

The role of acetylation of histone H3 lysine 27 in enhancer function

Kirill Jefimov

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2020

UNIVERSITY OF BERGEN



The role of acetylation of histone H3 lysine 27 in enhancer function

Kirill Jefimov



Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 09.12.2020

© Copyright Kirill Jefimov

The material in this publication is covered by the provisions of the Copyright Act.

Year: 2020

Title: The role of acetylation of histone H3 lysine 27 in enhancer function

Name: Kirill Jefimov

Print: Skipnes Kommunikasjon / University of Bergen

*“My guard stood hard when abstract threats
Too noble to neglect
Deceived me into thinking
I had something to protect
Good and bad, I define these terms
Quite clear, no doubt, somehow
Ah, but I was so much older then
I'm younger than that now”*

Bob Dylan - My Back Pages

Scientific environment

The work presented in this dissertation was carried out at The Department of Molecular Biology at the Faculty of Mathematics and Natural Sciences, University of Bergen, Bergen, Norway, between September 2013 and November 2016, and at The Department of Biosciences, Section for Biochemistry and Molecular Biology, University of Oslo, Oslo, Norway, between February 2019- February 2020. It was supervised by Professor Rein Aasland, Dr. Signe Värvi, Associate Professor Ragnhild Eskeland, and at the later stage Professor Rune Male. The project was funded by the Norwegian Cancer Society (grant to Aasland, #197477). The thesis training section was supported by The Norwegian Biochemical Society and The Norwegian Cancer Society (various travel grants to Jefimov). The study was conducted in collaboration with the labs of Professor Michiel Vermeulen and Professor Robin Andersson.

Acknowledgements

First, I would like to thank my principal supervisor, Professor Rein Aasland. Thank you for the opportunity to perform my PhD in your lab, and receive diverse experience and training in molecular biology. I am grateful for the advice and constructive criticism of the thesis and throughout my PhD. I would also like to thank my co-supervisor, Dr. Signe Värvi, for the support and constant feedback throughout all those years. Many thanks to my co-supervisor, Associate Professor Ragnhild Eskeland, for exciting discussions and providing me with valuable scientific resources. I would also like to thank my co-supervisor, Professor Rune Male, for the support at the department level.

I gratefully acknowledge the funding received towards my PhD from the Norwegian Cancer Society.

I greatly appreciate the support received through the collaborative work undertaken in Professor Michiel Vermeulen's group that opened my eyes to the potential of using mass-spectrometry. Special thanks go to Dr. Susan Kloet, who guided me during my stay in Nijmegen. I would also like to thank Dr. Arne Smith, Dr. Raghu R. Edupuganti, Pascal Jensen, Dr. Cristina Furlan, Marijke P. Baltissen, and the rest of the Vermeulen Lab.

My sincere thanks also go to Professor Robin Andersson's group for collaborative work and valuable discussions that formed my current view on enhancer-mediated regulation of gene expression. I want to pay my special regards to Nicolas Alvares for profound bioinformatical analysis.

I wish to express my deepest gratitude to Dr. Eivind Valen and his group for the opportunity to gain invaluable experience in methods of RNA biology. Exceptional gratitude goes to Dr. Teshome Tilahun Bizuayehu for guiding me throughout my stay at Valen Lab, the most exciting times developing new protocols and ideas, and motivating me to finish my PhD. I would also like to thank Dr. Kornel Labun and Gunnar Schulze for making gray, routine office days brighter.

I thank my fellow office mates Dr. Øyvind Strømland and Dr. Rhian Gaenor Jacobsen, for the support, being the best office mates ever, and harmonical co-existence within confined space.

I would also like to thank Professors Kristian Prydz and Pål Falnes, and my good colleagues Dr. Erna Davydova, Lisa Schroer, Frida Forsberg, Marta Hammerstad, Melanie Engelfriet, Dr. Julien Resseguier, Dr. Pooja Kumari, Dr. Mads Bengtsen and Alec Spera for the stimulating lunch discussions and warm welcome at the University of Oslo.

Finally, I would like to thank Maxim Brilkov, Dr. Martin Jakubec, and Dr. Evgueni Dinvay for friendship and lots of outdoor experience throughout my stay in Bergen.

Kirill Jefimov

Oslo, 2020

Table of Contents

Scientific environment	iv
Acknowledgements	v
Table of Contents	vii
List of publications	ix
Selected abbreviaions.....	x
Abstract	xiii
1. Introduction	1
1.1 <i>Transcription</i>	2
1.1.1 Transcription regulation.....	4
1.1.2 Enhancers drive specific gene expression program	7
1.2 <i>Chromatin and gene expression</i>	8
1.2.1 Chromatin – from discovery to regulatory significance	8
1.2.2 The nucleosome as basis for gene expression control.....	9
1.2.3 Organization of chromatin in 3D.....	24
1.3 <i>Enhancer chromatin</i>	27
1.4 <i>Post-translational modification of H3K27 is a switch between active and silent chromatin</i>	30
2. Aim of the study	35
3. Summary of the results.....	36
3.1 <i>The H3K27ac interactome</i>	36
3.2 <i>GBAF complex composition</i>	38
3.3 <i>GBAF localizes preferentially at intergenic/intronic regions enriched with H3K27ac</i>	39
3.4 <i>Effects of H3K27ac recognition inhibition by GBAF on histone peptide binding specificity and chromatin localization</i>	40
3.5 <i>Effects of GBAF H3K27ac recognition inhibition on transcription</i>	40
4. Discussion	44

4.1	<i>Identifying H3K27ac interacting proteins</i>	44
4.1.1	Histone-peptide pulldown MS vs ChIP-MS.....	44
4.1.2	H3K27ac is recognized by SEC and GBAF	45
4.1.3	Specificity of H3K27ac recognition.....	46
4.2	<i>GBAF complex composition</i>	47
4.2.1	GBAF – a non-canonical BAF complex.....	47
4.2.2	GLTSCR1L-mediated interactions and the GiBAF domain.....	48
4.2.3	BRD9-mediated GBAF targeting to chromatin	50
4.3	<i>GBAF localization in chromatin</i>	51
4.3.1	GBAF localizes preferentially at enhancers	51
4.3.2	Inhibition of H3K27ac recognition by BRD9 abrogates GBAF localisation in chromatin.....	53
4.4	<i>Transcriptional effects of H3K27ac recognition inhibition in GBAF</i>	55
4.5	<i>Possible role of H3K27ac in enhancer</i>	57
5.	Concluding remarks and future perspectives	58
	References	60

List of publications

The GBAF chromatin remodeling complex binds H3K27ac and mediates enhancer transcription

Kirill Jefimov, Nicolas Alcaraz, Susan L. Kloet, Signe Värvi, Siri Aastedatter Sakya, Christian Dalager Vaagenso, Michiel Vermeulen, Rein Aasland, and Robin Andersson

bioRxiv 445148; doi: <https://doi.org/10.1101/445148>

Sellected abbreviaions

ATAC-seq – assay for transposase-accessible chromatin using sequencing

ATP – adenosine triphosphate

BAF – BRG1- or BRM-associated factors

BRD9 – Bromodomain containing protein 9

CAGE – cap analysis of gene expression

CBP – CREB-binding protein

CDK – cyclin-dependent kinase

CREB – cAMP response element-binding protein

CTCF – CCCTC-binding factor

eRNA – enhancer RNA

ESC – embryonic stem cells

GLTSCR1L – Glioma tumor suppressor candidate region gene 1-like

GTF – general transcription factor

H3K27ac – acetylation of histone H3 at lysine 27

HAT – histone acetyl-transferase

HDAC – histone de-acetylase

hEGF – human epidermal growth factor

HMT – histone methyltransferase

IEG – immediate early genes

IP – immunoprecipitation

KDM – lysine demethylase

lncRNA – long non-coding RNA

miRNA – microRNA

MS – mass-spectrometry

ncRNA – non-coding RNA

NFR – nucleosome-free region

PAS – polyadenylation signal

PBAF – polybromo-associated BAF

PIC – pre-initiation complex

pTEF-b – positive transcription elongation factor beta

PTM – post-translational modification

RNAP II – RNA polymerase II

RNAP II CTD – C terminal domain of RNAP II

RSC – remodeling the structure of chromatin

RT-qPCR – reverse transcription quantitative polymerase chain reaction

siRNA – small interfering RNA

SMARCN1 – SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily X, member N

snRNA – small nuclear RNA

SWI/SNF – switch/sucrose non-fermentable

TAD – topologically associated domain

TES – transcription end site

TF – transcription factor

TFBS – transcription factor binding sites

TFIID – Transcription factor II D

TSS – transcription start site

UTR – untranslated region

VDR – vitamin D receptor

Abstract

The histone modification H3K27ac is a hallmark of active enhancers. However, its role in enhance-specific activity remains obscure. We applied mass spectrometry-based quantitative interaction proteomics to determine proteins that specifically bind H3K27ac. We identified GBAF, a non-canonical GLTSCR1L- and BRD9-containing SWI/SNF chromatin remodeling complex. GBAF was further systematically characterized in terms of protein composition and chromatin localization. A series of ChIP-seq experiments validated the interaction between GLTSCR1L and H3K27ac to be BRD9-dependent. Impairment of the H3K27ac recognition function of GBAF resulted in the dislocation of GLTSCR1L from its preferred binding sites and the genome-wide downregulation, specifically, of enhancer RNA transcription. We show that GBAF binds H3K27ac, and is an enhancer-specific chromatin remodeler involved in the transcriptional and regulatory activity of enhancers.

1. Introduction

Different organisms bear unique sets of genes encoded in their genomes. After cell division, daughter cells share the same DNA sequence. For unicellular organisms, the genome itself defines the organism. However, multicellular organisms contain many phenotypically distinct cell types that make up diverse tissues and organs using essentially the same genome. How is this possible?

All of the observed cell-type diversity develops from one single cell, the zygote, through the processes of cell proliferation and differentiation. Early embryonically dividing cells are supplied with maternally stored mRNAs and proteins. At this stage, zygotic transcription (the production of RNA molecules from a DNA template) is almost absent. After several divisions, maternal mRNA stocks get depleted, and the embryo undergoes the so-called maternal to zygotic transition, which is characterized by the activation of zygotic transcription (zygotic genome activation). Only at this point, does cell differentiation based on the zygote's own genome begin (Tadros and Lipshitz, 2009; Lee, Bonneau and Giraldez, 2014). By virtue of different regulatory mechanisms, spatially and temporally, coordinated waves of transcription differentiate the embryo throughout the development by uniquely expressed sets of genes.

Compaction of eukaryotic DNA into chromatin allows for the segregation of the genome broadly into transcriptionally active and inactive regions. These two states of chromatin create diverse transcriptional outcomes that define cell phenotypes. Transcription of genes happens in chromatin. Thus, the maintenance of chromatin conformation and transcription are directly interconnected and are regulated in coherence with each other.

Transcription is an immensely complicated process. RNA polymerase II (RNAP II) transcribes most of the genes in eukaryotic organisms. Prior to transcription, the RNAP II complex has to assemble correctly at regulatory regions, such as promoters and enhancers. Both enhancers and promoters are specific DNA sequences located near the transcribed segments of the genome. They contain context-dependent transcription factor (TF) binding sites and possess characteristic chromatin features.

Despite multiple common properties, transcription from promoter regions results in the production of mRNA molecules that are translated into protein. In contrast, transcription from enhancer elements produces enhancer RNAs (eRNA), which are usually rapidly degraded and not translated. Upon certain stimuli, chromatin undergoes structural and biochemical rearrangements which form an open conformation at those regulatory regions and bring the gene's promoter and its enhancer(s) into proximity.

An ensemble of transcription and chromatin remodeling factors shape the local chromatin environment and enable transcription (Sanyal *et al.*, 2012). Nucleosomes (DNA-histone proteins complex), the fundamental units of chromatin, may occlude the binding sites for TFs and create a barrier for RNAP II to read through. However, nucleosomes can be post-translationally modified to reduce this barrier and serve as a docking site for transcription cofactors that position and maintain them at the correct location relative to the transcription start site (TSS). Only after these chromatin rearrangements have taken place, may efficient and controlled transcription commence.

Among many potential histone post-translational modifications (PTMs), acetylation of histone H3 at lysine 27 (H3K27ac) is associated with the regulatory activity of enhancers (Creyghton *et al.*, 2010; Ernst *et al.*, 2011; Rada-Iglesias *et al.*, 2011; Rajagopal *et al.*, 2014). Although the association of H3K27ac with active regulatory elements has been known for some time, its role in enhancer function is still unclear. It is not clear, for instance, whether H3K27ac influences downstream processes of chromatin remodeling or transcription. This thesis focuses on discerning the molecular mechanisms associated with the H3K27ac histone mark in enhancer function.

1.1 Transcription

Transcription is the process of “reading” a stretch of DNA and followed by production of the corresponding RNA molecule. Transcribed RNAs can be broadly divided into two types:

-
- I. Coding or messenger RNA (mRNA) that is translated into proteins
 - II. Non-coding RNA (ncRNA) - that does not encode any proteins, but fulfills other essential functions, such as:
 1. transport of amino acids to the site of translation - transfer RNA (tRNA)
 2. structural/enzymatic - ribosomal RNA (rRNA)
 3. regulatory - microRNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), long non-coding RNA (lncRNA), enhancer RNA.

Eukaryotic transcription occurs in the cell nucleus and is accomplished by DNA-dependent RNA polymerases (RNAPs). There are three main types of RNAPs: I, II, and III. RNAP I is accountable for almost half of the total RNA production in the cell and transcribes all rRNA with the exception of 5S rRNA, which is transcribed by RNAP III. RNAP III also produces tRNAs and other small RNA molecules. RNAP II transcribes all mRNAs and most of the snRNAs and microRNAs. Transcription can occur in the 5'-3' direction along the DNA double helix on either strand, resulting in sense (+) or antisense (-) transcripts. mRNAs that are the products of eukaryotic RNAP II are typically monocistronic with short non-coding stretches before the coding region and longer ones after (5' and 3' untranslated regions, or UTRs). Mature mRNAs (and also long non-coding RNAs) undergo pre-mRNA processing: 5' capping (attachment of 7-methylguanylate cap to the first transcribed nucleotide), intron removal, 3' end formation (polyadenylation) and cleavage. In addition, RNAP II is also responsible for the transcription of eRNA.

RNAP II

Eukaryotic RNAP II is a 12-subunit holoenzyme complex. The largest subunit of RNAP II is RPB1, and it possesses the enzyme's catalytic activity. On its C-terminus, RPB1 has a motif that is repeated 52 times in mammals, called the "C terminal domain" or RNAP II CTD. The CTD motif is a highly conserved heptapeptide sequence: YSPTSPS (Egloff and Murphy, 2008) and is involved in the regulation of several steps of transcription and mRNA processing (Hirose and Manley, 2000; Maniatis and Reed, 2002; Proudfoot, Furger and Dye, 2002). All of the RNAP II CTD residues can be enzymatically modified either by phosphorylation of tyrosine, threonine, serine, or by isomerization of proline. The pattern of RNAP II

CTD modifications changes during different stages of transcription as the nascent pre-mRNA matures (Phatnani and Greenleaf, 2006; Buratowski, 2009; Perales and Bentley, 2009). Dynamic modification of the CTD is maintained mostly by kinases and phosphatases. The Serine 5 and Serine 2 phosphorylations (Ser5-P and Ser2-P) are the most conserved general marks of transcription that regulate RNAP II activity and correlate with the position of RNAP II on the transcribed genes. In yeasts and mammals, RNAP II phosphorylated at Ser5 generally accumulates at the 5' end of a gene, while Ser2-P gradually increases when polymerase approaches the 3' end.

1.1.1 Transcription regulation

Transcription is controlled at multiple stop- and check-points (De Almeida and Carmo-Fonseca, 2010) (Fig. 1):

1. assembly of RNAP II at the gene promoter (pre-initiation complex (PIC) formation)
2. initiation of RNA synthesis from TSS
3. abortive transcription
4. promoter escape
5. promoter-proximal pausing (transcriptional pausing or early elongation)
6. elongation along the gene body
7. termination after RNAP II reaches transcription end site (TES)
8. reinitiation

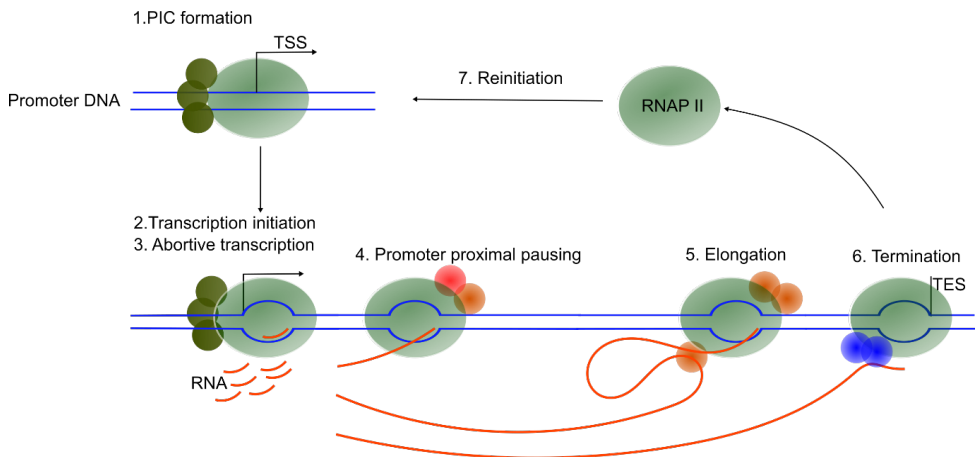


Figure 1. Transcriptional regulation. Transcription is regulated at several key steps associated with different factors (initiation-dark green, elongation-red and orange, termination-blue), that control RNAP II (light green) activity and process newly synthesized RNA molecule (orange line). PIC – pre-initiation complex. TSS – transcription start site. TES – transcription end site.

Initiation

Transcription initiation is preceded by the formation of PIC, which is composed of general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH; the Mediator complex; and RNAP II itself. Gene promoters contain specific sequences that facilitate the binding of GTFs and assembly of RNAP II at the TSS of a gene (Carninci *et al.*, 2006; Farnham, 2009).

After the PIC is assembled, RNAP II creates an open complex by unwinding ~14bp of the DNA double helix resulting in the production of short transcripts (~2-10bp long). This so-called abortive transcription (Goldman, Ebright and Nickels, 2009) continues until RNAP II escapes the promoter, Ser5 at the CTD gets phosphorylated by CDK7 (a component of the TFIIH), and the elongation phase begins.

Elongation

The phosphorylation of Ser5 at the CTD attracts 5' mRNA capping enzymes, and the nascent mRNA strand gets modified at the 5' end (5' cap or 7-methylguanosine cap/m7G-cap). The elongation phase starts after RNAP II departs from the promoter sequence leaving certain GTFs behind, transcribes a short stretch

of the coding sequence (CDS, 20-65 bp), and stops. This phenomenon is called transcriptional pausing (Gilmour and Lis, 1986; Rougvie and Lis, 1988; Weake and Workman, 2010). The paused state of RNAP II is the transition state between early and productive elongation and is thought to facilitate controlled immediate gene expression upon stimuli (Muse *et al.*, 2007; Zeitlinger *et al.*, 2007). Numerous genome-wide studies in higher eukaryotes show that most of the genes bound by RNAP II reside in a poised state for several minutes prior to productive elongation (Henriques *et al.*, 2013, 2018; Buckley *et al.*, 2014; Chen *et al.*, 2015).

Two main protein complexes are responsible for transcriptional pause control: DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF) (Renner *et al.*, 2001). DSIF consists of two proteins SPT5 and SPT4, that directly interact with RNAP II and NELF, connecting them. The NELF complex is a major mediator of transcriptional pausing. The phosphorylation pattern of CTD is dynamically changed during transcription elongation. The release from transcriptional pausing is possible only after CDK9 (a component of the pTEF-b¹: positive transcription elongation factor-beta that consists of CDK9 and CyclinT1/T2/K heterodimer; Peng *et al.*, 1998; Fu *et al.*, 1999) phosphorylates NELF, SPT5, and Ser2 at the CTD of RNAP II. Phosphorylation of SPT5 and NELF leads to the dissociation of the latter from RNAP II. Release from NELF turns DSIF into a positive transcription elongation factor, which remains bound to RNAP II until the end of a particular transcription round. Both Ser2-P and Ser5-P of the CTD are recognized by the RNA splicing machinery. Splicing is a co-transcriptional process of pre-mRNA editing that removes introns and joins exons together, resulting in mature mRNA (Allemand, Batsché and Muchardt, 2008; Pandit, Wang and Fu, 2008; Luco *et al.*, 2011).

Ser5-P is important for 5' mRNA capping (Schwer and Shuman, 2011); thus, it peaks at the 5', then gets gradually dephosphorylated by phosphatases towards the 3' end of the gene. Ser2-P levels increase towards the 3' end and begin to saturate close to 600 nucleotides (nt) downstream of the TSS, then sharply decrease at around 100

¹ Active pTEF-b associates with large protein complexes like Super elongation complex (SEC) or BRD4-containing elongation complex (BEC) (Bacon and D'Orso, 2019).

nt downstream of the poly(A) addition site (PAS) (Mayer *et al.*, 2010). A transition from Ser5-P to Ser2-P is observed at about 450 nt downstream of the TSS. Ser2-P interacts with a large complex of cleavage and polyadenylation specificity factors (CPSFs) that generate the 3' end of the mRNA.

Termination

RNAP II-transcribed genes vary in length from a few hundred bp to more than 100 kbp for small nuclear RNA (snRNA) and protein-coding genes, respectively. RNAP II is a highly processive and efficient enzyme. Thus its termination mechanisms need to be robust and reliable. Transcription termination failure may lead to severe consequences for the cell. Briefly, termination begins when the elongating RNAP II reaches the PAS. CPSFs cleave the nascent RNA strand, and a template-independent poly(A) polymerase adds ~200 adenine nucleotides to the newly synthesized RNA molecule. Then the mature mRNA is transported through nuclear pores to the cytoplasm and translated.

1.1.2 Enhancers drive specific gene expression program

Eukaryotic genomes contain diverse sets of cis-regulatory elements, such as enhancers, promoters, silencers, and insulators, which integrate cellular history² and extracellular environmental signals to mediate correct usage of the genetic information. Among these regulatory elements, enhancers drive cell type-specific gene expression, confer spatio-temporal specificity, and orchestrate gene expression patterns in response to environmental or developmental stimuli (Lam *et al.*, 2014). Enhancer elements are more numerous, and their expression patterns are more diverse and cell-type-specific as compared to coding regions and promoters (Bulger and Groudine, 2011; Ong and Corces, 2011). However, most of them are kept in a silent or poised state until an activating signal is received (Rada-Iglesias *et al.*, 2011). Strictly speaking, active enhancers possess two activities (Andersson and Sandelin, 2020):

² Cells of the organism have to 'remember' which cell type they belong to. Cellular memory or 'epigenetics' has been defined as 'heritable changes in genome function that occur without changes in the DNA sequence' (Russo, Martienssen and Riggs, 1996).

1. Regulatory – they modulate the transcription of the target promoters from a distance.
2. Transcriptional – enhancer flanking regions are transcribed and, as a result, so-called enhancer RNA produced.

Enhancers modulate transcription of their target genes at distances ranging from a few to hundreds of kilo base-pairs (Bulger and Groudine, 2011; Ong and Corces, 2011, 2012). Furthermore, multiple enhancers can act on a single promoter, and one enhancer can control multiple promoters. Both enhancer and promoter regions contain transcription factor binding sites (TFBS), which, if recognized by TFs, stimulate transcription either through direct or indirect (coactivator-mediated) contacts with RNAP II. In most cases, however, TFs recruit chromatin remodeler proteins that shape the local chromatin structure appropriate for PIC formation and transcription initiation (Li, Carey and Workman, 2007; Weake and Workman, 2010; Bell *et al.*, 2011; Spitz and Furlong, 2012).

1.2 Chromatin and gene expression

Although transcription is a complex process with multiple stop- and check-points, it is quite pervasive (Birney *et al.*, 2007). It has been shown that cryptic, non-coding transcripts can originate from nucleosome-free regions (NFR), although they do not seem to have any function and flood the transcript pool of a cell (Jensen, Jacquier and Libri, 2013). So, despite the RNAP II activity being tightly regulated, cells use another essential level of transcriptional regulation, which is provided by the chromatin.

1.2.1 Chromatin – from discovery to regulatory significance

After DNA and histone proteins were first purified from the cell nucleus (Miescher-Rüsch, 1871), Flemming introduced the term “chromatin” (Mayr, 1982). Later, in the first half of the 20th century, ideas about chromosome function, organization and replication started to appear. Kolcov proposed one of the first thoughts that resembled the complementarity principle known today in 1926 (Morange, 2011): “omnis molecula ex molecula”, which meant that each daughter

cell would need at least one copy of the chromosomes to replicate itself. Due to the simplicity of DNA (alphabet of four letters) and the complexity of proteins (alphabet of 20 letters), the idea that DNA can carry any hereditary information was rejected (Stedman and Stedman, 1947). In the 1940s and 50s, however, it became clear that DNA, but not proteins, is the carrier of genetic information (Avery, 1944; Hershey, 1952). Finally, the structure of the DNA double helix was discovered in 1953, and the complementarity principle was established explaining how genetic information could be stored, replicated, and inherited (Franklin and Gosling, 1953; Watson and Crick, F. H. C., 1953). In a way, the pendulum had swung back, and it seemed that the DNA sequence holds all the answers.

Histones tightly bind and condense the DNA, which makes the underlying DNA sequence inaccessible to RNA polymerases and other DNA-interacting proteins. It was initially believed that histones serve only structural purposes, such as folding/packaging/condensing of DNA, and are inhibiting DNA-dependent processes such as transcription and replication (Stedman and Stedman, 1950; Huang, 1962). However, in the 70s, it was shown that transcription from the DNA-histone complex was possible and histones were no longer considered to be the ultimate obstacle, but a means of transcriptional regulation (Cedar and Felsenfeld, 1973). Since then, histone–DNA interactions have been extensively studied, and are now quite well understood (Kornberg and Thonmas, 1974; Gina Arents *et al.*, 1991; Arents and Moudrianakis, 1993; Luger *et al.*, 1997). The current view that chromatin fulfills both structural and complex regulatory functions in the mammalian cell has been developed after a plethora of protein complexes, which modulate conformational, dynamic, and covalent changes of chromatin were discovered (Clapier and Cairns, 2009).

1.2.2 The nucleosome as basis for gene expression control

The human cell nucleus (about 10 μm in diameter) fits about 2 meters of DNA, which requires sophisticated compaction into chromatin. The nucleosome – a repeating unit of chromatin, is a complex of an octamer of four different histone proteins (two dimers of H2A-H2B, and a tetramer of H3-H4) and about 200 base

pairs of DNA (Kornberg and Thonmas, 1974) assembled *in vivo* in a stepwise fashion.

Histones are hydrophilic, basic proteins. The majority of the histone mass is an α -helical domain (histone fold domain), which forms a spool-like structure that is wrapped by DNA with 1.67 left-handed super-helical turns or \sim 150 bp long DNA (Richmond and Davey, 2003; Campos and Reinberg, 2009). Histones alone compact overall DNA volume by a factor of 5-10. Nucleosome arrays are organized as beads-on-a-string (10 nm fiber) with internucleosomal spacing of about 60 bp (Prieto and Maeshima, 2019).

Another type of histone proteins – linker histones (such as histone H1) which are not a part of the nucleosome, immobilize the nucleosome preventing it from sliding along the DNA and help to fold chromatin into higher-order structures (the so-called 30 nm fiber) (Thomas, 1984; Bednar *et al.*, 2017). The total compaction factor of highly ordered chromatin with histone H1 is \sim 50 (Brown, 2003). Compaction of the DNA into chromatin prevents cryptic transcription and protects DNA from the damage.

Inaccessibility of the DNA packaged into chromatin may seem like an obstacle for gene expression and transcription. However, mechanisms for chromatin regulation turn it into an advantage, facilitating controlled access to DNA. Chromatin remodeling is a process of altering DNA accessibility by altering positioning or physicochemical properties of nucleosomes and, consequently, the structure of the chromatin fiber by means of histone PTMs, ATP-dependent chromatin remodeling, histone variants and nucleosome turnover (Fig. 2). The interplay between these mechanisms ensures controlled access to DNA despite the protective nature of chromatin and is integral to every aspect of genome function.

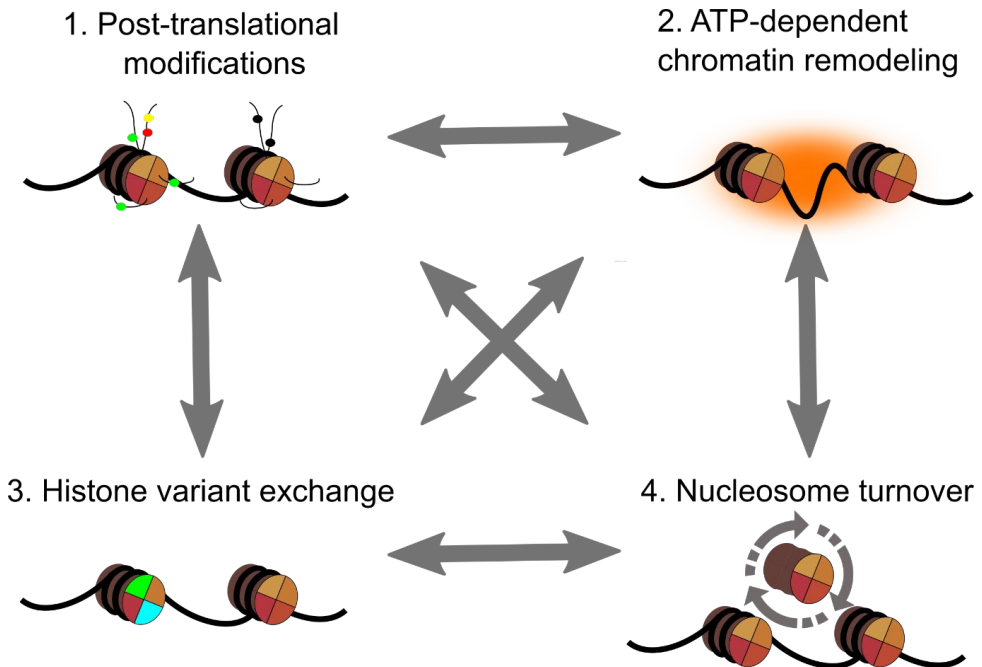


Figure 2. Mechanisms of chromatin regulation. DNA in black. Nucleosome octamer is depicted as spool, where histones in shades of yellow, dark orange, brown, and red. Post-translational modifications are visualized as small colored circles on histone tails. Histone variants are depicted as green and blue histones. The regulation of NFR length is depicted in orange.

Histone post-translational modifications

In the early 1960s, Vincent Allfrey and colleagues found that histones were modified by post-translational acetylation and methylation, and proposed that these modifications correlate with the control of gene expression (Allfrey, Faulkner and Mirsky, 1964). Most of the histone modifications reside at the so-called histone tails. Histone tails (Fig. 3A) comprise up to 38% of the histone mass and are seemingly unstructured (G. Arents *et al.*, 1991; Luger *et al.*, 1997). Initially, histone tails were defined by their sensitivity to proteases (Böhm and Crane-Robinson, 1984). These domains reside at the N-termini of all four core histone proteins and additionally for H2A at the C-terminus. The tails of histones H4 and H3 are evolutionarily highly conserved sequences and contain multiple lysine and arginine residues, which make them positively charged and facilitate their interaction with DNA. When fully extended, they protrude far beyond the DNA in the nucleosome (Fletcher and

Hansen, 1996). Remarkably, removal of the histone tails does not alter the overall core structure of nucleosomes (Dong, Hansen and van Holde, 1990; Hayes, Clark and Wolffe, 1991). They do, however, take part in compacting chains of nucleosomes into the 30 nm chromatin fiber (Garcia-Ramirez, Dong and Ausio, 1992; Garcia-Ramirez, Rocchini and Ausio, 1995) and higher-order chromatin structures (Tse and Hansen, 1997; Tse *et al.*, 1998).

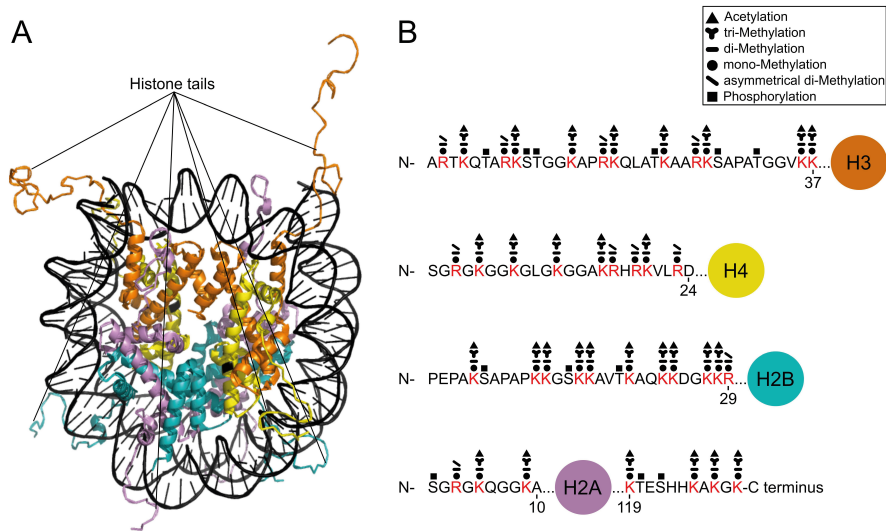


Figure 3. Histone tails and their PTMs. A. Structure of mononucleosome (PDB code 1KX5) with protruding histone tails. DNA in black. Histones H3, H4, H2A, and H2B, are colored orange, yellow, violet, and teal, respectively. B. Specific amino acid positions of histone tail PTMs.

Histone tail PTMs play a central role in regulatory mechanisms involving chromatin by attenuating intra- and inter-nucleosomal interaction potential of the histone tails which in turn influences the folding and functional state of the chromatin fiber, recruitment of the transcription factors, assembly of chromatin-modifying complexes, and modulation of other non-histone protein interactions (Grunstein, 1997; Howe *et al.*, 1999; Jenuwein, 2001; Bannister and Kouzarides, 2011). They affect chromatin structure, nucleosome dynamics, and serve as important binding sites for diverse chromatin-binding proteins (“readers of PTMs”) (Mersfelder and Parthun, 2006; Taverna *et al.*, 2007). The current repertoire of histone marks comprises at least 130 different PTMs (Tan *et al.*, 2011). The most studied histone

PTMs are acetylation, methylation, ubiquitination of lysines, and phosphorylation of serine and threonine residues (Kouzarides, 2007). Histone PTMs are introduced by different enzymes (PTM “writers”), such as kinases, histone acetyltransferases (HATs), and histone methyltransferases (HMTs). The vast majority of histone PTMs are reversible and are removed by the so-called “eraser” enzymes: phosphatases, histone deacetylases (HDACs), lysine demethylases (KDMs). The coordinated action of these PTM “writers” and “erasers” with other regulatory mechanisms ensures correct genome function through the cell cycle based on cellular memory³ and in response to external cues.

Histone tail PTMs occur at specific positions (Fig. 3B). Despite a wide variety of modifications, they can occur only in a limited number of combinations. The resulting patterns of histone tail PTMs are associated with different biological processes (for example, acetylation of histone H3/H4 is associated with transcription). It has been suggested that multiple PTMs act sequentially or combinatorially to regulate specific biological outcomes. Hence, combinations of histone modifications that are not mutually exclusive have the potential to form a so-called ‘histone code’ (Turner, 1993; Strahl and Allis, 2000).

Such chromatin states can be maintained stably through cell divisions, and histone PTMs are carriers of this cellular memory³. They also have to be flexible to allow changes in transcriptional programs during development or differentiation. Breakdown of any of these two properties is a characteristic feature of many diseases, such as cancer.

³ Histone variants and PTMs, as well as DNA methylation (not discussed here) can affect transcriptional outcome and can be heritably transmitted through mitosis or meiosis (Bird, 1978; Ng. and Gurdon, 2008; Hathaway *et al.*, 2012), and therefore serve as mediators of epigenetic/cellular memory. However, since the definition of ‘epigenetics’ implies that particular state or mark that defines cell identity has to be heritable and maintained, not all of the histone PTMs can be defined as ‘epigenetic’ marks (Berger *et al.*, 2009; Audergon *et al.*, 2015; Ragunathan, Jih and Moazed, 2015; Henikoff and Grealley, 2016).

Histone acetylation

Acetylation is a process of transfer of the acetyl group from Acetyl-coenzyme A to an amino acid residue side chain. There are two types of protein acetylation: N-terminal (that may happen at any amino acid if it is at the N-termini of the protein) and N- ϵ -lysine acetylation. Histones are subject to both types of acetylation. N-terminal acetylation is an irreversible co-translational modification, whereas N- ϵ -lysine acetylation is a reversible and dynamic PTM. Only a few studies have so far addressed the function of the N-terminal acetylation of histone proteins, and its role is yet unclear. Thus, I will focus only on the N- ϵ -lysine acetylation henceforth.

Histone proteins were the first proteins discovered to be acetylated.

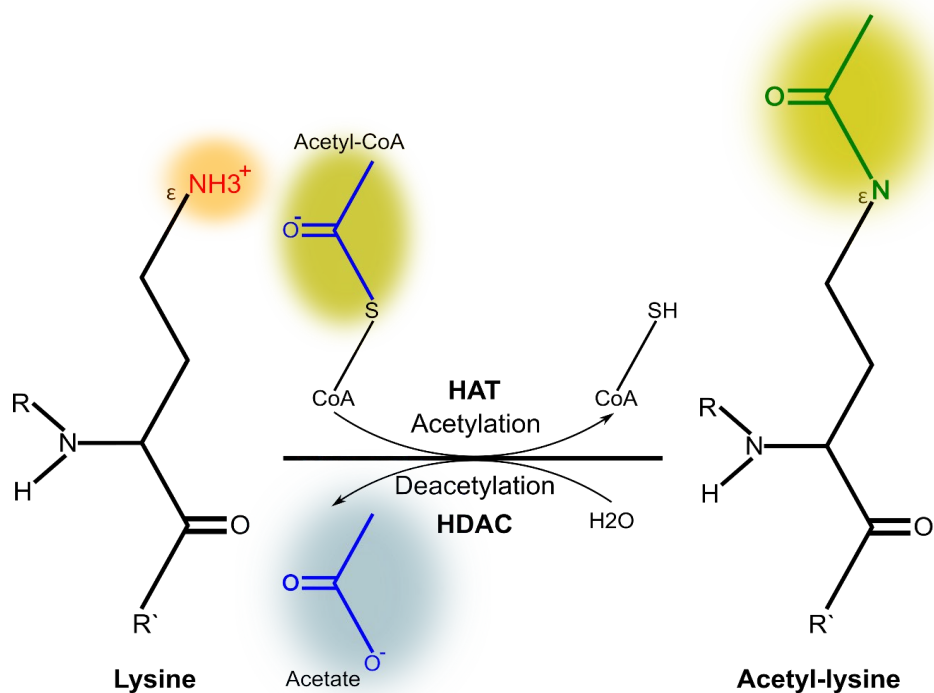


Figure 4. Lysine acetylation. Histone N- ϵ -lysine acetylation is conducted by histone acetyltransferase (HAT) enzymes. The positive charge of lysine, provided by its amine group (in red) is neutralized by transfer of the acetyl group (green) from Acetyl-CoA. Lysine acetylation is reversible and conducted by histone deacetylases (HDAC), releasing acetate (blue).

Histones contain multiple lysines and arginines (about 30% of all amino-acid residues), which are positively charged amino acids that facilitate DNA-histone interaction. Both histone tails and core regions are subject to lysine acetylation (Peterson and Laniel, 2004). Histone lysine acetylation (Fig. 4) is performed by HATs and reversed by HDACs.

Histone lysine acetylation neutralizes the positive charge of lysine, reducing the strength of internucleosomal interactions which relaxes the chromatin fiber (Tse *et al.*, 1998; Annunziato and Hansen, 2000; Shogren-Knaak *et al.*, 2006; Robinson *et al.*, 2008), DNA-histone interaction (Oliva *et al.*, 1990; Bresnick, John and Hager, 1991; Widlund *et al.*, 2000), and consequently relieves the overall nucleosomal barrier for transcription with little effect on the physical stability of the nucleosome. Histone acetylation was also shown to enhance the disassembly of the nucleosome by the histone chaperone Nap1, which facilitates the generation of NFR and, consequently, transcription initiation (Sharma and Nyborg, 2008).

Many chromatin interacting proteins that modulate nucleosome dynamics, maintain or alter histone modifications, conduct chromatin remodeling, and promote transcription can bind histone tails (and nucleosomes in general). Several protein domains, such as the Bromodomain, the YEATS, and some PHD fingers, are specific ‘readers’ of acetylated lysines in histones. These domains are often combined with other chromatin-binding and enzymatic domains in multidomain proteins or protein complexes (Mujtaba, Zeng and Zhou, 2007; Li *et al.*, 2014; Fujisawa and Filippakopoulos, 2017). In many of the chromatin-interacting proteins that contain a Bromodomain, it appears more than once. Thus these proteins are capable of binding more than one acetylated residue. Histone tails frequently bear several acetylation marks (Schwämmle *et al.*, 2014), and spacing between lysines in histone tails appears to be strikingly regular (Fig. 5). Multiple lysine-acetylations on histone tails can enhance the interaction with Bromodomains (for example, di-acetylation stabilizes the interaction of a single Bromodomain with a histone tail; Morinière *et al.*, 2009).



Figure 5. Regularity of lysines in histone H3.1 and H4 N-terminal tails. Sequences of H3.1 and H4 N-terminus and positions of the lysines (red). In Histone H3.1 lysines are regularly spaced with 3 or 4 residues (black numbers above top brackets) which also forms a pattern of lysines that are spaced with the step of 8 residues (black numbers below bottom brackets). Histone H4 has simple spacing of 3 residues between lysines.

Histone H3 is mostly acetylated at three sites: H3K14, H3K18 and H3K23 (Schwämmle *et al.*, 2014). These acetylated sites are enriched in euchromatin and are found in broad acetylation regions. Broad acetylation hinders chromatin folding into highly compact structures (Garcia-Ramirez, Rocchini and Ausio, 1995; Tse *et al.*, 1998), making it accessible for interacting proteins (Hebbes *et al.*, 1994; Krajewski and Becker, 1998). Albeit broad histone acetylation marks transcriptionally active chromosome regions, it is not tightly correlated with transcription *per se*, but considered as a general marker of euchromatin.

In contrast to broadly acetylated chromosome patches, regulatory and transcriptionally active regions such as promoters and enhancers, are labeled additionally by localized, targeted acetylation (Brown *et al.*, 2000; Forsberg and Bresnick, 2001; Litt *et al.*, 2001).

Even though general mechanisms of histone acetylation-dependent transcription activation are clear, specificity towards different regulatory regions of the targeted acetylation, its gene regulatory outcomes, and significance are far from being understood.

Histone variants

The eukaryotic genome contains multiple copies of histone genes, and their expression is tightly controlled (Marzluff, 2005; Kurat *et al.*, 2014). Canonical histone proteins are only produced during the S phase of the cell cycle and deposited across the genome in a replication-dependent manner. However, there are histone protein variants (paralogues) that have specific functions in different processes such as transcription, cell division, and DNA repair (Talbert and Henikoff, 2010). In contrast to canonical histones, synthesis of histone variants is cell cycle-independent and their deposition is targeted to specific sites, which differentiate the chromatin. Histone variants change the properties of nucleosomes they reside in, due to sequence variations at structurally and functionally important regions. Replacement of canonical histones by histone variants alters the chromatin landscape, affecting nucleosome dynamics and interaction with chromatin remodelers.

One of the most striking examples of how histone variants differentiate chromatin is CENP-A, a histone H3 variant that facilitates centromeric region formation and plays an essential role in chromosome segregation. CENP-A contains a histone fold and N-terminal tail domain; however, it shares only about 50% of sequence identity with conventional H3 and has different structural and PTM features (Sharma *et al.*, 2019). Another H3 variant H3.3 is deposited at transcriptionally active chromatin (Filipescu, Müller and Almouzni, 2014). H3.3 differs in only four residues from canonical H3.1, and three of those are explicitly meant for replication-independent (RI) deposition. Actively transcribed chromatin requires RI histone replacement since elongating RNAP II disassembles nucleosomes (Dion *et al.*, 2007), and chromatin reassembly itself changes the dynamics of chromatin (Schwartz and Ahmad, 2005). Also histone H2A has several variants that are associated with different processes. H2A.Z, similarly to H3.3, is observed at highly transcribed loci and regulatory regions. This histone variant reduces the stability of nucleosomes and thus lowers the transcriptional barrier (Weber, Henikoff and Henikoff, 2010). Histone variants as well, as their canonical paralogues, are subjects to numerous PTMs. Those that are similar to their canonical counterparts, such as H3.3, bear both similar and variant-specific modifications (McKittrick *et al.*, 2004; Chang *et al.*,

2015), the diverged histone variants, such as H2A.Z and CENP-A, have only variant-specific PTMs.

Nucleosome turnover

Nucleosomes are exchanged several times throughout the cell cycle (Dion *et al.*, 2007; Deal, Henikoff and Henikoff, 2010; Radman-Livaja *et al.*, 2011). During replication, transcription, and DNA repair, chromatin is disassembled and reassembled both with newly synthesized and parental nucleosomes. Although parental nucleosomes are transducers of cellular memory and chromatin differentiation, replication dilutes previously deposited histone variants and PTMs (Venkatesh and Workman, 2015). After DNA replication, old nucleosomes are reinstated close to their original positions, which preserves parental local chromatin state and possibly stimulates chromatin-modifying enzymes and remodelers to propagate similar PTMs and histone variants (Reverón-Gómez *et al.*, 2018). Replication dependent chromatin assembly pathway involves the action of numerous histone chaperones (such as chromatin assembly factor 1 (CAF-1) and anti-silencing function 1 (ASF1)) that can interact only with newly synthesized histone H3.1-containing nucleosomes. These nucleosomes are specifically acetylated at the H3K56 position, which makes them unstable and creates loosely packed chromatin, which is susceptible to inappropriate, cryptic transcription initiation and requires further differentiation by the chromatin remodelers (Li *et al.*, 2008; Li, Burgess and Zhang, 2012).

