiler™ software. Box-and-Whisker plots indicate the median and the 10–90 percentile from three independent experiments (**$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$, Student's t-test). (**C**) DRIP-qPCR analysis of RNA:DNA hybrid structures at the *synP* element upon recruitment of rTetR-LSD1 and under suppression of *Ddx19a*. Total nucleic acids were extracted from NIH/3T3 cells expressing the *synP-mCherry* reporter, the rTetR-LSD1 fusion protein and the indicated shRNAs at day 14 with and without DOX treatment and used as input for the IP with the S9.6 antibody. qPCR signals are shown relative to -DOX. Circles represent independent replicates ($n = 4$, mean ± s.e.m.; **$P \leq 0.01$, ***$P \leq 0.001$, n.s. = non-significant, Student's t-test). Pre-treatment of the input samples with RNAseH1 before DRIP was used as a negative control. (**D**) DRIP-qPCR analysis of RNA:DNA hybrid structures at selected endogenous loci under suppression of *Ddx19a* expression. Total nucleic acids were extracted from NIH/3T3 cells expressing the *synP-mCherry* reporter, the rTetR-LSD1 fusion protein and the indicated shRNAs for 10 days and used as input for the IP with the S9.6 antibody. qPCR signals are shown relative to *shControl*. Circles represent independent replicates ($n = 3$, mean ± s.e.m.; *$P \leq 0.05$, **$P \leq 0.01$, n.s. = non-significant, Student's t-test). (**E**) Representative result of an *in vitro* RNA:DNA hybrid unwinding assay using recombinant DDX19A. On the right of the image, the Cy5-labeled single strand DNA (ssDNA) and the RNA:DNA hybrid substrate are shown schematically. The RNA is colored blue, the DNA black and the Cy5 label is indicated with a red star.
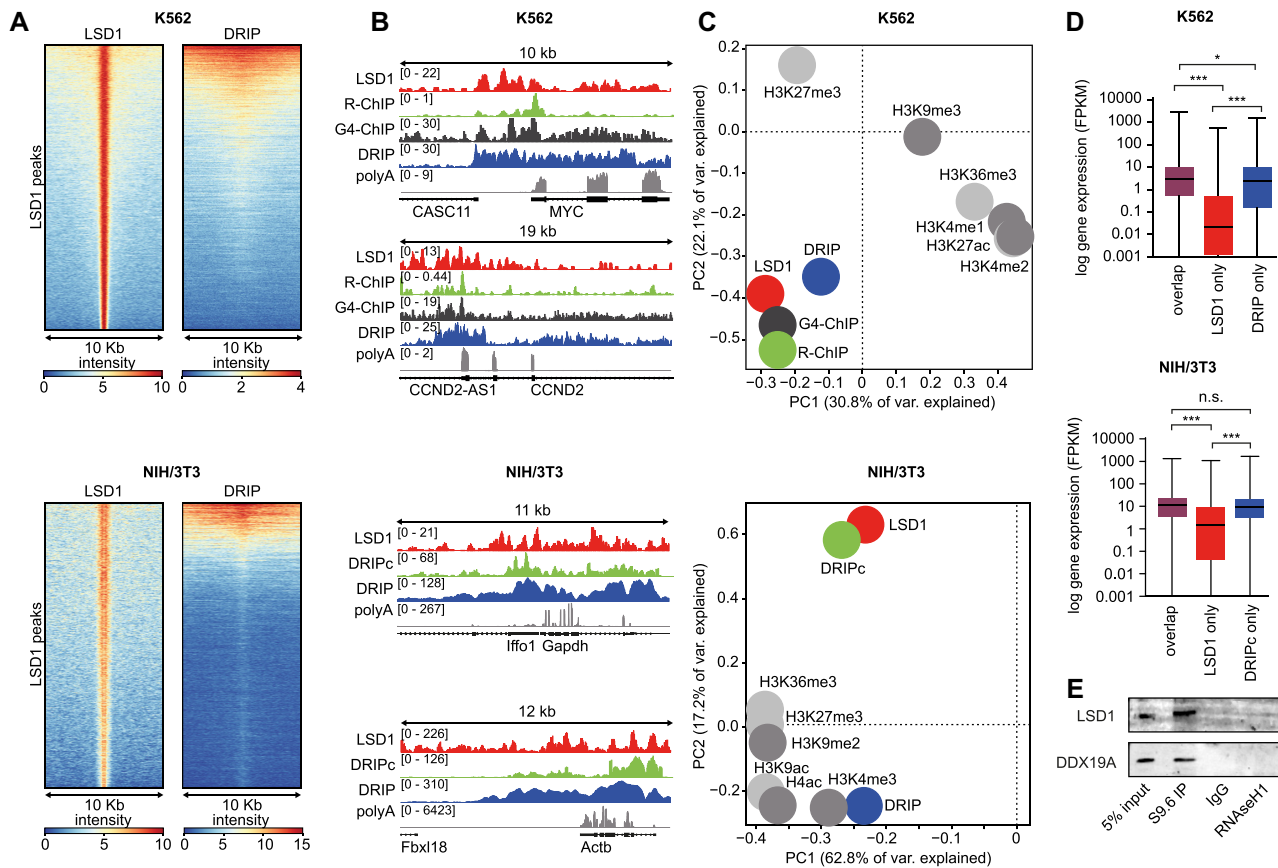
Figure 4. Regions of LSD1 occupancy and R-loop related features correlate genome wide in K562 and NIH/3T3 cells. (**A**) Heatmaps of LSD1-ChIP and DRIP signals in K562 (top) and NIH/3T3 cells (bottom). Signals are plotted on the heatmap within a 10 kb window around the peak centre. (**B**) Representative regions showing the occupancy of R-loop related features and LSD1 in K562 and NIH/3T3 cells. (**C**) Principle component analysis of deep sequencing data for K562 (top) and NIH/3T3 (bottom) cells. ChIP-seq, DRIP-seq, DRIPc-seq, R-ChIP-seq and G4-ChIP-seq datasets were obtained from the Gene Expression Omnibus (GEO) repository and the ENCODE project. R-loop related features are highlighted in blue, green and black. R-loop features analysed in panel B and C: RNA:DNA hybrid IP (DRIP); RNA:DNA hybrid IP followed by cDNA conversion (DRIPc); RNA:DNA hybrid IP using a catalytically inactive RNAseH1 (R-ChIP); Genome-wide mapping of endogenous G-quadruplex DNA structures (G4-ChIP). (**D**) Expression levels (mRNA, FPKM) of LSD1- and DRIP/DRIPc-associated genes in K562 and NIH/3T3 (*$P \leq 0.05$, **$P \leq 0.001$, n.s. = non-significant, ordinary one-way ANOVA with multiple comparisons test). (**E**) Representative Western blot showing co-precipitation of LSD1 and DDX19A with R-loops after IP using the S9.6 antibody. NIH/3T3 whole cell lysate was used as input. Pulldown with mouse IgG and treatment with RNAseH1 before IP with the S9.6 antibody was used as control.

K562 cells (94) and compared those to available DRIP-seq datasets that investigated the genome-wide distribution of R-loops using the S9.6 antibody (82). This analysis revealed a considerable colocalization of LSD1 with R-loops in both cell lines (Figure 4A). We further investigated the colocalization of LSD1 with additional R-loop features and used publicly available data of an IP with a catalytically dead RNAseH1 mutant that specifically binds to R-loops (R-ChIP-seq) (95) or a G-quadruplex (G4) specific antibody (G4-ChIP-seq) (96) and compared these with ChIP-seq datasets of different histone modifications from the Gene Expression Omnibus (GEO) repository (97) and the ENCODE project (98) for both cell lines (see Materials and Methods section for the corresponding GO numbers). Surprisingly, LSD1 binding showed a much stronger correlation with R-loop related features than with histone modifications previously reported to associate with LSD1, like methylated H3K4 (22) in both cell lines (Figure 4B, C and Supplementary Figure S4A, B).

Expression of the genes showing LSD1 binding and/or R-loop features was retrieved from previously published RNA-seq analysis for K562 (62) and publicly available RNA-seq data for NIH/3T3 cells from ENCODE. A detailed analysis of the expression of genes positioned in the vicinity of either LSD1 binding or R-loop structures showed that genes which were solely bound by LSD1 are expressed at a very low level. In contrast, loci which are only decorated with features of R-loops were associated with genes exhibiting high expression levels in both cell lines (Figure 4D and Supplementary Figure S4C–E). However, genomic regions covered with LSD1 and R-loops together were affiliated with genes expressed at an even higher level than genes associated with R-loops only (Figure 4D and Supplementary Figure S4C–E). Next, we performed a gene ontology (GO) enrichment analysis of these genes for the K562 cell line to analyse these genes according to their annotated biological process. We identified 'negative regulation of erythrocyte differentiation' among the top 3 en-

riched pathways ranked for the respective *P*-value (Supplementary Figure S4F), which is in line with the fact that the chronic myelogenous leukaemia (CML) cell line K562 bears resemblance with undifferentiated erythrocytes (99).

To support these global observations, we performed an immunoprecipitation of R-loops from NIH/3T3 cell lysate using the S9.6 antibody, followed by immunodetection of LSD1 and DDX19A, which indicated a strong physical association of both proteins with R-loops. Co-precipitation of LSD1 or DDX19A was attenuated if the sample was treated with RNAseH1 before the IP (Figure 4E). These findings imply that LSD1 and R-loops co-occupy highly transcribed genes, which are important to control cell type-specific developmental gene expression profiles and define cellular identity. This is in agreement with previous reports showing that, despite its mostly repressive function, LSD1 occupies enhancers and core promoters of a substantial fraction of actively transcribed genes (28) and suggests that the activity of LSD1 at these regions is repressed.

DDX19A plays a role in LSD1 induced silencing

To gain additional insights into the local consequences of R-loop accumulation, in particular the effects on downstream effector proteins of LSD1, we analysed histone modifications at the *synP* promoter element after suppression of *Ddx19a*, which led to an incomplete silencing of the *synP-mCherry* reporter after recruitment of LSD1 (Figures 2B, Supplementary Figure S2E and 5A). Despite a significantly diminished reduction of H3K4me2 (Figure 5B), which was comparable to the level induced by the catalytically inactive LSD1 mutant K661A (Figure 1D), we did not observe any alteration in H3K9me3 and H3K27ac histone modifications after suppression of *Ddx19a* expression compared to *shControl* (Figure 5B). In contrast, we noticed a distinct reduction in the increase of H3K27me3 signal at the *synP* element upon LSD1 recruitment and suppression of *Ddx19a* (Figure 5B). These results suggest, that the increase in R-loops at the *synP* promoter (Figure 3C) influences the removal of H3K4me2 and/or deposition of H3K27me3, thereby interfering with efficient reduction of gene expression. This effect can be explained by the inhibition of LSD1 (38) and PRC2 (39,40) via RNA that is anchored at the locus through R-loops. This interpretation is in line with the fact that suppression of *Ddx19a* expression did not affect the residual silencing activity of K661A (Figure 5C). Interestingly, we observed a substantial increase in R-loops at the *synP* element and representative endogenous regions upon treatment with GSK-LSD1, a selective LSD1 inhibitor (Figure 5D and Supplementary Figure S5), demonstrating that inhibition of LSD1 activity results in a gain of R-loops not only at our artificial promoter element but also representative endogenous loci, which suggests a mutual influence of LSD1 and R-loops on one another.

DDX19A binds to defined histone modifications associated with gene repression

Investigating the downstream consequences of impaired H3K4 demethylation at the *synP* promoter element, we noticed that the K661A inactive mutant was ineffective in inducing a gain of H3K27me3 signal in contrast to the LSD1 wt (Figure 1D). This led to the conclusion that the reduction of H3K4me2 allows PRC2 to bind to the *synP* promoter (59,60). The fact that suppression of *Ddx19a* expression did not affect residual gene silencing by the K661A mutant suggests that the activity of LSD1 is a prerequisite for DDX19A to facilitate a robust reduction of transcription. While, the exact mode how DDX19A is recruited to the *synP* promoter element remained unknown, the previous results suggested that the gain of H3K27me3 seemed to play an important role in this process (Figures 1D and 5B). Searching for the underlying mechanism by which LSD1 activity can induce the recruitment of DDX19A, we screened the binding properties of recombinant full-length DDX19A (Supplementary Figure S3F) to 384 different histone peptides containing 59 post-translational modifications of the N-terminal tails of histones H2A, H2B, H3 and H4. Although DDX19A does not contain any of the known domains specific for lysine methylation binding, we observed a remarkably distinct interaction of DDX19A with H3K27me3 and to a lesser extent H4K20me3 among all other tested histone modifications (Figure 6A, B and Supplementary Figure S6A). We validated this effect with additional modified histone tail peptide arrays containing selected methylated and unmethylated lysine residues on H3 and H4 (Supplementary Figure S6B). Using fluorescence anisotropy (FA), we determined the dissociation constant (K_D) of DDX19A with the H3K27me3 peptide with 82 ($\pm$2.9) nM while the corresponding unmethylated peptide showed a much weaker affinity with a K_D >1000 nM (Figure 6C).

The role of H3K27me3 in the recruitment or stimulation of DDX19A was confirmed by inhibition of the core PRC2 complex component EZH2 using an EZH2-specific inhibitor (GSK343), which induces a global reduction of H3K27me3 signal (Supplementary Figure S6C). Indeed, we observed increased R-loop occupancy at the *synP* element and representative endogenous loci after GSK343 treatment confirming the reduced activity of DDX19A at the *synP* element (Figure 6F, Supplementary Figures S3D and S6D). As a consequence of reduced DDX19A binding to H3K27me3 and elevated R-loop occupancy, GSK343 treatment resulted in a diminished LSD1 silencing capacity at the *synP* element (Figure 6D) comparable to the activity under suppression of DDX19A (Figure 5A). Furthermore, K661A did not reduce R-loops at the *synP* element as efficiently as the LSD1 wt (Figure 6E) after recruitment to the *synP* element, likely due to the lack of H3K4 demethylation preventing deposition of H3K27me3. As expected by the reduced DDX19A activity or recruitment under these conditions, residual silencing of the *synP* element by LSD1-K661A was not affected by the suppression of *Ddx19a* expression (Figure 5C). We conclude that the DEAD-box helicase DDX19A specifically interacts with histone modifications signaling repressive chromatin states, preferably H3K27me3, which is introduced by PRC2 upon reduction of H3K4 methylation by LSD1 (Figure 1D). This binding precedes the ATP-dependent helicase activity of DDX19A regulating the formation of R-loops (Figure 6D, E).
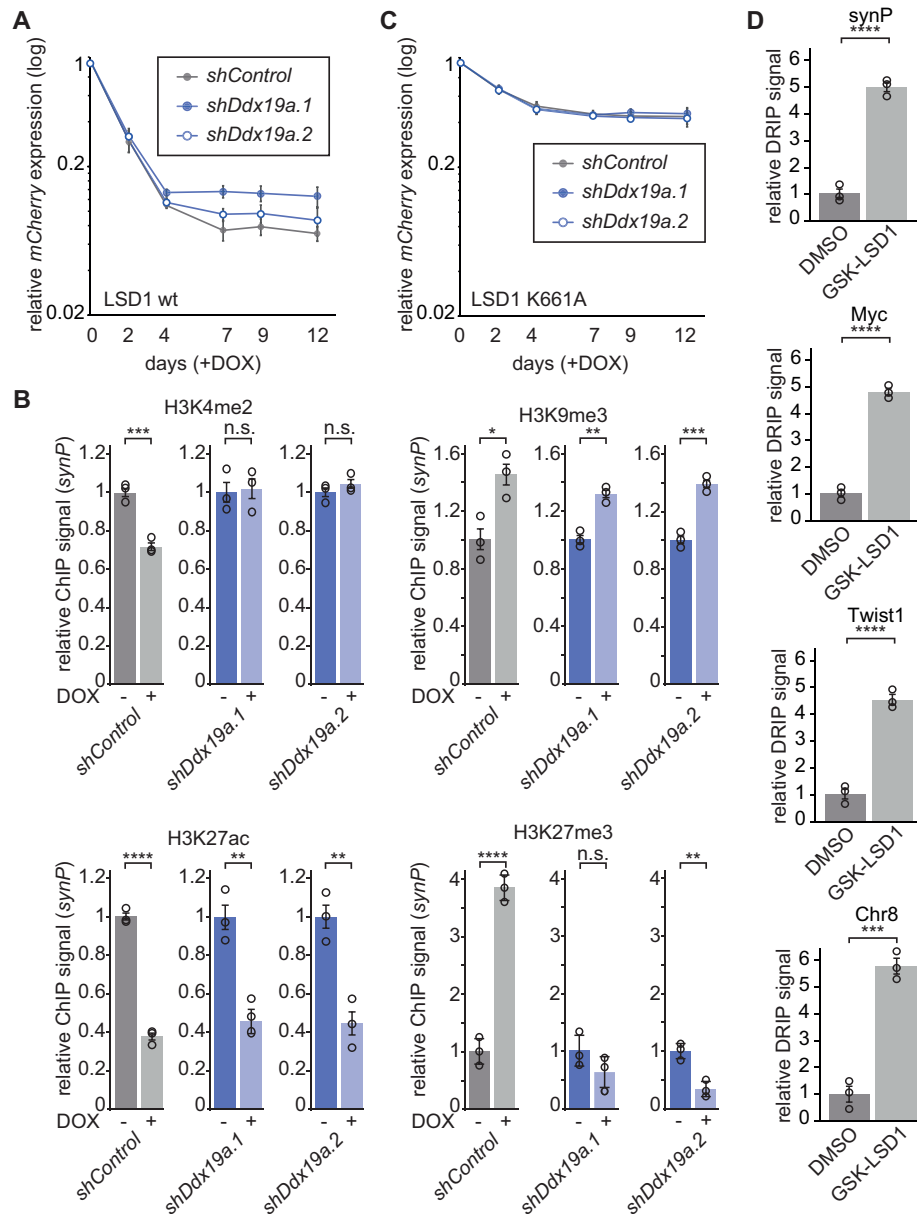
Figure 5. LSD1 and DDX19A act in concert to regulate transcription. (**A**) Time course of the *mCherry* expression in NIH/3T3 expressing the *synP-mCherry* reporter, the rTetR-LSD1 fusion protein and the indicated shRNAs under DOX treatment. Circles indicate the median *mCherry* expression measured by flow cytometry relative to the initial measurement ($n = 3$; mean $\pm$ s.e.m.). (**B**) ChIP-qPCR analysis of the indicated histone modifications at the *synP* element in cells expressing the LSD1 reporter system and the indicated shRNAs. IPs were performed with mononucleosomes isolated from NIH/3T3 reporter cell lines after 4 days (H3K4me2, H3K27ac) or 14 days (H3K9me3) of LSD1 recruitment. Bars are relative to -DOX ($n = 3$; mean $\pm$ s.e.m.; *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$, n.s. = non-significant; Student's *t*-test). (**C**) Time course of the *mCherry* expression in NIH/3T3 cells expressing the *synP-mCherry* reporter, the rTetR-K661A fusion protein and the indicated shRNAs under treatment with DOX. Circles indicate the median *mCherry* expression measured by flow cytometry relative to day 0 ($n = 3$; mean $\pm$ s.e.m.). (**D**) DRIP-qPCR analysis of RNA:DNA hybrid structures at selected endogenous loci under treatment with GSK-LSD1. Total nucleic acids were extracted from NIH/3T3 cells transduced with the *synP-mCherry* reporter after 3 days of treatment with 50 μM GSK-LSD1 or DMSO and used as input for the IP with the S9.6 antibody. qPCR signals are shown relative to DMSO. Circles represent independent replicates ($n = 3$, mean $\pm$ s.e.m.; ***$P \leq 0.001$, ****$P \leq 0.0001$, Student's *t*-test).

DISCUSSION

Transcription is regulated by a complex network of coregulators often with opposing functions, which act in concert to fine tune gene expression. LSD1 has been demonstrated to be an important regulator of developmental genes in embryonic stem cells and malignant cells (9,28) and it was shown to interact with a large variety of different coregula-

tor complexes. The mechanisms underlying the precise regulation of the pluripotency program and its response to developmental cues are still relatively unknown. The tightly controlled balance of opposing chromatin effector functions such as KMTs, lysine demethylases (KDMs) or histone acetyltransferases (HATs) and HDACs (100,101) constitutes a feasible mechanism to ensure developmental plas-
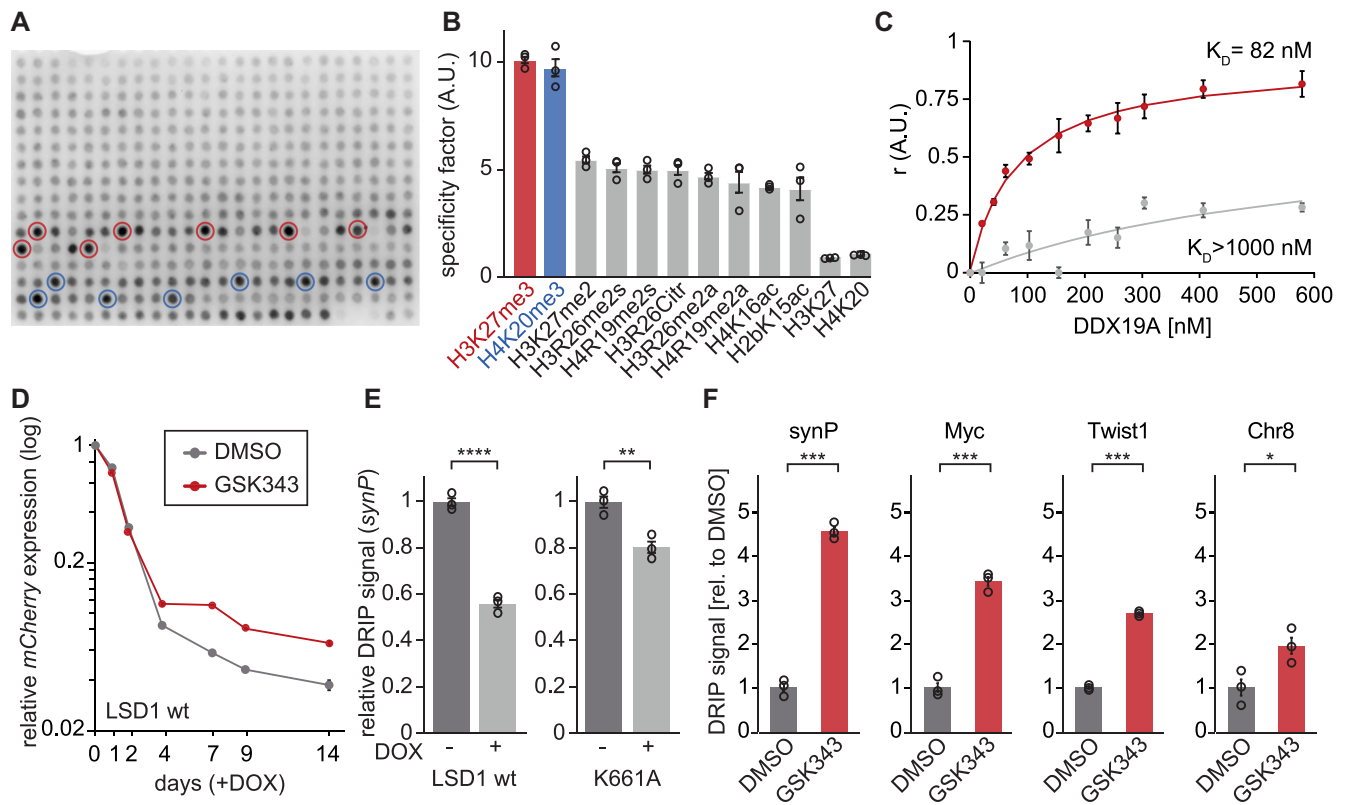
Figure 6. H3K27me3 provides a binding site for DDX19A and regulates the formation of R-loops. (**A**) Representative MODified™ Histone Peptide Array incubated with 50 nM DDX19A. Peptides featuring H3K27me3 are highlighted in red, peptides featuring H4K20me3 in blue. (**B**) Bar graph comparing signals of the top 10 histone modifications bound by DDX19A to binding of DDX19A to the unmodified peptides. DDX19A binding was quantified on independent MODified™ Histone Peptide Arrays using the Array Analyze software ($n = 3$; mean ± s.e.m.). 'H3K27' and 'H4K20' refers to the specificity factor of the corresponding unmodified peptides. [A.U.] = arbitrary unit. (**C**) Determination of the dissociation constant of DDX19A binding to H3K27me3 (red) or H3K27me0 (gray) by equilibrium peptide binding experiments. Fluorescence anisotropy measurement was performed with a fluorescein (FITC)-labeled peptide incubated with recombinant DDX19A. [A.U.] = arbitrary unit. (**D**) Time course of *mCherry* expression in NIH/3T3 expressing the *synP-mCherry* reporter during recruitment of the rTetR-LSD1 fusion protein via DOX in the presence of 5 μM GSK343 or DMSO. Circles indicate the median *mCherry* expression measured by flow cytometry relative to the initial measurement ($n = 3$; mean ± s.e.m.). (**E**) DRIP-qPCR at the *synP* element after recruitment of LSD1 wt or K661A via DOX addition. Total nucleic acids for the IP were isolated after 14 days of DOX treatment. Bars are relative to -DOX. Circles represent independent replicates ($n = 3$; mean ± s.e.m., **$P \leq 0.01$, ****$P \leq 0.0001$, Student's *t*-test). (**F**) DRIP-qPCR analysis of RNA:DNA hybrid structures at the indicated regions after the treatment of cells with 5 μM GSK343 or DMSO for 3 days. Total nucleic acids were extracted from NIH/3T3 cells expressing the *synP-mCherry* reporter and used as input for IP with the S9.6 antibody. qPCR signals are shown relative to DMSO. Circles represent independent replicates ($n = 3$; mean ± s.e.m.; *$P \leq 0.05$, ***$P \leq 0.001$, n.s. = non-significant, Student's *t*-test).

ticity ([102](#)) and several reports demonstrate that LSD1 and the PRC2 complex are associated with actively transcribed developmental genes ([28](#),[103](#),[104](#)). To explore these functional networks in detail we adopted the concept of the CiA assay ([44](#)) and investigated the function of epigenetic effector proteins in LSD1 mediated gene silencing. By fusing LSD1 to rTetR, we induced spatial proximity of LSD1 with a synthetic promoter element (*synP*) consisting of six *tetO* sites upstream of an *EF1a* promoter through the addition of DOX. The reporter system is modular, highly flexible and allows to study epigenetic effectors in different cell lines of various cell types in a time-resolved manner independent of a specific genomic locus (Figure [1](#) and Supplementary Figure S1). We combined the fluorescent reporter system for LSD1 activity with a comprehensive library of shRNAs targeting chromatin associated proteins and performed a <u>ch</u>romatin <u>e</u>ffector <u>c</u>oregulator <u>s</u>creen (ChECS) to

identify functional, essential and novel coregulators. Compared to screening methods that depend on the temporal and chemical stability of physical complexes of two or more proteins, the ChECS approach enables the identification of coregulatory factors on the basis of a functional connection between the coregulator and the target factor, in this study LSD1. This allows to detect novel dependencies in epigenetic networks (Figure [2](#)) in an unbiased fashion without the need for a pre-existing hypothesis. One of the top hits identified in our screen, the DEAD-box helicase DDX19A, had not been connected to LSD1 biology before. Following up on the identification of DDX19A, we showed that it plays an important role in LSD1 induced silencing and attributed this role to its function in R-loop homeostasis (Figure [3](#) and Supplementary Figure S3). Suppression of *Ddx19a* expression led to a strong accumulation of R-loops in the nuclei of NIH/3T3 cells leading to stabilization of the

fluorescent reporter protein expression after recruitment of LSD1 (Figure 2B, Supplementary Figure S2E and Figure 5A).