Extensive chromatin remodeling and control over the nucleosome turnover is required in all stages of transcription to grant RNAP II, and cofactors temporally appropriate access to the DNA and reassemble the chromatin (Williams and Tyler, 2007). Active transcription sites have elevated rates of nucleosome turnover, facilitated by the passage of RNAP II, site-specific histone variants (such as H2A.Z and H3.3), and acetylated nucleosomes (Lai and Pugh, 2017).

ATP-dependent Chromatin remodeling

Nucleosome remodeling is the process of altering histone-DNA interactions by disrupting, assembling, or moving nucleosomes along the DNA. Due to the intrinsic

stability of the nucleosomes, chromatin remodeling reactions require ATP hydrolysis. ATP-dependent remodelers are multi-subunit, large molecular complexes that regulate chromatin architecture (Wu, Lessard and Crabtree, 2009; Kadoch and Crabtree, 2015; Masliah-Planchon *et al.*, 2015). These complexes contain DNA translocases that hydrolyze ATP and alter histone-DNA interactions in bound nucleosomes. As the main catalytic unit, they all contain SNF2-related ATPase domain, which is related to a larger group of nucleic acid helicases (Flaus *et al.*, 2006; Flaus and Owen-Hughes, 2011). SNF2 family helicases bind double-stranded DNA and move along one strand separating the double helix. Nucleosome remodelers, however, are DNA translocases that move along the DNA without separating the double helix (Saha, Wittmeyer and Cairns, 2006; Gangaraju and Bartholomew, 2007). According to sequence similarities of Snf2p-related ATPase-domains (Flaus and Owen-Hughes, 2011) chromatin remodelers can be divided into four major families: the switch/sucrose non-fermenting (SWI/SNF), inositol requiring 80 (INO80), the imitation switch (ISWI), and chromodomain helicase DNA-binding (CHD/M2).

Table 1. Families of ATP-dependent chromatin remodelers, their remodeling activity and biological context.

Remodeler Family	ATP-ase subunit	Remodeling activity	Biological context
SWI/SNF	Brg1 or BRM (SMARCA2/4)	Nucleosome sliding ¹ , DNA unwinding ⁷ , dimer ejection ⁴ and octamer eviction ³ .	Transcription - NFR spacing, chromatin accessibility. DNA replication fork progression, DNA damage repair, recombination ⁹ .
ISWI	SNF2L or SNF2H	Nucleosome sliding ⁵ .	Transcription - NFR spacing, chromatin accessibility ⁹ .
CHD/M2	CHD1-9	Nucleosome sliding ⁶ .	Transcription - NFR spacing, chromatin accessibility, and editing (incorporating histone H3.3) ⁹ .
INO80	INO80, Tip60, SRCAP	Nucleosome sliding ⁸ , histone dimer exchange ² .	Transcription - NFR spacing, histone variant exchange. DNA replication fork progression, DNA damage repair, recombination ⁹ .

Source: 1-Whitehouse *et al.*, 1999; 2-Mizuguchi *et al.*, 2004; 3-Lorch, Zhang and Kornberg, 1999,4-Bruno *et al.*, 2003; Yang *et al.*, 2007, 5-Hamiche *et al.*, 1999; Längst *et al.*, 1999,6-Winger *et al.*, 2018; 7-Kassabov *et al.*, 2003, 8-Eustermann *et al.*, 2018; 9-Clapier *et al.*, 2017.

Members of each family frequently share particular domains and features other than the ATPase domain. For example, SWI/SNF family complexes contain Bromodomains that “read” acetylated lysine marks of histone and non-histone proteins, and CHD/M2 family members contain methylated lysine “reader” chromodomains (Yap and Zhou, 2011), and ISWI family contains SANT-SLIDE domains that bind the unmodified histone H3 tail (Boyer, Latek and Peterson, 2004). The histone PTM “reader” domains recognize histone PTMs or variants and modulate the activity of ATPase (Paul and Bartholomew, 2018).

ATP-dependent nucleosome remodeling mechanisms (Fig. 6A) may affect all levels of chromatin organization (Fig. 6B):

1. Nucleosome assembly or disassembly
2. Nucleosome sliding
3. Nucleosome unwinding
4. Nucleosome or histone dimer ejection
5. Nucleosome or histone dimer exchange

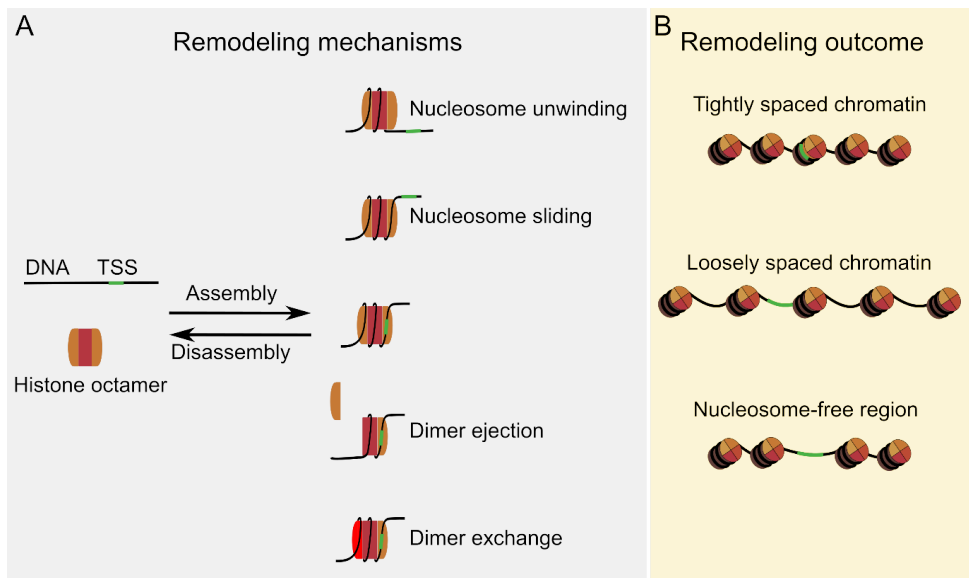


Figure 6. Chromatin remodeling mechanisms and outcomes. A. ATP-dependent chromatin remodeling mechanisms: DNA showed in black, Histone proteins in orange and bordo, histone variant in red, TSS in green. B. Some of the remodeling outcomes.

These five mechanisms result in *de novo* nucleosome assembly, histone eviction, nucleosome positioning, spacing, and histone variant exchange, which facilitates chromatin folding, specialization, and regulates DNA accessibility. One of their main functions is to remodel nucleosomes at regulatory regions such as promoters and enhancers, to allow for the establishment of gene expression programs while maintaining the integrity of the chromatin fiber.

Specification of the ATP-dependent remodeling complexes, their cofactors, and preference towards certain nucleosome substrates can define the outcome of chromatin remodeling reactions (Dann *et al.*, 2017). Despite many similarities, members within the same family of chromatin remodelers, can have distinct targets and functions due to a large number of paralogs among their subunits. The subunit composition of these complexes is frequently tissue-specific, and despite the transience of the reactions they catalyze, they can be involved in the formation and maintenance of stable (epigenetic) chromatin states (Becker and Workman, 2013). Probably the most versatile remodeling complex is mSWI/SNF, which is involved in regulating diverse gene expression programs (Hota *et al.*, 2019; Witwicka *et al.*, 2019).

SWI/SNF complex

The SWI/SNF complex was the first ATP-dependent chromatin remodeling complex to be discovered (Stern, Jensen and Herskowitz, 1984). It has been shown to play a vital role in the regulation of gene expression during animal development (Archacki *et al.*, 2009; Clapier and Cairns, 2009; Ho and Crabtree, 2010). Mammalian SWI/SNF (mSWI/SNF) is a family of large (1 to 1.5 MDa) heterogeneous, combinatorically assembled complexes. About 30 gene products generate diversity in the composition of mSWI/SNF complexes that have specialized functions in tissue-specific gene regulation during development (Wang *et al.*, 1996; Lessard *et al.*, 2007). mSWI/SNF complexes can function both as transcriptional activators or repressors. The same complex can even switch between two opposite modes of action at the same gene locus (Gkikopoulos *et al.*, 2011). Tissue-specific mSWI/SNF complexes interact with various transcription factors in different cell-types allowing these complexes to fulfill context-dependent functions (Shi *et al.*,

2013; Murakami *et al.*, 2017). Mutations of mSWI/SNF complex subunits are implicated in a wide range of diseases from cancer to neurological disorders (Kadoch *et al.*, 2013; Lopez and Wood, 2015; Bögershausen and Wollnik, 2018).

Mammalian SWI/SNF complexes are thought to be organized in a modular fashion consisting of a core module (SMARCC1⁴, SMARCD1/2/3, SMARCE1), complex-specific subunits, and a catalytic ATPase module (SMARCA2/4, BCL7A/B/C, ACTL6A) (Mashtalir *et al.*, 2018). The function of the catalytic module is to exert translocation of DNA bound by nucleosome, while the rest of the modules probably regulate (accelerators/brakes), or sense the DNA to allosterically regulate nucleosome mobilization (Paul and Bartholomew, 2018).

mSWI/SNF remodelers contain multiple DNA, protein-protein, and histone tail interaction domains. mSWI/SNF-DNA interactions are sequence-independent and thought to enhance nucleosome binding (Mohrmann and Verrijzer, 2005).

mSWI/SNF contains multiple histone recognition domains, such as Bromodomain (PBRM1, BRD7, BRD9, SMARCA2/4) and PHD finger (DPF1/2/3, PH10), which facilitate nucleosomal interactions. Histone tails are required for efficient recruitment of mSWI/SNF complexes to the chromatin and regulation of its remodeling activity. Different histone tail PTMs are found to either enhance or block nucleosome remodeling by SWI/SNF (Corona *et al.*, 2001; Ferreira, Flaus and Owen-Hughes, 2007; Chatterjee *et al.*, 2011). SWI/SNF can recognize acetylated nucleosomes, displace, or partially unwrap nucleosomal DNA, enabling transcription (Lorch and Kornberg, 2015).

The main catalytic subunit of all mSWI/SNF complexes is one of the two DNA-dependent ATPases, SMARCA4 (Brahma-related gene 1- BRG1) or SMARCA2 (Brahma - BRM). Apart from the ATPase domain SMARCA2/4 contain multiple other domains that facilitate protein-protein interactions (HSA domain and QLQ domain; Trotter *et al.*, 2008), acetylated nucleosome binding (Bromodomain; Tamkun *et al.*, 1992), and nonspecific DNA binding (AT-hook domain; Singh, D'Silva and Holak, 2006). Despite being highly similar proteins (86% similarity;

⁴ Gene names of SWI/SNF family members stand for SWI/SNF related, matrix associated, actin-dependent regulator of chromatin, subfamily X, member Y - SMARCC

Muchardt and Yaniv, 2001), they have both overlapping and specific targets in chromatin and can have distinct or similar effects on gene expression. The difference in N-terminal amino acid sequences between SMARCA2 and SMARCA4 allows interaction with distinct transcription factors and might be the reason for the observed differential targeting in chromatin (Kadam and Emerson, 2003; Raab *et al.*, 2017). These subunits are mutually exclusive in assembled complexes. The knockout of SMARCA4 is lethal, whereas SMARCA2 is dispensable for viability. SMARCA4 was found to be required for the zygotic gene activation (Bultman *et al.*, 2006) at the developmental stage when embryonic transcription begins. Furthermore, the SWI/SNF complex stays bound to RNAP II to evict nucleosomes during elongation (Schwabish and Struhl, 2007).

Three distinct types of mSWI/SNF complexes have been reported, and they can be divided further into many cell-type-specific assemblies. Two of the most studied canonical types of SWI/SNF assemblies are BAF (Brahma Associated Factors) and PBAF (Poly-Bromo, Brahma Associated Factors) (Ho and Crabtree, 2010). Recently another, a non-canonical BAF complex has been independently reported as GBAF by Alpsy and Dykhuizen, 2018 and this work, and ncBAF by Mashtalir *et al.*, 2018.

PBAF is homologous to RSC (Remodeling the Structure of Chromatin) complex in yeast (Cairns *et al.*, 1996). The defining components of this complex are ARID2, PHF10, BRD7, and, most importantly, PBRM1 (Protein polybromo-1). PBAF can utilize both SMARCA4 or SMARCA2 as the catalytic subunit (Simpson *et al.*, 2015). This complex has clear transcription activating functions in yeast. It occupies about 10% of all genes during normal conditions but is indispensable for transcription activation during stress response (Sudarsanam *et al.*, 2000; Shivaswamy and Iyer, 2008; Dutta *et al.*, 2014; Lorch and Kornberg, 2017). RSC in yeast is involved in the control of NFR architecture at the gene promoter, which is flanked by firmly positioned +1 and -1 nucleosomes, with the transcription start site typically 10–15 bp inside the +1 nucleosome. RSC binds to NFR, attaches to +1 nucleosomes, and translocates DNA to slide or eject the +1 promoter nucleosome to extend NFR further and expose TSS and facilitate transcription (Ye *et al.*, 2019). In addition, RSC is involved in numerous other chromosomal processes which require the same

fundamental activities of DNA unwrapping and DNA translocation. These processes include DNA repair, chromosome segregation, and other chromosomal DNA transactions.

The composition of the BAF complex is species-specific. In mammals, the BAF complex is more abundant than PBAF, may contain either SMARCA4/2 as the catalytic subunit, and several other complex-specific subunits, such as ARID1A/1B, DPF1/2/3, and SS18/18L (Simpson *et al.*, 2015). So far, it is not clear how and whether the catalytic activity of this complex differs from PBAF. Nevertheless, the BAF complex is involved in a wide range of chromatin-associated processes similar to PBAF, with opposing or overlapping functions.

GBAF complex is the smallest SWI/SNF assembly (Alpsoy and Dykhuizen, 2018; Mashtalir *et al.*, 2018; this work). It lacks DNA interaction domains and has Bromodomains only in SMARCA2/4 and BRD9 subunits. This feature makes GBAF potentially highly dependent on acetylated lysine recognition for targeting to the nucleosomes. However, the functions of this complex remain elusive.

1.2.3 Organization of chromatin in 3D

The genome is not randomly folded in the interphase nucleus. Each chromosome occupies a particular volume termed chromosome territory (Fig. 7) (Cremer and Cremer, 2010), overlapping with other chromosomes only at the borders of chromosome territories (Branco and Pombo, 2006). Chromosome territories are radially positioned in the cell nucleus, where gene-rich chromosome regions reside at the center and gene-poor at the periphery (Bolzer *et al.*, 2005). The radial arrangement of the chromosome territories allows spatio-temporal regulation of the cell nuclear processes (Bickmore and Van Steensel, 2013). Chromosome territories can be further divided into two main compartments (A and B) depending on chromatin compaction. Compartment A consists of accessible, transcriptionally active chromatin, whereas compartment B consists of dense, inactive chromatin (Lieberman-Aiden *et al.*, 2009; Rao *et al.*, 2014; Dixon *et al.*, 2015). Compartments A and B can be further subdivided into five principal types of chromatin (two types of the accessible and three types of the dense chromatin) that are conserved among metazoans and defined by unique yet overlapping combinations of bound proteins

and histone PTMs (Filion *et al.*, 2010). Compartment partitioning is cell type-specific and switches throughout cell differentiation (Dixon *et al.*, 2015). Compartments consist of several thousands of topologically associated domains (TADs) - fundamental units of genome organization (Nora *et al.*, 2012; Dekker and Heard, 2015; Dixon, Gorkin and Ren, 2016).

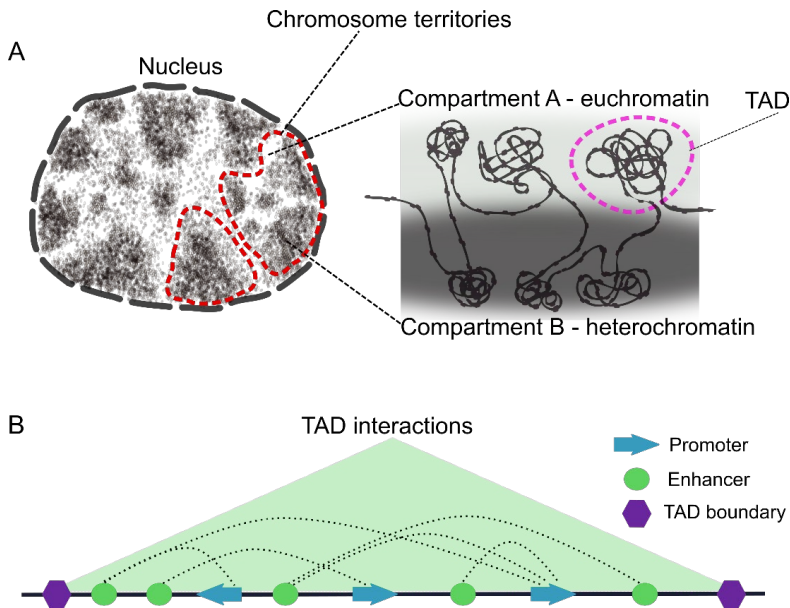


Figure 7. Chromatin organization in the nucleus. A. Individual chromosomes occupy confined space within the nucleus - chromosome territories. Chromosome territories are further split broadly into two compartments A and B, which contain active chromatin and repressed chromatin, respectively. B. Compartments are further organized into TADs, which confine locally interacting chromatin. Interactions are visualized by the dotted line.

In mammals, TADs are chromosomal domains with a median size of 880 kb (Dixon *et al.*, 2012). Positions of TADs are evolutionarily conserved (Dixon *et al.*, 2015), and each TAD represents a separate replication domain (Pope *et al.*, 2014). The frequency of genomic interactions within the same TAD is much higher than between different TADs. Gene regulatory elements, such as enhancers and their target promoters, are mostly located within the same TAD to ensure the specificity of gene expression by limiting undesired inter-TAD contacts. TADs have clear

boundaries, which are enriched with multiple features such as enrichment of highly transcribed genes, H3K36me3 and H3K4me1 histone marks, TSSs, housekeeping and tRNA genes, CTCF (CCCTC-binding factor) binding sites and some other factors (Dixon *et al.*, 2012). It is highly likely that CTCF and transcription are both involved in TAD formation (van Steensel and Furlong, 2019). Currently, there are two hypothetical models of TAD formation:

1. The loop extrusion model suggests that TADs emerge as a result of pushing chromatin fiber loop through Cohesin containing protein complex that resembles an eye of a needle (Sanborn *et al.*, 2015; Fudenberg *et al.*, 2016; Vian *et al.*, 2018).
2. The multivalent transcription factor model - highly similar to the loop extrusion model, but the role of the eye of a needle is taken by TFs that can interact with each other (Brackley *et al.*, 2016).

In both cases, it is not clear:

- i. which molecular mechanisms are pushing the DNA through
- ii. is loop formation an active process and whether it requires additional motors besides transcribing RNAP II.

Active RNAP II concentration is not uniform in the cell nucleus (Iborra *et al.*, 1996), which influences the nuclear spatial organization of genes (Sutherland and Bickmore, 2009). Recent studies suggest that transcription elongation itself stimulates loop formation (Cook and Marenduzzo, 2018; Rowley and Corces, 2018).

If TAD boundaries are disrupted, aberrant interactions between different domains lead to inappropriate enhancer-promoter pairs and, as a consequence, ectopic gene expression (Lupiáñez *et al.*, 2015; Hnisz *et al.*, 2016). Modular organization of chromosomes in the cell nuclei spatially organizes thousands of genes and regulatory regions, providing a framework to restrict an action of regulatory elements within the correct range.

1.3 Enhancer chromatin

Different cell-types acquire specific profiles of chromatin accessibility and histone PTMs configuration, which define regulatory regions (Thurman *et al.*, 2012; Stergachis *et al.*, 2013).

Two main hallmarks define active enhancers: the production of non-coding RNA transcripts known as enhancer RNAs (Kim *et al.*, 2010; Ørom *et al.*, 2010; Hah *et al.*, 2011; Wang *et al.*, 2011) and elevated Histone 3 lysine 27 acetylation (H3K27ac) levels (Creyghton *et al.*, 2010). Upon activation, enhancers physically interact with target promoters. It has been reported that eRNAs can promote looping between enhancers and promoters through recruitment of cohesin (Li *et al.*, 2013; Hsieh *et al.*, 2014), increase chromatin accessibility (Mousavi *et al.*, 2013), and recruit the mediator complex to promoters (Lai *et al.*, 2013). At poised genes, eRNAs can promote the release of paused RNAP II into the gene body (Schaukowitch *et al.*, 2014). Consistently, depletion of eRNAs inhibits transcription from enhancer-associated genes (Kim *et al.*, 2010; Lai *et al.*, 2013; Li *et al.*, 2013; Melo *et al.*, 2013; Mousavi *et al.*, 2013).

Enhancer-promoter interactions mostly occur within confined space of the genome since higher-order chromatin structure in the cell is quite static (Jin *et al.*, 2013), and do not change significantly during development (Ghavi-Helm *et al.*, 2014).

Various combinations of chromatin marks have been associated with different states of enhancer activity (Table 2).

Table 2. Enhancer associated chromatin marks

Enhancer state	Associated chromatin mark
Repressed/poised	H3K4me1/H3K27me3 (Rada-Iglesias <i>et al.</i> , 2011; Bonn <i>et al.</i> , 2012)
Primed	H3K4me1/no acetylation (Zentner, Tesar and Scacheri, 2011)
Active	H3K4me1/H3K27ac (Rada-Iglesias <i>et al.</i> , 2011; Zentner, Tesar and Scacheri, 2011)

However, these marks are not exclusive to enhancers (e.g., H3K4me1 can also be found at insulators and H3K27ac at promoters), and not all of the enhancers can be identified based on these marks. For example, a distinct class of active enhancers in *Drosophila* and mice have no H3K27ac (Bonn *et al.*, 2012; Pradeepa, 2017).

Although highly correlated with enhancer activity it is not, currently, clear whether these marks and chromatin properties are functionally required for enhancer activity or are just the consequence of it.

It is well established that active enhancers are depleted of nucleosomes; however, not all NFRs are enhancers (Thurman *et al.*, 2012). Also, while located at NFR and being active in one cell type, the same enhancer may be inactive in other cell types, still residing in the same NFR (Zentner, Tesar and Scacheri, 2011; Andersson, Gebhard, *et al.*, 2014).

Both active promoters and enhancers are flanked by nucleosomes enriched with H3K4me1 and H3K4me3 marks. The H3K4me1/me3 ratio was found to be higher at enhancers than at promoters. This observation was later refined, suggesting that despite H3K4me1 being observed at many enhancers, it does not distinguish between active and inactive enhancers (Bonn *et al.*, 2012; Andersson, 2015). Furthermore, it has been demonstrated that replacement of Mll3/Mll4 (proteins that deposit the H3K4me1 mark) with their catalytically dead mutants, is dispensable, whereas their presence is required for enhancer activity (Dorigi *et al.*, 2017).

Enhancers regulate the activity of their target promoters by attracting co-activator proteins such as acetyltransferases, CREB binding protein (CBP), and P300. These proteins are typically recruited by transcription factors, but also contain a Bromodomain, which enables them to bind acetylated nucleosomes on their own. CBP/p300 has multiple acetylation targets such as histones, transcription factors, and RNAP II (Schröder *et al.*, 2013). These proteins generally lead to transcription activation via the emergence of highly acetylated nucleosome regions (Roh, Cuddapah and Zhao, 2005; Roh *et al.*, 2007). CBP-mediated acetylation facilitates TF recruitment and promoter escape of RNAP II (Stasevich *et al.*, 2014). Enrichment of p300 at chromatin sites has been used as a signature feature of enhancers and resulted in the identification of highly cell-type-specific enhancers (Heintzman *et al.*,

2007, 2009). Targeting the p300-HAT domain alone to enhancers utilizing deactivated Cas9 nuclease (dCas9) and enhancer-specific guide RNAs (gRNAs) increases the expression of enhancer target genes (Hilton *et al.*, 2015). Furthermore, CBP has been shown to interact with eRNA, which, in turn, activates acetyltransferase activity (Bose *et al.*, 2017). This may constitute a positive feedback loop at active enhancers. One of the major CBP/p300 histone targets is acetylation of H3K27, which is known to be enriched at active enhancers (Tie *et al.*, 2009; Creighton *et al.*, 2010; F. Jin *et al.*, 2011). Interestingly, the treatment of p300 with a selective Bromodomain inhibitor led to the abolishment of H3K27ac deposition preferentially at enhancers and, consequently, reduction of eRNA transcription and enhancer activity. Unlike the H3K4me mark, impairing processes that maintain H3K27ac levels lead to inappropriate gene expression (Raisner *et al.*, 2018). However, the molecular mechanisms underlying this phenomenon are largely unknown.

1.4 Post-translational modification of H3K27 is a switch between active and silent chromatin

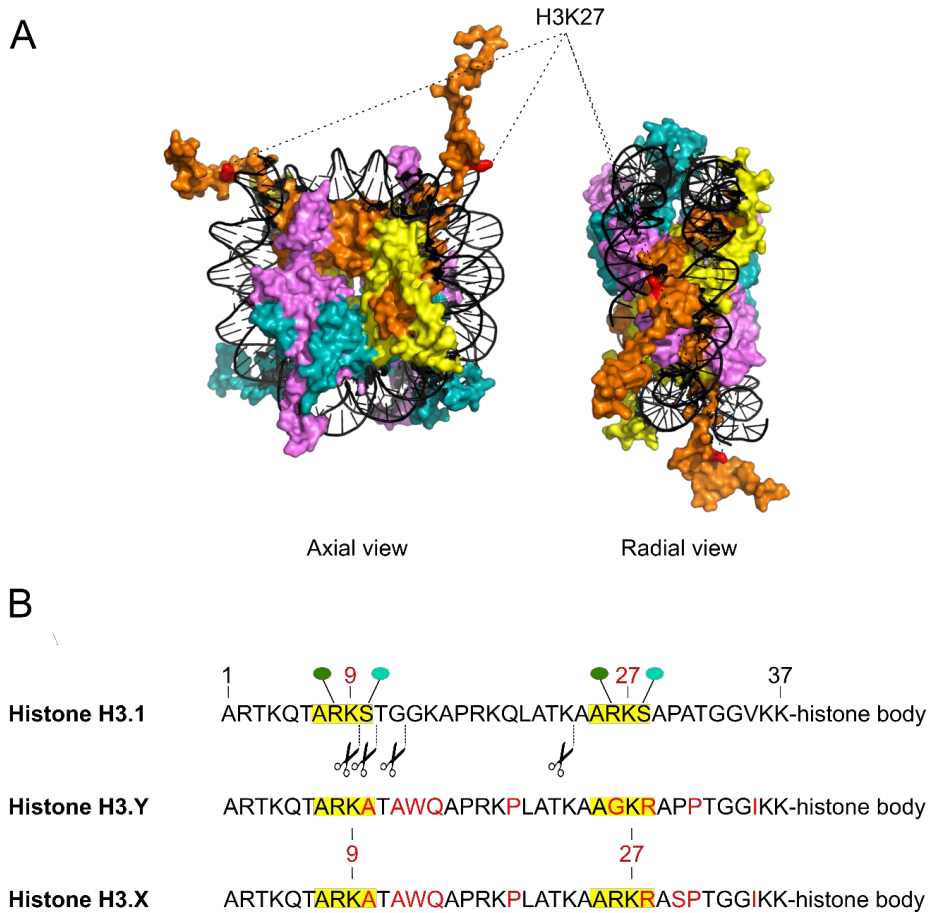


Figure 8. H3K27 position in a nucleosome (PDB code 1KX5) and the -ARKS- motif. A. The position of the H3K27 in a nucleosome. H3K27 labeled with red. B. N-terminal amino acid sequences of histone H3.1, H3.Y, and H3.X N-terminal tail. -ARKS- motifs are labeled with yellow. Scissors represent histone clipping sites. Different methylation states of H3R8 and H3R26 are labeled with green oval and phosphorylation of H3S10 and H3S28 with the blue oval. Residues that differ from canonical H3.1 are labeled with red.

H3K27 is one of the closest lysine residues to the DNA in the nucleosome (Fig. 8A). Methylation and acetylation of this residue has been found to have opposing effects on chromatin states (Wiles and Selker, 2017).

Acetylated H3K27 is found in open, transcriptionally active chromatin, and is one of the targeted acetylation sites specifically enriched at active enhancers. It is deposited by the CBP/p300, which is recruited by different transcription factors. Acetylation of the H3K27 neutralizes lysine's positive charge, which is disruptive to the DNA-histone tail interaction. Since the histone H3 protein resides at the DNA entry point of the nucleosome, H3K27ac may also affect nucleosome stability, lower the transcription activation barrier, facilitate chromatin remodeling, and serve as an important binding site for chromatin remodeling and transcription coactivator proteins that contain acetyl-lysine recognition modules. It is still not clear what specific “reader” proteins recognize this mark and what regulatory mechanisms are triggered to influence enhancer function.

Methylation of H3K27 is strongly associated with inactive, silent chromatin and introduced by the methyltransferase Enhancer of Zeste Homolog 2 (EZH2), a subunit of the polycomb⁵ repressive complex 2 (PRC2). H3K27me₃ is a binding site for various heterochromatin-associated proteins, such as PRC1/2.

It has been argued that H3K27ac is simply an H3K27me₃ antagonist and does not have a function on its own (Pengelly *et al.*, 2013). It is difficult to directly address the role of H3K27 modifications by mutation of that site, since, in higher eukaryotes, the histone H3 gene has multiple copies and variants. Even the brilliant approach taken by Pengelly *et al.*, who replaced all wild-type histones copies with mutant variants (H3.3K27R, H3.1K27R) in *Drosophila*, could not avoid the presence of heterogeneous nucleosomes containing both mutated and wild-type histone proteins,

⁵ Polycomb-group proteins (PcG) are a family of transcriptional repressors, first described in *Drosophila* (Slifer, 1942; Lewis, 1947), and named after adult males that had extra sex combs on the second and third pairs of legs, whereas wild-type only on the first pair. Later, after other mutants with developmental abnormalities had been discovered, it was proposed that impaired genes were global repressors of the homeotic (Hox) genes (Gellon and McGinnis, 1998), that drive development and differentiation (Lewis, 1978). This hypothesis initiated the search for positive regulators of Hox genes, which were found shortly after and now known as the Trithorax group (TrxG) genes (Ingham and Whittle, 1980; Capdevila and Garcia-Bellido, 1981; Forquignon, 1981; Ingham, 1981). The PcG and TrxG play a repressive and activating role in cellular memory, respectively. The TrxG contains multiple transcription activators such as SWI/SNF and Mediator subunits that maintain the ‘on’ state of the chromatin as opposed to PRC1/2 that maintains the ‘off’ state of the chromatin (Kingston and Tamkun, 2014).

due to maternally supplied wild-type histone proteins. Yet, this study demonstrated that flies with the K27R mutation die during the developmental stage because of impaired heterochromatin formation and PRC2 mediated silencing.

Another mutation of that residue, H3K27M is observed in >70% of pediatric gliomas with dominant PRC2 loss of function, and was reported to account for the cancerous phenotype formation and developmental perturbations (Schwartzentruber *et al.*, 2012; Wu *et al.*, 2012; Bender *et al.*, 2013; Funato *et al.*, 2014; Herz *et al.*, 2014). Unfortunately, none of these studies, including the latest study that applied single-cell transcriptomic analysis of H3K27M glioma cells (Filbin *et al.*, 2018), addressed the effects of this mutation on enhancer function. Notably, H3K27M in diffuse intrinsic pontine glioma cells is associated with increased levels of H3K27ac and the majority of the heterotypic H3K27M-K27ac nucleosomes were enriched with Bromodomain (BRD2/4) proteins at the actively transcribed loci. Moreover, the authors show that the application of BET family Bromodomain inhibitors efficiently inhibits tumor progression, which opposes previous conclusions about the dominant role of impaired heterochromatin function in cancer development (Piunti *et al.* 2017). It has also been shown that PRC2 is still able to recognize and bind H3K27M containing heterotypic nucleosomes and does not change the amount of PRC2 bound to chromatin (Tatavosian *et al.*, 2018).

Overall, this data implies that H3K27 fulfills an important function, regulating both silencing and activation of transcription by employing mechanisms that antagonize each other, and hence creating a switch between active and silent chromatin.

The ARKS repeated motif

H3K27 resides in a motif of four amino acids ARKS that is found twice on the histone H3 tail: around K9 (N-terminal repeat) and K27 (C-terminal repeat) (Fig. 8B).

Multiple proteolytic cleavage sites surround the N-terminal ARKS motif. When clipped at these sites, the histone H3 N-terminal tail is shortened, reducing the number of important docking sites for histone binding proteins (Tvardovskiy *et al.*, 2015). In that case, the C-terminal ARKS tail motif may still provide a binding site for essential interacting proteins. Indeed, PTMs of both ARKS motifs are correlated

with somewhat similar context and often have similar “readers” and “writers” (Table 3).

Table 3. Modifications of the ARKS motif.

Histone Mark	Associated chromatin	Regulatory regions	PTM writers	PTM readers	Crosstalk
<i>N-terminal ARKS repeat</i>					
H3R8me	Unclear	Repressed promoter regions ¹⁷	PRMT5 ¹⁶		Prevented by any modification of H3K9 ¹⁶
H3K9me2/3	Heterochromatin (constitutive) ¹	Bulk heterochromatin	SUV39H ³	HP1 ³ , PRC1*	Mutually exclusive with H3K27me3 and H3K9ac ⁷
H3K9ac	Euchromatin	Active promoters ^{6,8,10}	GCN5/PCAF ¹⁵	SEC (YEATS) ⁹	Promoted by H310/28ph ¹¹
H3S10ph	Heterochromatin (pericentromeric) ¹ 2	Regulatory regions of inducible genes ^{11,13}	Aurora-B ¹¹	14-3-3 ¹¹	Reduces H3K9me3/H3K27me3 ¹¹
<i>C-terminal ARKS repeat</i>					
H3R26me	Unclear	Unclear	CARM1 ¹⁷		Prevented by any modification of H3K27 ¹⁷
H3K27me2/3	Heterochromatin (facultative) ¹	Poised enhancers/promotes (in combination with H3K4me1/3) ^{6,7}	EZH2 ⁵ (PRC2)	PRC1 ⁴ EED (PRC2) ⁵	Mutually exclusive with H3K9me3 and H3K27ac ⁷
H3K27ac	Euchromatin	Active enhancers ^{6,8}	CBP/p300 ¹⁴	SEC (YEATS) ⁹	Promoted by H3S10/S28ph ¹¹
H3S28ph	Heterochromatin (pericentromeric) ¹	Regulatory regions of inducible genes ^{11,13}	Aurora-B ¹¹	14-3-3 ¹¹	Reduces H3K9me3/H3K27me3 ¹¹

Source: 1-Saksouk, Simboeck and Déjardin, 2015; 2-Campos and Reinberg, 2009; 3-Lachner *et al.*, 2001; 4-Fischle *et al.*, 2003; 5-Cao *et al.*, 2002; 6-Rada-Iglesias *et al.*, 2011; 7-Bonn *et al.*, 2012; 7-Zhang, Cooper and Brockdorff, 2015; 8-Zentner, Tesar and Scacheri, 2011; 9-Li *et al.*, 2014; 10-Karmodiya *et al.*, 2012; 11-Sawicka and Seiser, 2012; 12-Hendzel *et al.*, 1997; 13-Sawicka *et al.*, 2014; 14-Tie *et al.*, 2009; 15-Q. Jin *et al.*, 2011; 16-Pal *et al.*, 2004; 17-Bedford and Clarke, 2009; *(Bernstein *et al.*, 2006; Vermeulen *et al.*, 2010; Kaustov *et al.*, 2011; Eberl *et al.*, 2013; Tardat *et al.*, 2015)

Three out of four amino acids originating from -ARKS- repeat may be subject to PTM. High density and diversity of possible PTMs at these repeats provide a framework for the “histone code” to modulate the activity of effector proteins and, consequently, biological outcome. Interestingly, primate-specific variants of H3 (H3.Y and H3.X) bear variations in ARKS repeat residues (Fig. 8B) (Filipescu, Müller and Almozni, 2014), which could affect binding specificity of PTM readers

or histone tail clipping. One of the examples of the crosstalk between PTMs of the ARKS repeats is the so-called “methyl-phospho switch” (Sawicka and Seiser, 2012). In brief, phosphorylation of H3S10 and H3S28 abrogates binding to H3K9me3 and H3K27me3 respectively, which leads to the de-repression of transcriptionally silent chromatin by promoting acetylation of H3K9 and H3K27 (Fischle, Wang and Allis, 2003; Fischle *et al.*, 2005; Gehani *et al.*, 2010; Lau and Cheung, 2011a, 2011b).

Despite sharing many common features between PTMs of ARKS repeat residues, H3K9 and H3K27 are deposited by different complexes and are associated with different types of chromatin and regulatory regions (Table 3). H3K9me3 is a typical mark of constitutive⁶ heterochromatin, whereas H3K27me3 marks facultative heterochromatin generated by PRC2 (Saksouk, Simboeck and Déjardin, 2015). Acetylation of H3K9 is associated with active promoters (Karmodiya *et al.*, 2012), whereas H3K27ac is enriched at active enhancers (Creyghton *et al.*, 2010). Despite this observed specificity, neither of these acetylation marks are exclusive and can be enriched at both types of regulatory regions (Karmodiya *et al.*, 2012).

Due to apparent similarities between promoters and enhancers (Andersson, 2015; Andersson and Sandelin, 2020), it comes as no surprise that promoters and enhancers have similar activation requirements and protein interacting environments. However, impairment of p300 acetylation function and subsequent reduction in levels of H3K27ac was observed to affect only enhancer, but not promoter function (Raisner *et al.*, 2018). This fact underscores that despite similarities between promoters and enhancers, the cell has mechanisms that differentiate and conduct regulatory region-specific functions. On this basis, it is essential to find out how H3K27ac influences enhancer chromatin and function.

⁶ Constitutive and facultative heterochromatin are the two best-studied heterochromatin types. Constitutive chromatin is gene-poor, consists of different satellite repeats and transposon repeats that need to be silenced throughout the cell cycle. Facultative heterochromatin is gene-rich and consists of silenced genes that were active or yet to be expressed.

2. Aim of the study

So far, H3K27ac is the PTM, which is most strongly correlated with active enhancers. We hypothesize that the enhancer function of the H3K27ac may arise from its interaction with specific reader proteins. Selective inhibition or knock out/down of such a “reader” could lead us to identify the role of the H3K27ac.

The overall aim of the current study has, therefore, been to explore the underlying mechanisms of the H3K27ac histone mark in enhancer function.

A genetic approach for identifying the function of this mark is unattainable since mutation of H3K27 would perturb the action of H3K27me3 in transcription silencing and confounded by the presence of heterotypic nucleosomes that contain both mutant and wild-type H3.

To identify candidate H3K27ac-reader proteins, we decided to apply a histone-peptide pulldown approach based on quantitative proteomics (Vermeulen, 2012). Depending on the nature of the identified candidate proteins, their authenticity as H3K27ac-binding should be validated and further characterized by chromatin profiling, biochemical, and functional studies.

3. Summary of the results

The cell is a complex system, and enhancers are vital mediators of cell-type and context-specific gene expression patterns. Although H3K27ac is associated with active enhancers, the role of this histone modification in enhancer function and whether it differentiates enhancer elements from promoters is still not clear.

In the current study, we aimed to identify specific reader(s) of H3K27ac that may account for specificity towards enhancer function, and could further help us understand how this PTM influences gene expression.

3.1 The H3K27ac interactome

We investigated the interactome of the H3K27ac histone mark via **Stable Isotope Labeling by Amino acids in Cell culture (SILAC)** coupled with **Histone Peptide PullDown (HPPD)** quantitative mass spectrometry (MS). HPPD-MS has been successfully applied in multiple studies to identify specific readers of methylation marks (Vermeulen *et al.*, 2007, 2010) as well as several acetylated marks of histone H3 (Vicent *et al.*, 2009).

As ‘bait,’ we used biotinylated, synthetic histone-peptides derived from histone H3 (residues 15-36) acetylated at K23, K27, or both. As the control ‘bait,’ we used an unmodified histone peptide with the same amino acid sequence. For the HPPD-MS, we used HeLa and mESC cells as the source of nuclear extracts. Human cancer cells (HeLa) and mouse ESCs, represent aberrant and normal cell background, respectively. Binding was measured as SILAC ratios between peptide with the PTM of interest and the unmodified peptide.

Due to the known specificity of Bromodomains for acetylated lysines, we expected to see Bromodomain containing proteins and their associated proteins. Prior to our study, the primary candidate for the specific recognition of the H3K27ac was the BRD4 protein that contains a tandem pair of bromodomains (BD1 and BD2) (Dey *et al.*, 2000). BRD4 has previously been reported to be enriched at super-enhancers (Lovén *et al.*, 2013). BRD4 interacts with active CDK9 (forming so-called BRD4 containing Elongation Complex – BEC; Bacon and D’Orso, 2019), which promotes

RNAP II pause release into productive elongation through subsequent RNAP II CTD Ser2 phosphorylation (Moon *et al.*, 2005).

Despite being associated with the H3K27ac mark, we could not detect BRD4 protein in any of our HPPD-MS experiments. The majority of the identified interactome of H3 acetyl-lysine consisted of proteins that originate from two chromatin-interacting complexes, The Super Elongation Complex (SEC) (Luo, Lin and Shilatifard, 2012) and the SWI/SNF family complexes (PBAF and GBAF). The SEC components ENL, AFF1/4, and AF9, were identified as H3K27ac interactors in HeLa cells and additionally CDK9 in mESC with diacetylated histone peptide as bait. All known subunits of PBAF (in HeLa were identified with every histone-peptide tested; in mESC with H3K23ac and H3K23K27ac histone peptides). The complete GBAF complex was identified in HeLa with H3K27ac and H3K23K27ac; in mESC with every histone-peptide tested. Additionally, we detected NuA4 complex components DMAP1 and YEATS4 (Doyon *et al.*, 2004) and several subunits of the general transcription factor TFIID (Dymlacht, Hoey and Tjian, 1991). All these complexes contain known acetyl-lysine recognition domains, such as YEATS and Bromodomains.

Unexpectedly, GLTSCR1L, a poorly characterized protein, but not its paralog GLTSCR1, showed a clear binding preference towards H3K27ac containing peptides. The SEC components and the GLTSCR1L showed the highest SILAC ratios and were assigned as the strongest interactors of H3K27ac (Manuscript, Fig. 1b-c; Supplementary Fig. 1b-c). Moreover, the PBAF complex subunits showed a clear preference for the H3K23ac mark (Manuscript, Fig. 1a; Supplementary Fig. 1a).

Albeit we did not detect BRD4 as H3K27ac interacting protein, we identified the SEC that utilizes a similar CDK9-mediated mechanism for RNAP II CTD Ser2 phosphorylation (Moon *et al.*, 2005) and subsequent RNAP II pause release into productive elongation. However, SEC can interact with both H3K9ac and H3K27ac histone marks (Li *et al.*, 2014; Erb *et al.*, 2017). Hence both enhancers and promoters might use a common SEC-mediated RNAP II pause release mechanism. In the further work, we focused on GLTSCR1L and GBAF since GLTSCR1L protein has shown clear H3K27ac interaction preference in HPPD-MS (Manuscript, Fig. 1d;

Supplementary Fig. 1d), thus potentially could account for enhancer-specific activities.

3.2 GBAF complex composition

At the time our initial HPPD-MS experiments were conducted, the GBAF complex was unknown and we had to find out which complex GLTSCR1L belongs to. To test this, we developed transgenic HeLa-FRT cell lines with inducible expression of GFP-tagged GLTSCR1L protein, and used them in a series of immunoprecipitation experiments combined with quantitative mass spectrometry-based proteomics (IP-MS) experiments.

The observed interaction partners of full-length GLTSCR1L protein were SWI/SNF core proteins SMARCC1, SMARCD1, Bromodomain containing BRD9 protein, BCL7C, SS18 subunits, and ATPase SMARCA4 (Manuscript, Fig. 2a).

To validate identified interacting partners we developed similar transgenic HeLa FRT cell lines expressing GFP-tagged BRD9, SMARCD1 and SMARCC1 and performed reciprocal IP-MS experiments to corroborate that GLTSCR1L is a bona fide subunit of a new SWI/SNF complex named GBAF (Manuscript, Fig. 2b-c; Supplementary Fig. 3a). IP-MS also allows for relative abundance quantification of analyzed proteins. The measured abundance of GLTSCR1 or GLTSCR1L immunoprecipitated with SMARCD1 and SMARCC1 was below 2% (Manuscript, Fig. 2f; Supplementary Fig. 3f).

Taken together, conducted IP-MS experiments validate GLTSCR1L as a component of a novel SWI/SNF family complex, which we term GBAF. According to our data, GBAF is less abundant as compared to canonical BAF or PBAF complexes.

GLTSCR1L-mediated interactions and the GLTSCR domain

Next, we isolated and expressed, in a similar fashion as described above, the GLTSCR1 (GiBAF) domain, the most conserved region of the GLTSCR1L (Manuscript, Supplementary Fig. 2). We performed IP-MS experiments and found

that this domain precipitated all GBAF subunits and, additionally, BRD4 and AFF1 (Manuscript, Fig. 2d). Further investigation by immunofluorescence imaging showed that it is also responsible for the nuclear localization of the GLTSCR1L protein (Manuscript, Supplementary Fig. 3b).

IP-MS experiments with the GiBAF domain deletion mutant resulted in no significant SWI/SNF interactions. However, we detected the GiBAF domain deletion mutant to interact with TAF6 and TAF4, components of the TFIID complex were observed (Manuscript, Supplementary Fig. 3c).

Taken together, our data suggest that the GiBAF domain interacts with the GBAF complex subunits and additionally able to interact with BRD4 and AFF1 proteins (components of BEC/SEC). The deletion mutant of the GiBAF domain was found to interact only with TFIID complex components.

3.3 GBAF localizes preferentially at intergenic/intronic regions enriched with H3K27ac

We performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) to verify if the GBAF complex components (GLTSCR1L and BRD9) co-localize with H3K27ac in chromatin. We applied multiple bioinformatics approaches to validate their genomic colocalization with H3K27ac, other chromatin marks, and factors associated with enhancer activity in chromatin, and to annotate their genomic locations.

We observed a high degree of overlap between GLTSCR1L, BRD9, H3K27ac, chromatin marks, and factors associated with active chromatin (H3K4me1/me2/me3, P300, RNAP II, H3K9ac) (Manuscript, Fig. 3b, c). About 75% of the detected GLTSCR1L and BRD9 peaks reside within intronic/intergenic regions (Manuscript, Fig. 3d), which are known localization places of a large proportion of enhancer elements (Gerstein *et al.*, 2007; Khalil *et al.*, 2009; Chorev and Carmel, 2012). We also performed an analysis of enrichment of known transcription factor motifs and observed enrichment analysis of multiple transcription factor motifs such as JunB and CTCF (Appendix A-B).

Taken together, our data suggest that the GLTSCR1L-containing GBAF complex is enriched at enhancer regions in chromatin and sites that overlap with the H3K27ac histone mark.

3.4 Effects of H3K27ac recognition inhibition by GBAF on histone peptide binding specificity and chromatin localization

To probe the role of H3K27 acetyl-lysine binding by GBAF complex in chromatin, we impaired this function with a selective inhibitor of BRD9 Bromodomain (I-BRD9) (Theodoulou *et al.*, 2016) and performed HPPD followed by immuno-blotting, and a series of ChIP-seq experiments.

We observed a pronounced effect of I-BRD9 *in vitro*: GLTSCR1L interaction with acetylated H3 histone peptides was almost abrogated, whereas BRD9 lost specificity (the BRD9 signal was the same for all acetylated histone peptides tested) of interaction with histone peptides in HPPD (Manuscript, Fig. 3a). Treatment of HeLa cells with I-BRD9 *in vivo* induced changes in genomic occupancy of GLTSCR1L and BRD9. Interestingly, localization of BRD9 was only mildly affected, as compared to an almost complete abrogation of GLTSCR1L binding to chromatin. Following I-BRD9 treatment, we also observed that remaining GLTSCR1L peaks overlapped mostly with promoter-associated histone marks (H3Kme3, H3K9ac) and TSS regions (Manuscript, Fig. 3c, d). Thus inhibition of GBAF H3K27ac recognition by I-BRD9 leads to a loss of BRD9's specificity towards H3K27ac, dissociation of GLTSCR1L from chromatin, and most likely, disruption of GBAF complex assembly.

3.5 Effects of GBAF H3K27ac recognition inhibition on transcription

Since the SWI/SNF family of chromatin remodelers are implicated in transcription regulation and we observed enrichment of the GBAF complex at enhancer-associated genomic regions marked with H3K27ac, and dissociation of the GLTSCR1L from chromatin upon I-BRD9 treatment, we decided to check how I-

BRD9 treatment affects transcription. Enhancer transcription is correlated with enhancer regulatory activity and together with H3K27ac, is the best predictor of active enhancers (Melgar, Collins and Sethupathy, 2011; Wang *et al.*, 2011; Andersson, Refsing Andersen, *et al.*, 2014; Henriques *et al.*, 2018; Mikhaylichenko *et al.*, 2018; Rennie *et al.*, 2018).

First, we measured expression levels of the human epidermal growth factor (hEGF) inducible NR4A1 gene and its known enhancer (~80 kb downstream of the NR4A1 gene locus) in HeLa cells by RT-qPCR. NR4A1 and its enhancer are immediate early genes (IEG) that can be induced by hEGF (Lai *et al.* 2015). Measurements were taken before and at 30 minutes after treatment of the cells with hEGF in the presence or absence of I-BRD9 (applied for 6 hours before hEGF induction). We also performed ChIP-qPCR to validate that GLTSCR1L and BRD9 bind the NR4A1 TSS and enhancer, and to monitor how the induction of HeLa cells with hEGF affects this interaction in the presence or absence of I-BDR9.

ChIP-qPCR indicated that BRD9 is enriched at the NR4A1 enhancer compared to the TSS region. Upon hEGF induction, we detected a 2-fold increase in the BRD9 signal at the enhancer and, to a lesser extent, at the TSS region. GLTSCR1L was weakly associated with both NR4A1 TSS and enhancer, but upon hEGF induction, we observed an almost 10-fold increase of GLTSCR1L signal at NR4A1 enhancer and to a lesser extent at TSS region. Upon I-BRD9 treatment, we observed abrogation of GLTSCR1L binding to chromatin in both, hEGF induced and non-induced HeLa cells (Manuscript, Supplementary Fig. 5a). Thus, GLTSCR1L and BRD9 were recruited to the induced NR4A1 enhancer element in BRD9 bromodomain dependent fashion.

Furthermore, we observed that, upon I-BRD9 treatment and hEGF induction, the expression levels of NR4A1 promoter were slightly increased, while enhancer transcription was strongly downregulated as compared to non-treated with I-BRD9 cells (Manuscript, Supplementary Fig. 5b). Thus, impairment of GBAF's ability to bind to H3K27ac inhibits transcription of the NR4A1 enhancer, but not the promoter.

To assess the global effects on transcription of the GBAF and H3K27ac interaction inhibition by I-BRD9, we performed a series of Cap Analysis of Gene

Expression (CAGE) experiments, and in parallel, Assay for Transposase-Accessible Chromatin-sequencing (ATAC-seq) to focus on true transcription initiation events at open chromatin loci. With a combination of methods for the detection of histone marks, chromatin accessibility, and genome-wide assays that measure RNA levels, the activity of enhancer elements can be measured quite accurately. The measurements for the CAGE and ATAC-seq were taken at 0, 30, 60, and 240 minutes after hEGF treatment.

Consistent with the RT-qPCR results, we observed little or no effect on the NR4A1 gene and downregulation of its corresponding enhancer transcription upon I-BRD9 treatment. Compared aggregated fold change of NFR (nucleosome free region) expression across all hEGF time points revealed selective transcription downregulation of intronic/intergenic regions, whereas transcription from other genomic loci was not affected by I-BRD9 treatment (Manuscript, Fig. 4b). Furthermore, we compared NFR transcriptional events measured at each time point separately in control cells, and in I-BRD9 treated cells. We observed differential expression of both TSS and intronic/intergenic regions. However, a reduced fold change of expression between I-BRD9 treated and untreated cells was observed only at intronic/intergenic regions (Manuscript, Fig. 4c). Thus, inhibition of GBAF binding to H3K27ac by I-BRD9 leads to a preferential downregulation of enhancer transcription throughout time course after treatment of HeLa cells with hEGF. Although upon I-BRD9 treatment, GLTSCR1L disassociated from all genomic regions tested, we observed the downregulation of transcription exclusively at intronic/intergenic regions that lost GLTSCR1L, but no downregulation at those regions that retained GLTSCR1L (Manuscript, Fig. 4d). The ATAC-seq also revealed minor, but significant decrease in chromatin accessibility⁷ at GLTSCR1L-depleted enhancers (Manuscript, Supplementary Fig. 5e). Due to the observed differential gene expression after I-BRD9 treatment, using publically available Hi-C data derived from HeLa cells, we asked whether downregulated genes reside within the same TADs as downregulated enhancers or not. Remarkably, TADs that contain at least one

⁷ However, it has to be noted that in order to be confident in this result, ATAC-seq experiments should be performed at least as triplicates, which is not the case with the current data.

downregulated enhancer showed significant downregulation of promoter transcription as compared to TADs with no downregulated enhancers (Manuscript, Fig. 4e).

Taken together, HeLa cells treated with I-BRD9 showed selective downregulation of transcription at enhancers, which coincided with the dislocation of GLTSCR1L and reduced chromatin accessibility.

4. Discussion

In the work described above we have identified two protein complexes, SEC and GBAF, that both specifically recognize H3 histone peptides with the H3K27ac mark, the hallmark of active enhancers. It is likely that SEC facilitates eRNA transcription, while GBAF remodels nucleosomes in enhancer chromatin. Below I discuss the approach taken and the implications of our findings in the light of recent publications in the field.

4.1 Identifying H3K27ac interacting proteins

4.1.1 Histone-peptide pulldown MS vs ChIP-MS

So far, several studies have addressed the identification of H3K27ac-interacting proteins using chromatin immunoprecipitation, followed by mass-spectrometry (ChIP-MS) (Engelen *et al.*, 2015; Ji *et al.*, 2015). These studies have detected numerous transcription factors, chromatin remodelers, RNAP II, and other chromatin interacting proteins to be enriched at H3K27ac-marked genomic regions. As for ChIP-seq and ChIP-qPCR, ChIP-MS relies on the purification of crosslinked, sheared chromatin with pre-existing protein complexes that are pulled down with an antibody raised against a histone PTM or a chromatin-bound protein. None of these methods allows for identification of proteins that specifically bind the histone mark of interest. There are two main reasons for this: 1. the antibody recognizes the PTM, which is supposed to be bound by the interacting proteins, and 2. the crosslinked chromatin is left with bound sequence-specific TFs which can attract and lead to co-precipitation of different remodelers and other chromatin interacting proteins. Thus we cannot discern the contribution of H3K27ac mark alone in a DNA-independent context. To overcome these problems, we applied a histone peptide pulldown MS approach that does not rely on crosslinking chromatin and the use of the antibodies to enrich for the target chromatin regions bearing PTM of interest. Histone peptide pulldown has been successfully applied in multiple studies to identify specific readers of histone methylation marks (Vermeulen *et al.*, 2007, 2010) as well as several acetylation marks of histone H3 (Vicent *et al.*, 2009). The next step in this approach is the

demonstration of colocalization of candidate proteins with the PTM of interest by ChIP-seq.

4.1.2 H3K27ac is recognized by SEC and GBAF

Unexpectedly, we did not observe BRD4 in the interactome of any investigated acetyl-lysine marks. Instead, we identified several SEC complex subunits. The canonical SEC contains AF4/FMR2 family members 1 or 4 (AFF1/4), eleven-nineteen Lys-rich leukemia (ELL) proteins ELL1/2/3, the YEATS domain-containing MLLT3 (AF9) or ENL (MLLT1) proteins, and associates with a pTEF-b module (CDK9, CyclinT1/2) (Luo, Lin and Shilatifard, 2012). The SEC subunits detected in HPPD-MS indicate that fully active SEC (only CyclinT1/T2/K were undetected) can be attracted by H3K27ac containing nucleosomes. The YEATS domain of AF9 has previously been demonstrated to bind H3K9ac, H3K18ac, and H3K27ac⁸ (Li *et al.*, 2014). Also, the interaction between SEC and H3K27ac has been described in the context of a negative feedback regulatory loop with H3R26me (Zhang *et al.*, 2017). Therefore, one of the possible roles of the H3K27ac is to facilitate SEC mediated RNAP II elongation pause release, at both promoters and enhancers, a feature thus universal for both regulatory regions.

Another complex that indicated specificity towards H3K27ac containing peptides in our HPPD-MS assay is the GBAF complex, which was only recently characterized (Alpsy and Dykhuizen, 2018). GBAF belongs to the SWI/SNF family of chromatin remodelers that have a diverse range of functions related to DNA-dependent processes in the cell nucleus. The GBAF is a non-canonical BAF complex defined by the presence of BRD9 and one of the paralogues GLTSCR1 or GLTSCR1L proteins. Interestingly GBAF, unlike canonical SWI/SNF complexes, does not contain any known DNA-interaction modules (except in SMARCA2/4), which is likely to make it more dependent on binding to acetylated lysines for interaction with chromatin or other target proteins. Several studies have identified numerous SWI/SNF proteins associated with H3K27ac (Engelen *et al.*, 2015; Ji *et al.*, 2015; Zhang *et al.*, 2017), but none of them have detected GLTSCR1L. Interestingly,

⁸ However, the interaction with H9K9ac was significantly favored over H3K27ac.

only GLTSCR1L, but not GLTSCR1 protein, was observed to interact with H3K27ac in our study, which leads us to conclude that the GBAF complex might have different functions or targets depending on which of the GLTSCR1 paralogs reside in the complex.

4.1.3 Specificity of H3K27ac recognition

Since our screen had been performed in two different cell lines, we were able to observe a possible cell-type-specific effect of H3K27ac recognition. In HeLa cells, the SEC was the main interactor of the H3K27ac mark alone, and the GBAF interaction was more profound with di-acetylated H3K23acK27ac histone peptide. In mESC, we observed the opposite, as the histone peptide with H3K27ac alone was interacting with GBAF, whereas the strongest interactor for diacetylated H3K23acK27ac histone peptide was found to be SEC. The cell-type-specific H3K27ac recognition may arise from variations of protein abundances or differential (maybe posttranslational) regulation of the complexes. However, technical limitations of the mass-spectrometry, such as sensitivity in the quantification of low abundant peptides in complex samples, cannot be ruled out as a source of differences between cell types.

HPPD SILAC ratios of H3K27ac alone were generally lower compared to H3K23ac mark alone or to diacetylated H3K23acK27ac (Manuscript, Fig. 1b-d; Supplementary Fig. 1b-d), which suggests that H3K27ac is probably deposited and recognized in concert with other PTMs (Schwämmle *et al.*, 2016).

As mentioned earlier, H3K27ac resides in a sequence motif (ARKS) that appears twice on the N-terminal tail of histone H3.1. Furthermore, H3K9ac and H3K27ac show conserved negative crosstalk (Schwämmle *et al.*, 2016). Interestingly, similar HPPD-MS and HPPD-WB (PAGE-immunoblotting) experiments have detected preferential binding of PBAF complex to H3K9ac, rather than H3K27ac (Vicent *et al.*, 2009; Slaughter *et al.*, 2018). Hence it is possible that H3K9/K27 acetylation provides overlapping transcription stimulating functions as well as distinct features on promoters and enhancers, which might be exploited by PBAF and GBAF, respectively.

4.2 GBAF complex composition

4.2.1 GBAF – a non-canonical BAF complex

Our series of IP-MS experiments have demonstrated that GLTSCR1L is part of a SWI/SNF type complex since most of the detected interacting proteins are known subunits of that chromatin remodeler family. However, it was not clear whether the observed complex composition is complete since, initially, different assemblies of PBAF and BAF were known to exist, and we have not detected SMARCB1 and SMARCE1 (core SWI/SNF subunits), ARID1A/B/2 (PBAF and BAF core subunits), or PBRM1 (BAF180 - PBAF core subunit), which all contain DNA interaction domains (Mashtalir *et al.*, 2018). After subsequent reciprocal pulldown MS analysis with BRD9 and SMARCD1, it became apparent that we are looking at a new and relatively low-abundant assembly of SWI/SNF complex in HeLa cells, the GLTSCR1/1L containing BAF: GBAF. GBAF is defined by the presence of GLTSCR1 or GLTSCR1L and BRD9 proteins.

Recent studies have also described the composition of the GBAF complex (Alpsoy and Dykhuizen, 2018; Mashtalir *et al.*, 2018). Both studies used SWI/SNF subunits for initial IP and glycerol gradient sedimentation to separate SWI/SNF complexes of different sizes. In particular, the first study described the GBAF complex as a distinct non-canonical BAF complex (Alpsoy and Dykhuizen, 2018). After these authors confirmed GLTSCR1 to be a true subunit of catalytically active, non-canonical BAF complex, they performed a series of knockdown experiments to investigate whether the GLTSCR1 and GLTSCR1L proteins are interchangeable. It was found that upon GLTSCR1 knockdown, GLTSCR1L is able to substitute its paralog. Interestingly, this led to an increase of SMARCA4 in GBAF containing glycerol fractions, suggesting that GLTSCR1L containing GBAF forms more stably associated complexes than GLTSCR1 (Alpsoy and Dykhuizen, 2018). It has also been shown, using sequential salt extraction assay, that GBAF (containing GLTSCR1) interacts with chromatin as strongly as the BAF complex subunit ARID1A (contains DNA binding module), but slightly weaker than PBRM1 in

PBAF. Thus, even without multiple DNA interaction modules, the interaction of GBAF with chromatin is at least as strong as that of canonical BAF.

In our study, GLTSCR1L (but not GLTSCR1) and BRD9 were identified in the context of H3K27ac binding, which suggests an enhancer-specific function for these proteins. Due to the absence of DNA binding domain-containing proteins in GBAF, we hypothesize that targeting of GBAF to chromatin is more reliant on the recognition of acetylated histones (namely H3K27ac) or other acetylated lysine-containing proteins, as compared to canonical SWI/SNF complexes that all contain DNA-binding domains. Such domains can promote interaction with nucleosomal DNA, activate SMARCA2/4 ATPase activity, and induce chromatin remodeling (Phelan *et al.*, 1999; Sen *et al.*, 2017).

To conclude, our study is in agreement with recently published GBAF complex composition data (Alpsoy and Dykhuizen, 2018; Mashtalir *et al.*, 2018). Due to atypical composition with the presence of unique components, such as GLTSCR1/1L and BRD9, and multiple targeting mechanisms (discussed below), the function of the GBAF complex is likely to be different from canonical SWI/SNF family members.

4.2.2 GLTSCR1L-mediated interactions and the GiBAF domain

GLTSCR1 is a large (150kDA) protein, which was previously associated with the development of oligodendrogliomas (Yang *et al.*, 2005). GLTSCR1 has been previously detected in multiple proteomic analyses of the SWI/SNF chromatin remodeling complex but has never been validated as a SWI/SNF complex subunit (Ho *et al.*, 2009; Middeljans *et al.*, 2012; Hein *et al.*, 2015). GLTSCR1 was also found to interact with the ET domain of BRD4 protein (Rahman *et al.*, 2011).

GLTSCR1 and its paralog GLTSCR1L both contain a conserved GiBAF (GLTSCR) domain of an unknown function, that shares 63% sequence similarity (Alpsoy and Dykhuizen, 2018). GiBAF domain was earlier found to interact with SH3 containing proteins (Wu *et al.*, 2007). The function of the rest of the protein is unknown. Computational analysis of the protein sequence indicates that it contains several segments of highly disordered nature (as predicted by IUPred; Dosztányi *et al.*, 2005) (<http://iupred.enzim.hu>).

We performed IP-MS experiments with the isolated the GiBAF domain and observed enrichment of all GBAF subunits and additionally BRD4 and AFF1 (components of BEC and SEC complexes, respectively). Very similar results (apart from detected BRD4 and AFF1 interaction) were obtained in the study conducted by Michel *et al.* (2018). Although these experiments were performed with the whole C-terminal half, it was impossible to determine the exact GiBAF domain interaction profile.

Unlike GLTSCR1 (Han *et al.*, 2019), GLTSCR1L does not contain any known nuclear localization signal sequence, but it is still able to translocate to the nucleus. We have shown by IF microscopy, that the GiBAF domain is responsible for that function. Hence we suggest, that the GiBAF domain might be recognized and transported to the nucleus by a nuclear localization signal-independent mechanism (Fagotto, Glück and Gumbiner, 1998). It is an interesting question of why and how GLTSCR1L lost its NLS sequence and which proteins help to transport it into the nucleus.

Since IP-MS with a GiBAF domain deletion mutant protein did not enrich any SWI/SNF subunits, we propose that it can serve as the core scaffold for the GBAF complex. Interestingly, we detected multiple TFIIID complex components, such as TAF4 and TAF6, which link GLTSCR1L to the transcription initiation complex (Louder *et al.*, 2016) and reveals a possible interplay between chromatin remodeling by GBAF and transcription initiation.

Apart from the GiBAF domain, GLTSCR1/1L contains multiple intrinsically disordered regions, which is common for transcription factors. Active transcription sites have recently been associated with phase separation transition (Palikyras and Papanonis, 2019). Highly disordered regions in proteins may play an essential role in this process since they can undergo phase transitions in response to various stimuli (Ruff *et al.*, 2018). Phase separation and largely unexplored properties of unstructured chromatin interacting proteins may be a key mediator of spatial and temporal control over the diverse nuclear processes.

4.2.3 BRD9-mediated GBAF targeting to chromatin

BRD9 is one of the defining GBAF subunits. BRD9 contains a Bromodomain - a well established acetylated lysine interaction module, and a DUF3512 domain of the unknown function (probably protein-protein interaction). BRD9 is related to BRD7, which is PBAF specific subunit. Hence, it is tempting to speculate that PBAF and GBAF might have either similar or opposing functions.

In this work, we provide evidence that the GBAF complex specifically recognizes H3K27ac on chromatin. We show that the BRD9 Bromodomain mediates this interaction. Since, when this domain is blocked by the specific inhibitor (I-BRD9), GBAF can no longer bind to acetylated histone peptides (Manuscript, Fig. 3a). Recent studies have, however, suggested several alternative chromatin targeting mechanisms for BRD9 (Gatchalian *et al.*, 2018; Wei *et al.*, 2018). First, the use of I-BRD9 has been reported to result in concentration and time-dependent decrease in cell proliferation, pointing to a Bromodomain-dependent role for BRD9 in maintaining naive pluripotency of murine ESCs (Gatchalian *et al.*, 2018). In that study, BRD9 was shown to recognize an acetylated form of BRD4, which is thought to target GBAF to some of the BRD4 enriched sites on chromatin. This interaction was affected by inhibition of BRD9 Bromodomain with the I-BRD9 compound, which suggests that this interaction is also BRD9 Bromodomain-dependent. Interestingly BRD4 has been shown to have acetyltransferase activity (Devaiah *et al.*, 2016), which may deposit histone lysine-acetylation targets for BRD9. As a study by Wei *et al.* (2018) demonstrated that, BRD7/9 Bromodomains recognize acetylated lysine K91 of vitamin D receptor (VDR). This interaction facilitates the association of the PBAF and GBAF complexes in a ligand-dependent fashion to VDR target sites and modulates expression of the critical inflammatory response genes in human induced pluripotent stem (iPS) cell-derived B-like cells. VDR K91ac is located in the T-box domain (Quack *et al.*, 1998), which is juxtaposed to its DNA-binding zinc fingers. This proximity of an acetylation target to the DNA binding domain of the transcription factor VDR is reminiscent of acetylation targets on nucleosomal histone tails. Many other transcription factors have acetylated lysines near their DNA binding domains (Sterner and Berger, 2000; Choudhary *et al.*, 2009; Sykes *et al.*, 2009; Xie *et*

al., 2011). It is tempting to speculate that BRD7/9 requires, if not recognizes, a DNA context as well as an acetylated lysine in order to efficiently interact with chromatin. Interestingly, calcipotriol (cal) - a synthetic ligand of VDR receptor (vitamin D analog), when applied, reduces the BRD9-VDR interaction, which points to an important function of BRD9 in the regulation of unliganded VDR target gene expression. Inhibition of BRD9-binding with I-BRD9 as well as depletion of BRD9 by knockdown allowed re-localization of PBAF complex on previously BAF-occupied loci, and increased association of VDR with the PBAF complex and PCAF, which resulted in prolonged transcriptional activation (Wei *et al.*, 2018).

Taken together, recognition of H3K27ac or other acetylated lysines in histones, VDR, or BRD4 by BRD9, may substitute/compensate for the stimulating effect mediated by canonical DNA-binding SWI/SNF subunits.

4.3 GBAF localization in chromatin

4.3.1 GBAF localizes preferentially at enhancers

We performed a series of ChIP-seq experiments that showed overlap between genomic regions enriched with BRD9, GLTSCR1L, H3K27ac, and a panel of active chromatin marks (Manuscript, Fig. 3). Our finding of localization bias of GBAF to the intronic/intergenic region also supports our initial hypothesis that GBAF possibly facilitates an enhancer-specific function on chromatin. Our analysis of ChIP-seq data also indicated an enrichment of the enhancer-associated histone mark H3K4me1 rather than promoter-associated H3K4me3 (Manuscript, Fig. 3c). Several recent studies have reported similar ChIP-seq data and analyses of GBAF complex components. The study by Michel *et al.* (2018) also shows a significant overlap between binding sites of BRD9 and GLTSCR1 protein. However, these authors suggest that the GBAF complex is mostly associated with gene promoters and CTCF sites. Although we also observed enrichment to CTCF sites (Appendix A-B), our data show that GBAF is enriched at enhancers, rather than promoters. CTCF is a protein known to mark the TAD boundaries. It has also been shown that CTCF and TAD

boundaries are mostly located in intergenic regions (Cuddapah *et al.*, 2009), which may account for at least a fraction of the GBAF-bound intergenic loci. Nevertheless, our data differ from some of the data of Michel *et al.* (2018), which suggests promoter localization of the GBAF complex. Michel *et al.* (2018) and Gatchalian *et al.* (2018) have both performed annotation of BRD9-bound chromatin regions with chromatin marks: H3K4me1 (enhancer-associated), H3K4me3 (promoter-associated), H3K27ac (active transcription-associated). Recent studies have surprisingly shown that most of the active enhancers harbor H3K4me3 rather than H3K4me1. Furthermore, the level of H3K4 methylation corresponds to the transcription intensity at a given locus (Henriques *et al.*, 2018) and does not distinguish active enhancers from promoters. Based on the enrichment of those marks, both studies claim that BRD9 is associated with promoters. It has to be noted, however, that these chromatin marks are associations, not transcriptionally validated promoters or enhancers (Andersson, 2015).

On the other hand, a study conducted with sarcoma cell lines by Brien *et al.* (2018), who used genomic annotations, assigned most of the BRD9 peaks (65%) to intergenic/intronic regions, in agreement with our data. Interestingly, following I-BRD9 treatment, we observed a reduction of BRD9-bound intronic/intergenic regions and most of the GLTSCR1L peaks to remain at promoter regions. These observations support the idea that the GBAF complex is involved in chromatin remodeling activity when promoters and enhancers are interacting. The observed loss of GLTSCR1L from enhancers upon BRD9 bromodomain inhibition indicates that GBAF is specifically bound to enhancers while remaining bound unspecifically to promoters.

Due to a large number of paralogs, SWI/SNF family remodeling complexes might, in theory, form about ~1500 different assemblies (Mashtalir *et al.*, 2018). Some of the subunits that contain DNA/histone interaction domains (for example, SMARCA2/4) are universal for all of the SWI/SNF assemblies. This leads to the partial colocalization not even between different assemblies of the same type (for example, PBAF and nPBAF), but also between different types (PBAF partially overlaps with BAF) of the SWI/SNF complexes. On the other hand, type-specific subunits, which contain DNA/histone or other chromatin targeting domains, could

differentiate their activity. Moreover, the same subunits can endow different localization patterns depending on which type of SWI/SNF complex they reside in. For example, SMARCB1 - a subunit that is shared between both BAF and PBAF, has been shown to affect promoter-distal (enhancer) localization of BAF, but not the PBAF complex (Nakayama *et al.*, 2017). No systematic genome-wide study to map and differentiate localization of different types of SWI/SNF assemblies have yet been performed. Recent studies suggest, however, that different types of SWI/SNF complexes have specialized roles and complex-specific localization (Ho, Lloyd and Bao, 2019).

4.3.2 Inhibition of H3K27ac recognition by BRD9 diminishes GBAF localisation in chromatin

By impeding the H3K27ac recognition of GBAF with the selective inhibitor of BRD9 (I-BRD9), we did not destroy the GBAF complex itself, as would have happened with a knockdown/out. Thus when using I-BRD9, we did not impair the possible structural role of that complex.

Our HPPD-WB data indicates that upon I-BRD9 treatment, GLTSCR1L binding to histone-peptides is severely diminished. Interestingly we still detected the interaction of acetylated histone peptides with BRD9, but the specificity towards H3K27ac recognition was lost. These data have led us to draw two conclusions:

1. BRD9 can form a partial assembly with the ATPase module of SWI/SNF (Mashtalir *et al.*, 2018), that can interact with histone peptides in a BRD9 Bromodomain-independent fashion
2. upon H3K27ac binding, BRD9 may undergo a conformational change that facilitates GLTSCR1L-binding and the formation of the active GBAF complex. Alternatively, I-BRD9 may evoke a conformational change that inhibits BRD9's ability to interact with GLTSCR1L and non-ATPase subunits such as SMARCD1(BAF60a) and SMARCC1(BAF155)

Supporting data was obtained by ChIP-qPCR at the NR4A1 locus and globally by ChIP-seq. We observed abrogation of GLTSCR1L interaction with chromatin upon I-BRD9 treatment, whereas BRD9 remained on chromatin. Other inhibitors of

the BRD9 Bromodomain have shown similar effects on BRD9 (Hohmann *et al.*, 2016). The study by Brien *et al.* (2018) performed systematic analyses of BRD9 inhibitors and mutational inactivation of its Bromodomain and DUF3512 domains in different cancer cell lines and arrived at similar conclusions.

Our observed defects of GBAF assembly upon I-BRD9 treatment, suggest that acetyl-lysine recognition is essential for GBAF integrity. Acetyl-lysine recognition has been shown to induce a conformational change in the Bromodomain of the RSC complex (Skiniotis, Moazed and Walz, 2007) and proposed to stabilize certain interactions within the complex, which is a prerequisite for subsequent remodeling activity. Also, the Bromodomain of CBP/P300 was found to be involved in H3K27 acetylation regulation (Park *et al.*, 2017). Furthermore, impairment of H3K18 acetyl-lysine recognition by P300-specific Bromodomain inhibitor (GNE-049) severely reduced H3K27ac levels and, consequently, enhancer-specific activity (Raisner *et al.*, 2018). Regarding the previously observed incomplete dislocation of BRD9 from chromatin upon I-BRD9 treatment, BRD9 may bind to chromatin in a Bromodomain-independent manner. Furthermore, it may be possible that GBAF is sequentially assembled on chromatin, whereas recognition of H3K27ac (or other acetylated lysines) leads to the conformational change of BRD9 that facilitates the formation of fully active GBAF complex. Interestingly, Alpsy and Dyukhuiz (2018) have also observed that knockdown of GLTSCR1 or GLTSCR1L proteins reduced the BRG1-associated BRD9 levels, which suggests that BRD9 alone is not able to assemble catalytically active complex. Further structural analyses are needed to identify if BRD9 undergoes structural rearrangements upon acetyl-lysine recognition, which could facilitate signal-dependent protein assembly/activity regulation.

Altogether, our HPPD-WB and ChIP-seq data have shown that the recognition of H3K27ac by the Bromodomain of BRD9 is essential not only for targeting GBAF but also for the assembly of the complex.

4.4 Transcriptional effects of H3K27ac recognition inhibition in GBAF

Enhancer-promoter interactions are known to be regulated by the three-dimensional genome structure. These interactions typically reside within large chromosome TADs (Dixon *et al.*, 2012; Nora *et al.*, 2012), which are limited by boundaries (also-called insulator elements) formed by CTCF and cohesin (Parelho *et al.*, 2008; Wendt *et al.*, 2008), and which prevent interdomain contacts (Narendra *et al.*, 2015).

About 70% of CAGE-defined enhancers were validated in reporter assays (Andersson, Gebhard, *et al.*, 2014). The remaining fraction of transcribed elements may have other functions. Transcription has been reported not only for promoters and enhancers but also for insulators (Melgar, Collins and Sethupathy, 2011) and accessible DNA in general (Young *et al.*, 2017). We show that GBAF was found to co-localize with CTCF; thus, observed significant downregulation of intronic/intergenic transcription rates might at least partially arise from CTCF sites (insulators).

Apart from downregulated intronic/intergenic transcripts, we also observed differential expression of mRNA genes. Despite that the global level of transcription from intronic/intergenic regions was significantly reduced, transcription from TSS's was not affected by I-BRD9 treatment. To summarize, inhibition of the GBAF H3K27ac recognition impairs the correct gene expression globally through the downregulation of eRNA at transcribed enhancers.

While global mRNA expression levels were not affected by I-BRD9, transcription from TSS of genes residing within TADs with downregulated enhancers tended to be downregulated. However, at our model locus, NR4A1, the counterintuitive effect of I-BRD9 was observed. eRNA expression at the NR4A1 enhancer was severely downregulated while the NR4A1 promoter was upregulated following I-BRD9 treatment (Manuscript, Supplementary Fig. 5b). An increase in NR4A1 promoter upon I-BRD9 treatment might result from the formation of alternative DNA loops by the promoter within the TAD it resides in (Long, Prescott and Wysocka, 2016). As reported by Lai *et al.*, 2015, the NR4A1 locus has several

enhancers and promoters. When multiple promoters and enhancers are shared, enhancer-promoter targeting might be not additive, but sub-additive due to competition for the promoter(s) (Bothma *et al.*, 2015).

Furthermore, upregulation of promoter transcription and partial co-localization of GBAF and factors such as CTCF and cohesin in chromatin imply that the GBAF complex might be involved in the fine-tuning 3D organization of the NR4A1 TAD or smaller insulating chromatin loops within the TAD (Hao, Shearwin and Dodd, 2019; Kim *et al.*, 2019) ensuring the correct level of the expression of that gene.

However, it may also be that functionalities of enhancer and promoter in the case with NR4A1 locus may not be interdependent (Arnold *et al.*, 2017; Catarino, Neumayr and Stark, 2017; Dao *et al.*, 2017; Diao *et al.*, 2017; Mikhaylichenko *et al.*, 2018). Another possibility is that enhancer-independent signaling by hEGF directly to the promoter was sufficient to drive the NR4A1 gene expression under our experiments' conditions. If that was the case, it is reasonable to question the significance of enhancer transcription in target gene regulation and prompt the search for other non-trivial functions of eRNA expression. Nevertheless, at the NR4A1 locus, we show that GBAF is involved exclusively in enhancer transcription modulation.

Global transcription levels measured by CAGE, coupled with ATAC-seq (mostly for validating transcriptional events), have led us to conclude that the GBAF complex has enhancer-specific function noted by downregulation of eRNA transcription and corroborated by a minor, but significant decrease in chromatin accessibility at GLTSCR1L-depleted enhancers in I-BRD9-treated HeLa cells. However, the minor decrease in chromatin accessibility might also be explained by reduced transcription rates and not the activity of GBAF itself.

Taken together, we propose that GBAF is essential for both enhancer transcription and enhancer regulatory activity to drive distal transcription.

4.5 Possible role of H3K27ac in enhancer

According to our data, both the SEC and the GBAF complexes can recognize H3K27ac. These two complexes facilitate two processes, transcription elongation pause release and ATP-dependent chromatin remodeling. Moreover, we have found a link between these two processes: GiBAF domain of GLTSCR1L protein interacts with AFF1 from SEC, and BRD4 that, similarly to SEC, promotes CDK9 mediated RNAP II pause release. Hence, we suggest that transcription elongation pause release and ATP-dependent chromatin remodeling might be interconnected.

The SEC components AF-9 and ENL can recognize both H3K9ac and H3K27ac histone marks. However, both of these proteins preferentially associate with H3K9ac rather than H3K27ac (Li *et al.*, 2014; Erb *et al.*, 2017). Although H3K9ac and H3K27ac have been reported to be mutually exclusive chromatin marks on histone tails (Schwämmle *et al.*, 2016), both of them can be found at active enhancers and promoters (Karmodiya *et al.*, 2012). Since both of these regulatory elements are transcriptionally active, SEC-mediated promotion of transcription elongation can be utilized by both. However, the preference of the SEC towards H3K9ac binding could explain higher transcription rates at promoters than enhancers (Mikhaylichenko *et al.*, 2018).

In contrast, GBAF has been found to interact with H3K27ac only. The GBAF complex has only recently been described (Alpsoy and Dykhuizen, 2018; Brien *et al.*, 2018; Gatchalian *et al.*, 2018; Jefimov *et al.*, 2018; Mashtalir *et al.*, 2018; Michel *et al.*, 2018), and there is so far little data that can shed light on mechanisms employed by this complex. Our data shows that GBAF has enhancer-specific functions, but the exact mechanisms are yet to be elucidated.

Taken together, we propose a dual role for the H3K27ac. First, H3K27ac can recruit GBAF and facilitate ATP-dependent chromatin remodeling, preferentially at enhancers. Secondly, this mark promotes transcription elongation pause release by SEC/BEC mediated RNAP II CTD Ser2 phosphorylation.

5. Concluding remarks and future perspectives

Modern advances in methods for investigating chromatin biology have the potential to reveal the regulatory mechanisms of chromatin. In the current study, we have used a combined proteomic and functional genome-wide approaches to provide new insights into the role of the H3K27ac chromatin mark in enhancer function. In search of specific binders for the H3K27ac, we detected the GBAF complex that connects histone acetylation, chromatin remodeling, and enhancer-specific activities. By impeding H3K27ac recognition by GBAF with the selective inhibitor of BRD9 (I-BRD9), we induced transcriptional defects preferentially at enhancer elements and observed effects on gene expression. This indicates the enhancer-specific regulatory activity of the GBAF complex.

Although we have uncovered an important role of H3K27ac and GBAF in enhancer regulatory function, new questions arise. First of all, it is not clear exactly how GBAF exerts its enhancer-specific function and whether it takes part in the establishment of productive enhancer-promoter communication. It is, therefore, of great interest to analyze how the dysregulation of GBAF activity influences the 3D chromosome organization. It is of high importance to follow up this study by methods that can detect nucleosome remodeling activity, such as ATAC-seq, to assess how GBAF affects chromatin accessibility for TFs and other regulatory factors. Furthermore, identification of structural rearrangements in BRD9 upon acetyl-lysine binding could lead to a better understanding of the signal-induced activity of chromatin remodelers since it is likely a common regulatory mechanism.

Regulation of gene expression is a dynamic and continuous process. High-throughput methods, such as genome-wide chromatin profiling and proteomics, have significantly contributed to the recent advances in our understanding of gene regulation. However, due to a limited number of time points of measurements and scattered datasets from different model systems, we are still far from a holistic view of how mechanisms of gene expression regulation work. Further development, combination, and integration of different methods of molecular biology should pave the way for sophisticated and versatile analyses of chromatin-mediated gene

expression regulation. Optimization of already existing ChIP-seq and other chromatin-profiling protocols will be required to obtain a higher time-resolution to observe the dynamics of these processes. A combination of proteomics and chromatin-conformation capture techniques would also allow revealing the protein content of communicating chromatin regions. An integration of different types of datasets may facilitate the establishment of the link between different molecular events that have been previously elusive. This should lead us to a qualitative leap in our understanding of chromatin-mediated regulatory mechanisms.

References

- Allemand, E., Batsché, E. and Muchardt, C. (2008) 'Splicing, transcription, and chromatin: a ménage à trois', *Current Opinion in Genetics and Development*, 18(2), pp. 145–151. doi: 10.1016/j.gde.2008.01.006.
- Allfrey, V. G., Faulkner, R. and 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*, 51(5), pp. 786–794. doi: 10.1073/pnas.51.5.786.
- De Almeida, S. F. and Carmo-Fonseca, M. (2010) 'Cotranscriptional RNA checkpoints', *Epigenomics*, 2(3), pp. 449–455. doi: 10.2217/epi.10.21.
- Alpsoy, A. and Dykhuizen, E. C. (2018) 'Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes', *Journal of Biological Chemistry*, 293(11), pp. 3892–3903. doi: 10.1074/jbc.RA117.001065.
- Andersson, R., Gebhard, C., *et al.* (2014) 'An atlas of active enhancers across human cell types and tissues.', *Nature*, 507(7493), pp. 455–61. doi: 10.1038/nature12787.
- Andersson, R., Refsing Andersen, P., *et al.* (2014) 'Nuclear stability and transcriptional directionality separate functionally distinct RNA species.', *Nature communications*. Nature Publishing Group, 5, p. 5336. doi: 10.1038/ncomms6336.
- Andersson, R. (2015) 'Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model', *BioEssays*, 37(3), pp. 314–323. doi: 10.1002/bies.201400162.
- Andersson, R. and Sandelin, A. (2020) 'Determinants of enhancer and promoter activities of regulatory elements', *Nature Reviews Genetics*. Springer US, 21(2), pp. 71–87. doi: 10.1038/s41576-019-0173-8.
- Anunziato, A. T. and Hansen, J. C. (2000) 'Role of histone acetylation in the assembly and modulation of chromatin structures', *Gene Expression*, 9(1–2), pp. 37–61. doi: 10.3727/000000001783992687.
- Archacki, R. *et al.* (2009) 'Genetic analysis of functional redundancy of BRM ATPase and ATSWI3C subunits of Arabidopsis SWI/SNF chromatin remodelling complexes', *Planta*, 229(6), pp. 1281–1292. doi: 10.1007/s00425-009-0915-5.
- Arents, G. *et al.* (1991) 'The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix.', *Proceedings of the National Academy of Sciences*, 88(22), pp. 10148–10152. doi: 10.1073/pnas.88.22.10148.
- Arents, Gina *et al.* (1991) 'The nucleosomal core histone octamer at 3.1 Å resolution: A tripartite protein assembly and a left-handed superhelix', *Proceedings of the*

National Academy of Sciences of the United States of America, 88(22), pp. 10148–10152. doi: 10.1073/pnas.88.22.10148.

Arents, G. and Moudrianakis, E. N. (1993) ‘Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA’, *Pnas*, 90(November), pp. 10489–10493. doi: 10.1073/pnas.90.22.10489.

Arnold, C. D. *et al.* (2017) ‘Genome-wide assessment of sequence-intrinsic enhancer responsiveness at single-base-pair resolution’, *Nature Biotechnology*, 35(2), pp. 136–144. doi: 10.1038/nbt.3739.

Audergon, P. N. C. B. *et al.* (2015) ‘Restricted epigenetic inheritance of H3K9 methylation’, *Science*, 348(6230), pp. 132–135. doi: 10.1126/science.1260638.

Avery, O. T. (1944) ‘STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES: INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III’, *Journal of Experimental Medicine*, 79(2), pp. 137–158. doi: 10.1084/jem.79.2.137.

Bacon, C. W. and D’Orso, I. (2019) ‘CDK9: a signaling hub for transcriptional control’, *Transcription*. Taylor & Francis, 10(2), pp. 57–75. doi: 10.1080/21541264.2018.1523668.

Bannister, A. J. and Kouzarides, T. (2011) ‘Regulation of chromatin by histone modifications’, *Cell Research*. Nature Publishing Group, 21(3), pp. 381–395. doi: 10.1038/cr.2011.22.

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

Bedford, M. T. and Clarke, S. G. (2009) ‘Protein Arginine Methylation in Mammals: Who, What, and Why’, *Molecular Cell*. Elsevier Inc., 33(1), pp. 1–13. doi: 10.1016/j.molcel.2008.12.013.

Bednar, J. *et al.* (2017) ‘Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1’, *Molecular Cell*, 66(3), pp. 384–397.e8. doi: 10.1016/j.molcel.2017.04.012.

Bell, O. *et al.* (2011) ‘Determinants and dynamics of genome accessibility’, *Nature Reviews Genetics*. Nature Publishing Group, 12(8), pp. 554–564. doi: 10.1038/nrg3017.

Bender, S. *et al.* (2013) ‘Reduced H3K27me3 and DNA Hypomethylation Are Major Drivers of Gene Expression in K27M Mutant Pediatric High-Grade Gliomas’, *Cancer Cell*, 24(5), pp. 660–672. doi: 10.1016/j.ccr.2013.10.006.

Berger, S. L. *et al.* (2009) ‘An operational definition of epigenetics’, *Genes and Development*, 23(7), pp. 781–783. doi: 10.1101/gad.1787609.

-
- Bernstein, E. *et al.* (2006) ‘Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin’, *Molecular and Cellular Biology*, 26(7), pp. 2560–2569. doi: 10.1128/mcb.26.7.2560-2569.2006.
- Bickmore, W. A. and Van Steensel, B. (2013) ‘Genome architecture: Domain organization of interphase chromosomes’, *Cell*. Elsevier Inc., 152(6), pp. 1270–1284. doi: 10.1016/j.cell.2013.02.001.
- Bird, A. P. (1978) ‘Use of restriction enzymes to study eukaryotic DNA methylation. II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern’, *Journal of Molecular Biology*, 118(1), pp. 49–60. doi: 10.1016/0022-2836(78)90243-7.
- Birney, E. *et al.* (2007) ‘Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project’, *Nature*, 447(7146), pp. 799–816. doi: 10.1038/nature05874.
- Bögershausen, N. and Wollnik, B. (2018) ‘Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders’, *Frontiers in Molecular Neuroscience*, 11(August), pp. 1–18. doi: 10.3389/fnmol.2018.00252.
- Böhm, L. and Crane-Robinson, C. (1984) ‘Proteases as structural probes for chromatin: The domain structure of histones - Review’, *Bioscience Reports*, 4(5), pp. 365–386. doi: 10.1007/BF01122502.
- Bolzer, A. *et al.* (2005) ‘Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes’, *PLoS Biology*, 3(5), pp. 0826–0842. doi: 10.1371/journal.pbio.0030157.
- Bonn, S. *et al.* (2012) ‘Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development’, *Nature Genetics*, 44(2), pp. 148–156. doi: 10.1038/ng.1064.
- Bose, D. A. *et al.* (2017) ‘RNA Binding to CBP Stimulates Histone Acetylation and Transcription’, *Cell*. Elsevier, 168(1–2), pp. 135–149.e22. doi: 10.1016/j.cell.2016.12.020.
- Bothma, J. P. *et al.* (2015) ‘Enhancer additivity and non-additivity are determined by enhancer strength in the *Drosophila* embryo’, *eLife*, 4, pp. 1–14. doi: 10.7554/eLife.07956.
- Boyer, L. A., Latek, R. R. and Peterson, C. L. (2004) ‘The SANT domain: A unique histone-tail-binding module?’, *Nature Reviews Molecular Cell Biology*, 5(2), pp. 158–163. doi: 10.1038/nrm1314.
- Brackley, C. A. *et al.* (2016) ‘Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological

-
- domains', *Nucleic Acids Research*, 44(8), pp. 3503–3512. doi: 10.1093/nar/gkw135.
- Branco, M. R. and Pombo, A. (2006) 'Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations', *PLoS Biology*, 4(5), pp. 780–788. doi: 10.1371/journal.pbio.0040138.
- Bresnick, E. H., John, S. and Hager, G. L. (1991) 'Histone Hyperacetylation Does Not Alter the Positioning or Stability of Phased Nucleosomes on the Mouse Mammary Tumor Virus Long Terminal Repeat', *Biochemistry*, 30(14), pp. 3490–3497. doi: 10.1021/bi00228a020.
- Brien, G. L. *et al.* (2018) 'Targeted degradation of BRD9 reverses oncogenic gene expression in synovial sarcoma', *eLife*, 7, pp. 1–26. doi: 10.7554/eLife.41305.
- Brown, C. E. *et al.* (2000) 'The many HATs of transcription coactivators', *Trends in Biochemical Sciences*, 25(1), pp. 15–19. doi: 10.1016/S0968-0004(99)01516-9.
- Brown, D. T. (2003) 'Histone H1 and the dynamic regulation of chromatin function', *Biochemistry and Cell Biology*, 81(3), pp. 221–227. doi: 10.1139/o03-049.
- Bruno, M. *et al.* (2003) 'Histone H2A/H2B Dimer Exchange by ATP-Dependent Chromatin Remodeling Activities', *Molecular Cell*, 12(6), pp. 1599–1606. doi: 10.1016/S1097-2765(03)00499-4.
- Buckley, M. S. *et al.* (2014) 'Kinetics of promoter Pol II on Hsp70 reveal stable pausing and key insights into its regulation', *Genes and Development*, 28(1), pp. 14–19. doi: 10.1101/gad.231886.113.
- Bulger, M. and Groudine, M. (2011) 'Functional and mechanistic diversity of distal transcription enhancers', *Cell*. Elsevier Inc., 144(3), pp. 327–339. doi: 10.1016/j.cell.2011.01.024.
- Bultman, S. J. *et al.* (2006) 'Maternal BRG1 regulates zygotic genome activation in the mouse', *Genes and Development*, 20(13), pp. 1744–1754. doi: 10.1101/gad.1435106.
- Buratowski, S. (2009) 'Progression through the RNA Polymerase II CTD Cycle', *Molecular Cell*. Elsevier Inc., 36(4), pp. 541–546. doi: 10.1016/j.molcel.2009.10.019.
- Cairns, B. R. *et al.* (1996) 'RSC, an essential, abundant chromatin-remodeling complex', *Cell*, 87(7), pp. 1249–1260. doi: 10.1016/S0092-8674(00)81820-6.
- Campos, E. I. and Reinberg, D. (2009) 'Histones : Annotating Chromatin', pp. 559–601. doi: 10.1146/annurev.genet.032608.103928.
- Cao, R. *et al.* (2002) 'Role of histone H3 lysine 27 methylation in Polycomb group silencing', *Science*, 298(November), pp. 1039–1044. doi: 10.1126/science.1076997.
- Capdevila, M. P. and Garcia-Bellido, A. (1981) 'Genes involved in the activation of

the bithorax complex of *Drosophila*', *Wilhelm Roux's Archives of Developmental Biology*, 190(6), pp. 339–350. doi: 10.1007/BF00863271.