Our results expand the knowledge of the highly complex regulatory network surrounding LSD1 (Figure 7A) and identified a novel regulatory circuit embedded in this network, which enables robust downregulation of transcription upon an external stimulus. Our results indicate that LSD1 and PRC2 are prevented from gene silencing through R-loops or the associated nascent RNA, which is in line with recent publications that already described an inhibitory effect on LSD1 and the PRC2 complex through specific RNA structures (38–40). If a cell perceives an external signal, e.g. during development, the expression of developmental genes needs to be completely and robustly repressed in order to prevent uncontrolled proliferation (105,106). During this process R-loops, which are not only a consequence of transcription but also pose a regulatory function (80,107), need to be removed in order to allow robust gene silencing. Removal of R-loops can be triggered by the recruitment of DDX19A, or the activation of the already bound enzyme. Currently it is unclear what guides this initial recruitment, but the fact that DDX19 is a helicase able to bind to RNA (108) suggests that DDX19A can be recruited to sites of active transcription by the nascent RNA. The removal of R-loops by DDX19A activates LSD1 leading to the demethylation of H3K4, presumably in concert with KDM5b, another top hit of our screen. H3K4me2/3 demethylation is a prerequisite for the methylation of H3K27 by PRC2 in our system, which is in line with published literature (59,60). Subsequently, the introduced H3K27me3 signal provides a binding motif for DDX19A, which enhances DDX19A activity by reinforcing the binding to target loci or amplifying the enzymatic activity of DDX19A to effectively remove remaining R-loops (Figure 7B). This activity finally increases the activation of LSD1 and PRC2 establishing a self-enforcing feedback cycle. Therefore, the entire sequence of events is required to induce a strong and stable shut-down of gene expression and in case one of the components is lost or inactive, the silencing stimulus will not lead to a complete reduction of expression of the respective gene (Figure 7). In the ChECS system, reporter gene silencing was triggered by the recruitment of LSD1. During development or differentiation, silencing of endogenous loci could be a result of changes in the catalytic activity of LSD1, PRC2 or it could be triggered by changes in expression of associated genes.

This novel mechanism is supported by the following key observations described previously in detail: (i) The biochemical activity of DDX19A in resolving R-loops was demonstrated *in vitro* and silencing of *Ddx19a* expression was shown to lead to an increase in R-loop occupancy at the reporter gene and endogenous loci. (ii) We observed a correlation of LSD1 occupancy with the distribution of R-loop associated features genome-wide via the analysis of ChIP-seq data in two different cell lines (Figure 4 and Supplementary Figure S4), illustrating that LSD1 and R-loops co-occupy regions in the vicinity of highly expressed genes (Figure 4D and Supplementary Figure S4E) that are associated with essential developmental transcription programs (Supplementary Figure S4F). Furthermore, a direct phys-
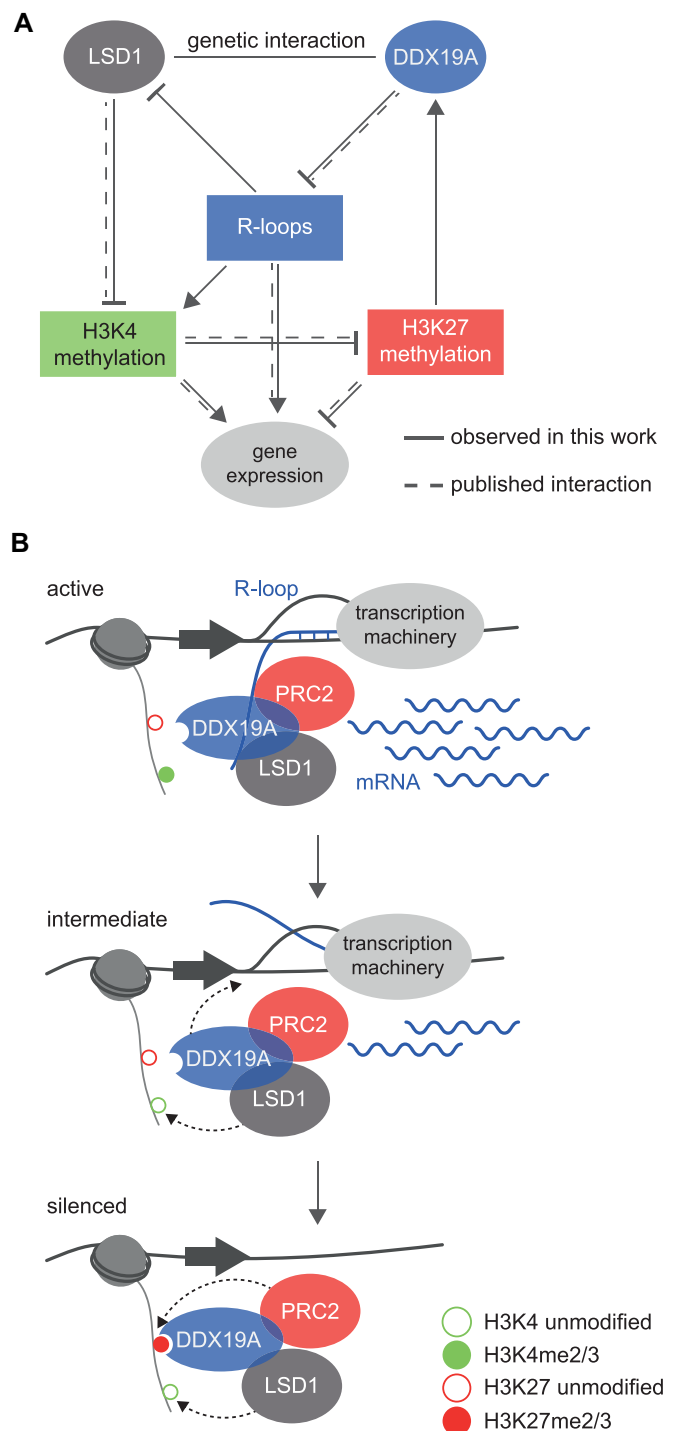
Figure 7. Expression of actively transcribed genes is regulated by a network of histone modifications and R-loops. (**A**) Regulatory network depicting the known and novel connections that interact to modulate gene expression. (**B**) Model of the regulatory cascade for transcriptional repression depending on the removal of transcription-associated R-loops downstream of LSD1 activity. LSD1 and PRC2 are localized at highly transcribed genes. Active transcription promotes the local presence of H3K4me2/3 and the formation of R-loops. These R-loops are balanced by specific helicases (e.g. DDX19A). Upon an external repressive stimulus LSD1 activity is increased leading to a reduction in H3K4 methylation, which enables the PRC2 complex to methylate H3K27. H3K27me3 serves as a binding motif for DDX19A, which then efficiently removes local R-loops and leading to robust silencing of transcription.

ical interaction of LSD1 and DDX19A with R-loops was confirmed by co-IP using a RNA:DNA hybrid-specific antibody (Figure 4E). (iii) R-loop enrichment by *Ddx19a* suppression affected the robust downregulation of transcription only if LSD1 was catalytically active suggesting that R-loops inhibit LSD1 enzymatic activity (Figure 5A, C). Suppression of *Ddx19a* expression did not only lead to an increase in R-loops but also to an impaired demethylation of H3K4me2 and reduced increase in H3K27me3. (iv) Reduction of LSD1 activity by recruitment of an inactive mutant or application of a small molecule inhibitor, not only led to reduced decline of H3K4me2, but also reduced introduction of H3K27me3 and increased R-loop occupancy. The observation that H3K27 methylation by PRC2 depends on H3K4me2 demethylation is in line with published literature (59,60). (v) Application of a PRC2 inhibitor not only prevented introduction of H3K27me3, but also led to an increase in R-loops and it impedes reporter gene silencing. (vi) H3K9me3 deposition and H3K27 deacetylation following LSD1 recruitment were not affected by suppression of *Ddx19a* suggesting that known complex partners of LSD1 like G9a and HDACs were still recruited (Figure 5B). (vii) We observed distinct binding of DDX19A to H3K27me3 and H4K20me3 histone peptides (Figure 6A–C and Supplementary Figure S6A, B) providing biochemical evidence for the interaction of DDX19A with H3K27me3.

Taken together, we describe a novel multiplexed approach to identify functional coregulators of chromatin effectors. ChECS enabled us to identify a so far unknown transcriptional regulatory cascade controlling developmental gene expression programs, which could contribute to oncogenesis in case of deregulation. Our study illustrates an interconnection of LSD1 and R-loop homeostasis, which provides novel insights into the biological functions of LSD1. Furthermore, we show for the first time that a DEAD-box helicase (DDX19A) contains a bona fide binding motif for selected histone modifications associated with gene repression and binds H3K27me3 with considerably high affinity (109). To date, DEAD-box helicases were not reported to interact with modified histone tails and do not harbour any of the so far known histone reading domains, a finding that might extend to additional RNA:DNA helicases, disclosing the potential for additional unknown regulatory pathways.

DATA AVAILABILITY

LSD1 ChIP-seq data (GSE158441) is available via Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/).

The following ChIP-seq tracks were obtained from published data sets in K562 cells: H3K27ac (Encode sample ENCFF384ZZM), H3K27me3 (Encode sample ENCFF936BVT), H3K9me3 (Encode sample ENCFF700FQH), H3K36me3 (Encode sample ENCFF223BKS), H3K4me1 (Encode sample ENCFF463AQS), H3K4me2 (Encode sample ENCFF778DNU), LSD1 (GEO sample GSM831002), R-ChIP (GEO sample GSM2551007/8), DRIP-seq (GEO sample GSM1720619), G4-ChIP-seq (GEO sample GSM2876090/1). K562 RNA-seq (GEO sample GSM1557077). The following ChIP-seq tracks were obtained from published data sets in NIH/3T3 cells:

H3K9ac (GEO sample GSM1246687), H3K27me3 (GEO sample GSM1246690), H3K36me3 (GEO sample GSM1246692), H3K9me2 (GEO sample GSM1246688), H4ac (GEO sample GSM1418787), H3K4me3 (GEO sample GSM1246686), DRIPc-seq (GEO sample GSM2104456), DRIP-seq (GEO sample GSM1720621). NIH/3T3 RNA-seq (Encode sample ENCFF001QSC).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Maiques-Diaz,A. and Somervaille,T.C. (2016) LSD1: biologic roles and therapeutic targeting. *Epigenomics*, **8**, 1103–1116.
2. Adamo,A., Sesé,B., Boue,S., Castaño,J., Paramonov,I., Barrero,M.J. and Belmonte,J.C.I. (2011) LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nat. Cell Biol.*, **13**, 652–659.
3. Foster,C.T., Dovey,O.M., Lezina,L., Luo,J.L., Gant,T.W., Barlev,N., Bradley,A. and Cowley,S.M. (2010) Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol. Cell. Biol.*, **30**, 4851–4863.
4. Takeuchi,M., Fuse,Y., Watanabe,M., Andrea,C.S., Takeuchi,M., Nakajima,H., Ohashi,K., Kaneko,H., Kobayashi-Osaki,M., Yamamoto,M. *et al.* (2015) LSD1/KDM1A promotes

hematopoietic commitment of hemangioblasts through downregulation of Etv2. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 13922–13927.