Carninci, P. *et al.* (2006) 'Genome-wide analysis of mammalian promoter architecture and evolution', *Nature Genetics*, 38(6), pp. 626–635. doi: 10.1038/ng1789.

Catarino, R. R., Neumayr, C. and Stark, A. (2017) 'Promoting transcription over long distances', *Nature Genetics*. Nature Publishing Group, 49(7), pp. 972–973. doi: 10.1038/ng.3904.

Cedar, H. and Felsenfeld, G. (1973) 'Transcription of chromatin in vitro', *Journal of Molecular Biology*, 77(2), pp. 237–254. doi: 10.1016/0022-2836(73)90334-3.

Chang, F. T. M. *et al.* (2015) 'CHK1-driven histone H3.3 serine 31 phosphorylation is important for chromatin maintenance and cell survival in human ALT cancer cells', *Nucleic Acids Research*, 43(5), pp. 2603–2614. doi: 10.1093/nar/gkv104.

Chatterjee, N. *et al.* (2011) 'Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms', *Nucleic Acids Research*, 39(19), pp. 8378–8391. doi: 10.1093/nar/gkr535.

Chen, F. *et al.* (2015) 'Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide', *Genes and Development*, 29(1), pp. 39–47. doi: 10.1101/gad.246173.114.

Chorev, M. and Carmel, L. (2012) 'The function of introns', *Frontiers in Genetics*, 3(APR), pp. 1–15. doi: 10.3389/fgene.2012.00055.

Choudhary, C. *et al.* (2009) 'Lysine acetylation targets protein complexes and co-regulates major cellular functions', *Science*, 325(5942), pp. 834–840. doi: 10.1126/science.1175371.

Clapier, C. R. *et al.* (2017) 'Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(7), pp. 407–422. doi: 10.1038/nrm.2017.26.

Clapier, C. R. and Cairns, B. R. (2009) 'The Biology of Chromatin Remodeling Complexes', *Annual Review of Biochemistry*, 78(1), pp. 273–304. doi: 10.1146/annurev.biochem.77.062706.153223.

Cook, P. R. and Marenduzzo, D. (2018) 'Transcription-driven genome organization: A model for chromosome structure and the regulation of gene expression tested through simulations', *Nucleic Acids Research*. Oxford University Press, 46(19), pp. 9895–9906. doi: 10.1093/nar/gky763.

Corona, D. F. V *et al.* (2001) 'Critical Role for the Histone H4 N Terminus in Nucleosome Remodeling by ISWI', *Molecular and cellular biology*, 21(3), pp. 875–883. doi: 10.1128/MCB.21.3.875.

-
- Cremer, T. and Cremer, M. (2010) 'Chromosome territories.', *Cold Spring Harbor perspectives in biology*, 2(3), pp. 1–22. doi: 10.1101/cshperspect.a003889.
- Creighton, M. P. *et al.* (2010) 'Histone H3K27ac separates active from poised enhancers and predicts developmental state.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), pp. 21931–21936. doi: 10.1073/pnas.1016071107.
- Cuddapah, S. *et al.* (2009) 'Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains.', *Genome research*, 19(1), pp. 24–32. doi: 10.1101/gr.082800.108.
- Dann, G. P. *et al.* (2017) 'ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference', *Nature*. Nature Publishing Group, 548(7669), pp. 607–611. doi: 10.1038/nature23671.
- Dao, L. T. M. *et al.* (2017) 'Genome-wide characterization of mammalian promoters with distal enhancer functions', *Nature Genetics*. Nature Publishing Group, 49(7), pp. 1073–1081. doi: 10.1038/ng.3884.
- Deal, R. B., Henikoff, J. G. and Henikoff, S. (2010) 'Genome-Wide Kinetics of Nucleosome', *Science*, 328(5982), pp. 1161–1165. doi: 10.1126/science.1186777.
- Dekker, J. and Heard, E. (2015) 'Structural and functional diversity of Topologically Associating Domains', *FEBS Letters*. Federation of European Biochemical Societies, 589(20), pp. 2877–2884. doi: 10.1016/j.febslet.2015.08.044.
- Devaiah, B. N. *et al.* (2016) 'BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin', *Nature Structural and Molecular Biology*. Nature Publishing Group, 23(6), pp. 540–548. doi: 10.1038/nsmb.3228.
- Dey, A. *et al.* (2000) 'A Bromodomain Protein, MCAP, Associates with Mitotic Chromosomes and Affects G2-to-M Transition', *Molecular and Cellular Biology*, 20(17), pp. 6537–6549. doi: 10.1128/mcb.20.17.6537-6549.2000.
- Diao, Y. *et al.* (2017) 'A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells', *Nature Methods*. Nature Publishing Group, 14(6), pp. 629–635. doi: 10.1038/nmeth.4264.
- Dion, M. F. *et al.* (2007) 'Dynamics of replication-independent histone turnover in budding yeast', *Science*, 315(5817), pp. 1405–1408. doi: 10.1126/science.1134053.
- Dixon, J. R. *et al.* (2012) 'Topological domains in mammalian genomes identified by analysis of chromatin interactions', *Nature*. Nature Publishing Group, 485(7398), pp. 376–380. doi: 10.1038/nature11082.
- Dixon, J. R. *et al.* (2015) 'Chromatin architecture reorganization during stem cell differentiation', *Nature*. Nature Publishing Group, 518(7539), pp. 331–336. doi: 10.1038/nature14222.

- Dixon, J. R., Gorkin, D. U. and Ren, B. (2016) 'Chromatin Domains: The Unit of Chromosome Organization', *Molecular Cell*. Elsevier Inc., 62(5), pp. 668–680. doi: 10.1016/j.molcel.2016.05.018.
- Dong, F., Hansen, J. C. and van Holde, K. E. (1990) '{DNA} and protein determinants of nucleosome positioning on sea urchin 5{S} r{RNA} gene sequences in vitro.', *Proc Natl Acad Sci U S A*, 87(15), pp. 5724–5728. doi: 10.1073/pnas.87.15.5724.
- Dorigi, K. M. *et al.* (2017) 'MII3 and MII4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation', *Molecular Cell*. Elsevier Inc., 66(4), pp. 568–576.e4. doi: 10.1016/j.molcel.2017.04.018.
- Dosztányi, Z. *et al.* (2005) 'IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content', *Bioinformatics*, 21(16), pp. 3433–3434. doi: 10.1093/bioinformatics/bti541.
- Doyon, Y. *et al.* (2004) 'Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans.', *Molecular and cellular biology*, 24(5), pp. 1884–96. doi: 10.1128/MCB.24.5.1884.
- Dutta, A. *et al.* (2014) 'Swi/Snf dynamics on stress-responsive genes is governed by competitive bromodomain interactions', *Genes and Development*, 28(20), pp. 2314–2330. doi: 10.1101/gad.243584.114.
- Dynlacht, B. D., Hoey, T. and Tjian, R. (1991) 'Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation', *Cell*, 66(3), pp. 563–576. doi: 10.1016/0092-8674(81)90019-2.
- Eberl, H. C. *et al.* (2013) 'A Map of General and Specialized Chromatin Readers in Mouse Tissues Generated by Label-free Interaction Proteomics', *Molecular Cell*. Elsevier Inc., 49(2), pp. 368–378. doi: 10.1016/j.molcel.2012.10.026.
- Egloff, S. and Murphy, S. (2008) 'Cracking the RNA polymerase II CTD code', *Trends in Genetics*, 24(6), pp. 280–288. doi: 10.1016/j.tig.2008.03.008.
- Engelen, E. *et al.* (2015) 'Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry.', *Nature communications*, 6(May), p. 7155. doi: 10.1038/ncomms8155.
- Erb, M. A. *et al.* (2017) 'Transcription control by the ENL YEATS domain in acute leukaemia', *Nature*. Nature Publishing Group, 543(7644), pp. 270–274. doi: 10.1038/nature21688.
- Ernst, J. *et al.* (2011) 'Mapping and analysis of chromatin state dynamics in nine human cell types', *Nature*. Nature Publishing Group, 473(7345), pp. 43–49. doi: 10.1038/nature09906.
- Eustermann, S. *et al.* (2018) 'Structural basis for ATP-dependent chromatin

- remodelling by the INO80 complex', *Nature*. Springer US, 556(7701), pp. 386–390. doi: 10.1038/s41586-018-0029-y.
- Fagotto, F., Glück, U. and Gumbiner, B. M. (1998) 'Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of β -catenin', *Current Biology*, 8(4), pp. 181–190. doi: 10.1016/S0960-9822(98)70082-X.
- Farnham, P. J. (2009) 'Insights from genomic profiling of transcription factors', *Nature Reviews Genetics*, 10(9), pp. 605–616. doi: 10.1038/nrg2636.
- Ferreira, H., Flaus, A. and Owen-Hughes, T. (2007) 'Histone Modifications Influence the Action of Snf2 Family Remodelling Enzymes by Different Mechanisms', *Journal of Molecular Biology*. Elsevier Ltd, 374(3), pp. 563–579. doi: 10.1016/j.jmb.2007.09.059.
- Filbin, M. G. *et al.* (2018) 'Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq', 335(April), pp. 331–335.
- Filion, G. J. *et al.* (2010) 'Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells', *Cell*, 143(2), pp. 212–224. doi: 10.1016/j.cell.2010.09.009.
- Filipescu, D., Müller, S. and Almouzni, G. (2014) 'Histone H3 Variants and Their Chaperones During Development and Disease: Contributing to Epigenetic Control', *Annual Review of Cell and Developmental Biology*, 30(1), pp. 615–646. doi: 10.1146/annurev-cellbio-100913-013311.
- Fischle, W. *et al.* (2003) 'Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by polycomb and HP1 chromodomains', *Genes and Development*, 17(15), pp. 1870–1881. doi: 10.1101/gad.1110503.
- Fischle, W. *et al.* (2005) 'Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation', *Nature*, 438(7071), pp. 1116–1122. doi: 10.1038/nature04219.
- Fischle, W., Wang, Y. and Allis, C. D. (2003) 'Histone and chromatin cross-talk', *Current Opinion in Cell Biology*, 15(2), pp. 172–183. doi: 10.1016/S0955-0674(03)00013-9.
- Flaus, A. *et al.* (2006) 'Identification of multiple distinct Snf2 subfamilies with conserved structural motifs', *Nucleic Acids Research*, 34(10), pp. 2887–2905. doi: 10.1093/nar/gkl295.
- Flaus, A. and Owen-Hughes, T. (2011) 'Mechanisms for ATP-dependent chromatin remodelling: The means to the end', *FEBS Journal*, 278(19), pp. 3579–3595. doi: 10.1111/j.1742-4658.2011.08281.x.
- Fletcher, T. M. and Hansen, J. C. (1996) 'The Nucleosomal Array: Structure/Function Relationships', *Critical Reviews □ in Eukaryotic Gene Expression*,

6(2–3), pp. 149–188. doi: 10.1615/CritRevEukarGeneExpr.v6.i2-3.40.

Forquignon, F. (1981) ‘A maternal effect mutation leading to deficiencies of organs and homeotic transformations in the adults of *Drosophila*’, *Wilhelm Roux's Archives of Developmental Biology*, 190(3), pp. 132–138. doi: 10.1007/BF00867798.

Forsberg, E. C. and Bresnick, E. H. (2001) ‘Histone acetylation beyond promoters: Long-range acetylation patterns in the chromatin world’, *BioEssays*, 23(9), pp. 820–830. doi: 10.1002/bies.1117.

Franklin, R. E. and Gosling, R. G. (1953) ‘Molecular configuration in sodium thymonucleate.’, *Nature*, 171(4356), pp. 740–1. doi: 10.1038/171740a0.

Fu, T. J. *et al.* (1999) ‘Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription’, *Journal of Biological Chemistry*, 274(49), pp. 34527–34530. doi: 10.1074/jbc.274.49.34527.

Fudenberg, G. *et al.* (2016) ‘Formation of Chromosomal Domains by Loop Extrusion’, *Cell Reports*. The Author(s), 15(9), pp. 2038–2049. doi: 10.1016/j.celrep.2016.04.085.

Fujisawa, T. and Filippakopoulos, P. (2017) ‘Functions of bromodomain-containing proteins and their roles in homeostasis and cancer’, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(4), pp. 246–262. doi: 10.1038/nrm.2016.143.

Funato, K. *et al.* (2014) ‘Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation’, *Science*, 346(6216), pp. 1529–1533. doi: 10.1126/science.1253799.

Gangaraju, V. K. and Bartholomew, B. (2007) ‘Mechanisms of ATP dependent chromatin remodeling’, *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 618(1–2), pp. 3–17. doi: 10.1016/j.mrfmmm.2006.08.015.

Garcia-Ramirez, M., Dong, F. and Ausio, J. (1992) ‘Role of the histone “tails” in the folding of oligonucleosomes depleted of histone H1’, *Journal of Biological Chemistry*, 267(27), pp. 19587–19595.

Garcia-Ramirez, M., Rocchini, C. and Ausio, J. (1995) ‘Modulation of chromatin folding by histone acetylation’, *Journal of Biological Chemistry*, pp. 17923–17928. doi: 10.1074/jbc.270.30.17923.

Gatchalian, J. *et al.* (2018) ‘A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells’, *Nature Communications*. Springer US, 9(1), p. 5139. doi: 10.1038/s41467-018-07528-9.

Gehani, S. S. *et al.* (2010) ‘Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation’, *Molecular Cell*. Elsevier

Inc., 39(6), pp. 886–900. doi: 10.1016/j.molcel.2010.08.020.

Gellon, G. and McGinnis, W. (1998) ‘Shaping animal body plans in development and evolution by modulation of Hox expression patterns’, *BioEssays*, 20(2), pp. 116–125. doi: 10.1002/(SICI)1521-1878(199802)20:2<116::AID-BIES4>3.0.CO;2-R.

Gerstein, M. B. *et al.* (2007) ‘What is a gene, post-ENCODE? History and updated definition’, *Genome Research*, 17(6), pp. 669–681. doi: 10.1101/gr.6339607.

Ghavi-Helm, Y. *et al.* (2014) ‘Enhancer loops appear stable during development and are associated with paused polymerase’, *Nature*. Nature Publishing Group, 512(1), pp. 96–100. doi: 10.1038/nature13417.

Gilmour, D. S. and Lis, J. T. (1986) ‘RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells.’, *Molecular and Cellular Biology*. Elsevier, 6(11), pp. 3984–3989. doi: 10.1128/MCB.6.11.3984.

Gkikopoulos, T. *et al.* (2011) ‘A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization’, *Science*, 333(6050), pp. 1758–1760. doi: 10.1126/science.1206097.

Goldman, S. R., Ebright, R. H. and Nickels, B. E. (2009) ‘Direct Detection of Abortive RNA Transcripts in Vivo’, *Science*, 324(5929), pp. 927–928. doi: 10.1126/science.1169237.

Grunstein, M. (1997) ‘Histone acetylation in chromatin structure and transcription’, *Nature*, 389(6649), pp. 349–352. doi: 10.1038/38664.

Hah, N. *et al.* (2011) ‘A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells’, *Cell*. Elsevier Inc., 145(4), pp. 622–634. doi: 10.1016/j.cell.2011.03.042.

Hamiche, A. *et al.* (1999) ‘ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF’, *Cell*, 97(7), pp. 833–842. doi: 10.1016/S0092-8674(00)80796-5.

Han, F. *et al.* (2019) ‘GLTSCR1 Negatively Regulates BRD4-Dependent Transcription Elongation and Inhibits CRC Metastasis’, *Advanced Science*, 1901114. doi: 10.1002/advs.201901114.

Hao, N., Shearwin, K. E. and Dodd, I. B. (2019) ‘Positive and Negative Control of Enhancer-Promoter Interactions by Other DNA Loops Generates Specificity and Tunability’, *Cell Reports*. Elsevier Company., 26(9), pp. 2419–2433.e3. doi: 10.1016/j.celrep.2019.02.002.

Hathaway, N. A. *et al.* (2012) ‘Dynamics and memory of heterochromatin in living cells’, *Cell*. Elsevier Inc., 149(7), pp. 1447–1460. doi: 10.1016/j.cell.2012.03.052.

Hayes, J. J., Clark, D. J. and Wolffe, A. P. (1991) ‘Histone contributions to the

structure of DNA in the nucleosome.’, *Proceedings of the National Academy of Sciences*, 88(15), pp. 6829–6833. doi: 10.1073/pnas.88.15.6829.

Hebbes, T. R. *et al.* (1994) ‘Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain.’, *The EMBO Journal*, 13(8), pp. 1823–1830. doi: 10.1002/j.1460-2075.1994.tb06451.x.

Hein, M. Y. *et al.* (2015) ‘A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances’, *Cell*, 163(3), pp. 712–723. doi: 10.1016/j.cell.2015.09.053.

Heintzman, N. D. *et al.* (2007) ‘Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome’, *Nature Genetics*, 39(3), pp. 311–318. doi: 10.1038/ng1966.

Heintzman, N. D. *et al.* (2009) ‘Histone modifications at human enhancers reflect global cell-type-specific gene expression’, *Nature*. Nature Publishing Group, 459(7243), pp. 108–112. doi: 10.1038/nature07829.

Hendzel, M. J. *et al.* (1997) ‘Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation’, *Chromosoma*, 106(6), pp. 348–360. doi: 10.1007/s004120050256.

Henikoff, S. and Gready, J. M. (2016) ‘Epigenetics, cellular memory and gene regulation’, *Current Biology*. Elsevier, 26(14), pp. R644–R648. doi: 10.1016/j.cub.2016.06.011.

Henriques, T. *et al.* (2013) ‘Stable pausing by rna polymerase II provides an opportunity to target and integrate regulatory signals’, *Molecular Cell*. Elsevier Inc., 52(4), pp. 517–528. doi: 10.1016/j.molcel.2013.10.001.

Henriques, T. *et al.* (2018) ‘Widespread transcriptional pausing and elongation control at enhancers’, *Genes and Development*, 32(1), pp. 26–41. doi: 10.1101/gad.309351.117.

Hershey, A. D. (1952) ‘INDEPENDENT FUNCTIONS OF VIRAL PROTEIN AND NUCLEIC ACID IN GROWTH OF BACTERIOPHAGE’, *The Journal of General Physiology*, 36(1), pp. 39–56. doi: 10.1085/jgp.36.1.39.

Herz, H. M. *et al.* (2014) ‘Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling’, *Science*, 345(6200), pp. 1065–1070. doi: 10.1126/science.1255104.

Hilton, I. B. *et al.* (2015) ‘Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers.’, *Nature biotechnology*, 33(5), pp. 510–7. doi: 10.1038/nbt.3199.

Hirose, Y. and Manley, J. L. (2000) ‘RNA polymerase II and the integration of

nuclear events', *Genes and Development*, 14(12), pp. 1415–1429. doi: 10.1101/gad.14.12.1415.

Hnisz, D. *et al.* (2016) 'Activation of proto-oncogenes by disruption of chromosome neighborhoods', *Science*, 351(6280), pp. 1454–1458. doi: 10.1126/science.aad9024.

Ho, L. *et al.* (2009) 'An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency.', *Proceedings of the National Academy of Sciences of the United States of America*, 106(13), pp. 5181–6. doi: 10.1073/pnas.0812889106.

Ho, L. and Crabtree, G. R. (2010) 'Chromatin remodelling during development', *Nature*, 463(7280), pp. 474–484. doi: 10.1038/nature08911.

Ho, P. J., Lloyd, S. M. and Bao, X. (2019) 'Unwinding chromatin at the right places: How BAF is targeted to specific genomic locations during development', *Development (Cambridge)*, 146(19), pp. 1–14. doi: 10.1242/dev.178780.

Hohmann, A. F. *et al.* (2016) 'Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition', *Nature Chemical Biology*. Nature Publishing Group, 12(9), pp. 672–679. doi: 10.1038/nchembio.2115.

Hota, S. K. *et al.* (2019) 'Dynamic BAF chromatin remodeling complex subunit inclusion promotes temporally distinct gene expression programs in cardiogenesis', *Development (Cambridge)*, 146(19). doi: 10.1242/dev.174086.

Howe, L. *et al.* (1999) 'Histone Acetyltransferase Complexes and Their Link to Transcription', *Critical Reviews □ in Eukaryotic Gene Expression*, 9(3–4), pp. 231–243. doi: 10.1615/CritRevEukarGeneExpr.v9.i3-4.80.

Hsieh, C. L. *et al.* (2014) 'Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation', *Proceedings of the National Academy of Sciences of the United States of America*, 111(20), pp. 7319–7324. doi: 10.1073/pnas.1324151111.

Huang, R. C. (1962) 'Proc., 20, 85 (1961). HISTONE, A SUPPRESSOR OF CHROMOSOMAL RNA SYNTHESIS*', pp. 1216–1222.

Iborra, F. J. *et al.* (1996) 'Active RNA polymerases are localized within discrete transcription "factories" in human nuclei', *Journal of Cell Science*, 109(6), pp. 1427–1436.

Ingham, P. W. (1981) 'Trithorax: A new homoeotic mutation of *Drosophila melanogaster* - II. The role of *trx*⁺ after embryogenesis', *Wilhelm Roux's Archives of Developmental Biology*, 190(6), pp. 365–369. doi: 10.1007/BF00863275.

Ingham, P. and Whittle, R. (1980) 'Trithorax: A new homoeotic mutation of *Drosophila melanogaster* causing transformations of abdominal and thoracic imaginal segments - I. Putative role during embryogenesis', *MGG Molecular & General*

Genetics, 179(3), pp. 607–614. doi: 10.1007/BF00271751.

Jefimov, K. *et al.* (2018) ‘The GBAF chromatin remodeling complex binds H3K27ac and mediates enhancer transcription’, *bioRxiv*. Available at: <http://biorxiv.org/content/early/2018/10/17/445148.abstract>.

Jensen, T. H., Jacquier, A. and Libri, D. (2013) ‘Dealing with Pervasive Transcription’, *Molecular Cell*. Cell Press, 52(4), pp. 473–484. doi: 10.1016/J.MOLCEL.2013.10.032.

Jenuwein, T. (2001) ‘Translating the Histone Code’, *Science*, 293(5532), pp. 1074–1080. doi: 10.1126/science.1063127.

Ji, X. *et al.* (2015) ‘Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions.’, *Proceedings of the National Academy of Sciences of the United States of America*, 112(12), pp. 3841–6. doi: 10.1073/pnas.1502971112.

Jin, F. *et al.* (2011) ‘Enhancers: Multi-dimensional signal integrators’, *Transcription*, 2(5), pp. 232–236. doi: 10.4161/trns.2.5.17712.

Jin, F. *et al.* (2013) ‘A high-resolution map of the three-dimensional chromatin interactome in human cells’, *Nature*. Nature Publishing Group, 503(7475), pp. 290–294. doi: 10.1038/nature12644.

Jin, Q. *et al.* (2011) ‘Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation’, *EMBO Journal*. Nature Publishing Group, 30(2), pp. 249–262. doi: 10.1038/emboj.2010.318.

Kadam, S. and Emerson, B. M. (2003) ‘Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes’, *Molecular Cell*, 11(2), pp. 377–389. doi: 10.1016/S1097-2765(03)00034-0.

Kadoch, C. *et al.* (2013) ‘Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy’, *Nature Genetics*. Nature Publishing Group, 45(6), pp. 592–601. doi: 10.1038/ng.2628.

Kadoch, C. and Crabtree, G. R. (2015) ‘Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics’, *Science Advances*, 1(5), pp. 1–18. doi: 10.1126/sciadv.1500447.

Karmodiya, K. *et al.* (2012) ‘H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells’, *BMC Genomics*, 13(1). doi: 10.1186/1471-2164-13-424.

Kassabov, S. R. *et al.* (2003) ‘SWI/SNF unwraps, slides, and rewraps the nucleosome’, *Molecular Cell*, 11(2), pp. 391–403. doi: 10.1016/S1097-2765(03)00039-X.

Kaustov, L. *et al.* (2011) ‘Recognition and specificity determinants of the human Cbx chromodomains’, *Journal of Biological Chemistry*, 286(1), pp. 521–529. doi: 10.1074/jbc.M110.191411.

Khalil, A. M. *et al.* (2009) ‘Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression’, *Proceedings of the National Academy of Sciences of the United States of America*, 106(28), pp. 11667–11672. doi: 10.1073/pnas.0904715106.

Kim, T. K. *et al.* (2010) ‘Widespread transcription at neuronal activity-regulated enhancers’, *Nature*. Nature Publishing Group, 465(7295), pp. 182–187. doi: 10.1038/nature09033.

Kim, Y. *et al.* (2019) ‘Human cohesin compacts DNA by loop extrusion’, *Science*, 366(6471), pp. 1345–1349. doi: 10.1126/science.aaz4475.

Kingston, R. E. and Tamkun, J. W. (2014) ‘Transcriptional Regulation by Trithorax-Group Proteins’, *Cold Spring Harbor Perspectives in Biology*, 6(10), pp. a019349–a019349. doi: 10.1101/cshperspect.a019349.

Kornberg, R. D. and Thonmas, J. O. (1974) ‘Chromatin Structure: Oligomers of the Histones’, *Science*, 184(4139), pp. 865–868. doi: 10.1126/science.184.4139.865.

Kouzarides, T. (2007) ‘Chromatin Modifications and Their Function’, *Cell*, 128(4), pp. 693–705. doi: 10.1016/j.cell.2007.02.005.

Krajewski, W. A. and Becker, P. B. (1998) ‘Reconstitution of hyperacetylated, DNase I-sensitive chromatin characterized by high conformational flexibility of nucleosomal DNA’, *Proceedings of the National Academy of Sciences of the United States of America*, 95(4), pp. 1540–1545. doi: 10.1073/pnas.95.4.1540.

Kurat, C. F. *et al.* (2014) ‘Regulation of histone gene transcription in yeast’, *Cellular and Molecular Life Sciences*, 71(4), pp. 599–613. doi: 10.1007/s00018-013-1443-9.

Lachner, M. *et al.* (2001) ‘Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins’, *Nature*, 410(6824), pp. 116–120. doi: 10.1038/35065132.

Lai, F. *et al.* (2013) ‘Activating RNAs associate with Mediator to enhance chromatin architecture and transcription’, *Nature*, 494(7438), pp. 497–501. doi: 10.1038/nature11884.

Lai, W. K. M. and Pugh, B. F. (2017) ‘Understanding nucleosome dynamics and their links to gene expression and DNA replication’, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(9), pp. 548–562. doi: 10.1038/nrm.2017.47.

Lam, M. T. Y. *et al.* (2014) ‘Enhancer RNAs and regulated transcriptional programs’, *Trends in Biochemical Sciences*. Elsevier Ltd, 39(4), pp. 170–182. doi: 10.1016/j.tibs.2014.02.007.

- Längst, G. *et al.* (1999) 'Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer', *Cell*, 97(7), pp. 843–852. doi: 10.1016/S0092-8674(00)80797-7.
- Lau, P. N. I. and Cheung, P. (2011a) 'Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing', *Proceedings of the National Academy of Sciences of the United States of America*, 108(7), pp. 2801–2806. doi: 10.1073/pnas.1012798108.
- Lau, P. N. I. and Cheung, P. (2011b) 'Unlocking polycomb silencing through histone H3 phosphorylation', *Cell Cycle*, 10(10), pp. 1514–1515. doi: 10.4161/cc.10.10.15433.
- Lee, M. T., Bonneau, A. R. and Giraldez, A. J. (2014) 'Zygotic Genome Activation During the Maternal-to-Zygotic Transition', *Annual Review of Cell and Developmental Biology*, 30(1), pp. 581–613. doi: 10.1146/annurev-cellbio-100913-013027.
- Lessard, J. *et al.* (2007) 'Article An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development', pp. 201–215. doi: 10.1016/j.neuron.2007.06.019.
- Lewis, E. B. (1978) 'A gene complex controlling segmentation in *Drosophila*', *Nature*, 276(5688), pp. 565–570. doi: 10.1038/276565a0.
- Li, B., Carey, M. and Workman, J. L. (2007) 'The Role of Chromatin during Transcription', *Cell*, 128(4), pp. 707–719. doi: 10.1016/j.cell.2007.01.015.
- Li, Q. *et al.* (2008) 'Acetylation of Histone H3 Lysine 56 Regulates Replication-Coupled Nucleosome Assembly', *Cell*, 134(2), pp. 244–255. doi: 10.1016/j.cell.2008.06.018.
- Li, Q., Burgess, R. and Zhang, Z. (2012) 'All roads lead to chromatin: Multiple pathways for histone deposition', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. Elsevier B.V., 1819(3–4), pp. 238–246. doi: 10.1016/j.bbagr.2011.06.013.
- Li, W. *et al.* (2013) 'Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation.', *Nature*, 498(7455), pp. 516–20. doi: 10.1038/nature12210.
- Li, Y. *et al.* (2014) 'AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation', *Cell*. Elsevier Inc., 159(3), pp. 558–571. doi: 10.1016/j.cell.2014.09.049.
- Lieberman-Aiden, E. *et al.* (2009) 'Comprehensive mapping of long-range interactions reveals folding principles of the human genome', *Science*, 326(5950), pp. 289–293. doi: 10.1126/science.1181369.

-
- Litt, M. D. *et al.* (2001) 'Correlation between histone lysine methylation and developmental changes at the chicken β -globin locus', *Science*, 293(5539), pp. 2453–2455. doi: 10.1126/science.1064413.
- Long, H. K., Prescott, S. L. and Wysocka, J. (2016) 'Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution', *Cell*. Elsevier, 167(5), pp. 1170–1187. doi: 10.1016/j.cell.2016.09.018.
- Lopez, A. J. and Wood, M. A. (2015) 'Role of nucleosome remodeling in neurodevelopmental and intellectual disability disorders', *Frontiers in Behavioral Neuroscience*, 9(APR), pp. 1–10. doi: 10.3389/fnbeh.2015.00100.
- Lorch, Y. and Kornberg, R. D. (2015) 'Chromatin-remodeling and the initiation of transcription', *Quarterly Reviews of Biophysics*, 48(4), pp. 465–470. doi: 10.1017/S0033583515000116.
- Lorch, Y. and Kornberg, R. D. (2017) 'Chromatin-remodeling for transcription', *Quarterly Reviews of Biophysics*, 50, pp. 1–15. doi: 10.1017/S003358351700004X.
- Lorch, Y., Zhang, M. and Kornberg, R. D. (1999) 'Histone octamer transfer by a chromatin-remodeling complex', *Cell*, 96(3), pp. 389–392. doi: 10.1016/S0092-8674(00)80551-6.
- Louder, R. K. *et al.* (2016) 'Structure of promoter-bound TFIID and model of human pre-initiation complex assembly', *Nature*. Nature Publishing Group, 531(7596), pp. 604–609. doi: 10.1038/nature17394.
- Luco, R. F. *et al.* (2011) 'Epigenetics in alternative pre-mRNA splicing', *Cell*. Elsevier Inc., 144(1), pp. 16–26. doi: 10.1016/j.cell.2010.11.056.
- Luger, K. *et al.* (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature*, 389(6648), pp. 251–260. doi: 10.1038/38444.
- Luo, Z., Lin, C. and Shilatifard, A. (2012) 'The super elongation complex (SEC) family in transcriptional control', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 13(9), pp. 543–547. doi: 10.1038/nrm3417.
- Lupiáñez, D. G. *et al.* (2015) 'Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions', *Cell*, 161(5), pp. 1012–1025. doi: 10.1016/j.cell.2015.04.004.
- Maniatis, T. and Reed, R. (2002) 'An extensive network of coupling among gene expression machines', *Nature*, 416(6880), pp. 499–506. doi: 10.1038/416499a.
- Marzluff, W. F. (2005) 'Metazoan replication-dependent histone mRNAs: A distinct set of RNA polymerase II transcripts', *Current Opinion in Cell Biology*, 17(3), pp. 274–280. doi: 10.1016/j.ceb.2005.04.010.
- Mashtalir, N. *et al.* (2018) 'Modular Organization and Assembly of SWI/SNF Family

Chromatin Remodeling Complexes', *Cell*. Elsevier Inc., pp. 1–17. doi: 10.1016/j.cell.2018.09.032.

Masliah-Planchon, J. *et al.* (2015) 'SWI/SNF Chromatin Remodeling and Human Malignancies', *Annual Review of Pathology: Mechanisms of Disease*, 10(1), pp. 145–171. doi: 10.1146/annurev-pathol-012414-040445.

Mayer, A. *et al.* (2010) 'Uniform transitions of the general RNA polymerase II transcription complex', *Nature Structural and Molecular Biology*. Nature Publishing Group, 17(10), pp. 1272–1278. doi: 10.1038/nsmb.1903.

Mayr, E. (1982) *The growth of biological thought: Diversity, evolution, and inheritance*. Harvard University Press.

McKittrick, E. *et al.* (2004) 'Histone H3.3 is enriched in covalent modifications associated with active chromatin', *Proceedings of the National Academy of Sciences of the United States of America*, 101(6), pp. 1525–1530. doi: 10.1073/pnas.0308092100.

Melgar, M. F., Collins, F. S. and Sethupathy, P. (2011) 'Discovery of active enhancers through bidirectional expression of short transcripts', *Genome Biology*, 12(11). doi: 10.1186/gb-2011-12-11-r113.

Melo, C. A. *et al.* (2013) 'ERNA's Are Required for p53-Dependent Enhancer Activity and Gene Transcription', *Molecular Cell*, 49(3), pp. 524–535. doi: 10.1016/j.molcel.2012.11.021.

Mersfelder, E. L. and Parthun, M. R. (2006) 'The tale beyond the tail: Histone core domain modifications and the regulation of chromatin structure', *Nucleic Acids Research*, 34(9), pp. 2653–2662. doi: 10.1093/nar/gkl338.

Michel, B. C. *et al.* (2018) 'A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation', *Nature Cell Biology*. doi: 10.1038/s41556-018-0221-1.

Middeljans, E. *et al.* (2012) 'SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes', *PLoS ONE*, 7(3). doi: 10.1371/journal.pone.0033834.

Miescher-Rüsch, F. (1871) *Ueber die chemische Zusammensetzung der Eiterzellen*.

Mikhaylichenko, O. *et al.* (2018) 'The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription', *Genes and Development*. doi: 10.1101/gad.308619.117.

Mizuguchi, G. *et al.* (2004) 'ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex', *Science*, 303(5656), pp. 343–348. doi: 10.1126/science.1090701.

-
- Mohrmann, L. and Verrijzer, C. P. (2005) 'Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes', *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1681(2–3), pp. 59–73. doi: 10.1016/j.bbaexp.2004.10.005.
- Moon, K. J. *et al.* (2005) 'The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription', *Molecular Cell*, 19(4), pp. 523–534. doi: 10.1016/j.molcel.2005.06.027.
- Morange, M. (2011) 'The attempt of Nikolai Koltzoff (Koltsov) to link genetics, embryology and physical chemistry', *Journal of Biosciences*, 36(2), pp. 211–214. doi: 10.1007/s12038-011-9075-4.
- Morinière, J. *et al.* (2009) 'Cooperative binding of two acetylation marks on a histone tail by a single bromodomain', *Nature*, 461(7264), pp. 664–668. doi: 10.1038/nature08397.
- Mousavi, K. *et al.* (2013) 'Article eRNAs Promote Transcription by Establishing Chromatin Accessibility at Defined Genomic Loci', *Molecular Cell*. Elsevier Inc., 51(5), pp. 606–617. doi: 10.1016/j.molcel.2013.07.022.
- Muchardt, C. and Yaniv, M. (2001) 'When the SWI/SNF complex remodels ... the cell cycle', *Oncogene*, 20(24), pp. 3067–3075. doi: 10.1038/sj.onc.1204331.
- Mujtaba, S., Zeng, L. and Zhou, M. M. (2007) 'Structure and acetyl-lysine recognition of the bromodomain', *Oncogene*, 26(37), pp. 5521–5527. doi: 10.1038/sj.onc.1210618.
- Murakami, A. *et al.* (2017) 'Context-dependent role for chromatin remodeling component PBRM1/BAF180 in clear cell renal cell carcinoma', *Oncogenesis*, 6(1). doi: 10.1038/oncsis.2016.89.
- Muse, G. W. *et al.* (2007) 'RNA polymerase is poised for activation across the genome', *Nature Genetics*, 39(12), pp. 1507–1511. doi: 10.1038/ng.2007.21.
- Nakayama, R. T. *et al.* (2017) 'SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters', *Nature Genetics*, (August). doi: 10.1038/ng.3958.
- Narendra, V. *et al.* (2015) 'Transcription. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation.', *Science (New York, N.Y.)*, 347(6225), pp. 1017–21. doi: 10.1126/science.1262088.
- Ng, R. K. and Gurdon, J. B. (2008) 'Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription', *Nature Cell Biology*, 10(1), pp. 102–109. doi: 10.1038/ncb1674.
- Nora, E. P. *et al.* (2012) 'Spatial partitioning of the regulatory landscape of the X-inactivation centre', *Nature*, 485(7398), pp. 381–385. doi: 10.1038/nature11049.

- Oliva, R. *et al.* (1990) 'Histone hyperacetylation can induce unfolding of the nucleosome core particle', *Nucleic Acids Research*, 18(9), pp. 2739–2747. doi: 10.1093/nar/18.9.2739.
- Ong, C. T. and Corces, V. G. (2011) 'Enhancer function: New insights into the regulation of tissue-specific gene expression', *Nature Reviews Genetics*. Nature Publishing Group, 12(4), pp. 283–293. doi: 10.1038/nrg2957.
- Ong, C. T. and Corces, V. G. (2012) 'Enhancers: Emerging roles in cell fate specification', *EMBO Reports*. Nature Publishing Group, 13(5), pp. 423–430. doi: 10.1038/embor.2012.52.
- Ørom, U. A. *et al.* (2010) 'Long noncoding RNAs with enhancer-like function in human cells', *Cell*. Elsevier Ltd, 143(1), pp. 46–58. doi: 10.1016/j.cell.2010.09.001.
- Pal, S. *et al.* (2004) 'Human SWI/SNF-Associated PRMT5 Methylates Histone H3 Arginine 8 and Negatively Regulates Expression of ST7 and NM23 Tumor Suppressor Genes', *Molecular and Cellular Biology*, 24(21), pp. 9630–9645. doi: 10.1128/MCB.24.21.9630-9645.2004.
- Palikyras, S. and Papantonis, A. (2019) 'Modes of phase separation affecting chromatin regulation', *Open biology*, 9(10), p. 190167. doi: 10.1098/rsob.190167.
- Pandit, S., Wang, D. and Fu, X. D. (2008) 'Functional integration of transcriptional and RNA processing machineries', *Current Opinion in Cell Biology*, 20(3), pp. 260–265. doi: 10.1016/j.ceb.2008.03.001.
- Parelho, V. *et al.* (2008) 'Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms', *Cell*, 132(3), pp. 422–433. doi: 10.1016/j.cell.2008.01.011.
- Park, S. *et al.* (2017) 'Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation', *Proceedings of the National Academy of Sciences*, 114(27), pp. E5335–E5342. doi: 10.1073/pnas.1703105114.
- Paul, S. and Bartholomew, B. (2018) 'Regulation of ATP-dependent chromatin remodelers: accelerators/brakes, anchors and sensors', *Biochemical Society Transactions*, (November), p. BST20180043. doi: 10.1042/BST20180043.
- Peng, J. *et al.* (1998) 'Identification of multiple cyclin subunits of human P-TEFb', *Genes and Development*, 12(5), pp. 755–762. doi: 10.1101/gad.12.5.755.
- Pengelly, A. R. *et al.* (2013) 'A histone mutant reproduces the phenotype caused by loss of histone-modifying factor polycomb', *Science*, 339(6120), pp. 698–699. doi: 10.1126/science.1231382.
- Perales, R. and Bentley, D. (2009) "'Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions', *Molecular Cell*. Elsevier Ltd, 36(2), pp. 178–191. doi: 10.1016/j.molcel.2009.09.018.

-
- Peterson, C. L. and Laniel, M. (2004) 'Peterson e Laniel. 2004. Histones and histone modifications.PDF', 14(14), pp. 546–551. doi: <https://doi.org/10.1016/j.cub.2004.07.007>.
- Phatnani, H. P. and Greenleaf, A. L. (2006) 'Phosphorylation and functions of the RNA polymerase II CTD', *Genes and Development*, 20(21), pp. 2922–2936. doi: 10.1101/gad.1477006.
- Phelan, M. L. *et al.* (1999) 'Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits', *Molecular Cell*, 3(2), pp. 247–253. doi: 10.1016/S1097-2765(00)80315-9.
- Pope, B. D. *et al.* (2014) 'Topologically associating domains are stable units of replication-timing regulation', *Nature*, 515(7527), pp. 402–405. doi: 10.1038/nature13986.
- Pradeepa, M. M. (2017) 'Causal role of histone acetylations in enhancer function', *Transcription*, 8(1), pp. 40–47. doi: 10.1080/21541264.2016.1253529.
- Prieto, E. I. and Maeshima, K. (2019) 'Dynamic chromatin organization in the cell', *Essays in Biochemistry*, 63(1), pp. 133–145. doi: 10.1042/EBC20180054.
- Proudfoot, N. J., Furger, A. and Dye, M. J. (2002) 'Integrating mRNA Processing with Transcription', *Cell*, 108(4), pp. 501–512. doi: 10.1016/S0092-8674(02)00617-7.
- Quack, M. *et al.* (1998) 'The role of the T-box for the function of the vitamin D receptor on different types of response elements', *Nucleic Acids Research*, 26(23), pp. 5372–5378. doi: 10.1093/nar/26.23.5372.
- Raab, J. R. *et al.* (2017) 'Co-regulation of transcription by BRG1 and BRM, two mutually exclusive SWI/SNF ATPase subunits', *Epigenetics and Chromatin*. BioMed Central, 10(1), pp. 1–15. doi: 10.1186/s13072-017-0167-8.
- Rada-Iglesias, A. *et al.* (2011) 'A unique chromatin signature uncovers early developmental enhancers in humans.', *Nature*. Nature Publishing Group, 470(7333), pp. 279–83. doi: 10.1038/nature09692.
- Radman-Livaja, M. *et al.* (2011) 'Patterns and mechanisms of Ancestral Histone protein inheritance in Budding yeast', *PLoS Biology*, 9(6). doi: 10.1371/journal.pbio.1001075.
- Ragunathan, K., Jih, G. and Moazed, D. (2015) 'Epigenetic inheritance uncoupled from sequence-specific recruitment', *Science*, 348(6230). doi: 10.1126/science.1258699.
- Rahman, S. *et al.* (2011) 'The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3.', *Molecular and cellular biology*, 31(13), pp. 2641–52. doi: 10.1128/MCB.01341-10.

-
- Raisner, R. *et al.* (2018) 'Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation', *Cell Reports*. Elsevier Company., 24(7), pp. 1722–1729. doi: 10.1016/j.celrep.2018.07.041.
- Rajagopal, N. *et al.* (2014) 'Distinct and Predictive Histone Lysine Acetylation Patterns at Promoters, Enhancers, and Gene Bodies', *G3 & Genes|Genomes|Genetics*, 4(11), pp. 2051–2063. doi: 10.1534/g3.114.013565.
- Rao, S. S. P. *et al.* (2014) 'A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping', *Cell*. Elsevier Inc., 159(7), pp. 1665–1680. doi: 10.1016/j.cell.2014.11.021.
- Renner, D. B. *et al.* (2001) 'A Highly Purified RNA Polymerase II Elongation Control System', *Journal of Biological Chemistry*, 276(45), pp. 42601–42609. doi: 10.1074/jbc.M104967200.
- Rennie, S. *et al.* (2018) 'Transcription start site analysis reveals widespread divergent transcription in *D. melanogaster* and core promoter-encoded enhancer activities', *Nucleic Acids Research*. Oxford University Press, (April), pp. 1–15. doi: 10.1093/nar/gky244.
- Reverón-Gómez, N. *et al.* (2018) 'Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication', *Molecular Cell*, 72(2), pp. 239–249.e5. doi: 10.1016/j.molcel.2018.08.010.
- Richmond, T. J. and Davey, C. A. (2003) 'The structure of DNA in the nucleosome core', *Nature*, 423(6936), pp. 145–150. doi: 10.1038/nature01595.
- Robinson, P. J. J. *et al.* (2008) '30 nm Chromatin Fibre Decompaction Requires both H4-K16 Acetylation and Linker Histone Eviction', *Journal of Molecular Biology*, 381(4), pp. 816–825. doi: 10.1016/j.jmb.2008.04.050.
- Roh, T., Cuddapah, S. and Zhao, K. (2005) 'Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping -- Roh *et al.* 19 (5): 542 -- Genes and Development', pp. 542–552. doi: 10.1101/gad.1272505.lation.
- Roh, T. Y. *et al.* (2007) 'Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns', *Genome Research*, 17(1), pp. 74–81. doi: 10.1101/gr.5767907.
- Rougvie, A. E. and Lis, J. T. (1988) 'The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged', *Cell*, 54(6), pp. 795–804. doi: 10.1016/S0092-8674(88)91087-2.
- Rowley, M. J. and Corces, V. G. (2018) 'Organizational principles of 3D genome architecture', *Nature Reviews Genetics*. Springer US, 13, p. 1. doi: 10.1038/s41576-018-0060-8.
- Ruff, K. M. *et al.* (2018) 'Advances in Understanding Stimulus-Responsive Phase

Behavior of Intrinsically Disordered Protein Polymers', *Journal of Molecular Biology*. The Authors, 430(23), pp. 4619–4635. doi: 10.1016/j.jmb.2018.06.031.