5. Thambyrajah,R., Mazan,M., Patel,R., Moignard,V., Stefanska,M., Marinopoulou,E., Li,Y., Lancrin,C., Clapes,T., Möröÿ,T. *et al.* (2016) GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nat. Cell Biol.*, **18**, 21–32.
6. Kerenyi,M.A., Shao,Z., Hsu,Y.J., Guo,G., Luc,S., O'Brien,K., Fujiwara,Y., Peng,C., Nguyen,M. and Orkin,S.H. (2013) Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. *Elife*, **2**, e00633.
7. Mosammaparast,N., Kim,H., Laurent,B., Zhao,Y., Lim,H.J., Majid,M.C., Dango,S., Luo,Y., Hempel,K., Sowa,M.E. *et al.* (2013) The histone demethylase LSD1/KDM1A promotes the DNA damage response. *J. Cell Biol.*, **203**, 457–470.
8. Hayami,S., Kelly,J.D., Cho,H.S., Yoshimatsu,M., Unoki,M., Tsunoda,T., Field,H.I., Neal,D.E., Yamaue,H., Ponder,B.A.J. *et al.* (2011) Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int. J. Cancer*, **128**, 574–586.
9. Amente,S., Lania,L. and Majello,B. (2013) The histone LSD1 demethylase in stemness and cancer transcription programs. *Biochim. Biophys. Acta - Gene Regul. Mech.*, **1829**, 981–986.
10. Magerl,C., Ellinger,J., Braunschweig,T., Kremmer,E., Koch,L.K., Höller,T., Büttner,R., Lüscher,B. and Gütgemann,I. (2010) H3K4 dimethylation in hepatocellular carcinoma is rare compared with other hepatobiliary and gastrointestinal carcinomas and correlates with expression of the methylase Ash2 and the demethylase LSD1. *Hum. Pathol.*, **41**, 181–189.
11. Lim,S., Janzer,A., Becker,A., Zimmer,A., Schüle,R., Buettner,R. and Kirfel,J. (2010) Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. *Carcinogenesis*, **31**, 512–520.
12. Abdel-Hafiz,H. (2017) Epigenetic mechanisms of tamoxifen resistance in luminal breast cancer. *Diseases*, **5**, 16.
13. Lv,T., Yuan,D., Miao,X., Lv,Y., Zhan,P., Shen,X. and Song,Y. (2012) Over-expression of LSD1 promotes proliferation, migration and invasion in non-small cell lung cancer. *PLoS One*, **7**, e35065.
14. Kauffman,E.C., Robinson,B.D., Downes,M.J., Powell,L.G., Lee,M.M., Scherr,D.S., Gudas,L.J. and Mongan,N.P. (2011) Role of androgen receptor and associated lysine-demethylase coregulators, LSD1 and JMJD2A, in localized and advanced human bladder cancer. *Mol. Carcinog.*, **50**, 931–944.
15. Serce,N., Gnatzy,A., Steiner,S., Lorenzen,H., Kirfel,J. and Buettner,R. (2012) Elevated expression of LSD1 (Lysine-specific demethylase 1) during tumour progression from pre-invasive to invasive ductal carcinoma of the breast. *BMC Clin. Pathol.*, **12**, 13.
16. Boulding,T., McCuaig,R.D., Tan,A., Hardy,K., Wu,F., Dunn,J., Kalimutho,M., Sutton,C.R., Forwood,J.K., Bert,A.G. *et al.* (2018) LSD1 activation promotes inducible EMT programs and modulates the tumour microenvironment in breast cancer. *Sci. Rep.*, **8**, 73.
17. Carnesecchi,J., Forcet,C., Zhang,L., Tribollet,V., Barenton,B., Boudra,R., Cerutti,C., Billas,I.M.L., Sérandour,A.A., Carroll,J.S. *et al.* (2017) ERRα induces H3K9 demethylation by LSD1 to promote cell invasion. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 3909–3914.
18. Ferrari-Amorotti,G., Chiodoni,C., Shen,F., Cattelani,S., Soliera,A.R., Manzotti,G., Grisendi,G., Dominici,M., Rivasi,F., Colombo,M.P. *et al.* (2014) Suppression of invasion and metastasis of triple-negative breast cancer lines by pharmacological or genetic inhibition of slug activity. *Neoplasia*, **16**, 1047–1058.
19. Pollock,J.A., Larrea,M.D., Jasper,J.S., McDonnell,D.P. and McCafferty,D.G. (2012) Lysine-specific histone demethylase 1 inhibitors control breast cancer proliferation in ERα-dependent and -independent manners. *ACS Chem. Biol.*, **3**, 1691–1707.
20. Rivero,S., Ceballos-Chávez,M., Bhattacharya,S.S. and Reyes,J.C. (2015) HMG20A is required for SNAI1-mediated epithelial to mesenchymal transition. *Oncogene*, **34**, 5264–5276.
21. Harris,W.J., Huang,X., Lynch,J.T., Spencer,G.J., Hitchin,J.R., Li,Y., Ciceri,F., Blaser,J.G., Greystoke,B.F., Jordan,A.M. *et al.* (2012) The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell*, **21**, 473–487.
22. Maiques-Diaz,A., Spencer,G.J., Lynch,J.T., Ciceri,F., Williams,E.L., Amaral,F.M.R., Wiseman,D.H., Harris,W.J., Li,Y., Sahoo,S. *et al.* (2018) Enhancer activation by pharmacologic displacement of

LSD1 from GFI1 induces differentiation in acute myeloid leukemia. *Cell Rep.*, **22**, 3641–3659.
23. Garcia-Bassets,I., Kwon,Y.S., Telese,F., Prefontaine,G.G., Hutt,K.R., Cheng,C.S., Ju,B.G., Ohgi,K.A., Wang,J., Escoubet-Lozach,L. *et al.* (2007) Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell*, **128**, 505–518.
24. Shi,Y., Lan,F., Matson,C., Mulligan,P., Whetstine,J.R., Cole,P.A., Casero,R.A. and Shi,Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, **119**, 941–953.
25. Lynch,J.T., Spencer,G.J., Harris,W.J., Maiques-Díaz,A., Ciceri,F., Huang,X. and Somervaille,T.C.P. (2014) Pharmacological inhibitors of LSD1 promote differentiation of myeloid leukemia cells through a mechanism independent of histone demethylation. *Blood*, **124**, 267.
26. Huang,J., Sengupta,R., Espejo,A.B., Lee,M.G., Dorsey,J.A., Richter,M., Opravil,S., Shiekhattar,R., Bedford,M.T., Jenuwein,T. *et al.* (2007) p53 is regulated by the lysine demethylase LSD1. *Nature*, **449**, 105–108.
27. Wang,J., Hevi,S., Kurash,J.K., Lei,H., Gay,F., Bajko,J., Su,H., Sun,W., Chang,H., Xu,G. *et al.* (2009) The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat. Genet.*, **41**, 125–129.
28. Whyte,W.A., Bilodeau,S., Orlando,D.A., Hoke,H.A., Frampton,G.M., Foster,C.T., Cowley,S.M. and Young,R.A. (2012) Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature*, **482**, 221–225.
29. Cai,C., He,H.H., Gao,S., Chen,S., Yu,Z., Gao,Y., Chen,S., Chen,M.W., Zhang,J., Ahmed,M. *et al.* (2014) Lysine-Specific demethylase 1 has dual functions as a major regulator of androgen receptor transcriptional activity. *Cell Rep.*, **9**, 1618–1628.
30. Cai,C., He,H.H., Chen,S., Coleman,I., Wang,H., Fang,Z., Chen,S., Nelson,P.S., Liu,X.S., Brown,M. *et al.* (2011) Androgen receptor gene expression in prostate cancer is directly suppressed by the androgen receptor through recruitment of lysine-specific demethylase 1. *Cancer Cell*, **20**, 457–471.
31. Metzger,E., Wissmann,M., Yin,N., Müller,J.M., Schneider,R., Peters,A.H.F.M., Günther,T., Buettner,R. and Schüle,R. (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature*, **437**, 25–28.
32. Bennesch,M.A., Segala,G., Wider,D. and Picard,D. (2016) LSD1 engages a corepressor complex for the activation of the estrogen receptor α by estrogen and cAMP. *Nucleic. Acids. Res.*, **44**, 8655–8670.
33. Laurent,B., Ruitu,L., Murn,J., Hempel,K., Ferrao,R., Xiang,Y., Liu,S., Garcia,B.A., Wu,H., Wu,F. *et al.* (2015) A Specific LSD1/KDM1A Isoform Regulates Neuronal Differentiation through H3K9 Demethylation. *Mol. Cell*, **57**, 957–970.
34. Toffolo,E., Rusconi,F., Paganini,L., Tortorici,M., Pilotto,S., Heise,C., Verpelli,C., Tedeschi,G., Maffioli,E., Sala,C. *et al.* (2014) Phosphorylation of neuronal Lysine-Specific Demethylase 1LSD1/KDM1A impairs transcriptional repression by regulating interaction with CoREST and histone deacetylases HDAC1/2. *J. Neurochem.*, **128**, 603–616.
35. Zibetti,C., Adamo,A., Binda,C., Forneris,F., Toffolo,E., Verpelli,C., Ginelli,E., Mattevi,A., Sala,C. and Battaglioli,E. (2010) Alternative splicing of the histone demethylase LSD1/KDM1 contributes to the modulation of neurite morphogenesis in the mammalian nervous system. *J. Neurosci.*, **30**, 2521–2532.
36. Kozub,M.M., Carr,R.M., Lomberk,G.L. and Fernandez-Zapico,M.E. (2017) LSD1, a double-edged sword, confers dynamic chromatin regulation but commonly promotes aberrant cell growth. *F1000Research*, **6**, 2016.
37. Shi,Y.J., Matson,C., Lan,F., Iwase,S., Baba,T. and Shi,Y. (2005) Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell*, **19**, 857–864.
38. Hirschi,A., Martin,W.J., Luka,Z., Loukachevitch,L. V. and Reiter,N.J. (2016) G-quadruplex RNA binding and recognition by the lysine-specific histone demethylase-1 enzyme. *RNA*, **22**, 1250–1260.
39. Zhang,Q., McKenzie,N.J., Warneford-Thomson,R., Gail,E.H., Flanigan,S.F., Owen,B.M., Lauman,R., Levina,V., Garcia,B.A., Schittenhelm,R.B. *et al.* (2019) RNA exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2. *Nat. Struct. Mol. Biol.*, **26**, 237–247.