Russo, V. E. A., Martienssen, R. A. and Riggs, A. D. (1996) *Epigenetic mechanisms of gene regulation*. Cold Spring Harbor Laboratory Press.

Saha, A., Wittmeyer, J. and Cairns, B. R. (2006) 'Chromatin remodelling: The industrial revolution of DNA around histones', *Nature Reviews Molecular Cell Biology*, 7(6), pp. 437–447. doi: 10.1038/nrm1945.

Saksouk, N., Simboeck, E. and Déjardin, J. (2015) 'Constitutive heterochromatin formation and transcription in mammals', *Epigenetics and Chromatin*, 8(1), pp. 1–17. doi: 10.1186/1756-8935-8-3.

Sanborn, A. L. *et al.* (2015) 'Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes', *Proceedings of the National Academy of Sciences of the United States of America*, 112(47), pp. E6456–E6465. doi: 10.1073/pnas.1518552112.

Santenard, A. *et al.* (2010) 'Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3', *Nature Cell Biology*. Nature Publishing Group, 12(9), pp. 853–862. doi: 10.1038/ncb2089.

Sanyal, A. *et al.* (2012) 'The long-range interaction landscape of gene promoters', *Nature*. Nature Publishing Group, 489(7414), pp. 109–113. doi: 10.1038/nature11279.

Sawicka, A. *et al.* (2014) 'H3S28 phosphorylation is a hallmark of the transcriptional response to cellular stress', *Genome Research*, 24(11), pp. 1808–1820. doi: 10.1101/gr.176255.114.

Sawicka, A. and Seiser, C. (2012) 'Histone H3 phosphorylation - A versatile chromatin modification for different occasions', *Biochimie*. Elsevier, 94(11), pp. 2193–2201. doi: 10.1016/j.biochi.2012.04.018.

Schaukowitch, K. *et al.* (2014) 'Enhancer RNA facilitates NELF release from immediate early genes', *Molecular Cell*. Elsevier Inc., 56(1), pp. 29–42. doi: 10.1016/j.molcel.2014.08.023.

Schröder, S. *et al.* (2013) 'Acetylation of RNA Polymerase II Regulates Growth-Factor-Induced Gene Transcription in Mammalian Cells', *Molecular Cell*, 52(3), pp. 314–324. doi: 10.1016/j.molcel.2013.10.009.

Schwabish, M. A. and Struhl, K. (2007) 'The Swi/Snf Complex Is Important for Histone Eviction during Transcriptional Activation and RNA Polymerase II Elongation In Vivo', *Molecular and Cellular Biology*, 27(20), pp. 6987–6995. doi: 10.1128/mcb.00717-07.

Schwämmle, V. *et al.* (2014) 'Large scale analysis of co-existing post-translational

modifications in histone tails reveals global fine structure of cross-talk', *Molecular and Cellular Proteomics*, 13(7), pp. 1855–1865. doi: 10.1074/mcp.O113.036335.

Schwämmle, V. *et al.* (2016) 'Systems Level Analysis of Histone H3 Post-translational Modifications (PTMs) Reveals Features of PTM Crosstalk in Chromatin Regulation', *Molecular & Cellular Proteomics*, 15(8), pp. 2715–2729. doi: 10.1074/mcp.M115.054460.

Schwartz, B. E. and Ahmad, K. (2005) 'Transcriptional activation triggers deposition and removal of the histone variant H3.3', *Genes and Development*, 19(7), pp. 804–814. doi: 10.1101/gad.1259805.

Schwartzentruber, J. *et al.* (2012) 'Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma', *Nature*, 482(7384), pp. 226–231. doi: 10.1038/nature10833.

Schwer, B. and Shuman, S. (2011) 'Deciphering the RNA Polymerase II CTD Code in Fission Yeast', *Molecular Cell*. Elsevier Inc., 43(2), pp. 311–318. doi: 10.1016/j.molcel.2011.05.024.

Sen, P. *et al.* (2017) 'Loss of Snf5 Induces Formation of an Aberrant SWI/SNF Complex', *Cell Reports*. Elsevier Company., 18(9), pp. 2135–2147. doi: 10.1016/j.celrep.2017.02.017.

Sharma, A. B. *et al.* (2019) 'SURVEY AND SUMMARY Centromeric and ectopic assembly of CENP-A chromatin in health and cancer: old marks and new tracks', *Nucleic Acids Research*. Oxford University Press, 47(3), pp. 1051–1069. doi: 10.1093/nar/gky1298.

Sharma, N. and Nyborg, J. K. (2008) 'The coactivators CBP/p300 and the histone chaperone NAP1 promote transcription-independent nucleosome eviction at the HTLV-1 promoter', *Proceedings of the National Academy of Sciences of the United States of America*, 105(23), pp. 7959–7963. doi: 10.1073/pnas.0800534105.

Shi, J. *et al.* (2013) 'Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation', *Genes and Development*, 27(24), pp. 2648–2662. doi: 10.1101/gad.232710.113.

Shivaswamy, S. and Iyer, V. R. (2008) 'Stress-Dependent Dynamics of Global Chromatin Remodeling in Yeast: Dual Role for SWI/SNF in the Heat Shock Stress Response', *Molecular and Cellular Biology*, 28(7), pp. 2221–2234. doi: 10.1128/mcb.01659-07.

Shogren-Knaak, M. *et al.* (2006) 'Histone H4-K16 acetylation controls chromatin structure and protein interactions', *Science*, 311(5762), pp. 844–847. doi: 10.1126/science.1124000.

Simpson, D. A. *et al.* (2015) 'SWI/SNF complexes are required for full activation of

the DNA-damage response', *Oncotarget*, 6(2), pp. 732–745. doi: 10.18632/oncotarget.2715.

Singh, M., D'Silva, L. and Holak, T. A. (2006) 'DNA-binding properties of the recombinant high-mobility-group-like AT-hook-containing region from human BRG1 protein', *Biological Chemistry*, 387(10–11), pp. 1469–1478. doi: 10.1515/BC.2006.184.

Skiniotis, G., Moazed, D. and Walz, T. (2007) 'Acetylated histone tail peptides induce structural rearrangements in the RSC chromatin remodeling complex', *Journal of Biological Chemistry*, 282(29), pp. 20804–20808. doi: 10.1074/jbc.C700081200.

Slaughter, M. J. *et al.* (2018) 'PBRM1 bromodomains variably influence nucleosome interactions and cellular function', *Journal of Biological Chemistry*, 293(35), pp. 13592–13603. doi: 10.1074/jbc.RA118.003381.

Slifer, E. H. (1942) 'A mutant stock of *Drosophila* with extra sex-combs', *Journal of Experimental Zoology*, 90(1), pp. 31–40. doi: 10.1002/jez.1400900103.

Spitz, F. and Furlong, E. E. M. (2012) 'Transcription factors: from enhancer binding to developmental control.', *Nature reviews. Genetics*. Nature Publishing Group, 13(9), pp. 613–26. doi: 10.1038/nrg3207.

Stasevich, T. J. *et al.* (2014) 'Regulation of RNA polymerase II activation by histone acetylation in single living cells', *Nature*, 516(7530), pp. 272–275. doi: 10.1038/nature13714.

Stedman, Edgar and Stedman, Ellen (1947) 'The chemical nature and functions of the components of cell nuclei', in *Cold Spring Harbor Symposia on Quantitative Biology*. Cold Spring Harbor Laboratory Press, pp. 224–236.

Stedman, Edgar and Stedman, Ellen (1950) 'Cell specificity of histones', *Nature*, 166(4227), pp. 780–781. doi: 10.1038/166780a0.

van Steensel, B. and Furlong, E. E. M. (2019) 'The role of transcription in shaping the spatial organization of the genome', *Nature Reviews Molecular Cell Biology*. Springer US, 20(6), pp. 327–337. doi: 10.1038/s41580-019-0114-6.

Stergachis, A. B. *et al.* (2013) 'Developmental fate and cellular maturity encoded in human regulatory DNA landscapes', *Cell*. Elsevier Inc., 154(4), pp. 888–903. doi: 10.1016/j.cell.2013.07.020.

Stern, M., Jensen, R. and Herskowitz, I. (1984) 'Five SWI genes are required for expression of the HO gene in yeast', *Journal of Molecular Biology*, 178(4), pp. 853–868. doi: 10.1016/0022-2836(84)90315-2.

Sterner, D. E. and Berger, S. L. (2000) 'Acetylation of Histones and Transcription-Related Factors', *Microbiology and Molecular Biology Reviews*, 64(2), pp. 435–459.

doi: 10.1128/membr.64.2.435-459.2000.

Strahl, B. D. and Allis, C. D. (2000) 'The language of covalent histone modifications', *Nature*, 403(6765), pp. 41–45. doi: 10.1038/47412.

Sudarsanam, P. *et al.* (2000) 'Whole-genome expression analysis of Snf/Swi mutants of *Saccharomyces cerevisiae*', *Proceedings of the National Academy of Sciences of the United States of America*, 97(7), pp. 3364–3369. doi: 10.1073/pnas.97.7.3364.

Sutherland, H. and Bickmore, W. A. (2009) 'Transcription factories: Gene expression in unions?', *Nature Reviews Genetics*. Nature Publishing Group, 10(7), pp. 457–466. doi: 10.1038/nrg2592.

Sykes, S. M. *et al.* (2009) 'Acetylation of the DNA binding domain regulates transcription-independent apoptosis by p53', *Journal of Biological Chemistry*, 284(30), pp. 20197–20205. doi: 10.1074/jbc.M109.026096.

Tadros, W. and Lipshitz, H. D. (2009) 'The maternal-to-zygotic transition: A play in two acts', *Development*, 136(18), pp. 3033–3042. doi: 10.1242/dev.033183.

Talbert, P. B. and Henikoff, S. (2010) 'Histone variants ancient wrap artists of the epigenome', *Nature Reviews Molecular Cell Biology*, 11(4), pp. 264–275. doi: 10.1038/nrm2861.

Tamkun, J. W. *et al.* (1992) 'brahma: A regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2 SWI2', *Cell*, 68(3), pp. 561–572. doi: 10.1016/0092-8674(92)90191-E.

Tan, M. *et al.* (2011) 'Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification', *Cell*. Elsevier Inc., 146(6), pp. 1016–1028. doi: 10.1016/j.cell.2011.08.008.

Tardat, M. *et al.* (2015) 'Cbx2 targets PRC1 to constitutive heterochromatin in mouse zygotes in a parent-of-origin-dependent manner', *Molecular Cell*. Elsevier Inc., 58(1), pp. 157–171. doi: 10.1016/j.molcel.2015.02.013.

Tatavosian, R. *et al.* (2018) 'Live-cell single-molecule dynamics of PcG proteins imposed by the DIPG H3.3K27M mutation', *Nature Communications*. Springer US, 9(1), p. 2080. doi: 10.1038/s41467-018-04455-7.

Taverna, S. D. *et al.* (2007) 'How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers', *Nature Structural and Molecular Biology*, 14(11), pp. 1025–1040. doi: 10.1038/nsmb1338.

Theodoulou, N. H. *et al.* (2016) 'Discovery of I-BRD9, a Selective Cell Active Chemical Probe for Bromodomain Containing Protein 9 Inhibition', *Journal of Medicinal Chemistry*, 59(4), pp. 1425–1439. doi: 10.1021/acs.jmedchem.5b00256.

Thomas, J. O. (1984) 'The higher order structure of chromatin and histone H1',

Journal of Cell Science, pp. 1–20. doi: 10.1242/jcs.1984.supplement_1.1.

Thurman, R. E. *et al.* (2012) ‘The accessible chromatin landscape of the human genome’, *Nature*. Nature Publishing Group, 489(7414), pp. 75–82. doi: 10.1038/nature11232.

Tie, F. *et al.* (2009) ‘CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing’, *Development*, 136(18), pp. 3131–3141. doi: 10.1242/dev.037127.

Trotter, K. W. *et al.* (2008) ‘The HSA Domain of BRG1 Mediates Critical Interactions Required for Glucocorticoid Receptor-Dependent Transcriptional Activation In Vivo’, *Molecular and Cellular Biology*, 28(4), pp. 1413–1426. doi: 10.1128/mcb.01301-07.

Tse, C. *et al.* (1998) ‘Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III.’, *Molecular and cellular biology*, 18(8), pp. 4629–4638. doi: 10.1128/MCB.18.8.4629.

Tse, C. and Hansen, J. C. (1997) ‘Hybrid trypsinized nucleosomal arrays: Identification of multiple functional roles of the H2A/H2B and H3/H4 N-termini in chromatin fiber compaction’, *Biochemistry*, 36(38), pp. 11381–11388. doi: 10.1021/bi970801n.

Turner, B. (1993) ‘Decoding the nucleosome’, *Cell*, 75(1), pp. 5–8. doi: 10.1016/0092-8674(93)90673-e.

Tvardovskiy, A. *et al.* (2015) ‘Top-down and middle-down protein analysis reveals that intact and clipped human histones differ in post-translational modification patterns.’, *Molecular & cellular proteomics : MCP*, 14(12), pp. 3142–3153. doi: 10.1074/mcp.M115.048975.

Venkatesh, S. and Workman, J. L. (2015) ‘Histone exchange, chromatin structure and the regulation of transcription’, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 16(3), pp. 178–189. doi: 10.1038/nrm3941.

Vermeulen, M. *et al.* (2007) ‘Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4’, *Cell*, 131(1), pp. 58–69. doi: 10.1016/j.cell.2007.08.016.

Vermeulen, M. *et al.* (2010) ‘Quantitative Interaction Proteomics and Genome-wide Profiling of Epigenetic Histone Marks and Their Readers’, *Cell*. Elsevier Inc., 142(6), pp. 967–980. doi: 10.1016/j.cell.2010.08.020.

Vermeulen, M. (2012) *Identifying chromatin readers using a SILAC-based histone peptide pull-down approach*. 1st edn, *Methods in Enzymology*. 1st edn. Elsevier Inc. doi: 10.1016/B978-0-12-391940-3.00007-X.

Vian, L. *et al.* (2018) ‘The Energetics and Physiological Impact of Cohesin

-
- Extrusion', *Cell*, 173(5), pp. 1165-1178.e20. doi: 10.1016/j.cell.2018.03.072.
- Vicent, G. P. *et al.* (2009) 'Two chromatin remodeling activities cooperate during activation of hormone responsive promoters', *PLoS Genetics*, 5(7). doi: 10.1371/journal.pgen.1000567.
- Wang, D. *et al.* (2011) 'Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA', *Nature*. Nature Publishing Group, 474(7351), pp. 390-397. doi: 10.1038/nature10006.
- Wang, W. *et al.* (1996) 'Diversity and specialization of mammalian SWI / SNF complexes', *Genes & Development*, 4(17), pp. 2117-2130. doi: 10.1101/gad.10.17.2117.
- Watson, J. D. and Crick, F. H. C. (1953) 'Molecular Structures of Nucleic Acids', *Nature*, 171(1), pp. 737-738. doi: 10.1038/171737a0.
- Weake, V. M. and Workman, J. L. (2010) 'Inducible gene expression: Diverse regulatory mechanisms', *Nature Reviews Genetics*. Nature Publishing Group, 11(6), pp. 426-437. doi: 10.1038/nrg2781.
- Weber, C. M., Henikoff, J. G. and Henikoff, S. (2010) 'H2A.Z nucleosomes enriched over active genes are homotypic', *Nature Structural and Molecular Biology*. Nature Publishing Group, 17(12), pp. 1500-1507. doi: 10.1038/nsmb.1926.
- Wei, Z. *et al.* (2018) 'Vitamin D Switches BAF Complexes to Protect β Cells', *Cell*. Elsevier Inc., 173(5), pp. 1135-1149.e15. doi: 10.1016/j.cell.2018.04.013.
- Wendt, K. S. *et al.* (2008) 'Cohesin mediates transcriptional insulation by CCCTC-binding factor', *Nature*, 451(7180), pp. 796-801. doi: 10.1038/nature06634.
- Whitehouse, I. *et al.* (1999) 'Nucleosome mobilization catalysed by the yeast SWI/SNF complex', *Nature*, 400(6746), pp. 784-787. doi: 10.1038/23506.
- Widlund, H. R. *et al.* (2000) 'DNA sequence-dependent contributions of core histone tails to nucleosome stability: Differential effects of acetylation and proteolytic tail removal', *Biochemistry*, 39(13), pp. 3835-3841. doi: 10.1021/bi9919571.
- Wiles, E. T. and Selker, E. U. (2017) 'H3K27 methylation: a promiscuous repressive chromatin mark', *Current Opinion in Genetics and Development*. Elsevier Ltd, 43, pp. 31-37. doi: 10.1016/j.gde.2016.11.001.
- Williams, S. K. and Tyler, J. K. (2007) 'Transcriptional regulation by chromatin disassembly and reassembly', *Current Opinion in Genetics and Development*, 17(2), pp. 88-93. doi: 10.1016/j.gde.2007.02.001.
- Winger, J. *et al.* (2018) 'A twist defect mechanism for ATP-dependent translocation of nucleosomal DNA', *eLife*, 7, pp. 1-23. doi: 10.7554/eLife.34100.

-
- Witwicka, H. *et al.* (2019) ‘Calcineurin Broadly Regulates the Initiation of Skeletal Muscle-Specific Gene Expression by Binding Target Promoters and Facilitating the Interaction of the SWI/SNF Chromatin Remodeling Enzyme’, *Molecular and Cellular Biology*, 39(19). doi: 10.1128/mcb.00063-19.
- Wu, C. *et al.* (2007) ‘Systematic identification of SH3 domain-mediated human protein – protein interactions by peptide array target screening’, pp. 1775–1785. doi: 10.1002/pmic.200601006.
- Wu, G. *et al.* (2012) ‘Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas’, *Nature Genetics*, 44(3), pp. 251–253. doi: 10.1038/ng.1102.
- Wu, J. I., Lessard, J. and Crabtree, G. R. (2009) ‘Understanding the Words of Chromatin Regulation’, *Cell*, 136(2), pp. 200–206. doi: 10.1016/j.cell.2009.01.009.
- Xie, S. *et al.* (2011) ‘Increased acetylation in the DNA-binding domain of TR4 nuclear receptor by the coregulator ARA55 leads to suppression of TR4 transactivation’, *Journal of Biological Chemistry*, 286(24), pp. 21129–21136. doi: 10.1074/jbc.M110.208181.
- Yang, P. *et al.* (2005) ‘Polymorphisms in GLTSCR1 and ERCC2 are associated with the development of oligodendrogliomas.’, *Cancer*, 103(11), pp. 2363–72. doi: 10.1002/cncr.21028.
- Yang, X. *et al.* (2007) ‘Swi3p controls SWI/SNF assembly and ATP-dependent H2A-H2B displacement’, *Nature Structural and Molecular Biology*, 14(6), pp. 540–547. doi: 10.1038/nsmb1238.
- Yap, K. L. and Zhou, M. M. (2011) ‘Structure and mechanisms of lysine methylation recognition by the chromodomain in gene transcription’, *Biochemistry*, 50(12), pp. 1966–1980. doi: 10.1021/bi101885m.
- Ye, Y. *et al.* (2019) ‘Structure of the RSC complex bound to the nucleosome.’, *Science (New York, N.Y.)*, 0033(October), pp. 1–10. doi: 10.1126/science.aay0033.
- Young, R. S. *et al.* (2017) ‘Bidirectional transcription initiation marks accessible chromatin and is not specific to enhancers’, *Genome Biology*. *Genome Biology*, 18(1), pp. 1–11. doi: 10.1186/s13059-017-1379-8.
- Zeitlinger, J. *et al.* (2007) ‘RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo’, *Nature Genetics*, 39(12), pp. 1512–1516. doi: 10.1038/ng.2007.26.
- Zentner, G. E., Tesar, P. J. and Scacheri, P. C. (2011) ‘Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions’, *Genome Research*, 21(8), pp. 1273–1283. doi: 10.1101/gr.122382.111.
- Zhang, T., Cooper, S. and Brockdorff, N. (2015) ‘The interplay of histone

modifications – writers that read’, *EMBO reports*, 16(11), pp. 1467–1481. doi: 10.15252/embr.201540945.

Zhang, Z. *et al.* (2017) ‘Crosstalk between histone modifications indicates that inhibition of arginine methyltransferase CARM1 activity reverses HIV latency’, *Nucleic Acids Research*. Oxford University Press, 45(16), pp. 9348–9360. doi: 10.1093/nar/gkx550.

The GBAF chromatin remodeling complex binds H3K27ac and mediates enhancer transcription

Kirill Jefimov^{1,#}, Nicolas Alcaraz^{2,#}, Susan L. Kloet^{3,#}, Signe Värvi⁴, Siri Aastedatter Sakya¹, Christian Dalager Vaagenso², Michiel Vermeulen^{3,*}, Rein Aasland^{4,*}, and Robin Andersson^{2,*}.

¹ Department of Biological Sciences, University of Bergen, 5008 Bergen, Norway, ² The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark,

³ Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Radboud University Nijmegen, 6525 Nijmegen, The Netherlands, ⁴ Department of Biosciences, University of Oslo, 0371 Oslo, Norway

#: equal contribution

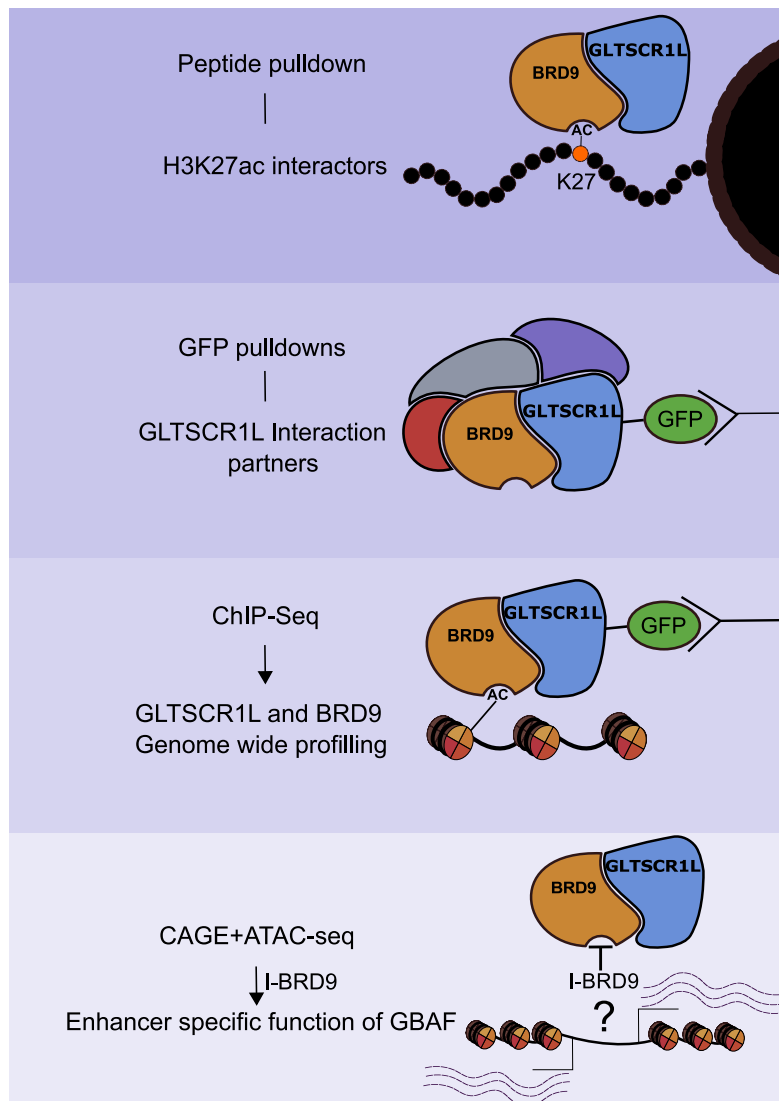
*: to whom correspondence should be addressed

(M.V.: michiel.vermeulen@science.ru.nl, R.Aa.: rein.aasland@ibv.uio.no, R.An.: robin@binf.ku.dk)

Abstract

H3K27ac is associated with regulatory active enhancers, but its exact role in enhancer function remains elusive. Using mass spectrometry-based interaction proteomics, we identified the Super Elongation Complex (SEC) and GBAF, a non-canonical GLTSCR1L- and BRD9-containing SWI/SNF chromatin remodeling complex, to be major interactors of H3K27ac. We systematically characterized the composition of GBAF and the conserved GLTSCR1/1L 'GiBAF'-domain, which we found to be responsible for GBAF complex formation and GLTSCR1L nuclear localization. Inhibition of the bromodomain of BRD9 revealed interaction between GLTSCR1L and H3K27ac to be BRD9-dependent and led to GLTSCR1L dislocation from its preferred binding sites at H3K27ac-associated enhancers. GLTSCR1L disassociation from chromatin resulted in genome-wide downregulation of enhancer transcription while leaving most mRNA expression levels unchanged, except for reduced mRNA levels from loci topologically linked to affected enhancers. Our results indicate that GBAF is an enhancer-associated chromatin remodeler important for transcriptional and regulatory activity of enhancers.

Graphical abstract



Regulatory events at gene promoters and transcriptional enhancers modulate cell-type specific gene activities and allow cells to respond to external cues (Heinz et al. 2015; Beagrie & Pombo 2016; Lenhard et al. 2012). Several processes take part in these actions, including ATP-dependent chromatin remodeling, transcription factor (TF) and co-activator binding, and the recruitment of general transcription factors (GTF) and RNA polymerase II (Pol II) (Vernimmen & Bickmore 2015). Integral to the activation of gene transcription is a favorable chromatin environment. Gene transcriptional activity is associated with permissive histone post-translational modifications (PTMs) (Li et al. 2007) and a three-dimensional folding of the genome that brings enhancers in close proximity with gene promoters (Sanyal et al. 2012; Rao et al. 2014), thereby allowing the regulation of target genes.

Genome-wide charting of histone PTMs has identified chromatin signatures associated with repressive and transcriptionally permissive states as well as enhancer activity (Jenuwein 2001; Heintzman et al. 2009; Kundaje et al. 2015). However, the mechanisms by which many histone PTMs exert their putative functions are elusive. These may include the modulation of nucleosome-DNA interaction strength or recruitment of protein complexes that execute specific functions (Kouzarides, 2007). Among many potential histone PTMs, acetylation of histone H3 at lysine 27 (H3K27ac) is associated with regulatory activity of both gene promoters and enhancers (Creyghton et al., 2010, Ernst et al., 2011, Rajagopal et al., 2014, Rada-Iglesias et al., 2011). H3K27ac is deposited by acetyltransferases CBP and P300 (Tie et al. 2009). Targeting of P300 fused with nuclease-null dCas9 has been reported to elevate H3K27ac levels and activate transcription from promoter-proximal as well as distal regulatory regions (Hilton et al. 2015).

Although the association of H3K27ac with active regulatory elements is known, its potential role in enhancer function is still unclear. Given its recognition by acetyl-lysine reader domains, such as BROMO and YEATS, it is possible that H3K27ac attracts effector proteins to enhancers. Previous work to identify H3K27ac-interacting proteins has relied on chromatin immunoprecipitation followed by identification of co-purified proteins by mass-spectrometry (ChIP-MS) (Engelen et al. 2015; Ji et al. 2015). Unfortunately, due to experimental setup (purification of sheared, crosslinked chromatin with pre-existing protein complexes), it is hard to discern the contribution

of H3K27ac to enhancer function from DNA-dependent activities, such as TF binding.

Here, we applied a SILAC-based histone-peptide pulldown approach to identify proteins interacting with H3K27ac in mouse embryonic stem cells and HeLa cells, also in combination with H3K23ac. These experiments revealed Super Elongation Complex (SEC) and BRM/BRG1 Associated Factors (BAF) as major interactors of H3K27ac. Furthermore, we identified GLTSCR1L protein to be a part of a H3K27ac-specific BAF complex, recently identified as GBAF (Alpsoy & Dykhuizen 2018), and we characterized the interaction partners of the GLTSCR1L protein and its conserved protein-interaction domain. Our findings led us to further investigate the function of the GLTSCR1L-BRD9 containing GBAF complex in transcription. Upon treatment with a BRD9 bromodomain inhibitor, we observed enhancer-specific transcriptional abnormalities, indicating an important role for GBAF in enhancer transcription and regulatory activity.

Results

SWI/SNF and Super Elongation Complex are major acetyl-lysine interacting complexes

We applied an established SILAC histone peptide pulldown approach (Vermeulen et al. 2007) to identify proteins that specifically bind to H3K27ac (Methods). To discriminate proteins that bind specifically to H3K27ac from proteins with a general affinity for acetyl-lysine-containing histone peptides, we performed the same experiment with peptides associated with the highly abundant (Sidoli et al. 2015; Tvardovskiy et al. 2015) H3K23ac mark (Fig. 1a,b). In addition, we included an experiment with histone peptides diacetylated at both H3K23 and H3K27 residues (Fig. 1e,f), since these acetylations frequently colocalize and are associated with active transcription (Wang et al. 2008). To understand how the recognition of H3K27ac compares across mammals and whether there are any differences between pluripotent and terminally differentiated states, we used both HeLa cells and mouse embryonic stem cells (mESCs) as sources of nuclear extracts for the pulldown experiments. Each peptide pulldown experiment resulted in the

identification of more than 2,000 proteins, of which 10 to 30 showed significant acetyl-lysine specific binding (Fig. 1a-c, Supplementary Fig. 1a-c, Supplementary Data 1). Our experimental design allowed us not only to identify acetyl-lysine binding complexes, but also to estimate their binding preferences (H3K27ac versus
5 H3K23ac). We averaged forward and reverse SILAC ratios of each individual pulldown separately and applied hierarchical clustering. Proteins from the same complexes were found to cluster together (Fig. 1d, Supplementary Fig. 1d). The majority of identified histone H3 acetyl-lysine interacting proteins are members of two distinct protein complexes: the SEC and the SWI/SNF family of chromatin
10 remodeling complexes (Wang et al. 1996; Lin et al. 2010). In addition, we detected NuA4 complex components DMAP1 and YEATS4 (Doyon et al. 2004) and components of the general transcription factor TFIID (Dymlach et al. 1991).

The SEC complex is important for the activation of transcription by release of paused
15 Pol II (Lin et al 2010; Luo et al, 2012). Although many SEC complex components were identified in pulldowns with all acetylated histone peptides, we observed differences in SILAC ratios between nuclear extracts. The highest SILAC ratio for the SEC complex subunits was observed with monoacetylated H3K27ac histone peptide with HeLa and diacetylated H3K23acK27ac peptide with mESC nuclear extracts. In
20 HeLa cells, we were able to detect the presence of AF9, AFF4, AFF1 and ENL, but not ELL family proteins nor P-TEFb (Fig. 1b), important mediators of SEC and Pol II interaction (Luo et al, 2012; Knutson et al, 2016). In mESCs, we detected nearly all SEC complex subunits, including AF9, AFF4, ELL3 and CDK9 (Supplementary Fig. 1c). AF9 contains a YEATS-domain that is able to recognize histone modifications
25 and has previously been found to bind H3K9ac and H3K27ac (Li et al, 2014).

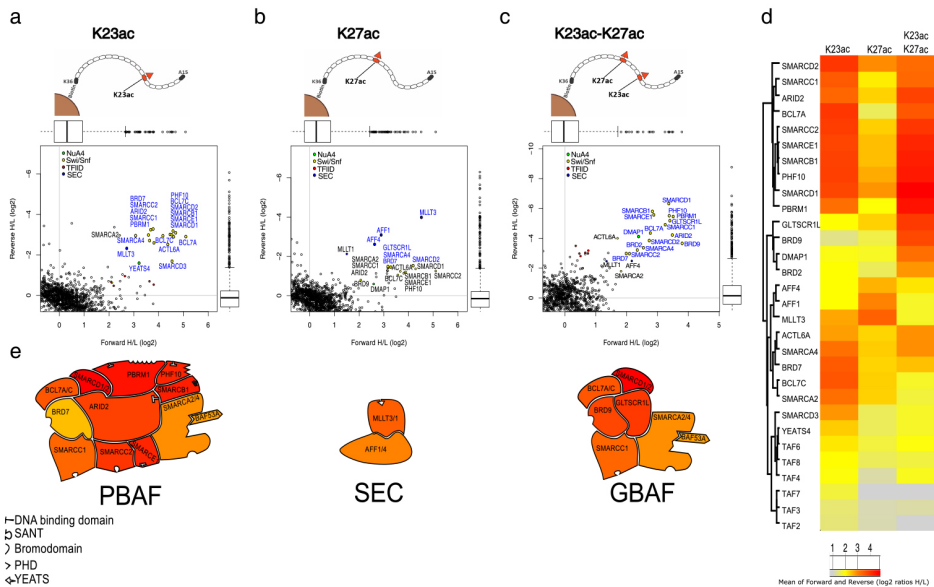


Figure 1. SILAC histone peptide pulldowns identify SEC and SWI/SNF complexes as major interactors of H3K23ac and H3K27ac. a-c Scatter plots of forward (horizontal axes) versus reverse (vertical axes) SILAC ratios from histone peptide pulldowns using (a) H3K23ac, (b) H3K27ac, or (c) diacetylated H3K23acK27ac baits with HeLa cell nuclear extracts (n=2 pulldowns for each bait). The lower and upper hinges of boxes correspond to the first and third quartiles of data, respectively, and the whiskers extend to the largest and smallest data points no further away than 1.5 times the interquartile range. **d** Heatmap of average forward and reverse SILAC ratios (\log_2). The ordering of rows in the heatmap and the associated dendrogram were derived from agglomerative hierarchical clustering of the H3K23- and K27-interacting proteins identified in the histone peptide pulldowns (a-c), using Euclidean distances. **e** Visual summary of the most enriched interactors of H3K23ac (left), H3K27ac (middle), and H3K23acK27ac (right) peptide pulldowns. Subunits are colored according to the \log_2 value of their enrichment in the pulldowns.

Mammalian SWI/SNF (or BAF) complexes are ATP-dependent chromatin remodelers implicated in a wide variety of processes in the cell nucleus (Hargreaves & Crabtree 2011). Different combinations of SWI/SNF subunits generate a diversity of functionally distinct complexes, including two canonical subfamilies (BAF, PBAF), their subcomplexes, and GBAF, a recently described non-canonical SWI/SNF complex (Phelan et al. 1999; Hohmann & Vakoc 2014; Alpsy & Dykhuizen 2018). We observed known core SWI/SNF subunits (SMARCA2/4, SMARCC1, SMARCC2) and shared subunits (SMARCB1, SMARCE1, SMARCD1/2/3, ACTL6A, BCL7A/C) in pulldowns with each of the acetylated histone peptide baits, with both HeLa and

mESC nuclear extracts. The complete PBAF complex, represented by ARID2 (BAF200), PHF10 (BAF45A), PBRM1 (BAF180) and BRD7, was found to prefer interaction with H3K23ac-containing peptides (both mono-acetylated H3K23ac and di-acetylated H3K23ack27ac peptides) (Fig. 1a,c, Supplementary Fig. 1a,c).
5 However, known BAF complex specific subunits (ARID1A/1B (BAF250A/B), DPF1/2/3 (BAF45B/C/D)) were absent from these pulldowns. Instead, we observed BRD9 and GLTSCR1L (BICRAL)¹ proteins, subunits of the GBAF complex (Alpsoy & Dykhuizen 2018). These proteins predominantly interact with H3K27ac-containing peptides. GLTSCR1L was the most enriched interactor of monoacetylated H3K27ac
10 in mESCs and one of the most enriched H3K23ack27ac interactors in HeLa cells (Fig.1c, Supplementary Fig. 1b).

Taken together, mass spectrometry-based interaction proteomics experiments identified SWI/SNF and SEC as major acetyl-lysine readers, which is consistent with
15 earlier findings (Li et al. 2014; Chandrasekaran & Thompson 2007; Chandy et al. 2006). Additionally, we identified GBAF components GLTSCR1L and BRD9 as prominent H3K27ac-interacting proteins.

GLTSCR1L is a subunit of a distinct non-canonical BAF complex

20 Due to the observed preferential binding of the GLTSCR1L protein to H3K27ac-containing histone peptides, we decided to identify GLTSCR1L-interacting proteins by quantitative interaction proteomic methods. We generated a HeLa cell line stably expressing doxycycline-inducible GFP-tagged GLTSCR1L and performed label-free GFP pulldowns followed by liquid chromatography-MS (LC-MS/MS, Methods) (Smits
25 et al. 2013). The full length GFP-GLTSCR1L protein interacted with core SWI/SNF subunits including SMARCD1 (BAF60a), SMARCC1 (BAF155), SMARCA4 (BRG1), and BAF-specific subunit SS18 (Fig. 2a). One of the strongest GFP-GLTSCR1L interactors was BRD9, which also clustered together with GLTSCR1L in our histone peptide pulldown analysis (Fig. 1d, Supplementary Fig. 1d). Additionally, BCL7C was

¹ GLTSCR1L and GLTSCR are paralogs of a group of proteins found in vertebrates. They are also known by the names BICRAL and BICRA (UniProt IDs Q6AI39 and Q9NZM4, respectively). GLTSCR1L and GLTSCR contain a highly conserved domain, which we refer to as 'GiBAF' (GLTSCR1/1L domain interacting with BAF complex; Supplementary Fig. 2).

identified as a novel interactor of GLTSCR1L, while other canonical BAF complex subunits, such as ARID1A, DPF1/2/3 (BAF45B/C/D), and SMARCE1 (BAF57), were absent in the pulldowns.

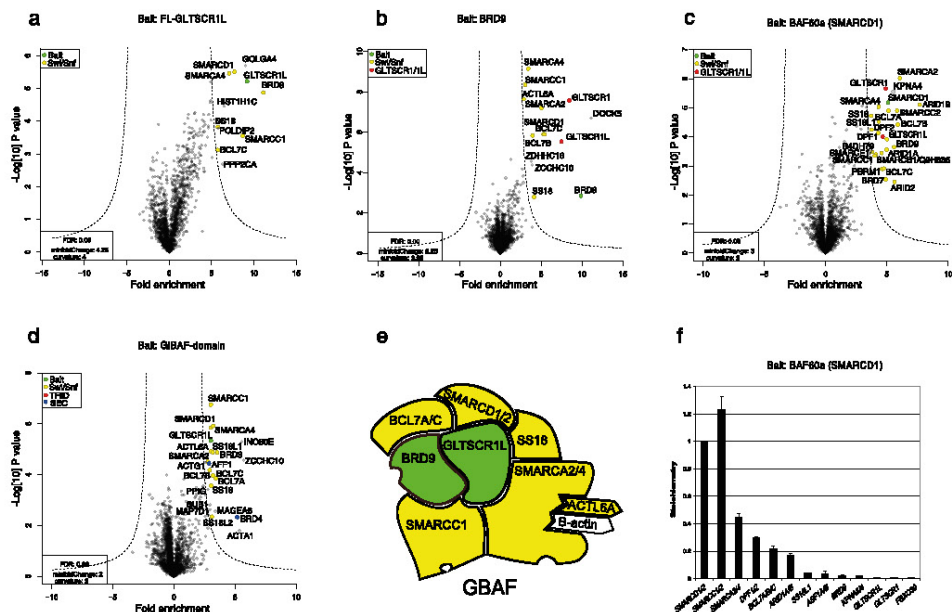


Figure 2. Quantitative MS analysis reveals the interacting proteins of GLTSCR1L protein and its GiBAF-domain. a-d Volcano plots displaying fold enrichments (\log_2 GFP fusion protein over mock, horizontal axes) versus t-test P values ($-\log_{10}$, vertical axes) of identified interactors of affinity-purified GFP-GLTSCR1L (a), GFP-BRD9 (b), GFP-BAF60a (SMARCD1) (c) and GFP-GiBAF-domain of GLTSCR1L (d) isolated from HeLa whole cell extracts (a-c) or nuclear extracts (d). Each volcano plot represents three independent pulldowns. e Schematic representation of the GiBAF complex. Proteins in green have been used as baits, proteins in yellow represent identified interactors. f Stoichiometry of GFP-BAF60a interactors. The intensity-based absolute quantification (iBAQ) value of each protein is divided by the iBAQ value of the bait and plotted relative to a bait value of 1. Data are shown as mean \pm s.d. ($n = 3$ pulldowns).

In order to validate GLTSCR1L as a *bona fide* interactor of SWI/SNF complexes, we performed reciprocal label-free GFP pulldowns. We generated three doxycycline-inducible HeLa cell lines expressing GFP-tagged baits of identified GLTSCR1L-interacting proteins: BRD9 (Fig. 2b), SMARCD1 (BAF60a) (Fig. 2c) and SMARCC1 (BAF155) (Supplementary Fig. 3a). As expected, we observed an enrichment of SWI/SNF complex subunits in all pulldowns. The highest number of identified SWI/SNF subunits was observed in pulldowns with SMARCC1 (BAF155) as a bait,

followed by SMARCD1 (BAF60a) and BRD9 pull-downs (Supplementary Table 1). Additionally, GLTSCR1L was detected as an interactor of each of these three bait proteins (Fig. 2b,c, Supplementary Fig. 3a). We also observed a nearly complete overlap of interactors in the GFP-BRD9 and GFP-GLTSCR1L pull-downs (Fig. 2a,b),
5 suggesting that these two proteins frequently co-occur in the same sub-complex. Interestingly, the homologous protein GLTSCR1 also appeared in all reciprocal pull-down experiments but not in pull-downs with GLTSCR1L, indicating that GLTSCR1 and GLTSCR1L proteins bind to SWI/SNF complexes in a mutually exclusive manner.

10

The observed binding of both GLTSCR1 and GLTSCR1L to SWI/SNF complexes led us to investigate the role of their shared, conserved domain in mediating the interaction with SWI/SNF family members. We refer to the domain as 'GiBAF', for GLTSCR1/1L domain interacting with BAF complex. First, we produced a TY1-tagged GLTSCR1L protein containing an in-frame deletion of the GiBAF-domain. By
15 means of immunofluorescence microscopy (IF) with FL and domain deletion mutant proteins, we found that the GiBAF-domain is responsible for the nuclear localization of GLTSCR1L (Supplementary Fig. 3b). Next, we generated two other HeLa cell lines with doxycycline-inducible GFP-tagged GLTSCR1L protein mutants expressing
20 only the GiBAF-domain or the full-length protein lacking the domain (▲GiBAF-domain). In pull-downs with the GiBAF-domain only, we observed all full length GFP-GLTSCR1L protein interactors in addition to some other BAF complex components (BCL7A/B/C, ACTL6A, SMARCA2, SS18/L1/L2) (Fig. 2d). GiBAF-domain interaction with BAF155, BAF60a and BRD9 was also confirmed by immunoblotting
25 (Supplementary Fig. 3g) (Alpsoy & Dykhuizen 2018). Additionally, we detected an interaction between the GiBAF-domain from GLTSCR1L protein and BRD4, in line with previous work demonstrating an interaction between GLTSCR1 and the BRD4 ET-domain, which contributes to transcriptional regulation of BRD4 target genes (Rahman et al. 2011). We also observed interaction with INO80E (Chen et al. 2011)
30 from the INO80 (Jin et al. 2005) chromatin remodeling complex, which has previously been shown to interact with SWI/SNF complex components (Cai et al. 2007; Yao et al. 2008). AFF1, a SEC complex subunit, was also found to interact with the GiBAF-domain of GLTSCR1L. Similar analysis performed with the GiBAF-domain deletion mutant of GLTSCR1L protein (▲GiBAF-domain) resulted in no

significant interactions with BAF complex subunits (Supplementary Fig. 3c), further supporting a role for the GiBAF-domain in mediating interactions with other SWI/SNF subunits. Instead, we observed a significant enrichment of TAF6 and TAF4 TFIIID components co-purified with ▲GiBAF-domain, indicating that other parts of the
5 GLTSCR1L protein may be responsible for non-SWI/SNF protein interactions.

Label-free GFP pulldowns allowed to use the intensity-based absolute quantification (iBAQ) algorithm (Schwanhäusser et al. 2011) to calculate the relative abundance (stoichiometry) of observed proteins. Stoichiometry values from the FL GLTSCR1L
10 and GiBAF-domain pulldowns (Supplementary Fig. 3d-e) confirmed a strong association between GLTSCR1L and BRD9. Interestingly, analyses of reciprocal pulldown-MS with BAF complex components SMARCD1 and SMARCC1 revealed very low levels of GLTSCR1L protein abundance (Fig. 2f, Supplementary Fig. 3f; ~1,6% of SMARCD1 and 0,066% of SMARCC1 are found together with
15 GLTSCR1/1L), indicating that GLTSCR1L is either a transient interactor or a component of a rare BAF complex in HeLa cells. Taken together, these results suggest that GLTSCR1L is a subunit of a distinct SWI/SNF chromatin remodeling subcomplex, whose interactions with SWI/SNF subunits and its nuclear localization is mediated by the GiBAF-domain.

20 Selective inhibition of the BRD9 bromodomain leads to disassembly of GLTSCR1L but not BRD9 from chromatin

Since the GLTSCR1L protein does not contain any known DNA/chromatin-binding domain, we hypothesized that, due to the interaction between GLTSCR1L and BRD9, the latter might facilitate targeting of GBAF to chromatin. In order to validate
25 the association between GLTSCR1L, BRD9 and H3K27ac, we performed histone-peptide pulldown assays using I-BRD9 (Theodoulou et al. 2016), a specific inhibitor of the BRD9 bromodomain. I-BRD9 does not impair any other function of BRD9 beyond its acetyl-lysine recognition, making it a suitable tool to investigate the role of H3K27ac binding by the BRD9 bromodomain with limited secondary effects. Both
30 GLTSCR1L and BRD9 were found to follow the trend of acetyl-lysine recognition observed in histone-peptide pulldown MS (H3K23ac<H3K27ac<H3K23acK27ac) (Fig. 1d). After treatment with I-BRD9, we detected a loss of specificity of BRD9

binding to H3K27ac-containing peptides. Although BRD9 still showed interaction with histone peptides while its bromodomain was blocked by I-BRD9, GLTSCR1L binding to H3K27ac-containing histone peptides was completely abrogated (Fig 3a).