40. Beltran,M., Tavares,M., Justin,N., Khandelwal,G., Ambrose,J., Foster,B.M., Worlock,K.B., Tvardovskiy,A., Kunzelmann,S., Herrero,J. et al. (2019) G-tract RNA removes Polycomb repressive complex 2 from genes. *Nat. Struct. Mol. Biol.*, **26**, 899–909.
41. Perillo,B., Tramontano,A., Pezone,A. and Migliaccio,A. (2020) LSD1: more than demethylation of histone lysine residues. *Exp. Mol. Med.*, **52**, 1936–1947.
42. Pandey,M.R. and Wang,E.S. (2019) What potential is there for LSD1 inhibitors to reach approval for AML? *Expert Opin. Emerg. Drugs*, **24**, 205–212.
43. Stanton,B.Z., Chory,E.J. and Crabtree,G.R. (2018) Chemically induced proximity in biology and medicine. *Science*, **359**, eaao5902.
44. Hathaway,N.A., Bell,O., Hodges,C., Miller,E.L., Neel,D.S. and Crabtree,G.R. (2012) Dynamics and memory of heterochromatin in living cells. *Cell*, **149**, 1447–1460.
45. Moussa,H.F., Bsteh,D., Yelagandula,R., Pribitzer,C., Stecher,K., Bartalska,K., Michetti,L., Wang,J., Zepeda-Martinez,J.A., Elling,U. et al. (2019) Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing. *Nat. Commun.*, **10**, 1931.
46. Headley,K.M., Kedziora,K.M., Alejo,A., Lai,E.Z.X., Purvis,J.E. and Hathaway,N.A. (2019) Chemical screen for epigenetic barriers to single allele activation of Oct4. *Stem Cell Res.*, **38**, 101470.
47. Stanton,B.Z., Hodges,C., Calarco,J.P., Braun,S.M.G., Ku,W.L., Kadoch,C., Zhao,K. and Crabtree,G.R. (2017) Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat. Genet.*, **49**, 282–288.
48. Kadoch,C., Williams,R.T., Calarco,J.P., Miller,E.L., Weber,C.M., Braun,S.M.G., Pulice,J.L., Chory,E.J. and Crabtree,G.R. (2017) Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat. Genet.*, **49**, 213–222.
49. Fellmann,C., Hoffmann,T., Sridhar,V., Hopfgartner,B., Muhar,M., Roth,M., Lai,D.Y., Barbosa,I.A.M., Kwon,J.S., Guan,Y. et al. (2013) An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.*, **5**, 1704–1713.
50. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
51. Ramírez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic. Acids. Res.*, **44**, W160–W165.
52. Welch,R.P., Lee,C., Imbriano,P.M., Patil,S., Weymouth,T.E., Smith,R.A., Scott,L.J. and Sartor,M.A. (2014) ChIP-enrich: gene set enrichment testing for ChIP-seq data. *Nucleic. Acids. Res.*, **42**, e105.
53. Halász,L., Karányi,Z., Boros-Oláh,B., Kuik-Rózsa,T., Sipos,É., Nagy,É., Mosolygó-L,Á., Mázló,A., Rajnavölgyi,É., Halmos,G. et al. (2017) RNA-DNA hybrid (R-loop) immunoprecipitation mapping: an analytical workflow to evaluate inherent biases. *Genome Res.*, **27**, 1063–1073.
54. Hodroj,D., Recolin,B., Serhal,K., Martinez,S., Tsanov,N., Abou Merhi,R. and Maiorano,D. (2017) An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. *EMBO J.*, **36**, 1182–1198.
55. Kudithipudi,S., Lungu,C., Rathert,P., Happel,N. and Jeltsch,A. (2014) Substrate specificity analysis and novel substrates of the protein lysine methyltransferase NSD1. *Chem. Biol.*, **21**, 226–237.
56. Lamprecht,M.R., Sabatini,D.M. and Carpenter,A.E. (2007) CellProfiler: Free, versatile software for automated biological image analysis. *BioTechniques*, **42**, 71–75.
57. Cao,K., Collings,C.K., Morgan,M.A., Marshall,S.A., Rendleman,E.J., Ozark,P.A., Smith,E.R. and Shilatifard,A. (2018) An Mll4/COMPASS-Lsd1 epigenetic axis governs enhancer function and pluripotency transition in embryonic stem cells. *Sci. Adv.*, **4**, eaap8747.
58. Lee,M.G., Wynder,C., Cooch,N. and Shiekhattar,R. (2005) An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature*, **437**, 432–435.
59. Schmitges,F.W., Prusty,A.B., Faty,M., Stützer,A., Lingaraju,G.M., Aiwazian,J., Sack,R., Hess,D., Li,L., Zhou,S. et al. (2011) Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell*, **42**, 330–341.
60. Voigt,P., LeRoy,G., Drury,W.J., Zee,B.M., Son,J., Beck,D.B., Young,N.L., Garcia,B.A. and Reinberg,D. (2012) Asymmetrically modified nucleosomes. *Cell*, **151**, 181–193.
61. Kolla,V., Naraparaju,K., Zhuang,T., Higashi,M., Kolla,S., Blobel,G.A. and Brodeur,G.M. (2015) The tumour suppressor CHD5 forms a NuRD-type chromatin remodelling complex. *Biochem. J.*, **468**, 345–352.
62. Rathert,P., Roth,M., Neumann,T., Muerdter,F., Roe,J.-S., Muhar,M., Deswal,S., Cerny-Reiterer,S., Peter,B., Jude,J. et al. (2015) Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature*, **525**, 543–547.
63. Petell,C.J., Alabdi,L., He,M., San Miguel,P., Rose,R. and Gowher,H. (2016) An epigenetic switch regulates de novo DNA methylation at a subset of pluripotency gene enhancers during embryonic stem cell differentiation. *Nucleic. Acids. Res.*, **44**, 7605–7617.
64. Wang,Y., Zhang,H., Chen,Y., Sun,Y., Yang,F., Yu,W., Liang,J., Sun,L., Yang,X., Shi,L. et al. (2009) LSD1 is a subunit of the NuRD complex and targets the metastasis Programs in breast cancer. *Cell*, **138**, 660–672.
65. Yang,Y., Huang,W., Qiu,R., Liu,R., Zeng,Y., Gao,J., Zheng,Y., Hou,Y., Wang,S., Yu,W. et al. (2018) LSD1 coordinates with the SIN3A/HDAC complex and maintains sensitivity to chemotherapy in breast cancer. *J. Mol. Cell Biol.*, **10**, 285–301.
66. Shiio,Y., Rose,D.W., Aur,R., Donohoe,S., Aebersold,R. and Eisenman,R.N. (2006) Identification and characterization of SAP25, a novel component of the mSin3 corepressor complex. *Mol. Cell. Biol.*, **26**, 1386–1397.
67. Li,Q., Shi,L., Gui,B., Yu,W., Wang,J., Zhang,D., Han,X., Yao,Z. and Shang,Y. (2011) Binding of the JmjC demethylase JARID1B to LSD1/NuRD suppresses angiogenesis and metastasis in breast cancer cells by repressing chemokine CCL14. *Cancer Res.*, **71**, 6899–6908.
68. Bindea,G., Mlecnik,B., Hackl,H., Charoentong,P., Tosolini,M., Kirilovsky,A., Fridman,W.H., Pagès,F., Trajanoski,Z. and Galon,J. (2009) ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, **25**, 1091–1093.
69. Zolotukhin,A.S., Uranishi,H., Lindtner,S., Bear,J., Pavlakis,G.N. and Felber,B.K. (2009) Nuclear export factor RBM15 facilitates the access of DBP5 to mRNA. *Nucleic. Acids. Res.*, **37**, 7151–7162.
70. Alcázar-Román,A.R., Tran,E.J., Guo,S. and Wente,S.R. (2006) Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nat. Cell Biol.*, **8**, 711–716.
71. Mikhailova,T., Shuvalova,E., Ivanov,A., Susorov,D., Shuvalov,A., Kolosov,P.M. and Alkalaeva,E. (2017) RNA helicase DDX19 stabilizes ribosomal elongation and termination complexes. *Nucleic Acids Res.*, **45**, 1307–1318.
72. Awasthi,S., Chakrapani,B., Mahesh,A., Chavali,P.L., Chavali,S. and Dhayalan,A. (2018) DDX39B promotes translation through regulation of pre-ribosomal RNA levels. *RNA Biol.*, **15**, 1157–1166.
73. Nakata,D., Nakao,S., Nakayama,K., Araki,S., Nakayama,Y., Aparicio,S., Hara,T. and Nakanishi,A. (2017) The RNA helicase DDX39B and its paralog DDX39A regulate androgen receptor splice variant AR-V7 generation. *Biochem. Biophys. Res. Commun.*, **483**, 271–276.
74. Mathew,R., Hartmuth,K., Möhlmann,S., Urlaub,H., Ficner,R. and Lührmann,R. (2008) Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nat. Struct. Mol. Biol.*, **15**, 435–443.
75. Allison,D.F. and Wang,G.G. (2019) R-loops: formation, function, and relevance to cell stress. *Cell Stress*, **3**, 38–46.
76. Chédin,F. (2016) Nascent connections: R-Loops and chromatin patterning. *Trends Genet.*, **32**, 828–838.
77. Crossley,M.P., Bocek,M. and Cimprich,K.A. (2019) R-loops as cellular regulators and genomic threats. *Mol. Cell*, **73**, 398–411.
78. Skourti-Stathaki,K. and Proudfoot,N.J. (2014) A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.*, **28**, 1384–1396.
79. Hartono,S.R., Korf,I.F. and Chédin,F. (2015) GC skew is a conserved property of unmethylated CpG island promoters across vertebrates. *Nucleic Acids Res.*, **43**, 9729–9741.

80. Niehrs,C. and Luke,B. (2020) Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol.*, **21**, 167–178.

81. Ginno,P.A., Lott,P.L., Christensen,H.C., Korf,I. and Chédin,F. (2012) R-Loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell*, **45**, 814–825.

82. Sanz,L.A., Hartono,S.R., Lim,Y.W., Steyaert,S., Rajpurkar,A., Ginno,P.A., Xu,X. and Chédin,F. (2016) Prevalent, dynamic, and conserved R-Loop structures associate with specific epigenomic signatures in mammals. *Mol. Cell*, **63**, 167–178.

83. Sehrawat,A., Gao,L., Wang,Y., Bankhead,A., McWeeney,S.K., King,C.J., Schwartzman,J., Urrutia,J., Bisson,W.H., Coleman,D.J. *et al.* (2018) LSD1 activates a lethal prostate cancer gene network independently of its demethylase function. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E4179–E4188.

84. Malovannaya,A., Lanz,R.B., Jung,S.Y., Bulynko,Y., Le,N.T., Chan,D.W., Ding,C., Shi,Y., Yucer,N., Krenciute,G. *et al.* (2011) Analysis of the human endogenous coregulator complexome. *Cell*, **145**, 787–799.

85. Boguslawski,S.J., Smith,D.E., Michalak,M.A., Mickelson,K.E., Yehle,C.O., Patterson,W.L. and Carrico,R.J. (1986) Characterization of monoclonal antibody to DNA•RNA and its application to immunodetection of hybrids. *J. Immunol. Methods*, **89**, 123–130.

86. Poole,C.J. and van Riggelen,J. (2017) MYC—master regulator of the cancer epigenome and transcriptome. *Genes (Basel)*, **8**, 142.

87. Hurlin,P.J. (2013) Control of vertebrate development by MYC. *Cold Spring Harb. Perspect. Med.*, **3**, a014332.

88. Yang,J., Mani,S.A., Donaher,J.L., Ramaswamy,S., Itzykson,R.A., Come,C., Savagner,P., Gitelman,I., Richardson,A. and Weinberg,R.A. (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*, **117**, 927–939.

89. Qin,Q., Xu,Y., He,T., Qin,C. and Xu,J. (2012) Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. *Cell Res.*, **22**, 90–106.

90. Grunseich,C., Wang,I.X., Watts,J.A., Burdick,J.T., Guber,R.D., Zhu,Z., Bruzel,A., Lanman,T., Chen,K., Schindler,A.B. *et al.* (2018) Senataxin mutation reveals how R-Loops promote transcription by blocking DNA methylation at gene promoters. *Mol. Cell*, **69**, 426–437.

91. Chen,P.B., Chen,H. V., Acharya,D., Rando,O.J. and Fazzio,T.G. (2015) R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.*, **22**, 999–1007.

92. Boque-Sastre,R., Soler,M., Oliveira-Mateos,C., Portela,A., Moutinho,C., Sayols,S., Villanueva,A., Esteller,M. and Guil,S. (2015) Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 5785–5790.

93. Skourti-Stathaki,K., Torlai Triglia,E., Warburton,M., Voigt,P., Bird,A. and Pombo,A. (2019) R-Loops enhance polycomb repression at a subset of developmental regulator genes. *Mol. Cell*, **73**, 930–945.

94. Ram,O., Goren,A., Amit,I., Shoresh,N., Yosef,N., Ernst,J., Kellis,M., Gymrek,M., Issner,R., Coyne,M. *et al.* (2011) Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell*, **147**, 1628–1639.

95. Chen,J.Y., Zhang,X., Fu,X.D. and Chen,L. (2019) R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1. *Nat. Protoc.*, **14**, 1661–1685.

96. Mao,S.Q., Ghanbarian,A.T., Spiegel,J., Martínez Cuesta,S., Beraldi,D., Di Antonio,M., Marsico,G., Hänsel-Hertsch,R., Tannahill,D. and Balasubramanian,S. (2018) DNA G-quadruplex structures mold the DNA methylome. *Nat. Struct. Mol. Biol.*, **25**, 951–957.

97. Clough,E. and Barrett,T. (2016) The Gene Expression Omnibus database. In: Mathé,E. and Davis,S. (eds). *Statistical Genomics: Methods in Molecular Biology*. Vol. **1418**, Humana Press, NY, pp. 1064–3745.

98. Sloan,C.A., Chan,E.T., Davidson,J.M., Malladi,V.S., Strattan,J.S., Hitz,B.C., Gabdank,I., Narayanan,A.K., Ho,M., Lee,B.T. *et al.* (2016) ENCODE data at the ENCODE portal. *Nucleic Acids Res.*, **44**, D726–D732.

99. Andersson,L.C., Nilsson,K. and Gahmberg,C.G. (1979) K562—a human erythroleukemic cell line. *Int. J. Cancer*, **23**, 143–147.

100. Wang,Z., Zang,C., Cui,K., Schones,D.E., Barski,A., Peng,W. and Zhao,K. (2009) Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell*, **138**, 1019–1031.

101. Clayton,A.L., Hazzalin,C.A. and Mahadevan,L.C. (2006) Enhanced histone acetylation and transcription: a dynamic perspective. *Mol. Cell*, **23**, 289–296.

102. Hu,G. and Wade,P.A. (2012) NuRD and pluripotency: a complex balancing act. *Cell Stem Cell*, **10**, 497–503.

103. Beltran,M., Yates,C.M., Skalska,L., Dawson,M., Reis,F.P., Viiri,K., Fisher,C.L., Sibley,C.R., Foster,B.M., Bartke,T. *et al.* (2016) The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res.*, **26**, 896–907.

104. Kaneko,S., Son,J., Shen,S.S., Reinberg,D. and Bonasio,R. (2013) PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat. Struct. Mol. Biol.*, **20**, 1258–1264.

105. Dambacher,S., Hahn,M. and Schotta,G. (2010) Epigenetic regulation of development by histone lysine methylation. *Heredity (Edinb)*, **105**, 24–37.

106. Flavahan,W.A., Gaskell,E. and Bernstein,B.E. (2017) Epigenetic plasticity and the hallmarks of cancer. *Science (80-.)*, **357**, eaal2380.