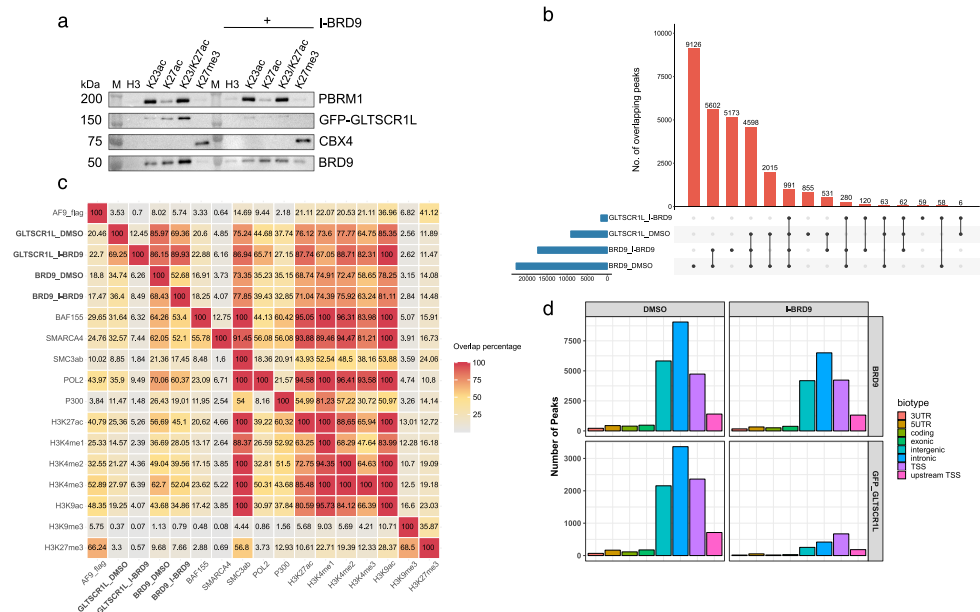


Figure 3. Inhibition of the bromodomain of BRD9 leads to disassembly of GLTSCR1L from chromatin. **a** Histone peptide pulldown assay using biotinylated histone H3 peptides (aa 15-36) and nuclear extract from HeLa-FRT-GFP-GLTSCR1L cells in the absence and presence of BRD9 bromodomain inhibitor I-BRD9. Unmodified (H3) and modified peptides were used. Input and affinity purified fractions were analyzed after SDS-PAGE by immunoblotting. PBRM1 and CBX4 serve as controls for I-BRD9 specificity. M: protein molecular weight standard. **b** Histogram attribute (upset) plot of ChIP-seq peak intersections between GLTSCR1L and BRD9 with and without I-BRD9. Vertical axis gives the frequency of overlaps between combinations of peak calls specified on the horizontal axis. **c** Heatmap of the overlap percentages between GLTSCR1L and BRD9 ChIP-seq peaks (bold) and ENCODE-called peaks of factors and chromatin marks associated with transcription. Each cell shows the percentage of ChIP-seq peaks for factors and marks in rows overlapping with ChIP-seq peaks for factors and marks in columns. **d** Genomic annotation of GLTSCR1L and BRD9 ChIP-seq peaks based on GENCODE-inferred (v19) genomic annotation biotypes.

To validate our histone-peptide pulldown results and to investigate whether GLTSCR1L, BRD9 and H3K27ac colocalize on chromatin, we performed chromatin IP of GLTSCR1L and BRD9 followed by high-throughput sequencing (ChIP-seq). We used a HeLa FRT cell-line expressing GFP-tagged GLTSCR1L and antibodies to GFP and wtBRD9, due to the unavailability of a ChIP-seq grade GLTSCR1L

antibody. In parallel, we conducted similar ChIP-seq experiments with I-BRD9 treated cells. A total of 22,559 and 9,155 peaks were called (Irreproducible Discovery Rate < 0.05) for BRD9 and GFP-GLTSCR1L, respectively. We observed a strong colocalization between BRD9, GLTSCR1L, and H3K27ac (Supplementary
5 Fig. 4a; permutation test, $P < 1 \times 10^{-4}$ for BRD9 and GLTSCR1L). In agreement with the histone-peptide binding assay, selective inhibition of the BRD9 bromodomain with I-BRD9 led to only a mild reduction of BRD9 binding to chromatin (22,559 versus 17,366 peaks for DMSO and I-BRD9, respectively), but severely abrogated GLTSCR1L chromatin binding (9,115 vs 1,639 peaks for DMSO and I-BRD9,
10 respectively) (Fig 3b).

We observed a large degree of overlap of BRD9 and GLTSCR1L binding sites with ChIP-seq peaks of cohesin (SMC3ab) and active chromatin marks, especially H3K4me1/2, H3K27ac, and H3K9ac (Fig. 3c, 70-85% of GLTSCR1L and BRD9
15 peaks). GLTSCR1L and BRD9 was also associated with Pol II, with 70% of Pol II peaks overlapping with those of BRD9 and 35% with GLTSCR1L, suggesting a role for GBAF in the control of Pol II activity. Reassuringly, we detected chromatin colocalizations of both GLTSCR1L and BRD9 with core SWI/SNF subunits (~32% of BAF155 and SMARCA4 peaks overlapped with GLTSCR1L peaks and ~65% with
20 BRD9). A small degree of colocalization was also observed between AF9 protein (a H3K9/K27ac reader), GLTSCR1L, and BRD9 (~20% of peaks for both proteins). However, only ~14% of GLTSCR1L and BRD9 peaks were found to overlap with H3K27me3 peaks compared to 41% for AF9, further pointing to a role of GBAF in transcriptionally active chromatin. Specifically, GLTSCR1L and BRD9 peaks were
25 primarily located in intronic or intergenic regions and, to a lesser extent, at gene promoters. Upon I-BRD9 treatment, BRD9 largely remained at intronic/intergenic regions, whereas the greatly reduced binding sites of GLTSCR1L were mostly found at active gene TSSs (Fig. 3c,d). Taken together, our results indicate that GBAF binds active chromatin and that the bromodomain of BRD9 is responsible for GLTSCR1L
30 binding to putative enhancers.

BRD9 inhibition globally reduces enhancer transcription

To assess the role of GLTSCR1L and the GBAF complex in transcription, we examined I-BRD9 treated HeLa cells stimulated with epidermal growth factor (EGF) to induce a rapid transcriptional response. First, we inspected GLTSCR1L and BRD9 binding as well as RNA expression levels of the EGF-inducible gene *NR4A1* and its known enhancer 80 kb downstream of the *NR4A1* gene locus (Lai et al. 2015) by qPCR. EGF treatment increased binding of both GFP-GLTSCR1L and BRD9 proteins at the enhancer, but no increase was observed in I-BRD9 treated cells (Supplementary Fig. 5a). I-BRD9 did not affect the induction of *NR4A1* gene expression by EGF, in line with a non-significant enrichment of GFP-GLTSCR1L and BRD9 proteins at the *NR4A1* gene promoter, but almost completely abrogated expression of the enhancer RNA (Supplementary Fig. 5b).

We next performed 5' end sequencing of capped RNAs, using Cap Analysis of Gene Expression (CAGE (Takahashi et al. 2012)), to assess the effect of EGF-induced transcription initiation events and enhancer activities (Andersson et al. 2014) genome-wide. CAGE experiments were performed in parallel with ATAC-seq across the EGF-response time course to focus on transcription initiation events at open chromatin loci such as enhancers and gene promoters. Using CAGE data in combination with ATAC-seq data, we were able to dissect the *NR4A1* super-enhancer into a set of five bidirectionally transcribed enhancers in open chromatin with similar transcriptional EGF-response patterns, and found all of them to be downregulated by I-BRD9 (Fig. 4a, focusing on three enhancer constituents). Genome-wide, we next focused on all transcribed nucleosome-free regions (NFRs), inferred from ATAC-seq peaks associated with CAGE expression above estimated background noise level (Methods). Transcribed NFRs corresponded to ~11% (23,639 out of 213,017) of all detected open chromatin loci and displayed clear differences in expression patterns between I-BRD9 and DMSO-treated samples (52% of variance explained) as well as between time points (Supplementary Fig. 5c). Time points 30 and 60 minutes displayed different transcriptional activities than time points 0 and 240 minutes after treatment, indicating that at 240 minutes after EGF induction most transcriptional activities had returned to baseline levels.

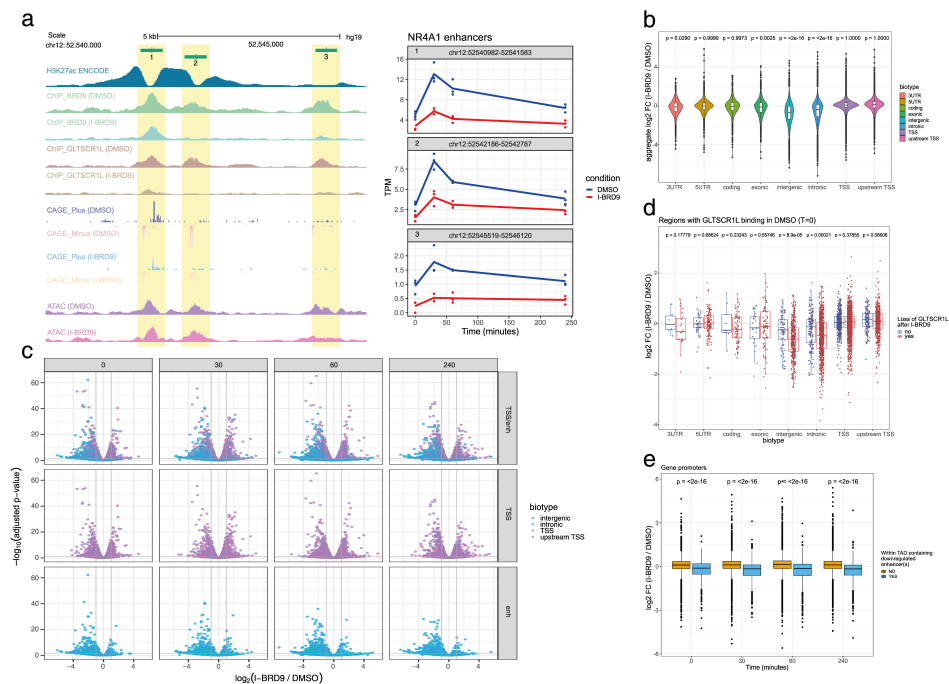


Figure 4. Inhibition of the bromodomain of BRD9 globally inhibits enhancer transcription. **a** Genome browser view (left) of the NR4A1 super-enhancer locus, showing tracks with pile-up signal from H3K27ac (ENCODE, HeLa S3), GLTSCR1L and BRD9 ChIP-seq, ATAC-seq, and CAGE data. Differences in ChIP-seq signal (except for H3K27ac), ATAC-seq signal, and CAGE expression between control (DMSO) and I-BRD9 treated cells are observable at three highlighted super-enhancer constituents before EGF stimulus. CAGE expression levels (TPM normalized) are shown (right) for the three enhancers across the HeLa EGF stimulation time course, comparing control (DMSO) and I-BRD9 treated cells. **b** Violin plots depicting densities of aggregated fold changes of expressed NFRs across the EGF treatment time course, broken up by genomic annotation biotype inferred from GENCODE (v19). **c** Volcano plots of NFR expression fold changes (horizontal axes) and adjusted P-values of differential expression (vertical axes, Methods), comparing I-BRD9 treated cells with non-treated (DMSO), broken by biotype and time point after EGF stimulation. **d** Box-and-whisker plot (as in Fig. 1) of expression fold changes (\log_2) comparing control (DMSO) and I-BRD9 inhibited cells at GLTSCR1L-bound expressed NFRs (before EGF stimulus) that lose or keep GLTSCR1L binding upon I-BRD9 treatment. **e** Box-and-whisker plot (as in Fig. 1) of fold changes (I-BRD9 versus control) of gene promoters within TADs that contain or do not contain down-regulated enhancers.

20

To investigate the effect of I-BRD9 on transcriptional responses, we compared the aggregated fold change of NFR expression levels in I-BRD9 treated cells and control cells across all EGF time points. Similar to our observation at the NR4A1 enhancer locus, the EGF responses of putative enhancers, as indicated by transcribed intergenic and intronic loci (75% and 25% of expressed intergenic and intronic regions overlap with FANTOM5 enhancers (Andersson et al. 2014), respectively), were significantly downregulated upon BRD9 inhibition (Wilcoxon signed-rank Test, lower tail, $P < 2.2 \times 10^{-16}$) in contrast to the responses of mRNA promoters, which were largely not affected (Fig. 4b). We further compared transcriptional events measured at each time point in control and I-BRD9 treated cells (Fig. 4c). Most downregulated events ($\log_2 FC < -1$, FDR adjusted $P < 0.05$) were detected at one hour after EGF induction (Supplementary Fig. 5d) and occurred mostly at intronic or intergenic regions (894 out of 1,335 NFRs compared to 336 mRNA promoters). Thus, I-BRD9 treatment leads to a preferential downregulation of enhancer transcription throughout the HeLa EGF induction time course. Notably, 86% of downregulated FANTOM5 enhancers were depleted of GFP-GLTSCR1L by I-BRD9. Reciprocally, GFP-GLTSCR1L depleted intergenic/intronic NFRs in general (Wilcoxon signed-rank test, $P = 8.9 \times 10^{-5}$ and $P = 2.1 \times 10^{-4}$ for intergenic and intronic NFRs, respectively, Fig. 4d) and FANTOM5 enhancers in particular (Wilcoxon signed-rank test, $P = 3.7 \times 10^{-9}$) were associated with downregulated expression. These results indicate that transcription of enhancers, but not mRNA genes, is affected by H3K27ac-recognition by GBAF through the bromodomain of BRD9.

Although mRNA expression levels were, to a large extent, not affected by BRD9 inhibition, we hypothesized that downregulation of enhancer expression levels reflects reduced enhancer regulatory activities. To test this hypothesis, we compared expression levels of putative enhancer-promoter pairs contained within the same topologically associating domains (TADs) of HeLa cells (Rao et al. 2014). Indeed, promoters within TADs containing at least one differentially downregulated enhancer showed significantly larger downregulation upon I-BRD9 treatment in all time points when compared to promoters in TADs not containing any downregulated enhancer (Fig. 4e). These results indicate that GBAF is important not only for enhancer transcription but also for enhancer regulatory activities.

We posited that GBAF may have a role in forming or maintaining permissive chromatin at enhancers. At putative intronic/intergenic enhancers associated with transcriptional downregulation and GFP-GLTSCR1L depletion by I-BRD9, we noted a significant reduction in chromatin accessibility (Wilcoxon signed-rank paired test, $P = 5.9 \times 10^{-9}$, Supplementary Fig. 5e). Taken together, our data indicate that I-BRD9 leads to a reduced rate of eRNA transcription at GLTSCR1L-depleted enhancers and a downregulation of their target genes. We hypothesize that this is the consequence of reduced GBAF chromatin remodeling activity at enhancers caused by disrupted recognition of H3K27ac by BRD9 through inhibition of its bromodomain.

10 Discussion

Although H3K27ac is frequently associated with active enhancers, its role in enhancer activity and its protein interaction environment have remained elusive. In this study, we applied a SILAC histone-peptide pulldown MS approach to identify protein complexes that interact with H3K27ac alone and in combination with H3K23ac. This approach allowed us to assess acetyl-lysine interaction preferences and to identify the SEC and GBAF complexes as major interactors of H3K27ac. Specifically, we identified H3K27ac interactions with both BRD9 and GLTSCR1L, definitive subunits of the GBAF complex (Alpsoy & Dykhuizen 2018), indicating a H3K27ac-associated function of GBAF.

20 Our systematic investigation of the subunits of GBAF revealed insights into their putative functions. Our data refines the composition of GBAF (Alpsoy & Dykhuizen 2018) by adding BCL7A/B/C as an additional subunit. Interestingly, none of the subunits of the GBAF complex contains known DNA binding domains, unlike those of canonical BAF/PBAF complexes (e.g. ARID2/ARID1/1a, PBRM1, SMARCE1, SMARCB1, PHF10 (Supplementary Table 1)). Hence, the GBAF complex has only two functional chromatin interaction domains, the bromodomains of SMARCA2/4 and BRD9. Due to this specific feature, we hypothesize that targeting of GBAF to chromatin is more sensitive to the recognition of acetylated histones (H3K27ac), than the canonical SWI/SNF complexes that contain sequence-independent DNA-binding subunits. Investigation of the function of the GiBAF-domain of GLTSCR1L revealed that it is capable of binding all inferred GLTSCR1L- and BRD9-interacting proteins as

well as BRD4 and AFF1. We also found that the GiBAF-domain is responsible for the nuclear localization of GLTSCR1L. Other parts of the GLTSCR1L protein may be involved in Pol II interactions or pre-initiation complex formation, as indicated by the enrichment of TFIID components in GFP pulldown MS results of a GiBAF-domain deletion mutant.

To investigate the role of acetyl-lysine recognition by BRD9, we made use of I-BRD9, a selective inhibitor of the BRD9 bromodomain. I-BRD9 is a useful tool since it doesn't impair any function of BRD9 beyond its acetyl-lysine recognition (Theodoulou et al. 2016). Our histone-peptide pulldown immunoblot experiments with I-BRD9 revealed that GLTSCR1L chromatin interaction is mediated by the bromodomain of BRD9 and is H3K27/23ac dependent. Upon I-BRD9 treatment, BRD9 lost specificity in acetyl-lysine recognition but remained bound to histone peptides. This is likely due to its participation in several BAF subcomplexes, where binding is mediated by acetyl-lysine recognition modules of other BAF subunits, such as the bromodomain of SMARCA2/4. Concomitantly with these proteomic interaction results, GLTSCR1L and BRD9 colocalized at NFRs flanked by H3K27ac. We observed an enrichment of BRD9 and GLTSCR1L binding at intronic and intergenic NFRs and to a lesser extent at gene promoters, and used I-BRD9 to verify the role of the BRD9 bromodomain in the targeting of GBAF complex to chromatin. Upon I-BRD9 treatment, in agreement with our histone peptide pulldown immunoblot results, GFP-GLTSCR1L was observed to dislocate from H3K27ac-associated intronic/intergenic chromatin and, in particular, from enhancers.

The observed loss of GLTSCR1L from enhancers upon BRD9 inhibition indicates that GBAF has a specific role at enhancers. This hypothesis is supported by a decrease in chromatin accessibility at GLTSCR1L-depleted enhancers and an overall downregulation of enhancer transcription initiation events in I-BRD9-treated HeLa cells stimulated with EGF. While overall mRNA expression levels were not affected by I-BRD9, genes residing within the same TADs as downregulated enhancers tended to follow the same trend. Therefore, we propose that GBAF is important for both enhancer transcription and enhancer regulatory activity. However, no downregulation was observed for *NR4A1* mRNA expression levels upon downregulation of eRNAs from its cognate enhancers by I-BRD9. This may be a

consequence of enhancer-independent EGF stimulation of *NR4A1* gene expression by promoter-proximal events. Nevertheless, as the timing of events in the communication between enhancers and promoters are poorly understood, we do not rule out that the *NR4A1* locus is in an alternative mode of regulation.

5

While several of our experiments point at an enhancer-associated function of GBAF, further investigations are needed to assess its exact role and the associated mechanisms. Although we observed a minor but significant decrease in chromatin accessibility at GLTSCR1L-depleted downregulated enhancers, we cannot rule out
10 that this is a consequence of reduced Pol II activity. Precise nucleosome positioning methods are needed to verify abnormalities in the maintenance of the enhancer NFR and nucleosomal occupancy at sites of transcriptional enhancers in the absence of a functional GBAF complex. Numerous studies have observed enhancer-specific effects of different SWI/SNF complexes, which are mostly related to chromatin
15 accessibility (Yu et al. 2013; Hodges et al. 2018; Nakayama et al. 2017; Alver et al. 2017; Wang et al. 2016). Given that GBAF is recruited to H3K27ac-marked enhancers, which are frequently associated with open chromatin, it is therefore possible that GBAF activity is secondary to chromatin opening and maintenance by other remodeling complexes. In line with its association with eRNA transcription, it is
20 conceivable that GBAF has a role in the positioning of enhancer TSS-proximal nucleosomes. The interactions of GBAF with transcriptional regulators (such as SEC, BRD4, and TFIID) indicate that GBAF may have additional roles at enhancer TSSs in Pol II pre-initiation complex formation or in the recruitment or assembly of factors needed for pause release and transcriptional elongation. Due to the critical
25 importance of correct enhancer activities in development and tissue homeostasis, we anticipate that dysfunction of GBAF or its targeting to chromatin may be connected with developmental abnormalities or complex diseases such as cancer.

Methods

Cell culture and inducible cell-line generation

For histone peptide pulldowns, HeLa S3 cells (ATCC) and IB10 mESCs (ATCC) were grown in DMEM SILAC medium without Lys and Arg (Silantes, 280001300), with dialyzed FBS (Silantes, 281000900) and supplemented either with light (R0K0) or heavy (R10K8) isotopes of lysine and arginine (Silantes, 282986440 Lys-0:HCl, 211604102 Lys-8:HCl, 282986444 Arg-0:HCl, 201604102 Arg-10:HCl). Additionally, IB10 mouse embryonic stem cells were cultured in the presence of 2i compounds as described (Kloet et al. 2016). For GFP pulldowns, ChIP-seq, and CAGE experiments, HeLa S3 cells were cultured in DMEM with 10% FBS, supplemented with glutamine and Pen-Strep. For induction and inhibition experiments, hEGF (Sigma-Aldrich, E9644; PeproTech, AF-100-15) was added to the culture medium at a concentration of 100 ng/ml and I-BRD9 (Tocris Bioscience, 5591) was added at a concentration of 10 μ M.

15

To generate inducible cell lines, 3×10^5 HeLa S3 cells containing an integrated FRT site (van Nuland et al. 2013) were seeded on 6-well plates and transfected after 24 h with two vectors: 1) pOG44 expressing FLP recombinase under the control of human cytomegalovirus (CMV) promoter and carrying blasticidin selection marker; 2) pcDNA5/FRT containing a hygromycin selection marker, FRT recognition sites, and an N-terminal GFP fusion protein in frame with the gene of interest (GLTSCR1L, BRD9, BAF60a, BAF155, GLTSCR1L GiBAF-domain, or GLTSCR1L with GiBAF-deleted under the control of a doxycycline-inducible (TET-ON) CMV promoter. 16 h after transfection, selection media supplemented with 3 μ g/ml blasticidin (Sigma) and 100 μ g/ml hygromycin (Invitrogen) was applied to cells. Single colonies that remained after 10 days of selection were picked and propagated in single 30 mm plates and subsequently tested for the expression of desired proteins after induction with doxycycline (1 μ g/ml).

30

Nuclear extract and whole cell lysate preparation

Nuclear extracts were prepared according to (Dignam et al., 1983). Cells were trypsinized, harvested, washed twice with PBS, and centrifuged at 400 g for 5 min at 4 °C. Resuspended cell pellets were incubated for 10 min at 4 °C in five volumes of
5 buffer A (10 mM HEPES-KOH, pH 7.9, 1.5 mM MgCl₂, and 10 mM KCl), then pelleted at 400 g for 5 min at 4 °C. Cells were resuspended in two volumes of buffer A supplemented with protease inhibitors and 0.15% NP-40. Cells were homogenized by 30–40 strokes with a type B pestle in Dounce homogenizer. After homogenization, lysates were spun at 3,200 g for 15 min at 4 °C. Nuclear pellet was
10 washed once with PBS and spun at 3,200 g for 5 min at 4 °C. Pellet was resuspended in two volumes of buffer C (420 mM NaCl, 20 mM HEPES-KOH, pH 7.9, 20% (v/v) glycerol, 2 mM MgCl₂, and 0.2 mM EDTA) with 0.1% NP-40, protease inhibitors, and 0.5 mM dithiothreitol and incubated with rotation for 1 h at 4 °C, then spun at 20 000 g for 30 min at 4 °C. The supernatant (nuclear extract) was aliquoted
15 and stored at –80 °C until further use.

Whole cell extracts were prepared by adding 5 cell pellet volumes of lysis buffer (0.5% NP40, 150 mM NaCl, 50 mM Tris pH 8.0, 10% Glycerol and 1 × Complete Protease Inhibitors). Cells were vortexed for 30 s and then incubated for 2 hr on a
20 rotation wheel. Samples were then centrifuged at 4,000 g in a swinging bucket rotor for 30 min, after which soluble whole extracts were aliquoted and stored at –80 °C until further use.

Histone-peptide pulldowns

25 Histone peptide pulldowns were performed as described in (Vermeulen 2012), with minor modifications. Briefly, biotinylated (modified and non-modified) histone H3 peptides (aa 15-36) were purchased from BIOSYNTAN GmbH. 50 µg of histone peptide per pull-down was incubated with 75 µl of MyOne Streptavidin C1 Dynabeads (Thermo Fisher, 65002) for 20 min at RT in peptide binding buffer (150
30 mM NaCl, 50 mM Tris–HCl, pH 8.0, 0.1% (v/v) NP40). Beads were washed three times with 1 ml protein binding buffer [150 mM NaCl, 50 mM Tris-HCl pH 8.0, 1% NP40, 0.5 mM DTT, 10 mM ZnCl₂ and complete protease inhibitors – EDTA free (Roche)]. 350-700 µg of nuclear extract (diluted to 0.6 mg/ml) was incubated with

immobilized histone peptides in protein binding buffer for two hours at 4 °C on a rotation wheel. Beads were washed five times with 1 ml of protein binding buffer containing 400 mM NaCl and finally twice with 1 ml of protein binding buffer. Beads from both pull-downs (with non-modified and modified peptide) were pooled, and bound proteins were eluted and visualized on 4 %–12 % SDS-PAA gradient gels (Invitrogen) by colloidal blue staining (Invitrogen). Lanes corresponding to Forward (heavy - modified peptide pulldown; light - non-modified peptide pulldown) and Reverse (light - modified peptide pulldown, heavy - non-modified peptide pulldown) experiments were divided into 6-8 pieces, sliced into small (~ 1 mm) fragments and then subjected to in-gel trypsin digestion essentially as described in (Shevchenko et al. 2007). Antibodies used for detection of histone peptide binding proteins from HeLa-FRT-GFP-GLTSCR1L nuclear extracts after SDS-PAAG using immunoblotting (Fig 3a; cells grown in non-SILAC medium) are listed in Supplementary Table 2.

15 **Mass spectrometry and data analysis of histone peptide pulldowns**

After trypsin digestion of gel slices, peptides were extracted, desalted using StageTips (Rappsilber et al., 2003), and separated using an EASY-nLC (Proxeon) connected online to a LTQ-Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). Scans were collected in data-dependent top speed mode with dynamic exclusion set at 60 seconds. Raw data were analyzed using MaxQuant version 1.5.1.0 with default settings and searched against the Uniprot mouse and human proteomes, release 2015_12. Analysis was performed using Perseus 1.5.5.3. After filtering, the mean value was calculated on the ratios from both forward and reverse experiments. Missing values were imputed with a normal distribution (settings: downshift = 1, window = 0) and scatter plots were made using R. Hierarchical clustering was used to generate heatmaps based on Euclidean distance (settings: linkage = average, number of clusters = 300, processing = k-means, iterations = 10, restarts = 1).

30

GFP affinity purification mass spectrometry (AP MS-MS)

Nuclear extracts (NE) or whole cell extracts (WCE) from doxycycline-induced (for 16 hours) and non-induced cells were subjected to GFP-affinity enrichment using GFP nanotrap beads (Chromotek) in triplicate. For each pull-down, 1 mg of NE or 3 mg of WCE was incubated with 7.5 μ l beads in incubation buffer (300 mM NaCl, 0.1 % NP-40, 0.5 mM DDT, 20 mM HEPES–KOH pH 7.9, 50 μ g/ml ethidium bromide) in a total volume of 400 μ l. Beads were washed twice with incubation buffer containing 0.5 % NP-40, twice with 1X PBS containing 0.5 % NP-40 and finally twice with 1X PBS. Affinity purified proteins were subject to on-bead trypsin digestion as described previously (Baymaz et al. 2014). Tryptic peptides were acidified and desalted using StageTips (Rappsilber et al. 2007) and separated with an online Easy-nLC 1000 (Thermo Scientific). Mass spectra were recorded on an LTQ-Orbitrap QExactive mass spectrometer (Thermo Fisher Scientific), selecting the top 10 most intense precursor ions for fragmentation, or on an LTQ-Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). Scans were collected in data-dependent top speed mode with dynamic exclusion set at 60 seconds.

LFQ peptide analysis and identification

Thermo RAW files from LFQ AP MS-MS were analyzed with MaxQuant version 1.5.1.0 using default settings and searching against the UniProt human proteome, release 2015_12. Additional options for match between runs, LFQ, and iBAQ were selected. The msVolcano Shiny application was used to produce volcano plots for GFP-affinity purification experiments (Singh et al. 2016). Stoichiometry calculations were produced essentially as described (Smits et al. 2013) using Perseus version 1.4.0.8 and in-house R scripts.

Chromatin preparation

Attached HeLa cells were double cross-linked, first with DSG (ThermoFisher) for 40 min, then washed with PBS followed by treatment with 1% formaldehyde in PBS for 10 min at room temperature with gentle shaking. Crosslinking was quenched with the addition of 1/10 volume 1.25 M glycine. Cells were washed with PBS, then harvested by scraping in buffer B (20 mM HEPES, 0.25 % Triton X-100, 10 mM EDTA, and 0.5 mM EGTA). Cells were pelleted by centrifugation at 600 g for 5 min at 4 °C. Cell pellets were resuspended in buffer C (150 mM NaCl, 50 mM HEPES, 1 mM EDTA,

and 0.5 mM EGTA) and rotated for 10 min at 4 °C. Cells were pelleted by centrifugation at 600 g for 5 min at 4 °C. The cell pellet was then resuspended in 1X incubation buffer (0.15% SDS, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES) at 15 million cells/mL. Cells were sheared in a Bioruptor Pico sonicator (Diagenode) at 4 °C with 7 cycles of 30 s on, 30 s off. Sonicated material was spun at 18,000 g for 10 min at 4 °C, then divided into aliquots and stored at -80 °C.

Chromatin immunoprecipitation

10 million cells were used as an input material. Chromatin was incubated overnight at 4 °C in 1X incubation buffer (0.15 % SDS, 1 % Triton X-100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES) supplemented with protease inhibitors and 0.1 % BSA. Antibody amounts and catalog numbers are listed in Supplementary Table 2. A 50:50 mix of Protein A and G Dynabeads (Invitrogen) were added the next day and incubated for 90 min. The beads were washed twice with wash buffer 1 (0.1 % SDS, 0.1 % sodium deoxycholate, 1 % Triton X-100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES), once with wash buffer 2 (wash buffer 1 with 500 mM NaCl), once with wash buffer 3 (250 mM LiCl, 0.5% sodium deoxycholate, 0.5 % NP-40, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES), and twice with wash buffer 4 (1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES). After washing steps, beads were rotated for 20 min at room temperature in elution buffer (1 % SDS and 0.1 M NaHCO₃). The supernatant was de-crosslinked with 200 mM NaCl and 100 µg/mL proteinase K for 4 h at 65 °C. De-crosslinked DNA was purified with MinElute PCR Purification columns (Qiagen). DNA amounts were determined with Qubit fluorometric quantification (ThermoFisher Scientific).

Chromatin immunoprecipitation sequencing and data analysis

Libraries were prepared with a Kapa Hyper Prep Kit for Illumina sequencing (Kapa Biosystems) according to the manufacturer's protocol with the following modifications. 5 ng DNA was used as input, with NEXTflex adapters (Bioo Scientific) and ten cycles of PCR amplification. Post-amplification cleanup was performed with QIAquick MinElute columns (Qiagen), and size selection was performed with an E-gel (300-bp fragments) (ThermoFisher Scientific). Size-selected samples were analyzed for purity with a High Sensitivity DNA Chip on a Bioanalyzer 2100 system

(Agilent). Samples were sequenced on an Illumina HiSeq2000 or NextSeq500. Reads were mapped to the reference human genome hg19 with the Burrows–Wheeler Alignment tool (BWA), allowing one mismatch. Only uniquely mapped reads were used for data analysis and visualization. Each ChIP-seq bam file was first
5 converted to tag align files using gawk and bedtools and peak calling was performed using MACS (version 2.1.0). Afterwards, to get a set of confident peaks, only peaks with Irreproducible Discovery Rate (IDR) < 0.05 were kept.

CAGE library preparation, sequencing and mapping

10 HeLa S3 cells were grown on 60 mm plates and treated with DMSO or I-BRD9 (10 μ M) 6 h prior to induction with EGF (100 ng/ml) for 30 min. Total RNA was isolated using TRI Reagent® (Ambion) according to manufacturer's recommendations. RNA from each of the biological triplicates were quality controlled using a Bioanalyzer. RIN scores were between 9.6 and 10. CAGE libraries were prepared using the
15 protocol by (Takahashi et al. 2012) with an input of 3 μ g of total RNA. Prior to sequencing, four CAGE libraries with different barcodes were pooled and applied to the same sequencing lane. Libraries were sequenced on a Illumina HiSeq 2000 at the National High-Throughput DNA Sequencing Centre, University of Copenhagen. To compensate for the low complexity of 5' ends in the CAGE libraries, 30% Phi-X
20 spike-ins were added to each sequencing lane, as recommended by Illumina. CAGE reads were assigned to their respective originating sample according to identically matching barcodes. Using the FASTX Toolkit, assigned reads were 5'-end trimmed to remove linker sequences (9+2 bp to account for the CAGE protocol G-bias), 3'-end trimmed to a length of 25 bp, and filtered for a minimum sequencing quality of
25 Q30 in 50% of the bases. Reads matching to reference rRNA sequences were discarded using rRNA dust. Mapping to the human genome (hg19) was performed using BWA (version 0.7.10). Only the 5' ends of mapped reads were considered in subsequent analyses.

30 ATAC-seq library preparation and processing

ATAC-seq was performed on approximately 50,000 cells as described in (Buenroostro et al. 2015) with three modifications. First, the total volume of the tagmentation reaction with in-house made Tn5 enzyme was halved. Second, the tagmentation reaction was stopped with 44 mM EDTA, 131 mM NaCl, 0.3% SDS, and 600 μ g/ml

proteinase K. Lastly, a reverse-phase 0.65× SPRI beads (Ampure) DNA purification was done after the first PCR. Libraries were sequenced on an Illumina HiSeq 2000. Paired-end 50-bp sequencing reads were aligned to hg19 with BWA (version 0.7.10) allowing one mismatch. To produce a consensus set of ATAC-seq peaks, bam files were first converted to tag align files using gawk and bedtools and peak calling was performed on all the samples (pooled and individually) using MACS (version 2.1.0) after shifting tag alignments (+4 bp for plus strand and -5bp for minus strand) to account for TN5 insertion. Afterwards, narrow peaks were identified from the pooled list overlapping at least two individual peak lists by at least 50% of bps (FDR < 1%). This resulted in a preliminary list of 229,091 peaks.

Open chromatin loci as focus points for transcription initiation and expression quantification

Open chromatin loci (also referred to as NFRs) were used as focus points for characterizing transcription initiation events as described elsewhere (Andersson et al. 2014), with minor modifications. Instead of focusing on DNase-seq signal summits, center points were defined from ATAC-seq peak signal summits. Open chromatin loci were filtered to not overlap any other open chromatin loci strand-specifically with respect to these windows. This resulted in a final set of 213,017 well-defined open chromatin loci. NFR-associated expression were quantified by counting of CAGE tags in genomic windows of 300 bp immediately flanking ATAC-seq peak summits. An average of 79% of all CAGE tags were covered by the filtered set of open chromatin loci.

For robust assessment of lowly expressed loci, CAGE genomic background noise levels were estimated as described elsewhere (Rennie et al. 2018). First, the CAGE mappability of the hg19 reference genome was calculated by mapping each 25-sized subsequence of the reference genome back to itself, using the same mapping approach as for real CAGE data. Then, the number of CAGE 5' ends from each CAGE library mapping to each of two strand-specific genomic windows genomic regions of size 300 bp was quantified, similarly to the expression quantification of NFRs (above). Genomic windows were required to be uniquely mappable (as determined by the mappability track) in at least 50% of its potential TSS positions (unique bps). Regions that were proximal (within 500bp) of GENCODE (v19) gene

TSSs, transcript ends, or midpoints of ENCODE DHSs (ENCODE January 2011 integration data), or overlapping GENCODE gene exons were discarded. Based on the empirical distribution of CAGE expression noise from annotation-distal genomic regions, the 99th percentile for each library was used as a threshold to call regions significantly expressed in subsequent analyses, if fulfilled in at least 2 out of 3 replicates. This resulted in a set of 20,303 and 22,771 open chromatin loci expressed in the I-BRD9 and DMSO time courses respectively.

Differential expression analysis was performed on all expressed open chromatin loci using DESeq2. To find differential dynamic open chromatin, comparisons were made between all time points $T>0$ and $T=0$. This resulted in a total of 1,709 differentially dynamic open chromatin loci ($FDR<5\%$, $|\log_2 FC| > 1$) in I-BRD9 and 1,922 in DMSO (control) time course. Differential expression analysis was also performed between conditions on the same time points, resulting in 5,633 open chromatin loci with expression differences in at least one time point, of which 710 were also differentially dynamic.

Data availability

Sequencing data (ChIP-seq, ATAC-seq, CAGE) generated in this study have been deposited in GEO under accession number GSE121351. Proteomics data generated in this study have been deposited in the PRIDE archive of the ProteomeXchange consortium under accession number PXD011376.

Acknowledgements

We would like to thank Ragnhild Eskeland for rewarding discussions and for kindly sharing the pEF1neo-TY1-tag plasmid. We would also like to thank members of the Vermeulen, Aasland, and Andersson labs for insightful comments and suggestions. Research in Vermeulen lab was supported by 4DCellFate within the European Commission's Seventh Framework Program (277899), and the European Research Council (StG no. 309384)., The Vermeulen lab is part of the Oncode Institute, which is partly funded by the Dutch Cancer Society (KWF). Research in the Aasland lab was supported by the Norwegian Cancer Society (#197477). Research in the Andersson lab was supported by Independent Research Fund Denmark (6108-

00038B) and the European Research Council under the European Union's Horizon 2020 research and innovation programme (StG no. 638173).

Author contributions

K.J. and R.Aa conceived the project. M.V., R.Aa, and R.An. supervised the project.
5 K.J., S.L.K., M.V., and R.Aa. designed proteomics experiments and K.J., M.V.,
R.Aa., and R.An. designed sequencing experiments. K.J. and S.L.K. performed
mass spectrometry experiments. S.V. performed qPCR experiments and MS data
validation by immunoblotting. S.A.S. performed immunofluorescence experiments.
S.L.K. performed ATAC-seq experiments. S.V. and C.D.V. performed CAGE
10 experiments. K.J. and S.L.K. performed CHIP-seq experiments. K.J. and S.L.K.
analyzed mass spectrometry data. N.A. analyzed ATAC, CHIP, and CAGE
sequencing data and performed integrative bioinformatics analyses. K.J. and R.An.
wrote the manuscript with input from all authors.

References

- 15 Alpsy, A. & Dykhuizen, E.C., 2018. Glioma tumor suppressor candidate region
gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin
remodeling subcomplexes. *Journal of Biological Chemistry*, 293(11), pp.3892–
3903.
- 20 Alver, B.H. et al., 2017. The SWI/SNF chromatin remodelling complex is required for
maintenance of lineage specific enhancers. *Nature Communications*, 8,
p.14648.
- Andersson, R. et al., 2014. An atlas of active enhancers across human cell types
and tissues. *Nature*, 507(7493), pp.455–61.
- 25 Baymaz, H.I. et al., 2014. MBD5 and MBD6 interact with the human PR-DUB
complex through their methyl-CpG-binding domain. *Proteomics*, 14(19),
pp.2179–2189.
- Beagrie, R.A. & Pombo, A., 2016. Gene activation by metazoan enhancers: Diverse
mechanisms stimulate distinct steps of transcription. *BioEssays*, 38(9), pp.881–
893.
- 30 Buenrostro, J.D. et al., 2015. ATAC-seq: A method for assaying chromatin
accessibility genome-wide. *Current Protocols in Molecular Biology*,
2015(January), p.21.29.1-21.29.9.
- Cai, Y. et al., 2007. YY1 functions with INO80 to activate transcription. *Nature
Structural & Molecular Biology*, 14(9), pp.872–874.

- Chandrasekaran, R. & Thompson, M., 2007. Polybromo-1-bromodomains bind histone H3 at specific acetyl-lysine positions. *Biochemical and Biophysical Research Communications*, 355(3), pp.661–666.
- Chandy, M. et al., 2006. SWI/SNF displaces SAGA-acetylated nucleosomes. *Eukaryotic Cell*, 5(10), pp.1738–1747.
- 5 Chen, L. et al., 2011. Subunit organization of the human INO80 chromatin remodeling complex: An evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *Journal of Biological Chemistry*, 286(13), pp.11283–11289.
- 10 Creighton, M.P. et al., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), pp.21931–21936.
- Dignani, J.D., Lebovitz, R.M. & Roeder, R.G., 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research*, 11(5), pp.1475–1489.
- 15 Doyon, Y. et al., 2004. Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Molecular and cellular biology*, 24(5), pp.1884–96.
- 20 Dynlacht, B.D., Hoey, T. & Tjian, R., 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell*, 66(3), pp.563–576.
- Engelen, E. et al., 2015. Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nature communications*, 6(May), p.7155.
- 25 Ernst, J. et al., 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, 473(7345), pp.43–49.
- Hargreaves, D.C. & Crabtree, G.R., 2011. ATP-dependent chromatin remodeling : genetics , genomics and mechanisms. *Nature Publishing Group*, 21(3), pp.396–420.
- 30 Heintzman, N.D. et al., 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459(7243), pp.108–112.
- Heinz, S. et al., 2015. The selection and function of cell type-specific enhancers. *Nature reviews. Molecular cell biology*, 16(3), pp.144–54.
- 35 Hilton, I.B. et al., 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature biotechnology*, 33(5), pp.510–7.
- Hodges, H.C. et al., 2018. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. *Nature Structural and Molecular Biology*, 25(1), pp.61–72.
- 40 Hohmann, A.F. & Vakoc, C.R., 2014. A rationale to target the SWI/SNF complex for cancer therapy. *Trends in Genetics*, 30(8), pp.356–363.
- Jenuwein, T., 2001. Translating the Histone Code. *Science*, 293(5532), pp.1074–1080.

- Ji, X. et al., 2015. Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proceedings of the National Academy of Sciences of the United States of America*, 112(12), pp.3841–6.
- Jin, J. et al., 2005. A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *Journal of Biological Chemistry*, 280(50), pp.41207–41212.
- Kloet, S.L. et al., 2016. The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nature Structural & Molecular Biology*, 23(7), pp.682–690.
- 10 Knutson, B.A. et al., 2016. Super elongation complex contains a TFIIIF-related subcomplex. *Transcription*, 7(4), pp.133–140.
- Kouzarides, T., 2007. Chromatin Modifications and Their Function. *Cell*, 128(4), pp.693–705.
- Kundaje, A. et al., 2015. Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539), pp.317–330.
- 15 Lai, F. et al., 2015. Integrator mediates the biogenesis of enhancer RNAs.
- Lenhard, B., Sandelin, A. & Carninci, P., 2012. Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. *Nature Reviews Genetics*, 13(4), pp.233–245.
- 20 Li, B., Carey, M. & Workman, J.L., 2007. The Role of Chromatin during Transcription. *Cell*, 128(4), pp.707–719.
- Li, Y. et al., 2014. AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell*, 159(3), pp.558–571.
- Lin, C. et al., 2010. AFF4, a Component of the ELL/P-TEFb Elongation Complex and a Shared Subunit of MLL Chimeras, Can Link Transcription Elongation to Leukemia. *Molecular Cell*, 37(3), pp.429–437.
- 25 Luo, Z., Lin, C. & Shilatifard, A., 2012. The super elongation complex (SEC) family in transcriptional control. *Nature Reviews Molecular Cell Biology*, 13(9), pp.543–547.
- 30 Nakayama, R.T. et al., 2017. SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nature Genetics*, (August).
- Phelan, M.L. et al., 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Molecular Cell*, 3(2), pp.247–253.
- 35 Rada-Iglesias, A. et al., 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, 470(7333), pp.279–83.
- Rahman, S. et al., 2011. The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Molecular and cellular biology*, 31(13), pp.2641–52.
- 40 Rajagopal, N. et al., 2014. Distinct and Predictive Histone Lysine Acetylation Patterns at Promoters, Enhancers, and Gene Bodies. *Genes & Genomes Genetics*, 4(11), pp.2051–2063.
- Rao, S.S.P. et al., 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7), pp.1665–1680.

- Rappsilber, J., Mann, M. & Ishihama, Y., 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, 2(8), pp.1896–1906.
- Rennie, S. et al., 2018. Transcription start site analysis reveals widespread divergent transcription in *D. melanogaster* and core promoter-encoded enhancer activities. *Nucleic Acids Research*, (April), pp.1–15.
- 5 Sanyal, A. et al., 2012. The long-range interaction landscape of gene promoters. *Nature*, 489(7414), pp.109–113.
- Schwanhäusser, B. et al., 2011. Global quantification of mammalian gene expression control. *Nature*, 473(7347), pp.337–342.
- 10 Shevchenko, A. et al., 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, 1(6), pp.2856–2860.
- Sidoli, S. et al., 2015. Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH) Analysis for Characterization and Quantification of Histone Post-translational Modifications. *Molecular cell Proteomics*, 14(9), pp.2420–2428.
- 15 Singh, S., Hein, M.Y. & Stewart, A.F., 2016. msVolcano: A flexible web application for visualizing quantitative proteomics data. *Proteomics*, 16(18), pp.2491–2494.
- Smits, A.H. et al., 2013. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. *Nucleic Acids Research*, 41(1), pp.1–8.
- 20 Takahashi, H. et al., 2012. 5' End-Centered Expression Profiling Using Cap-Analysis Gene Expression and Next-Generation Sequencing. *Nature Protocols*, 7(3), pp.542–561.
- 25 Theodoulou, N.H. et al., 2016. Discovery of I-BRD9, a Selective Cell Active Chemical Probe for Bromodomain Containing Protein 9 Inhibition. *Journal of Medicinal Chemistry*, 59(4), pp.1425–1439.
- Tie, F. et al., 2009. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development*, 136(18), pp.3131–3141.
- 30 Tvardovskiy, A. et al., 2015. Top-down and middle-down protein analysis reveals that intact and clipped human histones differ in post-translational modification patterns. *Molecular & cellular proteomics : MCP*, 14(12), pp.3142–3153.
- Vermeulen, M., 2012. *Identifying chromatin readers using a SILAC-based histone peptide pull-down approach* 1st ed., Elsevier Inc.
- 35 Vermeulen, M. et al., 2007. Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell*, 131(1), pp.58–69.
- Vernimmen, D. & Bickmore, W.A., 2015. The Hierarchy of Transcriptional Activation: From Enhancer to Promoter. *Trends in Genetics*, 31(12), pp.696–708.
- 40 Wang, W. et al., 1996. Diversity and specialization of mammalian SWI / SNF complexes. *Genes & Development*, 4(17), pp.2117–2130.
- Wang, X. et al., 2016. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nature Genetics*, (December).
- Wang, Z. et al., 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics*, 40(7), pp.897–903.

Yao, T. et al., 2008. Distinct Modes of Regulation of the Uch37 Deubiquitinating Enzyme in the Proteasome and in the Ino80 Chromatin-Remodeling Complex. *Molecular Cell*, 31(6), pp.909–917.

5 Yu, Y. et al., 2013. Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell*, 152(1–2), pp.248–261.

Supplementary materials and methods for:

The GBAF chromatin remodeling complex binds H3K27ac and mediates enhancer transcription

Kirill Jefimov^{1,#}, Nicolas Alcaraz^{2,#}, Susan L. Kloet^{3,#}, Signe Värnv⁴, Siri Aastedatter Sakya¹, Christian Dalager Vaagenso², Michiel Vermeulen^{3,*}, Rein Aasland^{4,*}, and Robin Andersson^{2,*}.

¹ Department of Biological Sciences, University of Bergen, 5008 Bergen, Norway, ² The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark,

³ Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Radboud University Nijmegen, 6525 Nijmegen, The Netherlands, ⁴ Department of Biosciences, University of Oslo, 0371 Oslo, Norway

#: equal contribution

*: to whom correspondence should be addressed

(M.V.: michiel.vermeulen@science.ru.nl, R.Aa.: rein.aasland@ibv.uio.no, R.An.: robin@binf.ku.dk)

Contents

Supplementary Figures 1-5

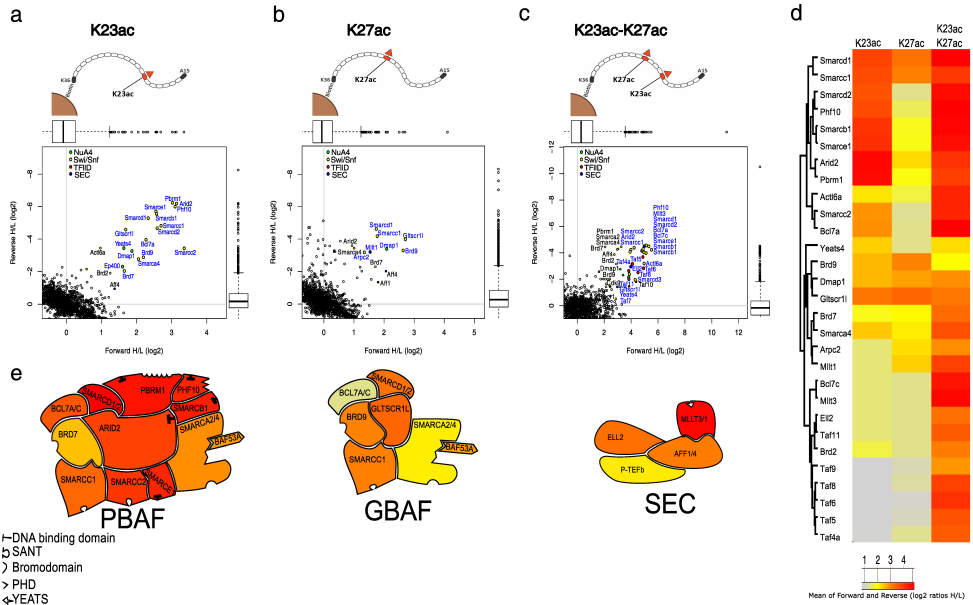
Supplementary Tables 1-4

Supplementary Data legend 1

Supplementary Methods

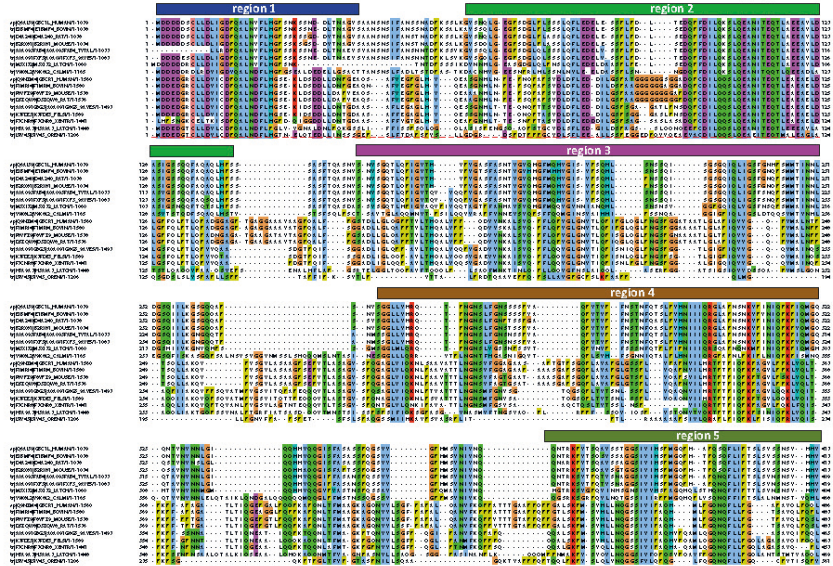
Supplementary References

Supplementary Figures

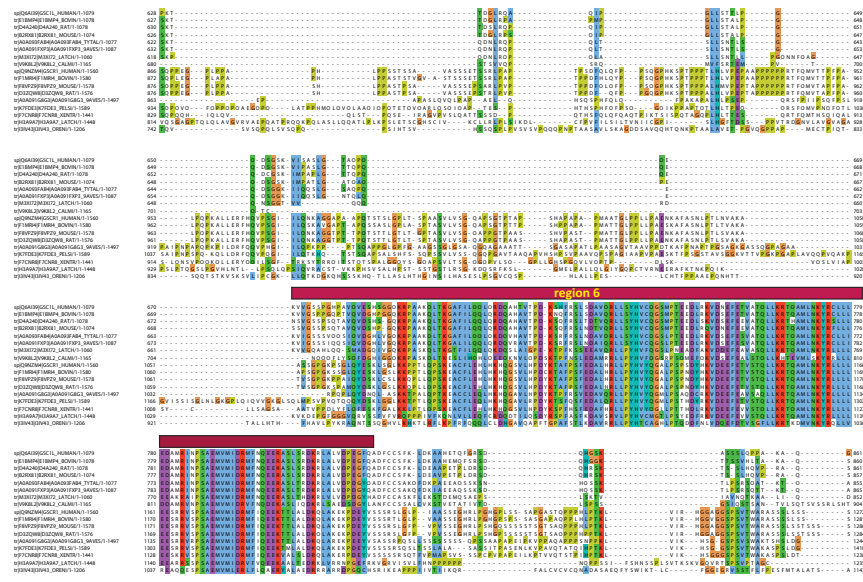
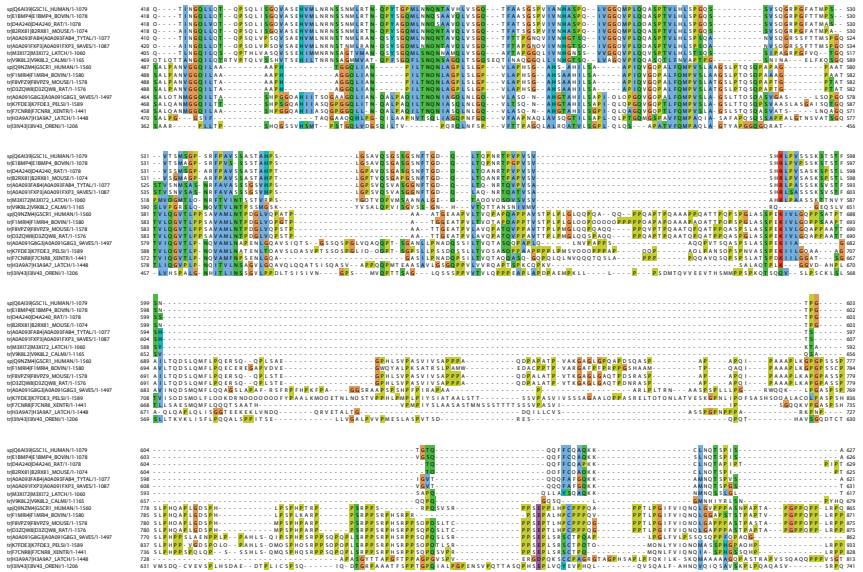


Supplementary Figure 1. SILAC histone peptide pulldowns identify SEC and SWI/SNF complexes to be major interactors of H3K23ac and H3K27ac in mESCs. **a-c** Scatter plots of forward (horizontal axes) versus reverse (vertical axes) SILAC ratios from histone peptide pulldowns using **(a)** H3K23ac, **(b)** H3K27ac, or **(c)** diacetylated H3K23acK27ac baits with mESC nuclear extracts (n=2 pulldowns for each bait). The lower and upper hinges of boxes correspond to the first and third quartiles of data, respectively, and the whiskers extend to the largest and smallest data points no further away than 1.5 times the interquartile range. **d** Heatmap of average forward and reverse SILAC ratios (log₂). The ordering of rows in the heatmap and the associated dendrogram were derived from agglomerative hierarchical clustering of the H3K23- and K27-interacting proteins identified in the histone peptide pulldowns **(a-c)**, using Euclidean distances. **e** Visual summary of the most enriched interactors of H3K23ac (left), H3K27ac (middle), and H3K23acK27ac (right) peptide pulldowns. Subunits are colored according to their enrichment in the pulldowns.

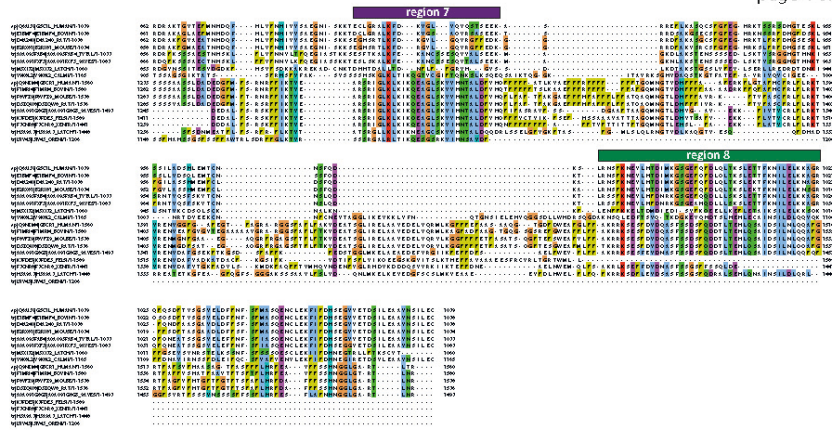
GLTSCR1L family multiple sequence alignment



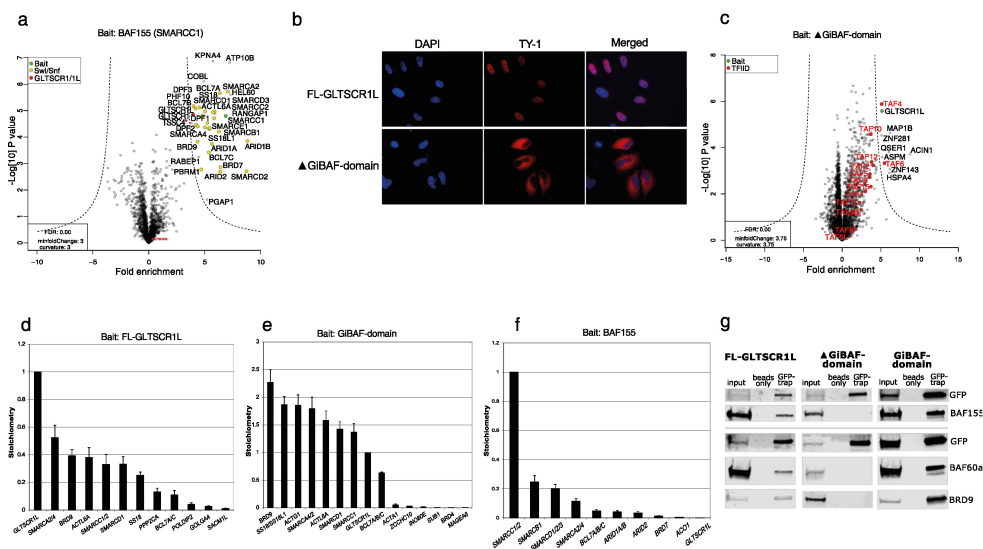
Supplementary Figure 2 (page 1 of 4). GLTSCR1L family multiple sequence alignment. Multiple sequence alignment of a selected set of full-length sequences of the vertebrate GLTSCR1L-family. The alignment was generated with Clustal Omega color coded with the ClustalX coloring scheme. The upper and lower half of the alignment contains GLTSCR1L and GLTSCR1 homologs, respectively. Conserved regions (R1-R8) are indicated with colored bars. The most strongly conserved region (R6) corresponds to the Pfam family PF15249 (Conserved region of unknown function on GLTSCR protein) and termed 'GiBAF'-domain (GLTSCR1/1L domain interacting with BAF complex). Database sources, accession numbers and Uniprot IDs are shown. GLTSCR1L and GLTSCR1 are GSC1L_HUMAN and GSCR1_HUMAN, respectively.



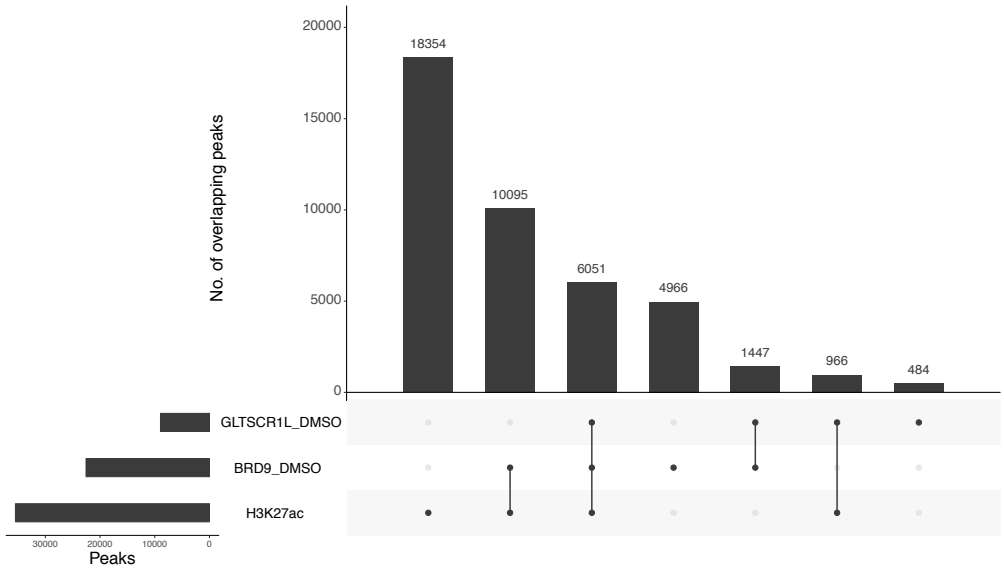
Supplementary Figure 2 (page 2 and 3 of 4).



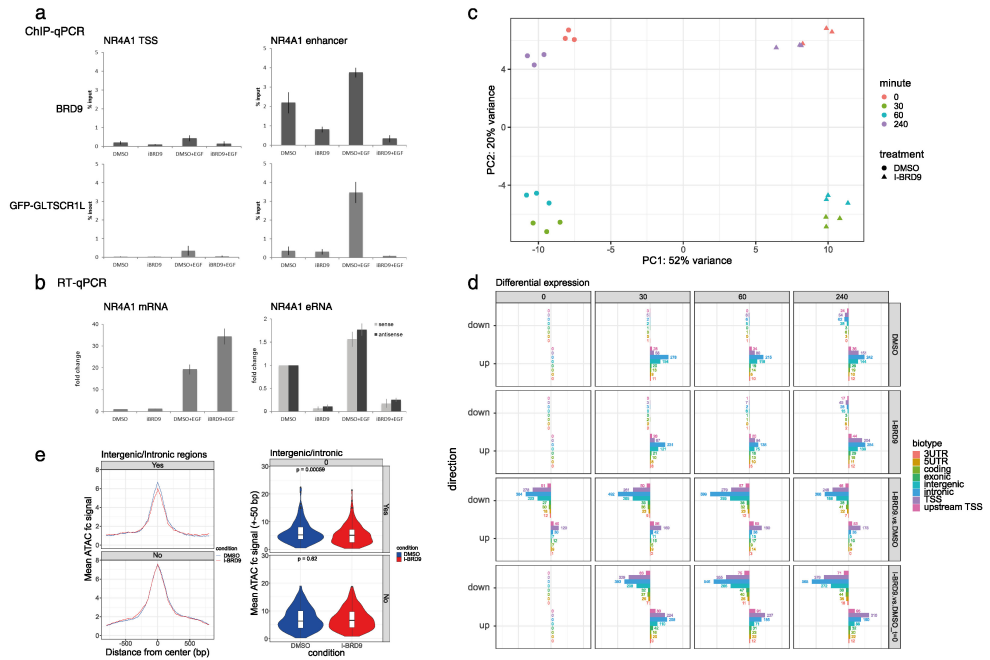
Supplementary Figure 2 (page 4 of 4).



Supplementary Figure 3. GiBAF-domain facilitates GLTSCR1L nuclear localization and interaction with BAF complex. **a** Quantitative MS analysis of GFP-BAF155 protein and its interactors isolated from HeLa nuclear extracts. Volcano plots displaying fold enrichments (\log_2 GFP fusion protein over mock, horizontal axes) versus t-test P values ($-\log_{10}$, vertical axes) from three independent pulldowns ($n=3$). **b** Immunofluorescence of the intracellular localization of wild-type (FL-GLTSCR1L) and mutant (Δ GiBAF-domain) GLTSCR1L in HeLa cells. Cell nuclei (blue) were stained with DAPI and GLTSCR1L proteins (red) were visualized using anti-TY1 antibody followed by treatment with anti-mouse Alexa Fluor 594. Cells were permeabilized 48 hr after transfection. 100x magnification. **c** Interactors of affinity-purified GFP- Δ GiBAF-domain isolated from HeLa whole cell extracts. Statistical analysis as in (a). **d-f** Stoichiometry of GFP-FL-GLTSCR1L (d), GFP-GiBAF-domain (e) and GFP-BAF155 (f) interactors. The intensity-based absolute quantification (iBAQ) value of each protein is divided by the iBAQ value of the bait and plotted relative to a bait value of 1. Data are shown as mean \pm s.d. ($n = 3$ pulldowns). **g** Validation of GFP-affinity enrichment-MS data with immunoblotting. Nuclear or whole cell extracts from doxycycline-inducible HeLa-FRT cells expressing GFP-tagged FL-GLTSCR1L or its mutants (Δ GiBAF-domain and GiBAF-domain only) were incubated either with empty or GFP-nanotrap conjugated beads. Input (10 %) and eluate fractions were analyzed after SDS-PAGE by immunoblotting.



Supplementary Figure 4. Peak overlap analysis shows strong co-localization of BRD9, GLTSCR1L and H3K27ac. Histogram attribute (upset) plot showing the number of overlapping peaks between BRD9, GLTSCR1L, and ENCODE HeLa H3K27ac. Two peaks were considered overlapping if their intersection is at least 1bp.



Supplementary Figure 5. BRD9 bromodomain inhibition leads to downregulation of EGF-inducible enhancer transcription.

a Chromatin immunoprecipitation performed with nuclear extracts from doxycycline-induced HeLa-FRT-GFP-GLTSCR1L cells. BRD9 and GLTSCR1L protein occupancy was detected on NR4A1 gene promoter and enhancers before and after treatment with BRD9 bromodomain inhibitor I-BRD9 (10 μ M) and 30 min induction of transcription with EGF (100 ng/ml). Error bars indicate standard error of means of three independent experiments. **b** NR4A1 promoter RNA (TSS) and enhancer RNA (eRNA) levels in cells treated as in (a). RNA expression values were normalized against *Actinb* and are given relative to values in cells treated with DMSO. Error bars indicate standard error of means of four (eRNA) or three (TSS) independent experiments. **c** Principal Component Analysis plot of the top 500 NFRs with most variance based on CAGE expression. **d** Number of differentially expressed NFRs for DMSO and I-BRD9 time courses (T>0 vs T=0), between I-BRD9 vs DMSO (all time points), and between I-BRD9 T>0 vs DMSO T=0. **e** Left figures: Mean ATAC-seq signal over intronic/intergenic THS regions that are down-regulated (according to CAGE expression) and have Loss of GLTSCR1L (Yes panel) and No loss of GLTSCR1L (No panel). Right figures: Mean Signal ATAC-seq summarized signal over +50 bp around THS center (Same regions as left figures). Wilcoxon-test was performed between DMSO and I-BRD9 conditions.

Supplementary Tables

Supplementary Table 1. BAF subunits identified as interactors of SMARCC1 (1), SMARCD1 (2), BRD9 (3), GLTSCR1L (4), or GLTSCR1L GiBAF-domain (5).

Gene name	BAF subunit	Chromatin/DNA interaction domain	SWI/SNF Complex	Bait
SMARCA2	BRM	DNA-dependent ATPase, bromodomain	BAF/PBAF/GBAF	1,2,3,4,5
SMARCA4	BRG1	DNA-dependent ATPase, bromodomain	BAF/PBAF/GBAF	1,2,3,4,5
SMARCC1	BAF155	No	BAF/PBAF/GBAF	2,3,4,5
SMARCC2	BAF170	No	BAF/PBAF	1,2
SMARCD1	BAF60A	No	BAF/PBAF/GBAF	1,3,4,5
SMARCD2	BAF60B	No	BAF/PBAF	1
SMARCD3	BAF60C	No	BAF/PBAF	1
SMARCB1	BAF47	SNF5	BAF/PBAF	1,2
SMARCE1	BAF57	HMG box	BAF/PBAF	1,2
ACTL6A	BAF53A	No	BAF/PBAF/GBAF	1,3,5
BCL7A	BCL7A	No	BAF/PBAF/GBAF	5
BCL7B	BCL7B	No	BAF/PBAF/GBAF	1,2,3,5
BCL7C	BCL7C	No	BAF/PBAF/GBAF	1,2,3,5
PBRM1	BAF180	bromodomain, HMG, zing finger	PBAF	1,2
ARID2	BAF200	ARID/BRIGHT	PBAF	1,2
PHF10	BAF45A	zinc finger	PBAF	1,2
BRD7	BRD7	bromodomain	PBAF	1,2
DPF1	BAF45B	zinc finger	BAF	1,2
DPF2	BAF45C	zinc finger	BAF	1,2
DPF3	BAF45D	zinc finger	BAF	1,2
ARID1A	BAF250A	ARID/BRIGHT	BAF	1,2
ARID1B	BAF250B	ARID/BRIGHT	BAF	1,2
BRD9	BRD9	bromodomain	BAF/GBAF	1,2,4,5
GLISTR1L	BICRAL	No	GBAF	1,2,3
GLTSCR1	BICRA	No	GBAF	1,2,3
SS18	SSXT	No	BAF/GBAF	1,2,3,4,5
SS18L1	CREST	No	BAF/GBAF	2,5
SS18L2	SS18L2	No	BAF/GBAF	5

GFP-Pulldown interaction baits: 1-SMARCC1; 2-SMARCD1; 3-BRD9; 4-GLTSCR1L; 5-GiBAF-domain

Supplementary Table 2. Antibodies used.

ANTIBODY	SOURCE	DILUTION for immunoblotting	Amount per ChIP
anti-BRD9 ; rabbit polyclonal	Active Motif 61537, 61538	1:1000	2 μ l
anti-BAF60a ; clone 23/BAF60A, mouse monoclonal	BD Transduction Laboratories, 611728	1:1000	-
anti-SMARCC1/BAF155 ; rabbit monoclonal	Cell Signaling, 11956	1:1000	-
anti-PBRM1 ; rabbit polyclonal	Sigma, HPA015629	1:250	-
MPC2 (T-20) (anti-CBX4); goat polyclonal	Santa Cruz, sc-19299	1:1000	-
anti-GFP ; rabbit polyclonal	Abcam, ab290	1:2500	2 μ l
anti-GFP (JL-8) ; mouse monoclonal	Takara/Living colours, 632380	1:1000	-
anti-Ty1 Tag ; mouse monoclonal	Diagenode (Thermo Scientific), MA5-23513	1:5000	-
Normal Rabbit IgG ; polyclonal	Millipore, 12-370	-	2 μ g

Supplementary Table 3. Primers used for RT-qPCR

NR4A1-S-qfor1/eRNA sense (<i>Lai et al 2015</i>)	cctccttctaagcctgaact
NR4A1-S-qrev2/eRNA sense (<i>Lai et al 2015</i>)	tgctccctggaacagagat
NR4A1-As-qfor1/eRNA antisense (<i>Lai et al 2015</i>)	ttcctcacaacacctactcc
NR4A1-As-qrev1/eRNA antisense (<i>Lai et al 2015</i>)	ccctccgtcactctcaaatg
NR4A1 codseq F1/mRNA	agaagatccctggcttgct
NR4A1 codseq R1/mRNA	cagggacatcgacaagcaag
actin-F	ctacaatgagctgcgtgtggc
actin-R	caggtccagacgcaggatggc

Supplementary Table 4. Primers used for ChIP qPCR

NR4A1 R4 prom F4/promoter region	catcagcattacagtcaccctt
NR4A1 R4 prom R4/promoter region	cacgtttgaactgtgtaggcca
R_ENH_F/NR4A1 enhancer	tgggtgtgcctgtatgtgac
R_ENH_R/NR4A1 enhancer	ctgtgagtgtggcgggtgat

Supplementary Data legend

Supplementary Data 1.

Acetyl-lysine interacting proteins identified by SILAC histone-peptide pulldown MS. All proteins that were not found in both (Forward and Reverse) pulldowns were removed. Proteins are sorted according to their SILAC H/L ratio in the forward pulldown. Outliers are labelled as significant and colored green. Proteins that were close to outliers, and potentially interacting with acetylated-lysines are colored yellow.

Supplementary Methods

RNA extraction, reverse transcription, qPCR

HeLa S3 cells were seeded on 6-well plates, treated with DMSO or I-BRD9 (10 μ M) for 6 h before 30 min induction with EGF (100 ng/ml). Total RNA was isolated using TRI Reagent® (Ambion) according to manufacturer's recommendations. cDNA synthesis was performed in 50 μ l volume from 3–4 μ g of total RNA using 100 U M-MuL V Reverse Transcriptase (Thermo Scientific) and 200 pmol of Random Hexamers (Thermo Scientific) in the presence of 20 U of RiboLock (Thermo Scientific). For one quantitative PCR reaction in LightCycler®96 (Roche Life Science), 2 μ l of cDNA and 5x HOT FIREPol®EvaGreen®qPCR Supermix (Solis BioDyne) was used. NR4A1 TSS RNA and eRNA values were calculated by normalization to *Actb*, using the comparative CT method. Primers used are listed in Supplementary Table 3.

ChIP-qPCR

For one chromatin immunoprecipitation sample, nuclear extract (NE) from 15 million HeLa-FRT-GFP-GLTSCR1L cells after 16 h induction with doxycycline (1 μ g/ml) was collected. ChIP was performed as described in (Kloet et al. 2016) with minor modifications. Briefly, cells were washed twice with 1X PBS; cross-linked for 40 min with 2,5 mM DSG and after washing with 1 X PBS, 10 min with 1% formaldehyde. Cross-linking was quenched with the addition of 1/10 volume of 2 M glycine. After washing with 1 X PBS, cells were collected in Buffer B (20 mM HEPES, 0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA), pelleted by centrifugation at 600 g for 5 min at 4 °C. Cell pellets were resuspended in buffer C (150 mM NaCl, 50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA) and incubated for 10 min with rotation at 4 °C. Cells were pelleted as before and resuspended in incubation buffer (0.15% SDS, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES, 1X Roche protease inhibitor cocktail) at 15 million cells/ml. Chromatin was sheared with Bioruptor Pico sonicator (Diagenode) using 6 times the "30 s ON/30 s OFF" cycle. Sonicated material was centrifuged first for 5 min, followed by 15 min centrifugation at 4 °C, 16 000 g. Supernatant was stored at -80 °C.

An aliquot of 1 ml of NE was used as input material and incubated overnight at 4 °C with 2 μ l of anti-BRD9 (Active Motif, 61538), anti-GFP (Abcam, ab290) or normal rabbit IgG (Millipore, 12-370) antibodies. A 50:50 mix of Protein A and G Dynabeads was added and incubated for 90 min at 4 °C with rotation. The beads were washed once with 1,3 ml of wash buffer 1 (0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES), wash buffer 2 (wash buffer 1 with 500 mM NaCl), with wash buffer 3 (50 mM LiCl, 0.5% sodium deoxycholate, 0.5% NP-50, 1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES) and twice with wash buffer 4 (1 mM EDTA, 0.5 mM EGTA, and 20 mM HEPES). DNA-protein complexes were eluted from the beads with 200 μ l of elution buffer (100 mM Tris, pH 7.8, 10 mM EDTA, 1% SDS, 400 mM NaCl). 200 μ l of water with 20 mg of proteinase K was added to the eluate and de-crosslinking took place for 4 h at 65 °C. DNA was extracted using phenol/chloroform. qPCR analysis of ChIP DNA was performed with 5x HOT FIREPol®EvaGreen®qPCR Supermix (Solis BioDyne) on LightCycler® 96 (Roche Life Science). Enrichment values were calculated as percentage of input using comparative CT method. Primers used are listed in Supplementary Table 4.

Immunofluorescence

HeLa-S3 cells were seeded onto glass slides in 6-well plates and transfected with 3 µg of pEF1Neo-FL-GLTSCR1L and pEF1Neo-ΔGiBAF-domain constructs using X-tremeGENE 9 (Sigma-Aldrich) according to manufacturer's instructions. 48 h post-transfection cells were fixed with 4 % ice-cold formaldehyde in 1 X PBS, incubated for 15 min, washed with 1 X PBS and permeabilized with 0,25 % TritonX-100 in PBS at room temperature for 15 min. After washing with 1 X PBS, cells were blocked with DMEM supplemented with 10 % (v/v) FCS at room temperature for 1 h and incubated with primary antibody recognizing TY1 tag at 4 °C overnight. Coverslips were washed twice with 0,1 % TritonX-100 in PBS and 1 X PBS before 1 h incubation at 4 °C with anti-mouse Alexa Fluor 594 secondary antibody. Nuclei were stained with 0,5 µg/ml of DAPI in PBS for 10 min at room temperature and cells were washed as previously described. After washing steps Leica DFC350 FX microscope was used for imaging.

Multiple sequence alignment

Multiple sequence alignment of vertebrate homologs of GLTSCR1 and GLTSCR1L was generated with Clustal Omega (Sievers et al. 2013) with default parameters using the Jalview environment (Waterhouse et al., 2009). The alignment was color coded with the ClustalX coloring scheme (Thompson et al., 1997).

Supplementary References

Kloet SL, Makowski MM, Baymaz HI, van Voorthuijsen L, Karemaker ID, Santanach A, Jansen PWTC, Di Croce L, Vermeulen M. (2016) The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nat Struct Mol Biol* 23:682-690

Lai F, Gardini A, Zhang A, Shiekhattar R. (2015) Integrator mediates the biogenesis of enhancer RNAs. *Nature* 525:399-403

Sievers F, Dineen D, Wilm A, Higgins DG. (2013) Making automated multiple alignments of very large numbers of protein sequences. *Bioinformatics* 29:989-995

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876-4882

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189-1191

Appendix

As described above, GLTSCR1L and BRD9 co-localize with H3K27ac and other regulatory chromatin marks preferentially at intronic/intergenic regions (see Results 3.3). Sequence-specific TFs are also known to bind regulatory elements. Thus, the possible overlap between GLTSCR1L/BRD9-enriched chromatin sites and sequence-specific transcription factor binding motifs could further corroborate our hypothesis that GBAF is associated with genome regulatory elements. Known transcription factor motif enrichment was performed with the package Homer (Heinz *et al.*, 2010) and JASPAR (Fornes *et al.*, 2020) database using default settings on THS regions bound by BRD9 (Appendix A) or GFP-GLTSCR1L (Appendix B) and using all transcribed THS regions as background.

32		RUNX(Runt)/HPC7-Runx1-ChIP-Seq(GSE22178)/Homer	1e-10	-2.395e+01	0.0000	1187.0	7.58%	783.8	6.29%	mol file (mbl)
33		Klf4(Zf)/mES-Klf4-ChIP-Seq(GSE11431)/Homer	1e-10	-2.372e+01	0.0000	1294.0	8.27%	862.8	6.92%	mol file (mbl)
34		Fli1(ETS)/CD8-FL1-ChIP-Seq(GSE20898)/Homer	1e-9	-2.295e+01	0.0000	2892.0	18.48%	2064.0	16.56%	mol file (mbl)
35		Maz(Zf)/HepG2-Maz-ChIP-Seq(GSE31477)/Homer	1e-9	-2.266e+01	0.0000	4085.0	26.10%	2982.9	23.93%	mol file (mbl)
36		Atf3(GZIP)/GfBM-ATF3-ChIP-Seq(GSE33912)/Homer	1e-9	-2.260e+01	0.0000	2566.0	16.39%	1818.0	14.59%	mol file (mbl)
37		Atf2(GZIP)/J3T3L1-Atf2-ChIP-Seq(GSE56872)/Homer	1e-9	-2.254e+01	0.0000	766.0	4.89%	484.0	3.88%	mol file (mbl)
38		ERG(ETS)/CaP-ERG-ChIP-Seq(GSE14097)/Homer	1e-9	-2.235e+01	0.0000	3148.0	20.11%	2263.8	18.16%	mol file (mbl)
39		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq(Unpublished)/Homer	1e-9	-2.108e+01	0.0000	1440.0	9.20%	980.2	7.86%	mol file (mbl)
40		X-box(HTH)/NPC-H3K4me1-ChIP-Seq(GSE16256)/Homer	1e-9	-2.094e+01	0.0000	232.0	1.48%	121.2	0.97%	mol file (mbl)
41		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-8	-1.971e+01	0.0000	1759.0	11.24%	1224.6	9.82%	mol file (mbl)
42		RUNX2(Runt)/PCa-RUNX2-ChIP-Seq(GSE33889)/Homer	1e-8	-1.966e+01	0.0000	1172.0	7.49%	788.3	6.32%	mol file (mbl)
43		Six2(Homeobox)/NephronProgenitor-Six2-ChIP-Seq(GSE39837)/Homer	1e-7	-1.838e+01	0.0000	1677.0	10.71%	1169.3	9.38%	mol file (mbl)
44		Rfx1(HTH)/NPC-H3K4me1-ChIP-Seq(GSE16256)/Homer	1e-7	-1.802e+01	0.0000	396.0	2.53%	236.0	1.89%	mol file (mbl)
45		GABPA(ETS)/Jurkat-GABPa-ChIP-Seq(GSE17954)/Homer	1e-7	-1.622e+01	0.0000	2184.0	13.95%	1564.7	12.55%	mol file (mbl)
46		OCT/OC1(POU1.Homeobox)/NPC-Bmi1-ChIP-Seq(GSE35496)/Homer	1e-7	-1.615e+01	0.0000	11.0	0.07%	1.8	0.01%	mol file (mbl)
47		AP-1(bZIP)/ThoMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-6	-1.605e+01	0.0000	2624.0	16.76%	1901.9	15.26%	mol file (mbl)
48		NRF1(NRF)/MCF7-NRF1-ChIP-Seq(Unpublished)/Homer	1e-6	-1.562e+01	0.0000	822.0	5.25%	547.1	4.39%	mol file (mbl)
49		FOXK1(Forkhead)/HEK293-FOXK1-ChIP-Seq(GSE51673)/Homer	1e-6	-1.496e+01	0.0000	1587.0	10.14%	1119.8	8.98%	mol file (mbl)

67		Smad3(MAD)/NPC-Smad3-ChIP-Seq(GSE36673)/Homer	1e-4	-1.071e+01	0.0001	4196.0	26.81%	3163.9	25.38%	mol file (msl)
68		Nanog(Homeobox)/mES-Nanog-ChIP-Seq(GSE11724)/Homer	1e-4	-1.052e+01	0.0001	5579.0	35.64%	4251.6	34.11%	mol file (msl)
69		Ap4(bHLH)/AML-Tfap4-ChIP-Seq(GSE45738)/Homer	1e-4	-1.030e+01	0.0002	2030.0	12.97%	1486.2	11.92%	mol file (msl)
70		FOXA1(Forkhead)/MCF7-FOXA1-ChIP-Seq(GSE26831)/Homer	1e-4	-1.007e+01	0.0002	1476.0	9.43%	1064.0	8.54%	mol file (msl)
71		Egr2(Zf)/Thymocytes-Egr2-ChIP-Seq(GSE34254)/Homer	1e-4	-9.888e+00	0.0003	528.0	3.37%	354.7	2.85%	mol file (msl)
72		NF1(FOXAI(CTF,Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	1e-4	-9.348e+00	0.0004	145.0	0.93%	83.3	0.67%	mol file (msl)
73		Arnt(Ahr(bHLH)/MCF7-Amt-ChIP-Seq(L_o_et_al)/Homer	1e-3	-9.056e+00	0.0006	1162.0	7.42%	832.4	6.68%	mol file (msl)
74		NF-E2(bZIP)/K562-NFE2-ChIP-Seq(GSE31477)/Homer	1e-3	-8.965e+00	0.0006	332.0	2.12%	215.6	1.73%	mol file (msl)
75		Nr2f2(bZIP)/Lymphoblast-Nrf2-ChIP-Seq(GSE37589)/Homer	1e-3	-8.864e+00	0.0007	270.0	1.73%	171.1	1.37%	mol file (msl)
76		Foxf1(bForkhead)/Lung-Foxf1-ChIP-Seq(GSE77951)/Homer	1e-3	-8.581e+00	0.0009	1422.0	9.09%	1033.2	8.29%	mol file (msl)
77		NRF(NRF)/Promoter/Homer	1e-3	-8.536e+00	0.0009	788.0	5.03%	553.9	4.44%	mol file (msl)
78		MyoG(bHLH)/C2C12-MyoG-ChIP-Seq(GSE36024)/Homer	1e-3	-8.224e+00	0.0013	1733.0	11.07%	1274.5	10.23%	mol file (msl)
79		Tcf21(bHLH)/ArterySmoothMuscle-Tcf21-ChIP-Seq(GSE61369)/Homer	1e-3	-8.117e+00	0.0014	1659.0	10.60%	1218.4	9.78%	mol file (msl)
80		FOXA1(Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	1e-3	-8.034e+00	0.0015	1731.0	11.06%	1274.3	10.22%	mol file (msl)
81		ZNF519(Zf)/HEK293-ZNF519.GFP-ChIP-Seq(GSE5834)/Homer	1e-3	-7.973e+00	0.0015	633.0	4.04%	440.9	3.54%	mol file (msl)
82		PBX1(Homeobox)/MCF7-PBX1-ChIP-Seq(GSE28007)/Homer	1e-3	-7.860e+00	0.0017	120.0	0.77%	69.5	0.56%	mol file (msl)
83		ZNF143(STAF/Zf)/CUTLL-ZNF143-ChIP-Seq(GSE29600)/Homer	1e-3	-7.752e+00	0.0019	598.0	3.82%	416.0	3.34%	mol file (msl)
84		Nr5a2(NR)/mES-Nr5a2-ChIP-Seq(GSE19019)/Homer	1e-3	-7.427e+00	0.0026	632.0	4.04%	443.0	3.55%	mol file (msl)



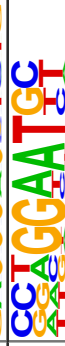














1102		Phx3 (Homeobox)/GMI2878-PBX3-ChIP-Seq (GSE32465)/Homer	1e-2	-5.461e+00	0.0152	317.0	2.03%	217.7	1.75%	mol file (msd)
1103		CLOCK (bHLH)/Liver-Clock-ChIP-Seq (GSE39860)/Homer	1e-2	-5.454e+00	0.0152	913.0	5.83%	667.4	5.35%	mol file (msd)
1104		Foxa3 (Forkhead)/Liver-Foxa3-ChIP-Seq (GSE77670)/Homer	1e-2	-5.436e+00	0.0153	539.0	3.44%	383.5	3.08%	mol file (msd)
1105		E-box (bHLH)/Promoter/Homer	1e-2	-5.354e+00	0.0164	292.0	1.87%	199.2	1.60%	mol file (msd)
1106		NFKB-p50-p52 (RHD)/Monocyte-p50-ChIP-Chip (Schreiber_et_al)/Homer	1e-2	-5.234e+00	0.0183	256.0	1.64%	173.2	1.39%	mol file (msd)
1107		ZNF16 (ZF)/HEK293-ZNF16-GFP-ChIP-Seq (GSE58341)/Homer	1e-2	-5.209e+00	0.0186	10.0	0.06%	3.0	0.02%	mol file (msd)
1108		Egr1 (ZF)/K562-Egr1-ChIP-Seq (GSE32465)/Homer	1e-2	-5.129e+00	0.0200	1641.0	10.48%	1231.2	9.88%	mol file (msd)
1109		Foxh1 (Forkhead)/hESC-FOXH1-ChIP-Seq (GSE29422)/Homer	1e-2	-4.937e+00	0.0240	726.0	4.64%	528.3	4.24%	mol file (msd)
1110		Stat3 (Stat)/mES-Stat3-ChIP-Seq (GSE11431)/Homer	1e-2	-4.863e+00	0.0256	950.0	6.07%	700.6	5.62%	mol file (msd)
1111		Bapx1 (Homeobox)/VertebralCol-Bapx1-ChIP-Seq (GSE36672)/Homer	1e-2	-4.853e+00	0.0256	2607.0	16.66%	1987.2	15.94%	mol file (msd)
1112		NeuroD1 (bHLH)/Islet-NeuroD1-ChIP-Seq (GSE30298)/Homer	1e-2	-4.842e+00	0.0256	1251.0	7.99%	932.6	7.48%	mol file (msd)
1113		NeuroG2 (bHLH)/Fibroblast-NeuroG2-ChIP-Seq (GSE75910)/Homer	1e-2	-4.728e+00	0.0285	2157.0	13.78%	1637.2	13.14%	mol file (msd)
1114		PAX5 (Paired, Homeobox), condensed/GMI2878-PAX5-ChIP-Seq (GSE32465)/Homer	1e-2	-4.727e+00	0.0285	196.0	1.25%	131.1	1.05%	mol file (msd)
1115		FXR (NR)/IRL/Liver-FXR-ChIP-Seq (Chong_et_al)/Homer	1e-2	-4.668e+00	0.0297	554.0	3.54%	399.9	3.21%	mol file (msd)
1116		Nr5a2 (NR)/Pancreas-LRHI-ChIP-Seq (GSE34295)/Homer	1e-2	-4.651e+00	0.0300	820.0	5.24%	602.2	4.83%	mol file (msd)
1117		HRE (HSF)/HepG2-HSF1-ChIP-Seq (GSE31477)/Homer	1e-2	-4.643e+00	0.0300	175.0	1.12%	116.3	0.93%	mol file (msd)











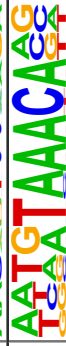






B.