107. Chakraborty,P. (2020) New insight into the biology of R-loops. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.*, **821**, 111711.

108. Hodroj,D., Serhal,K. and Maiorano,d, D. (2017) Ddx19 links mRNA nuclear export with progression of transcription and replication and suppresses genomic instability upon DNA damage in proliferating cells. *Nucleus*, **8**, 489–495.

109. Justin,N., Zhang,Y., Tarricone,C., Martin,S.R., Chen,S., Underwood,E., De Marco,V., Haire,L.F., Walker,P.A., Reinberg,D. *et al.* (2016) Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nat. Commun.*, **7**, 11316.

9.2. Appendix 2

A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation.

Sabine Pinter[1], Franziska Knodel[1], Michel Choudalakis[1], Philipp Schnee[1], Carolin Kroll[1], Marina Fuchs[1], Alexander Broehm[1], Sara Weirich[1], Mareike Roth[3], Stephan A. Eisler[2], Johannes Zuber[3,4], Albert Jeltsch[1], Philipp Rathert[1*]

Supplementary Material

Supplementary Figure S1: The synP-mCherry reporter system is stably expressed in different cell lines and can be manipulated by external stimuli.

Supplementary Figure S2: Suppression of DDX19A expression interferes with LSD1 activity.

Supplementary Figure S3: DDX19A is involved in R-loop homeostasis and Ddx19a suppression leads to global accumulation of R-loops.

Supplementary Figure S4: Regions of LSD1 occupancy correlate with R-loops genome wide in K562 and NIH/3T3 cells.

Supplementary Figure S5: DRIP RNAseH1 controls with and without GSK-LSD1 treatment.

Supplementary Figure S6: DDX19A specifically binds trimethylated H3K27 and H4K20.

Supplementary Figure: Uncropped Western Blots.ai

Supplementary Tables S1-5.xlsx: Supplementary Table 1: ChECS primary data; Supplementary Table 2: ChECS gene scores; Supplementary Table 3: ChECS sequencing primer; Supplementary Table 4: shRNA guide sequences; Supplementary Table 5: qPCR primer sequences
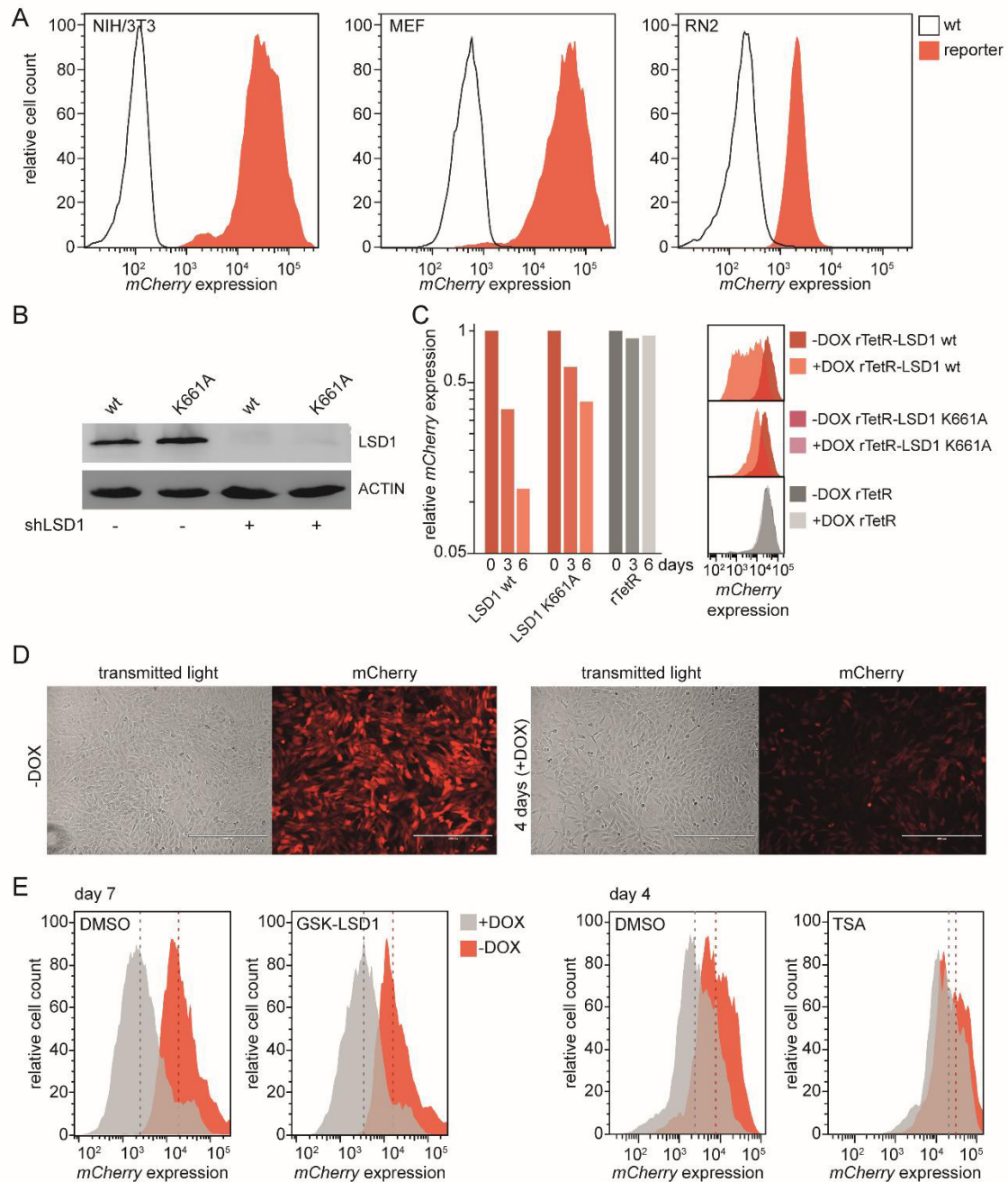
Supplementary Figure S1: The *synP-mCherry* reporter system is stably expressed in different cell lines and can be manipulated by external stimuli. (A) Flow-cytometric analysis of *mCherry* expression after transduction and selection of the *synP-mCherry* reporter cassette in different cell lines (red) compared to untransduced cells (grey). Signals were measured by flow cytometry and are normalized to the highest cell count. MEF=mouse embryonic fibroblasts; RN2=murine MLL–AF9;Nras[G12D] AML cells. **(B)** Immunodetection of the expression levels of the rTetR-LSD1 wt or K661A with and without suppression by *shLSD1*. Whole-cell lysate of NIH/3T3 cells expressing the *synP-mCherry* reporter and the rTetR-LSD1 fusion protein was loaded. ACTIN was used as a loading control. **(C)** Flow cytometry analysis of NIH/3T3 during recruitment of rTetR-LSD1 wt, K661A or rTetR alone to the *synP* promoter. Bar graphs represent the median *mCherry* expression relative to the initial

measurement. The histograms show *mCherry* expression profiles corresponding to the bar graphs at day 0 (-DOX) and day 6 (+DOX) of DOX treatment. One representative experiment is shown. **(D)** Representative fluorescence microscopy images of NIH/3T3 cells stably expressing the *synP-mCherry* reporter cassette together with rTetR-LSD1. Left: untreated cells (-DOX). Right: Cells with 4 days of LSD1 recruitment (+DOX). Scale bars are 400µm. **(E)** Representative *mCherry* expression profiles of NIH/3T3 cells after treatment with 10 µM GSK-LSD1, 400 nM TSA or DMSO and recruitment of rTetR-LSD1 by the addition of DOX for the indicated days. Grey: rTetR-LSD1 recruitment (+DOX), red: no recruitment (-DOX). Dotted lines represent the median *mCherry* signal of the respective population.
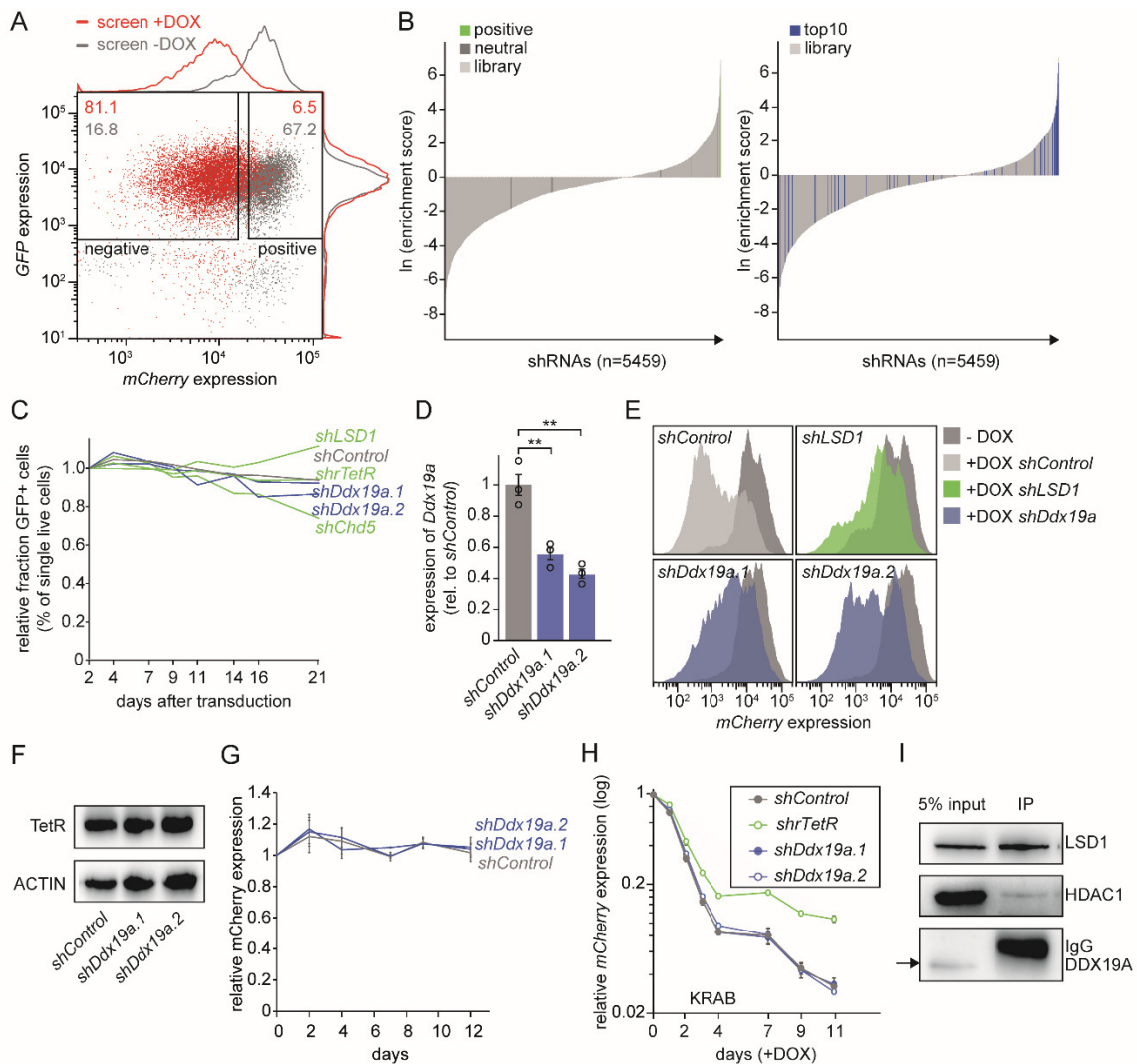
Supplementary Figure S2: Suppression of DDX19A expression interferes with LSD1 activity. (A) Scatter plot illustrating the gating strategy used in FACS. Red: representative cell population transduced with the shRNA library after 14 days +DOX. Grey: control population transduced with the shRNA library without the addition of DOX. Numbers indicate the percentage of cells that were sorted using the respective gate. Histograms of the respective populations are indicated. **(B)** Enrichment scores for all shRNAs tested in the screen. Enrichment scores for each shRNA were calculated by dividing the geometric mean of the read counts of all independent replicates (n=5) in the positive population by the geometric mean of the read counts of the negative population and are plotted as the natural logarithm (ln) in ascending order. Positive and neutral control shRNAs are highlighted in green and grey (left). shRNAs (4-6 per gene) targeting the top10 hits of the screen are highlighted in blue (right). **(C)** Competitive proliferation assays of NIH/3T3 cells expressing the indicated shRNAs. Shown is the fraction of GFP+/shRNA+ cells relative to the initial measurement. **(D)** qPCR analysis of *Ddx19a* mRNA levels in NIH/3T3 cells expressing *shDdx19a* or *shControl* for 10 days. Bars are relative to the *shControl* (n=3, mean±*s.e.m.*, p**≤0.01, *Student´s* t-test). **(E)** Validation of two shRNAs targeting *Ddx19a* (*shDdx19a.1* and *2*). Histograms showing representative *mCherry* expression profiles in rTetR-LSD1 reporter cell populations treated with DOX for 7 days. **(F)** Immunoblotting of NIH/3T3 cells expressing

248

the rTetR-LSD1 fusion protein and the indicated shRNAs. Detection was performed with a TetR-specific antibody. ACTIN was used as a loading control. **(G)** Expression of the *mCherry* reporter in NIH/3T3 cells co-expressing the *synP-mCherry* reporter, the rTetR-LSD1 fusion protein under suppression of *Ddx19a* expression. Plotted is the median *mCherry* signal measured by flow-cytometry in cells expressing the indicated shRNAs without DOX treatment for the indicated number of days relative to the signal at day 0. The mean of individual replicates is plotted (n=3, mean±*s.e.m.*). **(H)** Time course of the *mCherry* signal in NIH/3T3 expressing the *synP-mCherry* reporter, a rTetR-KRAB fusion protein and the indicated shRNAs. Circles indicate the median *mCherry* expression under recruitment of rTetR-KRAB (+DOX) measured by flow cytometry relative to the initial measurement (n=3; mean±*s.e.m.*). **(I)** Immunoblotting of proteins co-precipitated with LSD1 from NIH/3T3 lysate using an antibody against LSD1. The strong band on the lowest blot originates from the antibody used for IP. The DDX19A band is highlighted with an arrow.
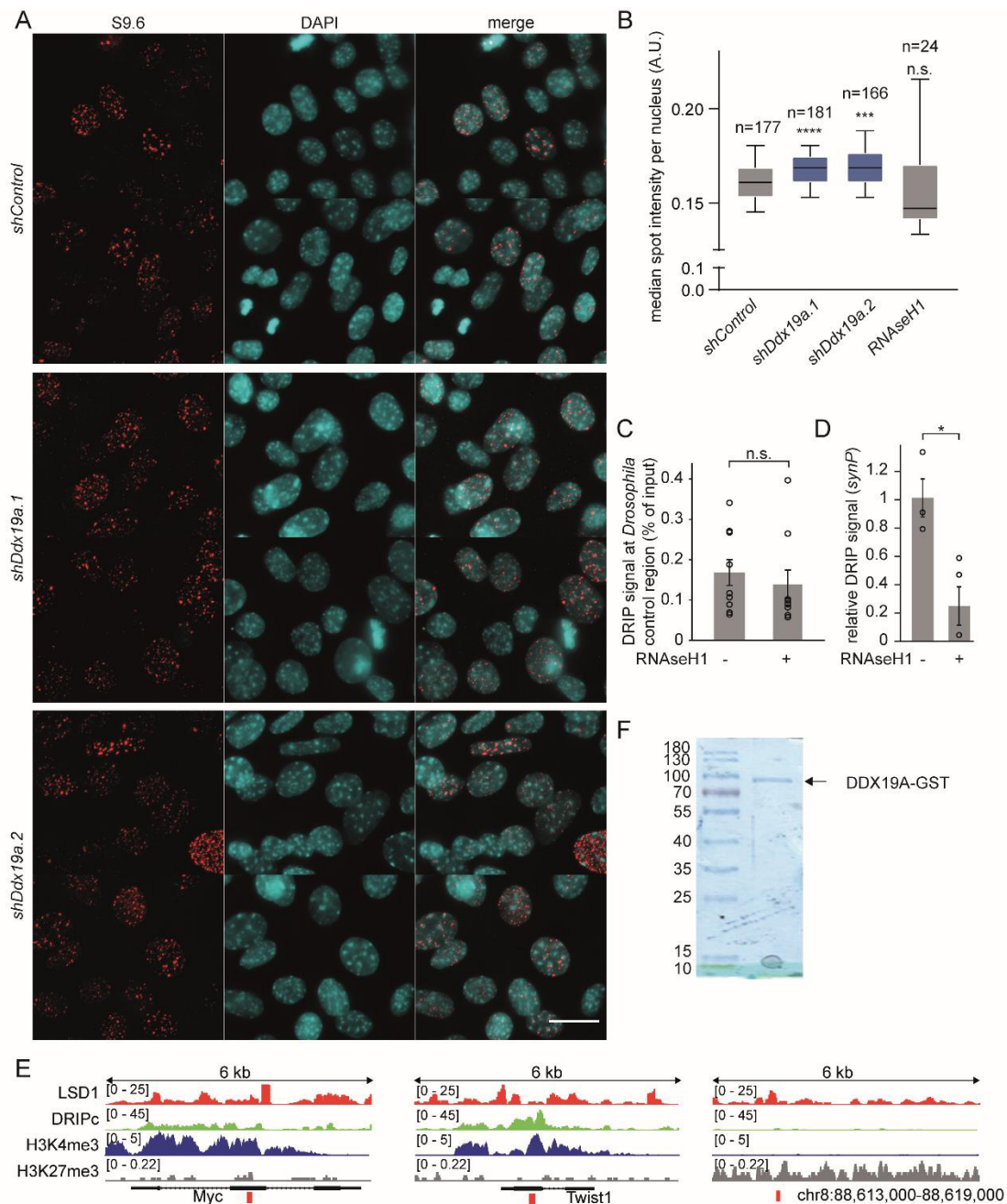
Supplementary Figure S3: DDX19A is involved in R-loop homeostasis and Ddx19a suppression leads to global accumulation of R-loops. (A) Representative immunofluorescence microscopy images of NIH/3T3 cells stably expressing the indicated shRNAs for 10 days and stained with the S9.6 antibody specific for R-loops as shown in Figure 3A. To capture more cells, a lower magnification than in Figure 3A is shown. Images are maximum intensity projections of a Z-stack covering the whole nucleus. Scale bar is 20 μm. **(B)** Box and whisker plot showing the analysis of the R-loop spot intensity per nucleus in NIH/3T3 cells from three independent experiments, stained with the S9.6 antibody as shown in (**A**). Images were analysed using CellProfiler™ software. Box-and-Whisker plots indicate the median and the 10-90 percentile (n=3; mean±*s.e.m.*, p***≤0.001, p****≤0.0001, n.s.=non-significant, *Student´s* t-test). [A.U.]=arbitrary unit. **(C)** Bar graph comparing the enrichment of spiked-in chromatin with the S9.6 antibody from samples that were treated with RNAseH1 before spike-in of *D.melanogaster*

chromatin compared to the enrichment of spike-in *D.melanogaster* chromatin from samples that were not RNAseH1 treated indicating that the IP efficiency is comparable between RNAseH1 treated and untreated samples. (n=9; mean±*s.e.m.*, n.s.=not significant, Student´s *t*-test). Spike-in chromatin was not treated with RNAseH and used for the normalization of pulldown efficiency. **(D)** RNaseH1 control for the DRIP at the synP reporter element shown in Figure 3C. Nucleic acids were treated with RNAseH1 before performing the IP in parallel to the samples without RNAseH1 treatment. qPCR signals are shown relative to -RNAseH1 (n=3, p*<0.05, Student´s t-test). **(E)** Selected ChIP-seq occupancy profiles from NIH/3T3 cells at representative endogenous loci shown in Figure 3D. The regions amplified by qPCR after DRIP are indicated as red bars. **(F)** Coomassie-stained polyacrylamide gel after SDS-PAGE of DDX19A-GST purified with affinity chromatography. Numbers on the left indicate the size of the molecular weight marker in kDa. Arrow indicates the expected size of the full-length recombinant protein.
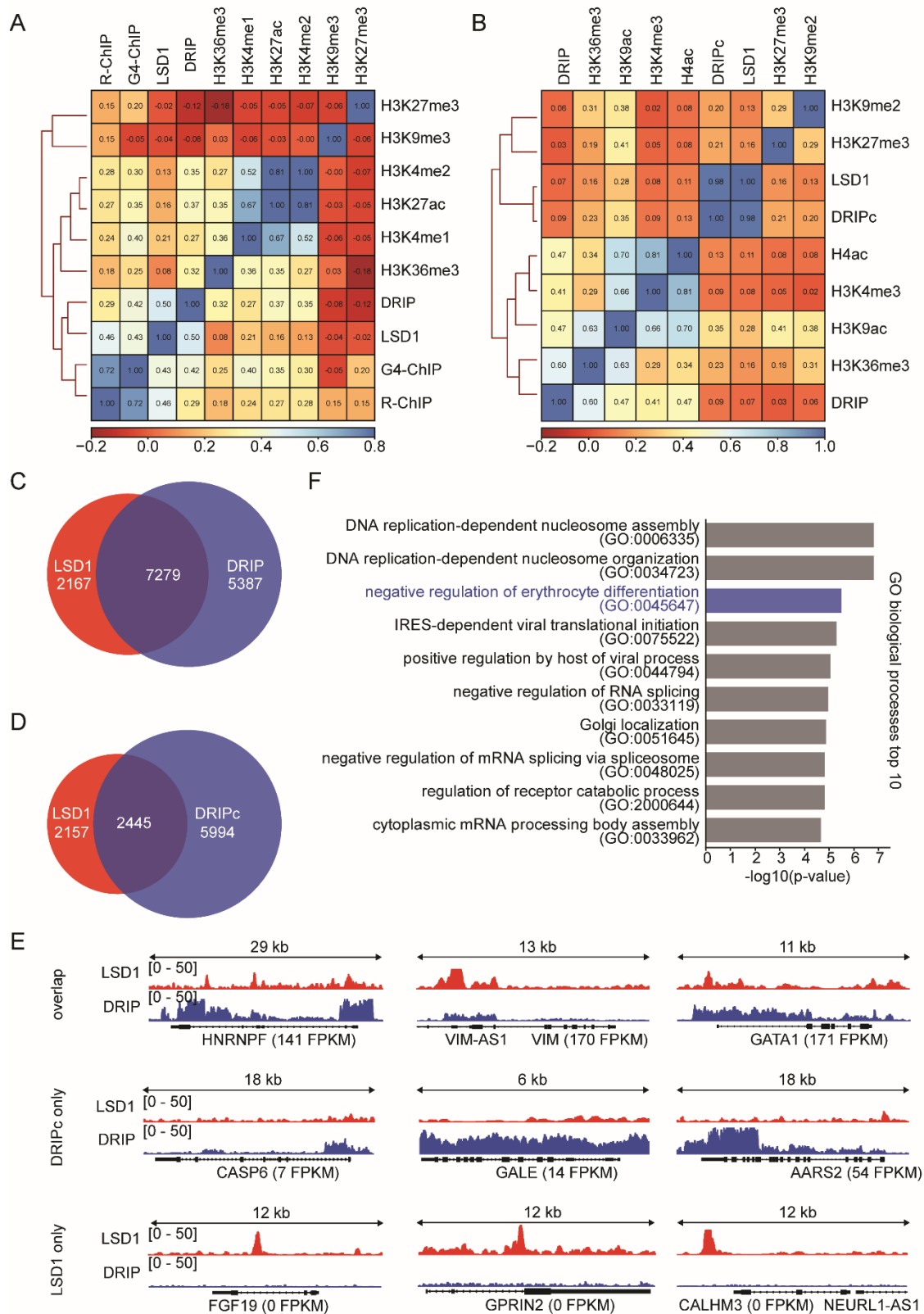
Supplementary Figure S4: Regions of LSD1 occupancy correlate with R-loops genome wide in K562 and NIH/3T3 cells. (A), (B) Correlation matrix of selected histone modifications and R-loop features in K562 (**A**) or NIH/3T3 cells (**B**) underlying the data shown in Figure 5C. The tracks were clustered according to their Pearson correlation coefficients. **(C)** and **(D),** Venn diagrams illustrating