Homer Known Motif Enrichment Results (GFP_GLTSCR1L_DMSO_homer_motifs/)

Homer de novo Motif Results
 Gene Ontology Enrichment Results
 Known Motif Enrichment Results (txt file)
 Total Target Sequences = 8225, Total Background Sequences = 40528


















Rank	Motif	Name	P-value	log P-value	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif	Motif File	SVG
1		BATF(bZIP)/Th1-7-BATF-ChIP-Seq(GSE39756)/Homer	1e-470	-1.083e+03	0.0000	1395.0	16.96%	1544.1	3.81%	motif file (matrix)	svg
2		JunB(bZIP)/DendriticCells-Jumb-ChIP-Seq(GSE36099)/Homer	1e-469	-1.082e+03	0.0000	1276.0	15.51%	1285.8	3.17%	motif file (matrix)	svg
3		Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46166)/Homer	1e-465	-1.071e+03	0.0000	1305.0	15.87%	1361.5	3.36%	motif file (matrix)	svg
4		Atf3(bZIP)/GBM-ATF3-ChIP-Seq(GSE33912)/Homer	1e-464	-1.068e+03	0.0000	1404.0	17.07%	1584.3	3.91%	motif file (matrix)	svg
5		Fra2(bZIP)/Striatum-Fra2-ChIP-Seq(GSE43429)/Homer	1e-460	-1.060e+03	0.0000	1198.0	14.57%	1149.8	2.84%	motif file (matrix)	svg
6		Fos2(bZIP)/3T3L1-Fos2-ChIP-Seq(GSE56872)/Homer	1e-454	-1.047e+03	0.0000	974.0	11.84%	738.4	1.82%	motif file (matrix)	svg
7		Jun-AP1(bZIP)/K562-clum-ChIP-Seq(GSE31477)/Homer	1e-439	-1.011e+03	0.0000	818.0	9.95%	511.2	1.26%	motif file (matrix)	svg
8		AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-430	-9.915e+02	0.0000	1432.0	17.41%	1764.1	4.36%	motif file (matrix)	svg
9		Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	1e-199	-4.595e+02	0.0000	505.0	6.14%	455.8	1.13%	motif file (matrix)	svg
10		Sp5(Zf)/mES-Sp5-Flag-ChIP-Seq(GSE72989)/Homer	1e-159	-3.669e+02	0.0000	2316.0	28.16%	6604.0	16.31%	motif file (matrix)	svg
11		KLF3(Zf)/MEF-KLF3-ChIP-Seq(GSE44748)/Homer	1e-157	-3.632e+02	0.0000	1327.0	16.13%	2965.5	7.32%	motif file (matrix)	svg
12		Sp1(Zf)/Promoter/Homer	1e-155	-3.581e+02	0.0000	1259.0	15.31%	2756.9	6.81%	motif file (matrix)	svg
13		KLF5L5(Zf)/LoVo-KLF5-ChIP-Seq(GSE49402)/Homer	1e-148	-3.425e+02	0.0000	2520.0	30.64%	7571.1	18.69%	motif file (matrix)	svg

14		TEAD(TEA)/Fibroblast-PU.1-ChIP-Seq (Unpublished)/Homer	1e-128	-2.959e+02/0.0000	846.0	10.29%	1642.7	4.06%	inmotif file (matrix)	svg
15		KLF6(Zf)/PDAC-KLF6-ChIP-Seq (GSE64557)/Homer	1e-124	-2.871e+02/0.0000	2137.0	25.98%	6367.9	15.72%	inmotif file (matrix)	svg
16		TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq (GSE37350)/Homer	1e-122	-2.830e+02/0.0000	1035.0	12.58%	2291.7	5.66%	inmotif file (matrix)	svg
17		Atf1(bZIP)/K562-ATF1-ChIP-Seq (GSE31477)/Homer	1e-118	-2.730e+02/0.0000	941.0	11.44%	2019.3	4.99%	inmotif file (matrix)	svg
18		NFY (CCAAT)/Promoter/Homer	1e-118	-2.727e+02/0.0000	1146.0	13.93%	2710.0	6.69%	inmotif file (matrix)	svg
19		Klf9(Zf)/GBM-Klf9-ChIP-Seq (GSE62211)/Homer	1e-113	-2.621e+02/0.0000	977.0	11.88%	2178.2	5.38%	inmotif file (matrix)	svg
20		Klf14(Zf)/HEK293-Klf14.GFP-ChIP-Seq (GSE38341)/Homer	1e-108	-2.503e+02/0.0000	2976.0	36.18%	10177.3	25.13%	inmotif file (matrix)	svg
21		TEAD2(TEA)/Py2f-Tead2-ChIP-Seq (GSE55709)/Homer	1e-101	-2.343e+02/0.0000	659.0	8.01%	1260.7	3.11%	inmotif file (matrix)	svg
22		Nf-E2(bZIP)/K562-NFE2-ChIP-Seq (GSE31477)/Homer	1e-89	-2.056e+02/0.0000	175.0	2.13%	115.1	0.28%	inmotif file (matrix)	svg
23		Klf4(Zf)/mES-Klf4-ChIP-Seq (GSE11431)/Homer	1e-88	-2.033e+02/0.0000	770.0	9.36%	1719.8	4.25%	inmotif file (matrix)	svg
24		Atf7(bZIP)/3T3L1-Atf7-ChIP-Seq (GSE56872)/Homer	1e-86	-1.983e+02/0.0000	663.0	8.06%	1390.7	3.43%	inmotif file (matrix)	svg
25		Fli1(ETS)/CD8-Fl1-ChIP-Seq (GSE20898)/Homer	1e-86	-1.982e+02/0.0000	1608.0	19.55%	4838.2	11.95%	inmotif file (matrix)	svg
26		Elk4(ETS)/HeLa-Elk4-ChIP-Seq (GSE31477)/Homer	1e-84	-1.949e+02/0.0000	1153.0	14.02%	3108.2	7.67%	inmotif file (matrix)	svg
27		Nr2f2(bZIP)/Lymphoblast-Nr2f2-ChIP-Seq (GSE37589)/Homer	1e-82	-1.895e+02/0.0000	158.0	1.92%	101.7	0.25%	inmotif file (matrix)	svg
28		Ets1(ETS)/Jurkat-Ets1-ChIP-Seq (GSE17954)/Homer	1e-78	-1.815e+02/0.0000	1324.0	16.10%	3835.6	9.47%	inmotif file (matrix)	svg
29		ETV1(ETS)/GIST48-ETV1-ChIP-Seq (GSE22441)/Homer	1e-77	-1.789e+02/0.0000	1676.0	20.38%	5248.2	12.96%	inmotif file (matrix)	svg
30		GABPA(ETS)/Jurkat-GABPa-ChIP-Seq (GSE17954)/Homer	1e-71	-1.639e+02/0.0000	1229.0	14.94%	3581.9	8.84%	inmotif file (matrix)	svg

31		Atf2(bZIP)/3T3L1-Atf2-ChIP-Seq(GSE56872)/Homer	le-70	-1.619e+02 0.0000	494.0	6.01%	984.6	2.43%	inmotif file (matrix)	SVG
32		Elk1(ETS)/HeLa-Elk1-ChIP-Seq(GSE31477)/Homer	le-69	-1.594e+02 0.0000	1104.0	13.42%	3130.3	7.73%	inmotif file (matrix)	SVG
33		Bach1(bZIP)/K562-Bach1-ChIP-Seq(GSE31477)/Homer	le-68	-1.576e+02 0.0000	147.0	1.79%	109.6	0.27%	inmotif file (matrix)	SVG
34		c-Jun-CRE(bZIP)/K562-cJun-ChIP-Seq(GSE31477)/Homer	le-64	-1.477e+02 0.0000	429.0	5.22%	831.6	2.05%	inmotif file (matrix)	SVG
35		CRE(bZIP)/Promoter/Homer	le-63	-1.473e+02 0.0000	498.0	6.05%	1046.0	2.58%	inmotif file (matrix)	SVG
36		ERG(ETS)/VCaP-ERG-ChIP-Seq(GSE14097)/Homer	le-63	-1.464e+02 0.0000	1709.0	20.78%	5644.2	13.94%	inmotif file (matrix)	SVG
37		ETS(ETS)/Promoter/Homer	le-60	-1.403e+02 0.0000	664.0	8.07%	1628.8	4.02%	inmotif file (matrix)	SVG
38		ELF1(ETS)/Junkrat-ELF1-ChIP-Seq(SRA014231)/Homer	le-60	-1.393e+02 0.0000	967.0	11.76%	2735.0	6.75%	inmotif file (matrix)	SVG
39		MatK(bZIP)/C2C12-MatK-ChIP-Seq(GSE36030)/Homer	le-60	-1.387e+02 0.0000	302.0	3.67%	490.4	1.21%	inmotif file (matrix)	SVG
40		Eiv2(ETS)/ES-ER71-ChIP-Seq(GSE59402)/Homer(0.967)	le-56	-1.293e+02 0.0000	1070.0	13.01%	3196.3	7.89%	inmotif file (matrix)	SVG
41		FoxL2(Forkhead)/Ovary-FoxL2-ChIP-Seq(GSE60858)/Homer	le-55	-1.273e+02 0.0000	779.0	9.47%	2106.3	5.20%	inmotif file (matrix)	SVG
42		FOXA1(Forkhead)/MCF7-FOXA1-ChIP-Seq(GSE26831)/Homer	le-54	-1.260e+02 0.0000	872.0	10.60%	2459.5	6.07%	inmotif file (matrix)	SVG
43		FOXA1(Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	le-53	-1.232e+02 0.0000	1023.0	12.44%	3056.7	7.55%	inmotif file (matrix)	SVG
44		Foxfl(Forkhead)/Lung-Foxfl-ChIP-Seq(GSE77951)/Homer	le-53	-1.224e+02 0.0000	822.0	9.99%	2295.0	5.67%	inmotif file (matrix)	SVG
45		Foxo3(Forkhead)/U2OS-Foxo3-ChIP-Seq(EMTAB-2701)/Homer	le-53	-1.221e+02 0.0000	710.0	8.63%	1883.0	4.65%	inmotif file (matrix)	SVG
46		Maz(Z)/HepG2-Maz-ChIP-Seq(GSE31477)/Homer	le-52	-1.208e+02 0.0000	2327.0	28.29%	8563.9	21.15%	inmotif file (matrix)	SVG
47		NF1(CTF)/LNCAP-NF1-ChIP-Seq(Unpublished)/Homer	le-50	-1.159e+02 0.0000	547.0	6.65%	1337.8	3.30%	inmotif file (matrix)	SVG


















48		EWS-FLI1-fusion(ETS)/SK_N_MC-EWS:FLI1-ChIP-Seq(SRA014231)/Homer	le-47	-1.104e+02 0.0000	743.0	9.03%	2072.3	5.12%	inmotif file (matrix) SVG
49		FOXM1(Forkhead)/MCF7-FOXMI-ChIP-Seq(GSE72977)/Homer	le-44	-1.034e+02 0.0000	854.0	10.38%	2539.2	6.27%	inmotif file (matrix) SVG
50		Six2(Homeobox)/Nephron Progenitor-Six2-ChIP-Seq(GSE39837)/Homer	le-44	-1.028e+02 0.0000	850.0	10.33%	2528.7	6.24%	inmotif file (matrix) SVG
51		Smad3(MAD)/NPC-Smad3-ChIP-Seq(GSE36673)/Homer	le-44	-1.022e+02 0.0000	2264.0	27.53%	8504.2	21.00%	inmotif file (matrix) SVG
52		KLF10(ZF)/HEK293-KLF10-GFP-ChIP-Seq(GSE58341)/Homer	le-44	-1.022e+02 0.0000	778.0	9.46%	2256.7	5.57%	inmotif file (matrix) SVG
53		Six1(Homeobox)/Myoblast-Six1-ChIP-Chip(GSE20150)/Homer	le-44	-1.020e+02 0.0000	298.0	3.62%	579.6	1.43%	inmotif file (matrix) SVG
54		NRF1(NRF)/MCF7-NRF1-ChIP-Seq(Unpublished)/Homer	le-40	-9.406e+01 0.0000	553.0	6.72%	1469.7	3.63%	inmotif file (matrix) SVG
55		FOXP1(Forkhead)/H9-FOXP1-ChIP-Seq(GSE31006)/Homer	le-39	-9.186e+01 0.0000	420.0	5.11%	1012.9	2.50%	inmotif file (matrix) SVG
56		NRF(NRF)/Promoter/Homer	le-39	-9.179e+01 0.0000	539.0	6.55%	1431.5	3.53%	inmotif file (matrix) SVG
57		JunD(bZIP)/K562-JunD-ChIP-Seq/Homer	le-38	-8.833e+01 0.0000	196.0	2.38%	323.4	0.80%	inmotif file (matrix) SVG
58		FOXK1(Forkhead)/HEK293-FOXK1-ChIP-Seq(GSE31673)/Homer	le-37	-8.680e+01 0.0000	870.0	10.58%	2726.8	6.73%	inmotif file (matrix) SVG
59		CEBP(bZIP)/ThioMac-CEBPb-ChIP-Seq(GSE21512)/Homer	le-35	-8.231e+01 0.0000	607.0	7.38%	1739.8	4.30%	inmotif file (matrix) SVG
60		Foxo3(Forkhead)/Liver-Foxo3-ChIP-Seq(GSE77670)/Homer	le-33	-7.777e+01 0.0000	313.0	3.81%	714.5	1.76%	inmotif file (matrix) SVG
61		EKLf(ZF)/Erythrocyte-Klf1-ChIP-Seq(GSE20478)/Homer	le-33	-7.671e+01 0.0000	272.0	3.31%	585.1	1.44%	inmotif file (matrix) SVG
62		CEBP:AP1(bZIP)/ThioMac-CEBPb-ChIP-Seq(GSE21512)/Homer	le-32	-7.522e+01 0.0000	643.0	7.82%	1925.9	4.76%	inmotif file (matrix) SVG
63		Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq(fan_et_al)/Homer	le-32	-7.438e+01 0.0000	1439.0	17.50%	5219.4	12.89%	inmotif file (matrix) SVG
64		RUNX(Runb)/HPC7-Runx1-ChIP-Seq(GSE22178)/Homer	le-30	-7.096e+01 0.0000	594.0	7.22%	1767.7	4.36%	inmotif file (matrix) SVG

65		MITF(bHLH)/Masi/Cells-MITF-ChIP-Seq(GSE48085)/Homer	le-30	-6.931e+01 0.0000	832.0	10.12%	2722.1	6.72%	inmotif file (matrix)	svg
66		MafA(bZIP)/MafA-ChIP-Seq(GSE30298)/Homer	le-29	-6.768e+01 0.0000	711.0	8.64%	2250.6	5.56%	inmotif file (matrix)	svg
67		Aif4(bZIP)/MEF-Aif4-ChIP-Seq(GSE35681)/Homer	le-29	-6.730e+01 0.0000	295.0	3.59%	699.6	1.73%	inmotif file (matrix)	svg
68		Ust2(bHLH)/C2C12-Ust2-ChIP-Seq(GSE36030)/Homer	le-27	-6.436e+01 0.0000	374.0	4.55%	989.5	2.44%	inmotif file (matrix)	svg
69		HLF(bZIP)/HSC-HLF-Flag-ChIP-Seq(GSE69817)/Homer	le-27	-6.412e+01 0.0000	660.0	8.02%	2077.2	5.13%	inmotif file (matrix)	svg
70		NF1-FOXA1(CTF,Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	le-27	-6.372e+01 0.0000	87.0	1.06%	99.3	0.25%	inmotif file (matrix)	svg
71		Tlx?(NR)/NPC-H3K4me1-ChIP-Seq(GSE16256)/Homer	le-27	-6.224e+01 0.0000	423.0	5.14%	1180.2	2.91%	inmotif file (matrix)	svg
72		EHF(ETS)/LoVo-EHF-ChIP-Seq(GSE49402)/Homer	le-26	-6.044e+01 0.0000	1192.0	14.49%	4323.3	10.67%	inmotif file (matrix)	svg
73		RUNX1(Runt)/Jurkat-RUNX1-ChIP-Seq(GSE29180)/Homer	le-25	-5.890e+01 0.0000	754.0	9.17%	2501.3	6.18%	inmotif file (matrix)	svg
74		FoxEbox(Forkhead,bHLH)/Panel1-Foxa2-ChIP-Seq(GSE47459)/Homer	le-24	-5.641e+01 0.0000	686.0	8.34%	2248.6	5.55%	inmotif file (matrix)	svg
75		Nanog(Homobox)/mES-Nanog-ChIP-Seq(GSE11724)/Homer	le-23	-5.473e+01 0.0000	3040.0	36.96%	12830.0	31.68%	inmotif file (matrix)	svg
76		Chop(bZIP)/MEF-Chop-ChIP-Seq(GSE35681)/Homer	le-23	-5.421e+01 0.0000	228.0	2.77%	532.6	1.31%	inmotif file (matrix)	svg
77		SPDEF(ETS)/VCaP-SPDEF-ChIP-Seq(SRA014231)/Homer	le-23	-5.378e+01 0.0000	942.0	11.45%	3336.5	8.24%	inmotif file (matrix)	svg
78		FOXK2(Forkhead)/U2OS-FOXK2-ChIP-Seq(EMTAB-2204)/Homer	le-23	-5.347e+01 0.0000	505.0	6.14%	1553.7	3.84%	inmotif file (matrix)	svg
79		CDX4(Homobox)/ZebrafishEmbryos-Cdx4-Myc-ChIP-Seq(GSE48254)/Homer	le-22	-5.261e+01 0.0000	701.0	8.52%	2345.3	5.79%	inmotif file (matrix)	svg
80		GFY-Statf(Zf)/Promoter/Homer	le-22	-5.258e+01 0.0000	147.0	1.79%	281.7	0.70%	inmotif file (matrix)	svg
81		Zfp281(Zf)/ES-Zfp281-ChIP-Seq(GSE81042)/Homer	le-22	-5.211e+01 0.0000	342.0	4.16%	944.3	2.33%	inmotif file (matrix)	svg

82		RUNX2(Runt)/PCa-RUNX2-ChIP-Seq(GSE33889)/Homer	le-22	-5.192e+01 0.0000	582.0	7.08%	1870.7	4.62%	inmotif file (matrix)	svg
83		RUNX-AML(Runt)/CD4+PolII-ChIP-Seq(BarSKI_et_al)/Homer	le-21	-4.901e+01 0.0000	527.0	6.41%	1676.4	4.14%	inmotif file (matrix)	svg
84		MatB(bZIP)/BMM-MatB-ChIP-Seq(GSE75722)/Homer	le-21	-4.850e+01 0.0000	309.0	3.76%	845.4	2.09%	inmotif file (matrix)	svg
85		Stat3(Stat)/mES-Stat3-ChIP-Seq(GSE11431)/Homer	le-20	-4.832e+01 0.0000	524.0	6.37%	1670.9	4.13%	inmotif file (matrix)	svg
86		HOXB13(Homeobox)/ProstateTumor-HOXB13-ChIP-Seq(GSE56288)/Homer	le-20	-4.733e+01 0.0000	800.0	9.73%	2811.7	6.94%	inmotif file (matrix)	svg
87		E-box(bHLH)/Promoter/Homer	le-20	-4.710e+01 0.0000	189.0	2.30%	433.5	1.07%	inmotif file (matrix)	svg
88		TFES(bHLH)/MEF-TFES-ChIP-Seq(GSE75757)/Homer	le-19	-4.513e+01 0.0000	150.0	1.82%	315.1	0.78%	inmotif file (matrix)	svg
89		Rom1n(THAP)/ES-Thap1-ChIP-Seq(GSE51522)/Homer	le-18	-4.308e+01 0.0000	121.0	1.47%	233.8	0.58%	inmotif file (matrix)	svg
90		Foxa2(Forkhead)/Liver-Foxa2-ChIP-Seq(GSE25694)/Homer	le-18	-4.277e+01 0.0000	590.0	7.17%	1988.2	4.91%	inmotif file (matrix)	svg
91		Ets1-distal(ETS)/CD4+PolII-ChIP-Seq(BarSKI_et_al)/Homer	le-18	-4.244e+01 0.0000	282.0	3.43%	784.0	1.94%	inmotif file (matrix)	svg
92		Stat3+i12(Stat)/CD4-Stat3-ChIP-Seq(GSE19198)/Homer	le-18	-4.238e+01 0.0000	590.0	7.17%	1992.7	4.92%	inmotif file (matrix)	svg
93		ETS:RUNX(ETS,Runt)/Jurkat-RUNX1-ChIP-Seq(GSE17954)/Homer	le-17	-4.071e+01 0.0000	142.0	1.73%	306.7	0.76%	inmotif file (matrix)	svg
94		ELF5(ETS)/T47D-ELF5-ChIP-Seq(GSE30407)/Homer	le-17	-4.021e+01 0.0000	683.0	8.30%	2401.0	5.93%	inmotif file (matrix)	svg
95		Smad4(MAD)/ESC-SMAD4-ChIP-Seq(GSE29422)/Homer	le-16	-3.873e+01 0.0000	1354.0	16.46%	5347.1	13.20%	inmotif file (matrix)	svg
96		NF1-halfsite(CTF)/NCaP-NF1-ChIP-Seq(Unpublished)/Homer	le-16	-3.861e+01 0.0000	1505.0	18.30%	6029.9	14.89%	inmotif file (matrix)	svg
97		Nur77(NR)/K562-NR4A1-ChIP-Seq(GSE31363)/Homer	le-16	-3.783e+01 0.0000	164.0	1.99%	390.3	0.96%	inmotif file (matrix)	svg
98		Ptx1(Homeobox)/Chicken-Ptx1-ChIP-Seq(GSE38910)/Homer	le-15	-3.674e+01 0.0000	2834.0	34.46%	12248.1	30.24%	inmotif file (matrix)	svg

99		GFY (?)/Promoter/Homer	1e-15	-3.621e+01 0.0000	139.0	1.69%	314.3	0.78%	motif file (matrix)	SVG
100		ZNF143/STAF(Zf)/CUTLL-ZNF143-ChIP-Seq(GSE29600)/Homer	1e-15	-3.565e+01 0.0000	337.0	4.10%	1039.3	2.57%	motif file (matrix)	SVG
101		Smad2(MAD)/ES-SMAD2-ChIP-Seq(GSE29422)/Homer	1e-15	-3.554e+01 0.0000	1330.0	16.17%	5295.5	13.08%	motif file (matrix)	SVG
102		HOXA9/Homeobox)/HSC-Hoxa9-ChIP-Seq(GSE33509)/Homer	1e-15	-3.483e+01 0.0000	500.0	6.08%	1700.8	4.20%	motif file (matrix)	SVG
103		EWS-ERG-fusion(ETS)/CADO_ES1-EWS-ERG-ChIP-Seq(SRA014231)/Homer	1e-14	-3.443e+01 0.0000	580.0	7.05%	2036.1	5.03%	motif file (matrix)	SVG
104		CTCF(Zf)/CD4+CTCF-ChIP-Seq(Barski_et_al)/Homer	1e-14	-3.377e+01 0.0000	176.0	2.14%	451.8	1.12%	motif file (matrix)	SVG
105		E2F4(E2F)/K562-E2F4-ChIP-Seq(GSE31477)/Homer	1e-14	-3.256e+01 0.0000	926.0	11.26%	3548.3	8.76%	motif file (matrix)	SVG
106		YY1(Zf)/Promoter/Homer	1e-13	-3.208e+01 0.0000	154.0	1.87%	383.6	0.95%	motif file (matrix)	SVG
107		USF1(bHLH)/GM12878-Usf1-ChIP-Seq(GSE32465)/Homer	1e-13	-3.154e+01 0.0000	498.0	6.05%	1727.6	4.27%	motif file (matrix)	SVG
108		Unknown/Homeobox)/Limb-p300-ChIP-Seq/Homer	1e-13	-3.152e+01 0.0000	490.0	5.96%	1694.1	4.18%	motif file (matrix)	SVG
109		Cdx2/Homeobox)/mES-Cdx2-ChIP-Seq(GSE14586)/Homer	1e-13	-3.135e+01 0.0000	539.0	6.55%	1900.9	4.69%	motif file (matrix)	SVG
110		Hoxe9/Homeobox)/Aimv15-Hoxe9-ChIP-Seq(GSE21812)/Homer	1e-12	-2.937e+01 0.0000	391.0	4.75%	1310.6	3.24%	motif file (matrix)	SVG
111		BORIS(Zf)/K562-CTCF1-ChIP-Seq(GSE32465)/Homer	1e-12	-2.931e+01 0.0000	358.0	4.35%	1177.3	2.91%	motif file (matrix)	SVG
112		TATA-Box(TBP)/Promoter/Homer	1e-12	-2.856e+01 0.0000	877.0	10.66%	3396.5	8.39%	motif file (matrix)	SVG
113		Bcl6(Zf)/Liver-Bcl6-ChIP-Seq(GSE31578)/Homer	1e-12	-2.822e+01 0.0000	984.0	11.96%	3876.2	9.57%	motif file (matrix)	SVG
114		GFX(?) /Promoter/Homer	1e-12	-2.819e+01 0.0000	55.0	0.67%	86.2	0.21%	motif file (matrix)	SVG
115		HOXD13/Homeobox)/Chicken-Hoxd13-ChIP-Seq(GSE38910)/Homer	1e-12	-2.786e+01 0.0000	769.0	9.35%	2933.9	7.24%	motif file (matrix)	SVG

116		Ap4(bHLH)/AML-T1ap4-ChIP-Seq(GSE45738)/Homer	1e-11	-2.609e+01	0.0000	1070.0	13.01%	4300.5	10.62%	mmotif file (matrix)	svg
117		Amt:Ahr(bHLH)/MCF7-Amt-ChIP-Seq(Lo_et_al)/Homer	1e-10	-2.507e+01	0.0000	659.0	8.01%	2496.8	6.16%	mmotif file (matrix)	svg
118		NFRB-p65-Rel(RHD)/ThioMac-LPS-Expression(GSE23622)/Homer	1e-10	-2.507e+01	0.0000	70.0	0.85%	137.8	0.34%	mmotif file (matrix)	svg
119		STAT4(Stat4)/CD4-Stat4-ChIP-Seq(GSE22104)/Homer	1e-10	-2.483e+01	0.0000	704.0	8.56%	2696.0	6.66%	mmotif file (matrix)	svg
120		Tcf21(bHLH)/ArterySmoothMuscle-Tcf21-ChIP-Seq(GSE61369)/Homer	1e-10	-2.473e+01	0.0000	876.0	10.65%	3457.1	8.54%	mmotif file (matrix)	svg
121		PU.1(ETS)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-10	-2.451e+01	0.0000	394.0	4.79%	1374.4	3.39%	mmotif file (matrix)	svg
122		Egr2(Zf)/Thymocytes-Egr2-ChIP-Seq(GSE34254)/Homer	1e-10	-2.409e+01	0.0000	331.0	4.02%	1119.3	2.76%	mmotif file (matrix)	svg
123		CLOCK(bHLH)/Liver-Clock-ChIP-Seq(GSE39860)/Homer	1e-10	-2.372e+01	0.0000	525.0	6.38%	1937.1	4.78%	mmotif file (matrix)	svg
124		Isl1(Homeobox)/Neuron-Isl1-ChIP-Seq(GSE31456)/Homer	1e-10	-2.348e+01	0.0000	1305.0	15.87%	5422.8	13.39%	mmotif file (matrix)	svg
125		NeuroG2(bHLH)/Fibroblast-NeuroG2-ChIP-Seq(GSE75910)/Homer	1e-9	-2.278e+01	0.0000	1142.0	13.88%	4693.7	11.59%	mmotif file (matrix)	svg
126		E2F3(E2F)/MEF-E2F3-ChIP-Seq(GSE71376)/Homer	1e-9	-2.278e+01	0.0000	1094.0	13.30%	4475.6	11.05%	mmotif file (matrix)	svg
127		COUP-TFII(NR)/Artria-Nr2f2-ChIP-Seq(GSE46497)/Homer	1e-9	-2.208e+01	0.0000	1281.0	15.57%	5345.0	13.20%	mmotif file (matrix)	svg
128		ELF3(ETS)/PDAC-ELF3-ChIP-Seq(GSE64557)/Homer	1e-9	-2.196e+01	0.0000	583.0	7.09%	2213.4	5.47%	mmotif file (matrix)	svg
129		Sox17(HMG)/Endoderm-Sox17-ChIP-Seq(GSE61475)/Homer	1e-9	-2.087e+01	0.0000	474.0	5.76%	1758.4	4.34%	mmotif file (matrix)	svg
130		Sox10(HMG)/SciaticNerve-Sox3-ChIP-Seq(GSE35132)/Homer	1e-8	-2.041e+01	0.0000	1052.0	12.79%	4335.6	10.71%	mmotif file (matrix)	svg
131		MYB(HTH)/ERMYPB-Myb-ChIPSeq(GSE22095)/Homer	1e-8	-2.038e+01	0.0000	1415.0	17.20%	6003.0	14.82%	mmotif file (matrix)	svg
132		Otx2(Homeobox)/EpiLC-Otx2-ChIP-Seq(GSE56098)/Homer	1e-8	-2.026e+01	0.0000	504.0	6.13%	1896.3	4.68%	mmotif file (matrix)	svg

133		Sox3(HMG)/NPC-Sox3-ChIP-Seq(GSE33059)/Homer	1e-8	-1.980e+01	0.0000	1120.0	13.62%	4659.5	11.50%	m motif file (matrix)	svg
134		BMAL1(bHLH)/Liver-Bmal1-ChIP-Seq(GSE39860)/Homer	1e-8	-1.958e+01	0.0000	1265.0	15.38%	5330.3	13.16%	m motif file (matrix)	svg
135		Esrrb(NR)/mES-Esrrb-ChIP-Seq(GSE11431)/Homer	1e-8	-1.847e+01	0.0000	446.0	5.42%	1672.9	4.13%	m motif file (matrix)	svg
136		BMYB(HTH)/HeLa-BMYB-ChIP-Seq(GSE27030)/Homer	1e-7	-1.776e+01	0.0000	1107.0	13.46%	4648.4	11.48%	m motif file (matrix)	svg
137		AMYB(HTH)/Testes-AMYB-ChIP-Seq(GSE44588)/Homer	1e-7	-1.763e+01	0.0000	1140.0	13.86%	4803.4	11.86%	m motif file (matrix)	svg
138		bHLHE40(bHLH)/HepG2-BHLHE40-ChIP-Seq(GSE31477)/Homer	1e-7	-1.725e+01	0.0000	342.0	4.16%	1245.0	3.07%	m motif file (matrix)	svg
139		GSC(Homeobox)/FrogEmbryos-GSC-ChIP-Seq(DRA000576)/Homer	1e-7	-1.698e+01	0.0000	711.0	8.64%	2866.9	7.08%	m motif file (matrix)	svg
140		Foxh1(Forkhead)/hESC-FOXH1-ChIP-Seq(GSE29422)/Homer	1e-7	-1.647e+01	0.0000	387.0	4.71%	1447.6	3.57%	m motif file (matrix)	svg
141		ZNF467(ZF)/HEK293-ZNF467-GFP-ChIP-Seq(GSE58341)/Homer	1e-7	-1.639e+01	0.0000	959.0	11.66%	4003.6	9.89%	m motif file (matrix)	svg
142		Sox15(HMG)/CPA-Sox15-ChIP-Seq(GSE62909)/Homer	1e-7	-1.639e+01	0.0000	676.0	8.22%	2721.5	6.72%	m motif file (matrix)	svg
143		STAT1(Stat)/HeLaS3-STAT1-ChIP-Seq(GSE12782)/Homer	1e-7	-1.619e+01	0.0000	225.0	2.74%	769.0	1.90%	m motif file (matrix)	svg
144		PU.1(IRF/ETS-IRF)/Beel1-PU.1-ChIP-Seq(GSE21512)/Homer	1e-6	-1.606e+01	0.0000	926.0	11.26%	3860.6	9.53%	m motif file (matrix)	svg
145		HIF-1b(HLH)/T47D-HIF1b-ChIP-Seq(GSE59937)/Homer	1e-6	-1.586e+01	0.0000	1097.0	13.34%	4650.2	11.48%	m motif file (matrix)	svg
146		Sox2(HMG)/mES-Sox2-ChIP-Seq(GSE11431)/Homer	1e-6	-1.578e+01	0.0000	584.0	7.10%	2322.4	5.73%	m motif file (matrix)	svg
147		NFATc1(API)(RHD,bZIP)/Jurkat-NFATc1-ChIP-Seq(Jolma_et_al)/Homer	1e-6	-1.527e+01	0.0000	125.0	1.52%	378.3	0.93%	m motif file (matrix)	svg
148		CArG(MADS)/PUER-Srf-ChIP-Seq(Sullivan_et_al)/Homer	1e-6	-1.524e+01	0.0000	202.0	2.46%	684.5	1.69%	m motif file (matrix)	svg
149		CRX(Homeobox)/Retina-Crx-ChIP-Seq(GSE20012)/Homer	1e-6	-1.510e+01	0.0000	1416.0	17.22%	6154.6	15.20%	m motif file (matrix)	svg

150		Atoh1 (bHLH)/Cerebellum-Atoh1-ChIP-Seq (GSE22111)/Homer	1e-6	-1.508e+01	0.0000	869.0	10.57%	3623.2	8.95%	mmioff file (matrix)	svg
151		NeuroD1 (bHLH)/Islet-NeuroD1-ChIP-Seq (GSE30298)/Homer	1e-6	-1.471e+01	0.0000	606.0	7.37%	2441.0	6.03%	mmioff file (matrix)	svg
152		ZBTB33 (ZF)/GM12878-ZBTB33-ChIP-Seq (GSE32465)/Homer	1e-6	-1.400e+01	0.0000	116.0	1.41%	353.2	0.87%	mmioff file (matrix)	svg
153		E2F1 (E2F)/HeLa-E2F1-ChIP-Seq (GSE22478)/Homer	1e-6	-1.391e+01	0.0000	498.0	6.05%	1975.0	4.88%	mmioff file (matrix)	svg
154		HIF-1a (bHLH)/MCF7-HIF1a-ChIP-Seq (GSE28352)/Homer	1e-5	-1.376e+01	0.0000	309.0	3.76%	1151.5	2.84%	mmioff file (matrix)	svg
155		Nkx6.1 (Homeobox)/Islet-Nkx6.1-ChIP-Seq (GSE40975)/Homer	1e-5	-1.363e+01	0.0000	1722.0	20.94%	7642.6	18.87%	mmioff file (matrix)	svg
156		Egr1 (ZF)/K562-Egr1-ChIP-Seq (GSE32465)/Homer	1e-5	-1.358e+01	0.0000	911.0	11.08%	3853.8	9.52%	mmioff file (matrix)	svg
157		PAX6 (Paired-Homeobox)/Forebrain-Pax6-ChIP-Seq (GSE6961)/Homer	1e-5	-1.349e+01	0.0000	75.0	0.91%	203.9	0.50%	mmioff file (matrix)	svg
158		Lhx3 (Homeobox)/Neuron-Lhx3-ChIP-Seq (GSE31456)/Homer	1e-5	-1.332e+01	0.0000	1142.0	13.88%	4929.2	12.17%	mmioff file (matrix)	svg
159		Olig2 (bHLH)/Neuron-Olig2-ChIP-Seq (GSE30882)/Homer	1e-5	-1.315e+01	0.0000	1292.0	15.71%	5634.4	13.91%	mmioff file (matrix)	svg
160		Pdx1 (Homeobox)/Islet-Pdx1-ChIP-Seq (SRA008281)/Homer	1e-5	-1.312e+01	0.0000	635.0	7.72%	2605.1	6.43%	mmioff file (matrix)	svg
161		Gfi1b (ZF)/HPC7-Gfi1b-ChIP-Seq (GSE22178)/Homer	1e-5	-1.312e+01	0.0000	414.0	5.03%	1618.9	4.00%	mmioff file (matrix)	svg
162		E2F6 (E2F)/HeLa-E2F6-ChIP-Seq (GSE31477)/Homer	1e-5	-1.309e+01	0.0000	934.0	11.36%	3972.7	9.81%	mmioff file (matrix)	svg
163		NPAS (bHLH)/Liver-NPAS-ChIP-Seq (GSE39860)/Homer	1e-5	-1.283e+01	0.0000	1106.0	13.45%	4776.2	11.79%	mmioff file (matrix)	svg
164		STAT5 (Stat)/mCD4+Stat5-ChIP-Seq (GSE12346)/Homer	1e-5	-1.260e+01	0.0000	241.0	2.93%	877.5	2.17%	mmioff file (matrix)	svg
165		SpiB (ETS)/OCILY3-SpiB-ChIP-Seq (GSE56857)/Homer	1e-5	-1.248e+01	0.0000	178.0	2.16%	615.4	1.52%	mmioff file (matrix)	svg
166		GATA (Z), IR3/Treg-Gata3-ChIP-Seq (GSE20898)/Homer	1e-5	-1.247e+01	0.0000	108.0	1.31%	334.8	0.83%	mmioff file (matrix)	svg

167		Sox6(HMG)/Myotubes-Sox6-ChIP-Seq(GSE32627)/Homer	1e-5	-1.234e+01	0.0000	998.0	12.13%	4289.9	10.59%	mmotif file (matrix)	svg
168		SCL(bHLH)/HPC7-Scl-ChIP-Seq(GSE13511)/Homer	1e-5	-1.207e+01	0.0000	3411.0	41.47%	15834.1	39.10%	mmotif file (matrix)	svg
169		ZBTB18(Z)/HEK293-ZBTB18.GFP-ChIP-Seq(GSE38341)/Homer	1e-5	-1.202e+01	0.0000	371.0	4.51%	1448.3	3.58%	mmotif file (matrix)	svg
170		Rbpj(?)Panc1-Rbpj1-ChIP-Seq(GSE47459)/Homer	1e-5	-1.186e+01	0.0000	1023.0	12.44%	4419.4	10.91%	mmotif file (matrix)	svg
171		bZIP:IRF(bZIP,IRF)/Th17-Batf-ChIP-Seq(GSE39756)/Homer	1e-5	-1.155e+01	0.0000	258.0	3.14%	964.2	2.38%	mmotif file (matrix)	svg
172		E2F7(E2F)/Hela-E2F7-ChIP-Seq(GSE32673)/Homer	1e-5	-1.152e+01	0.0000	253.0	3.08%	943.7	2.33%	mmotif file (matrix)	svg
173		Max(bHLH)/K562-Max-ChIP-Seq(GSE31477)/Homer	1e-4	-1.133e+01	0.0000	529.0	6.43%	2166.6	5.35%	mmotif file (matrix)	svg
174		Sox4(HMG)/proB-Sox4-ChIP-Seq(GSE50066)/Homer	1e-4	-1.118e+01	0.0000	544.0	6.61%	2237.2	5.52%	mmotif file (matrix)	svg
175		NFKB-p65(RHD)/GM12877-p65-ChIP-Seq(GSE19485)/Homer	1e-4	-1.107e+01	0.0000	378.0	4.60%	1496.8	3.70%	mmotif file (matrix)	svg
176		Unknown-ESC-element(?)/mES-Nanog-ChIP-Seq(GSE11724)/Homer	1e-4	-1.093e+01	0.0000	502.0	6.10%	2053.9	5.07%	mmotif file (matrix)	svg
177		MyoG(bHLH)/C2C12-MyoG-ChIP-Seq(GSE36024)/Homer	1e-4	-1.083e+01	0.0000	902.0	10.97%	3888.0	9.60%	mmotif file (matrix)	svg
178		THRb(NR)/Liver-NR1A2-ChIP-Seq(GSE52613)/Homer	1e-4	-1.078e+01	0.0000	2712.0	32.97%	12501.6	30.87%	mmotif file (matrix)	svg
179		Tgf2(Homeobox)/mES-Tgf2-ChIP-Seq(GSE55404)/Homer	1e-4	-1.054e+01	0.0001	2153.0	26.18%	9819.7	24.25%	mmotif file (matrix)	svg
180		Me2b(MADS)/HEK293-Me2b.V5-ChIP-Seq(GSE67450)/Homer	1e-4	-1.029e+01	0.0001	511.0	6.21%	2108.9	5.21%	mmotif file (matrix)	svg
181		NFAT(RHD)/Jurkat-NFATC1-ChIP-Seq(Jolma_et_al)/Homer	1e-4	-9.736e+00	0.0001	594.0	7.22%	2499.7	6.17%	mmotif file (matrix)	svg
182		E2F7(E2F)/Hela-CellCycle-Expression/Homer	1e-4	-9.487e+00	0.0002	99.0	1.20%	325.0	0.80%	mmotif file (matrix)	svg
183		Lhx2(Homeobox)/HEPSC-Lhx2-ChIP-Seq(GSE48068)/Homer	1e-4	-9.426e+00	0.0002	714.0	8.68%	3059.8	7.56%	mmotif file (matrix)	svg

184		MaafF(bZIP)/HepG2-MaafF-ChIP-Seq(GSE31477)/Homer	1e-4	-9.378e+000.0002	172.0	2.09%	627.6	1.55%	inmotif file (matrix)	SVG
185		Zic3(Zf)/mES-Zic3-ChIP-Seq(GSE37889)/Homer	1e-4	-9.309e+000.0002	715.0	8.69%	3067.1	7.57%	inmotif file (matrix)	SVG
186		MeZd(MADS)/Retina-MeZd-ChIP-Seq(GSE61391)/Homer	1e-3	-9.141e+000.0002	131.0	1.59%	458.8	1.13%	inmotif file (matrix)	SVG
187		Srebp1a(bHLH)/HepG2-Srebp1a-ChIP-Seq(GSE31477)/Homer	1e-3	-9.105e+000.0002	141.0	1.71%	500.3	1.24%	inmotif file (matrix)	SVG
188		n-Myc(bHLH)/mES-nMyc-ChIP-Seq(GSE11431)/Homer	1e-3	-8.971e+000.0002	619.0	7.53%	2634.1	6.50%	inmotif file (matrix)	SVG
189		Bapx1(Homobox)/VertebralCol-Bapx1-ChIP-Seq(GSE36672)/Homer	1e-3	-8.776e+000.0003	1332.0	16.19%	5978.6	14.76%	inmotif file (matrix)	SVG
190		MeZc(MADS)/GM12878-MeZc-ChIP-Seq(GSE32465)/Homer	1e-3	-8.703e+000.0003	275.0	3.34%	1084.4	2.68%	inmotif file (matrix)	SVG
191		Rfx5(HTH)/GM12878-Rfx5-ChIP-Seq(GSE31477)/Homer	1e-3	-7.891e+000.0007	232.0	2.82%	909.9	2.25%	inmotif file (matrix)	SVG
192		p63(p53)/Keratinocyte-p63-ChIP-Seq(GSE17611)/Homer	1e-3	-7.381e+000.0012	222.0	2.70%	874.1	2.16%	inmotif file (matrix)	SVG
193		Sox9(HMG)/Limb-SOX9-ChIP-Seq(GSE73225)/Homer	1e-3	-7.328e+000.0012	491.0	5.97%	2091.6	5.16%	inmotif file (matrix)	SVG
194		e-Myc(bHLH)/LNCaP-eMyc-ChIP-Seq(Unpublished)/Homer	1e-3	-7.315e+000.0012	569.0	6.92%	2451.9	6.05%	inmotif file (matrix)	SVG
195		Nkx3.1(Homobox)/LNCaP-Nkx3.1-ChIP-Seq(GSE28264)/Homer	1e-3	-7.282e+000.0013	1474.0	17.92%	6719.2	16.59%	inmotif file (matrix)	SVG
196		Tbx20(T-box)/Heart-Tbx20-ChIP-Seq(GSE29636)/Homer	1e-3	-7.240e+000.0013	155.0	1.88%	584.8	1.44%	inmotif file (matrix)	SVG
197		p53(p53)/Saos-p53-ChIP-Seq(GSE15780)/Homer	1e-3	-6.911e+000.0018	55.0	0.67%	172.1	0.42%	inmotif file (matrix)	SVG
198		p53(p53)/Saos-p53-ChIP-Seq(Homer)	1e-3	-6.911e+000.0018	55.0	0.67%	172.1	0.42%	inmotif file (matrix)	SVG
199		GATA(Zf),IR4/Treg-Gata3-ChIP-Seq(GSE20898)/Homer	1e-2	-6.574e+000.0026	44.0	0.53%	132.1	0.33%	inmotif file (matrix)	SVG
200		ZNF711(Zf)/SH5Y-ZNF711-ChIP-Seq(GSE20673)/Homer	1e-2	-6.550e+000.0026	2141.0	26.03%	9963.9	24.60%	inmotif file (matrix)	SVG

201		Barx1 (Homeobox)/Stomach-Barx1.3xFlag-ChIP-Seq(GSE69483)/Homer	1e-2	-6.542e+000.0026	316.0	3.84%	1312.8	3.24%	mmioff file (matrix)	svg
202		Tgif1 (Homeobox)/mES-Tgif1-ChIP-Seq(GSE55404)/Homer	1e-2	-6.417e+000.0029	1899.0	23.09%	8803.8	21.74%	mmioff file (matrix)	svg
203		Tbox-Smad(T-box, MAD)/EScd5-Smad2_3-ChIP-Seq(GSE29422)/Homer	1e-2	-6.394e+000.0030	105.0	1.28%	382.9	0.95%	mmioff file (matrix)	svg
204		Irf8(Irf)/BMDM-Irf8-ChIP-Seq(GSE77884)/Homer	1e-2	-6.359e+000.0031	215.0	2.61%	862.7	2.13%	mmioff file (matrix)	svg
205		Hrf(Hsf)/HepG2-Hsf1-ChIP-Seq(GSE31477)/Homer	1e-2	-6.305e+000.0032	100.0	1.22%	362.6	0.90%	mmioff file (matrix)	svg
206		Trf4(Nr)/DR1/HeLa-Trf4-ChIP-Seq(GSE24685)/Homer	1e-2	-6.253e+000.0034	80.0	0.97%	279.7	0.69%	mmioff file (matrix)	svg
207		Tcf4(Hmg)/Het116-Tcf4-ChIP-Seq(SRA012054)/Homer	1e-2	-6.099e+000.0039	260.0	3.16%	1069.6	2.64%	mmioff file (matrix)	svg
208		Pax5(Paired, Homeobox), condensed/GM12878-Pax5-ChIP-Seq(GSE32465)/Homer	1e-2	-5.951e+000.0046	102.0	1.24%	375.4	0.93%	mmioff file (matrix)	svg
209		Erra(Nr)/HepG2-Erra-ChIP-Seq(GSE31477)/Homer	1e-2	-5.913e+000.0047	1503.0	18.27%	6927.6	17.11%	mmioff file (matrix)	svg
210		Npas2(bHLH)/Liver-NPAS2-ChIP-Seq(GSE39860)/Homer	1e-2	-5.756e+000.0055	758.0	9.22%	3388.1	8.37%	mmioff file (matrix)	svg
211		Hoxb4(Homeobox)/ES-Hoxb4-ChIP-Seq(GSE34014)/Homer	1e-2	-5.711e+000.0057	115.0	1.40%	434.3	1.07%	mmioff file (matrix)	svg
212		Gata3(Zf)/Treg-Gata3-ChIP-Seq(GSE20898)/Homer	1e-2	-5.696e+000.0058	802.0	9.75%	3598.3	8.88%	mmioff file (matrix)	svg
213		Ppare(Nr)/DR1/3T3L1-Pparg-ChIP-Seq(GSE13511)/Homer	1e-2	-5.668e+000.0059	623.0	7.57%	2757.8	6.81%	mmioff file (matrix)	svg
214		Lxre(Nr)/DR4/RAW-LXRb,biotin-ChIP-Seq(GSE21512)/Homer	1e-2	-5.643e+000.0060	36.0	0.44%	108.1	0.27%	mmioff file (matrix)	svg
215		Myf5(bHLH)/GM-Myf5-ChIP-Seq(GSE24852)/Homer	1e-2	-5.502e+000.0069	533.0	6.48%	2342.6	5.78%	mmioff file (matrix)	svg
216		Pax8(Paired, Homeobox)/Thyroid-Pax8-ChIP-Seq(GSE26938)/Homer	1e-2	-5.419e+000.0075	268.0	3.26%	1121.9	2.77%	mmioff file (matrix)	svg
217		Tbx5(T-box)/HL1-Tbx5,biotin-ChIP-Seq(GSE21529)/Homer	1e-2	-5.156e+000.0097	2150.0	26.14%	10094.4	24.92%	mmioff file (matrix)	svg

218 **AGGTCATGACCTT**
 219 **AGGCTAGG**

218	FXR(NR)IR1/Liver-FXR-ChIP-Seq(Chong_et_al)/Homer	1e-2	-4.698e+00	0.0152	301.0	3.66%	1292.0	3.19%	motif file (matrix)	SYG
219	ZFX(ZF)/mES-Zfx-ChIP-Seq(GSE11431)/Homer	1e-2	-4.665e+00	0.0157	1310.0	15.93%	6071.8	14.99%	motif file (matrix)	SYG

Appendix references

Fornes, O. *et al.* (2020) ‘JASPAR 2020: update of the open-access database of transcription factor binding profiles’, *Nucleic acids research*, 48(D1), pp. D87–D92. doi: 10.1093/nar/gkz1001.

Heinz, S. *et al.* (2010) ‘Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities’, *Molecular Cell*. Elsevier, 38(4), pp. 576–589. doi: 10.1016/j.molcel.2010.05.004.



Graphic design: Communication Division, UIB / Print: Skjipes Kommunikasjon AS



uib.no

ISBN: 9788230864371 (print)
9788230866764 (PDF)