overlap of genes in the vicinity of LSD1 binding and/or R-loops in K562 (**C**) or NIH/3T3 cells (**D**). **(E)** Representative genomic regions in K562 cells showing the occupancy of LSD1 and R-loops at genes associated with the categories shown in Figure 4D. Relative expression is indicated for each gene as FPKM value. **(F)** Gene Ontology analysis of genes bound by LSD1 and R-loops. Genes were identified from LSD1 ChIP-seq and DRIPc-seq data from K562 cells and categorized for their annotated biological processes.
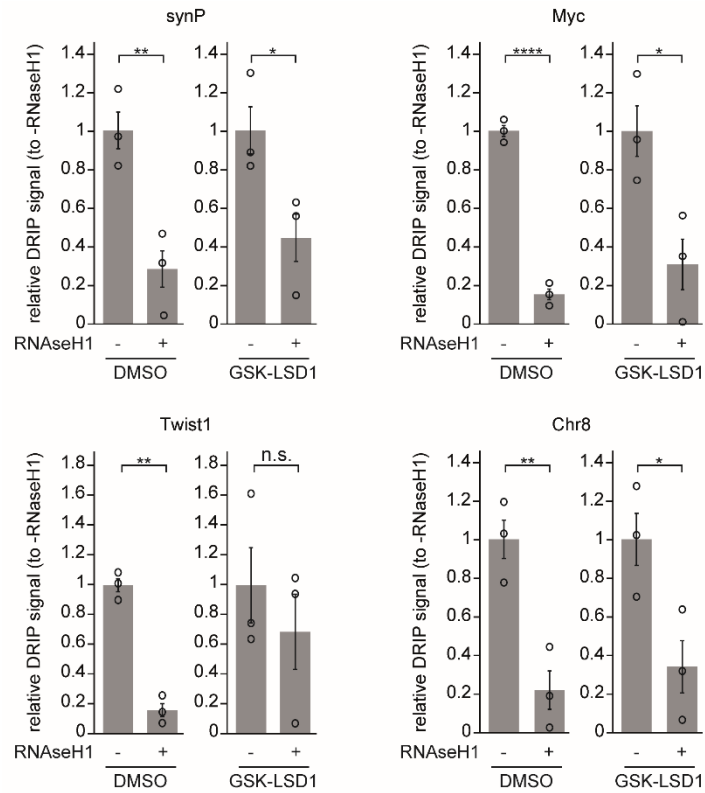
Supplementary Figure S5: DRIP RNAseH1 controls with and without GSK-LSD1 treatment. RNaseH1 controls for the DRIP analysis shown in Figure 5D at the indicated regions. Nucleic acids were treated with RNAseH1 before performing the IP in parallel to the samples without RNAseH1 treatment. qPCR signals are shown relative to -RNAseH1 (n=3, mean±*s.e.m.*, p*<0.05, p**<0.01, p****<0.0001, n.s.=not significant, Student´s t-test).
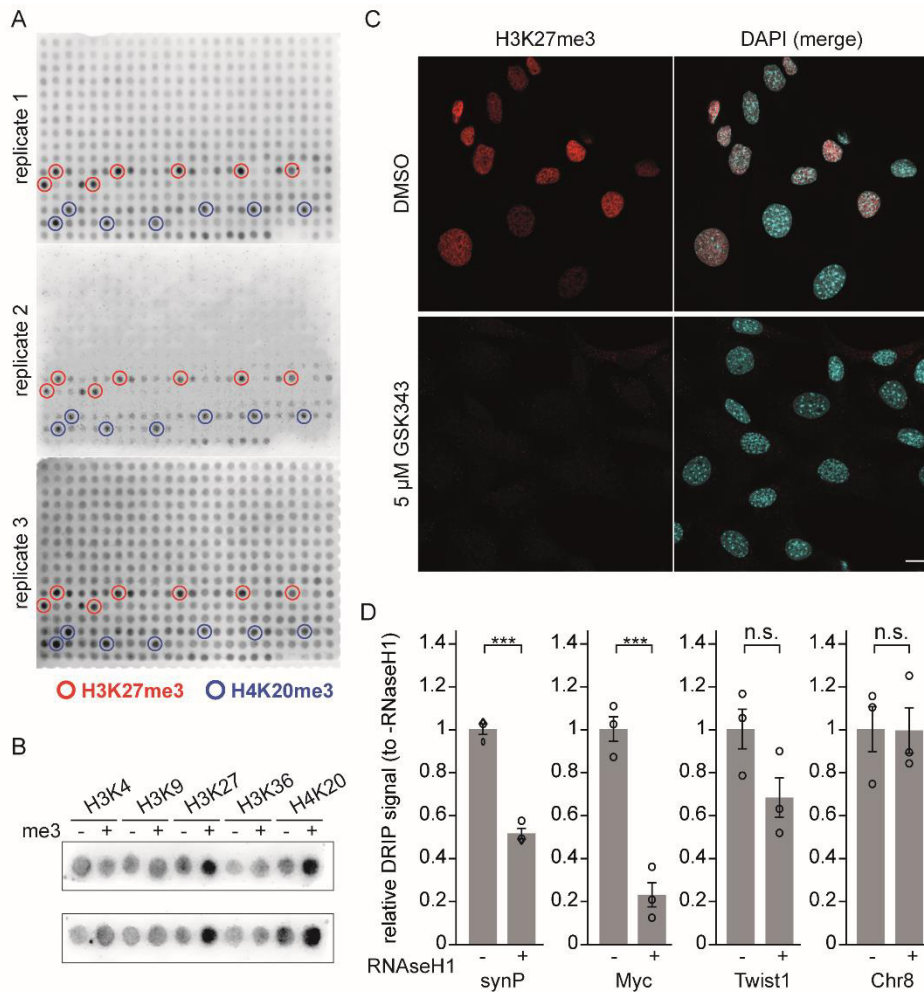
Supplementary Figure S6: DDX19A specifically binds trimethylated H3K27 and H4K20. (A) Independent replicates of the MODified™ Histone Peptide Arrays incubated with 50 nM DDX19A-GST. Binding was detected using an anti-GST antibody and an HRP-coupled secondary antibody. All peptides featuring H3K27me3 are highlighted in red, all peptides featuring H4K20me3 are highlighted in blue. The duplicates on the right side were left unlabelled for better visualization. **(B) Two** representative peptide SPOT arrays featuring the indicated unmodified and trimethylated histone peptides. Arrays were incubated with 50 nM DDX19A-GST. **(C)** Representative immunofluorescence images of NIH/3T3 cells treated with 5 µM GSK-343 or DMSO for 3 days and stained with an H3K27me3 specific antibody. Scale bar is 10 µm. DAPI was used as a nuclear marker. **(D)** RNaseH1 controls for DRIP at the indicated regions shown in Figure 6F. Nucleic acids were treated with RNAseH1 before performing the IP in parallel to the samples without RNAseH1 treatment. The qPCR signals are shown relative to -RNAseH1 (n=3, mean±*s.e.m.*, p***<0.001, n.s.=not significant, Student´s *t*-test).
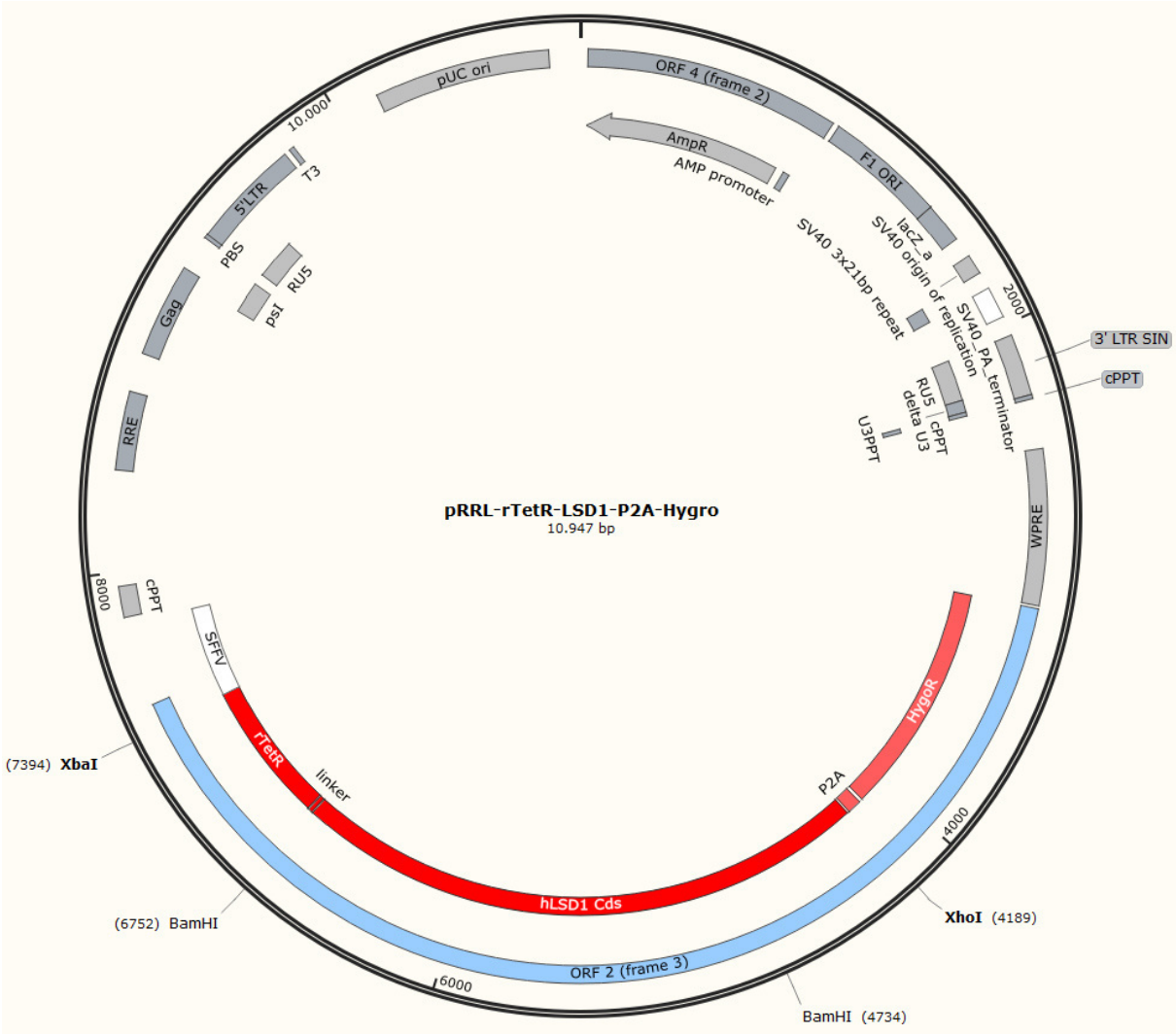
Appendix Figure 1: Plasmid map showing the viral expression vector for the rTetR-effector fusion construct. The construct containing LSD1 as a fusion protein is shown as an example.
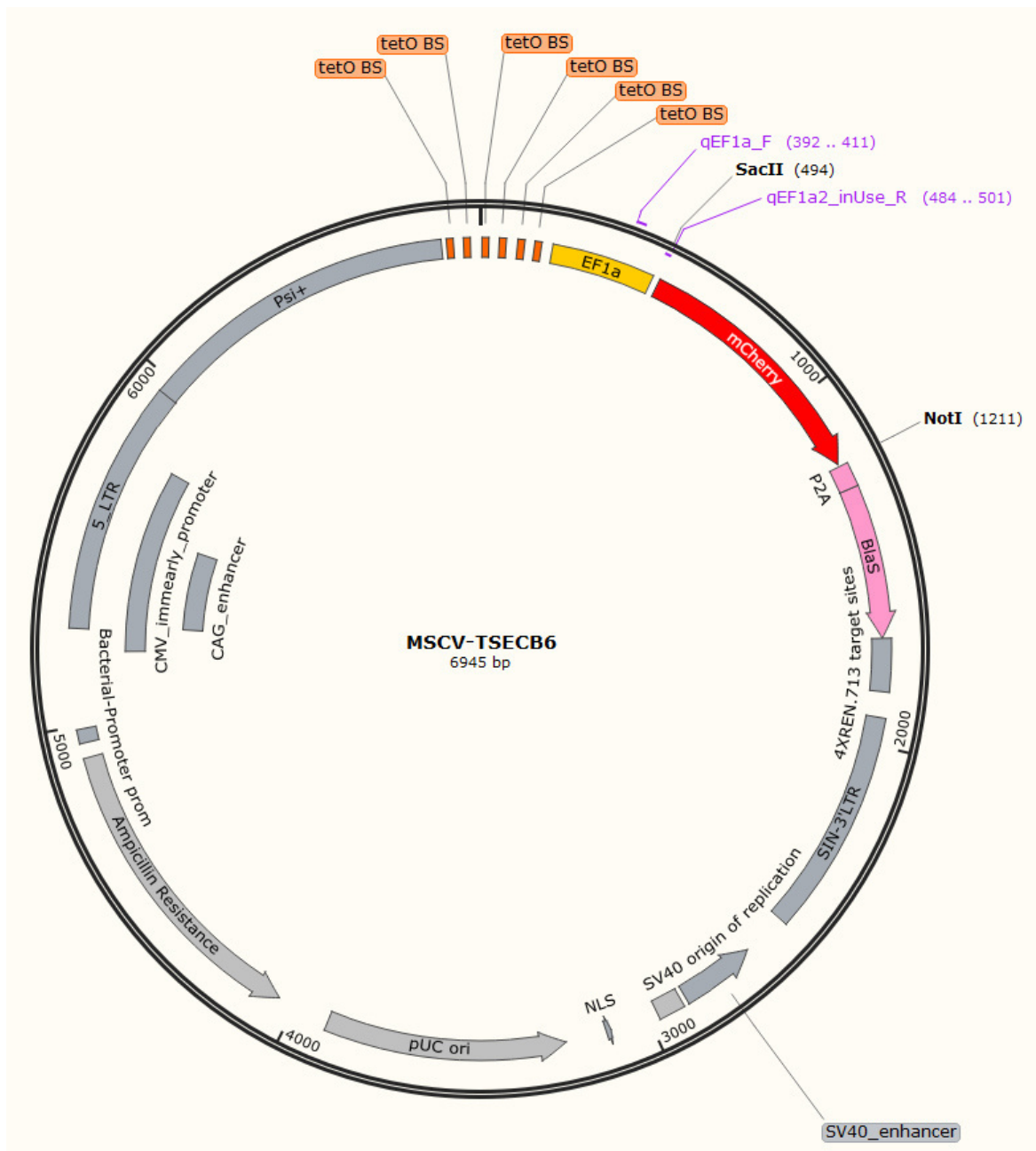
Appendix Figure 2: Plasmid map showing the viral construct for the *synP-mCherry* reporter.